

FILE COPY

New England Economic Review

Federal Reserve Bank of Boston

May/June 1992

Norman S. Fieleke

*One Trading World, or Many: The Issue
of Regional Trading Blocs*

*Joe Peek and
Eric S. Rosengren*

The Capital Crunch in New England

*Richard E. Randall
and Richard W. Kopcke*

*The Financial Condition and Regulation of
Insurance Companies: An Overview*

Peter Fortune

*The Municipal Bond Market, Part II:
Problems and Policies*

President

Richard F. Syron

**Senior Vice President and
Director of Research**

Alicia H. Munnell

*Research Department:***International**

Norman S. Fieleke, Vice President
Jane Sneddon Little

Monetary and Financial

Richard W. Kopcke, Vice President
Richard E. Randall, Vice President
Eric S. Rosengren, Vice President
Katerina Simons
Robert Tannenwald

National Business Conditions

Stephen K. McNees, Vice President
Joe Peek
Geoffrey M. B. Tootell

Regional

Lynn E. Browne, Vice President and
Deputy Director of Research for
Regional Affairs
Katharine L. Bradbury, Assistant Vice
President
Yolanda K. Henderson

The *New England Economic Review* is produced in the Research Department and is edited by Joan T. Poskanzer. Heather Schofield coordinates the design and production. The views expressed are those of the individual authors, and do not necessarily reflect official positions of the Federal Reserve Bank of Boston or the Federal Reserve System. The authors will be glad to receive comments on their articles. The Review is available without charge. Requests to be placed on the mailing list or for additional copies should be sent to the Research Library-D, Federal Reserve Bank of Boston, P.O. Box 2076, Boston, MA 02106-2076. Articles may be reprinted if the source is credited. Please send copies of the reprinted materials to the editor.

Contents

One Trading World, or Many: The Issue of Regional Trading Blocs

Norman S. Fieleke

Over the past several decades, more and more countries have entered into preferential trading arrangements, provoking concern that the benefits of free trade are being sacrificed to growing discrimination. Just how widespread is this discrimination in international trade, and is it "legitimate" under the codes of international behavior to which countries generally subscribe? What does economic theory tell us about the likely consequences of such discrimination, and why do so many nations engage in it?

The author finds that most of the preferential trading arrangements, accounting for about two-thirds of world trade, have increasingly resembled "trading blocs," in that their trade has become oriented more inward, among the members, and less outward, with the rest of the world. Over the long run, he points out, nondiscriminatory reductions in trade barriers are clearly preferable to discriminatory reductions. But should global negotiations fail, blocs that truly liberalized trade among themselves could improve the general welfare. To set the best example for the rest of the trading world, they should be receptive to new members, for the ideal free trade area is worldwide in scope. 3

The Capital Crunch in New England

Joe Peek and Eric S. Rosengren

The increase in real estate lending was a major reason for the rapid expansion of New England banks during the 1980s. When nominal real estate prices began to decline in New England, collateral became impaired and many loans stopped performing. The consequent increased provision for expected loan losses (loan loss reserves) caused a rapid deterioration in bank capital throughout the region.

Having just lost a significant proportion of their capital, many banks tried to satisfy their capital/asset ratio requirements by shrinking their institutions. This article discusses why banks facing binding capital constraints will shrink more than unconstrained banks when an adverse capital shock occurs. It shows that New England banks with low capital/asset ratios are in fact shrinking their institutions faster than better capitalized institutions, and that this behavior has been particularly apparent in those liability categories that are marginal sources of funds for most banks. 21

Contents

The Financial Condition and Regulation of Insurance Companies: An Overview

*Richard E. Randall and
Richard W. Kopcke*

The Municipal Bond Market, Part II: Problems and Policies

Peter Fortune

In October 1990 questions were raised about real estate problems in the life insurance industry after the ninth largest life company sustained a major loss as a consequence of a write-down of real-estate-related assets. The value of insurance company stocks declined as the financial community began to take a hard look at the recent changes that had taken place. During the spring of 1991 the press increasingly focused on the industry, once it became evident that the life subsidiaries of First Executive and First Capital were impaired as a consequence of substantial investments in junk bonds.

In the summer of 1991, the Federal Reserve Bank sponsored a conference to examine the dramatic changes that have transformed the seemingly stable insurance industries into industries that could arouse public anxieties. How pervasive are the weaknesses that have shown up in a few large insurers? Is there a danger that widespread liquidity pressures could develop? What changes should be made in regulation or in arrangements to protect customers of insurance companies? These are some of the primary questions addressed in the conference proceedings. 32

Why does Congress allow municipal interest payments to be exempted from federal income taxes in the face of a very large chronic deficit in the federal budget, even though no constitutional provision requires that this tax policy continue? The rhetoric of tax exemption is philosophical, appealing to notions of appropriate intergovernmental relations and, in particular, to the doctrine of reciprocal immunity: no level of government should use its taxing authority to impose harm on another level.

But the true force behind tax exemption is that it provides states and local governments with a valuable subsidy, which can be enjoyed at their discretion. This article identifies the problems posed by tax exemption, and assesses some alternatives. It asks whether the results of tax exemption represent an appropriate outcome, and questions whether tax exemption is really necessary to achieve the benefits stated in its favor. 47

One Trading World, or Many: The Issue of Regional Trading Blocs

Over the past several decades, more and more countries have entered into preferential trading arrangements, provoking concern that the benefits of free trade are being sacrificed to growing discrimination. Just how widespread is this discrimination in international trade, and is it "legitimate" under the codes of international behavior to which countries generally subscribe? What does economic theory tell us about the likely consequences of such discrimination, and why do so many nations engage in it? Are patterns of trade being seriously distorted by the emergence of discriminatory "trading blocs"? The answers offer little indication that the sky is falling, but neither do they provide grounds for complacency.

The Prevalence of Preferential Trading

To our knowledge, no comprehensive compilation of preferential trading arrangements has previously been published. And the listing in Table 1 may be less than complete, although we invested much research in its preparation. The list is formidable. Included in the 23 arrangements identified are 119 countries, accounting for some 82 percent of the world's international trade in goods. No region is free from such arrangements; indeed, one would be hard-pressed to find even one country that does not receive from, or grant to, other countries some form of explicit preferential treatment in international trade, although that treatment might take some mode other than participation in a multilateral arrangement such as those included in the table.

Preferential trading arrangements take several forms, but all favor the trade of the participants over that of nonparticipating countries. In the most casual arrangement, the trade preference association, each member establishes lower governmental barriers against imports of goods from other members than against comparable imports from

Norman S. Fieleke

Vice President and Economist, Federal Reserve Bank of Boston. Kenneth S. Neuhauser provided research assistance.

Table 1
Preferential Trading Arrangements, by Geographic Region and Year Launched

Region, Title and Membership	Year Launched	Type of Trade Arrangement
<u>Africa:</u>		
<u>Communauté Economique de l'Afrique de l'Ouest (CEAO), or West African Economic Community:</u> Benin, Burkina Faso, Côte d'Ivoire, Mali, Mauritania, Niger, Senegal	1959	Customs union
<u>Union Douanière et Economique de l'Afrique Centrale (UDEAC), or Economic and Customs Union of Central Africa:</u> Cameroon, Central African Republic, Chad, Congo, Equatorial Guinea, Gabon	1964	Customs union
<u>Southern African Customs Union (SACU):</u> Bophuthatswana, Botswana, Ciskei, Lesotho, Namibia, South Africa, Swaziland, Transkei, Venda	1969	Customs union
<u>Mano River Union (MRU):</u> Guinea, Liberia, Sierra Leone	1973	Customs union
<u>Economic Community of West African States (ECOWAS):</u> Benin, Burkina Faso, Cape Verde, Côte d'Ivoire, The Gambia, Ghana, Guinea, Guinea-Bissau, Liberia, Mali, Mauritania, Niger, Nigeria, Senegal, Sierra Leone, Togo	1975	Common market
<u>Preferential Trade Area for Eastern and Southern African States (PTA):</u> Burundi, Comoros, Djibouti, Ethiopia, Kenya, Lesotho, Malawi, Mauritius, Mozambique, Rwanda, Somalia, Swaziland, Tanzania, Uganda, Zambia, Zimbabwe	1981	Trade preference association
<u>Communauté Economique des Etats de l'Afrique Centrale (CEEAC), or Economic Community of Central African States:</u> Burundi, Cameroon, Central African Republic, Chad, Congo, Equatorial Guinea, Gabon, Rwanda, Sao Tome and Principe, Zaire	1981	Common market
<u>Arab Maghreb Union (AMU):</u> Algeria, Libya, Mauritania, Morocco, Tunisia	1989	Common market
<u>Asia:</u>		
<u>Association of Southeast Asian Nations (ASEAN):</u> Brunei, Indonesia, Malaysia, Philippines, Singapore, Thailand	1967	Free trade area
<u>Bangkok Agreement:</u> Bangladesh, India, Laos, South Korea, Sri Lanka	1976	Trade preference association
<u>Europe:</u>		
<u>European Community (EC):</u> Belgium, Denmark, France, Germany, Greece, Ireland, Italy, Luxembourg, Netherlands, Portugal, Spain, United Kingdom	1957	Common market
<u>European Free Trade Association (EFTA):</u> Austria, Finland, Iceland, Liechtenstein, Norway, Sweden, Switzerland	1960	Free trade area
<u>European Community and European Free Trade Association:</u> Member countries of the EC and EFTA	1972	Industrial free trade area
<u>Latin America:</u>		
<u>Central American Common Market (CACM):</u> Costa Rica, El Salvador, Guatemala, Honduras, Nicaragua	1960	Customs union
<u>Andean Common Market (ANCOM):</u> Bolivia, Colombia, Ecuador, Peru, Venezuela	1969	Common market
<u>Caribbean Common Market (CARICOM):</u> Antigua and Barbuda, Bahamas, Barbados, Belize, Dominica, Grenada, Guyana, Jamaica, Montserrat, St. Kitts-Nevis, St. Lucia, St. Vincent and the Grenadines, Trinidad and Tobago	1973	Common market

Table 1 *continued*

Region, Title and Membership	Year Launched	Type of Trade Arrangement
<u>Latin American Integration Association (LAIA):</u> Argentina, Bolivia, Brazil, Chile, Colombia, Ecuador, Mexico, Paraguay, Peru, Uruguay, Venezuela	1980	Trade preference association
<u>Organization of Eastern Caribbean States (OECS):</u> Antigua and Barbuda, Dominica, Grenada, Montserrat, St. Kitts-Nevis, St. Lucia, St. Vincent and Grenadines, Virgin Islands U.K.	1981	Customs union
<u>Southern Cone Common Market (MERCOSUL or MERCOSUR):</u> Argentina, Brazil, Paraguay, Uruguay	1991	Common market
<u>Middle East:</u>		
<u>Gulf Cooperation Council (GCC):</u> Bahrain, Kuwait, Oman, Qatar, Saudi Arabia, United Arab Emirates	1981	Common market
<u>Middle East and Africa:</u>		
<u>Arab Common Market (ACM):</u> Egypt, Iraq, Jordan, Lebanon, Libya, Mauritania, Syria	1964	Common market
<u>North America:</u>		
<u>Canada-United States Free Trade Agreement:</u> Canada, United States	1989	Free trade area
<u>Oceania:</u>		
<u>Australia-New Zealand Closer Economic Relations Trade Agreement (ANZCERT):</u> Australia, New Zealand	1983	Free trade area
<u>Other:</u> Other preferential arrangements include various bilateral free trade agreements, such as between Israel and the EC, Israel and the United States, and Chile and Mexico, and also preferential treatment for imports from less developed countries by many countries, including the EC and the United States.		

Source: See the Appendix.

nonmember countries. In the free trade area, members go a step further and completely eliminate governmental barriers against goods imports from other members, but, as in the trade preference association, maintain their individual barriers against imports from nonmembers. Establishment of a customs union requires that members not only eliminate government barriers against merchandise imports from one another, but also establish identical barriers—in particular, a common tariff barrier shared by all—against imports from nonmembers. Finally, the customs union becomes a common market with the removal of artificial or governmental impediments to all transactions between members, including transfers of labor, capital, and services as well as goods.

Is It Legal?

A code of law for international trade is set forth in the General Agreement on Tariffs and Trade (GATT), which is applied as a treaty obligation among countries that subscribe, or “contract,” to the Agreement. In addition, GATT entails an organization, or forum, in which countries discuss and negotiate issues of international trade, such as the multilateral liberalizations under consideration in the current Uruguay Round. GATT entered into force in 1948, with 23 original contracting parties, a number that had grown to 103 by November 1991, with an additional 29 countries applying the agreement de facto (GATT 1980, 1992).

The very first paragraph of the very first article of

the General Agreement lays down a broad prohibition against the use of preferential tariff rates: "With respect to customs dues . . . any advantage, favor, privilege or immunity granted by any contracting party to any product originating in or destined for any other country shall be accorded immediately and unconditionally to the like product originating in or destined for the territories of all other contracting parties." (GATT 1969, p. 2) This language gives expression to the unconditional most-favored-nation (MFN) principle, the principle that each contracting party must grant to every other contracting party treatment as favorable as it grants to any country. Long before the GATT, this principle of nondiscrimination had been generally observed by many nations.

However, the very second paragraph of this same first article allows an exception to this principle, permitting the continued application of many preferential tariff rates that were in effect at the time the GATT was adopted. And a much more significant exception is to be found in Article XXIV, which spells out the conditions under which GATT signatories may form customs unions and free trade areas (and interim arrangements leading to them). The formation of such preferential trading arrangements is allowed as long as the following conditions are met: (1) trade barriers are eliminated on substantially all trade among members; (2) the trade barriers remaining against nonmembers are not higher or more restrictive than those previously in effect (in the case of a customs union, not "on the whole" higher or more restrictive); and (3) interim arrangements lead-

No region is free from preferential trading arrangements, favoring the trade of the participants over that of nonparticipating countries.

ing to the free trade area or customs union are employed for only a reasonable length of time.

Provided these three rather ambiguous conditions are satisfied, arrangements such as those listed in Table 1 do not violate the legal obligations assumed by members of the GATT. Nonmembers are free to discriminate without satisfying any such conditions (unless they have limited that freedom through other treaty obligations), according to the prevailing inter-

pretation of international law (Jackson 1989, p. 134). Of the countries participating in the discriminatory arrangements identified in the table, only 18 neither belong to GATT nor apply its provisions de facto.

Under GATT it has been possible to rationalize, if not justify, the proliferation of preferential trading arrangements because of the ambiguous language used in specifying the three conditions that these arrangements must meet. Phrases such as "substantially all trade," "not on the whole higher or more restrictive," and "reasonable length of time" have allowed much latitude for interpretation. This latitude has been exploited. The GATT has been notified of more than 70 preferential trading arrangements—some establishing very loose preferences as "interim agreements" with no date for completing a free trade area—but GATT has never formally disapproved any of them (OECD 1990, p. 18; Jackson 1989, p. 141).

Although preferential arrangements of the sort indicated in Table 1 are the subject of this article, it should be noted that they are by no means the only mode of discrimination in international trade. One variant close to the types listed in the table is the granting of preferences to imports from less developed countries under schemes such as the Generalized System of Preferences (GSP), which was sanctioned by a waiver of the GATT most-favored-nation principle from 1971 to 1981 and thereafter by an international agreement (Jackson 1989, p. 141; Carl 1986, pp. 7–8). Another prominent form of discrimination is the widespread use of so-called "voluntary" export restraints, under which a country agrees under pressure to limit its exports of a certain good to a particular importing country or countries. The restraints imposed by Japan over its automobile exports to the United States are a well-known example. Other forms of significant discrimination could be adduced. All in all, more than one-fourth of world trade fails to observe the MFN principle, according to one recent estimate (Kostecke 1987, pp. 425 ff.).

Does It Matter?

Few words arouse more revulsion than "discrimination." But is discrimination in international trade a harmful practice? And if so, why is it so prevalent?

Perhaps the strongest case ever made against discrimination in international trade is Jacob Viner's classic work, *The Customs Union Issue* (1950), in which Viner focused on the effect of customs unions. His conclusion was that, "with respect to most customs

union projects the protectionist is right and the free trader is wrong in regarding the project as something, given his premises, which he can logically support" (p. 41). Viner based this conclusion on the belief that formation of a preferential trading area such as a customs union would usually shift purchases of traded goods primarily from lower to higher money cost sources (excluding tariffs), rather than from higher to lower cost sources.

In briefest outline, the underlying reasoning is as follows. Once the customs union has been completed, members will import from one another some

Viner believed that the formation of a preferential trading area would usually shift purchases of traded goods primarily from lower to higher money cost sources (excluding tariffs).

commodities that previously they did not import at all because of their tariffs (now eliminated on trade among the members). This "trade creation" is efficient and desirable, as it entails a shift from a higher cost (domestic) source to a lower cost (foreign) source. But because the union maintains a tariff barrier against imports from nonmembers, the members will also now import from one another some commodities that previously they had imported from nonmembers who had supplied the items at the lowest cost including the tariffs then levied (at rates that were the same for both members and nonmembers). This "trade diversion" from a lower cost (non-member) source to a higher cost (member) source is inefficient and undesirable. It was Viner's opinion that in the construction of the typical customs union high priority would be given to protecting domestic industries, so that trade diversion would outweigh trade creation. He believed this undesirable outcome would be even more likely in the case of preferential trading arrangements short of full customs union, on the grounds that the participants would select preferences that were predominantly protective of their own industries.

To this case against preferential trading arrangements, Viner recognized one significant qualification.

Within the enlarged protected market formed by such an arrangement, the expansion of output in various industries might be accompanied by economies of scale, with lower costs per unit of output, and this gain in efficiency might offset the losses stemming from net trade diversion. This beneficial outcome he considered unlikely, arguing that in most industries plants can attain their most efficient scale even if the industry is not large.

Viner also appreciated that by adopting a common tariff against imports from nonmembers, and by negotiating as a unit on trade issues with the rest of the world, the countries forming a customs union could exercise greater economic leverage than if they acted individually. Thus, the union might find it possible to improve the terms on which it traded with other countries. Any such gain for the union would, of course, represent a loss for the rest of the world.

While a number of significant refinements have been made in Viner's analysis, the concepts of trade creation, trade diversion, and economies of scale have remained central to empirical studies of preferential trading arrangements.¹ Empirical studies are very important, because both Viner's and subsequent theorizing have made one thing clear, namely, it is not possible to say a priori that customs unions—or, more generally, preferential trading arrangements—will inevitably either enhance or diminish world efficiency or world welfare. Each case must be examined carefully on its own merits. This is not to argue that preferential trading arrangements may be more efficient than perfectly free trade among all countries. But given that trade is less than free, the formation of a preferential area may or may not represent an improvement.

Thus, it is not only the GATT, but economic theory as well, that is somewhat ambiguous on the issue of preferential trading areas. It would be wrong to conclude from this shared ambiguity that GATT was based on the theorizing outlined in this section, as that theory blossomed only after GATT was founded. In the drafting of the GATT, economic theory probably was less influential than the harsh lessons of the years between the world wars, a period that witnessed a proliferation of bilateral and other discriminatory arrangements inimical to world trade and responsible for worsening both the Great Depression and international relations. In terms of sub-

¹ For developments in theory following Viner's work, see Gunter (1989); Wonnacott and Lutz (1989); Kowalczyk (1990); and the references cited in these works.

sequent economic theorizing, the three conditions required of preferential trading areas under GATT Article XXIV can be viewed as tending to restrain the trade-diverting aspects of such schemes while encouraging their trade-creating aspects. Of course, as with any agreement, the GATT is no better than its interpretation and enforcement.

Empirical Analyses

Since empirical analysis is crucial to evaluation of a preferential trading area, what has been revealed by such studies? Are they helpful in resolving the ambiguity of theory?

In recent years surprisingly little of the empirical research on preferential trading areas has provided estimates of trade creation and diversion. What is not surprising is that such estimates have been chiefly for the European Community (EC), the major customs union to develop after World War II. These EC estimates generally agree that for manufactured products trade creation exceeded trade diversion. On the other hand, EC policy toward agriculture has been highly protectionist and has likely generated substantial trade diversion. Thus, one recent survey concludes, "it is not obvious from . . . the empirical studies whether the volume of trade created outweighed that of trade diverted, whether there was any *external* trade creation by which *non-members* benefited from the increased EC market size, or whether the customs union among the original EC members improved global allocative efficiency" (Pomfret 1988, p. 131, emphasis added).

Similar conclusions are drawn regarding research into two other effects noted by Viner, namely, effects on economies of scale and terms of trade: "the available evidence shows that increased scale economies have been realized in some EC sectors since the establishment of the customs union, but gives little indication of a causal relationship or of the magnitude of any allocative efficiency gains Finally, although it is widely agreed that the EC customs union has involved terms of trade effects, there are few estimates of their magnitude In sum, the EC customs union seems to have involved small (and perhaps even negative) static welfare gains, possible but unproven dynamic benefits, and a welfare transfer from non-members" (Pomfret 1988, pp. 133–135).

Thus, empirical analysis has done little to resolve the ambiguity of theory on the question of whether preferential trading areas such as the EC serve to

enhance or diminish global efficiency and welfare. If in fact the benefits are so dubious, why are such arrangements so widespread? Before tackling this puzzle, we should note that by no means have all preferential trading arrangements turned out to be viable. Indeed, even some of those identified in Table 1 are, at this writing, more nearly nominal than functional. Because of this mixed record, much attention has been given to the question of what promotes the viability or demise of preferential arrangements, especially customs unions and free trading areas. A review of the conclusions is not only interesting in its own right but also sheds some light on the puzzle of why so many of these arrangements are launched.

This research on the viability of customs unions and free trade areas has yielded somewhat more definitive results than has research on trade creation and diversion and on welfare effects. The arrangements that endure and that seem to foster trade among the participants often display the following characteristics:²

- (1) the member countries have relatively similar levels of per capita GNP and relatively similar economic structures;
- (2) the member countries have compatible laws and policies governing international trade flows and adopt an across-the-board rather than product-by-product approach toward liberalizing trade among themselves;
- (3) the member countries are not located vast distances apart.

