FEDERAL RESERVE BANK OF KANSAS CITY

Economic Review

November/December 1991

FEDERAL RESERVE BANK OF KANSAS CITY

Economic Review

November/December 1991

Contents

Policy Implications of Trade and Currency Zones: A Summary of the Bank's 1991 Symposium 37

By George A. Kahn

The world trading system may be coalescing into a set of geographic free trade zones. Europe 1992, the U.S.-Canada Free Trade Agreement, and the initiatives to include Mexico and Latin America in a Western Hemisphere free trade zone provide recent examples of efforts to remove tariff and nontariff barriers to trade among countries in geographic regions. If accompanied by currency zones—the adoption within regions of fixed exchange rates or a common currency—this move toward trade zones could bring major changes in the international monetary system and in domestic economic policies.

To explore possible ramifications of trade and currency zones, the Federal Reserve Bank of Kansas City invited distinguished central bankers, academics, and industry representatives to a symposium entitled "Policy Implications of Trade and Currency Zones." The symposium was held August 22-24, 1991, in Jackson Hole, Wyoming.

Kahn summarizes the symposium papers and the discussions they stimulated. In general, most of the program participants supported the move to a trade and currency zone in Europe, although some expressed doubt about the benefits of trade and currency zones in other parts of the world.

Check-Cashing Outlets in the U.S. Financial System 53

By John P. Caskey

In the current debate over banking reform, some policymakers and consumer advocates have expressed concern that many lower-income Americans have lost access to basic payment services provided by banks. Reports of branch closings in low-income areas and increased service charges have led to proposals that banks be required to provide basic banking services to all consumers.

Most discussions of this issue are incomplete, however, because they overlook check-cashing outlets, a principal alternative to banks for many low and moderate-income consumers. Despite evidence of rapid growth over the past decade, relatively little is known about the check-cashing industry.

Caskey examines the role of check-cashing outlets. He provides new evidence on the costs of payment services and suggests that check-cashing outlets may play a role in the debate over basic banking services.

The Move Toward Free Trade Zones

By Paul Krugman

Paul Krugman is professor of economics at the Massachusetts Institute of Technology. The article is based on his presentation at the symposium on "Policy Implications of Trade and Currency Zones," sponsored by the Federal Reserve Bank of Kansas City at Jackson Hole, Wyoming, August 22-24, 1991. The views expressed in the article are the author's and do not necessarily reflect the views of the Federal Reserve Bank of Kansas City.

From World War II until about 1980, regional free trade agreements and global trade negotiations under the GATT could reasonably be seen as complements rather than substitutes—as two aspects of a broad march toward increasingly open international markets. Since then, however, the two have moved in opposite directions. The 1980s were marked by stunning and unexpected success for regional trading blocs. In Europe, the EC not only enlarged itself to include the new democracies of Southern Europe, but made a lunge for an even higher degree of economic unity with the cluster of market-integrating measures referred to as "1992." In North America, Canada ended a century of ambivalence about regional integration by signing a free trade agreement (which is also to an important extent an investment agreement) with the United States; even more startlingly, the reformist Salinas government in Mexico has sought, and appears likely to get, the same thing. And in East Asia, while formal moves toward regional free trade are absent, there was after 1985 a noticeable increase in Japanese investment in and imports from the region's new manufacturing exporters.

Meanwhile, however, the multilateral process that oversaw the great postwar growth in world trade seems to have run aground. The major multilateral trade negotiation of the decade, the Uruguay Round, was supposed to be concluded in late 1990. Instead, no agreement has yet been reached. And while some kind of face-saving document will probably be produced, in reality the round has clearly failed either to significantly liberalize trade or to generate goodwill that would help sustain further rounds of negotiation.

The contrast between the successes of regional free trade agreements and the failure of efforts to liberalize trade at the global level has raised disparate reactions. Official pronouncements, of course, call for renewed progress on all fronts. In practice, however, choices of emphasis must be made. Some politicians and economists despair of the multilateral process under the GATT, and would like to see further effort focused on regional or bilateral negotiations that seem more likely to get somewhere. Others, seeing the multilateral process as ultimately more important, fear that regional deals may undermine multilateralism. It is possible to find respected and influential voices taking fairly extreme positions on either side. For example, MIT's Rudiger Dornbusch has not only been a strong partisan of a U.S.-Mexico free trade pact, but has called for a U.S. turn to bilateral deals even with countries far from North America, such as South Korea. On the other side, Columbia's Jagdish Bhagwati, now a special adviser to the GATT, not only advocates remaining with the traditional process but has actually condemned the prospective U.S.-Mexico deal.

How can reasonable and well-informed people disagree so strongly? The answer lies in part in the inherent ambiguity of the welfare economics of preferential trading arrangements; it lies even more in the peculiarly contorted political economy of international trade negotiations.

Even in terms of straightforward welfare economics, the welfare effects of the creation of free trade areas are uncertain; indeed, it was precisely in the study of customs unions that the principle of the "second best," which says that half a loaf may be worse than none, was first formulated. A customs union, even if it only reduces trade barriers, may worsen trade distortions; moreover, consolidation of nations into trading blocs may lead even intelligent governments with the welfare of their citizens at heart to adopt more protectionist policies toward the outside world, potentially outweighing the gains from freer trade with their neighbors.

Worse yet, however, the motives of governments as they engage in trade negotiations are by no means adequately described by the idea that they maximize national welfare. In general, trade policy (like any microeconomic policy) is very much influenced by pressure from organized interest groups; the traditional framework of trade negotiation under the GATT channels these political pressures in a way that has generally led toward freer trade, but from an economist's point of view this framework has led to the right results for the wrong reasons. Given this, it is very difficult to decide whether a shift in the domain of negotiations will be a good or a bad thing.

Should the move toward free trade areas be applauded or condemned? The purpose of this paper is to help clarify the issues in a fundamentally murky debate. It is primarily a discussion of conceptual issues rather than a survey of actual recent moves toward free trade areas, although since the key questions about that move are inherently empirical, some appeal to facts and cases is necessary.

The paper is in three parts. The first part reviews the relatively straightforward economics of preferential trading arrangements. The second is an attempt to describe and analyze the political economy of trade negotiations, and the reasons why changes in this political economy have recently pushed the world in the direction of regional free trade areas. The third part tries to pull the economics and politics together, for a general discussion of the problem of free trade areas versus multilateralism.

The Economics of Trading Blocs

In spite of the major rethinking of the theory of international trade that has taken place over the past dozen years, few economists would disagree with the proposition that a world with free trade will be better off than under any other plausible set of trade policies. Yet preaching the virtues of global free trade somehow does not seem to get us there, and it often seems easier to negotiate free trade or at least trade liberalization on a more local basis. Indeed, in spite of the growing ease of international communication, the 1980s saw a shift of emphasis away from global trade negotiations toward regional deals.

The apparent conflict between what economists say should be in everyone's interest and what actually seems to happen politically should be a warning flag—it suggests that whatever is going on in international trade negotiations, it is not welfare maximization. And as I will argue in the second part of this paper, any assessment of the move toward free trade areas depends critically on understanding what governments actually do as well as what they should do. Still, suppose one takes it as a given that for some reason it is possible to negotiate a degree of trade liberalization among subsets of countries that goes beyond what is possible at a world level. The question is then, should trade liberalization be permitted to proceed at two speeds? Or should one try to ban special deals and insist that countries offer to everyone the same terms they offer to anyone?

A naive view would be that since free trade is a good thing, any move toward freer trade should be welcomed. Unfortunately, the case is not that simple. At least three (not entirely unrelated) objections may be offered to preferential free trade agreements:

(i) *Trade diversion*: Trade liberalization among a subset of countries, even if it is not accompanied by an increase in protectionism against extra-bloc imports, may create perverse incentives that lead to specialization in the wrong direction.

(ii) *Beggar-thy-neighbor effects*: The formation of free trade areas may well hurt countries outside those areas, even without any overt increase in protectionism.

(iii) *Trade warfare*: Regional trading blocs, being larger than their components, will have more market power in world trade; this may tempt them to engage in more aggressive trade policies, which damage the trade between blocs and may (through a kind of Prisoners' Dilemma) leave everyone worse off.

The analysis of the effects of preferential trading arrangements is the subject of a huge and intricate literature. We can, however, quickly survey some of the main results that seem to be relevant to the current problem of regionalism in world trade.

Trade creation vs. trade diversion

In a classic analysis, Jacob Viner (1950) pointed out that a move to free trade by two nations who continue to maintain tariffs against other countries could leave them worse rather than better off. Viner's insight remains fundamental to all analysis of preferential trading arrangements, and is worth restating.

The essential idea can be seen from a numerical example (Table 1).[1] Imagine that one country—which, not entirely innocently, we call Spain—can produce wheat for itself, import it from France, or import it from Canada. We suppose that the cost to Spain of producing a bushel of wheat for itself is 10, that the cost of a bushel of wheat bought from France is 8, and that the cost of a bushel bought from Canada is only 5.

Suppose initially that Spain has a tariff that applies equally to all imported wheat. If it imports wheat in spite of the tariff, it will buy it from the cheapest source, namely Canada. This case is illustrated in the table by the column labelled "Tariff = 4." If the tariff is high enough, however—as in the case where it equals 6—Spain will grow its own wheat.

Now suppose that Spain enters a customs

union with France, so that French wheat can enter free of tariff. Is this a good thing or a bad thing?

If the tariff was initially 6, the customs union is a good thing: Spain will replace its expensive domestic production with cheaper imported French wheat, freeing its own resources to do more useful things. If, however, the tariff was initially 4, the customs union will cause Spain to shift from Canadian wheat to more expensive French wheat, shifting from a low-cost to a high-cost source. In that case the customs union may well lower welfare.

As Viner pointed out, in the first, favorable case the customs union causes Spain to replace high-cost domestic production with imports; it thus leads to an increase in trade. In the unfavorable case, by contrast, Spain shifts from a foreign source outside the free trade area to another source inside. Thus Viner suggested that "trade creating" customs unions, in which increased imports of trading bloc members from one another replace domestic production, are desirable; "trade diverting" customs unions, in which imports are diverted from sources outside the union to sources inside, are not. Loosely speaking, if the extra trade that takes place between members of a trading bloc represents an addition to world trade, the bloc has raised world efficiency; if the trade is not additional, but represents a shift away from trade with countries outside the bloc, world efficiency declines.

This simple criterion is extremely suggestive, and makes it easy to understand how regional trade liberalization can actually reduce rather than increase world efficiency. Perhaps the most obvious real-world example, as the illustration itself suggested, is the effect of EC enlargement on agricultural trade. The Southern European countries are induced, by their entry into the EC, to buy grain and other cold-climate products from costly European sources rather than the low-cost suppliers on the other side of the Atlantic. Meanwhile, the northern European countries are now induced to buy Mediterranean products like wine and oil (and perhaps also labor-intensive manufactured goods) from Southern Europe rather than potentially cheaper suppliers elsewhere, e.g. in North Africa. It is by no means implausible to suggest that because of these trade-diverting effects on agriculture, EC enlargement reduced rather than increased world efficiency.

Table 1

Hypothetical example of free trade area

	Tariff rate		
	0	**4**	**6**
Cost of wheat from:			
Spain	10	10	10
France, before customs union	8	12	14
France, after customs union	8	8	8
Canada	5	9	11

While the creation/diversion idea captures the essence of the problem, however, its suggestion that customs unions are about as likely to cause harm as good is somewhat too pessimistic. For both theoretical and empirical reasons, one needs to bear in mind that the simple creation/diversion idea misses some potential gains from customs unions, even ones that are mostly trade-diverting.

First and least interesting of these additional gains is the reduction of consumption distortions. Even if Spain's initial tariff does not prevent it from importing Canadian wheat, the tariff will still distort consumer incentives. And shifting to free trade with France will reduce this consumer distortion even while diverting trade.

A second gain from regional free trade, which is very important in practice, comes from the increased size and hence both productive efficiency and competitiveness of oligopolistic markets subject to economies of scale. When the European Common Market was formed in 1958, substantial trade diversion seemed a likely outcome. What turned the arrangement into a strong economic success was the huge intra-industry trade in manufactures, and the associated rationalization of production, that the Treaty of Rome made possible.[2]

Finally, a third gain from formation of a customs union is that regional integration characteristically improves a region's terms of trade at the rest of the world's expense.

This last effect is obviously something less than an unmitigated good thing. It makes a regional trade deal more attractive, but it also suggests that such deals can in effect be beggar-thy-neighbor policies.

The beggar-thy-neighbor effect

Imagine a world consisting of three countries, A, B, and C. It is easiest to imagine that each country is specialized in the production of a different set of goods. Also suppose initially that all three countries maintain the same tariff rate against all imported goods. Now suppose that A and B form a customs union, eliminating the tariff on goods shipped to each other, while maintaining their tariffs on goods imported from C. What happens to C?

The presumption is that C is made worse off, through a deterioration of its terms of trade. To see why, consider what would happen as the result of the customs union if the prices of all goods remained the same. Then A and B would each tend to buy more of each others' products, substituting away from consumption both of their own products and from consumption of goods imported from C. The net effect on the demand for A's and B's goods would be ambiguous, because each country would buy less of its own goods but sell more to the other. The demand for C's products, however, would unambiguously fall. Thus to clear markets, the relative price of C's goods will normally have to fall; unless there is too much asymmetry, the prices of both A's and B's products will rise in terms of C's.

This terms of trade loss will increase the benefits of a customs union to A and B. Indeed, a customs union may well be desirable from their point of view even if it leads primarily to trade diversion rather than trade creation — because it is precisely trade diversion, that is, a shift of demand away from imports from the outside world, that leads to the improvement in the terms of trade. The extra gain will, however, come at the rest of the world's expense. The point is that even if formation of a customs union does not involve any increase in external tariffs, it can still in effect be a beggar-thy-neighbor policy.

Again, this is not an abstract point. The United States has been concerned that the enlargement of the EC deprives its agricultural exporters in particular of traditional markets, and has sought offsetting *reductions* in EC protection against agricultural products. And indeed this is what must happen if a customs union is not to be a de facto beggar-thy-neighbor policy: formation of the union must be accompanied by a reduction in external tariffs.

A customs union that also reduces tariffs on imports from outside can still be beneficial, through the normal gains from trade and specialization. Indeed, the idea that one could adjust tariffs so as to keep a customs union's trade with the outside world unchanged is the basis of a well-known demonstration that a customs union is always potentially beneficial to its members (Kemp and Wan 1976). But will a group of countries forming a trade area normally lower their external tariff sufficiently to avoid any trade diversion?

This depends on their motivations in forming the customs union in the first place. In practice, trading areas are formed for a variety of reasons, in which a careful assessment of costs and benefits is not usually high on the list. In the messy world of motivations discussed in the second part of this paper, it is possible either that a trading area might offer the rest of the world concessions in order to mollify it, or that the new bloc might have economically irrational autarkic tendencies as a way of emphasizing the political content of integration. For example, in the context of fairly amicable trade relations, one could imagine the EC cutting tariffs and subsidies in order to compensate the United States for any loss of markets due to increased European integration. In another context, one could imagine the emergence of a political context in which Fortress Europe shows a preference for self-sufficiency even beyond the beggar-thy-neighbor point.

Before we turn to political economy, however, let us at least ask what the economically rational action would be. And it is fairly obvious: not only would it not normally be in the interest of a trading bloc to throw away all of its terms of trade gain by reducing external tariffs, it would normally be in the bloc's interest to *raise* its external tariffs.

The reason is that a trading bloc will normally have more monopoly power in world trade than any of its members alone. The standard theory of the optimal tariff tells us that the optimal tariff for a country acting unilaterally to improve its terms of trade is higher, the lower the elasticity of world demand for its exports. So for a trading bloc attempting to maximize the welfare of its residents, the optimal tariff rate will normally be higher than the optimal tariff rates of its constituent countries acting individually.

This implies that the adjustment of external tariffs following formation of a regional trading bloc will not only not eliminate the beggar-thy-neighbor aspect, it will tend to worsen it.

Trading blocs and trade war

An individual trading bloc will tend to gain even in the face of trade diversion by improving its terms of trade at the rest of the world's expense. If one goes from envisioning a single bloc to imagining a world of trading blocs, however, the blocs may beggar each other. That is, formation of blocs can in effect set off a beggar-all trade war that leaves everyone worse off.

Imagine a world of *four* countries, A, B, C, and D. Imagine also that A and B enter negotiations to form a free trade area. They find that the area will primarily produce trade diversion rather than trade creation, but that it will still increase their welfare by improving their terms of trade at C and D's expense. Thus A and B will, correctly, form a free trade area; and this area will have an incentive to act as a trading bloc and raise its tariffs on imports from C and D. But suppose that C and D make the same calculation. Then both blocs will raise tariffs in an effort to exploit their market power. Obviously both cannot succeed; one bloc's terms of trade will actually deteriorate, while the other's will improve less than if it were acting on its own. Meanwhile, trade diversion will be taking its toll on world efficiency. The result of the tariff warfare may therefore be to leave all four countries worse off than they would have been had the trading blocs not been formed. And yet the members of each bloc are better off than they would have been if they had not joined their bloc, and thus left themselves at the mercy of the other bloc. So the game of free trade area formation itself may (though it need not) be a form of Prisoners' Dilemma, in which individually rational actions lead to a bad collective result.

This hypothetical example provides a

simple justification for those who fear that the indirect costs of the move toward free trade areas will exceed the direct benefits. While it is an extremely stylized picture, it captures at least some of the concern of critics of regional trading arrangements, like Jagdish Bhagwati. The basic logic here is that the regional deals undermine the multilateral system, and that the gains in intra-regional trade are more than offset by losses of inter-regional trade. In effect, bilateralism or regionalism leads to global trade diversion.[3]

Of course this is only a possibility, not a certainty. Indeed, it is perfectly possible that the gains from free trade between the pairs greatly outweigh the losses from multilateral trade diversion. This is essentially an empirical question, but it is one on which some numerical exercises can shed at least some light.

Trading blocs and world welfare

In an earlier paper (Krugman 1991) I offered a way of making a suggestive back-of-the-envelope calculation regarding the effects of a move toward the formation of regional trading blocs. The formal model is in the appendix to this paper; here I sketch out the approach and its results.

The basic idea is to examine how world welfare changes as a highly stylized world economy is organized into progressively fewer, progressively larger trading blocs. A trading bloc is envisaged as consisting of a large number of small geographic units ("provinces"), each specialized in the export of a different good. (Countries, which presumably themselves consist of one or more provinces, play no explicit role in the analysis.) Each trading bloc chooses an external tariff to maximize the welfare of its members, taking other blocs' tariffs as given.[4]

How does world welfare change as the number of blocs is reduced? There are two effects. On one side, the smaller the number of blocs, the more potential trade is unencumbered by tariffs; in the limit, with only one trading bloc, we have global free trade. On the other side, every time one merges blocs into larger blocs, there will be trade diversion; this effect will be reinforced by the fact that bigger blocs will have more market power and thus normally set higher external tariffs.

Which effect dominates? We know that free trade is best, so as the number of blocs goes from 2 to 1 welfare must rise. On the other hand, in a world of many small blocs nobody would have much market power, and since most of each bloc's consumption would be imported and hence subject to the same external tariff, there would be little trade diversion. So a fall in the number of blocs from a very large number to a somewhat smaller number might well reduce welfare. We would therefore expect a U-shaped relationship between the number of blocs and world welfare: while the best of all possible worlds has only one bloc, the worst is not a totally fragmented world but one with a moderate number.