It is one thing to observe these characteristics and another to explain why they seem to promote the success of a preferential arrangement. Presumably, similar levels of per capita GNP reduce the likelihood of disagreements over trade flow adjustments that generate unemployment or, more generally, that redistribute income. Firms in poorer countries fear the superior technology, managerial skills, and capitalization of firms in richer countries while workers in the richer countries fear the competition of lower-paid laborers in the poorer countries. These fears, and the disputes they provoke, may be less intense among countries with similar levels of development. It also happens that such countries commonly trade more heavily with each other, even in the absence of preferential understandings, than do countries of widely differing economic structure and per capita income.

² See Schott (1991, pp. 2–3); Wonnacott and Lutz (1989, pp. 74–83); Thoumi (1989); and Carl (1986, pp. 13–34).

In the realm of government policy, two points seem especially relevant. First, conflicting national laws and policies toward international commerce obviously inhibit its development. Second, barriers to trade between the members of a preferential arrangement are generally reduced more expeditiously if the successive reductions are applied to virtually all barriers, across the board, rather than if reductions of varying degrees are negotiated for different products. Successful opposition by affected interest groups is more likely to arise against reductions proposed product by product than those undertaken across the board.

With respect to the last of the characteristics common to successful preferential arrangements, the lesser is the distance between countries, the lower are the costs of transportation and communication that encumber their trade, other things equal. Other things are not always equal, of course, and geographic proximity is no guarantee of success. But customs unions or free trade areas are seldom, if ever, even attempted among nations that are poles apart.

The characteristics that make for success are clearly present in the case of the largest and longest-lived of the customs unions, the European Community (EC). Near the other extreme are the less successful Latin American Integration Association (LAIA)—and especially its predecessor, the Latin American Free Trade Area (LAFTA)—and the Economic Community of West African States (ECOWAS). Progress within ECOWAS has been impeded by the participation of its members in other preferential trading arrangements that impose obligations incompatible with those assumed under ECOWAS. The resulting inconsistency of the members' commercial policies has contributed to virtual paralysis in the mutual reduction of trade barriers (Agyemang 1990, esp. pp. 67 and 79). In the case of LAFTA, the product-by-product negotiating approach posed a formidable obstacle. The same approach has slowed the reduction of trade barriers within the Association of Southeast Asian Nations (ASEAN).³ More generally, progress within a number of integration schemes launched by less developed countries, such as LAFTA, has been slowed by disagreements over how the perceived gains and losses might be distributed, and over compensation to be provided by countries that gain to poorer, losing countries.⁴ These disagreements, of course, illustrate the likelihood of friction where per capita incomes and economic structures differ very much among the member countries.

Overall, preferential trading arrangements have a mixed record, with few approaching their announced goals; and even for the more successful arrangements it is hard to prove demonstrable benefits for the world at large or even for the members. Such being the case, why are these arrangements so common? A definitive answer is elusive, but the motivations behind many arrangements are fairly readily discerned.

Why So Many Preferential Trading Arrangements?

A Vinerian answer to this question might run along the following lines. Trade-diverting preferential arrangements are promoted by the producers who benefit, while the consumers who are injured offer little opposition. The explanation for this activism on the part of producers and passivity on the part of consumers is that the gains from diversion reaped by protected industries are concentrated enough to inspire them to lobby for preferential arrangements, while the injury done to consumers is spread too

*Motives besides protectionism—
some economic, others more
nearly political—lie behind the
formation of preferential trading
arrangements.*

thinly among them to provoke their strong protestations.

This rather traditional explanation of protectionist successes no doubt has some validity, but it is only part of the story. Other motives besides vulgar self-interested protectionism are also at work. Some are economic in nature, others more nearly political. A political motivation played an important role in the formation of the EC. It was hoped that economic integration would strengthen political ties among the West European countries and reduce the likelihood of

³ See Carl (1986, pp. 15–17); Wonnacut and Lutz (1989, pp. 74–77); and Balasubramanyam (1989, pp. 173–74).

⁴ See Carl (1986, pp. 16, 21–22, 28); Wonnacut and Lutz (1989, pp. 82–83); and Pomfret (1988, pp. 145–47).

⁵ See, for example, MacBean and Snowden (1981, pp. 145–46).

conflicts among them, such as had led to World War II, while also providing a democratic bulwark against the Soviet communist bloc.⁵ The prospective expansion of the EC to include some of the recently liberated countries of Central and Eastern Europe is similarly justified on the grounds that their inclusion will enhance the stability of their struggling new market economies and democratic institutions.

Another motivation at least partly political in nature is frustration with the slow advance of trade liberalization under GATT-sponsored negotiations. A highly influential example of this glacial pace is the current Uruguay Round—negotiations that were launched by more than 70 nations in September 1986 and that, at this writing, have yet to produce an overall agreement. One reason for this lack of progress is that trade negotiations these days are tackling more complex issues, such as intellectual property rights, dispute settlement techniques, nontariff barriers, and trade in services. Not surprisingly, agreements on matters of this nature are more readily reached among relatively few countries whose relevant policies are already fairly similar, a fact that inclines countries to strike a (preferential) bargain with just a few other countries rather than endure the lengthy and dubious multilateral GATT negotiations.

Still another reason that some countries opt for preferential rather than MFN agreements conducted through GATT is to avoid giving a "free ride," in the form of liberalizing or market-opening measures, to countries that fail to reciprocate (Belous and Hartley 1990, p. 13). For example, several countries might agree to extend to each other's banking firms the right to open branches within each other's borders, but be reluctant to extend the same right to banks in another country that did not offer that right.

Among less developed countries an important goal of preferential trading arrangements may be to reduce dependence on industrial countries, especially as a source of manufactured goods, by fostering trade and integration among the parties to the arrangement. A closely related goal may be to enhance the bargaining power of the parties vis-à-vis the industrial countries (Agyemang 1990, pp. 57–58).

Finally, countries may seek inclusion in preferential trading arrangements not because they expect significant gains but to avoid losses from the trade diversion to which they would be exposed as outsiders. Even if a preferential arrangement creates more trade than it diverts, the nonmembers experience only diversion—a reduction in their exports to the members—unless, over the long run, the arrange-

ment serves to accelerate the economic growth (and demand for imports) of the member countries. In the case of a customs union, such as the EC, another potential loss to be avoided through joining is the less favorable terms of international trade that the union, to its own advantage, may be able to impose upon the outside world. For these and other reasons, countries often fear the consequences of being "left out" of preferential arrangements. And countries that are left out may form their own arrangements, partly in "self-defense" and partly merely in imitation of other arrangements; the power of example can be powerful indeed.

Are Discriminatory Blocs Capturing and Transforming the World's Trade?

Table 2 reports, in order of magnitude, the percentage shares of world merchandise trade attributable to preferential trading groups identified in Table 1. As indicated in the first row of data, countries belonging to these arrangements account for more than 80 percent of all international trade. And, as noted at the bottom of Table 1 and elsewhere in this article, these arrangements by no means exhaust the catalog of preferential trading schemes in use.

The EC's share of world trade is much larger than that of any other trading group. Because the EC and EFTA have formed a free trade area for manufactured goods, the Canada-United States Free Trade Area is an even more distant second than indicated by the percentages reported. While some of the trading groups account for inconsequential shares of world trade, the aggregate for all the groups is most impressive.

But are these trading groups turning into trading blocs that promote trade among themselves at the expense of trade with the rest of the world? The mere finding that members of a group trade more intensively with each other than with other countries would not establish that the group had "turned inward." For one thing, the members are often closer geographically to one another than to nonmember countries and would be expected to trade more intensively, other things equal, if only because of lower transportation and communications costs. Moreover, if the members enjoy more rapid growth in production and overall trade than the rest of the world, trade among them would ordinarily be expected to grow more rapidly than their trade with the rest of the world.

Table 2
Preferential Trading Groups' Percentage Shares of World Merchandise Trade, 1989 and 1990

Trading Group	Exports		Imports		Total Trade	
	1989	1990	1989	1990	1989	1990
Total of countries in trading groups listed below	79.93	81.26	82.34	82.90	81.16	82.10
European Community (12)	38.99	40.96	38.89	41.04	38.94	41.00
Canada-United States Free Trade Area	16.63	15.70	20.44	18.54	18.56	17.14
European Free Trade Association	6.41	6.71	6.47	6.60	6.44	6.65
Association of Southeast Asian Nations	4.17	4.22	4.22	4.65	4.20	4.44
Latin American Integration Association	3.46	3.47	2.45	2.61	2.95	3.03
Bangkok Agreement	2.72	2.47	2.88	2.85	2.80	2.66
Gulf Cooperation Council	2.31	2.57	1.45	1.48	1.87	2.02
Australia-New Zealand Closer Economic Relations Trade Agreement	1.58	1.45	1.79	1.50	1.68	1.48
Southern Cone Common Market	1.54	1.42	.87	.85	1.20	1.13
Arab Common Market	.92	.95	.99	.95	.95	.95
Arab Maghreb Union	.79	.93	.83	.84	.81	.89
Andean Common Market	.85	.90	.59	.55	.72	.72
Southern African Customs Union ^a	.78	.76	.63	.53	.71	.65
Economic Community of West African States	.68	.68	.53	.58	.61	.63
Preferential Trade Area for Eastern and Southern African States	.24	.22	.35	.34	.30	.28
Caribbean Common Market	.15	.16	.25	.24	.20	.20
Economic Community of Central African States	.21	.22	.16	.16	.19	.19
West African Economic Community	.17	.16	.19	.18	.18	.17
Central American Common Market	.16	.14	.21	.20	.19	.17
Economic and Customs Union of Central Africa	.15	.18	.11	.11	.13	.14
Mano River Union	.10	.08	.12	.15	.11	.12
Organization of Eastern Caribbean States ^b	.01	.01	.03	.03	.02	.02

Note: The trade (exports, imports or total trade) for each group is the sum of the trade of the individual members of the group, including trade with countries both within and outside the group. Trading groups' shares of world trade do not sum to the total for all countries in the listed groups (first line), as some countries are members of more than one group and data for each country are counted only once in this total.

^aNot including member territories that are not countries.

^bNot including the British Virgin Islands.

Source: International Monetary Fund, *Direction of Trade Statistics Yearbook*, 1991.

This reasoning is applied in Table 3.⁶ Assume (arbitrarily) that in 1948 the trade among the countries that were later to form the EC was fairly "neutral," or relatively free of government preferences that succeeded in fostering trade among them.⁷ Then the trade share reported as neutral in column 5 is equivalent to the actual share recorded in column 4. For following years, the trade reported as neutral is that which would have taken place among EC members if the share of intra-EC trade in EC total trade had risen (or fallen) by the same percentage as the EC share of world trade. In fact, the share of intra-EC

trade in the EC total has risen faster than the EC share of world trade, with the result that, by 1990, the share of EC trade taking place within the group was 23.5 percentage points greater than if that trade had increased neutrally, or free of any growing bias toward doing business within the group ($59.2 - 35.7 = 23.5$).

⁶ Compare Frankel (1991, pp. 5-9).

⁷ In fact, it is not crucial that trade have been free from such preferences in 1948, for we are interested in examining the *change* after 1948, or whether trade became *less* neutral as time passed.

Table 3

European Community (12) Trade with Selected Areas, 1948 to 1990

Year	In Billions of U.S. Dollars			As Percent of EC Total Trade			EC Total Trade as Percent of World Trade (7)
	Total Trade (1)	Intra-EC Trade (2)	Trade with the United States (3)	Actual Intra-EC Trade (4)	Neutral Intra-EC Trade ^a (5)	Actual Trade with the United States (6)	
1948	35.2	9.6	4.6	27.4	27.4	13.0	31.4
1949	37.5	11.5	4.4	30.6	31.6	11.8	36.3
1950	37.4	13.0	3.6	34.7	31.3	9.7	36.0
1951	50.8	16.4	5.3	32.3	29.8	10.4	34.3
1952	49.8	16.4	4.9	33.0	29.8	9.9	34.3
1953	49.8	17.4	4.2	35.0	30.4	8.5	34.9
1954	53.7	19.3	4.4	35.9	31.5	8.2	36.2
1955	61.2	22.4	5.8	36.7	30.1	9.5	34.6
1956	67.1	24.6	7.0	36.7	29.7	10.5	34.1
1957	73.4	26.6	8.1	36.2	29.9	11.1	34.4
1958	71.5	26.2	7.0	36.6	34.6	9.8	39.8
1959	76.9	29.6	8.0	38.4	34.8	10.4	40.0
1960	89.7	35.7	9.5	39.8	35.4	10.6	40.7
1961	95.8	40.5	9.6	42.3	35.9	10.0	41.2
1962	102.8	45.5	10.4	44.3	36.6	10.1	42.1
1963	113.5	52.1	11.2	45.9	37.0	9.9	42.5
1964	127.0	59.6	12.5	47.0	36.3	9.9	41.7
1965	139.4	66.4	13.9	47.6	36.6	10.0	42.1
1966	152.5	73.3	15.6	48.0	36.3	10.2	41.7
1967	158.7	76.7	16.0	48.3	35.7	10.1	41.1
1968	176.7	87.0	18.8	49.2	35.2	10.6	40.5
1969	208.4	106.8	20.3	51.2	36.5	9.7	41.9
1970	240.4	124.8	23.2	51.9	36.3	9.6	41.7
1971	272.1	144.3	24.7	53.0	37.1	9.1	42.7
1972	326.3	178.3	27.9	54.7	37.5	8.5	43.1
1973	451.8	247.5	36.9	54.8	37.3	8.2	42.9
1974	605.6	311.3	47.1	51.4	34.7	7.8	39.8
1975	634.3	326.1	45.7	51.4	35.0	7.2	40.2
1976	709.0	372.2	50.1	52.5	34.3	7.1	39.5
1977	809.3	424.8	56.2	52.5	34.3	6.9	39.4
1978	964.5	518.7	68.8	53.8	35.4	7.1	40.7
1979	1235.6	675.2	86.7	54.6	35.9	7.0	41.3
1980	1463.7	766.6	104.7	52.4	34.8	7.2	40.0
1981	1323.7	668.8	102.9	50.5	30.0	7.8	34.5
1982	1270.7	660.1	99.2	51.9	30.8	7.8	35.4
1983	1226.9	648.9	97.6	52.9	30.7	8.0	35.2
1984	1249.7	656.3	110.2	52.5	29.5	8.8	33.9
1985	1313.7	702.8	118.3	53.5	30.3	9.0	34.8
1986	1577.9	896.7	131.1	56.8	33.5	8.3	38.5
1987	1914.3	1114.0	149.6	58.2	34.6	7.8	39.8
1988	2147.4	1257.6	164.0	58.6	33.9	7.6	39.0
1989	2305.2	1346.5	176.9	58.4	33.7	7.7	38.7
1990	2784.0	1648.7	201.8	59.2	35.7	7.2	41.1

Note: Trade is defined as the sum of reported exports (fob) and imports (cif). The EC includes Belgium, Denmark, France, Germany (West), Greece, Ireland, Italy, Luxembourg, the Netherlands, Portugal, Spain, and the United Kingdom.

^aEC percent of world trade (col. 7) in current year multiplied by the ratio [(intra-EC trade as a percent of EC total trade (col. 4) in 1948)/(EC percent of world trade (col. 7) in 1948)].

Source: International Monetary Fund, Direction of Trade Statistics through U.S. Department of Commerce, Compro data retrieval system.

What is of interest in such data is not so much the position in any one year but the change that has occurred over the years, because what part of trade is truly neutral in any year is highly debatable. What is beyond debate is that the members of the EC have come to trade much more intensively with one another than would be expected from their trade with the entire world, a fact that is reflected in the widening gap between the shares reported for actual and neutral intra-EC trade. As the customs union was completed and its membership has expanded, the bias toward intra-group trade has grown. Thus, the EC has increasingly assumed at least this characteristic of a trading bloc.

Because the EC and EFTA now comprise a very sizable free trade area for manufactured products, that arrangement also merits special attention. As can be seen in Table 4, EC-EFTA, like the EC, has become much more inward-oriented in its trade. Three of the EFTA countries have applied for full membership in the EC, and their inclusion seems highly likely.

Comparable data for the Canada-United States Free Trade Area in Table 5 lend only weak support to the perception of a developing trade bloc. To be sure, the share of actual trade between the two nations exceeded the neutral share by about 10 percentage points in 1990, but that differential has not risen appreciably in years and, indeed, was notably smaller in 1990 than in the early 1970s. Data that could be obtained on this differential for other preferential trading arrangements may be found in Table 6.

Whether a particular arrangement has tended to become a trading bloc is clearer in some cases than in others. Our own assessment, based on Table 6 and presented in Table 7, is that most of them have shown this tendency in recent years. Those that have account for about two-thirds of the world's trade.

Not only the direction but the composition of trade flows is influenced by preferential trading arrangements. At any time a nation will have developed certain comparative, or relatively competitive, advantages and disadvantages in world markets that will be exhibited in the composition of its exports and imports if markets are relatively free and efficient. The nation's "revealed comparative advantage" may be measured by the ratio of the nation's net exports (exports minus imports) in each commodity category to the sum of the nation's total exports and imports in that category (Balassa and Noland 1989, p. 175). This ratio, or index, can take on any value between -1 and $+1$; the larger the algebraic value for a commodity category relative to the values for other categories,

the greater is the country's revealed comparative advantage in the commodity category concerned. Thus, one can measure how the revealed comparative advantage of a country or of a preferential trading arrangement has changed over the years.

This analytic technique is applied here to the EC and the EFTA. It is well known that the EC has pursued a highly protectionist policy for its agriculture, so it might be expected that the EC's relative competitive position in agricultural commodities would have (artificially) improved over the years, as

The members of the EC have come to trade much more intensively with one another than would be expected from their trade with the entire world, thereby assuming at least this characteristic of a trading bloc.

exports were promoted through subsidies and imports were impeded through the variable tariff levy and other devices. One might also wonder if a preferential trading area such as the EC would provide inducements that boosted net exports of sophisticated manufactures in an effort to stimulate technological advance within the member countries.

These hypotheses receive no support from the statistics in Table 8. The rankings for the commodity groups listed in this table are based on revealed comparative advantage ratios computed for the two-year periods 1962-63 (the earliest for which the desired data were readily available) and 1988-89. To be sure, the EC's greatest revealed comparative advantage in 1988-89 was in beverages and tobacco, and the comparative advantage of that category had improved markedly since 1962-63. But the ranking, or relative competitiveness, of food and live animals deteriorated over the same span of years, and the opposite would have happened if the EC had actually succeeded in strengthening its competitive position in these agricultural items. Similarly, the EC's revealed comparative advantage weakened in machines and transport equipment, the reverse of what would have occurred under policies successfully promoting these sophisticated manufactures in world markets.

Table 4
European Community (12) and European Free Trade Association (EC-EFTA) Trade with Selected Areas, 1948 to 1990

Year	In Billions of U.S. Dollars			As Percent of EC-EFTA Total Trade			Total EC-EFTA Trade as Percent of World Trade (7)
	Total Trade (1)	Intra-EC-EFTA Trade (2)	Trade with the United States (3)	Actual Intra-EC-EFTA Trade (4)	Neutral Intra-EC-EFTA Trade ^a (5)	Actual Trade with the United States (6)	
1948	42.6	16.8	5.7	39.4	39.4	13.4	38.0
1949	44.5	18.7	5.3	42.1	44.6	11.9	43.1
1950	44.2	20.5	4.4	46.4	44.0	9.9	42.6
1951	60.9	27.9	6.3	45.8	42.6	10.4	41.2
1952	59.6	27.7	6.0	46.5	42.4	10.0	41.0
1953	58.9	28.1	5.1	47.7	42.7	8.7	41.3
1954	63.9	31.6	5.3	49.4	44.5	8.2	43.1
1955	72.7	36.6	6.8	50.3	42.4	9.4	41.0
1956	80.1	40.4	8.3	50.4	42.1	10.3	40.7
1957	85.5	42.4	9.3	49.6	41.4	10.9	40.0
1958	82.9	41.6	8.0	50.1	47.7	9.6	46.1
1959	89.1	46.0	9.1	51.7	48.0	10.2	46.4
1960	106.3	57.0	11.0	53.7	49.9	10.3	48.3
1961	113.7	64.2	11.0	56.4	50.6	9.6	49.0
1962	122.0	70.9	11.8	58.2	51.7	9.7	50.0
1963	134.2	79.5	12.8	59.2	52.0	9.5	50.2
1964	150.4	90.5	14.3	60.2	51.1	9.5	49.4
1965	165.0	100.2	15.8	60.7	51.6	9.6	49.9
1966	180.1	109.1	17.8	60.6	50.9	9.9	49.2
1967	187.6	113.7	18.1	60.6	50.2	9.7	48.5
1968	207.8	126.6	21.2	60.9	49.3	10.2	47.6
1969	244.5	153.6	22.9	62.8	50.9	9.4	49.2
1970	283.5	180.5	26.1	63.7	50.9	9.2	49.2
1971	318.9	205.2	27.8	64.3	51.8	8.7	50.1
1972	381.6	251.2	31.5	65.8	52.2	8.2	50.4
1973	527.6	348.0	41.5	66.0	51.8	7.9	50.1
1974	708.1	440.2	53.1	62.2	48.2	7.5	46.6
1975	743.4	460.5	51.9	61.9	48.7	7.0	47.1
1976	829.6	519.8	56.7	62.7	47.8	6.8	46.2
1977	945.8	592.6	63.8	62.7	47.7	6.8	46.1
1978	1122.1	715.8	78.3	63.8	49.0	7.0	47.3
1979	1434.6	926.5	98.3	64.6	49.6	6.9	47.9
1980	1703.2	1071.2	118.3	62.9	48.1	6.9	46.5
1981	1539.7	935.9	116.7	60.8	41.5	7.6	40.1
1982	1476.0	915.2	112.2	62.0	42.6	7.6	41.2
1983	1428.5	903.7	111.6	63.3	42.5	7.8	41.0
1984	1456.5	918.0	125.8	63.0	40.9	8.6	39.5
1985	1532.7	981.7	134.5	64.0	42.0	8.8	40.6
1986	1847.5	1253.1	149.7	67.8	46.6	8.1	45.1
1987	2242.2	1552.0	171.2	69.2	48.2	7.6	46.6
1988	2508.3	1737.5	187.9	69.3	47.1	7.5	45.5
1989	2686.4	1850.4	203.4	68.9	46.7	7.6	45.1
1990	3234.8	2255.0	232.1	69.7	49.4	7.2	47.7

Note: Trade is defined as the sum of reported exports (fob) and imports (cif). EFTA is here defined to include Austria, Finland, Iceland, Norway, Sweden, and Switzerland.

^aEC-EFTA percent of world trade (col. 7) in current year multiplied by the ratio [(intra-group trade as a percent of EC-EFTA total trade (col. 4) in 1948)/(EC-EFTA percent of world trade (col. 7) in 1948)].

Source: International Monetary Fund, Direction of Trade Statistics through U.S. Department of Commerce, Compro data retrieval system.

Table 5
Canada-United States Free Trade Area (Can-US) Trade with Selected Areas, 1948-90

Year	In Billions of U.S. Dollars		As Percent of Can-US Trade with World		Can-US Total Trade as Percent of World Trade (5)
	Total Trade with World (1)	Can-US Trade (2)	Actual Can-US Trade (3)	Neutral Can-US Trade ^a (4)	
1948	26.4	7.3	27.8	27.8	23.6
1949	25.2	7.3	28.9	28.8	24.4
1950	25.7	8.2	31.7	29.2	24.8
1951	36.6	11.7	32.1	29.1	24.7
1952	35.5	11.3	31.9	28.8	24.5
1953	36.3	12.0	33.0	30.0	25.5
1954	34.5	11.3	32.8	27.4	23.2
1955	37.1	12.8	34.5	24.7	20.9
1956	43.9	14.9	34.0	26.3	22.3
1957	46.4	14.9	32.1	25.6	21.7
1958	43.4	13.9	32.0	28.4	24.1
1959	46.4	15.2	32.8	28.5	24.2
1960	48.8	14.5	29.8	26.1	22.2
1961	49.0	14.9	30.3	24.9	21.1
1962	51.8	16.0	30.8	25.0	21.2
1963	55.1	16.6	30.2	24.3	20.6
1964	62.2	18.9	30.5	24.1	20.4
1965	67.6	21.7	32.2	24.1	20.4
1966	78.5	26.4	33.6	25.3	21.5
1967	82.3	29.8	36.2	25.1	21.3
1968	95.4	35.4	37.1	25.8	21.9
1969	104.7	40.5	38.7	24.8	21.1
1970	116.7	41.7	35.7	23.9	20.2
1971	127.4	47.8	37.5	23.6	20.0
1972	149.5	56.6	37.9	23.3	19.8
1973	195.3	69.2	35.4	21.9	18.5
1974	276.0	89.2	32.3	21.4	18.2
1975	281.8	91.0	32.3	21.0	17.9
1976	325.7	106.7	32.8	21.4	18.1
1977	364.2	116.7	32.1	20.9	17.7
1978	419.3	129.6	30.9	20.8	17.7
1979	513.2	152.2	29.7	20.2	17.1
1980	612.6	163.8	26.7	19.7	16.7
1981	654.6	182.6	27.9	20.1	17.1
1982	600.8	169.4	28.2	19.7	16.8
1983	616.6	193.3	31.3	20.9	17.7
1984	732.9	237.6	32.4	23.4	19.9
1985	752.1	244.4	32.5	23.5	19.9
1986	785.7	242.4	30.9	22.6	19.2
1987	874.5	268.4	30.7	21.4	18.2
1988	1016.7	310.6	30.5	21.8	18.5
1989	1106.9	335.1	30.3	21.9	18.6
1990	1168.2	354.9	30.4	20.3	17.2

Note: Trade is defined as the sum of reported exports (fob) and imports (cif).

^aCan-US percent of world trade (col. 5) in current year multiplied by the ratio [(Can-US trade as a percent of Can-US total trade with world (col. 3) in 1948)/(Can-US percent of world trade (col. 5) in 1948)].

Source: International Monetary Fund, Direction of Trade Statistics through U.S. Department of Commerce, Compro data retrieval system.

Table 6
Excess of Actual over Neutral Intragroup Trade, as Percentage of Group Total Trade, for Various Trading Groups

Year	Trading Arrangement and Percentage Excess ^a										
	ANZCERT	CACM	Can.-U.S. FTA	EC (12)	EC (12)- EFTA	EFTA	MERCOSUR	ANCOM	ASEAN	LAIA	GCC
1948	0	0	0	0	0	0					
1949	0	-.3	.1	-1.0	-2.5	-.6					
1950	-.2	-1.7	2.6	3.4	2.3	-.9					
1951	-.5	-1.0	3.0	2.4	3.2	-1.1					
1952	.3	-2.1	3.1	3.2	4.1	.1					
1953	1.0	-2.3	3.0	4.6	5.0	-.2					
1954	1.4	-2.2	5.5	4.4	4.9	-.2					
1955	1.6	-2.0	9.8	6.6	7.9	.2					
1956	2.5	-1.8	7.7	7.0	8.3	-.1					
1957	3.6	-1.6	6.5	6.3	8.2	1.9	0				
1958	3.9	-1.4	3.5	2.0	2.4	1.4	.5				
1959	2.7	1.2	4.4	3.6	3.7	1.5	-1.5				
1960	2.6	1.9	3.7	4.4	3.7	.2	-1.7				
1961	3.2	3.4	5.4	6.4	5.8	.7	-3.8				
1962	3.3	3.6	5.8	7.7	6.5	.9	-2.1				
1963	4.1	6.0	5.9	8.9	7.3	1.4	-.8				
1964	3.6	9.4	6.4	10.7	9.1	2.1	3.2				
1965	3.4	10.8	8.1	11.0	9.1	2.7	5.2				
1966	3.7	14.2	8.4	11.8	9.7	3.4	3.5				
1967	3.4	16.8	11.1	12.6	10.4	4.5	3.4				
1968	3.5	20.2	11.3	14.0	11.7	4.7	4.2				
1969	3.1	20.1	13.8	14.8	11.9	5.5	4.0				
1970	4.1	20.7	11.9	15.6	12.8	6.3	3.6				
1971	4.4	18.5	14.0	15.9	12.5	6.7	2.5				
1972	4.9	18.4	14.6	17.2	13.7	6.9	1.4	0	0	0	
1973	5.4	18.4	13.6	17.5	14.1	6.7	1.1	.7	-3.2	.8	
1974	5.3	17.5	10.9	16.7	14.0	7.4	-.2	.3	-7.1	-.6	
1975	4.7	16.6	11.3	16.5	13.2	7.4	-.4	1.2	-5.6	.4	
1976	4.9	16.2	11.4	18.1	14.9	6.7	1.4	1.9	-5.5	2.1	
1977	5.2	13.3	11.1	18.2	15.0	6.2	1.6	2.4	-6.2	3.2	
1978	5.2	16.4	10.1	18.4	14.8	4.9	2.1	1.1	-6.5	1.9	
1979	5.4	15.4	9.5	18.7	15.0	5.1	4.4	1.6	-6.3	2.5	0
1980	5.1	18.5	7.0	17.6	14.8	4.6	3.3	1.2	-9.0	.7	-1.1
1981	5.2	16.8	7.8	20.5	19.3	5.8	1.9	1.6	-9.4	.6	-.7
1982	4.9	16.4	8.4	21.1	19.4	5.7	2.3	2.3	-9.1	2.2	.1
1983	5.5	16.4	10.5	22.2	20.8	5.3	1.1	2.3	-9.3	1.2	.8
1984	5.9	14.0	9.0	23.0	22.1	5.5	2.1	1.8	-9.9	1.3	2.0
1985	5.4	11.0	9.0	23.2	22.0	5.3	1.8	1.9	-6.2	.7	3.4
1986	5.2	8.0	8.3	23.4	21.2	4.9	5.4	1.9	-4.6	4.3	4.6
1987	6.3	10.1	9.3	23.6	21.0	4.7	4.6	2.7	-5.4	4.1	4.5
1988	6.1	9.8	8.8	24.6	22.2	4.5	4.5	2.9	-8.7	3.8	4.8
1989	5.8	10.0	8.4	24.7	22.2	4.5	6.4	3.1	-11.6	4.7	5.3
1990	6.1	10.0	10.1	23.5	20.4	4.1	6.8	2.5	-13.2	3.6	4.9

^aMinus sign indicates excess of neutral over actual intragroup trade.

Note: As available, data are shown beginning with 1948. Lesotho and Swaziland are not included in PTA, nor is SACU included in this table, because of lack of data.

Source: International Monetary Fund, Direction of Trade Statistics through U.S. Department of Commerce, Compro data retrieval system.

Table 6 *continued*

Trading Arrangement and Percentage Excess ^a											
Year	ACM	AMU	Bangkok	CARICOM	CEAO	CEEAC	ECOWAS	MRU	OECS	PTA	UDEAC
1981	0	0	0	0	0	0	0	0	0	0	0
1982	-1.9	-.1	.1	1.0	1.3	-.6	1.7	.2	4.2	-.1	-1.1
1983	-2.3	-.1	.3	1.1	.4	-.9	2.3	.6	4.7	.3	-1.0
1984	-1.3	.4	.7	.5	-.6	-1.1	4.2	.2	4.1	.6	-1.5
1985	-.1	.6	-.1	1.8	0	-.7	3.3	.2	2.6	-.1	-.8
1986	.9	1.2	-.1	1.1	-.7	-.3	5.8	.2	1.3	.2	.1
1987	2.3	1.4	-.6	1.6	1.2	.7	5.3	.1	-.1	.7	1.2
1988	1.1	1.5	-.9	2.9	2.0	1.1	4.5	-.4	-.8	1.1	1.6
1989	1.8	2.0	-.8	4.4	3.3	.9	4.1	-.6	-1.4	1.1	1.7
1990	1.4	1.8	-.7	4.3	3.8	1.2	4.2	-.6	-.7	1.3	1.7

Table 7
Trends toward or away from Bloc Formation in International Trade by Preferential Trading Arrangements

Trade Becoming More Bloc-like	Trade Becoming Less Bloc-like	No Persistent Strong Tendency
AMU	ASEAN	ACM
ANCOM	CACM (since 1970)	Bangkok
ANZCERT	Can-US FTA (1972-80)	MRU
CACM (through 1970)	ECOWAS (since 1986)	
Can-US FTA (1948-72 and 1980-90)	EFTA (since 1975)	
CARICOM	OECS (since 1983)	
CEAO (since 1986)		
CEEAC (since 1984)		
EC		
ECOWAS (through 1986)		
EC-EFTA		
EFTA (through 1975)		
GCC		
LAIA (since 1985)		
MERCOSUR (since 1983)		
PTA (since 1985)		
UDEAC (since 1986)		

Source: Table 6.

Table 8
Ranking by Commodity Group of Revealed Comparative Advantage for EC (12) and EFTA Countries' Trade with the Rest of the World, 1962-63 and 1988-89

Major SITC Commodity Group	EC (12)		EFTA	
	1962-1963	1988-1989	1962-1963	1988-1989
Food and live animals	6	7	7	7
Beverages and tobacco	5	1	8	9
Crude materials excluding fuels	9	8	1	3
Mineral fuels, etc.	8	9	9	2
Animal, vegetable oil, fat	7	6	5	6
Chemicals	3	2	4	4
Basic manufactures	4	4	2	1
Machines, transport equipment	1	3	6	5
Misc. manufactured goods	2	5	3	8

Note: In the rankings 1 is most positive, 9 most negative. EFTA is here defined to include Austria, Finland, Iceland, Norway, Sweden, and Switzerland.

Source: U.S. Department of Commerce, Compro system, UN data base.

This is not to say that EC policies to protect agriculture or to promote sophisticated manufactures were without effect. Absent any policies of this nature, the deteriorations in rankings of food and live animals and of machines and transport equipment could have been even greater. What can be said here is that deteriorations did take place in spite of any such policies.

EFTA is much less tightly organized than the EC, especially with respect to agriculture. While the statistics in Table 8 would tend to support a hypothesis that EFTA has promoted certain categories of manufactures, further research beyond the scope of this article would be required to confirm that hypothesis.

Where Do We Go from Here?

In spite of Viner's classic critique, it would be hard to demonstrate that the preferential trading arrangements now in operation have had a significantly deleterious impact on the world economy. To demonstrate the opposite would be almost as difficult, however (because of the complexity of the question and the data required), and much concern has been voiced that the international economy is tending to fracture into estranged, if not hostile, discriminatory trading blocs. The EC constitutes one bloc, and is expanding, perhaps eventually to encompass virtually all of Europe. Another bloc may be forming around the free trade area formed by Canada and the United States, which are negotiating with Mexico to establish a North American Free Trade Area—which, in turn, could be extended to all of Latin America. Finally, some foresee the development of a third major bloc in East Asia centered about Japan, although countries in the region have shown no inclination for such an arrangement.

In any event, preferential liberalization of trade is clearly inferior to global liberalization. Although the global approach pursued under GATT may be complex and slow, any nation that undertakes to negotiate a series of preferential trade agreements—including a series of bilateral agreements eventually forming a free trade area—will find that this alternate approach is far from swift and simple. Under the preferential approach, each new agreement must take into account bargains struck in prior agreements, and prior agreements may well have to be renegotiated to accommodate the interests of all parties. And paradoxically, to negotiate liberalization across a broad range of trade may be more difficult for just a

few countries than for many, because when many offer sweeping reductions in barriers the odds may be greater that each party will perceive some considerable gain. Partly for this reason, GATT negotiations are much more likely than bilateral negotiations to yield liberalization of the most pernicious and intractable nontariff barriers, especially within highly sensitive areas such as agriculture and textiles and apparel.

But if global negotiations fail, blocs that genuinely liberalized trade among themselves could improve the general welfare and set a good example. In particular, they should welcome new members, for the best free trade area is worldwide in scope. In this connection, some encouragement should be taken from the recent collapse of the most discriminatory trading bloc of them all—COMECON, the Council for Mutual Economic Assistance, whose membership was drawn from the former Soviet communist bloc.

Conclusion

A multiplicity of preferential arrangements has permeated the trading world. In recent years most of these arrangements, accounting for about two-thirds of world trade, have increasingly resembled "trading blocs," in that their trade has become oriented more inward, among bloc members, and less outward, with the rest of the world. This outcome is hardly surprising, since the essence of a preferential trading arrangement is to discriminate in favor of the trade of participants over that of nonparticipating countries.

Certain types of preferential arrangements are sanctioned by the international codes to which most countries subscribe. This law is rather vague and has been loosely interpreted or applied, so that governments have felt relatively free to discriminate in international trade without much risk of retaliation from the countries that are disadvantaged.

Like the law, the standard theory of international trade is somewhat ambiguous regarding preferential arrangements. To be sure, the theory asserts that free trade is more efficient than discriminatory trade. But theory also acknowledges that in a world with less than free trade a discriminatory (or preferential) reduction of trade barriers can enhance efficiency in certain circumstances. Unfortunately, because of the complexity of the issue, very little has been learned about the actual impact of preferential trading on world efficiency and welfare, although a number of empirical studies have been undertaken.

Not all preferential trading arrangements have endured. They seem more likely to flourish if the members are fairly similar in their economic structures, per capita incomes, and policies toward international trade, if they are fairly close geographically, and if they adopt an across-the-board rather than product-by-product approach toward liberalizing trade among themselves.

That so many preferential arrangements have been launched may seem puzzling, in view of their frequent failure to attain their goals and their dubious impact on world efficiency and welfare. More than one motivation has been at work. Producers who expect to gain greater protection may lobby for preferential arrangements, while the injured consumers may offer little resistance. Also, frustration with the slow progress of global trade liberalization and with

the "free ride" taken by some countries that benefit but contribute little has inspired growing interest in preferential arrangements limited to "kindred souls," all of whom contribute. Or nations may join such arrangements to enhance their bargaining power vis-à-vis the rest of the world or to avoid being "left out" and becoming victims of increased discrimination. Sometimes the motivation is partly to form a more stable political area, as was the case for the E.C.

Over the long run, nondiscriminatory reductions in trade barriers are clearly preferable to discriminatory reductions. But should global negotiations fail, blocs that truly liberalized trade among themselves could improve the general welfare. To set the best example for the rest of the trading world, they should be receptive to new members, for the ideal free trade area is worldwide in scope.

References

- Agyemang, Augustus Asante. 1990. "Trade Liberalization Under the Treaty of the Economic Community of West African States (ECOWAS): Some Preliminary Highlights of Legal and Economic Problems." *Journal of World Trade*, vol. 24, no. 6 (December), pp. 57-79.
- Balassa, Bela, and Marcus Noland. 1989. "The Changing Comparative Advantage of Japan and the United States." *Journal of the Japanese and International Economies*, vol. 3 (June), pp. 174-88.
- Balasubramanyam, V.N. 1989. "ASEAN and Regional Trade Cooperation in Southeast Asia." In Greenaway, Hyclak, and Thornton, eds., *Economic Aspects of Regional Trading Arrangements*, pp. 167-87. New York: New York University Press.
- Belous, Richard S. and Rebecca S. Hartley. 1990. "Regional Trading Blocs: An Introduction." In Belous and Hartley, eds., *The Growth of Regional Trading Blocs in the Global Economy*, pp. 1-21. Washington, D.C.: National Planning Association.
- Bruce, James. 1991. "Latin Nations in Mercosul Bloc to Begin Duty Cuts Nov. 29." *The Journal of Commerce*, November 4, p. 5A.
- Carl, Beverly May. 1986. *Economic Integration Among Developing Nations: Law and Policy*. Praeger Special Studies, Praeger Scientific. New York: Praeger Publishers.
- El-Kuwaiz, Abdullah Ibrahim. 1987. "Economic Integration of the Cooperation Council of the Arab States of the Gulf: Challenges, Achievements and Future Outlook." In J. Sandwizk, ed. *The Gulf Cooperation Council*, pp. 71-83. London: Westview Press for the American-Arab Affairs Council.
- Frankel, Jeffrey A. 1991. "Is a Yen Bloc Forming in Pacific Asia?" In Richard O'Brien, ed., *Finance and the International Economy: 5, The AMEX Bank Review Prize Essays in Memory of Robert Marjolin*, pp. 5-20. New York: Oxford University Press.
- General Agreement on Tariffs and Trade. 1969. *Basic Instruments and Selected Documents*, vol. IV, March. Geneva: The Contracting Parties to the General Agreement on Tariffs and Trade.
- . "GATT: What It Is, What It Does." 1980. GATT Information Service.
- . 1992. *International Trade 1990-91*. Geneva: GATT.
- Guma, X.P. 1990. "The Revised Southern African Customs Union Agreement: An Appraisal." *The South African Journal of Economics*, vol. 58, no. 1, pp. 63-77.
- Gunter, Frank R. 1989. "Customs Union Theory: Retrospect and Prospect." In Greenaway, Hyclak, and Thornton, eds., *Economic Aspects of Regional Trading Arrangements*, pp. 1-30. New York: New York University Press.
- International Monetary Fund, General Economy Division of the Bureau of Statistics. 1991. *Direction of Trade Statistics Yearbook*, 1991. Washington, D.C.: International Monetary Fund.
- Jackson, John H. 1989. *The World Trading System: Law and Policy of International Economic Relations*. Cambridge, Massachusetts: The MIT Press.
- Jennings, Thomas. 1991. "United States-Canada Free-Trade Agreement: After 2 Years, How Does It Shape Up?" *International Economic Review*, United States International Trade Commission, April, pp. 7-9.
- Kostecke, M.M. 1987. "Export-restraint Arrangements and Trade Liberalization." *World Economy*, vol. 10, no. 4, pp. 425-53.
- Kowalczyk, Corsten. 1990. "Welfare and Customs Unions." National Bureau of Economic Research. Working Paper no. 3476, October. Cambridge, MA: NBER.
- Krugman, Paul. 1991. "The Move Toward Free Trade Zones." *Economic Review*, Federal Reserve Bank of Kansas City, November/December, pp. 5-25.
- Lewis, Paul. 1984. "Western Europe Ends Industrial Trade Curbs." *The New York Times*, January 2, p. 39.
- MacBean, A.I. and P.N. Snowden. 1981. *International Institutions in Trade and Finance*. London: George Allen & Unwin Ltd.
- Manifold, Diane. 1992. "Regional Organizations in Southeast Asia." *International Economic Review*, United States International Trade Commission, April, pp. 12-17.
- Moore, Paul. 1992. "Approval of Andean Trade Preference Act Will Create Opportunities for U.S. Business." *Business America*, January 13, p. 15.
- Organisation for Economic Cooperation and Development, Trade Committee. 1990. "Recent Developments in Regional Trading Arrangements among OECD Countries: Main Implications for Third Countries and for the Multilateral Trading System." Note by the Secretariat, December 21, 1990. Paris: OECD.

- Pomfret, Richard. 1988. *Unequal Trade: The Economics of Discriminatory International Trade Policies*. New York: Basil Blackwell, Inc.
- Pura, Raphael. 1992. "ASEAN Leaders See Need to Create Unified Market to Avoid Loss of Trade." *The Wall Street Journal*, January 28, p. A10.
- Schott, Jeffrey J. 1991. "Trading Blocs and the World Trading System." *The World Economy*, vol. 14, no. 1 (March), pp. 1-17.
- Stamps, James. 1991a. "Laying the Groundwork for Free Trade in the Western Hemisphere: An Update." *International Economic Review*, United States International Trade Commission, October, pp. 9-11.
- . 1991b. "Liberalizing Trade in Latin America and the Caribbean: An Update." *International Economic Review*, United States International Trade Commission, April, pp. 6-7.
- Thoumi, Francisco E. 1989. "Bilateral Trade Flows and Economic Integration in Latin America and the Caribbean." *World Development*, vol. 17, no. 3, pp. 421-29.
- Thurston, Charles W. 1991. "Another Trade Zone Formed as Chile, Mexico Implement Pact." *The Journal of Commerce*, December 31, pp. 1A-2A.
- Union of International Associations. 1991. *Yearbook of International Organizations 1991-1992*. 28th ed., vols. 1 & 3. Brussels: Union of International Associations.
- United Nations, Department of International Economic and Social Affairs. 1990. *World Economic Survey 1990: Current Trends and Policies in the World Economy*. New York: United Nations.
- . 1991. *World Economic Survey 1991: Current Trends and Policies in the World Economy*. New York: United Nations.
- U.S. Department of State, Bureau of Public Affairs, Office of the Historian. 1990. "European Community Integration, 1951-1990." April.
- U.S. Department of State. 1991. "Fact Sheet: Association of South East Asian Nations (ASEAN)." Dispatch, August 5, p. 564.
- Viner, Jacob. 1950. *The Customs Union Issue*. New York: Carnegie Endowment for International Peace. Reprint, Washington, D.C.: Anderson Kramer Associates, 1961.
- Wonnacott, Paul, and Lutz, Mark. 1989. "Is There a Case for Free Trade Areas?" In Jeffrey J. Schott, ed., *Free Trade Areas and U.S. Trade Policy*, pp. 59-95. Washington, D.C.: Institute for International Economics.