In the simplest version of this story, all provinces stand in symmetric relationship to one another, so that there are no "natural" trading blocs. In this case, as is shown in the appendix, there are only two parameters: the number of blocs and the elasticity of substitution between the products of any two provinces. Figure 3, in the appendix, shows the relationship between the number of blocs and world welfare for three values of this elasticity: 2, a number that implies very large monopoly power in trade (although it is still high compared with empirical estimates, which tend to be not much greater than 1); 4; and 10. Remarkably, for this wide range of elasticities we consistently get the same answer: world welfare is minimized for a world of three trading blocs. The resemblance to the apparent current trend makes this an extremely interesting result!

It is a result that should, however, be treated with considerable caution. Like any abstract model, this one makes a large number of simplifying assumptions; perhaps the most objectionable in this case is the assumption that under free trade any arbitrary pair of "provinces" would have the same volume of trade as any other. This amounts to assuming away geography, the extent to which some countries would be each others' major trading partners even in the absence of preferential trading arrangements. If trading blocs are formed, not with arbitrary membership, but among countries that would be each others' main markets anyway, the consolidation of the world into a limited number of such blocs is less likely to be harmful.

The importance of "natural" trading blocs

If transportation and communication costs lead to a strong tendency of countries to trade with their neighbors, and if free trade areas are to be formed among such good neighbors, then the likelihood that consolidation into a few large trading blocs will reduce world welfare is much less than suggested by the simple numerical example in Figure 3. The reason is straightforward: the gains from freeing intra-regional trade will be larger, and the costs of reducing inter-regional trade, than the geography-free story suggests.

Imagine, for example, a world of six countries, which may potentially form into three trading blocs. If these countries are all symmetric, then three blocs is the number that minimizes world welfare, and hence this consolidation will be harmful. Suppose, however, that each pair of countries is on a different continent, and that intercontinental transport costs are sufficiently high that the bulk of trade would be between continental neighbors even in the absence of tariffs. Then the right way to think about the formation of continental free trade areas is not as a movement from 6 to 3, but as a movement of each continent from 2 to 1—which is beneficial, not harmful.

In practice the sets of countries that are now engaging in free trade agreements are indeed "natural" trading partners, who would have done much of their trade with one another even in the absence of special arrangements. A crude but indicative measure of the extent to which countries are especially significant trading partners comes from comparison of their trade with what would have been predicted by a "gravity" equation, which assumes that trade between any two countries is a function of the product of their national incomes.

Even casual inspection of such gravity-type relations reveals the strong tendency of countries to focus their trade on nearby partners; that is, in spite of modern transportation and communications, trade is still largely a neighborhood affair.

The magnitude of the strength of natural trading blocs can be crudely calculated from a regression of the following form:

$$\ln(T_{ij}) = \alpha + \beta \ln(Y_i Y_j) + \sum_z \gamma^z D^z{}_{ij} ,$$

where T_{ij} represents the value of trade (exports plus imports) between some pair of countries i and j; and Y_i, Y_j represent the two countries' national incomes. We suppose that the countries belong to several groups that are or might become trading blocs, and we index these groups by z, with D^{zij} equal to 1 if the pair of countries i and j belong to group z, 0 otherwise. Then we would say that a potential trading bloc is natural to the extent that the estimated γ is strongly positive for that z.

The simplest regression of this kind that one can perform uses the G7 countries (which after all account for most of world output in any case) and defines the two groupings as z=1: the United States and Canada, z=2: Europe. The results of that regression are shown in Table 2. To nobody's surprise, they point out very strongly the local bias of trade: the United States and Canada, according to the regression, do 13 times as much trade as they would if they were not neighbors, while the four major European countries do seven times as much.

Of course these results are in part due to the fact that there are already special trading arrangements between the United States and Canada, on one side, and within the EC on the other. Yet the results are so strong that they make it overwhelmingly clear that distance still matters and still creates natural trading blocs.

To reemphasize why this matters: if a disproportionate share of world trade would take place within trading blocs even in the absence of any preferential trading arrangement, then the gains from trade creation within blocs are likely to outweigh any possible losses from external trade diversion.

While the coincidence between potential trading blocs and natural blocs helps allay fears of global immiserization, however, it also raises a warning flag about the indiscriminate use of the free trade agreement as a weapon of policy. U.S.-Canada free trade is almost certainly a good thing, not just because we like each other, but because the two countries plus Mexico clearly form a natural bloc. U.S.-Korea or U.S.-Israel free trade, to take examples of less neighborly proposals that have been floated, do not share the same virtue; indeed, Israel is if anything a natural member of the European bloc. Such "unnatural" free trade areas are highly likely to cause trade diversion rather than creation.

On the whole, however, the fact that geography has already given international trade a strong regional bias makes the concern that allowing free trade agreements at a regional level will lead to a Prisoners' Dilemma a minor one. That is, if governments maximized the welfare of their citizens, prospective moves toward regional free trade would almost surely do more good than harm to the members of the free trade areas.

Table 2
A G7 gravity regression

	Estimated value	T-statistic
α	-8.4302	-6.894
β	0.7387	8.966
γ^1	2.6092	6.576
γ^2	1.9823	9.479

$R^2 = 0.7796$

The major problem with this optimistic statement is, of course, that governments do no such thing. Before turning to the political economy of trade, however, we should also note an important point: while most of the world's output is generated by countries that appear likely to be inside one or another big free trade area, most people live outside. And it is these non-neighbors who are most likely to be beggared.

The innocent bystander problem

A turn to increased protectionism against outsiders by groups of countries that have formed free trade areas and as a result start behaving as a bloc toward the outside world is unlikely to leave the members of the blocs worse off. It can, however, quite easily do a lot of damage to countries that for whatever reason do not get inside the blocs.

Consider the following back-of-the-envelope example. Imagine that the world's industrialized countries plus a few developing countries were in fact to consolidate into three blocs, consisting of Europe, North America, and an East Asian collection centered on Japan. On average, these three blocs currently import about 10 percent of gross bloc product from outside themselves. Leaving aside agriculture, the average tariff equivalent they impose on these extra-bloc imports is currently fairly low; call it 10 percent.

Now suppose that because the blocs have more market power than their constituent nations, and in general behave more belligerently, they increase their external tariff equivalent to 30 percent. Given typical estimated elasticities, the effect of such a tariff rise would be to reduce extra-bloc imports by about 20 percent. We can use standard methods to come up with an estimate of the welfare loss from this tariff increase. The implied efficiency loss is the average of the initial and final tariff rates, multiplied by the fall in imports: 0.2 times 2 percent of gross bloc product, or 0.4 percent. This is a small though not negligible cost; more to the point, it could easily be outweighed by the gains from free trade within the trading blocs.[5]

But consider the same situation from the point of view of a nation that is not part of one of the blocs. This nation simply sees an increase in the tariff its exports must pay to enter the world's major markets. It will therefore suffer a terms of trade loss, which may be close to the size of the tariff increase. For example, a country that exports 15 percent of its GNP to the OECD nations, faced with a 20 percent rise in the external tariffs of the newly formed blocs, could suffer a real income loss of close to 3 percent—with no compensating gain in market access elsewhere. The point, then, is that the biggest costs of a consolidation of the world into a few large trading blocs would likely be borne not by the countries in the blocs but by those left out in the cold.

Summary

The purely economic analysis of free trade areas suggests that in principle formation of such areas might hurt rather than help the world economy. Trade diversion could outweigh trade creation even with external protectionism unchanged; and the increased market power that countries gain by consolidating into trading blocs could lead optimizing but noncooperative governments to raise tariffs increasing the cost.

While some moves toward free trade surely do produce costly trade diversion, however, it seems unlikely that the net effect on world efficiency will be negative. The reason is geography: the possibly emergent trading blocs consist of more or less neighboring countries, who would be each others' main trading partners even without special arrangements. As a result, the potential losses from trade diversion are limited, and the potential gains from trade creation are large.

The main concern suggested by this economic analysis is distributional: inward-turning free trade areas, while doing little damage to themselves or each other, can easily inflict much more harm on economically smaller players that for one reason or another are not part of any of the big blocs.

The Political Economy of Free Trade Areas

In a fundamental sense the issue of the desirability of free trade areas is a question of political economy rather than of economics proper. While one could argue against the formation of free trade areas purely on the grounds that they might produce trade diversion, in practice (as argued above) the costs of trade diversion are unlikely to outweigh the gains from freer trade within regions. The real objection is a political judgement: fear that regional deals will undermine the delicate balance of interests that supports the GATT. Implicit in this concern is the idea that government do not set tariffs to maximize national welfare, that they are instead ruled by special interest politics disciplined and channeled by an international structure whose preservation is therefore a high priority.

To discuss the political economy of free trade areas, it is necessary to offer at least a rough outline of how trade policy actually works, and of why free trade areas rather than multilateral agreements seem to be the current trend. Only then can we ask whether such preferential agreements will help or hurt the overall prospects for trade.

GATT-think and trade negotiations

International trade policy has many horror stories. Examples of outrageous policy, like the sugar quota that for a time led U.S. producers to extract sucrose from imported pancake mix, are easy to come by. All microeconomic policy areas, however, offer similar stories of government actions that disregard efficiency and cater to organized interests. Indeed, one may argue that the surprising thing about trade policy is how good it is. Think of the way that the U.S. government handles water rights in the West, or tries to control pollution: these show a disregard for even the most elementary considerations of economic logic or social justice that make trade policy seem clean and efficient. Arguably trade policy is one of our best microeconomic policy areas—largely because it is disciplined by international treaties that have over time led to a progressive dismantling of many trade barriers.

One might be inclined to ascribe credit for this to the economists. After all, economists have for nearly two centuries preached the virtues of free trade. It seems natural to think of the GATT, and the relatively free trading system built around the GATT, as the result of the ideology of free trade.

Yet if one examines the reality of international trade negotiations, one discovers that the GATT is not built on a foundation laid by economic theory. That is not to say that there are no principles. On the contrary, one can make a great deal of sense of trade negotiations if one adopts a sort of working theory about the aims and interests of the participants, a theory that is built into the language of the GATT itself. The problem is that this underlying theory has nothing to do with what economists believe.

There is no generally accepted label for the theoretical underpinnings of the GATT. I like to refer to it as "GATT-think": a simple set of principles that is entirely consistent, explains most of what goes on in the negotiations, but makes no sense in terms of economics.

The principles of GATT-think. To make sense of international trade negotiations, one needs to remember three simple rules about the objectives of the negotiating countries:

1. Exports are good.
2. Imports are bad.
3. Other things equal, an equal increase in imports and exports is good.

In other words, GATT-think is enlightened mercantilism. It is mercantilist in that it presumes that each country, acting on its own, would like to subsidize exports and restrict imports. But it is enlightened in that it recognizes that it is destructive if everyone does this, and it is a good thing if everyone agrees to expand trade by accepting each others' exports.

GATT-think is also, to an economist, nonsense. In the first place, general equilibrium theory tells us that the trade balance has very little to do with trade policy: a country that restricts imports will indirectly be restricting its exports as well. So even if one agreed with principles 1 and 2, one would argue that countries gain nothing from import restriction.

Nor do economists agree that exports are good and imports bad. The point of trade is to get useful things from other countries, i.e., imports, which are a benefit, not a cost; the unfortunate necessity of sending other countries useful things in return, i.e. exports, is a cost rather than a benefit.

Moreover, standard trade theory does not see export subsidies and import restrictions as similar policies. On the contrary, in general equilibrium an import tariff is equivalent to an export *tax*. Furthermore, in standard trade theory an export subsidy is a stupid policy but not a malicious one, since it generally worsens a country's terms of trade, and thus benefits the rest of the world. As Avinash Dixit once put it, when the Commerce Department ascertained that European nations had been subsidizing steel exports to the United States, its appropriate response should have been to send a note of thanks.

Finally, standard trade theory generally argues that free trade is the best *unilateral* policy, regardless of whether other countries do the same. That is, in standard theory one does not need to justify free trade in the context of international agreements. (The qualification is the optimal tariff argument, which generally plays no part at all in real-world trade discussion.)

In effect, GATT-think sees the trade policy problem as a Prisoners' Dilemma: individually, countries have an incentive to be protectionist, yet collectively they benefit from free trade. Standard trade theory does not agree: it asserts that it is in countries' unilateral interest to be free traders—as Bastiat put it, to be protectionist because other countries are is to block up one's own harbors because other countries have rocky coasts.

Yet although GATT-think is economic nonsense, it is a very good model of what happens. Indeed, it is embedded in the very language of the negotiations. Suppose that the United States succeeds in pressuring the European Community to stop exporting wheat that costs it three times the world market price to produce, or Japan to take a little rice at one-tenth the cost of domestic production. In GATT parlance these would represent European and Japanese "concessions": things that they would do unwillingly (and at present appear unwilling to do at all). That is, as GATT-think predicts, countries seem to treat exports—almost any exports, at almost any price—as desirable, and imports—no matter how much better or cheaper than the domestic substitute—as undesirable.

Moreover, over the years a trading system based on the principles of GATT-think has on the whole done very well. No amount of lecturing by economists on the virtues of free trade could have achieved the extraordinary dismantling of trade barriers accomplished by lawyers in the 30 years following World War II. If there are problems with the system now, they have more to do with perceptions that some countries are not playing by the rules than with a dissatisfaction of the political process with the rules themselves.

GATT-think, then, is very wrong yet somehow turns out mostly right. Why?

The hidden logic of GATT-think. GATT-think is not, presumably, the product of a continuing mercantilist tradition, preserved by legislators and lawyers in defiance of economists—although it is probably true that a more or less mercantilist view of trade comes more naturally to the untutored than the economist's blanket endorsement of free trade. The reason why GATT-think works is, instead, that it captures some basic realities of the political process.

Trade policy is a policy of details. Only a tiny fraction of the U.S. electorate knows that we have a sugar import quota, let alone keeps track of such crucial issues (for a few firms) as the enamel-on-steel-cookware case. What Mancur Olson (1965) taught us is that in such circumstances we should not expect government policy to reflect any reasonable definition of the public interest. Political pressure is a public good, and tends to be supplied on behalf of small, well-organized groups. In the case of trade policy, with few exceptions this means *producers*: producers of exported goods, producers of import-competing goods. The consumers who might have benefited from cheap imports, or the lower prices that would prevail if firms were not subsidized to provide goods to foreigners rather than themselves, count for very little.

This explains the first two principles of GATT-think: we need only append the words "for export producers" and "for import-competing producers," and one has statements with which economists can agree. Add that trade policy is set one industry at a time, so general equilibrium is disregarded, and that consumers are not at the table, and the mercantilist tone of trade negotiations is explained.

The third principle is more complicated. One would like to think that it reflects a residual concern with efficiency. Maybe it does. But it is also true that on average a dollar of exports adds more domestic value added than a dollar of imports subtracts, simply because not all imports compete directly with domestic goods. So perhaps the idea of gains from trade plays no role at all.

Yet the result of applying the principles of GATT-think has up to now been pretty good. The reason is the process of multilateral negotiation, which in effect sets each country's exporting interests as a counterweight to import-competing interests; as trade negotiators bargain for access to each others' markets, they move toward free trade despite their disregard for the gains from trade as economists understand them. (Notice also that in this context the GATT's harsh attitude toward export subsidies makes a great deal of sense: without such subsidies, export interests become a force for free trade; with free access to subsidies, they are not.)

During the 1980s, unfortunately, the effectiveness of the GATT process seemed to wane, with the focus shifting to regional free trade agreements. We must next ask why.

The erosion of the multilateral process

Everyone who thinks about it has his own list of problems with the GATT process. I would list four main factors that have eroded the effectiveness of the GATT mechanism at channeling special interests.

First is the decline of the U.S. leadership role. There is considerable disagreement among political scientists about the extent to which international policy coordination requires a hegemonic power. What is clear is that the dominant position of the United States in the early postwar period was helpful as a way of limiting free rider problems: the United States could and did both twist arms and offer system-sustaining concessions as a way of helping the GATT process work. With the United States accounting for a progressively smaller share of gross world product, and with U.S. dominance in productivity and technology progressively eroded, the United States has been losing both the means and the desire to serve as global trade hegemon.[6]

A second long-term trend that has undermined the GATT process is the growing subtlety of the issues that must be dealt with. Increasingly, trade negotiations must deal with problems for which regulating the policies imposed by nations at their borders are insufficient. The manufactured goods that enter world trade are increasingly knowledge-intensive; this implies both that traditional criteria for "unfair" trade practices are inappropriate and that domestic policies in support of R&D become issues of trade conflict. The growing role of direct investment blurs the lines between trade policy, which is subject to GATT discipline, and investment policy, which is not. And the role of government itself, and its intrusiveness into the economy, has (in spite of conservative ideological triumph) grown to a point where the distinction between international and domestic policies is difficult to draw.

A third problem is the changing character of protectionism itself, based on the creativity of bureaucrats. In the early postwar period, protectionism was a matter of explicit, unilateral government policies: tariffs, quotas, exchange controls. The great postwar liberalization steadily ratcheted these measures down, to the point where except in agriculture they are now fairly unimportant. But the new protectionism that emerged with increasing force after the mid-1970s was more slippery, exploiting the weaknesses of the system. "Voluntary" export restraints, orderly marketing agreements, harassment by countervailing duty cases, red tape barriers, etc., have all proved much more difficult to police than straightforward tariffs and quotas.

Finally, the legitimacy of the GATT system has been undermined by the growing importance of new players in the world economy—above all Japan—who are institutionally different enough from the original players to raise questions about what is being negotiated. The GATT is a system largely imposed by the United States, and created in our own image. That is, it is a legalistic system that focuses on process rather than results. Whatever the facts of the (much disputed) case, the widespread perception is that such legalisms are ineffective when dealing with Japan; that the Japanese economy may be as open de jure as one likes, and yet that the collusive institutional structure of Japan's economy will continue to produce an economy that is de facto highly protectionist.

From the economist's point of view, none of these trends should affect the desirability of free trade. Leaving aside some of the recent strategic trade policy arguments, the basic economic argument is still that unilateral free trade is the best policy; it doesn't matter whether there is a hegemon to enforce the rules, whether the rules are inadequate to the new game, whether players have become more adept at cheating, or whether there are new players for whom the rules are meaningless. Given the real political factors that underlie GATT-think, however, these factors do matter very much. And if the evidence of the 1980s is anything to go by, the cumulative effect of these problems has been to erode the effectiveness of the GATT process to the point where further progress has effectively ground to a halt.

The regional answer

The same checklist of frustrations with the GATT process helps explain why regional free trade agreements have gained so much force as an alternative.

First, the decline of the hegemonic role of the United States at a global level can be ignored in regional agreements where there either is a local hegemon or a special correlation of forces that makes such a hegemon unnecessary. In North America the United States obviously remains and will remain for the indefinite future the overwhelmingly dominant player; and U.S. political interest in helping Mexican reformers gives the U.S.-Mexico deal, at least, some of the national security gloss that used to be attached to the idea that free trade helped fight Communism. In Europe the case is somewhat more complex: in effect the idea of a single market is being pushed by a Franco-German entente, in which Germany for historical reasons needs to be seen as a good European nation, and France sees its national influence best served by being part of a European whole. In the EC enlargement, as in the U.S. embrace of Mexico, politics played a large part: the wealthy EC nations wanted to reward and safeguard the Southern European transition to democracy.

Our second and third problems with the GATT—the complexities of dealing with modern trade and with modern trade barriers—are also, on the evidence, more easily dealt with at a regional level than at a global level. Europe's 1992 is not so much a trade agreement as an agreement to coordinate policies that have historically been regarded as domestic. That is, it is in effect a mutual sacrifice of national sovereignty. The Canada-U.S. FTA also involves significantly more than free trade: it is a pact over investment rules, and involves creation of dispute settlement mechanisms that limit the ability of the countries to act unilaterally.