Appendix: Sources for Table 1

- Agyemang, Augustus Asante. 1990. "Trade Liberalization Under the Treaty of the Economic Community of West African States (ECOWAS): Some Preliminary Highlights of Legal and Economic Problems." *Journal of World Trade*, vol. 24, no. 6 (December), pp. 57-79.
- Balasubramanyam, V.N. 1989. "ASEAN" and Regional Trade Cooperation in Southeast Asia." In Greenaway, Hyclak, and Thornton, eds., *Economic Aspects of Regional Trading Arrangements*, pp. 169-70. New York: New York University Press.
- Bruce, James. 1991. "Latin Nations in Mercosul Bloc to Begin Duty Cuts Nov. 29." *The Journal of Commerce*, November 4, p. 5A.
- El-Kuwaiz, Abdullah Ibrahim. 1987. "Economic Integration of the Cooperation Council of the Arab States of the Gulf: Challenges, Achievements and Future Outlook." In J. Sandwizk, ed., *The Gulf Cooperation Council*, pp. 71-83. London: Westview Press for the American-Arab Affairs Council.
- Guma, X.P. 1990. "The Revised Southern African Customs Union Agreement: An Appraisal." *The South African Journal of Economics*, vol. 58, no. 1, pp. 63-77.
- Jennings, Thomas. 1991. "United States-Canada Free-Trade Agreement: After 2 Years, How Does It Shape Up?" *International Economic Review*. United States International Trade Commission, April pp. 7-9.
- Lewis, Paul. 1984. "Western Europe Ends Industrial Trade Curbs." *The New York Times*, January 2, p. 39.
- MacBean, A. I. and P. N. Snowden. 1981. *International Institutions in Trade and Finance*. London: George Allen and Unwin Ltd, pp. 167-68.
- Manifold, Diane. 1992. "Regional Organizations in Southeast Asia." *International Economic Review*, United States International Trade Commission, April, p. 13.
- Moore, Paul. 1992. "Approval of Andean Trade Preference Act Will Create Opportunities for U.S. Business." *Business America*, January 13, p. 15.
- Organisation for Economic Cooperation and Development, Trade Committee. 1990. "Recent Developments in Regional Trading Arrangements among OECD Countries: Main Implications for Third Countries and for the Multilateral Trading System." Note by the Secretariat December 21, 1990, p. 19.
- Pomfret, Richard. 1988. *Unequal Trade: The Economics of Discriminatory International Trade Policies*. New York: Basil Blackwell Inc., pp. 84-89.
- Pura, Raphael. 1992. "ASEAN Leaders See Need to Create Unified Market to Avoid Loss of Trade." *The Wall Street Journal*, January 28, p. A10.
- Stamps, James. 1991. "Laying the Groundwork for Free Trade in the Western Hemisphere: An Update." *International Economic Review*, United States International Trade Commission. October, pp. 9-11.
- . 1991. "Liberalizing Trade in Latin America and the Caribbean: An Update." *International Economic Review*, United States International Trade Commission. April, pp. 6-7.
- Thurston, Charles W. 1991. "Another Trade Zone Formed as Chile, Mexico Implement Pact." *The Journal of Commerce*, December 31, p. 1A-2A.
- Union of International Associations. 1991. *Yearbook of International Organizations 1991-1992*. 28th ed., vols. 1 & 3. Brussels: Union of International Associations.
- United Nations, Department of International Economic and Social Affairs. 1990. *World Economic Survey 1990. Current Trends and Policies in the World Economy*. New York: United Nations, pp. 67-68, 70.
- U.S. Department of State, Bureau of Public Affairs, Office of the Historian. 1990. "European Community Integration, 1951-1990." April.
- U.S. Department of State. 1991. "Fact Sheet: Association of South East Asian Nations (ASEAN)." Dispatch, August 5, 1991, p. 564.

The Capital Crunch in New England

New England banks, among the healthiest in the nation during the 1980s, have recently experienced failures at a rate unprecedented in the postwar period. An average of less than one New England bank insured by the Federal Deposit Insurance Corporation (FDIC) closed each year between 1979 and 1989. Forty-six FDIC-insured banks failed here in 1991, and of those still operating, many are struggling to meet capital requirements. The primary cause of this collapse was the extensive bank exposure to real estate loans.

Real estate lending had been a principal reason for the rapid expansion of New England banks. While banks nationwide substantially increased their exposure to real estate during the 1980s, real estate portfolios at New England banks grew at twice the national rate. When nominal real estate prices began to decline in New England, collateral became impaired and many loans stopped performing. The consequent increased provision for expected loan losses (loan loss reserves) caused a rapid deterioration in bank capital throughout the region.

The timing of this decline in bank capital was most inopportune, occurring just as regulators, in response to new legislation and international agreements, increasingly emphasized the importance of adequate bank capital/asset ratios. Having just lost a significant proportion of their capital, many banks tried to satisfy their capital/asset ratio requirements by shrinking their institutions.

Banks' attempts to shrink can have serious ramifications for the rest of the regional economy. If banks tighten credit conditions, call loans, and discourage new business, bank-dependent borrowers will face serious constraints as their access to external funds is restricted. These bank-dependent borrowers are most likely to be small and medium-sized firms that do not have access to national credit markets or even to banks outside the region.

The first section of this article documents the critical role played by real estate in the loss of bank capital and the way bank capital regulation

*Joe Peek and
Eric S. Rosengren*

Professor of Economics, Boston College and Visiting Economist, Federal Reserve Bank of Boston, and Vice President and Economist at the Boston Fed, respectively. Robert Chicoski provided valuable research assistance.

has exacerbated the problem for New England banks. The second section discusses why banks facing binding capital constraints will shrink more than unconstrained banks when an adverse capital shock occurs. The third section shows that New England banks with low capital/asset ratios are in fact shrinking their institutions faster than better-capitalized institutions, and that this behavior has been particularly apparent in those liability categories that are the marginal sources of funds for most banks. The final section summarizes the findings.

I. The Role of Real Estate in the New England Capital Crunch

During the 1980s, New England experienced a real estate boom. With the New England unemployment rate below 3.5 percent and personal income more than 20 percent above the national average, housing prices and construction activity grew at an unprecedented rate. Between 1980 and 1988, employment in the construction sector grew by 84 percent, while the population of New England grew by only 5 percent. Such a rapid expansion in construction could be sustained only with substantial growth in real estate lending, which the banking sector provided in large part.

Table 1 shows the rapid increase in bank assets at FDIC-insured commercial and savings banks from 1984 to 1989. While total assets in FDIC-insured institutions in the nation grew by a little over one-

Much of the growth in New England banking assets in the 1980s was the result of the rapid expansion of real estate lending.

third from 1984 to 1989, assets in New England institutions more than doubled.¹

Much of the growth in New England banking assets was the result of the rapid expansion of real estate lending. In 1984, New England commercial banks had 16.6 percent of their assets in real estate loans, virtually the same percentage as their FDIC-insured counterparts elsewhere in the country. They had slightly higher concentrations in business lend-

ing compared to FDIC-insured institutions nationally, with larger portfolio shares held in commercial and industrial loans, commercial real estate, and leases.

In contrast, New England savings banks had 53.8 percent of their assets in real estate loans in 1984, primarily in mortgages on one- to four-family homes. This difference reflects the historical role of savings banks, which traditionally had specialized in real estate loans and particularly one- to four-family mortgages, much like savings and loans elsewhere in the country.² Savings banks also held a larger proportion of multifamily and commercial real estate loans than did commercial banks in New England, but they had many fewer construction loans, consumer loans, and commercial and industrial loans. Savings banks also were much better capitalized than commercial banks in New England or FDIC-insured institutions nationwide.

With the boom in New England real estate, both commercial and savings banks increased their exposure to real estate lending. The largest increase for both types of institutions was in the construction loan category, which grew 332.1 percent for commercial banks and an astounding 921.5 percent for savings banks (though some of the increase in savings bank assets reflects increases in the number of FDIC-insured banks). By 1989, savings banks had a greater share of their assets in construction loans than either commercial banks in New England or all FDIC-insured institutions nationwide. Furthermore, both commercial banks and savings banks in New England increased their commercial real estate and multifamily mortgage lending by over 250 percent. Although during this period much of the increase in assets occurred in the more risky categories, savings banks were profitable enough to raise their capital ratios by more than did commercial banks.

Other lending categories were growing rapidly as well. At New England's commercial banks, commercial and industrial loans grew by 95.2 percent and consumer loans by 62 percent, both above the growth rate for FDIC-insured institutions nationwide. Savings banks exhibited even more rapid growth, increasing their lending in areas they had traditionally

¹ From 1984 to 1987 the sharp increase is due in part to the greater increase in the number of FDIC-insured institutions in New England than in the rest of the nation.

² In 1989, only 6.6 percent of the total assets of New England depository institutions were held by S & L institutions supervised by the Office of Thrift Supervision.

Table 1
*Percentage Change and Allocation of Assets and Liabilities of FDIC-Insured Banks,
 New England and the United States, 1984 to 1989*

	New England						United States ^a		
	Commercial Banks			Savings Banks			Commercial and Savings Banks		
	Percent Change	Percent of Total Assets		Percent Change	Percent of Total Assets		Percent Change	Percent of Total Assets	
	1989	1984	1989	1989	1984	1989	1989	1984	1989
Assets	94.6						34.1		
C&I	95.2	19.8	19.8	263.7	4.2	5.8	25.4	16.3	15.2
Consumer	62.0	10.4	8.7	97.0	6.6	4.9	55.6	9.6	11.1
Real Estate	269.9	16.6	31.4	223.2	53.8	65.8	99.4	16.5	24.6
Construction	332.1	2.8	6.2	921.5	1.6	6.3	99.5	2.8	4.2
1-4 Family	257.4	8.2	15.0	181.3	41.3	44.0	90.2	8.7	12.3
Multifamily	269.3	.4	.8	278.8	2.8	4.0	81.0	.7	1.0
Commercial	257.4	5.1	9.4	277.5	8.1	11.5	127.3	3.9	6.7
Leases	301.1	1.3	2.6	n.a.	n.a.	.2	106.9	.6	.9
Securities	87.9	14.0	13.5	n.a.	n.a.	14.8	37.0	16.1	16.5
Liabilities									
Total Deposits	87.9	77.3	74.4	140.4	88.3	80.4	32.5	78.1	77.1
NOWs	n.a.	n.a.	4.8	n.a.	n.a.	4.3	n.a.	n.a.	5.7
MMDAs	72.9	15.7	13.9	n.a.	n.a.	10.6	41.8	9.9	10.5
CDs	167.7	8.9	12.2	n.a.	n.a.	7.9	45.6	10.1	11.0
Capital	114.4	5.5	6.0	218.0	7.3	8.8	42.7	6.1	6.4
Memo: Total Assets (\$ billions)		\$93	\$181		\$38	\$100		\$2577	\$3457

^aSavings banks are not broken out nationally because they represented only 6.9% of assets for all FDIC-insured institutions in the United States in 1989.

n.a. = not available.

Source: Call Report data for FDIC-insured institutions.

left to commercial banks. Commercial and industrial loans held by New England savings banks grew by 263.7 percent, dwarfing the 25.4 percent growth rate for all FDIC-insured commercial and savings banks nationwide.

By 1989, New England commercial banks also had much higher concentrations of commercial and industrial loans and real estate loans than banks in the rest of the nation. These concentrations were particularly large in areas viewed as risky, such as construction lending and commercial real estate loans. Savings banks were even more aggressive, with both construction and commercial real estate loans representing a higher percentage of their assets than in New England commercial banks. Initially quite profitable, these loan concentrations were to pose a serious problem when real estate prices began to fall.

Some of the increased lending to real estate was at the expense of investments in other lending categories. The share of assets held in securities by New England banks decreased, although it increased nationally. Similarly, the share of assets in consumer loans decreased in New England, but increased nationally. This difference could also have been related to the rapid rise in New England house prices, which created equity that could be used to substitute home equity loans for consumer loans. Home equity loans were typically a less expensive alternative, especially since the tax-deductibility of interest on consumer loans was phased out. Thus, as a result of the real estate boom, banks in New England moved far more aggressively into real estate lending than banks in the rest of the nation.

On the other side of the balance sheet, deposits were growing almost as fast as assets, and bank

capital grew even faster than assets. The aggressiveness of New England banks can be seen in the pattern of deposit growth rates. For commercial banks, growth in certificates of deposit (CDs) was nearly double the growth rate of total deposits, raising the CD share of liabilities from 8.9 to 12.2 percent, in part by attracting funds from outside the region.³ At the same time, the share of money market deposit accounts (MMDAs) was declining. At the national level, the increase in the CD share was much smaller, and the MMDA share rose rather than fell.

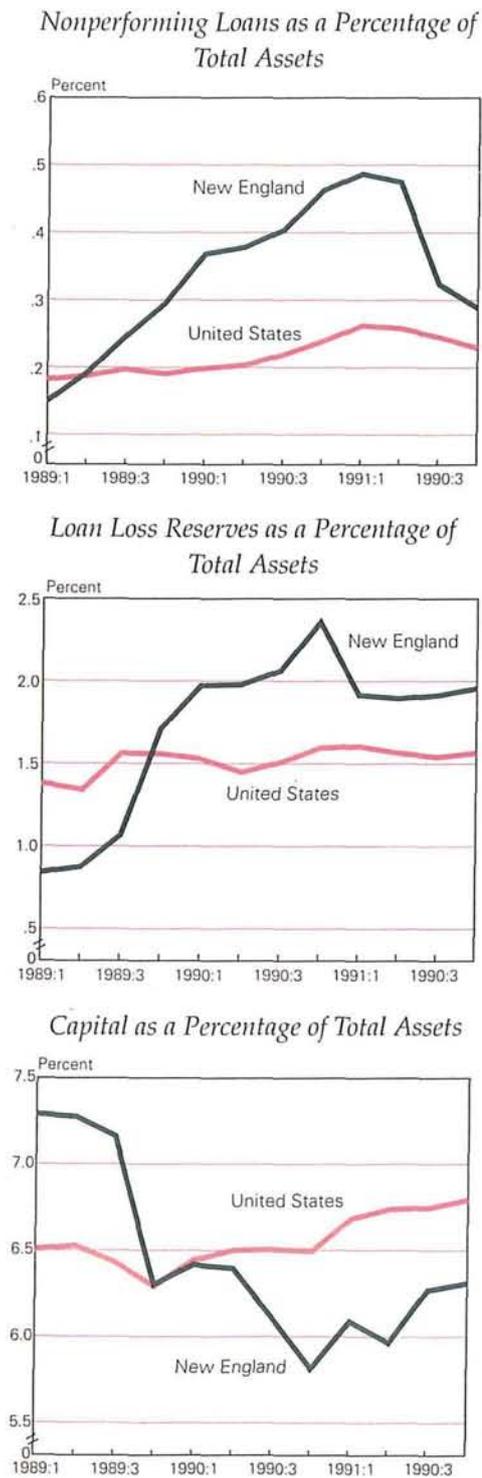
The Impact of Falling Real Estate Prices on Bank Capital

In 1989, it became clear that the real estate boom had ended. The median sales price of a single-family home in Boston, \$96,000 in the second quarter of 1984, peaked at \$186,000 in the second quarter of 1989 before beginning its fall to \$171,000 by the second quarter of 1991. As prices for homes, commercial property, and land began to fall, and as the economy deteriorated and the unemployment rate rose, the risk in most banks' loan portfolios increased significantly. Cash-flow problems and diminished collateral values caused many borrowers to stop making their loan payments.

The top panel of Figure 1 shows that in 1989, New England banks suffered a rapid increase in nonperforming loans, defined here as the sum of loans 90 days past due and nonaccruing loans. (See the box for the accounting treatment of loan losses for banks.) As banks realized that loan losses would be substantially greater than anticipated, they increased their loan loss reserves rather dramatically (middle panel of Figure 1); this, in turn, seriously depleted their capital (bottom panel). The pattern was very different for all U.S. FDIC-insured banks. They experienced substantially slower increases in nonperforming loans and have not increased loan loss reserves as much as New England institutions, and they have actually experienced increases in capital.

Table 2 shows the growth rates of assets, liabilities, and capital for New England commercial and savings banks and for all U.S. FDIC-insured institutions during a more recent period, the two years ending in the third quarter of 1991 (the most recent

Figure 1



Source: Call report data for FDIC-insured institutions.

³ Some historical data are not available from the call report data, either because a subset of institutions did not report that item, or because the definition of that item was significantly different in previous call reports.

Troubled Loans

A loan more than 30 days behind in payments of principal and interest is considered past due. If the loan remains past due, the lending bank will anticipate that some or all of the loan will not be repaid, and it will add to its loan loss reserves. Loan loss reserves are a contra balance sheet account; total loans outstanding minus the loan loss reserve equals the total of loans expected to repay interest and principal. Once a loan is no longer expected to be fully repaid, the expected

loss is charged off, that is, the amount of the expected loss is removed from the loan loss reserve and the loan no longer accrues interest. For these nonaccruing loans, any additional payments by the borrower are subtracted from the principal rather than credited to earnings. Increases in loan loss reserves cause a decrease in earnings. Loan charge-offs have no direct effect on earnings unless the loan loss reserve is replenished. For details on the accounting of problem loans, see Walter (1991).

Table 2

Percentage Change in Assets and Liabilities of FDIC-Insured Banks, New England and the United States, 1989:III to 1991:III

	New England						United States ^a		
	Commercial Banks			Savings Banks			Commercial and Savings Banks		
	Percent Change	Percent of Total Assets		Percent Change	Percent of Total Assets		Percent Change	Percent of Total Assets	
	91:III 89:III	89:III	91:III	91:III 89:III	89:III	91:III	91:III 89:III	89:III	91:III
Assets	-14.3			-11.9			5.4		
C&I	-29.2	20.1	16.6	-42.4	5.8	3.8	-8.7	15.3	13.2
Consumer	-37.8	8.7	6.3	-32.2	4.9	3.7	-2.4	11.2	10.3
Real Estate	-19.0	32.0	30.2	-17.7	65.8	61.1	10.0	25.0	26.0
Construction	-65.2	6.3	2.6	-66.8	6.3	2.4	-22.0	4.2	3.1
1-4 Family	-4.7	15.4	17.1	-9.2	43.8	44.9	17.8	12.5	14.0
Multifamily	-39.0	.9	.6	-31.3	4.3	3.3	5.0	1.0	1.0
Commercial	-9.7	9.4	9.9	-18.4	11.4	10.5	16.1	6.7	7.4
Leases	-43.9	2.8	1.8	-33.1	.3	.2	-2.5	1.0	.9
Securities	21.1	13.1	18.5	22.3	14.7	20.2	19.1	16.5	18.6
Liabilities									
Total Deposits	-9.3	73.7	77.9	-5.8	80.4	85.5	7.1	76.8	78.0
NOWs	13.5	4.7	6.2	14.1	4.2	5.4	22.9	5.6	6.5
MMDAs	-5.7	13.8	15.2	2.1	10.3	11.9	19.0	10.4	11.8
CDs	-39.0	12.4	8.8	-43.7	7.9	5.0	-19.0	11.0	8.4
Capital	-21.5	6.3	5.7	-27.3	8.8	7.2	10.6	6.4	6.7
Memo: Total Assets (\$ billions)		\$182	\$156		\$101	\$89		\$3461	\$3648

^aSavings banks are not broken out nationally because they represented only 6.9% of assets for all FDIC-insured institutions in the United States in 1989.

Source: Call Report data for FDIC-insured institutions.

data available). Capital for all FDIC-insured banks rose by 10.6 percent, compared to declines of 21.5 percent and 27.3 percent for New England commercial and savings banks, respectively. Total assets of U.S. banks also rose over this two-year period, although by less than capital, resulting in an improvement in their average capital/asset ratio. The same was not true for New England banks, however. Assets declined by 14.3 percent for commercial banks and 11.9 percent for savings banks. Because capital for both types of institutions decreased by a greater percentage than their assets, their capital/asset ratios showed sharp declines.

The pattern of lending by New England banks has changed substantially over the past two years.

How much of the contraction in the New England banking system reflected a drop in the demand for bank services, and how much was the result of constraints on the banking system?

Just as New England commercial and savings banks had expanded more aggressively than banks nationwide into construction, commercial real estate, and commercial and industrial loans during the boom, they contracted these same loan categories more rapidly during the decline. For example, the amount of construction loans held by banks in New England contracted by over 65 percent during this two-year period. Construction loans showed a substantial but smaller decrease at the national level, and this was the only loan category to shrink by more than 9 percent. In New England, every bank lending category declined, all but one by more than 9 percent. At the same time, securities holdings in New England banks increased at a rate slightly higher than observed nationwide.

While New England banks grew faster than banks nationwide during the boom, and contracted faster than banks nationwide during the bust, it is possible that banking is merely a microcosm of the overall economy of the region. The New England economy also grew faster than the nation during the boom, and contracted faster during the recession. It is

important to determine how much of the contraction in the New England banking system reflected a drop in the demand for bank services, and how much was the result of constraints on the banking system. The next section examines the particular importance of such supply constraints during this period.

II. Implications of Depleted Bank Capital for Future Lending

New England is not the first region to experience significant problems in its banking sector. During the 1980s, banks in the Midwest with agricultural loans, banks in the Southwest with oil loans, and money center banks with Third World loans all experienced serious erosions of bank capital. The New England experience is distinctive because bank capital was lost during a period of increasing regulatory scrutiny, making forbearance much less feasible. The increased emphasis on bank capital reflects national concerns with the huge costs of the savings and loan debacle and new international and domestic standards.

New international capital requirements, the Basle Accord, were agreed upon in December 1987. The accord set uniform international capital standards for all commercial banks, requiring them to maintain minimum capital ratios based upon the degree of credit risk in their asset portfolios. The standards focused on credit risk for broad asset categories and ignored interest rate risk, liquidity risk, and portfolio risk. While only a rough approximation of the riskiness of most bank portfolios, it was a first step in attempting to set bank capital standards related to a bank's ability to weather future potential losses.

A second standard was set by U.S. regulators, the leverage ratio, which required banks to maintain minimum capital standards without directly weighing the credit risks of the assets. This provided a floor for acceptable capital that all U.S. banks were expected to satisfy. Both capital ratios have been phased in and will be completely operative by the end of 1992.

This emphasis on capital ratios occurred at the same time as a substantial erosion of the capital base of New England banks (Syron 1991). Banks with capital/asset ratios below the required level had to either increase equity capital or shrink their asset portfolios. However, these banks could not raise capital from retained earnings in the face of large, continuing loan losses. The other capital-raising al-

ternative, issuing new shares, also was not feasible for many institutions; investors required a large risk premium, making it difficult for sound banks to issue new shares at what they deemed to be a "fair" price. (This is the "lemons problem." See, for example, Myers and Majluf 1984.) Consequently, the only remaining option for many New England banks was to shrink.

A regulatory incentive to shrink can have serious ramifications for the economy as banks tighten credit standards and refuse to renew loans. While large firms typically have alternatives, most small and medium-sized firms rely on banks to meet their demands for credit. Because of their knowledge of local firms and local economic conditions, banks specialize in this segment of the loan market, where their intermediary services are most valuable. (See, for example, Gertler and Gilchrist 1991; Elliehausen and Wolken 1990.)

To determine whether bank shrinkage is the result of a weak economy or low capital ratios, the response of undercapitalized banks to a decrease in capital must be modeled. Because the assets of a bank always equal its liabilities plus capital, a reduction in capital, given liabilities, will result in a reduction of assets. However, bank liabilities are unlikely to remain constant, and, in fact, respond differently depending on whether capital is constrained (Peek and

ferences between the constrained and the unconstrained cases would be expected when examining deposits rather than assets. Consequently, this study will focus on the liability side of bank balance sheets rather than on bank assets, as is common in previous studies (for example, Bernanke and Lown 1992 and King 1986).

The capital crunch hypothesis includes the empirical prediction that, other things equal, poorly capitalized institutions will shrink liabilities much more than well-capitalized institutions. The next section examines the role of bank capital in the shrinkage of New England banks.

III. Empirical Evidence of a Capital Crunch

Section I shows that capital, assets, and liabilities of New England banks all decreased significantly over the past two years. This two-year period includes a recession that was particularly severe in New England. The aggregate bank data cannot distinguish between the decrease in the demand for bank services that normally occurs in a recession and the shrinkage in bank assets and liabilities caused by binding capital regulations. In an attempt to separate these two explanations of New England bank behavior, this section focuses on a cross section of banks from the First Federal Reserve District (New England) in order to determine if the observed bank shrinkage was associated with bank capital positions.

If decreases in assets and liabilities of banks during recessions were solely due to decreased demand, the degree of contraction should be unrelated to capital/asset ratios. If, however, the capital crunch hypothesis is correct, the shrinkage of liabilities and assets should be greater, the lower the capital/asset ratio of the bank.

The Data

The sample includes 419 commercial banks and savings banks in New England from the first quarter of 1990 to the first quarter of 1991. The first quarter of 1990 coincides with the announcement of serious problems in the real estate portfolio of Bank of New England. At that point, bank management, bank examiners, and bank analysts began scrutinizing all New England banks for problems associated with declining real estate prices. Because seasonal factors distort bank balance sheets, comparisons must be made over periods that are a multiple of four quar-

The regulatory emphasis on bank capital occurred at the same time as a substantial erosion of the capital base of New England banks.

Rosengren 1992). In the unconstrained case, a negative capital shock will cause banks to substitute deposits for some of the diminished capital. Thus bank deposits are increased, mitigating the overall shrinkage of the bank. In the constrained case, banks cannot substitute deposits for capital because of the binding capital constraint. Thus, they must shrink deposits, resulting in a more severe shrinkage of assets compared to the unconstrained case. Loans shrink in either case, but shrink by more when a bank is capital constrained. Therefore, more dramatic dif-

ters. Thus, at this time the study can consider only one full year of data.

The sample of banks analyzed comprises all mature New England commercial and savings banks that operated continuously over the period. Any bank that showed a pattern between capital and lending that did not reflect mature bank behavior was omitted. For example, newly formed banks were not included because initially they will expand much more rapidly than mature banks. Thus, inclusion of new banks would reflect expansion due to new formation rather than regular bank operations, providing a spurious positive relationship between capital and lending for the whole sample.⁴

FDIC-insured institutions that merged between January 1989 and the first quarter of 1991 were combined into a single institution for the sample. That is, they were treated as if the merger were consummated at the beginning rather than in the middle of the sample period. Otherwise, merged institutions would have to be dropped and acquiring institutions would experience large increases in liabilities as a result of the acquisition. A separate file omitting institutions involved in nonaffiliate acquisitions was maintained to ensure that this assumption did not significantly affect the results.⁵

Another potential problem is the definition of capital. Capital regulation includes a variety of definitions that use differing measures of capital and assets and different treatments of intangible assets.⁶ Rather than attempt to test all the different definitions of capital, this study uses total equity capital divided by total assets. This definition most closely conforms

⁴ This study also omitted failed institutions, banks that acquired assets of failed OTS-supervised institutions, institutions that consistently maintained a capital/asset ratio above 20 percent, and institutions with either no loan loss reserves, no commercial and industrial loans, or no demand deposits. Data from these institutions would not be comparable to the "mature" banks remaining in the sample.

⁵ The full sample included 419 banks: 49 large commercial banks, 146 small commercial banks, 81 large savings banks, and 143 small savings banks. After excluding those banks that merged with institutions outside of their holding company during the 1989-1990 period, the sample size was reduced to 404 banks. Essentially the same empirical results were obtained with this "clean" sample.

⁶ The risk-based ratios required by regulators are 8 percent for the ratio of total capital to risk-weighted assets and 4 percent for tier 1 capital to risk-weighted assets. The leverage ratio (tier 1 capital to total assets) is 3 percent for banks with the most favorable bank rating of 1. All other banks are expected to maintain capital ratios 100 to 200 basis points above the minimum. Tier 1 capital consists of common equity, qualifying preferred stock, and minority interest in consolidated subsidiaries less goodwill. In practice, tier 1 capital is frequently calculated net of all intangible assets.

to the leverage ratio, the capital standard that is generally the most binding on banks.⁷

The Empirical Test

The capital crunch hypothesis predicts that poorly capitalized institutions will shrink deposits more rapidly than better capitalized institutions, other things (including loan demand) equal. This hypothesis is tested by estimating the following equation, with a positive predicted sign for a_1 .

$$(1) \text{ DEP}_i = a_0 + a_1 \frac{K_i}{A_i} + a_2 \log(A_i) + a_3 \text{FEE}_i + a_4 \frac{\text{CI}_i}{A_i} + a_5 \frac{\text{RE}_i}{A_i} + \epsilon_i$$

The dependent variable is the percentage change in total deposits (DEP) from the first quarter of 1990 to the first quarter of 1991.⁸ The beginning-of-period capital-to-asset ratio (K/A) is calculated using first-quarter 1990 data for total equity and assets.

Limiting the sample to New England banks greatly reduces (though it may not totally eliminate) the variations in loan demand shocks across banks in our sample. It is possible, however, that banks specializing in particular types of loans may experience different demand shocks. Consequently, the regression includes several control variables in order to try to capture potential differences in demand: the logarithm of assets (A), as of the first quarter of 1990; and 1989 calendar year values for the remaining three variables, the ratio of fee income to the sum of total interest and fee income (FEE), the ratio of commercial and industrial loans (CI) to total assets, and the ratio of real estate loans (RE) to total assets. These control variables are intended to capture changes in demand across banks that otherwise might be attributed incorrectly to the capital/asset ratio.⁹

⁷ As of June 30, 1991, of the 20 largest First District commercial and savings banks, none violated tier 1 risk-based guidelines, seven violated total risk-based guidelines, and nine violated a 5 percent leverage ratio.

⁸ DEP (total deposits) is calculated as a bank's total liabilities excluding its total equity; it is composed primarily but not exclusively of deposits.

⁹ These control variables may capture demand shocks that may not be evenly distributed across all banking markets. For example, asset size could be important if large firms are more severely affected by the recession and tend to use large banks as their primary lender. Similarly, servicing fees, commercial lending, and real estate lending may have experienced different demand shocks.

The sample is further segmented in order to verify that it is controlled for possible differences across banks in the degree to which they are affected by demand shocks. Because New England savings banks generally have been less active in lending to businesses, institutions are categorized by whether they have a commercial or a savings bank charter. This provides a further check on whether CI captures differences in demand shocks across institutions. The sample is further split into large bank and small bank categories. (Large is defined as any institution with at least \$300 million in assets, consistent with the classification used in call reports.)

Table 3 reports the results of estimating equation (1) for all FDIC-insured banks in New England and for the four subcategories: large commercial banks, large savings banks, small commercial banks, and small savings banks. The results provide substantial support for the capital crunch hypothesis. Capital ratios are a statistically significant determinant of deposit growth in four of the five regressions, with the estimated capital ratio coefficient significant at the 1 percent confidence level in the large savings bank and the all banks samples. A decrease of 1 percentage point in a bank's capital/asset ratio corresponds to a decline of more than 1 percent in its deposit growth rate for the small savings bank and all banks samples, and an even more dramatic 1.47 percent drop for the large commercial bank sample.

Asset size has a statistically significant negative estimated coefficient in all five regressions, with coefficients significant at the 1 percent confidence level for the all banks and the two savings banks regressions. Fee income has a positive sign in four of the five regressions, although none are statistically significant. This is consistent with the hypothesis that banks relying heavily on fee income were more insulated from the recent demand shocks. Banks with substantial commercial and industrial loans and real estate loans do not appear to have experienced significantly different demand shocks, with real estate loans having a statistically significant effect only in the small savings bank sample.

The results shown in Table 3 support the capital crunch hypothesis: institutions with lower capital ratios grew more slowly (shrank more rapidly) to try to satisfy capital requirements. Furthermore, the results are fairly consistent across types and sizes of banks. The next section examines how the decrease was distributed across categories of deposits.

Changes in the Composition of Deposits

If banks are shrinking to satisfy capital requirements, presumably they will choose to shrink the most expensive accounts, while trying to leave unchanged deposits that provide low-cost funds. In addition, core deposits, such as NOW accounts, pay

Table 3
Determinants of the Percentage Change in Total Bank Deposits^{a,b} at New England Banks, 1990:I to 1991:I

Institution	Constant	K/A a ₁	Assets a ₂	FEE a ₃	C&I a ₄	RE a ₅	n	R ²	SEE
Large Commercial Banks	.21 (.19)	1.47* (.72)	-.03* (.01)	.29 (.17)	.04 (.14)	-.10 (.08)	49	.15	.080
Small Commercial Banks	.31 (.20)	.81 (.53)	-.03* (.01)	.17 (.25)	.03 (.13)	.04 (.12)	146	.01	.120
Large Savings Banks	.58** (.16)	.93** (.22)	-.05** (.01)	-.35 (.24)	-.10 (.12)	.01 (.07)	81	.44	.056
Small Savings Banks	.50** (.13)	1.08* (.47)	-.04** (.01)	.45 (.58)	-.04 (.15)	-.18* (.08)	143	.15	.084
All Banks	.38** (.06)	1.03** (.24)	-.03** (.00)	.11 (.14)	-.03 (.07)	-.07 (.05)	419	.23	.093

^aTotal bank deposits are defined here as total bank liabilities less bank capital.

^bEstimated with a White correction for heteroskedasticity; standard errors in parentheses.

*Significant at 5% confidence level.

**Significant at 1% confidence level.

Table 4
Effect of Capital/Asset Ratios on Deposit Growth, by Category of Deposit,^a 1990:I to 1991:I

	NOW Accounts	MMDAs	Large CDs
Large Commercial Banks	4.18** (1.07)	4.86* (2.40)	7.77* (3.74)
Small Commercial Banks	6.44 (4.93)	3.27* (1.31)	3.31** (1.19)
Large Savings Banks	2.00* (.86)	2.13** (.77)	3.23** (.81)
Small Savings Banks	-1.82 (2.67)	.36 (.96)	5.07 (3.07)
All Banks	1.99 (2.25)	1.82** (.55)	4.40** (1.27)

^aThe equations have been estimated with a White correction for heteroskedasticity and include the same set of explanatory variables as the equations in Table 3. Standard errors in parentheses.

*Significant at 5% confidence level.

**Significant at 1% confidence level.

the same rate to all depositors, while certificates of deposit (CDs) pay different rates depending on prevailing market conditions at the time of issue. Thus, by using CDs as their marginal source of funds, banks are able to segment the deposit market.

Three categories of deposits are examined: NOW accounts, money market deposit accounts (MMDAs), and large CDs. The average interest rates paid nationally in 1990 for these accounts were 4.58, 6.29, and 7.99 percent, respectively (Brunner, Duca, and McLaughlin 1991). If poorly capitalized banks are shrinking to satisfy capital requirements, the greatest shrinkage can be expected to occur in CDs, the marginal and most expensive source of funds, and the least shrinkage in NOW accounts, with the shrinkage in MMDAs between the two extremes.

Equation (1) was reestimated with growth rates by deposit category replacing the growth rate of total deposits. The results, indicating the sensitivity of deposit growth by category to a bank's capital position, are reported in Table 4. The capital crunch hypothesis would imply that the capital/asset ratio would have a larger positive sign, the more costly the deposit account and the more the deposit type serves as the marginal source of funds.

The results in Table 4 support the hypothesis that banks have been reducing the most costly ac-

counts. For the all banks category, large CDs have an estimated coefficient more than twice the size of either of the less costly accounts, and the estimated coefficient is statistically significant at the 1 percent confidence level. While capital ratios have a statistically significant effect on MMDA growth in the all-banks sample, their effect is not significant for NOW accounts. Even though MMDAs and NOW accounts have similar responses, the MMDA response is measured with much greater precision.

The point estimate of the sensitivity of large CD growth to the capital ratio is greater than that for MMDAs in each bank category, and greater than that for NOW accounts in all but the small commercial bank category. In three of the four bank subcategories (small savings being the exception), the capital/asset ratio has a significant effect on both CD and MMDA growth. On the other hand, only large commercial and large savings banks show a significant response to the NOW account equations. This evidence is consistent with banks passively accepting changes in such transactions accounts, implying that their changes would not necessarily be related to the capital/asset ratio of the institution. Thus, the general pattern is confirmed, whereby the more managed accounts such as CDs grow more slowly (shrink more) than NOW accounts when institutions become poorly capitalized. The hypothesis that CDs may be distinguishable as the marginal source of funds for many institutions is also confirmed. In fact, a 1 percentage point decline in the capital/asset ratio implies a 4.4 percent decline in the growth rate of large CDs for the all banks sample and a decline of nearly 8 percent in the large commercial bank sample.

IV. Conclusions

New England banks have experienced substantial shrinkage over the past two years. Some contraction was inevitable after the bursting of the New England real estate bubble and the slowdown in the New England (and national) economy. However, their effects were aggravated by an increased emphasis on bank capital regulation at the same time that New England banks experienced a substantial reduction in bank capital. This article has shown that the shrinkage of banks in New England has not been uniform, as might have been expected if it were related solely to the economic downturn. Rather, poorly capitalized banks have contracted more rapidly than well-capitalized banks. Furthermore, banks

have contracted deposits most in those categories that serve as their marginal source of funds.

To date, the capital crunch appears to be concentrated in New England, although large losses in bank capital have occurred in the mid-Atlantic and some western states. Had we had a national banking system, bank capital would have been able to flow more freely across geographical regions, substantially reducing capital shortages. Because banks typically have had asset concentrations in their own region, the disruption of the lending behavior of banks due to regional economic problems tends to be concentrated in that region. With more nationwide banks, a given bank's capital would be less eroded by problems in any one region of the country, and well-capitalized institutions would be available to fill any gap caused by shrinkage of poorly capitalized institutions.

To determine whether the current capital crunch in New England has resulted in a credit crunch, the analysis must be extended to assets. Banks have several options available to reduce their assets. One possibility is selling securities, leaving the loan portfolio unchanged. Assets would shrink, but the size of the loan portfolio would be unaffected. Alternatively, banks can shrink their loan portfolios, either by selling or securitizing loans, by calling loans, or by tightening credit standards. Loan sales should be preferred by banks because they do not disrupt historical lending relationships. It has also become relatively easy to sell certain categories of loans. For

example, an active secondary market exists for residential mortgages that conform to secondary market standards. It has also become common to sell consumer loans. These loan sales can reduce the stock of loans in a bank's portfolio without affecting its flow of new lending. In that case, credit availability for new loans would be unaffected, despite a large decline in loans reported on a bank's balance sheet.

If banks choose to shrink by tightening credit standards and calling loans, borrowers will be affected only if alternative sources of credit are not available. Large firms with access to national credit markets will be insulated from many disruptions in bank lending. Similarly, firms in the middle market may have alternative sources of funds, such as foreign banks or banks inside or outside the region that are not capital constrained. In addition, insurance companies, venture capital firms, and finance companies have expanded operations to lend in markets traditionally serviced by banks. Therefore, even if banks in one region reduce their lending, credit availability becomes a problem only for those firms that must rely on local banks for their credit, either because they are too small to go outside the region or because banks outside the region and nontraditional lenders are not available.

Unfortunately, the data required to adequately address this question are not yet available. However, current research by the authors will develop a more refined data set and investigate further the link between a capital crunch and a credit crunch.

References

- Bernanke, Ben, and Cara Lown. 1992. "The Credit Crunch." *Brookings Papers on Economic Activity*, No. 2, pp. 205-39.
- Brunner, Allen D., John V. Duca, and Mary M. McLaughlin. 1991. "Recent Developments Affecting the Profitability and Practices of Commercial Banks." *Federal Reserve Bulletin*, July, pp. 505-27.
- Elliehausen, Gregory E. and John D. Wolken. 1990. "Banking Markets and the Use of Financial Services by Small and Medium-Sized Businesses." Staff Studies No. 160, Board of Governors of the Federal Reserve System, September.
- Gertler, Mark and Simon Gilchrist. 1991. "Monetary Policy, Business Cycles and the Behavior of Small Manufacturing Firms." Board of Governors of the Federal Reserve System.
- King, Stephen R. 1986. "Monetary Transmission: Through Bank Loans or Bank Liabilities?" *Journal of Money, Credit, and Banking*, vol. 18, no. 3, pp. 290-303.
- Myers, Stewart C. and Nicholas S. Majluf. 1984. "Corporate Financing and Investment Decisions When Firms Have Information that Investors Do Not Have." *Journal of Financial Economics*, vol. 13, no. 2, pp. 187-221.
- Peek, Joe and Eric S. Rosengren. 1992. "The Capital Crunch: Neither a Borrower Nor a Lender-Be." Federal Reserve Bank of Boston.
- Syron, Richard F. 1991. "Are We Experiencing a Credit Crunch?" *New England Economic Review*, July/August, pp. 3-10.
- Walter, John R. 1991. "Loan Loss Reserves." Federal Reserve Bank of Richmond *Economic Review*, July/August, pp. 20-30.
- White, Halbert. 1980. "A Heteroskedasticity-Consistency Covariance Matrix Estimator and a Direct Test for Heteroskedasticity." *Econometrica*, vol. 48, no. 4, pp. 817-38.

The Financial Condition and Regulation of Insurance Companies: An Overview

Throughout much of the twentieth century, the large insurance companies have been popular symbols of unquestioned strength and stability. The image was not much different for professionals in the financial community: the risks were perceived to be modest in large, diversified insurance companies; managements were considered conservative; and ratings generally ranged from superior to excellent.

A crack appeared in the facade in 1988 when the fourth largest life insurance company sustained well-publicized losses that ate deeply into surplus, but this was considered to be an isolated situation. However, in October 1990 questions were raised about real estate problems in the life insurance industry after the ninth largest life company sustained a major loss as a consequence of a write-down of real-estate-related assets. The value of insurance company stocks declined in late 1990 as the financial community began to take a hard look at the recent changes that had taken place.

During the spring of 1991 the press increasingly focused on the industry, once it became evident that the life subsidiaries of First Executive and First Capital were impaired as a consequence of substantial investments in junk bonds. The seizure of these relatively large life companies by regulators brought to the fore the issues of guaranty fund protection and liquidity runs.

In the summer of 1991, the Federal Reserve Bank sponsored a conference to examine the dramatic changes in risk factors that have transformed the seemingly stable and dependable insurance industries into industries that could arouse widespread public anxieties. How pervasive are the weaknesses that have shown up in a few large insurers? Is there a danger that widespread liquidity pressures could develop? What changes should be made in regulation or in arrangements to protect customers of insurance companies? These are some of the primary questions addressed. Although the immediate concerns have been largely associated with life insurance companies, the confer-

*Richard E. Randall
and Richard W. Kopcke*

Vice President, and Vice President and Economist, respectively, Federal Reserve Bank of Boston. This article summarizes the Bank's economic conference held in June 1991. Information on how to obtain the proceedings appears on p. 45.

ence also devoted considerable attention to property-liability insurance, which perhaps is inherently more risky.

Six papers were presented, each with two or three discussants. The first paper considers insurance companies as financial intermediaries, examining their role in credit markets and the risks inherent in the balance sheets of both life and property-liability companies. The next two papers analyze the structure, conduct, and regulation of domestic life and property-liability insurers. The fourth paper discusses the structure of insurance companies abroad. The final two papers evaluate public policy questions relating to domestic life and property-liability insurers.

A major issue is the quality of the assets currently held by life insurance companies. Some participants stress that the outlook for commercial real estate is negative in a number of regions and that several large companies are heavily exposed. The inadequacy of the capital cushion relative to potential losses is noted. Industry representatives argue, however, that the nature of their commercial real estate assets is distinguishable from that of assets held by commercial banks, and that problems are limited to a few institutions and not systemic to the industry, as was the case with the thrifts. It is generally agreed that no solvency threat is impending for the property-liability industry, although various areas of vulnerability are discussed, including potential exposure to environmental catastrophes. Much attention is focused on the ability of state guaranty funds to function effectively in large failures, and on the nature and degree of protection that should be provided to customers.

Industry representatives and some academics see little need for a federal role in supervision. Some participants argue for a limited federal role, with reinsurance and international activities examples of areas appropriate for federal regulation. Others argue for a more extensive federal role in solvency regulation, although no one advocates eliminating state regulation. With respect to property-liability insurers, however, some argue for phasing out state rate regulation and placing reliance on competitive forces to control prices.