Why can regional pacts do what global negotiations cannot? The answer appears to be that neighbors understand and trust one another to negotiate at a level of detail and mutual intrusiveness that parties to global negotiations cannot. One does not hear U.S. businessmen raising the arguments against free trade with Canada that they raise against Japan—nobody claims that Canada is so institutionally different from the United States, so conspiratorial a society, that negotiated agreements are worthless and ineffective. We think that we understand and can trust the Canadians; apparently the European nations have reached a similar point of mutual understanding and trust. North Americans and Europeans have not reached a comparable state with regard to one another, and both deeply distrust the Japanese.

And this is the final point. Whether or not Japan is really a radically different kind of player from other advanced nations,[7] the perception that it is has done a great deal to undermine the perceived effectiveness and legitimacy of the GATT in the United States and Europe. So the great advantage of regional pacts is that they can exclude Japan.

One could argue that the surge of interest in regional free trade agreements is actually a godsend to world trade. Given the loss of momentum in global trade negotiations, regional pacts offer a route through which trade can still increase. Of course this trade increase might in principle be diversion rather than creation, and hence make the world worse rather than better off. As argued in the first part of this paper, however, the importance of natural blocs is such that this is unlikely.

The real case against free trade agreements is that they may undermine the effort to deal with the problems of the multilateral system.

Free trade agreements and the international system

In the past two years there has been a schizophrenic mood in Washington regarding trade policy. On one side, the dismal prospects for the Uruguay Round, and the perceived lack of public spirit by the Europeans, have led to disillusionment with the prospects for the GATT—and, to at least some extent, a resigned acceptance of the likelihood of greater U.S. protectionism against Japan. On the other side, prospects for free trade with Mexico have brought out the traditional export sector support for liberalization with full force. It has been noted by a number of observers that the U.S. business community has put much more effort into supporting Mexican free trade than into any other trade area, even though Mexico remains a considerably smaller market than either the EC or Japan.

European enthusiasm over 1992 has similarly gone hand in hand with a rather sour attitude toward trade with non-European nations, and in particular with a fairly notable failure to make any concessions on agriculture that would help make the Uruguay Round a success and thus help sustain the GATT's credibility.

Suppose that one could make the following two-part argument:

(i) By focusing on regional free trade, the United States and the EC have diverted political energies away from working on the problems of the GATT.

(ii) Had they committed themselves to working within a multilateral framework, they could have achieved a solution to the GATT's difficulties that would have led to better results than the local solutions they have achieved instead.

If one believed this argument, one could then believe that the rise of free trade agreements has had an overall negative effect.

Part (i) of the argument clearly has some validity. Free trade agreements in Europe and North America have diverted some political, administrative, and intellectual capital away from the multilateral negotiating process. They have also reduced the sense of urgency about getting on with that process.

But would the GATT process really have done much better in the absence of moves toward regional free trade? This does not seem too plausible. The GATT's problems are deep-seated; it is hard to imagine achieving anything at the global level remotely approaching what the EC and the Canadian/U.S. pact have accomplished. And the problem of Japan seems extremely intractable.

It is understandable that economists and trade negotiators who have grown up in a world in which multilateral negotiations were the centerpiece of trade policy would be disturbed by a shift in emphasis toward regional agreements, especially if that shift seems to impair the effectiveness of the multilateral process—which it does. But while the move to free trade areas has surely done the multilateral process some harm, it is almost surely more a symptom than a cause of the decline of the GATT.

The Impact of the Move Toward Free Trade Zones

An unsophisticated view would see Europe 1992 and the move toward North American free trade as unadulterated good things. Global free trade would be better still, but these moves at least are in the right direction. And even if one is dismayed by the disappointments of the Uruguay Round, one may still take comfort in the continuing integration of markets at a more local level.

A more sophisticated view sees both economic and political shadows. Free trade areas are not necessarily a good thing economically, because they may lead to trade diversion rather than trade creation. In the highly imperfect politics of international trade, regional free trade zones could upset the balance of forces that has allowed the creation of a fairly liberal world trading system.

The basic message of this paper is that the unsophisticated reaction is wrong in theory but right in practice. The prospects of trade diversion from free trade areas are limited, because the prospective trading blocs mostly fall along the lines of "natural" trading areas, countries that in any case do a disproportionate amount of their trade with one another. While regionalism does to some extent probably undermine the political force behind multilateral trade negotiations, the problems of the GATT are so deep-seated that it is unlikely that a world without regional free trade agreements would do much better.

The world may well be breaking up into three trading blocs; trade within those blocs will be quite free, while trade between the blocs will at best be no freer than it is now and may well be considerably less free. This is not what we might have hoped for. But the situation would not be better, and could easily have been worse, had the great free trade agreements of recent years never happened.

Appendix

Trading Blocs and World Welfare

This appendix lays out a simple model of the relationship between the number of trading blocs in the world economy and world welfare. It is based on Krugman (1991); as discussed in the text, it is intended as a guide to framing the issue rather than as a realistic tool for calculating the effects of free trade zones.

We imagine a world whose basic units are geographic units that we will refer to as "provinces." There are a large number N of such provinces in the world. A country in general consists of a large number of provinces. For the analysis here, however, we ignore the country level, focusing instead on "trading blocs" that contain a number of countries and hence a larger number of provinces. There will be assumed to be $B < N$ trading blocs in the world. They are symmetric, each containing N/B provinces (with the problem of whole numbers ignored). In this simplified world, the issue of free trade zones reduces to the following: how does world welfare depend on B?

Each province produces a single good that is an imperfect substitute for the products of all other provinces. We choose units so that each province produces one unit of its own good, and assume that all provincial goods enter symmetrically into demand, with a constant elasticity of substitution between any pair of goods. Thus, everyone in the world has tastes represented by the CES utility function

$$U = \left[\sum_{i=1}^{N} C_i^{\theta}\right]^{1/\theta}, \tag{1}$$

where C_i is consumption of the good of province i, and the elasticity of substitution between any pair of products is

$$\sigma = \frac{1}{1-\theta}. \tag{2}$$

A trading bloc is a group of provinces with internal free trade and a common external ad valorem tariff. We ignore the realistic politics of trade policy, and simply assume that each bloc sets a tariff that maximizes welfare, taking the policies of other trading blocs as given. This is a standard problem in international economics: the optimal tariff for a bloc is

$$t^* = \frac{1}{\varepsilon - 1}, \tag{3}$$

where ε is the elasticity of demand for the bloc's exports.

In a symmetric equilibrium in which all blocs charge the same tariff rate, it is possible to show that (see Krugman 1991)

$$\varepsilon = s + (1-s)\,\sigma, \tag{4}$$

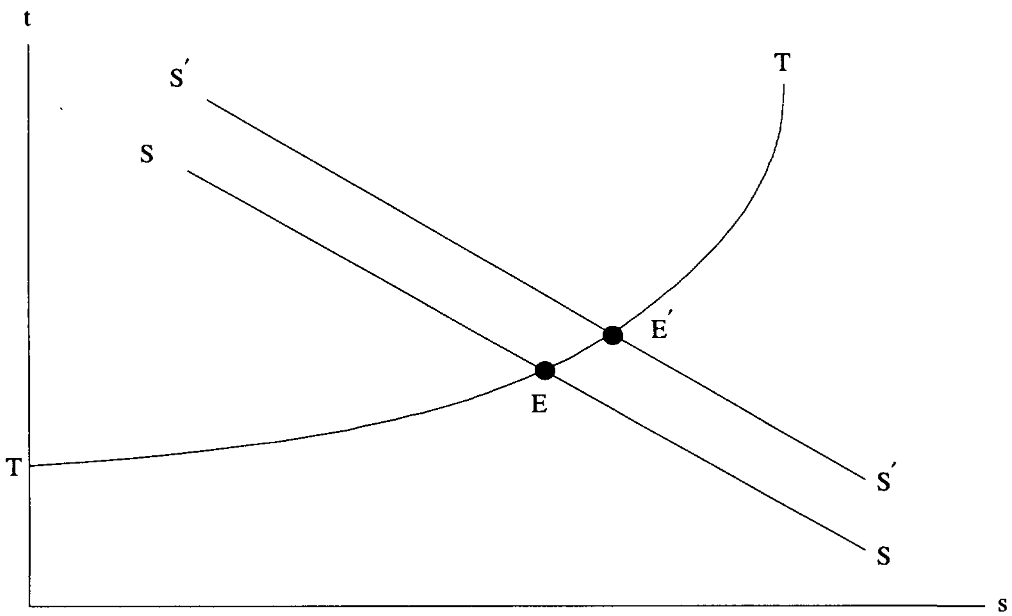
where s is the share of each bloc in the rest of the world's income measured at world prices. The optimal tariff is therefore

$$t^* = \frac{1}{(1-s)(\sigma-1)}. \tag{5}$$

It is apparent from (5) that the larger the share of each bloc's exports in the income of the world outside the bloc, the higher will be the level of tariffs on intra-bloc trade. This immediately suggests that a consolidation of the world into fewer, larger blocs will lead to higher barriers on inter-bloc trade.

One cannot quite stop here, however, because the share of each bloc in the rest of the world's spending depends both on the number of blocs and on the worldwide level of tariffs.

Figure 1



Again after some algebra it is possible to show that this share equals

$$s = \frac{1}{(1+t)^{\sigma} + B - 1}, \tag{6}$$

so that the share of each bloc's exports in the rest of the world's income is decreasing in both the tariff rate and the number of blocs.

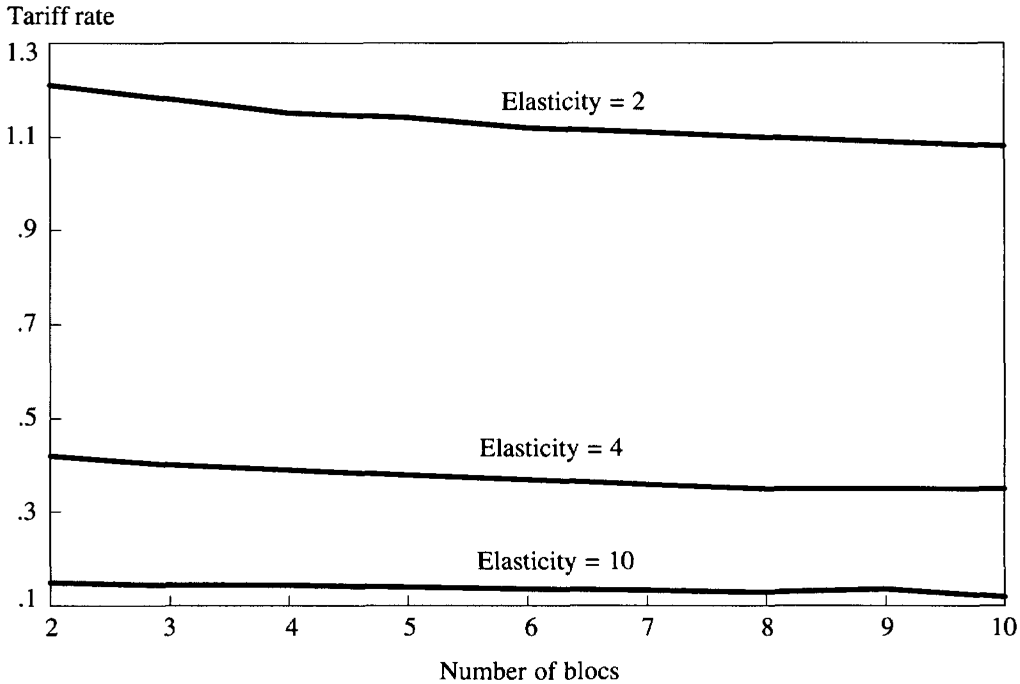
Equations (5) and (6) simultaneously determine the tariff rate and the export share for a given number of blocs B. In Figure 1, the downward-sloping curve *SS* represents (6); it shows that the higher is the worldwide level of tariffs, the lower the share of each bloc in the spending of other blocs. The curve *TT* represents (5); it shows that the optimal tariff rate is higher, the smaller that export share. Equilibrium is at point E, where each bloc is levying the unilaterally optimal tariff.

Now suppose that there is a consolidation of the world into a smaller number of blocs. We see from (6) that for any given tariff rate, the effect of the reduction in B is increase s; thus *SS* shifts up to $S'S'$. As a result, tariff rates rise, as equilibrium shifts from E to E'

Clearly this change will reduce the volume of trade between any two provinces that are in different blocs. Even at an unchanged tariff, the removal of trade barriers between members of the expanded bloc would divert some trade that would otherwise have taken place between blocs. This trade diversion would be reinforced by the rise in the tariff rate.

We now turn to welfare. Given the utility function (1), it is possible to calculate the welfare of a representative province as a function of the total number of provinces N, the number

Figure 2



of blocs B, and the tariff rate t on inter-bloc trade. Since N plays no role in the analysis, we can simplify matters somewhat by normalizing N to equal 1. Again after considerable algebra, given in Krugman (1991), we find that the utility of a representative province is

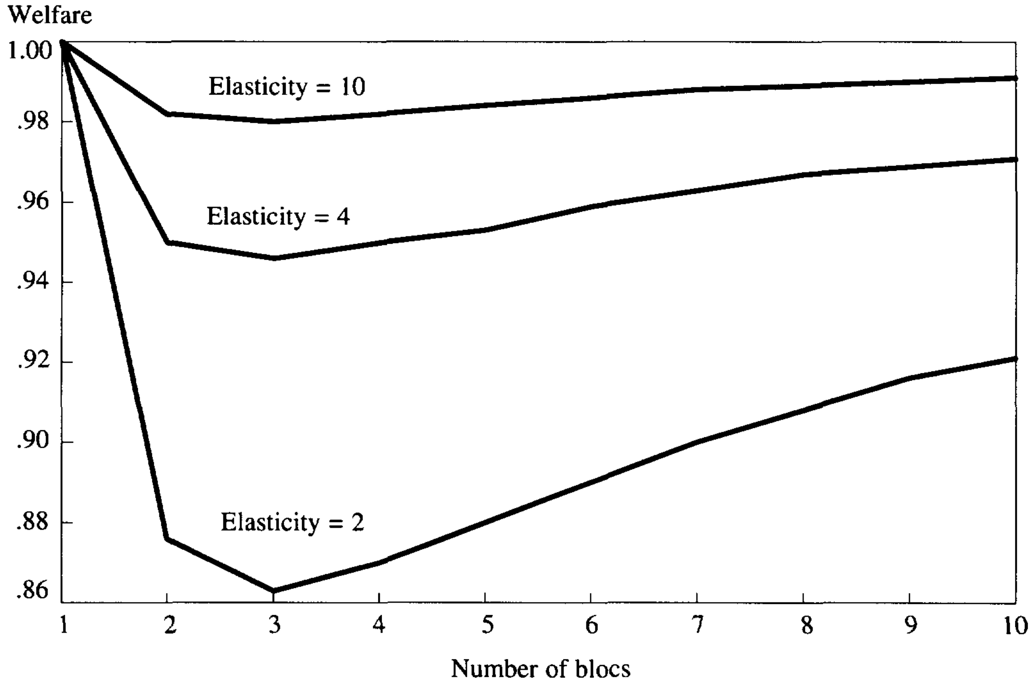
$$U = \left[\frac{B}{(1+t)^{\sigma}+b-1}\right]\left[(1-B^{-1}) + B^{-1}(1+t)^{\sigma\theta}\right]^{1/\theta}. \tag{7}$$

If trade were free, this would imply a utility of 1. Since the tariff rate t is also a function of B, we can use (5), (6), and (7) together to determine how world welfare varies with the number of trading blocs.

The easiest way to proceed at this point is to solve the model numerically. This grossly over-simplified model has only two parameters, the number of trading blocs and the elasticity of substitution between any pair of provinces; it is therefore straightforward to solve first for tariffs as a function of B given several possible values of the elasticity, and then to calculate the implied effect on world welfare. Here the values of ε considered are 2, 4, and 10.

Figure 2 shows how world tariff rates vary with the number of blocs. Two points are worth noting. First, the relationship between tariff rates and the number of blocs is fairly flat. The reason is that when there are fewer blocs, trade diversion tends to reduce interbloc trade, and thus leads to less of a rise in each bloc's share of external markets than one might have expected. Second, except in the case of an implausibly high elasticity of demand, predicted tariff rates are much higher than one actually observes among advanced nations. This is not an artifact of the

Figure 3



economic model: virtually all calculations suggest that unilateral optimum tariff rates are very high. What it tells us, therefore, is that actual trade relationships among advanced countries are far more cooperative than envisaged here.

Finally, we calculate welfare. Figure 3 shows the results. World welfare is of course maximized when there is only one bloc, in other words, global free trade. As suggested informally in the text, however, the relationship between welfare and the number of trading blocs is not monotonic but U-shaped: world welfare reaches a minimum when there are a few large blocs, and would be higher if there were more blocs, each with less market power.

The figure also shows a startling result: for the full range of elasticities considered, world welfare is minimized when there are three blocs.

As pointed out in the text, however, this result is an artifact of the assumption that under free trade any two provinces will trade as much as any other pair. That is, it ignores geography, which gives rise to natural trading blocs; as argued there, in practice the strength of this natural linkage is strong enough to make it unlikely that consolidation of the world into regional blocs would actually reduce welfare.

Endnotes

[1] Indeed, this is one of those concepts that tends to get lost if one uses anything more high-powered than a numerical example.

[2] Hopes for large benefits from both the U.S.-Canada free trade agreement and Europe 1992 rest largely on increased competition and rationalization. In the North American case, the estimates of Harris and Cox (1984), which attempt to take account of competitive/industrial organization effects, suggest a gain for Canada from free trade that is about four times as large as those of standard models. In Europe, the widely cited although controversial figure of a 7 percent gain due to 1992, presented in the Cecchini report (Commission of the European Communities (1988)) rests primarily on estimates by Alasdair Smith and Anthony Venables of gains from increased competition and rationalization.

[3] Bhagwati and others have, of course, a much subtler view than this. They are not so much concerned with the fear that trading blocs will pursue optimal tariff policies as with the fear that regional trade negotiations will shift political resources away from the task of defending global trade against special interest politics. So this approach is only a rough metaphor for a real political story to be described in the paper's second part.

[4] This setup is clearly both too cynical and not cynical enough about the political economy of trade. The internal politics of trade are not nearly this benign: governments do not simply (or ever) maximize the welfare of their citizens. At the same time, the external politics of trade show far more cooperation than this. An attempt at more realism follows later in the paper.

[5] The cost of an increase in protection here may seem surprisingly small. It is a familiar proposition to those who work with quantitative trade models, however, that the estimated costs of protection usually turn out to be embarrassingly small.

[6] It is surely also not irrelevant that with the collapse of the Soviet empire the national security argument for fostering free trade among U.S. allies has suddenly lost its force.

[7] I believe that concerns that Japan is fundamentally different, and that negotiated trade liberalization is largely ineffective for Japan, are justified; but what is important here is not what is true but what is believed.

References

Commission of the European Communities. 1988. *The Economics of 1992, European Economy*, no. 35. Commission of the European Communities, Luxembourg.

Harris, R. and Cox, D. 1985. "Trade Liberalization and Industrial Organization: Some Estimates for Canada," *Journal of Political Economy*, pp. 115-45.

Kemp, M. and Wan, H. 1976. "An Elementary Proposition Concerning the Formation of Customs Unions," International Economic Review, pp. 95-97.

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

Viner, J. 1950. *The Customs Union Issue*. New York: Carnegie Endowment for International Peace.