A difference of opinion is apparent between those who would place more responsibility on regulators to prevent excessive risk concentrations from developing, and those who would limit guaranty fund protection in order to enhance market discipline as a constraint on industry risk-taking. Several par-

ticipants note weaknesses in accounting and the difficulty outsiders have in trying to evaluate risk in insurance companies. Some also draw attention to the risk of liquidity runs on life insurance companies thought to be insolvent, illiquid, or weaker than their competitors.

The papers are rich in the variety of matters discussed beyond the major solvency issues mentioned here. Among these are the wisdom of removing rate regulation and/or antitrust immunity in property-liability insurance, federal tax policy with respect to the savings element in various life products, the shrinking presence of U.S. insurers in world markets, mark-to-market accounting, the appropriateness of retroactive loss loading in property-liability underwriting, and the prospects for industry consolidation.

Insurance Companies as Financial Intermediaries: Risk and Return

The paper by Richard Kopcke and Richard Randall was presented as a catalyst to discussion of the evolving risk profile of the industry and the supervisory challenges recent changes entail. It focuses on the implications for risk of the increasing role of life companies in offering investment products, and the vulnerability of both life and property-liability companies to rising interest rates, declining property

Kopcke and Randall stress the need to deal promptly with dangerous risk concentrations and to support investment and other risk with adequate capital.

values, and disappointing corporate profits. It stresses the need to deal promptly with dangerous risk concentrations and to support investment and other risk with adequate capital.

The authors begin by noting the importance of insurers as holders of corporate bonds and commercial mortgages. A number of life companies recently have been funding a significant portion of such assets with relatively short-term liabilities, mostly guaranteed investment contracts (GICs), thus raising both

interest sensitivity and liquidity concerns. Property-liability companies are also vulnerable to increases in interest rates, since their claims are relatively short-term and irregular. Higher interest rates lower the value of their assets, which may have to be sold to meet claims.

The capitalization of property-liability companies has fallen significantly in the past 30 years, while their risks have not diminished. Capital ratios of life companies have remained essentially constant, but many life companies have undertaken investments that are riskier with respect to both possible default and vulnerability to interest rate increases. The paper documents the extent to which life companies with weak capital ratios hold particularly risky assets. The nature of some of the riskier investments of life companies, such as commercial real estate joint ventures, commercial mortgages, and leveraged buy-outs, is such that outsiders have great difficulty in assessing the risk of individual companies.

The recent failures of a few relatively large life companies, and the widely reported vulnerability of additional companies to the depressed state of commercial real estate, warrant a review of how these dangers arose and how they could have been avoided. The authors present several case studies that show characteristics in common with the extraordinary asset quality problems experienced by large banks in recent years.

In general, risk concentrations developed over several years, during which time the institutions appeared to be in sound condition. A turning point occurred, adversely affecting the areas of risk concentration, and it soon became apparent that the institutions were severely, often fatally, damaged. With respect to both banks and insurance companies, supervisory action would have to have been directed at the risk concentration *before* the triggering economic event (disruption of the junk bond market, crash of real estate values, or the like). While the analysis by Kopcke and Randall does not equate the degree of the insurers' problems with those of banks, it does suggest that supervisory restraints on excessive risk-taking are equally appropriate in both industries.

Jeffrey Cohen sees a regulators' dilemma in the Kopcke/Randall proposal for early intervention to limit risk concentrations. He notes that the circumstances may not be clear when managements take actions that get them into trouble, and questions whether regulators should substitute their judgments for those of management or the markets. He also notes that regulators have a conflict between promot-

ing solvency of the company and keeping insurance affordable to the consumer.

Cohen sees the fundamental industry problem as insufficient profitability, leading to greater risk-taking and weaker capital ratios. He attributes this in part to the presence of too many companies, and he would remove barriers to consolidation and not allow banks to enter the field. Cohen believes that life insurers are not profiting from the issuance of GICs because they write them at too narrow a spread between the yields they receive on their investments and the yields they pay on GICs, not allowing for an adequate risk premium.

He attributes the decline in property-liability insurers' capital ratios to a shift from property to liability lines, which permit a longer earning period before claims must be paid. He argues that the property-liability industry is not sufficiently profitable to support its present capitalization. Cohen calls for more mark-to-market disclosure and action to make the demutualization process easier.

In his comments, Thomas Maloney reviews the transformation of the larger life insurance companies over the past 20 years into multi-line financial companies. He finds that the majority of companies have adapted well to the more competitive environment. The larger companies are generally safer because of geographic and product diversification, and failures have generally involved small companies.

While a number of life companies underpriced products in recent years and overpaid to attract funds, most have rectified their mistakes. The few large life failures involved levels of risk-taking well above that of the rest of the industry, and the likelihood of widespread failures across the industry is low because of diversification and relatively high asset quality. Insurance companies perform better in a downturn than banks, a result of their greater geographic diversification and the character of their assets.

In reviewing current "reform" proposals, Maloney predicts that the outcome of the federal versus state regulation issue will depend on how quickly the states can strengthen supervision. He notes one fault of the current guaranty system: the prudent companies are burdened with the eventual losses incurred by their overly aggressive competitors. He also foresees industry consolidation in order to meet capital requirements.

Frederick Townsend's comments focus on the asset risks of life insurers, particularly the junk bonds that forced some rapidly growing companies into

conservatorship and the real-estate-related assets that are creating capital losses in some of the large, established life companies. He emphasizes the poor credit quality of the junk bonds acquired, particularly by Executive Life, and he argues that the recent failures might not have occurred if regulations had limited junk bond concentrations.

Townsend points out that analysis of insurance companies must distinguish between the operating companies and the parent. He cites instances of damaged life companies with strong parents, and others where the problem was largely in the parent.

He notes the importance of product design and duration matching in avoiding runs by policyholders. Townsend also notes that while high capital ratios increase the odds of survival, they do not guarantee it. He concurs with Kopcke and Randall that capital ratios decline in the problem realization phase, not in the earlier, risk-taking phase.

The Structure, Conduct, and Regulation of the Life Insurance Industry

Kenneth Wright presents an account of financial conditions in the life insurance industry and the changed environment and competitive pressures that have so altered the industry in recent years. He reviews prior instances of liquidity pressures, the disintermediation periods of 1966, 1969, and 1979–81. He traces the development of new instruments, particularly universal life, variable life, flexible premium variable life, single-payment annuities, and GICs, and the corresponding shifts in investment strategies.

Wright finds the measurement of industry profitability difficult, but presents data suggesting a significant decrease in the 1979–87 period. He shows that capital ratios have declined in recent years, unless security valuation reserves are included in capital, in which case they have been virtually unchanged for the past decade.

Wright estimates that the life insurance industry holds \$60 billion to \$70 billion in junk bonds, but notes that the historical default record on corporate bonds has been favorable, and an important offset to the increased holding of riskier bonds has been greater holdings of Treasury and agency securities. With respect to commercial mortgages, Wright notes the rising delinquency numbers, but points out they have not yet reached the peak levels of 1976.

The industry is greatly concerned about the solvency issue even though it believes that serious

problems are limited to relatively few companies. An insurance company failure exposes even healthy firms to the danger of runs, and the integrity of life insurance products may be called into question, deterring purchases.

Guaranty fund assessments are also an issue, although these payments can often be passed along to the states in the form of tax credits. The industry has supported efforts to modernize state solvency regulation and improve coordination between states through the work of the National Association of Insurance Commissioners (NAIC).

Wright concludes that the industry is not as

Wright sees the life insurance industry as having been forced by competitive pressures to accept higher risks, while state regulators have had to struggle to stay abreast of marketplace developments.

financially sound as it was a dozen years ago, as a result of reduced profitability and greater financial risks. He sees the industry as having been forced by competitive pressures to accept higher risks, while the state regulators have had to struggle to stay abreast of marketplace developments. Wright sees the troubles of a few companies as presenting real problems for the industry and its regulators.

In his discussion of Wright's paper, Terence Lennon contrasts the environment for life insurers that existed in previous decades with the one that emerged in the late 1970s and early 1980s as a result of the destabilization of interest rates. Insurance customers were transformed from savers to investors, and life companies developed new products that met customer demands but increased interest rate risk and credit risk for the insurers.

A decline in margins—the difference between the yields earned on assets and those paid on liabilities—depressed capital ratios somewhat; more importantly, various accounting innovations such as securitization and financial reinsurance diminished the validity of book capital. The cushion that had long

existed because of the industry's conservative accounting disappeared.

Lennon uses the Executive Life case to illustrate that aggregate limits can work for insurance companies, but do little good if imposed after companies have overinvested in risky assets. Lennon believes that conditions now are right for the adoption of a risk-based capital measure. He anticipates some federal regulatory role, and suggests greater conservatism could be induced in the industry through federal tax policy. Lennon foresees a 20 percent reduction in the number of life companies during the 1990s.

Kenneth Pinkes directs his comments to the fundamental forces he sees at work in the financial services industry. His message is that business risk will continue to rise as the successful innovators become more efficient and stronger and the weak become weaker. Financial institutions, including insurance companies, will become more susceptible to shocks.

Pinkes identifies two groups of fundamental forces, the effects of information technology and changes in the regulatory and public policy environment. The first set of forces will result in product unbundling, economies of scale in a broader range of products, and managerial complexity. Among the second group of forces will be greater tolerance for concentration, greater willingness to subordinate regulatory sovereignty for common global or regional standards, greater acceptance of the blurring of boundaries between regulated and nonregulated sectors, and greater insistence on market discipline. These forces will place increased demands on managements already under severe testing.

Roberi Schneider challenges Wright's conclusion that the life industry is not as financially sound as it was a dozen years ago. He notes that the introduction of interest-sensitive products permits companies to compete on the basis of volatile interest rates without providing overly risky guarantees with respect to rates in the distant future.

For mutual companies, participating whole life policies are able to compete with newer products such as universal life because the dividend paid to policyholders has always included a significant contribution from interest earned in excess of the guaranteed rate. It was primarily the stock companies that had to redesign their products to compete in the environment of the 1980s. While annuity products, both single-premium deferred annuities and GICs, generate more investment risks, they have little or no mortality risk. The use of sophisticated investment

management techniques can insulate an insurer fairly well from interest rate risk. The recent shift toward greater holdings of liquid assets has mitigated the increased liquidity risks of GICs.

The level of public concern over life insurance companies' holdings of junk bonds is misplaced except with respect to a very few companies, Schneider states. Most holdings are in the least risky category of junk bonds, and much of what is classified as junk is private placements with greater security than the stereotypical junk issue. Mortgages and real estate investments represent a more significant asset in most life companies, but even here concerns seem overstated. The character of insurance company real estate loans is quite different from the construction loans held by banks. Schneider considers the severity of the real estate problems of life companies to be comparable to those of the 1975-76 period, which did not threaten company solvency.

The Structure, Conduct, and Regulation of the Property-Liability Insurance Industry

J. David Cummins and Mary Weiss address a number of complaints, accusations, and expressions of concern with reference to property-liability insurers. For the most part they find little legitimate basis for these particular areas of dissatisfaction with the industry, but they do identify some serious problems that need to be examined.

The authors find the industry to be competitively structured in most business lines, with numerous firms, relatively easy entry, and satisfactory concentration levels. Much of the blame for premium inflation is put on factors beyond the control of the industry. They find the organizational structure of the industry, including its distribution systems, to be logical. They examine cash flow underwriting—that is, reducing prices during periods of high interest rates in order to increase cash flow and have more investable funds—and conclude that it is a natural practice in competitive markets.

The authors also discuss retroactive loss loading, where insurers price new policies to help absorb past losses. They present an argument that insurers can, and perhaps must, price in this way in situations where a number of insurers incur abnormal losses at about the same time.

Cummins and Weiss find internal rates of return and returns on equity to be reasonable, despite complaints by some that profits are excessive and

protests by the industry that profits are insufficient to support an adequate surplus. However, they do see supply problems in the auto and workers' compensation lines if profitability is not improved, and they note the correlations between inadequate pricing of certain lines and intensive rate regulation.

The authors do not see any clear indication of an impending insolvency crisis among property-liability insurers. However, they express unease with the level of reinsurance receivables to surplus and with the fact that many reinsurers are virtually unregulated. They are also nervous about the quality of bond portfolios, fearing that some companies have invested a substantial portion of their assets in bonds of near-junk quality. In general, Cummins and Weiss consider solvency surveillance by regulators to be inadequate. They call for improved statutory statements both to facilitate improved surveillance and to permit more sophisticated research on the underwriting cycle and the causes of insurance crises.

Roger Joslin reinforces the Cummins and Weiss arguments that the property-liability insurance industry is intensely competitive, and that much of the rhetoric concerning affordability, availability, insurance cycles, and profitability is unjustified. Joslin emphasizes the political demagoguery associated with much rate regulation, and clearly sees little justification for such regulation or for barriers to firms exiting a state or line of business.

He does not see the industry facing a solvency crisis, and he argues that most failures of property-liability companies are preventable, or at least containable if laws are enforced and regulatory action is timely. Joslin sees a need to improve insurance accounting, to hold reinsurance to a high standard, to be skeptical of particularly rapid growth, and to defer the booking of underwriting profit until well after the close of the accident year. Joslin would also reduce the profit opportunities and increase the risk of loss to insider manipulators through a broader definition of voidable preferences and easier reversal of detrimental transactions with financially interested parties.

James Stone applauds the Cummins/Weiss paper for the issues it raises, but wishes the authors had gone further in developing answers to the difficult questions they raised. On the subject of competition, Stone notes that direct response insurance marketing can produce the lowest distribution costs, as a result of economies of scale. Under regulatory schemes that look only at cost and ignore the level of service provided, direct writing would be favored over indepen-

dent agents. This could lead to a more highly concentrated industry, to the detriment of competition.

Since the authors do not identify the cause of commercial insurance cycles, Stone offers his own theory. He attributes such cycles to market signaling, or use of competitors' price movements as a basis for

Cummins and Weiss call for improved solvency surveillance by regulators and more sophisticated research on the underwriting cycle and the causes of insurance crises.

a firm's price changes. This phenomenon exists because of a dearth of hard evidence on which to base pricing decisions, and will continue as long as underwriters lack the necessary information.

With respect to solvency, Stone disagrees with the authors' suggestion that, without further research, the solvency threat to the property-liability insurance industry cannot be distinguished from the savings and loan disaster. Investment returns are a sufficiently small component of price, and market shares sufficiently price inelastic in the short run, to keep the industry's risk exposure within bounds. A number of firms in the industry are likely to fail in the coming years, however, and the authors' complaints about obsolete accounting and weak reinsurance are valid.

Stone notes the authors' statement that availability and affordability of auto insurance are beyond the control of the insurance industry. He believes that it is in the industry's self interest to serve as a catalyst for change, lessening dependence on the tort mechanism, tightening fraud control, and reexamining the notion of compulsory insurance. He favors a tempering of rate spreads between high-cost urban areas and low-cost suburban areas.

The Structure and Regulation of Insurance Markets Abroad

Sotirios Kollias describes the insurance industries and regulatory regimes of the major industrialized countries and discusses the dramatic changes

taking place in conjunction with European integration. Most European insurance markets have historically been national markets separated by restrictive regulation and other obstacles to entry. An exception is reinsurance, for which an international market exists. Insurance markets have been most highly developed in the United Kingdom, the Netherlands, Japan, and the United States, somewhat less so in France and Germany, and much less developed in the southern European Community (EC) nations. Kollias estimates that rates of return on investments by insurance firms have been highest in the United Kingdom because of U.K. companies' relative freedom to invest in equities. Some measures indicate that companies in the United States and Japan are less efficient than companies in some of the EC countries.

Nonlife companies in most EC countries have been losing money on underwriting but have continued to show profits as a result of sharp increases in asset values. Life companies in Europe have generally been profitable, but Kollias did point out that the five big composite (multi-line) companies in the United Kingdom lost more than \$1 billion in 1990. These companies have, nonetheless, been involved in less damaging competition than their counterparts in the United States.

The separation of European insurance markets began to erode in 1988, and since then a series of changes have been underway. Kollias discusses the principal EC agreements, the Single European Act of 1987 which included a program of financial integration, and proposals for harmonization of supervision of investment services. Integration of insurance activities has followed two separate paths, with nonlife large commercial risk and individual life policies being sold abroad under home country control, but "mass risk" life and nonlife insurance being sold under host country regulations. More recent proposals are expected to permit the free supply of insurance under home country rules.

The lowering of international barriers and deregulation are rapidly producing a much more competitive environment for insurance activities in Europe. Important structural changes are also taking place through mergers, joint ventures, cross-sector subsidiaries, bank/insurance conglomerates, and network distribution alliances.

In most European countries banks have not been able to underwrite insurance, and life and nonlife companies have been segregated. This separation is likely to be ended soon. Banks have been allowed to

distribute insurance products, although insurance companies have generally not been allowed to distribute non-insurance products.

The European integration of banking and insurance in the form of mergers, establishment of subsidiaries, and cross-participation contrasts with the strict limitations on such operations in the United States and the prohibitions in Japan. EC draft directives call for the close cooperation of insurance and bank regulators if a bank or holding company controls an insurance company, however.

Henry Parker points out that the insurance market in the United States, while still the world's largest, is slipping rapidly in its share of world premium volume. He criticizes the domestic industry because so few companies participate aggressively in the expanding overseas markets. While substantial impediments to entry exist in some national markets, it can be done and it is getting easier as a result of federal efforts toward freer international trade.

Parker sees 1995 as the earliest date for real insurance market uniformity in the EC. He anticipates some very substantial reductions in insurance prices in several countries, citing Italy, France, and Luxembourg as examples of the wide variations in premiums for identical exposures. He also sees advantages in terms of expense reduction, product

Kollias notes that the lowering of international barriers and deregulation are rapidly producing a much more competitive environment for insurance activities in Europe.

innovation, and achievement of critical mass. Distribution systems will be altered, with more insurance sold through branches of affiliated banks and other financial service providers. An important stumbling block to rapid completion of the insurance directive is agreement on uniform accounting practices.

One concern for U.S. companies expanding into Europe is the possible reemergence of protectionism, particularly if transition problems severely damage long-protected European companies. There is some

risk that a reciprocity standard might replace national treatment, to the detriment of U.S. companies.

Parker notes the importance and potential of the insurance market along the Pacific rim. He also calls attention to the acquisitions of U.S. insurance companies by foreign insurers.

Steven Skalicky reviews insurance market structure in Asia, Latin America, and Eastern Europe to complement Kollias's analysis, which focused primarily on the EC. He makes it clear that barriers that preserve fragmented national markets are under attack around the world.

Asia has the potential to be the fastest-growing market in the 1990s. Japan, the dominant market in Asia, is characterized by a relatively few large companies, including most of the top 10 insurance companies in the world. Japanese companies have been strictly supervised and limited as to their range of investments. Proposals would liberalize the asset restrictions, and greater flexibility in premium rates was permitted recently.

While the Japanese market is technically open to foreign competition, entry has been difficult. Japanese insurers have not been aggressive in overseas operations, but have the potential for being so. The attraction of Asian countries is not current premium volume, but the potential for growth as they become more industrialized.

In Latin America, Skalicky is most optimistic about Mexico, where the insurance industry is growing rapidly and restrictions on outside ownership have been liberalized. The transition from state control in Eastern Europe eventually will also provide opportunities, as reforms permit foreign participation and ownership and economic changes produce growth.

Skalicky sees unprecedented challenges to the insurance companies, consumers, and regulators. Large insurers that have the capital and resources to penetrate rapidly growing insurance markets may, if successful, survive the global consolidation of the industry. Consumers should benefit from less expensive insurance, but will face increasing risks of insurer insolvency. Insurers' reliance on growth in the value of real estate and securities to offset underwriting losses eventually leads to problems. The challenge to insurance regulators to anticipate and deal with problems in foreign markets is formidable.

Public Policy and Life Insurance

Gerard Brannon proposes a framework for evaluating regulatory and tax policies in the life insurance

market. He begins by distinguishing between the risk coverage and the savings elements in the products of life companies, noting the significant tax benefits of the savings component. He presents historical data to show that since 1955, life company reserves have shifted from life insurance to pension and annuity products and life insurance reserves have declined as a percentage of household financial assets. Life insurance in force as a percentage of personal income has increased, however, as consumers shifted from whole life policies, which have a large savings element and require greater reserves, to term insurance. Despite this trend, evidence suggests that consumers still buy too little life insurance.

State regulation of life companies requires the maintenance of adequate reserves and limits the investment risk that can be assumed. In the late 1980s, the historic redundancy in reserves appears to have eroded and investment restrictions failed to protect policyholders from the risk of new financial

Brannon would support a guarantee of the ability of insurance companies to fulfill term life insurance contracts, but would not support the protection of savings.

innovations or the danger of disintermediation. The recent development of variable and universal life policies has been accompanied by higher-risk investments, but also the opportunity for the investors to make risk choices.

State regulators provide limited solvency guarantees for policyholders, funded by levies on competing companies. In some states insurance companies may apply such levies as credits against premium taxes, effectively transferring losses from the industry to the states. Brannon notes the relatively small volume of guaranty fund assessments in the period from 1975 to 1989 and expresses the view that solvency problems currently facing life insurers are clearly not in the same league as the solvency problems of banks and thrifts.

Brannon points out that the Pension Benefit Guaranty Corporation (PBGC) and state guaranty

funds are competitors. When a company purchases an irrevocable contract for an annuity to cover pension liabilities, the guarantee shifts from the PBGC to a state fund. This may work to the benefit of the employer but to the detriment of workers, who have no say in the choice of an insurer. Nonetheless, Brannon argues against federal support of such annuity obligations, using the First Executive case to illustrate his point.

If it is in the public interest to encourage life insurance purchases for the protection of dependents of breadwinners, Brannon would support a guarantee of the ability of insurance companies to fulfill term life insurance contracts, and he would expect such a guaranty program to be successful. However, he would not support the protection of savers and he deplores the current tax advantages that encourage the intermingling of insurance and investment features, complicating the development of an appropriate guaranty scheme for insurance.