Commentary: The Move Toward Free Trade Zones

By C. Fred Bergsten

C. Fred Bergsten is Director at the Institute for International Economics. The article is based on his commentary at the symposium on "Policy Implications of Trade and Currency Zones," sponsored by the Federal Reserve Bank of Kansas City at Jackson Hole, Wyoming, August 22-24, 1991. The views expressed in the article are the author's and do not necessarily reflect the views of the Federal Reserve Bank of Kansas City.

There are three central strands to Paul Krugman's analysis of free trade areas (FTAs):

- that they are considerably better in practice than in theory;
- that this is particularly the case when they are viewed as *alternatives* to multilateral trade liberalization because "half a loaf is better than none";
- which is how they should be viewed because of the demise of the GATT and the poor prospects for the Uruguay Round (UR).

My view is fundamentally different on all three counts:

- that FTAs are considerably less desirable than Krugman suggests, especially in practice;
- that this is *particularly* true if they are seen as alternatives to an effective global trading system;
- but that, fortunately, they need *not* be seen as alternatives because the UR is quite likely to succeed, thereby restoring the credibility and central role of the GATT and making the world safe for FTAs which, as *complements* to such a global system, are acceptable and even desirable.

The Big Picture

Before discussing FTAs in detail, it is essential to place the issue of trade (and currency) zones—the topic of this conference—within the context of the sweeping structural changes that will dominate the world economy in the 1990s and beyond.

The first of these historic transformations is the onset of full economic tripolarity. By sometime in this decade, if not already, the three economic superpowers—uniting Europe, Japan, and the United

States—will look much more alike than different:

– In terms of absolute economic size, Europe will be bigger than the United States. Japan, already the largest creditor country and most competitive national economy, will match the GNP of both early in the twenty-first century even on conservative assumptions concerning growth rates and exchange rates.

– In terms of economic openness, defined as the share of exports plus imports of goods and services in GNP, the three are already virtually identical. This ratio has changed very little over the past two decades for Japan and the EC as a group but has risen sharply for the United States.

Hence *there is no longer any economic basis for American hegemony*, rightly described by Allen Meltzer in his paper as a major element in bringing relative prosperity and stability to the postwar world.

This outcome is reinforced by the second historic transformation: the end of the Cold War. The Cold War provided a "security blanket" over the trans-Atlantic and trans-Pacific relationships for almost half a century, suppressing potential conflicts on economic and other issues in the overriding interest of maintaining firm alliances against the Soviet threat. That security blanket has now been pulled aside: neither Europe nor Japan any longer needs much American protection, and the United States no longer needs to strengthen its allies—who also happen to be its chief economic competitors. Hence *there is no longer any security basis for American hegemony either.*[1]

The Gulf War reinforces the conclusion that America's economic dominance is a thing of the past despite its new status as the only military superpower. The United States had to insist that other countries pay for the war—the first admission of such economic dependence by a military leader in modern history. And, despite American efforts to lever its leadership of the Gulf coalition into greater foreign cooperation in the Uruguay Round and G-7 policy coordination, there appears to date to be zero transferability of military power into economic payoff (beyond the payments for the war itself).

The Policy Choice: Globalism or Blocs?

Economic (and other issues) are now much more likely to produce conflict among the Big Three because of the onset of equal tripolar economic power and the elimination of the Cold War glue that bound the allies together. Hence these historic transformations are central to the question of trade and currency zones. In broad strategic terms, the Big Three—who together will clearly dominate the world economy for at least the next few decades—can evolve in only two directions:

– into an informal steering committee (G-3) to revitalize and subsequently maintain a *globally oriented economic system* based largely on the existing institutional framework or

– into the poles of *regional blocs* where, for the reasons posited by Crockett in his paper, the dynamics would move from trade arrangements into deeper economic integration and then monetary zones and, as Krugman himself notes, the resulting entities would likely become exclusionary and discriminatory.

Hence the issue of trade (and currency) zones is far more important than welfare triangles or even dynamic gains from trade. The outcome of the current trade debate will go far to shape the course of the world economy for the coming decades. There will be significant political effects as well—especially if, as Krugman suggests, trade zones were pursued partly to discriminate against a major economic actor (Japan).

There is a widespread view around the world, sufficiently powerful that it is rapidly becoming a self-fulfilling prophecy, that we are headed toward the second outcome: regional blocs. The deepening and widening of European economic unity—toward "completion of the single market" in 1992, Economic and Monetary Union (EMU) and the addition of more members and associates (including Eastern Europe)—generate defensive reactions in the Americas and Asia. Initiatives by the United States toward a North American Free Trade Area (NAFTA) and the Enterprise for the Americas Initiative (EAI) produce Asian fears that "the Western Hemisphere is going regional too," both generating proposals for exclusionary regional groupings there (notably Malaysia's East Asian Economic Grouping) and making it harder for Japan and others to resist such calls. The resulting "evidence" of burgeoning Asian regionalism reinforces advocacy of similar steps in the Americas. Some Europeans then cite both to justify the inward-looking focus of their own initiatives. The critical importance of renewing the postwar momentum of trade liberalization on a *global* basis, the only alternative to eventual realization of the prophecy, is shunted aside in the rush toward regionalism. Krugman's paper unfortunately supports this spiral by prematurely writing an obituary for the Uruguay Round, which he rightly suggests is essential to restoring momentum and credibility for the multilateral system.

A revitalized global system managed collectively by the Big Three is far superior to a devolution into regionalism. Within such a system, regional arrangements would still take place but they would complement the global order rather than substitute for it. I believe that it is still eminently possible to forge such a global approach:

– As pointed out by both Meltzer and Frenkel-Goldstein, the trade patterns of the Americas and Asia are quintessentially multilateral. They have experienced *no* long-run trends toward increased reliance on intra-regional trade. Indeed, the trade patterns of both the United States and Japan—the core countries of the supposed blocs—are split into almost equal thirds. These countries have no interest in substituting regional for global arrangements. Europe is now the only bloc but the share of extra-regional trade in its GDP is even greater than for the Americas or Asia so it too needs a multilateral world.[2]

– The markets of the three economic superpowers (and much of the rest of the world) are deeply intertwined. There would be enormous economic costs from any significant erosion of global trade and financial openness, and resulting political costs for those who let it happen.

– The Big Three are democracies, have been allies for over four decades, and have a habit of working closely together on economic issues. Despite the absence of historical precedents for effective cooperative leadership, they should be able to provide it.

– Though the UR has clearly faltered, multilateral trade negotiations always resemble "the Perils of Pauline." The prospects for both the Kennedy and Tokyo Rounds looked extremely grim at key points before their eventual successes. The "failure" to conclude the UR at Brussels in December 1990 should have come as no surprise because the only real deadline for such talks is the expiration of the negotiating authority extended to the U.S. Administration by the Congress—a deadline set for June *1993* by the Trade Act of 1988 and duly reaffirmed by extension of the "fast track" authority in May 1991. The Round is quite likely to achieve major success, probably greater than either the Kennedy or Tokyo Rounds, if only because the costs of failure would be so high in both economic and political (especially United States-Europe) terms.

Does It Matter?

The central issue is whether global or regional trade liberalization is superior and, in particular, whether there need be any conflict between them. Krugman recognizes that trade blocs are decidedly second best because they generate trade diversion and because "they would upset the balance of forces that has allowed the creation of a fairly liberal world trading system." He attacks the "proposals" for "unnatural" (i.e., non-neighborly) free trade agreements, such as U.S.-Israel (which has been in place since 1985) and U.S.-Korea. He himself points out that "world welfare is *minimized* (my emphasis) for a world of three trading blocs."

But Krugman goes on to endorse blocs, arguing that prospective diversion is modest because they are likely to take place mainly among geographical neighbors and thus the blocs "mostly fall along the lines of 'natural' trading areas." This is an empirical question on which Krugman offers little supportive evidence. There are four reasons why I believe the view is flawed.

First, the impact of geography on trade has declined dramatically in recent decades. Geographical propinquity is no longer central to trading patterns.[3] For example, American trade is much denser with Korea and Taiwan—"unnatural trading partners" in Krugman's view—than with Argentina and Brazil, even adjusting for the different size of the respective economies.

Second, partly as a result (and as already noted), there are no major "natural trading areas" anyway except for Europe and possibly NAFTA. United States and Japanese trade is split into almost equal thirds. The Americas and Asia as a whole are highly diversified. The concept of "natural trading areas" rationalizes the EC and NAFTA but provides no guidance beyond.

Third, it must be candidly recognized that trade diversion is a *goal* of many contemporary proposals for trade blocs. Canada sought primarily to achieve preferential treatment (i.e., exceptions) under any new protectionist steps by the United States. Mexico is driven importantly by a similar motive. As already noted, and stressed by Krugman, anti-Japanese sentiment lies near the surface of many FTA initiatives. Such a desire for discrimination suggests that it could very well occur.

Fourth, Krugman's supposition that neighboring countries would be the primary beneficiaries of trade liberalization anyway—so why *not* proceed on a regional basis?—does not stand up in practice, at least in the case of the Western Hemisphere. The hypothesis can be tested by assuming U.S. liberalization on an MFN basis and asking which countries would "naturally" get the business. Tariffs are already so low that their elimination would not make much difference. Hence the outcome would be determined primarily by the new trade patterns generated by liberalization of the seven large U.S. import quota regimes:[4]

1. Textiles and apparel: East Asian, South Asian, and some other developing countries are far more competitive than Latin America.
2. Steel: Brazil and Mexico could take some advantage but the major increases would accrue to Europe and Japan.
3. Automobiles: Mexico and perhaps Brazil could expand sales of parts but the overwhelming increases would come from Japan, Korea, and possibly Europe.
4. Machine tools: virtually all new imports would come from Europe, Japan, and Taiwan.
5. Dairy products: the bulk of the increased trade would emanate from Australia, New Zealand, and Europe.
6. Sugar: several Latin American countries could compete effectively if U.S. quotas were lifted but Australia and several others outside the Hemisphere would also be major beneficiaries.
7. Meat: several Latin American countries could gain markets but the bulk of the increased imports would derive from Australia and New Zealand.[5]

The lesson is that trade liberalization by the

United States on a regional basis would almost certainly generate much more trade with uncompetitive countries than with efficient suppliers. Economic welfare would be reduced to the extent that current (efficiently produced) imports were supplanted by less efficiently produced imports. For example, the United States has already unilaterally increased Mexico's share of its textile quotas while deducting a like amount from the quotas of Hong Kong and other Asian suppliers. Since the latter are considerably more efficient, the shift has further increased the welfare costs of the textile quotas to the American economy.[6]

Beyond this central point, there are a number of additional reasons why Krugman's advocacy of trade zones does not stand up in practice:

– His conceptual case for free trade arrangements, akin to the optimal tariff argument, is that they can strengthen the region's terms of trade by increasing its weight in the global economy and permitting it to extract better prices from its trading partners. This would, however, by definition hurt other countries. Moreover, it has very little to do with the contemporary world: "EC 1992" and EMU emphasize deepening rather than broadening of Europe's economic zone, and the creation of NAFTA would add less than 15 percent to the weight already exercised by the United States in the world economy.

– He is simply wrong to argue that regional trade deals produce bigger results than global deals. The U.S.-Canada FTA, contrary to his assertion, was a mouse in terms of liberalization: on the biggest issues, like agriculture and subsidies, the countries explicitly deferred to the Uruguay Round because there was not enough benefit on other issues in the bilateral context to justify taking on the domestic opponents.[7] Is it conceivable that the America's textile quotas could be liberalized more meaningfully in NAFTA than in GATT, where the offsetting "gains" (in GATT-think terms) would at least give the effort a fighting chance?

– He notes that, to an economist, unilateral liberalism is best but fails to observe that countries all over the world are practicing it: Australia and New Zealand, Eastern Europe and many in both Asia and Latin America. A successful Uruguay Round can induce these countries to bind their new regimes and thus obviate the risk of reversal. On the other hand, a withering of GATT would make it much easier for them to reverse gears—and could even compel them to try to strike defensive deals with one or another bloc instead that would include the erection of new barriers against outsiders.

– This would clearly include "unnatural" alignments of the type that Krugman himself denounces. In particular, few Asian countries want to join a bloc led by Japan. The United States is unlikely to "settle for" Latin America, both because such insulation from the most dynamic world markets would erode its own competitiveness over time (as Britain's preferences within the Empire and, later, Commonwealth undermined its economic strength) and because all the other countries in the Hemisphere are also debtors and cannot help the United States improve its trade balance. Moreover, the United States could hardly push for a Western Hemisphere bloc and oppose Japan's pushing for an Asian bloc—as it clearly would—without offering the Asians a place in its own "regional" arrangement. Hence there would almost certainly be an "unnatural" trans-Pacific dimension to a world of trading areas.

– The increasingly central global role of multinational enterprises adds to the potential for a negative dynamic if a world of blocs were ever to get seriously underway: once positioned within each bloc to hedge themselves, the companies would enjoy relative gains from the erosion of inter-bloc trading freedom and would, at a minimum, no longer espouse global liberalization. Other constituencies within member countries of a bloc also acquire a distaste for global liberalization and thus add to the exclusionary dynamic.

– Krugman strangely ignores the historical absence of any successful free trade agreements between industrial and developing countries, despite the centrality of this issue to any meaningful construction of blocs in Asia and the Americas. The difficulties in combining Japan and China, or even the United States and Mexico, loom considerably larger than meshing Greece and Portugal with the EC—and even that arrangement includes transfers of public capital equal to 5 percent of the GNP of the LDC partners.

– Indeed, as Krugman notes, the biggest losers from a world of regional blocs would be those left outside— which in practice would be primarily the poorest developing countries which could least afford it.

– One can only cringe when Krugman argues that "the great advantage of regional pacts is that they can exclude Japan." Many Americans and Europeans certainly do "deeply distrust the Japanese," as he asserts. It does not take much knowledge either of history or of contemporary thinking in Japan, however, to conclude that steps to institutionalize, rather than combat, that distrust would run enormous risks. History teaches that failure to accommodate rising powers in the systemic structure is a sure recipe for serious conflict.

– On the political economy plane, both the United States and Japan have sufficient national power to be world leaders without forming blocs around them. No individual European country does; hence bloc creation was essential to restore that area as a global player but such considerations hold nowhere else.

– It would be particularly tragic if the countries that created and nurtured the global trading system and the GATT, notably the United States and to a degree the EC, were to turn their backs on it now when (a) virtually all of the countries which have heretofore rejected that regime are now clamoring to get in (the USSR, China, East Europe, and most of Latin America), and (b) the developing countries have, in the Uruguay Round, for the first time become active participants in it.

Would FTAs Undermine Globalism?

Regional trading arrangements are clearly going to happen: further deepening and eventually broadening in Europe, NAFTA, and perhaps the Enterprise for the Americas Initiative in this hemisphere, Australia-New Zealand, and even conceivably an East Asian Economic Grouping per the current Malaysian proposal. Another possibility is a Pacific Basin construct, growing out of the recent Asia-Pacific Economic Cooperation initiative.

The crucial question is whether these arrangements take place within the context of an effective and credible *global* system. If so, they will be—and will be viewed as—*supplements* to that system between countries that choose to liberalize further together, perhaps providing a constructive challenge for emulation at the global level.

Indeed, it is the existence of tariff bindings under GATT (along with the proscriptions of Article XXIV itself) that prevent bloc members from raising barriers toward the outside world to exploit the potential gains described by Krugman. Even more importantly, it was the major liberalizing negotiations under the GATT—the Kennedy Round in response to the creation of the Common Market itself and the Tokyo Round in response to its broadening to include the United Kingdom and others—that achieved the reductions in the common external tariff of the EC that, as he correctly notes, were essential to convert the Community from a beggar-thy-neighbor arrangement into a positive force for the world economy.[8] At a minimum, a strong GATT system is essential to avoid the costs that Krugman acknowledges are quite likely to result from FTAs.

If there is no effective GATT system, FTAs would almost certainly come to be viewed as *alternatives* to globalism. In that case, they would almost certainly evolve over time—as Krugman suggests—in an exclusionary and eventually discriminatory direction. The economic costs would be significant and growing. The political effects would, at a minimum, be worrisome.

The present stalemate in the Uruguay Round has sharply raised the prospect of the regional path. If the Round were to fail, the trend toward regionalism almost certainly *would* accelerate. And it will be much harder to avoid "failure" of this multilateral negotiation than in the past because a modest agreement that tries to paper over the major problems would be denounced as such by the growing corps of proponents of regionalism as well as others; world leaders and trade officials can no longer "declare victory and go home."

The United States usually plays the pivotal role on international trade issues. It will do so even more so in this case: Europe is already a bloc and Asia is clearly not, so the United States will tip the balance. It is thus imperative for the United States to continue to make clear that its priority is a successful outcome to the Round.

The United States was motivated to negotiate the FTA with Canada, in the wake of the failed GATT Ministerial of late 1982, primarily to spur the launch of what became the Uruguay Round.[9] It fully intended to complete the Round before negotiating NAFTA, reaffirming the primacy of the global system. It has held back on any substantive negotiations with Latin American countries, other than Mexico, despite the eagerness of Chile and others to commence such talks.

The "failure" at Brussels in December 1990, however, means that NAFTA may now be concluded before—or simultaneously with—the Round. Hence the United States will be characterized as "joining the rush toward regionalism." This will reinforce the self-fulfilling prophecy, as noted above, making it harder for Japan and others in Asia to resist blandishments such as Malaysia's to pursue defensive arrangements of their own.

As important as continued American fealty to a successful Uruguay Round is full support for such an outcome from Europe and Japan. Europe bears a special responsibility in this context. As the only trade bloc, it has done much to stimulate similar developments in other parts of the world. Its current inward orientation, while unlikely to produce a "Fortress Europe," has raised anxieties elsewhere and intensified the risk of realization of the self-fulfilling prophecy. The EC has been the key partner of the United States in achieving successful outcomes of the last two global trade negotiations; it has both a major interest in, and major responsibility for, doing so again.

The stakes are even higher than the future of the international trading system, however. As noted at the outset, the overarching issue for world economic policy in the decade or more ahead is whether the Big Three can effectively co-manage a reinvigorated global order. The Uruguay Round is one of the first test cases. If the Big Three cannot deal with a few farmers and other recalcitrant interest groups, they will hardly be able to provide global leadership on the wide array of issues—including money, macroeconomic cooperation, energy, and the environment, as well as trade and the GATT—where it will be needed.

The Monetary Dimension

Finally, it is necessary to note that the one monetary bloc now extant and potentially expanding in the near future—again, in Europe—could also raise significant problems for the global system.

A successful move to EMU will convert Europe from a series of small and medium-sized open economies into one large and much less open economy. This change alone will have several effects:

– It will tend to increase the extent of currency fluctuations among Europe, America, and Japan—generating greater international financial instability and potentially misalignments that would distort trade and add further to the tendencies toward trade protection outlined above.

– It will tempt Europe to practice "benign neglect" from time to time, as the other large and relatively closed economy has done, or at least to try to force the costs of adjustment onto others as the United States has also done.

– If it fails to achieve a unified fiscal policy to go with its unified monetary policy, there will be a strong possibility of a Europe-wide repetition of Reaganomics from the early 1980s and the German policy mix of the early 1990s: large fiscal stimulus, very tight money, a sharp appreciation of the currency, big trade deficits, and resultant protectionism.

– Without a political master, the European Central Bank will be particularly likely to foster such an outcome. This will be especially true in its early years, as it seeks to prove its fealty to the goal of price stability and to discipline recalcitrant governments into fiscal rectitude.

Moreover, achievement of EMU—even without the final step of a single currency, but especially with it—will propel the ECU to a central role in a new multiple reserve currency system. This will both reflect and produce a substantial portfolio adjustment from (mainly) dollars into ECU, reinforcing the likely appreciation of European currencies with attendant trade balance and protectionist problems. This effect would be further accelerated if the EMU pooled Europe's monetary reserves and attempted to dispose of some of the "excess," identified by the EC Commission as on the order of $200 billion.[10]

The policy implication is that the United States and Japan should engage Europe in negotiations on the global monetary system while the latter works out its regional arrangements—particularly as both of the basic blueprints for EMU, the report of the Delors Commission[11] and Karl Otto Pöhl's design for a Eurofed,[12] totally ignored the external dimension thereof. American strategy in the trade area has been to engage Europe in a global negotiation at each key milestone in its evolution: the Kennedy Round when the Common Market was created, the Tokyo Round when it expanded to bring in the United Kingdom and others, the Uruguay Round as it moved toward "1992." A similar approach is needed in the monetary area to avoid the risk that EMU will destabilize global arrangements and that, once its details have been put in place, it will be too late. This should be feasible now that the G-7, by successfully placing a floor under the dollar in February 1991 and (so far) effectively capping the dollar in July 1991, seems to be returning at least de facto to reference ranges among the major currencies *a la* Louvre.

Epilogue

There is still time to restore the effectiveness and credibility of the global approach to world economic policy and its existing institutional framework. Contrary to Krugman's assertions, the Uruguay Round is still alive—and, if not totally well at this juncture, with reasonable prospects for meaningful success. EMU can still be channeled in directions that are fully compatible with global monetary stability.

The Big Three must seize leadership on both issues (and several others) and make a conscious effort to restore a global focus, however, or the regional drift will continue and perhaps accelerate. The costs of permitting such an outcome could be extremely high in both economic and political terms. Reversing it is the first major test the Big Three face in the tripolar, post-Cold War world economy of the 1990s and beyond.

Endnotes

[1] A detailed analysis of these historic transformations on the world economy can be found in C. Fred Bergsten, "The World Economy After the Cold War," *Foreign Affairs*, Summer 1990, pp. 96-112.

[2] Robert Z. Lawrence, "Emerging Regional Arrangements: Building Blocks or Stumbling Blocks?" mimeo, June 1991.

[3] The downsizing of economic activity accelerates that trend. See Alan Greenspan, "Remarks at the Annual Dinner of the Japan Society," New York, May 22, 1989.

[4] See the discussion of the impact of FTAs on nontariff barriers in Paul Wonnacott and Mark Lutz, "Is There a Case for Free Trade Areas?" in Jeffrey J. Schott, ed., *Free Trade Areas and U.S. Trade Policy*, Washington: Institute for International Economics, 1989, p. 64.

[5] These U.S. quota regimes are largely implemented through "voluntary" restraints administered by the exporting countries which are enormously costly to the United States because they transfer most of the quota rents to the exporting countries and firms. See C. Fred Bergsten, Kimberly Ann Elliott, Jeffrey J. Schott, and Wendy E. Takacs, *Auction Quotas and United States Trade Policy*, Washington: Institute for International Economics, September 1987.

[6] These costs are estimated in William R. Cline, *The Future of World Trade in Textiles and Apparel, Washington: Institute for International Economics,* Revised 1990.

[7] Jeffrey J. Schott and Murray G. Smith, eds., *The Canada-United States Free Trade Agreement: The Global Impact.* Washington, DC: Institute for International Economics, 1988.

[8] Excluding of course its agricultural component which, however, resulted from the EC's decision to create a customs union rather than an FTA and where liberalization was severely hampered by the United States decision to take agriculture out of the GATT in 1955 to protect its own farmers.

[9] The difference between that situation, when a regional initiative by the United States clearly promoted global liberalization, and that of today is emphasized in Jeffrey J. Schott, *More Free Trade Areas*? Washington: Institute for International Economics, 1989.

[10] See *One Market, One Money: An Evaluation of the Potential Benefits and Costs of Forming an Economic and Monetary Union*, Brussels: Commission of the European Communities, European Economy, No. 44, October 1990, Chapter 7.

[11] *Report on Economic and Monetary Union in the European Community*, prepared by the Committee for the Study of Economic and Monetary Union, April 1989.

[12] "Basic Features of a European Monetary Order," a Lecture by Karl Otto Pöhl, President of the Deutsche Bundesbank, organized by Le Monde, Paris, January 16, 1990.

Policy Implications of Trade and Currency Zones: A Summary of the Bank's 1991 Symposium

By George A. Kahn

The world trading system may be coalescing into a set of geographic free trade zones. Europe 1992, the U.S.-Canada Free Trade Agreement, and the initiatives to include Mexico and Latin America in a Western Hemisphere free trade zone provide recent examples of efforts to remove tariff and nontariff barriers to trade among countries in geographic regions. If accompanied by currency zones—the adoption within regions of fixed exchange rates or a common currency—this move toward trade zones could bring major changes in the international monetary system and in domestic economic policies.

The move toward trade and currency zones comes at a time of great change in the world economy. International financial markets have become increasingly deregulated. International trade in goods and services has increased. The world economy has moved closer to a tripolar monetary system with the U.S. dollar, German mark, and Japanese yen serving as principal currencies. And multilateral negotiations to promote free trade, such as the General Agreement on Tariffs and Trade (GATT), have stalled.

To explore possible ramifications of trade and currency zones, the Federal Reserve Bank of Kansas City invited distinguished central bankers, academics, and industry representatives to a symposium entitled "Policy Implications of Trade and Currency Zones." The symposium was held August 22-24, 1991, in Jackson Hole, Wyoming. In opening comments, Federal Reserve Chairman Alan Greenspan underscored the importance of the topic and surveyed the issues to be addressed. Acknowledging our limited experience with trade and

George A. Kahn is a senior economist at the Federal Reserve Bank of Kansas City. Carol Manthey, a research associate at the bank, assisted with the preparation of the article.

currency zones, he argued that answers would have to come both "from the abstract world of economic models and from the ongoing experience gained in the cases of European economic and monetary union and the North American Free Trade area that are already being planned." He also stated that "insights into the economic implications we can expect from trade and currency zones should guide us in choosing appropriate macroeconomic policies now and in the future—whether we are 'inside' or 'outside' a zone."

This article summarizes the symposium papers and the discussions they stimulated. In general, most of the program participants supported the move to a trade and currency zone in Europe, although some expressed doubt about the benefits of trade and currency zones in other parts of the world. The first section of the article discusses whether the move toward trade and currency zones will promote trade among countries. The second section describes financial market and macroeconomic policy implications of trade and currency zones. The third section explores global implications.

Will Trade and Currency Zones Promote World Trade?

Two areas of heated debate at the symposium were whether the move toward free trade zones will promote world trade and whether currency zones will be necessary to achieve the full benefits of trade zones. Participants agreed that if the move to free trade zones is accompanied by further progress on the GATT, trade zones will help foster world trade. But participants disagreed sharply about the effects of trade zones in the absence of further progress in reducing trade barriers on a multilateral basis. Participants also disagreed about whether currency zones will be necessary to realize the full benefits of trade zones.

Effects of free trade zones on world trade

Conference participants disagreed about the effects of free trade zones on world trade. Paul Krugman argued that free trade zones will foster trade regardless of whether further progress is made at the global level. C. Fred Bergsten countered that free trade zones will impede world trade unless they are accompanied by further progress toward global trade liberalization.

Bad in theory, good in practice. Krugman acknowledged the solid theoretical arguments against free trade zones but still argued that such areas will probably help rather than hurt the world economy. Moreover, problems with the GATT negotiations are so deep seated that further progress is unlikely. As a result, regional free trade zones are a promising alternative to multilateral negotiations for promoting free trade.

Krugman's central point was that free trade zones are bad in theory but good in practice. He indicated free trade zones are bad in theory because they potentially divert trade from low-cost to high-cost suppliers. Trade diversion occurs when a member of a free trade zone imports a good or service from a country inside its zone rather than from a lower cost, nonmember country. He also indicated free trade zones can harm nonmember countries, not only by reducing the demand for their exports, but also by reducing the relative prices of their exported products. The decline in prices in nonmember countries relative to prices in member countries—a "beggar-thy-neighbor" effect—reduces nonmember country welfare. Additionally, he held that trade zones potentially impede trade by promoting trade warfare.

Krugman nonetheless argued that, in practice, free trade zones are likely to help more than they hurt the world economy, largely because they increase the size of markets. Larger markets lead to greater productive efficiency and competitiveness. Thus, trade zones are likely to create more trade than they divert.

Moreover, he stated that trade zones seem to be forming along "natural" geographic boundaries. Countries naturally tend to trade more with their neighbors than with distant countries because transporting goods and services and communicating over long distances is costly. As a result, free trade zones among neighboring countries may, in practice, be good for the world. The gains from freeing trade within regional zones will be larger and the costs of reducing trade across zones smaller than implied by moving to zones that are not based on natural geographic boundaries.

Finally, Krugman argued that moving toward global trade liberalization through the GATT process is hopelessly stalled, making free trade zones the only viable alternative. Among the reasons Krugman cited for the demise of the GATT are the decline of the United States as the principal world economic power, the increasing importance of such non-tariff barriers as domestic regulatory and investment policies, and the growth of new players in the world economy, such as the Japanese, who arguably play by a different set of rules.

Bad in theory, bad in practice. In sharp contrast to Krugman, C. Fred Bergsten claimed that moving toward free trade zones was bad in both theory and practice. Moreover, Bergsten argued that free trade zones are particularly bad when viewed as an alternative to further progress toward global free trade. Finally, Bergsten maintained that free trade zones need not be viewed as an alternative to globalism, because the GATT negotiations are still viable.

Bergsten cited a number of reasons to support his view that free trade zones are bad in practice. First, he argued that geography is not nearly as important as in the past as a determinant of "natural" trade regions. With technological advance, transportation and communications costs are no longer central to trading patterns. Second, while Europe and possibly North America may be "natural trading areas," no other such areas exist. Third, trade diversion may not be simply a consequence of trade zones but, in some cases, a goal. And finally, a Western Hemisphere free trade zone is likely to divert trade from lower-cost producers in Europe, Asia, Australia, and New Zealand to higher-cost producers in the Western Hemisphere.

Assuming that the movement toward free trade areas is likely to continue, Bergsten argued that the movement should occur in the context of an effective and credible global trade system. One way to ensure that the movement toward free trade zones supplements rather than replaces globalism is to enforce and expand the GATT. The GATT process can still work, according to Bergsten. Trade patterns in the Americas and in Asia remain "quintessentially multilateral." The markets of the three economic superpowers—Europe, Japan, and the United States—remain deeply intertwined. The superpowers have worked closely together on economic issues in the past and should be able to cooperate in the future. And although recent GATT negotiations have stalled, the GATT process has always been a messy one, filled with false starts and stops.

Are currency zones necessary?

In addition to differing on the net benefits of trade zones, conference participants expressed a range of views on currency zones. Although participants agreed that moving to currency zones will make it harder to conduct independent national monetary policies, they disagreed about whether this cost of currency zones exceeded the benefits. Martin Feldstein argued that currency zones are unnecessary and potentially harmful. Miguel Mancera argued that while the benefits of currency zones might be "impressive," floating exchange rates are more desirable. Other participants, including David Laidler, Michael Emerson, and Salvatore Zecchini, argued that currency zones might be beneficial to some, such as the Europeans, but not to others.

According to Feldstein, the cost of currency zones is high relative to their benefits. The primary economic benefit of currency zones is the boost to trade from eliminating uncertainty about exchange rate fluctuations. Exchange rate fluctuations inhibit businesses from importing inputs because unanticipated exchange rate movements in the wrong direction can potentially eliminate profits. Thus, eliminating exchange rate fluctuations would reduce uncertainty about the value of international transactions and, thereby, promote international trade. Feldstein argues, however, that these benefits are likely to be small. Econometric studies have failed to detect an adverse effect of exchange rate volatility on international trade. Moreover, businesses can hedge exchange rate risk through futures markets for foreign exchange.

In contrast, the costs of currency zones are possibly quite large. The primary economic cost of currency zones is the loss of independent national monetary policies. Under fixed exchange rates, central banks use the tools of monetary policy to keep exchange rates constant. As a result, these tools are unavailable for pursuing other national economic objectives. Under a single currency, countries surrender policy autonomy to a supranational monetary authority.

For example, with a freely floating currency, national monetary policymakers can counter a decline in the demand for a country's products by stimulating monetary growth and reducing interest rates. This response to a decline in demand is not possible if there are no national currencies or if exchange rates are irrevocably fixed. And without such a response, the output and employment costs of adverse demand shocks could be high.

Why then has Europe moved toward a currency zone? Feldstein argued that the reasons are more political than economic. Proponents of a currency zone believe a single European monetary authority could limit the ability of national governments to pursue inflationary monetary policies. More important, however, a single European currency would accelerate the political unification of Europe which, in turn, would result in greater centralization of fiscal policies.

Miguel Mancera took a more eclectic view of currency zones. Mancera recognized significant benefits from currency zones, including reduced investment risks, the equalization of interest rates across countries, and lower international transactions costs. Nevertheless, because inflation rates vary widely within and among countries, Mancera questioned the advisability of currency zones. Under floating exchange rates, a country can potentially insulate itself from inflationary shocks affecting other countries. In a currency zone, these shocks might spread to all countries. Mancera indicated that for this and other reasons Mexico could not possibly participate in a currency zone, although it probably will participate in a trade zone.

Other conference participants viewed currency zones somewhat more favorably, especially in the case of Europe. Salvatore Zecchini argued that a move to currency zones could be beneficial because without them businesses might face significant exchange rate risk. In contrast to Feldstein, Zecchini argued that futures markets in foreign exchange were too thin and underdeveloped to sufficiently reduce exchange rate risk. In addition, political institutions must be in place to ensure that smaller countries retain some influence over the policies of the trade or currency zone. This influence over policy should be viewed as compensation for the loss of political autonomy. While Zecchini felt that these conditions did not apply in North America, he felt they did apply in some of the countries of the European Community.

David Laidler viewed the formation of a currency zone as possibly good for Europe but definitely bad for North America. Like Feldstein, Laidler viewed the move toward currency zones as a political as well as an economic development. The move to either a common currency or irrevocably fixed exchange rates implies a loss of national sovereignty. Any move to give up national currencies must be viewed in part as a move toward political unity.

Although countries could maintain national currencies under a system of irrevocably fixed exchange rates, Laidler suggested that this form of currency zone also reduces political autonomy. The choice of an inflation rate is a political as well as economic decision. Moving to fixed exchange rates—or to a common currency—takes the issue of inflation out of the national political arena. It also removes from political accountability any national authority that might otherwise be responsible for a country's inflation performance.

Laidler added that while the move to a trade zone in Europe has been accompanied by closer political ties, no such political movement has occurred in North America. European countries have already surrendered considerable authority to European political entities, but no such surrender has occurred or is likely to occur in North America. Therefore, while a currency zone might work in Europe, it would not likely work in North America.

Michael Emerson agreed that while the political and economic prerequisites for a currency zone were probably in place in Western Europe, they are not well established in other regions of the world.[1] For example, before joining a trade or currency zone, the Eastern European countries must first join the world economy. They must adopt convertible currencies and world price structures. Only as a second stage of development can they consider regional trade and currency agreements. Even then they must work toward economic convergence with the rest of Europe before considering economic integration. Likewise, the USSR must grapple with its own problems of currency convertibility and determine whether its new federalist structure makes a compelling case for a currency zone. Finally, the Pacific region appears to be more interested in open trade on a global basis than in integration along economic, monetary, or political lines.

National Policy Implications of Trade and Currency Zones

The move toward trade and currency zones has implications not only for world trade in goods and services but also for national financial markets and macroeconomic structure. For example, financial markets within a trade zone may need to be harmonized so that capital, as well as goods and services, flows freely across countries. In addition, as monetary policy becomes more harmonized across countries in a trade zone, monetary policy will increasingly be determined at a supranational level. National fiscal policies could play a more important role in economic fluctuations at the national level and therefore may need to be harmonized to ensure fiscal discipline.

Financial market implications

Andrew Crockett and John Heimann examined four questions relating to the financial market implications of trade and currency zones. Do trade zones lead to increasing financial market integration across countries? Does economic integration lead to changes in financial market structure? What supervision and regulation will be required to ensure the efficiency and safety of financial markets? And, how will financial relationships across major trade zones be managed?

Financial market integration. Crockett argued that realizing the full benefits of trade zones requires liberalizing financial flows. As a result, trade zones create an incentive to liberalize finance. Removing international barriers to trade in banking, insurance, and other financial services results in greater specialization and competition in the supply of these services. As a result, costs of supplying financial services decline. By increasing competitive pressures, financial market liberalization also promotes productivity growth and innovation in the financial services industry. Finally, removing capital controls improves the flow of funds from savers to investors and channels investment funds to their most profitable use.

Liberalizing capital flows, in turn, requires closer harmonization of exchange rate policies. Large capital movements can undermine exchange rate stability. If capital liberalization leads to speculation and wide swings in exchange rates, it may undermine the benefits of trade zones. As a result, Crockett suggested that financial market liberalization may call for closer cooperation on exchange rate policies and, possibly, currency zones.

In discussing Crockett's paper, Heimann agreed that trade zones lead to financial market liberalization, which in turn leads to closer cooperation on exchange rate policies. Heimann also pointed out that these tendencies have been at work at the global level. In particular, as the G-7 countries have become more economically integrated, international capital flows have increased. At the same time, increasing speculation in capital markets has led to exchange rate volatility. This increased volatility of exchange rates underlies the management of exchange rates by the G-7 countries since the Louvre accord was reached in 1987.

Financial market structure. Crockett argued that financial market structures are likely to evolve slowly in response to freer capital flows. A variety of different structures coexist in the world today and freer financial markets are likely to have only a gradual effect in harmonizing these structures. While least-cost producers of financial services will tend to displace higher-cost producers, it is not clear that market structures will change dramatically.

Most studies of financial markets have shown that structure has little effect on efficiency. For example, economies of scale in financial services are small relative to the size of financial markets. As a result, many small firms can supply these services as efficiently as a few large firms. Thus, despite significant international differences in financial market structure, little movement toward homogenization can be expected in the short run. And, even in the long run, complete homogenization of financial markets is unlikely.

Heimann agreed with Crockett's assessment of the short-run effect of trade zones on financial market structure. Over the longer run, however, Heimann sees the financial system evolving into two tiers—global markets served by global institutions and regional and national markets served by regional and national institutions. This development represents the continuation of events that have been going on for years.

Regulatory and supervisory issues. Financial market regulation and supervision grows more complex as financial firms reach across national boundaries. Crockett gave three guiding principles for regulating and supervising financial markets in trade and currency zones. First, let financial institutions offer financial services throughout the trade zone. Second, issue firms a single license so they may operate freely across national boundaries. The license should be issued by either a supranational regulatory authority or, providing mutual recognition by other countries, by a national regulatory authority. Third, regulators should concern themselves more with harmonizing capital standards for credit institutions than with harmonizing market practices. Given limited information about the optimal structure of securities markets, alternative structures should be allowed to coexist and compete.

Heimann echoed Crockett's views on supervision and regulation. Specifically, Heimann argued for an "international supervisory system of harmonized standards" and urged regulators to closely supervise capital market activities.

Financial relationships between trade zones. With the world trading system moving toward several trade zones, Crockett suggested that negotiations between zones for market access in the financial sector will become increasingly important. Two main approaches are possible. The "mirror image" approach would require "identical conditions of establishment for financial institutions in different markets." In contrast, "national treatment" would require a zone to apply the same rules and regulations to all financial institutions within its borders—but the zone would not have to offer the same privileges and regulations as other zones. Of the two approaches, national treatment holds the greater promise as a basis for financial relationships between countries or trade zones.