Joseph Belth confines his discussion to the issue of federal income taxation of the inside interest in cash-value life insurance and life annuities. Individuals tend to postpone the distressing subject of life insurance, and therefore a major expense for insurance companies is the commission paid to agents to perform the "anti-procrastination" function. Because natural premiums for life insurance are very low for young purchasers, companies do not receive sufficient revenue to compensate agents. Furthermore, the very high premiums in later years tend to produce adverse selection as healthier members drop insurance. Both of these problems can be mitigated by level-premium, cash-value insurance, which creates a savings component. The federal income tax on the inside interest is generally deferred. Life annuities, which provide regular payments over an individual's lifetime, make sense only in periods of low interest rates, because one can obtain almost as high a return investing principal directly during high-rate periods without destroying the principal, as happens with an annuity. A life annuity may have a lengthy accumulation period before the beginning of the liquidation period, and here again federal income taxation on inside interest is generally deferred.

A theoretical argument can be made that deferred tax treatment of inside interest in these two situations can no longer be justified. Cash-value life insurance is of increasing benefit to high-income individuals, and life annuities are increasingly used solely because of tax considerations. Nevertheless, Belth argues that current taxation of the inside inter-

est would have a "devastating impact on the life insurance industry and would threaten its very survival." He also believes the industry has sufficient political clout to discourage any legislative attempt to impose current taxation.

Earl Pomeroy brings a regulator's perspective to the issues raised by Brannon. He contends that the sophistication of regulatory oversight has been improved in response to the lower capitalization levels, slimmer profit margins, and higher risks found in the life insurance industry today. Pomeroy cites the improved system for bond evaluation, a model law covering bond concentrations, limits on junk bonds, and progress toward reserve requirements and limitations on other higher-risk investments. While such regulatory activity has the necessary effect of lowering investment returns and restricting capital flows to particular activities, it is wholly appropriate because solvency protection is the regulator's first priority.

Pomeroy discusses such consumer protection regulations as required disclosures of product characteristics and minimum product quality standards. He chides Congress for attempting to achieve social goals through the imposition of costly market restrictions.

With respect to guaranty funds, Pomeroy agrees with Brannon that they can dull consumer sensitivity to insurer risk exposure, but finds that they serve a critical role. Despite assessment limitations, Pomeroy is reasonably hopeful that the guaranty fund mechanism has sufficient capacity, on a state-by-state basis, to handle a major life insurance failure.

After briefly reviewing the history of state insurance regulation, including recent activities of the National Association of Insurance Commissioners (NAIC), Pomeroy lists several concerns state regulators have with federal regulation of insurance. He maintains that federal officials tend to overstate the solvency problem, because of their sensitivity to the thrift failures and because they view the Executive Life case as a harbinger of trouble for the life industry generally. Newly implemented state reforms should be given time to work. Pomeroy argues that political pressures could lead to a situation where federal solvency regulation is imposed alongside state regulation of rates with the two sets of regulators pursuing conflicting objectives. Pomeroy does not expect a specific federal regulatory proposal to have much political appeal, even though the general concept might.

Warren Wise challenges Brannon's characterization of the cash value in permanent life insurance as

being equivalent to a savings account. He argues that it arises from the leveling of premiums and is an integral part of providing lifetime protection at an acceptable price. The tax-free inside buildup is a subsidy to encourage life insurance protection, not savings.

Wise acknowledges that the industry is more vulnerable to failure than it once was, although his proposals for dealing with the problem are at odds with Brannon's. Rather than limit protection to death benefits, as Brannon would do, Wise would cover all policyholders. However, he would want all interested parties to share in losses when an insurer fails, including insurance sales representatives, policyholders, and state governments.

Guaranty fund assessments should be risk-based and collected on a regular basis so that the heaviest impact will fall on those insurers most likely to fail. Sales representatives should have an incentive to recommend safe companies, and states should have an incentive to devote adequate resources to solvency regulation. State contributions could be in the form of the tax offset for guaranty fund assessments that already exists in several states. Insurance consumers should share the burden by recovering less than the full amount due them.

Wise would improve regulation by linking capital requirements to risk, strengthening investment restrictions, improving accounting practices, and better controlling reinsurance transactions. Regulators must be provided sufficient resources to carry out their responsibilities.

The question remains of who should administer solvency regulation, and Wise would prefer that it be done without federal involvement if the states can adopt and enforce strong, uniform solvency standards. However, if a federal role proves to be necessary, he would prefer that federal involvement be limited to the setting of minimal standards, oversight, and the ensuring of compliance.

Public Policy and Property-Liability Insurance

Scott Harrington makes some very specific recommendations as to what changes should, and should not, be made to property-liability insurance regulation. He would like to reduce guaranty fund coverage in order to increase market discipline. He does not think a case has been made for a federal regulatory role, and believes that federal supervision could actually increase total insolvency costs. Harrington would like to see the abandonment of state

rate-setting, but would not alter the industry's anti-trust exemption.

With respect to guaranty funds, Harrington argues that guarantees result in policyholders having reduced incentives to buy coverage from safe insurers; the market collectively has more information and knowledge than the regulators, and the spreading of insolvency losses through guaranty funds can reduce pressure on government to commit adequate resources to solvency monitoring. It would be desirable to require a large co-payment from the policyholders, especially those who are best able to monitor insolvency. Harrington also makes a case for post-insolvency assessments being superior to an accumulated

Harrington argues that rate regulation of property-liability insurance has little or no justification, and would limit the regulatory role to requiring appropriate information disclosure.

fund. The arguments presented against federal regulation of property-liability insurers draw heavily on the thrift experience, and particularly the role of Congress in condoning forbearance for insolvent institutions.

Harrington argues that rate regulation of property-liability insurance has little or no justification, and he would limit the regulatory role to requiring appropriate information disclosure. The industry is highly competitive, with ease of entry, and market forces can most efficiently determine rates. Harrington contrasts the industry to public utilities, where rate regulation is necessary. Rate regulation can result in insurers exiting certain lines or states, reducing net worth and thereby increasing insolvency risk; it can also result in insurers being less innovative. Regulation can directly increase expenses and distract management as a result of the rate hearing process.

Harrington sees the cooperative development of policy forms and sharing of loss data as entirely constructive, lowering costs, easing entry, and increasing forecast accuracy. He sees the forecasting of

future losses by advisory organizations as serving a useful function to the extent that they improve individual insurer forecasts. He is concerned that a substantial change in the industry's antitrust exemption could lead to higher prices and less stability, and result in a surge of costly litigation.

J. Robert Hunter vigorously challenges Harrington's characterization of the property-liability insurance market as highly competitive, as well as his proposal to remove rate regulation while preserving the industry's exemption from antitrust laws. Hunter presents evidence that the public does not have sufficient information to select insurance companies on the basis of cost or service quality. He also cites findings that collusion on rates has been the norm, not the exception, in the industry. Hunter reviews the mechanism by which the Insurance Services Office, an industry service organization, provides insurers with advisory rates. He argues that, even with plans to exclude expense factors from the rate data, some critical components of the rate formula will still be provided that instead should be calculated independently by individual insurers, if collusion is to be prevented.

Hunter could agree to easing or even phasing out rate regulation, but only if all anticompetitive forces were eliminated. Specifically, he mentions the antitrust exemption, the anti-rebate laws, the anti-group laws, the barriers to entry by banks, the information gap, and the underwriting selection problem.

With respect to solvency, Hunter challenges Harrington's proposal to decrease guaranty fund coverage in order to improve market discipline. He would expand coverage for personal lines and small businesses. Even with respect to large commercial customers, he notes that loss of insurance protection could have secondary effects on the public when the business, as well as the insurance company, fails. Hunter calls for federal minimum standards for solvency regulation, and direct federal regulation of alien reinsurance and alien surplus lines markets.

Robert Litan agrees with most of Harrington's points, but he would not reject a federal solvency role and would draw different lessons from the thrift crisis. Litan faults the state regulators for their performance in connection with the larger failures of property-liability insurance companies in recent years. He attributes recent efforts by the NAIC to improve state regulation to the threat of federal regulation. Litan proposes creating a federal regulatory program and a national guaranty fund system as an alternative to state regulation and guaranty funds.

Insurers that chose the federal system would no longer be subject to rate regulation. While Litan acknowledges some adverse selection problems with his proposal, he sees it as a way of forcing reform of the state systems, or having property-liability insurance regulation gravitate to the federal level.

Litan draws on his interpretation of the thrift crisis to support the idea that a pre-funded guaranty system would be superior to the usual post-insolvency assessment procedure. He points out that thrift regulators engaged in forbearance largely because of insufficient funds to resolve failed institutions.

Litan is concerned that major exogenous events pose a substantial threat to the industry, citing specifically a potential major earthquake and possible court rulings making insurance companies responsible for the cost of cleaning up hazardous waste sites. He suggests steps that could be taken in advance to protect the industry from being overwhelmed by such calamities.

Richard Stewart briefly outlines what he sees as the major issues in rate regulation and in dealing with the underwriting cycle. He then turns to the issue of solvency and argues that insolvency is a natural outcome for a property-liability insurer.

It is the liabilities of the insurer, not the assets, that are of most concern, and these liabilities extend far into the future. In Stewart's view, the future is not going to be like the past, and therefore it is nearly impossible to estimate the extent of these liabilities for pricing or reserving purposes. In the general liability line the threats are systemic, further adding to the industry's susceptibility to catastrophes on the liability side. Moreover, the industry is intensely competitive, and the incentives and rewards are concentrated on the front end of a transaction, with willingness and ability to pay claims coming much later.

If it is the duty of the regulator to prevent insolvencies, it is very hard to accomplish this by early detection and swift action because of the uncertainty about the extent of the liabilities. However, it is easy to forbear and avoid recognition of insolvency for several years, thereby escaping responsibility. In Stewart's view, this perverse incentive for the regulator increases the risk of even greater losses.

Our system of compensation for accidents functions through an insured civil liability procedure. In the event of insurance company insolvency, the victims include not only direct policyholders but large groups of individuals, whose only link may be the use of a common product or exposure to a form of

pollution, and who are terribly hurt by the insurance company insolvency. We should not think only of corporate America in considering guaranty fund protection surrounding the property-liability insurance system.

Stewart believes that state regulation, with improvements such as those currently in process, can do a satisfactory job of detecting and acting against emerging insolvencies. However, liquidation and guarantees for large-scale general liability insolvencies should be managed at the national level.

Conclusions

The ability of domestic insurance companies to meet their obligations is vital not only to the welfare of their customers but also to the economy and social fabric of the country. In recent years the structure of the life insurance industry has changed in a way that has increased the risk of major insurers becoming insolvent or illiquid. Capital ratios have not increased in response. At the same time the property-liability

insurance industry has become more leveraged and perhaps more vulnerable to large-scale losses.

Opinions differ widely as to the extent and duration of the current weaknesses in the asset quality of life insurers, but it is generally agreed that state regulation and the system of guaranty funds are being materially strengthened by various initiatives. Experts disagree, however, about the ability of even strengthened state systems to avert solvency problems or to safeguard policyholders and others in the event of failures of major insurers. Agreement on the desirability and extent of protection to be provided for policyholders, pensioners, and savers dependent on an insurance company's ability to pay, would facilitate determination of what, if any, federal role is desirable in regulation or in administering guaranty funds.

Congressional interest in examining the insurance industry, continuing downgrades in ratings of individual companies, and the prospects for a prolonged period of depressed commercial real estate values, all suggest that insurance industry solvency issues will be with us for some time.

Conference Series

The Federal Reserve Bank of Boston Conference Series

No. 1	Controlling Monetary Aggregates	June 1969
No. 2	The International Adjustment Mechanism	October 1969
No. 3	Financing State and Local Governments in the Seventies	June 1970
No. 4	Housing and Monetary Policy (out of print)	October 1970
No. 5	Consumer Spending and Monetary Policy: The Linkages	June 1971
No. 6	Canadian-United States Financial Relationships	September 1971
No. 7	Financing Public Schools (out of print)	January 1972
No. 8	Policies for a More Competitive Financial System	June 1972
No. 9	Controlling Monetary Aggregates II: The Implementation	September 1972
No. 10	Issues in Federal Debt Management	June 1973
No. 11	Credit Allocation Techniques and Monetary Policy	September 1973
No. 12	International Aspects of Stabilization Policies	June 1974
No. 13	The Economics of a National Electronic Funds Transfer System	October 1974
No. 14	New Mortgage Designs for Stable Housing in an Inflationary Environment	January 1975
No. 15	New England and the Energy Crisis (out of print)	October 1975
No. 16	Funding Pensions: Issues and Implications for Financial Markets	October 1976
No. 17	Minority Business Development	November 1976
No. 18	Key Issues in International Banking	October 1977
No. 19	After the Phillips Curve: Persistence of High Inflation and High Unemployment	June 1978
No. 20	Managed Exchange-Rate Flexibility: The Recent Experience	October 1978
No. 21	The Regulation of Financial Institutions	October 1979
No. 22	The Decline in Productivity Growth	June 1980
No. 23	Controlling Monetary Aggregates III	October 1980
No. 24	The Future of the Thrift Industry	October 1981
No. 25	Saving and Government Policy	October 1982
No. 26	The Political Economy of Monetary Policy: National and International Aspects	July 1983
No. 27	The Economics of Large Government Deficits	October 1983
No. 28	The International Monetary System: Forty Years After Bretton Woods	May 1984
No. 29	Economic Consequences of Tax Simplification	October 1985
No. 30	Lessons from the Income Maintenance Experiments	September 1986
No. 31	The Merger Boom	October 1987
No. 32	International Payments Imbalances in the 1980s	October 1988
No. 33	Are the Distinctions between Debt and Equity Disappearing?	October 1989
No. 34	Is There a Shortfall in Public Capital Investment?	June 1990
No. 35	The Financial Condition and Regulation of Insurance Companies	June 1991

Conference Series



The Financial Condition and Regulation of Insurance Companies, proceedings of a conference sponsored by the Federal Reserve Bank of Boston, is now available. Each year since 1969 the Bank has gathered a distinguished group from universities, government, business, and finance to discuss important issues affecting economic policy. The Bank's June 1991 economic conference examined recent trends in the insurance industry, considering both life and property-liability companies. A group of academics, economists, government regulators, and industry officials met to consider the competitive pressures, the dramatic changes in risk factors, and the decline in asset quality currently facing the industry. The proceedings include all the conference papers and discussions; the volume begins with an overview of the conference, prepared by the editors. Among the authors and discussants are Joseph M. Belth, Gerard M. Brannon, Jeffrey Cohen, J. David Cummins, Scott E. Harrington, J. Robert Hunter, Terence Lennon, Robert E. Litan, Earl R. Pomeroy, Richard E. Stewart, James M. Stone, Frederick S. Townsend, Jr., and Kenneth M. Wright.

Copies of conference volume no. 35, *The Financial Condition and Regulation of Insurance Companies*, may be obtained without charge on request to the Research Library—D, Federal Reserve Bank of Boston, P.O. Box 2076, Boston, Massachusetts 02106-2076.

Please type or print

Name _____ Title _____

Organization _____

Address _____

Please check

- Please send Conference Series No. 35 *The Financial Condition and Regulation of Insurance Companies*.
- Please send back-issue Conference Series No. _____

Note: Single volumes will be mailed without charge, but a \$10.00 payment (check drawn on a branch of a U.S. bank) will be required for more than 10 volumes or more than 10 copies of the same volume.

PLACE
STAMP
HERE

Research Library—D
Federal Reserve Bank of Boston
P.O. Box 2076
Boston, Massachusetts 02106-2076

The Municipal Bond Market, Part II: Problems and Policies

Fifty years ago Henry C. Simons challenged the concept of tax exemption when he remarked (1938):

The exemption of the interest payments on an enormous amount of government bonds . . . is a flaw of major importance. It opens the way to deliberate avoidance on a grand scale . . . the exemption not only undermines the program of progressive personal taxation but also introduces a large measure of differentiation in favor of those whose role in our economy is merely that of rentiers.

While the “program of progressive personal taxation” appears to have been left behind, Simons’ criticism of the exemption is still widely held. The purpose of this article is to identify the problems posed by tax exemption, and to assess some alternatives. The analysis goes well beyond the issue of equity, which is the heart of Simons’ complaint. This study asks whether the results of tax exemption represent an appropriate outcome, and questions whether tax exemption is really necessary to achieve the benefits stated in its favor.

This article is a companion to an earlier one (Fortune 1991) that examined the effects of the income tax code on the market for municipal bonds, concluding that the municipal bond market is a creature of tax policy. That article explored the history of the exemption, reviewed the relevant tax legislation, presented a theoretical model of the municipal bond market, and employed econometric methods to determine the role played by the tax code.

The first section of this article addresses three major problems of municipal bond market performance: market instability, vertical equity, and financial efficiency. These problems have driven the debate about reform of the market. The second section discusses several approaches to mitigating these problems. The third section focuses on an aspect of tax exemption that has received very little attention, its impact on

Peter Fortune

Professor of Economics at Tufts University. The author did the basic research for this article while he was a Visiting Scholar at the Federal Reserve Bank of Boston in 1990–91. He is grateful to the Federal Reserve Bank of Boston for research support.

resource allocation and economic efficiency. The section estimates the loss in economic output due to the exemption, concluding that while the loss is small relative to the size of the economy, it is, nevertheless, worthy of attention. The last section summarizes the article and its conclusions.

I. Municipal Bond Market Performance

Why does Congress allow municipal interest payments to be exempted from federal income taxes in the face of a very large chronic deficit in the federal budget, even though it is now clear that no constitutional provision requires that this tax policy continue? The rhetoric of tax exemption is philosophical, appealing to notions of appropriate intergovernmental relations and, in particular, to the doctrine of reciprocal immunity: no level of government should use its taxing authority to impose harm on another level. But the true force behind tax exemption is that it provides state and local governments with a valuable subsidy, which can be enjoyed at their discretion. Political support for the exemption is very strong, and it will continue unless a better way can be found to structure a subsidy to state and local governments.

An assessment of the economics of tax exemption, which is a subsidy of capital costs, suggests that the case for it is weak. The economic argument must

The true force behind tax exemption is that it provides state and local governments with a valuable subsidy, which can be enjoyed at their discretion.

rest on the view that, in the absence of a capital cost subsidy, state and local governments will produce an inadequate amount of public services with insufficient capital intensity. While the final word on this issue is not yet spoken, the debate continues in the current discussion about public infrastructure, such as highways, schools, and solid waste facilities. For

example, Munnell (1990) finds a high marginal productivity of infrastructure, suggesting that an inadequate amount is available, while Hulten and Schwab (1991) find no indication of inadequate infrastructure.

However, even if infrastructure is insufficient, it can be argued that better methods than tax exemption can be used to achieve these goals. Three fundamental criticisms of tax exemption have received the most attention. The first says that tax exemption induces unnecessary volatility into municipal bond yields. According to this "market instability" argument, tax exemption narrows the market for municipal bonds and makes that market more sensitive to changes in the distribution of investable funds between individuals and financial institutions, as well as to other factors that affect financial markets. The result is that municipal bond yields are more volatile than yields on comparable taxable bonds, introducing cyclical variations in the cost of capital for state and local governments. This also introduces variability into the value of the capital-cost subsidy enjoyed by municipalities.

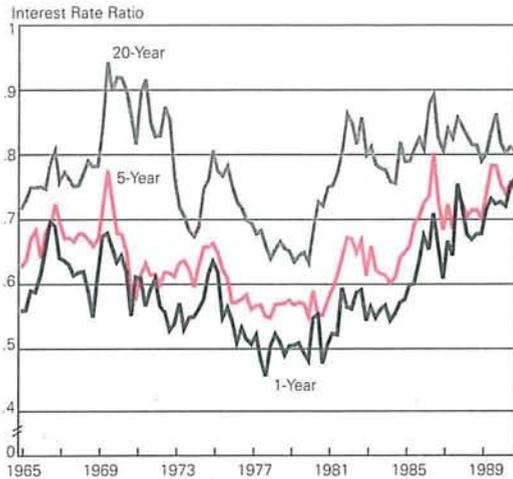
The second criticism, echoing Simons' complaint, is that tax exemption is inequitable; it confers upon the wealthy a valuable opportunity to increase their after-tax income, and it erodes the degree of vertical equity in our tax system by allowing the wealthy to avoid taxation in ways not available to the less affluent. This criticism is the most common in popular discussions of tax exemption.

The third criticism is that tax exemption is *financially* inefficient because it imposes greater costs on federal taxpayers than the benefits it confers upon state and local governments.¹ Still another criticism is that tax exemption fails to encourage *economic* efficiency. Instead, it is argued, tax exemption encourages overproduction of public services as well as overuse of capital by the public sector. A corollary is that the private sector has inadequate capital with which to produce goods and services. This view is based on the assumption that a competitive market economy, unfettered by government intervention in prices, will induce an appropriate allocation of resources. This issue will be discussed in the third section of the article.

¹ Note that the word "efficiency" in this context is used quite differently from the engineering context (getting the most for any given amount of inputs) or the economic context (Pareto-Optimality, or making each person as well off as possible given the positions of all other people). The focus of *financial* efficiency is on the very narrow question of how much benefit is received by lower levels of government per dollar of cost to the federal Treasury.

Figure 1

*Interest Rate Ratios for Selected Terms,
Prime Municipal Bonds vs.
U.S. Treasury Bonds*



Source: Salomon Brothers, Inc.

Market Instability

Figure 1 shows the interest rate ratio for municipal bonds of one-year, five-year, and 20-year maturities. For each maturity, this ratio is the yield to maturity on high-quality municipal bonds (R_m) (Salomon Brothers prime grade) over the yield on U.S. Treasury bonds (R_T) of the same maturity. Much of the movement in these interest rate ratios can be explained by changes in the income tax code (Fortune 1991).

It is clear that the interest rate ratio is highly variable for each maturity. From high ratios in the early 1970s, the ratios declined sharply until the early 1980s, after which they rose again. Thus, municipal bond yields are more volatile than are yields on U.S. Treasury bonds. It is interesting to note, however, that much of this volatility disappeared in the last half of the 1980s. The reduction in volatility in the 1980s was largely the result of the reduced progressivity of the tax system, as well as of tax policies that reduced commercial bank incentives to hold municipal bonds (Fortune 1991).

The interest rate ratio can be interpreted as determined by the tax rate of the marginal investor in

tax-exempt bonds; indeed, this implicit tax rate can be inferred from interest rate data as $t_m = 1 - (R_m/R_T)$, or equal to one minus the interest rate ratio for municipal bonds. The implicit tax rate (t_m) is also the rate of subsidy of state and local capital costs as a result of tax exemption. For example, if the marginal investor's tax rate is 30 percent, then state and local governments face a cost of capital that is only 70 percent of the cost associated with issuing taxable bonds. Thus, the variation in the interest rate ratio (R_m/R_T) translates into variation in the rate of subsidy.