In regulating market access, Heimann cited proposals for strengthening the regulatory system using three sets of regulations—home country, host country, and harmonized rule. These regulations underlie the principles of national treatment, mutual recognition, and effective market access.

Macroeconomic implications

Trade and currency zones have important consequences, not only for financial market policy, but also for domestic macroeconomic policies. Jacob Frenkel and Morris Goldstein, focusing on the implications of currency zones, discussed both monetary and fiscal policy. They argued that price stability is the appropriate goal of monetary policy and recommended a two-speed approach to currency union. They also stressed the importance of adopting mechanisms to ensure fiscal discipline. Michael Mussa and Tommaso Padoa-Schioppa, panelists in a session on macroeconomic policy implications, largely agreed with Frenkel and Goldstein on monetary and fiscal policy.

Monetary policy. Frenkel and Goldstein argued that the principal goal of monetary policy in a currency union should be price stability. In Europe and elsewhere a consensus has formed that only by achieving price stability can other goals of macroeconomic policy, such as high employment and economic growth, be achieved over the long run. This view has led to proposals that the monetary authority for the proposed European currency zone have an explicit mandate to pursue price stability as its primary goal. To ensure that the monetary authority carries out this mandate, the authority should have a significant degree of political independence and should be prohibited from issuing credit to the public sector.

Frenkel and Goldstein also addressed the issue of how countries in a trade zone should handle the transition to a currency zone. Frenkel and Goldstein recommended a two-speed approach in which one subgroup of countries takes a fast approach, while another subgroup takes a slow approach. Countries that have achieved low inflation rates and share other economic characteristics might move quickly toward a currency zone. Such a fast track approach would give "maximum credibility to exchange rate stability by eliminating exchange rates within the union," reduce or eliminate instability caused by capital mobility and divergent national monetary policies, and allow fast-track countries to realize all of the efficiency gains from having a single currency.

Countries with disparate economic performance would move more slowly toward membership in the currency zone. The slow approach would allow these countries to remain a part of the move toward monetary union without having to converge at a faster-than-desired pace to the economic performance levels of the fast-track countries. Thus, the two-speed approach would preserve momentum in the move to a currency zone.

While generally agreeing with Frenkel and Goldstein on the monetary policy implications of currency zones, Mussa emphasized the role of politics in determining monetary arrangements. Mussa argued that currency zones have historically been closely associated with areas of political authority. Thus, closer monetary ties and tighter exchange rate agreements come not just from a desire for greater economic unity but also from a desire for greater "political solidarity." The success of the European Community in establishing a currency zone, according to Mussa, depends more on the strength of shared political views than on a tally of economic costs and benefits.

In his discussion of monetary policy implications, Padoa-Schioppa emphasized monetary relationships between currency zones. Padoa-Schioppa argued that a European currency zone would lead to a "genuine multi-currency reserve system based on a tripolar relationship." Despite the fixity of exchange rates within currency zones, the exchange rate regime governing the three main reserve currencies—the dollar, yen, and European currency unit—should remain one of a mildly managed float.

Fiscal policy. Conference participants agreed that fiscal discipline was critical to the success of a currency union. Frenkel and Goldstein observed that, so far at least, moves toward currency union had not improved fiscal discipline in European countries. If sound fiscal policies are not forthcoming in a currency zone, the very objectives of the currency zone could be threatened.

Given the importance of sound fiscal policies, Frenkel and Goldstein described several mechanisms for ensuring fiscal discipline in a currency zone. One mechanism would be the marketplace itself. Member countries running excessive deficits with no recourse to finance deficits through money creation would face a rising default premium on government debt. The rising cost of government borrowing, along with reduced credit availability, would force governments to improve fiscal policies. Another mechanism would be fiscal policy rules. For example, rules might be enacted that place an upper limit on the size of budget deficits and government debt relative to GNP. Yet another mechanism would be peer group, multilateral surveillance. Under this mechanism, constraints on national fiscal policies would be more flexibly applied to discourage irresponsible fiscal policies of member countries.

Of these mechanisms, Frenkel and Goldstein prefer a combination of market discipline and peer-group surveillance. Given the right institutional setting, market discipline could be used as the primary mechanism to keep member countries' fiscal policies sound. Peer-group surveillance could be used as a supplement to encourage countries to solve pre-existing fiscal problems, preferably before they enter the currency zone. Peer-group surveillance could also be used to prevent "large fiscal policy excesses" in member countries.

Mussa agreed with this assessment of fiscal policy, but added that the most important mechanism for imposing discipline occurs when member countries get into fiscal crunches. At such times, both creditors and debtors need to know they will bear part of the cost of a financial crisis. Debtors must know they will bear a cost so that they will avoid irresponsible behavior. Creditors must know they will bear a cost so that they will "pull the plug" on excessive borrowing by the government.

Padoa-Schioppa went somewhat further in advocating the need for fiscal policy discipline. He argued that fiscal policy rules were desirable *per se* to reduce the budgetary discretion of member countries. He also argued that, in the case of the European Community, countries should give up some of their fiscal policy independence to a central fiscal authority. This transfer of responsibility should not take the form of Community control over national budgets, but rather the form of a more flexible use of the Community budget.

Global Implications of Trade and Currency Zones

Just as trade and currency zones will alter economic relationships within geographic regions, so will trade and currency zones alter relationships among regions of the world economy. One result of these changing relationships could be a tripolar monetary and trade system. Such a system could either enhance economic cooperation or foster hostile economic relationships among regions. This issue of a tripolar system was taken up by Allan Meltzer, Leonhard Gleske, and Kumiharu Shigehara. Related broad issues were addressed by Lawrence Summers, Jacques de Larosiere, Charles Carlisle, Pedro Aspe, Paul Volcker, and John Crow.

The emerging tripolar system

Meltzer argued that the world economy needs a new set of rules to maintain and enhance economic stability. Without new rules, the economic progress of the postwar period will not be sustained. Meltzer emphasized the importance of rules for maintaining trade and monetary stability.

Trade rules. Although the GATT remains in place, its rules are not being enforced. The lack of enforcement mechanisms has led to three responses. One response has been a move to managed, or "fair," trade in which producers form cartels to divide up markets for their products. Other responses include unilateral actions and bilateral and multilateral negotiations. But with the latest round of the GATT negotiations stalling, another mechanism has emerged—the move toward trade zones.

Meltzer argued that the development of trade zones is not a viable alternative to multilateral trade agreements, despite the failure of current GATT rules. With the formation of trade zones, trade *within* zones will increase at the expense of trade among zones. Grouping countries into three zones—Europe, the Americas, and Asia—Meltzer emphasized the importance of trade among zones. In the Americas and Asia, free trade among zones, or interzone trade, is greater than free trade within zones, or intrazone trade. Hence, developing intrazone trade "as a substitute for open, international trade" would not be in the interests of Japan and the United States. The European Community is the exception to this rule. Unlike the American and Asian zones, the European Community trades more within its zone than with the other two zones combined.

Shigehara shared Meltzer's concerns about the formation of trade zones. He suggested that the resulting industrial reorganization in Europe may be costly to firms outside of Europe. As bigger firms begin to exploit economies of scale, smaller firms will come under competitive pressure. As a result, European governments may attempt to keep high-cost firms in business by using protectionist measures against competing firms outside of Europe.

Monetary stability. Meltzer, Shigehara, and Gleske agreed that most countries will continue to rely on the dollar, mark, and yen as reserve currencies. Meltzer, however, emphasized that continued use of these currencies as major reserve currencies will require the United States, Germany, and Japan to keep price levels stable. If the United States maintains price stability, Meltzer believed the dollar would provide a store of value for many foreigners, remain the primary reserve currency, and continue to be used as the currency for pricing and purchasing commodities. Gleske agreed that, given domestic price stability, the dollar would likely remain the world's principal reserve currency.

Given price stability in the major world economies, Meltzer argued that a tripolar monetary system would provide international monetary stability. Countries with flexible exchange rates would experience greater stability of prices and exchange rates. Moreover, smaller countries could avoid inflation by fixing their exchange rates to one or more of the major reserve currencies.

Meltzer, Shigehara, and Gleske agreed that, while the European Community will probably form a currency zone, North America and Asia will not. European countries have more in common economically, socially, historically, and politically than do countries in Asia or North America. For example, Shigehara argued that in East Asia, countries were characterized by different stages of economic and financial development and different historical, cultural, and institutional backgrounds. These factors would limit the monetary integration of the Asian economies. Furthermore, Asian governments show little interest in relying on the yen as a reserve currency—the dollar still accounts for over half of the reserves of Asian governments. In addition, Shigehara argued that monetary union is a step toward political union, which is a goal in Europe but not in Asia.

Unlike North America and Asia, Europe is likely to adopt a currency zone. Gleske argued that Europe will benefit from this development. As Europe organizes its currency zone, Europe's real economy will become less susceptible to fluctuations in foreign exchange rates. The share of foreign trade in the "GNP" of Europe will fall sharply relative to the share of foreign trade in the GNP of many individual European countries. As a result, foreign exchange fluctuations will have less of an adverse effect on the European economy. In fact, the effect has already been reduced by the exchange rate mechanism and gradual stabilization of exchange relationships within the European Monetary System.

Overview remarks

Conference participants making broad overview comments expressed a range of views about the benefits of the move to trade and currency zones. Lawrence Summers and Jacques de Larosiere were optimistic about the trend. Charles Carlisle and Pedro Aspe had mixed feelings.[2] And Paul Volcker and John Crow were pessimistic.

The optimistic view. Summers argued that further progress was needed in liberalizing world trade. Toward that end, he supported any move to reduce barriers to trade, whether it be unilateral, bilateral, or multilateral. In particular, Summers said that most prospective trade zones were "likely to involve natural trading barriers and therefore to increase trade by more than they divert trade." And even if trade diversion occurs, it will be more likely to increase welfare rather than to reduce it. Moreover, trade zones will probably improve the domestic policies of member countries. And finally, trade zones could help accelerate the move to global trade liberalization.

De Larosiere, providing a European point of view, favored the move to trade and currency zones in Europe. He claimed that the move to a European trade zone has stimulated member countries' economic growth and trade. In the process, trade has increased not only among member countries but also with the rest of the world. As Europe has moved to a trade zone, exchange rates have stabilized, economic performance in member countries has converged, and monetary union now appears likely. Finally, de Larosiere argued that Europe's move to trade and currency zones does not imply isolation from the rest of the world. The European Community's economic integration will continue to benefit nonmember countries.

The mixed view. Carlisle argued that trade and currency zones could be either a positive or negative development. First, GATT statistics show that trade is not becoming more regionalized. Second, political realities make it unlikely that the world will coalesce into more than two great trade zones. Third, trade zones are not necessarily inconsistent with multilateral trade liberalization. Fourth, given that trade zones are going to develop, they must supplement, not replace, global trade liberalization. Finally, if trade zones replace global trade liberalization, all countries will be hurt.

Aspe agreed that membership in trade and currency zones could be extremely beneficial, especially to a small economy, so long as progress continues to be made at the global level. To this end, Mexico has joined the GATT and has expressed a willingness to join in various Western Hemisphere trade zones. Aspe argued that countries should be willing to act unilaterally, multilaterally, or as a part of a trade zone to reduce tariff and nontariff barriers to trade.

The pessimistic view. Volcker expressed concern about the trend toward trade zones. Siding with Bergsten, Volcker felt that regional trade zones would erect barriers to trade against the outside world and divert trade from nonmember countries. Moreover, he argued that trade zones could lead to greater interregional volatility in exchange rates. In response to the move to trade zones, Volcker suggested that Article 24 of the GATT, which restricts trade zones from taking protectionist actions, be more vigorously enforced. Although the article has been violated, particularly by the erection of nontariff barriers, remedial actions have not been taken.

Crow agreed with Volcker. Because of the dangers of trade and currency zones erecting protectionist barriers, Crow argued that further progress should be made on the GATT. Eastern Europe, the Soviet Union, and many developing countries are all striving to join the global trade system, and nothing should be done to prevent these emerging market-oriented economies from joining the GATT. In addition, Crow agreed with Meltzer that maintaining price stability is the best way to ensure the efficiency of world trade and payments.

Conclusions

The world economy may be moving toward trade and currency zones. Conference participants generally agreed that the move would be beneficial if it occurred along with further progress toward global trade liberalization. Participants also agreed that trade and currency zones would have profound effects on domestic financial, monetary, and fiscal policies and on trade and monetary relationships among regions of the world economy.

Conference participants disagreed about whether the move toward trade and currency zones would impede further multilateral trade liberalization or be beneficial without further multilateral progress. Participants also had different views about whether currency zones were necessary to achieve the full advantages of trade zones. From the discussions, though, it was clear that Europe would proceed toward establishing both trade and currency zones. Participants concurred that, of all of the proposed trade and currency zones, Europe is best suited to benefit from both.

Endnotes

1 Although Emerson was unable to attend the symposium, he contributed a paper.

2 Although Carlisle was unable to attend the symposium, he contributed a paper.

Policy Implications of Trade and Currency Zones

A symposium sponsored by the Federal Reserve Bank of Kansas City

August 22-24, 1991

Session I Chairman

ALAN GREENSPAN

Chairman, Board of Governors of the Federal Reserve System

The Move Toward Free Trade Zones

PAUL KRUGMAN, Professor of Economics, Massachusetts Institute of Technology

Commentary

C. FRED BERGSTEN, Director, Institute for International Economics

The Relationship Between Trade and Currency Zones

MICHAEL EMERSON, Ambassador to the USSR, European Community

MARTIN FELDSTEIN, President, National Bureau of Economic Research

DAVID LAIDLER, Professor of Economics, University of Western Ontario

MIGUEL MANCERA, Director General, Bank of Mexico

SALVATORE ZECCHINI, Assistant Secretary-General, Organisation for Economic Co-operation and Development

Financial Market Implications of Trade and Currency Zones

ANDREW D. CROCKETT, Executive Director, Bank of England

Commentary

JOHN HEIMANN, Chairman, Global Financial Institutions Group, Merrill Lynch & Co., Inc.

The U.S.-Canadian-Mexican Free Trade Agreement

PEDRO ASPE, Finance Minister of Mexico

Session II Chairman

JOHN W. CROW
Governor, Bank of Canada

Macroeconomic Policy Implications of Trade and Currency Zones

JACOB A. FRENKEL, Governor, Bank of Israel

MORRIS J. GOLDSTEIN, Deputy Director, Department of Research, International Monetary Fund

Commentary

MICHAEL MUSSA, Professor of Economics, University of Chicago

TOMMASO PADOA-SCHIOPPA, Deputy Director General, Bank of Italy

Global Implications of Trade and Currency Zones

ALLAN MELTZER, Professor of Political Economy and Public Policy, Carnegie-Mellon University

Commentary

LEONHARD GLESKE, Former Member of the Directorate, Deutsche Bundesbank

KUMIHARU SHIGEHARA, Director, Institute for Monetary and Economic Studies, Bank of Japan

Overview Panel

CHARLES R. CARLISLE, Deputy Director-General, The General Agreement on Tariffs and Trade

JACQUES DE LAROSIERE, Governor, Bank of France

LAWRENCE H. SUMMERS, Vice President and Chief Economist, World Bank

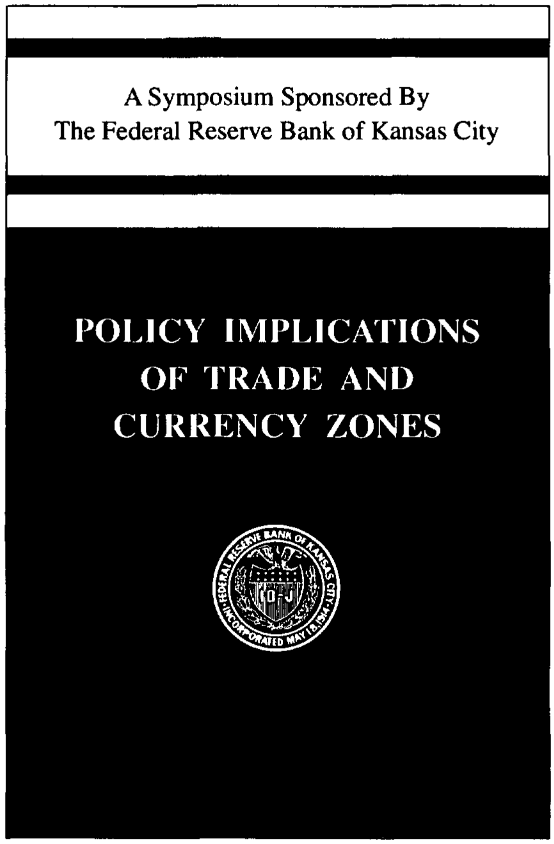
PAUL VOLCKER, Chairman, James D. Wolfensohn, Inc.

Policy Implications of Trade and Currency Zones

Is the world trading system moving toward free trade zones and currency zones? Would such zones usher in major changes in the international monetary system? To help policymakers examine these questions, the Federal Reserve Bank of Kansas City hosted a symposium on Policy Implications of Trade and Currency Zones, at Jackson Hole, Wyoming, on August 22-24, 1991. The symposium proceedings, available soon, discuss the move toward free trade zones and currency zones, their relation to one another, and their effect on the global economy.

For a copy of the current or past symposium proceedings, please write:

Public Affairs Department
Federal Reserve Bank of Kansas City
925 Grand Avenue
Kansas City, Missouri 64198



Past Symposiums

Central Banking Issues in Emerging Market-Oriented Economies (1990)

Monetary Policy Issues in the 1990s (1989)

Financial Market Volatility (1988)

Restructuring the Financial System (1987)

Debt, Financial Stability, and Public Policy (1986)

Competing in the World Marketplace: The Challenge for American Agriculture (1985)

The U.S. Dollar–Recent Developments, Outlook, and Policy Options (1985)

Price Stability and Public Policy (1984)

Industrial Change and Public Policy (1983)

Monetary Policy Issues in the 1980s (1982)

Check-Cashing Outlets in the U.S. Financial System

By John P. Caskey

John P. Caskey is a visiting scholar at the Federal Reserve Bank of Kansas City. The Russell Sage Foundation provided funding for this study. The views expressed in the article are solely those of the author and do not necessarily reflect the views of the Russell Sage Foundation, the Federal Reserve Bank of Kansas City, or the Federal Reserve System.

In the current debate over banking reform, some policymakers and consumer advocates have expressed concern that many lower-income Americans have lost access to basic payment services provided by banks. Reports of branch closings and increased service charges have led to proposals that banks be required to provide basic banking services to all consumers.

Most discussions of this issue are incomplete, however, because they overlook existing alternatives to banks for those who cannot or choose not to use banks to meet their payments needs. This article examines the role of check-cashing outlets (CCOs), a principal alternative to banks for many low and moderate-income consumers.[1] Despite evidence of rapid growth over the past decade, relatively little is known about the check-cashing industry. Understanding who uses CCOs and why provides new insight into the costs of payment services and adds a new dimension to the debate over basic banking services.

The first section of the article provides an overview of the check-cashing industry, including its services, fees, structure, and recent growth. The second section examines who uses CCOs and why, and offers possible explanations for recent growth. The final section addresses the regulation of CCOs and their possible role in providing basic banking services to low-income consumers.

An Overview of the Check-Cashing Industry

The check-cashing industry began in the 1930s as a response to banking problems during the Depression and to changes in employer payment practices. CCOs originally specialized in cashing payroll checks but over the years have evolved to provide a variety of payments services. Largely unregulated, the check-cashing industry has grown rapidly in the past decade, expanding beyond its traditional base in urban areas.

Services provided by CCOs

Nonfinancial businesses have cashed consumers' checks for many decades. Traditionally, this role was filled by bars, grocery stores, or other businesses that would cash third-party checks for regular customers or for customers making purchases. Such establishments rarely charged an explicit fee for cashing checks. The cost of the service was covered by the additional sales it generated.[2]

It is difficult to establish exactly when firms began to specialize in check-cashing and to levy a fee for the service. Most evidence suggests that CCOs evolved from other businesses that cashed checks on the side. CCOs apparently first appeared in Chicago and New York in the 1930s and spread to other large urban areas.

Most accounts cite widespread banking problems and changing employer payment practices as the principal factors motivating the early development of CCOs. For example, in Chicago, specialized check-cashing firms arose to provide payments services during the banking crisis of the 1930s (Illinois Department of Financial Institutions 1980). In addition, CCOs were stimulated by firms converting from cash payrolls to payroll checks during the 1930s and 1940s (Wolf).

The core business of a contemporary CCO is cashing checks for a fee. The fee is intended to provide the check-casher a profit after covering expenses, which include the cost of maintaining a storefront and insurance and personnel costs. Moreover, because the check-casher advances funds on checks that must subsequently be cleared through the banking system, CCOs incur interest expenses on the funds advanced. And, CCOs run the risk that some cashed checks will be uncollectible because of insufficient funds or fraud.[3]

Because of the risks associated with advancing money on checks, many outlets cash only customers' payroll or government entitlement checks. Some CCOs also cash personal checks but typically charge a higher fee for this service to cover the higher risk that the check will bounce. Many CCOs cash personal checks only after they have confirmed with the bank it is drawn on that there are sufficient funds.

In some states, CCOs make "payday" loans. They do this by cashing a customer's personal check, which is sometimes postdated, and agreeing to hold it until the customer's payday. Since this amounts to making an unsecured loan, check-cashers generally charge much higher fees for this service. It is generally offered only to customers with stable employment records who have maintained bank accounts in good standing for several months.[4]

While most CCOs derive most of their revenue from check-cashing fees, almost all CCOs do more than just cash checks.[5] They typically offer a range of financial and nonfinancial services—they may sell money orders, make wire transfers of cash, and handle telephone and utility bill payments. In some states, they sell lottery tickets and public transportation passes, offer income-tax preparation services, and distribute welfare payments and food stamps. In addition, many sell cigarettes and candy or buy and sell gold jewelry.

Fees charged by CCOs

CCO fees for cashing checks are usually expressed as a percentage of the face value of the check. In most states, check-cashers can charge whatever the market will bear; however, seven states currently set ceilings on check-cashing fees (Table 1).[6] As shown in the table,

Table 1

Maximum Check-Cashing Fees in Regulated States

(Rates are a percentage of the face value of the check)

State	Legal ceiling rate
Connecticut	2% for non-public aid checks and 1.0% for state public aid checks. (Ceiling fees set in 1990.)
Delaware	1% or $4.00, whichever is greater. (Ceiling fee set in 1989. The previous ceiling rate was 0.5% or $0.25.)
Georgia	The larger of $5.00 or 3% for public aid checks, 10% for personal checks, and 5% for all other checks (payroll). (Ceiling fees set in 1990.)
Illinois	1.2% plus $0.90. (Ceiling fee set in 1986. The previous ceiling rate was 1.1% plus $0.75.)
Minnesota	2.5% for public aid checks above $500 (5% for a first-time customer), no limit on personal checks but the rate must be filed with the state Commerce Department and be "reasonable," 3.0% on all other checks (6% for a first-time customer). (Ceiling fees set in 1991.)
New Jersey	1% for in-state checks and 1.5% for out-of-state checks or $0.50, whichever is greater. (Ceiling fees set in 1979. The previous ceiling rates were 0.75% on in-state checks and 1.0% on out-of-state checks, or $0.35.)
New York	0.9% or $0.50, whichever is greater. (Ceiling fee set in 1988. The previous ceiling rate was 0.75%.)

Source: State regulatory agencies.

the maximum permissible fee sometimes varies, depending on whether the check is drawn on an in-state or an out-of-state bank or is a government entitlement, payroll, or personal check. The different ceilings on fees across categories reflect the different speeds with which checks clear, different default risks, and the desire to limit the fees that public aid recipients pay for cashing their entitlement checks.

Outside of these seven states, commercial check-cashing fees vary widely. In 1989, the Consumer Federation of America (CFA) conducted a survey of the fees levied at check-cashing outlets in 20 major cities across the United States (Table 2). This survey suggests that CCOs charge roughly similar fees for payroll and government support checks.[7] For both types of checks, fees range from about 1.0 percent to 3.0 percent of the face value of the check, with an average rate of about 1.75 percent.[8]

About a third of the check-cashing outlets contacted by the CFA were willing to cash personal checks. Not surprisingly, given the default risk, they charge far more for this service. In the survey group, fees ranged from 1.66 percent to 20 percent of the face value of the check and averaged 7.7 percent.[9]

Table 2
National Check-Cashing Fees

Service	Minimum charge	Maximum charge	Average
Payroll checks	.9%	3.0%	1.74%
Government checks	.9%	3.25%	1.73%
Personal checks	1.66%	20%	7.7%
Money orders ($50)	$.19	$.99	$.55

Source: Consumer Federation of America (1989).

CCOs also levy fees for the other financial services they provide, such as selling money orders or making wire transfers. These services are largely used to pay bills by customers who do not have checkable bank deposits. The data suggest that many CCOs set low prices on these services. For example, the CFA survey found that the average charge for a $50 money order was $0.55, and many CCOs charged a flat fee independently of the size of the money order. This compares favorably to the $0.75 charged by the U.S. postal system for money orders up to $700.[10]

Structure of the industry

An examination of the structure of the check-cashing industry indicates commercial check-cashing is a relatively large industry, dominated mainly by local owner-operators. Historically, CCOs have been regulated extensively in only a few states. However, this picture is changing as national chains begin to develop and as more states consider regulating CCOs.

CCOs are currently regulated in only eight states. Seven states set ceilings on check-cashing fees and require that CCOs be licensed and abide by other regulations. These regulations generally require check-cashers to post their fees in a prominent location in the outlet and to provide customers with receipts. Often, the regulations require the CCO owner to meet a minimum bonding or capital requirement. Some states prohibit newly opened outlets from locating within a specified distance of existing CCOs. All states specify record-keeping requirements for the firms, and several of the states require check-cashers to report large sales of money orders or large wire transfers. This is to prevent check-cashing firms from being used in a money laundering process. Typically, the state banking department is responsible for issuing licenses and enforcing the regulations.

Because only a few states regulate the commercial check-cashing business, it is impossible to know exactly how many check-cashing firms are currently operating. However, across the United States there were 4,289 yellow-page listings of check-cashing firms in early 1991. This count is a lower-bound estimate of the total number of commercial check-cashing outlets nationally. In six of the eight states that require CCOs to be licensed, for example, the yellow-page count closely approximates the number of licenses outstanding. However, the yellow-page count understates the number of licensed outlets in New York by about 20 percent and by almost 50 percent in Georgia.

Given the sparse information on the industry, any estimate of the size of the industry in dollar terms is subject to a large margin of error. However, a conservative estimate indicates that the industry cashed about 150 million checks in 1990 with a combined face value of $45 billion. From this activity, the check-cashing industry earned approximately $790 million in fees.[11]

The vast majority of CCOs across the country appear to be owned by local independent operators, many of whom own three to ten outlets in a given area. There is evidence, however, that large national chains are developing. For example, one check-cashing company owns over 100 stores in the Northeast and is publicly traded on the over-the-counter stock market. And some check-cashing franchise operations have grown rapidly in the past few years. Recently, Western Union, which has provided money-wiring services to many check-cashers, announced plans to develop a national network of check-cashing outlets (*Wall Street Journal*).

The growth and location of the check-cashing industry

Data on the check-cashing industry are sparse but nevertheless indicate that the industry is growing rapidly. Moreover, the evidence suggests that the industry is beginning to expand beyond its traditional concentration in lower-income urban areas.

In interviews, check-cashers who have been in the business many years said that the industry grew slowly until the early or mid-1980s and then expanded rapidly. Unfortunately, there is not sufficient data to confirm this view.[12] However, American Business Information (ABI), a firm that tracks yellow-page listings of businesses, reported 4,289 listings of check-cashing (or currency exchange) outlets nationally in July of 1991. In 1987, the earliest year it provided data, ABI reported just 2,151 national listings. Thus, in four years, the industry appears to have doubled, a phenomenal growth rate.

Existing CCOs are disproportionately located in major urban areas, generally in low and moderate-income neighborhoods. For example, in eight states fewer than 10 percent of the CCOs are located in cities of less than 100,000.[13] The Illinois Department of Financial Institutions (1980, p. 107) reported that of 624 licensed check-cashers in the state in 1985, 90 percent were located in the Chicago area. And, a study for the New York State Banking Department found that 69 percent of all check-cashing outlets in New York City in 1990 were located in low-income census tracts (Kemlage and Renshaw).

The evidence suggests that the recent growth in CCOs has been uneven, with especially rapid growth outside of the few major urban areas where check-cashing establishments have long existed. For example, yellow-page listings from late 1988 to early 1991 show growth rates for Illinois, New Jersey, and New York of below 20 percent. Over that same period of time, the number of listed check-cashers grew by 85 percent in Florida, 195 percent in Georgia, 96 percent in Missouri, 293 percent in North Carolina, 80 percent in Texas, and 87 percent in Washington.

In states with early and well-developed check-cashing industries, recent growth has occurred mainly outside of the traditional inner-city areas. For example, the Illinois Department of Financial Institutions (1989, p. 5) reported that from 1985 to 1989, 108 new check-cashing licenses were granted but only 13 of these were for locations in Chicago; 75 were for locations in the Chicago suburbs and the remaining 20 were for downstate locations.

Explaining the Use and Growth of CCOs

Understanding the reasons behind the recent growth of check-cashing firms requires knowledge of who uses them and why. This section compares the cost and types of services offered by banks and CCOs, presents recent survey evidence on usage of CCOs, and examines factors behind their recent growth.

Comparing banks and CCOs

Since both banks and CCOs provide basic payments services, a key question is why consumers use CCOs rather than banks. One possible explanation is that CCOs are cheaper than banks. Or, perhaps CCOs are more convenient than banks or provide a type of service that banks are unable or unwilling to provide.[14]

The information on fees presented earlier can help provide an estimate of the cost to a household of meeting its payment needs through a CCO. For example, assume a family cashes its paychecks or government entitlement checks at a check-cashing firm charging a 1.5 percent fee and buys six money orders a month at an average price of $0.50 per money order. In this situation, a family with a $10,000 yearly income (about 75 percent of the 1990 official poverty level for a family of four) would spend $186 annually on basic financial transactions. Since check-cashing fees are a fixed percentage of the value of a check, a family with higher income would pay more. Thus, in this example a family with $24,000 annual income would spend $396 annually for financial services.[15]

The cost of obtaining similar services from a bank would be somewhat less, according to a 1990 national survey of bank fees by the Consumer Federation of America (1990). In estimating the cost of a checking account based on its survey data, the CFA assumed that a family maintains an average balance of under $400 in the account and that the account balance falls below $200 only once a month. In addition, the CFA assumed the family writes ten checks, makes four ATM withdrawals, and two deposits monthly and, over the year, the family bounces two checks and deposits one check that fails to clear. Based on this behavioral pattern, the CFA estimated that a family would pay $107.96 a year to maintain a noninterest-bearing checking account and would pay $111.39 a year to maintain an interest-bearing NOW account.

Regardless of the type of account maintained, it appears a family would save significant out-of-pocket costs by conducting its financial transactions through a bank rather than a CCO.[16] Because the fees for cashing checks at a CCO are assessed as a percentage of the face value of the check, the difference can be small for very low-income households. For example, a family earning $10,000 a year would save only about $80 annually by using a checking account rather than a CCO, while a family earning $24,000 a year would save almost $300. However, the very poorest households may be least able to afford the additional cost.

Two explanations account for the success of the check-cashing industry in the face of this cost disadvantage. One explanation is that out-of-pocket expenses do not measure the full cost of using a financial institution. Convenience, quality, and type of service also matter. In these aspects, CCOs may have an advantage for many consumers since most CCOs have much longer opening hours than do banks and are located more conveniently for some consumers. Also, CCOs may be faster with the range of simple financial transactions in which they specialize.

Another explanation for the success of CCOs is that bank services do not fully substitute for CCO services. Most important, while CCOs are willing to assume the risk that a check they cash will bounce, banks generally will not. Most banks require a consumer to maintain a deposit account in order to cash checks, even government checks with a negligible default risk.[17] For depositors, most banks require the customer either maintain sufficient funds in an account to cover the check or wait a few days for the check to clear. If the check fails to clear and the bank has cashed the check for a customer with sufficient funds to cover it, the customer's account is docked for the amount of the check. Moreover, many banks charge the customer for the bank's cost of handling a "returned" deposit.

Because of these differences in check-cashing policy, consumers without bank accounts may be forced to take their business to CCOs. Moreover, even if they maintain a bank account, consumers may not be able to cash a paycheck or government assistance check because the amount exceeds their account balance. Although these consumers could save money by depositing their check in a bank and waiting for it to clear, they may prefer to pay a fee to have the cash immediately.

Evidence on CCO use

Surveys of who uses commercial check-cashing firms and why they choose to do so suggest that most customers are either low-income to lower-middle income workers cashing payroll checks or recipients of government transfer payments. Relative to the population as a whole, a disproportionate percentage of CCO customers are young, nonwhite, and do not have bank accounts. Limited access to banking services and the convenience of CCOs appear to be the most important factors governing their use.

This profile of CCO customers is drawn from two recent surveys. One, a survey by the Consumers Banking Association (CBA), focused on consumers cashing paychecks. A second survey, conducted by the New Jersey Department of the Public Advocate, concentrated on those cashing public assistance and social security checks.[18]

The CBA survey found that CCO customers were younger and poorer than the general population and more likely to be a racial minority. Thirty-seven percent of respondents were between the ages of 18 and 30, and 29 percent reported a household income of less than $15,000 a year. The median reported household income in the survey was $20,400 as compared with a 1985 national median family income of $28,906. While 33 percent of respondents were white, 47 percent were black and 18 percent hispanic.

The survey found that customers' reasons for using a CCO revolved around their access to bank services. Two-thirds of customers surveyed had deposit accounts at banks or other financial institutions. Only 13 percent of these customers used CCOs regularly, citing convenience and ready access to cash. In contrast, the one-third of CCO customers without bank accounts made more regular use of CCOs. For those customers, lack of funds to maintain bank minimum balances and high bank service charges were cited as the main reasons for use of CCOs.[19]

The study by the New Jersey Department of the Public Advocate provides a somewhat different portrait of the customer base of the check-cashing industry because it focuses on those cashing public assistance and social security checks.[20] The Department interviewed 750 recipients of government transfer payments. In contrast to the CBA survey, 92 percent of those interviewed said that they did not have a bank account. Fifty-seven percent were cashing Aid to Families with Dependent Children (AFDC) checks. Another 20 percent, were cashing social security checks, and the rest were cashing unemployment benefits, veterans assistance, or state disability checks.

In the New Jersey survey, 79 percent of those interviewed stated that they never go to a bank to cash their government checks and, of these, 61 percent said they only go to CCOs. When asked why they were using a CCO to cash their government check, respondents cited lack of access to bank services and the convenience of CCOs.

Factors behind CCO growth

Knowledge of who uses CCOs and why is important for understanding the rapid growth in the industry during the 1980s. Changes in the economic situation of households may have led to an increased demand for check-cashing services. At the same time, regulatory changes may have increased the cost of banking services.

One factor contributing to the growth of CCOs may have been the strong growth in payroll employment following the 1982 recession. From 1983 to 1989, total civilian employment increased 16 percent (*Economic Report of the President*). Unlike the economic expansions of the 1960s and 1970s, however, employment growth in the 1980s was accompanied by a fall in employees' real incomes. For example, average weekly earnings of private sector, nonagricultural, industrial workers fell from $408 in 1978 to $346 in 1990.[21] Because the customer base of CCOs is disproportionately low-wage and moderate-wage workers, lower real incomes may have contributed to the demand for CCO services.

More generally, the 1980s saw a fall in the standard of living for many low-income families. From 1979 to 1988, the mean real family income of families in the lowest income quintile fell 5.4 percent (Bradbury, p.26). And, the number of families falling below the poverty line rose from 24.5 million in 1978 to 31.9 million in 1989 (*Economic Report of the President*). To the extent that poorer families had increased difficulty in accumulating financial savings to maintain bank balances, they may have had an increased incentive to use CCOs.

The 1980s also saw changes in the cost and supply of banking services. In 1980, the federal government enacted the Depository Institutions Deregulation and Monetary Control Act. Among other things, this act began a phaseout of ceilings on the interest rates banks could pay on deposits. The Act also required the Federal Reserve System to begin charging banks for a number of services it had previously provided for free.

Another factor was a change in the attitude of bank regulators at the federal and state levels toward competition among banks. Prior to 1980, regulators often looked unfavorably on a proposed branch that would be located in a community already well-served by other bank branches. However, after 1980, in an atmosphere much more favorable to free-market competition, regulators began to consider the increased competition provided by an additional community bank to be a positive factor in approving new bank branch applications (Spong).

Following these changes, banking became a much more competitive business. Banks reacted by pricing services based on the costs of providing those services. Thus, they began to charge for accounts with high transactions volume and small balances, significantly raising the cost of using banks for many low and moderate-income consumers (U.S. GAO). Bankers also reacted to the increased competition by closing branches in unprofitable or marginally profitable areas, which were often low-income areas, and opening branches in the more desirable, higher-income areas already served by other banks.[22] Combined, these changes worked to make banks both more expensive and less convenient for many low-income and moderate-income consumers, and likely contributed to a growing demand for commercial check-cashers' services.

Finally, the rapid growth in the check-cashing industry in the 1980s may have been stimulated by an increased awareness of the market potential of the millions of Americans who do not regularly use the banking system for their financial transactions. Beginning in the mid-1980s, journalists, academics, and policy analysts began to write about bank closings in low-income neighborhoods and the large number of households not using banks.[23] These reports may have captured the imagination of entrepreneurs and fed the expansion of nonconventional financial institutions serving those whose needs were poorly met by banks.

Public Policy Issues

Recognizing that CCOs are playing a more important role in the U.S. financial system raises a number of public policy issues concerning CCOs and the delivery of affordable financial services to low-income households. This section considers the trade-offs in regulating CCOs and the role they could play in the financial system.

Regulation of CCOs

Bank closings in low-income communities, increases in bank fees on small deposit accounts, and the rapid growth of the check-cashing industry have made the policies of CCOs far more relevant than the policies of banks for many segments of the population.

This observation has led to suggestions that the check-cashing industry be more widely regulated. Those advocating that more states, or perhaps even the federal government, should regulate the industry point out that many check-cashing customers are relatively unsophisticated consumers, with little social or economic power. These customers might be grossly overcharged by an unscrupulous operator, some of whom may have local monopoly power. Thus, there is concern that many poor and moderate-income individuals could spend a large percentage of their limited disposable incomes for basic financial transactions.

Indeed, evidence supports the concern that some check-cashing firms levy relatively high fees. For example, the survey by the Consumer Federation of America (1989) found that 11 percent of the firms charge 3 percent or more for cashing government entitlement checks. In New Jersey, for example, check-cashers are limited by law to charging 1.0 percent on in-state checks and 1.5 percent on out-of-state checks. Of 662 customers there who reported the amount of the check they cashed and the amount of fee they paid, 49 percent were charged more than the legal maximum (New Jersey Department of the Public Advocate, p. 29). On average, check-cashers overcharged by about 44 percent of the ceiling rate, and in some cases the excess charge was substantial. To cite two examples from the report: a Hispanic woman who could not speak English was charged $25 for cashing a $268 social security check, and another woman was charged $16 for cashing her $525 AFDC check.[24]

Interestingly, in its response to the study by the Department of the Public Advocate, the New Jersey Department of Banking, which oversees check-cashing outlets, reported that it had received only one check-cashing complaint over two years (GAO, p. 9). It appears, therefore, that the vast majority of people who were charged more than the legal maximum in New

Jersey did not complain to the oversight agency, perhaps because they were unaware of the overcharge, felt a complaint would be ineffective, or did not know how to file an official complaint or felt that the effort was greater than the cost of the overcharge.[25]

Those who favor limits on check-cashing fees need to be aware of possible consequences, however. Mandating very low check-cashing fees could kill the industry and hurt the low and moderate-income people who have no realistic alternatives for cashing their checks. Prior to 1989, for example, Delaware limited check-cashing outlets to charging a fee of 0.5 percent of the face value of the check or $0.25, whichever was greater. In 1989, the state raised the limit to 1.0 percent or $4.00, whichever is greater, noting that no CCOs were operating in the state under the old law.

On the other hand, it is clear that CCOs can flourish in urban areas when the ceiling rate is around 1.0 to 2.0 percent.[26] In New York, for example, the ceiling rate is 0.9 percent or $0.50, whichever is greater. Yet over 400 check-cashing outlets operate in the state. Illinois, which permits check-cashers to charge up to 1.2 percent of the face value of the check plus $0.90, has more CCOs per capita than any other state.[27]

The evidence suggests, therefore, that if regulation of CCOs is deemed desirable, states can set limits on check-cashing fees to protect consumers against the highest charges and yet permit the industry to flourish. The evidence from New Jersey also suggests, however, that the state must devote resources to enforcing compliance with the statute. In New York and Illinois, where the state banking departments conduct annual on-site surveys of CCOs, firms do not appear to charge more than the legal maximum. Presumably, annual license fees from CCOs can provide the states with the revenue to cover the costs of monitoring the industry and enforcing state legislation.

The role of CCOs in the financial system

The 1980s have seen increased emphasis on the access lower-income households have to affordable basic financial services. Legislatively, this concern has been expressed in congressional hearings or proposals to force banks to cash government entitlement checks for free and to offer "basic," or "life-line," bank accounts (U.S. Senate, U.S. House 1989).[28] Such accounts would permit a consumer to conduct a limited range of basic financial transactions for a very small fee or no fee. Regulators and community activists have also used the Community Reinvestment Act and other means to bring pressure on banks to keep branches open in low-income areas and to improve banking services in these communities.[29]

However, the possible cost or effectiveness of these proposals has also caused concern. For example, if banks are forced to provide these services without sufficient compensation, the burden might not be shared equally among banks. Indeed, banks with existing branches in low-income areas could be most affected. Moreover, imposing such policies on banks but not their competitors could place banks at a competitive disadvantage and, perhaps, lead to an acceleration of bank branch closings.

Recognition of the growing importance of CCOs, however, suggests that they might play a role in providing basic financial services to low-income households. CCOs specialize in delivering a narrow range of payments services. With experience, they have learned which financial services are most in demand by lower-income households and have learned to minimize the cost of providing these services. CCOs already compete for locations that are most convenient for the low-income and moderate-income households that make up their customer base.

By viewing CCOs as an integral part of the financial system, federal, state, and local governments may be able to work with them to ensure that they deliver affordable basic payments services. Indeed, a number of states already appear to be taking this approach, using CCOs in the distribution of public benefits and services. For example, residents of New York City and Chicago can elect to receive their AFDC payments or food stamps through local CCOs. In New York, the state pays the CCO to distribute AFDC benefits in cash. In Illinois, the CCOs handle the distribution of AFDC checks for free, but if the recipients cash their checks at the CCO, they pay the regulated state fee. And, in Illinois, many CCOs have the right to handle automobile registrations and title transfers.

The suggestion that CCOs be used as delivery points for government services is linked with the view that they be more widely regulated. This is true for two reasons. First, in a state where CCO fees and services are regulated, the industry is likely to have a better public image and therefore is more likely to be trusted for distributing public services. Second, because permitting CCOs to distribute AFDC payments, handle automobile registrations, or provide other public functions is profitable for CCOs, such opportunities can be traded for lower ceilings on the fees CCOs levy for basic financial services.

Realistically, however, advocating broader regulation and reliance on CCOs for the delivery of basic financial services does not require abandoning efforts to improve the accessibility of banks for lower-income households. While CCOs provide some basic payment services, they are not substitutes for banks. CCOs do not take deposits, so residents of a community served only by CCOs would not have a safe and convenient outlet for their savings. And, CCOs also do not make loans, so the economic development of a community served only by CCOs may suffer.

Summary

This article has surveyed the role check-cashing outlets play in the financial system. CCOs provide basic financial transaction services to many low-income and moderate-income households. And, measured by the number of outlets, CCOs may be the most rapidly growing segment of the financial system. Households that consistently use CCOs appear to devote a larger fraction of their incomes on average to pay for financial transactions than do families that rely on banks. Some use of CCOs appears to be voluntary. Consumers may turn to them rather than to banks because CCOs have a more convenient location or longer hours of operation. However, some consumers may turn to CCOs because they cannot afford to meet minimum balance requirements at banks.

For many moderate-income and low-income households in urban areas, a CCO may be the most important financial institution in their daily lives. This observation has led an increasing number of states to regulate CCOs and suggests that CCOs might be employed in the delivery of basic financial services and government benefits.

Endnotes

The author would like to thank Jennifer Ekert and Bart Yavorosky for excellent research assistance. The three of us are grateful to the commercial check-cashers and state regulators who took the time to explain the business and its regulatory environment to us. The article also benefited from conversations about this research with Gordon Sellon, Hyman Minsky, and my colleagues at Swarthmore College. The author is also grateful to Kristin Lewis of American Business Information for providing some of the data used in this article.

1 In Indiana, Illinois, Minnesota, and Wisconsin, firms that cash customers' checks for a fee are said to be in the "currency-exchange" business. The more widely used term "check-cashing" business is used to avoid confusion with foreign exchange transactions.

2 For lack of data, the article does not attempt to examine recent trends in check-cashing by these nonfinancial businesses. It also excludes from the analysis mobile payroll services.

3 As is clear from this explanation, all check-cashing outlets must work closely with at least one bank. This is because a CCO needs a bank to clear the large volume of checks the firm cashes. Moreover, most CCOs rely on bank lines of credit to meet their periodic, substantial needs for cash.

4 In the few states that regulate the check-cashing business, it is illegal for check-cashers to make such loans. In some unregulated states, payday loans are effectively illegal because the fees violate state usury laws.

5 Data from the Department of Financial Institutions in Illinois show check-cashing firms earn about 67 percent of their revenue from check-cashing fees and about 11 percent from sales of money orders.

6 Other states partially regulate the industry or have legislation pending. For example, Wisconsin has long required check-cashers to be licensed but does not otherwise the state regulate check-cashers' activities. Washington state recently established extensive regulations of the check-cashing industry that will take effect in 1992, but the regulations do not set ceilings on check-cashing fees. At the time of this writing, Ohio and Pennsylvania have regulatory legislation pending. Legislation to regulate the industry was also recently introduced in a few other states, but failed to pass. Illinois and New York were the first to establish such regulations, enacting legislation in 1943 and 1944, respectively. Delaware and New Jersey began to regulate CCO fees in the 1950s, and in the past two years, Connecticut, Georgia, and Minnesota have also done so.

7 It is also common for check-cashing firms to levy additional charges for first-time customers. Check-cashers say these charges are to cover the costs of issuing the customer an identification card or registering the customer.

8 The CFA survey suggests that CCOs charge slightly more for cashing AFDC (welfare) checks then for social security checks.

9 A 1991 telephone survey, by the author, of 42 check-cashing firms in several states found fees broadly agreeing with those found by the Consumer Federation of America. In the unregulated states, most firms charged between 1.5 and 3.0 percent to cash government and local payroll checks. Three outlets charged rates as high as 5 to 6 percent. Those that accepted personal checks charged from 4 to 15 percent. A small number of the firms permitted a customer to cash a post-dated personal check. For a check that was to be held up to one month, the customer typically was charged 20 to 35 percent of the amount advanced.

10 Check-cashers want to promote money order sales because a check-casher selling numerous money orders will not need to use as much of his own capital or tap a relatively expensive bank credit line to obtain cash for check-cashing customers. The check-casher simply hands out the cash he receives from selling the money orders. In addition, check-cashers can earn float (i.e., interest on money being transferred to someone else) from money order sales, for the check-casher normally pays the money order company with a slight delay (Gagerman).

11 In arriving at these estimates, it is assumed that there were 4,250 check-cashing outlets operating in 1990, each cashing an average 35,000 checks. This estimate of the average number of checks cashed is below the scale of operation of most check-cashing outlets in Illinois, New Jersey, and New York, as reported by the regulatory agencies in those states. However, outlets in these three states must do a greater volume of business than the national average to survive because these states have regulated fees lower than those charged elsewhere. Interviews with check-cashers in the unregulated states suggest most outlets handle between 25,000 and 40,000 checks annually.

These estimates also assume that the average check has a face value of $300 and the average cashing fee is 1.75

percent. The $300 estimate is consistent with the data for Illinois, New Jersey, and New York and was thought reasonable by check-cashers in the unregulated states. The 1.75 percent fee agrees with the national average reported by the Consumer Federation of America (1989).

[12] Data are available for Illinois, New Jersey, and New York, but the trends in these states may well have been affected by unique factors. For example, in both Illinois and New York there was a sharp increase in the average annual growth rate in the number of licensed check-cashing outlets in the second half of the 1980s as compared to the first half of the decade. However, both of these states in the second half of the 1980s raised the ceiling on the fees check-cashers were allowed to charge. Moreover, New York, at the end of 1985, stopped considering distance between competing check-cashing locations as a factor in approving applications for licenses (Renshaw, p. 8). In New Jersey, the number of licensed check-cashing outlets grew strongly throughout the 1980s, rising from 69 in 1980 to 88 in 1989. However, the growth in the early part of the decade may have been aided by a 1979 increase in the fee check-cashers in New Jersey could charge. Finally, the trends in these states are unlikely to be nationally representative because Illinois, New Jersey, and New York, unlike almost all other states, have had well-developed check-cashing industries for over 40 years. In fact, the study by Reeb and others concludes that check-cashing in New York City is a mature industry with limited future growth possibilities for its core services.

[13] This result is based on the author's survey of CCOs in eight states.

[14] CCOs might also be used by those who do not want to create deposit-account records because of tax reasons, immigration status, etc.

[15] This example assumes no taxes or withholding. The low-income family pays $150 ($10,000 x .015) for check-cashing and an additional $36 for money orders. The moderate-income family pays $360 for check-cashing and $36 for money orders.

[16] While this example appears to be based on reasonable assumptions, other assumptions could change relative costs. For example, since the CFA study found that banks charged $15.11 per bounced check, the cost of using a bank would increase if the family's account were overdrawn more frequently.

[17] In 1988, the Consumer Federation of America (1988) surveyed 110 banks and 84 thrifts located primarily in the urban areas of 15 states and the District of Columbia. It found that of the 191 financial institutions responding to the survey, 71 percent would not cash government checks for nondepositors at any price. Fourteen percent would cash non-depositors' government checks for free, and 15 percent would do so for a fee, averaging $3.88 for a $300 check. Outside of urban areas, banks are apparently more willing to cash government checks for nondepositors (U.S. GAO 1988, pp. 13-14).

The study (GAO 1988, pp. 16-17) suggests that banks that refuse to cash government checks for free for non-depositors do so because banks incur costs in handling checks, they do not want to crowd their lobbies with government aid recipients who only want to cash their entitlement checks, and they fear that some fraudulent checks might be cashed for which the government would not reimburse them.

[18] There are several reasons that neither the Consumer Bankers Association's survey nor the New Jersey Department of the Public Advocate's survey is alone likely to be broadly representative of the customer base of the check-cashing industry. The Department of the Public Advocate survey limited its study to the use of CCOs by recipients of government aid programs and ignored people cashing payroll checks. In the case of the Consumer Bankers Association (CBA) survey, customers who visit a CCO during a heavy payroll period are unlikely to be representative of customers generally; that is, they are more likely to be employed and have higher education and income levels. They are probably also more likely to maintain a deposit account.

[19] For additional evidence on reasons consumers may not use banks, see Canner and Maland.

[20] According to the data in Appendices P through S of the study, CCOs in three New Jersey counties (Camden, Essex, and Mercer counties) cashed about 1.5 million checks in 1986, about 13 percent of which were AFDC checks. By examining 4,842 canceled AFDC checks from three counties, the Department found that 47 percent of them were cashed at banks, 32 percent were cashed at CCOs, 12 percent were cashed at local businesses, and 9 percent were cashed by friends, relatives, or landlords. Of the AFDC checks cashed at banks, 75 percent were cashed at banks that serve as depositories of county funds and are required to cash AFDC checks for nondepositors without a fee.

[21] Both figures are expressed in 1990 dollars.

[22] For evidence on branch closings, see Obermiller, and Avery.

[23] For example, the U.S. GAO (1988, p. 19) estimated that about 16 million American families did not have banking accounts of any type in 1985. Also see the articles by Canner and Maland, Gross, Zamba, Lueck, Obermiller, and Bartlett.

[24] In a survey, by the author, of 42 check-cashing outlets

across several states, a few charged 5 to 6 percent to cash government and payroll checks. When asked why competition would not drive firms that charge more out of business, check-cashers said many of their customers just want their money as fast as possible and pay no attention to a difference of a few percentage points in the fee charged. In addition, customer transportation costs may limit competition among check-cashing outlets.

[25] The New Jersey Department of Banking told the New Jersey Department of the Public Advocate (p. 68) in 1987 that it relied on the "honor system" to assure compliance with state limits on check-cashing fees. A 1991 telephone survey, by the author, indicated that check-cashing firms in the state are now complying with the law, perhaps because the Department of Banking increased the resources it devoted to enforcement after the report by the Department of the Public Advocate.

The author called several other state consumer advocate agencies and state banking departments to find out if there had been complaints against check-cashing outlets. In no state was this the case. However, in unregulated states, it was often difficult to locate anyone in a state agency who knew where one would go to file such a complaint or how it would be classified by the consumer advocacy agency.

[26] If outlets are to cash very small checks or personal checks, a higher fee may need to be permitted in these cases.

[27] Other states should not automatically assume they can adopt the New York or Illinois ceilings on check-cashing fees without adversely affecting the industry. Both of these states use check-cashing outlets to distribute welfare payments, which brings additional business to the outlets. Also, in both states, check-cashing outlets are almost exclusively found in the dense urban areas. States with less concentrated populations may find check-cashing firms cannot function profitably with a 1.0 percent ceiling.

[28] Some policy analysts have also suggested reviving the U.S. postal savings system to ensure all communities have convenient access to a deposit-taking financial institution. In fact, perhaps a major reason check-cashing outlets do not exist in Europe is because most European countries have postal savings systems with giro accounts.

[29] See the 1986 policy statement on basic banking by The Federal Financial Institutions Examination Council in Canner and Maland. In 1989, federal financial institution regulators made provision of basic banking services a part of a bank's CRA rating.

References

Bartlett, Sarah. 1989. "Bank Closings Discriminate, Report Asserts," *New York Times*, January 30.

Bradbury, Katharine L. 1990. "The Changing Fortunes of American Families in the 1980s," *New England Economic Review*, July/August.

Canner, Glenn and Ellen Maland. 1987. "Basic Banking," *Federal Reserve Bulletin*, April.

Consumer Bankers Association. 1989. "Check Cashing Services Study," monograph of study conducted by the Roper Organization Inc., December.

Consumer Federation of America. 1987. "National Survey of Check Cashing Outlets," September.

____________. 1988. "Bank Fees on Consumer Accounts: The Fifth Annual National Survey," manuscript, June.

____________. 1989. "Check Cashing Outlet Fees: Still High and Climbing," December.

____________. 1990. "Ten Years after Financial Deregulation: The Sixth Annual National Bank Fee Survey," manuscript, June.

Economic Report of the President. 1991. U.S. Government Printing Office.

Gagerman, Jerome S. 1990. "Professional Check Cashers and the Compliance Implications of the Money Laundering Control Act," manuscript, August 1.

Gross, Laura. 1987. "Branch Cuts Seen Hurting Minorities," *American Banker*, February 23.

Illinois Department of Financial Institutions. 1980. *Annual Report*.

____________. 1989. *Annual Report*.

Kemlage, Donald J. 1991. "More about Branch Banking and Check-Cashing Outlets in New York State," Appendix G in Reeb et al, *Economic Profile of the Check Cashers' Industry*, A Report to the New York State Banking Department, May.

____________, and Edward Renshaw. 1991. "Branch Bank Closings in Low-Income Areas of NYC and Check-Cashing Outlets," Appendix F in Reeb and others. *Economic Profile of the Check Cashers' Industry*, A Report to the New York State Banking Department, May.

Lueck, Thomas. 1988. "Banks Shut in Poor Areas Stir Worries," *New York Times*, August 17.

New Jersey Department of the Public Advocate. 1988. "Who's Checking?" January.

Obermiller, Phillip J. 1988. "Banking at the Brink: The Effects of Banking Deregulation on Low-Income Neighborhoods," *Business and Society*, Spring.

Reeb, Donald and others. 1991. *Economic Profile of the Check Cashers' Industry*, A Report to the New York State Banking Department, May.
Renshaw, Edward. 1991. "Regulation of the Dedicated Check-Cashing Industry," Appendix A in Reeb and others. *Economic Profile of the Check Cashers' Industry*, A Report to the New York State Banking Department, May.
Spong, Kenneth. 1990. *Banking Regulation: Its Purposes, Implementation, and Effects*, Federal Reserve Bank of Kansas City.
U.S. General Accounting Office. 1987. "Banking Services: Changes in Fees and Deposit Account Interest Rates Since Deregulation," July.
____________. 1988. "Government Check Cashing Issues," October.
U.S. House of Representatives. 1989. "Consumer Access to Basic Financial Services," Hearings before Subcommittee on Consumer Affairs and Coinage of the Committee on Banking, Finance, and Urban Affairs, October 17.
U.S. Senate. 1989. "Government Check-Cashing, Lifeline Checking and the Community Reinvestment Act," Group before The Subcommittee on Consumer and Regulatory Affairs of the Committee on Banking, Housing, and Urban Affairs, Law G-7.
Wolf, Irving J. 1975. *The Licensed Check Cashing Industry In New York City*, MBA thesis at Pace University, September.
Zamba, Michael J. 1987. "Closed Banks Worry Neighbors," *Christian Science Monitor*, May 8.

1991 Index
Economic Review

Economic Review

November/December 1991, Vol. 76, No. 6

The *Economic Review* (ISSN0161-2387) is published bimonthly by the Federal Reserve Bank of Kansas City, 925 Grand Avenue, Kansas City, Missouri 64198-0001. Subscriptions and additional copies are available without charge. Send requests to the Public Affairs Department, Federal Reserve Bank of Kansas City, 925 Grand Avenue, Kansas City, Missouri 64198-0001. Second-class postage paid at Kansas City, Missouri.

POSTMASTER: Send address changes to *Economic Review*, Public Affairs Department, Federal Reserve Bank of Kansas City, 925 Grand Avenue, Kansas City, Missouri 64198-0001.

Economic Review
Federal Reserve Bank of Kansas City
Kansas City, Missouri 64198-0001
November/December 1991, Vol. 76, No. 6