Financial Efficiency and Equity

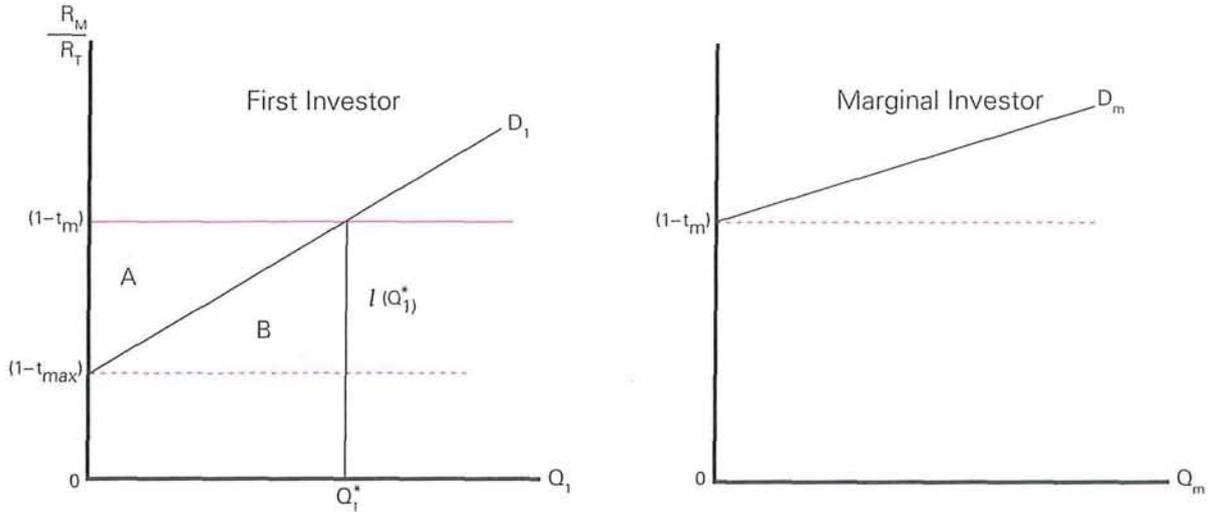
In order to assess the financial efficiency and equity problems, this study will use the model of the municipal bond market developed in Part I (Fortune 1991). Assuming that municipal bonds and taxable bonds are substitutes in investors' portfolios, each investor will choose an amount of municipal bonds based on her tax rate and on her assessments of the nonpecuniary advantages or disadvantages of municipal bonds. Among these nonpecuniary factors are differences in call features, tax rate uncertainty, duration, and liquidity. The optimal holding of municipal bonds will be that quantity for which $(R_m/R_T) = \ell + (1 - t)$, where t is her tax rate and ℓ is the "risk premium" required by the investor; the risk premium is the investor's compensation for nonpecuniary characteristics. While the tax rate is exogenous to the investor's decision, the risk premium is endogenous: as an investor contemplates increasing the amount she invests in municipal bonds, she will require a higher interest rate ratio to compensate for the increased risk of municipal bonds.

Assuming that the risk premium is zero for the first dollar of municipal bonds held by an investor, then if an investor holds no municipals, she considers the first dollar of municipals to be equivalent to a dollar of taxable bonds. This means that for infra-marginal investors, the interest rate ratio will exceed the value $(1 - t)$ by the risk premium required to induce them to hold municipal bonds. But for the marginal investor, who holds a small amount of municipal bonds, the interest rate ratio is $(1 - t_m)$, where t_m is the marginal investor's tax rate.

Figure 2 shows the demand functions for municipal bonds of two investors: the "first investor," whose tax rate, t_{max} , is the highest, and the "marginal investor," with tax rate t_m . The quantity of municipal bonds acquired is along the horizontal axis, and the

Figure 2

Individual Investors in the Market for Municipal Bonds



vertical axis shows the interest rate ratio. The broken horizontal lines at $(1 - t_{max})$ and $(1 - t_m)$, respectively, show each investor's demand function for municipal bonds if tax-exempt and taxable bonds are perfect substitutes. The upward-sloping solid lines labeled D_1 and D_m are the actual demand functions, with the vertical distance to the broken line representing the risk premium required to induce the investor to hold each quantity of municipal bonds.

Figure 2 assumes that the bond markets have settled into an equilibrium in which the interest rate ratio is just sufficient to induce a marginal investor with tax rate t_m to buy a small amount of tax-exempt bonds. The equilibrium interest rate ratio is $(1 - t_m)$, which is high enough to induce the first investor to hold Q_1^* in tax-exempt bonds. For each investor, the interest rate ratio has two parts. The first is the ratio required to give tax-exempts the same after-tax return as taxable bonds; for the first investor this is $(1 - t_{max})$. The second part is the risk premium required to induce the first investor to hold the quantity of tax-exempts he chooses. For the first investor the risk component is $\ell(Q_1^*)$, but for the marginal investor the risk component is (by assumption) zero.

Following an unfortunate convention, the term "windfall income" will be used to designate any

income from tax-exempts that is in excess of the income required to break even on an after-tax basis. Thus, for the first investor the amount of windfall income is given by the sum of areas A and B, multiplied by the taxable interest rate, or $(A + B) * R_T$. However, $(\text{area B}) * R_T$ is not really a windfall, for it is the amount of extra income required to induce the investor to hold Q_1^* . The only true excess income is measured by $(\text{area A}) * R_T$; this is the "investor's surplus," which exists because the investor earns interest on his infra-marginal investment in excess of the amount required. Note that in the case of a linear demand function, the investor's surplus will be 50 percent of the investor's windfall income.

Figure 3 shows the municipal bond market. The vertical line labeled SS is the supply function, showing the quantity of municipal bonds outstanding at each interest rate ratio. In order to focus attention solely on the demand function, it is assumed that this is not interest-elastic.² The upward-sloping schedule

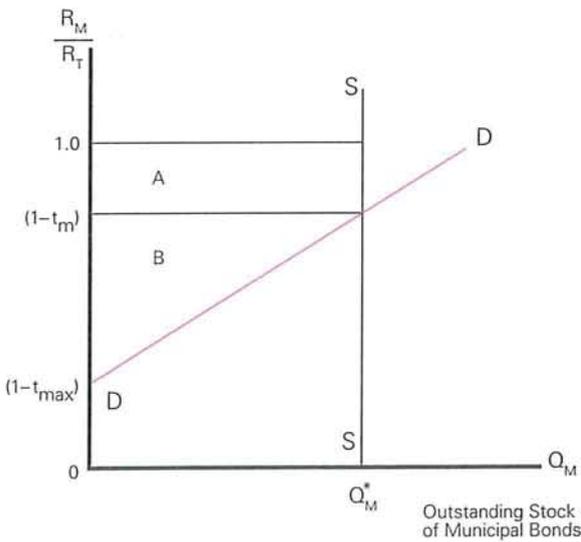
² Considerable evidence suggests that, in the long run, the amount of debt issued to finance capital outlays is not interest-sensitive, though the timing of debt issue is influenced by the interest rate cycle. Recent evidence does suggest, however, that arbitrage activity does induce some interest sensitivity to the supply of municipal bonds (Metcalf 1990, 1991).

DD shows the demand for municipal bonds as a function of the interest rate ratio; this is the horizontal summation of each investor's demand function.

DD rises because, as the amount of bonds outstanding increases, the interest rate ratio must rise by enough to induce infra-marginal investors to switch some portion of their portfolios from taxable to tax-exempt bonds, as well as to induce new marginal investors to enter the market as the original marginal investors become infra-marginal investors. For each quantity of municipal bonds outstanding, the vertical distance to DD is $(1 - t')$, where t' is the tax rate of

Figure 3

Costs and Benefits of Tax Exemption



the investor who buys the last dollar of municipal bonds. Thus, for the quantity actually outstanding (Q_M^*) the tax rate of the marginal investor is t_m and, recalling the assumption that $\ell = 0$ for the marginal investor, the equilibrium interest rate ratio is $R_M/R_T = (1 - t_{max})$. The marginal investor is receiving exactly the interest rate ratio he requires to be induced to hold municipal bonds. But all infra-marginal investors are receiving windfall income, a portion of which is investor's surplus.

Consider the first few dollars of municipal bonds issued. These will be sold to the investors with the highest tax rate (t_{max}); these investors would be

willing to buy municipal bonds if the interest rate ratio were as low as $(1 - t_{max})$, but because Q_M^* of municipal bonds are sold, the interest rate ratio must be $(1 - t_m)$. The windfall income for the highest-bracket investors—per unit of taxable interest paid—is, therefore, $[R_M/R_T - (1 - t_{max})] * Q_M^* = (t_{max} - t_m) * Q_M^*$; the dollar amount of the windfall is this times the taxable interest rate, or $(t_{max} - t_m) * (Q_M^* R_T)$. Note again that this "windfall" is not all unearned: some portion of it (approximately half) is a necessary reward for risk.

If this analysis is extended to compute the total windfall income for investors with higher tax rates than the marginal investor, windfall income is then represented (per unit of R_T) in Figure 3 by area B; the dollar value of the total windfall is $R_T * (\text{area B})$. In practice, one can estimate the total windfall income using the following formula:

$$(1) \text{ Windfall Income} = (\bar{t} - t_m) R_T Q_M$$

$$= (\bar{t} - t_m) [R_M Q_M] / (R_M / R_T).$$

In this formula \bar{t} is the "average marginal tax rate," the average of tax rates paid by all investors in municipal bonds,³ and t_m is the marginal investor's tax rate, calculated from the observed interest rate ratio as $t_m = 1 - (R_M/R_T)$. Windfall income is the difference $(\bar{t} - t_m)$ multiplied by total interest paid on municipal bonds, $R_M Q_M$, and divided by the interest rate ratio; in Figure 3 this amount is shown as (area B) * R_T .

The equity problem is inextricably connected to the financial efficiency problem. In order to assess the degree of financial efficiency, the federal tax revenues lost because of tax exemption must be calculated and compared with the interest payments saved by state and local governments. Consider first the interest savings experienced by states and municipalities. In the absence of tax exemption, municipalities would pay an interest rate ratio of 1.0, but because of tax exemption they pay a rate ratio of $(1 - t_m)$, thereby reducing the rate ratio by $[1 - (1 - t_m)] = t_m$.⁴ Interest

³ The average marginal tax rate would be the sum of each investor's marginal tax rate weighted by the proportion of total municipal bonds outstanding that he holds, or $\bar{t} = \sum t_i s_i$ where i is an index over investors, s_i is the share of municipal bonds owned by the i th investor, and t_i is the i th investor's tax rate.

⁴ For expository convenience, it is assumed that $\ell = 0$ if tax exemption is not allowed; that is, that all nonpecuniary factors that lead to different pricing of municipal and private bonds are due to the exemption. This is clearly not true, and as a result this analysis tends to understate the interest savings of state and local governments.

taking place in conjunction with European integration. Most European insurance markets have historically been national markets separated by restrictive regulation and other obstacles to entry. An exception is reinsurance, for which an international market exists. Insurance markets have been most highly developed in the United Kingdom, the Netherlands, Japan, and the United States, somewhat less so in France and Germany, and much less developed in the southern European Community (EC) nations. Kollias estimates that rates of return on investments by insurance firms have been highest in the United Kingdom because of U.K. companies' relative freedom to invest in equities. Some measures indicate that companies in the United States and Japan are less efficient than companies in some of the EC countries.

Nonlife companies in most EC countries have been losing money on underwriting but have continued to show profits as a result of sharp increases in asset values. Life companies in Europe have generally been profitable, but Kollias did point out that the five big composite (multi-line) companies in the United Kingdom lost more than \$1 billion in 1990. These companies have, nonetheless, been involved in less damaging competition than their counterparts in the United States.

The separation of European insurance markets began to erode in 1988, and since then a series of changes have been underway. Kollias discusses the principal EC agreements, the Single European Act of 1987 which included a program of financial integration, and proposals for harmonization of supervision of investment services. Integration of insurance activities has followed two separate paths, with nonlife large commercial risk and individual life policies being sold abroad under home country control, but "mass risk" life and nonlife insurance being sold under host country regulations. More recent proposals are expected to permit the free supply of insurance under home country rules.

The lowering of international barriers and deregulation are rapidly producing a much more competitive environment for insurance activities in Europe. Important structural changes are also taking place through mergers, joint ventures, cross-sector subsidiaries, bank/insurance conglomerates, and network distribution alliances.

In most European countries banks have not been able to underwrite insurance, and life and nonlife companies have been segregated. This separation is likely to be ended soon. Banks have been allowed to

distribute insurance products, although insurance companies have generally not been allowed to distribute non-insurance products.

The European integration of banking and insurance in the form of mergers, establishment of subsidiaries, and cross-participation contrasts with the strict limitations on such operations in the United States and the prohibitions in Japan. EC draft directives call for the close cooperation of insurance and bank regulators if a bank or holding company controls an insurance company, however.

Henry Parker points out that the insurance market in the United States, while still the world's largest, is slipping rapidly in its share of world premium volume. He criticizes the domestic industry because so few companies participate aggressively in the expanding overseas markets. While substantial impediments to entry exist in some national markets, it can be done and it is getting easier as a result of federal efforts toward freer international trade.

Parker sees 1995 as the earliest date for real insurance market uniformity in the EC. He anticipates some very substantial reductions in insurance prices in several countries, citing Italy, France, and Luxembourg as examples of the wide variations in premiums for identical exposures. He also sees advantages in terms of expense reduction, product

Kollias notes that the lowering of international barriers and deregulation are rapidly producing a much more competitive environment for insurance activities in Europe.

innovation, and achievement of critical mass. Distribution systems will be altered, with more insurance sold through branches of affiliated banks and other financial service providers. An important stumbling block to rapid completion of the insurance directive is agreement on uniform accounting practices.

One concern for U.S. companies expanding into Europe is the possible reemergence of protectionism, particularly if transition problems severely damage long-protected European companies. There is some

risk that a reciprocity standard might replace national treatment, to the detriment of U.S. companies.

Parker notes the importance and potential of the insurance market along the Pacific rim. He also calls attention to the acquisitions of U.S. insurance companies by foreign insurers.

Steven Skalicky reviews insurance market structure in Asia, Latin America, and Eastern Europe to complement Kollias's analysis, which focused primarily on the EC. He makes it clear that barriers that preserve fragmented national markets are under attack around the world.

Asia has the potential to be the fastest-growing market in the 1990s. Japan, the dominant market in Asia, is characterized by a relatively few large companies, including most of the top 10 insurance companies in the world. Japanese companies have been strictly supervised and limited as to their range of investments. Proposals would liberalize the asset restrictions, and greater flexibility in premium rates was permitted recently.

While the Japanese market is technically open to foreign competition, entry has been difficult. Japanese insurers have not been aggressive in overseas operations, but have the potential for being so. The attraction of Asian countries is not current premium volume, but the potential for growth as they become more industrialized.

In Latin America, Skalicky is most optimistic about Mexico, where the insurance industry is growing rapidly and restrictions on outside ownership have been liberalized. The transition from state control in Eastern Europe eventually will also provide opportunities, as reforms permit foreign participation and ownership and economic changes produce growth.

Skalicky sees unprecedented challenges to the insurance companies, consumers, and regulators. Large insurers that have the capital and resources to penetrate rapidly growing insurance markets may, if successful, survive the global consolidation of the industry. Consumers should benefit from less expensive insurance, but will face increasing risks of insurer insolvency. Insurers' reliance on growth in the value of real estate and securities to offset underwriting losses eventually leads to problems. The challenge to insurance regulators to anticipate and deal with problems in foreign markets is formidable.

Public Policy and Life Insurance

Gerard Brannon proposes a framework for evaluating regulatory and tax policies in the life insurance

market. He begins by distinguishing between the risk coverage and the savings elements in the products of life companies, noting the significant tax benefits of the savings component. He presents historical data to show that since 1955, life company reserves have shifted from life insurance to pension and annuity products and life insurance reserves have declined as a percentage of household financial assets. Life insurance in force as a percentage of personal income has increased, however, as consumers shifted from whole life policies, which have a large savings element and require greater reserves, to term insurance. Despite this trend, evidence suggests that consumers still buy too little life insurance.

State regulation of life companies requires the maintenance of adequate reserves and limits the investment risk that can be assumed. In the late 1980s, the historic redundancy in reserves appears to have eroded and investment restrictions failed to protect policyholders from the risk of new financial

Brannon would support a guarantee of the ability of insurance companies to fulfill term life insurance contracts, but would not support the protection of savings.

innovations or the danger of disintermediation. The recent development of variable and universal life policies has been accompanied by higher-risk investments, but also the opportunity for the investors to make risk choices.

State regulators provide limited solvency guarantees for policyholders, funded by levies on competing companies. In some states insurance companies may apply such levies as credits against premium taxes, effectively transferring losses from the industry to the states. Brannon notes the relatively small volume of guaranty fund assessments in the period from 1975 to 1989 and expresses the view that solvency problems currently facing life insurers are clearly not in the same league as the solvency problems of banks and thrifts.

Brannon points out that the Pension Benefit Guaranty Corporation (PBGC) and state guaranty

funds are competitors. When a company purchases an irrevocable contract for an annuity to cover pension liabilities, the guarantee shifts from the PBGC to a state fund. This may work to the benefit of the employer but to the detriment of workers, who have no say in the choice of an insurer. Nonetheless, Brannon argues against federal support of such annuity obligations, using the First Executive case to illustrate his point.

If it is in the public interest to encourage life insurance purchases for the protection of dependents of breadwinners, Brannon would support a guarantee of the ability of insurance companies to fulfill term life insurance contracts, and he would expect such a guaranty program to be successful. However, he would not support the protection of savers and he deplores the current tax advantages that encourage the intermingling of insurance and investment features, complicating the development of an appropriate guaranty scheme for insurance.

Joseph Belth confines his discussion to the issue of federal income taxation of the inside interest in cash-value life insurance and life annuities. Individuals tend to postpone the distressing subject of life insurance, and therefore a major expense for insurance companies is the commission paid to agents to perform the "anti-procrastination" function. Because natural premiums for life insurance are very low for young purchasers, companies do not receive sufficient revenue to compensate agents. Furthermore, the very high premiums in later years tend to produce adverse selection as healthier members drop insurance. Both of these problems can be mitigated by level-premium, cash-value insurance, which creates a savings component. The federal income tax on the inside interest is generally deferred. Life annuities, which provide regular payments over an individual's lifetime, make sense only in periods of low interest rates, because one can obtain almost as high a return investing principal directly during high-rate periods without destroying the principal, as happens with an annuity. A life annuity may have a lengthy accumulation period before the beginning of the liquidation period, and here again federal income taxation on inside interest is generally deferred.

A theoretical argument can be made that deferred tax treatment of inside interest in these two situations can no longer be justified. Cash-value life insurance is of increasing benefit to high-income individuals, and life annuities are increasingly used solely because of tax considerations. Nevertheless, Belth argues that current taxation of the inside inter-

est would have a "devastating impact on the life insurance industry and would threaten its very survival." He also believes the industry has sufficient political clout to discourage any legislative attempt to impose current taxation.

Earl Pomeroy brings a regulator's perspective to the issues raised by Brannon. He contends that the sophistication of regulatory oversight has been improved in response to the lower capitalization levels, slimmer profit margins, and higher risks found in the life insurance industry today. Pomeroy cites the improved system for bond evaluation, a model law covering bond concentrations, limits on junk bonds, and progress toward reserve requirements and limitations on other higher-risk investments. While such regulatory activity has the necessary effect of lowering investment returns and restricting capital flows to particular activities, it is wholly appropriate because solvency protection is the regulator's first priority.

Pomeroy discusses such consumer protection regulations as required disclosures of product characteristics and minimum product quality standards. He chides Congress for attempting to achieve social goals through the imposition of costly market restrictions.

With respect to guaranty funds, Pomeroy agrees with Brannon that they can dull consumer sensitivity to insurer risk exposure, but finds that they serve a critical role. Despite assessment limitations, Pomeroy is reasonably hopeful that the guaranty fund mechanism has sufficient capacity, on a state-by-state basis, to handle a major life insurance failure.

After briefly reviewing the history of state insurance regulation, including recent activities of the National Association of Insurance Commissioners (NAIC), Pomeroy lists several concerns state regulators have with federal regulation of insurance. He maintains that federal officials tend to overstate the solvency problem, because of their sensitivity to the thrift failures and because they view the Executive Life case as a harbinger of trouble for the life industry generally. Newly implemented state reforms should be given time to work. Pomeroy argues that political pressures could lead to a situation where federal solvency regulation is imposed alongside state regulation of rates with the two sets of regulators pursuing conflicting objectives. Pomeroy does not expect a specific federal regulatory proposal to have much political appeal, even though the general concept might.

Warren Wise challenges Brannon's characterization of the cash value in permanent life insurance as

being equivalent to a savings account. He argues that it arises from the leveling of premiums and is an integral part of providing lifetime protection at an acceptable price. The tax-free inside buildup is a subsidy to encourage life insurance protection, not savings.

Wise acknowledges that the industry is more vulnerable to failure than it once was, although his proposals for dealing with the problem are at odds with Brannon's. Rather than limit protection to death benefits, as Brannon would do, Wise would cover all policyholders. However, he would want all interested parties to share in losses when an insurer fails, including insurance sales representatives, policyholders, and state governments.

Guaranty fund assessments should be risk-based and collected on a regular basis so that the heaviest impact will fall on those insurers most likely to fail. Sales representatives should have an incentive to recommend safe companies, and states should have an incentive to devote adequate resources to solvency regulation. State contributions could be in the form of the tax offset for guaranty fund assessments that already exists in several states. Insurance consumers should share the burden by recovering less than the full amount due them.

Wise would improve regulation by linking capital requirements to risk, strengthening investment restrictions, improving accounting practices, and better controlling reinsurance transactions. Regulators must be provided sufficient resources to carry out their responsibilities.

The question remains of who should administer solvency regulation, and Wise would prefer that it be done without federal involvement if the states can adopt and enforce strong, uniform solvency standards. However, if a federal role proves to be necessary, he would prefer that federal involvement be limited to the setting of minimal standards, oversight, and the ensuring of compliance.

Public Policy and Property-Liability Insurance

Scott Harrington makes some very specific recommendations as to what changes should, and should not, be made to property-liability insurance regulation. He would like to reduce guaranty fund coverage in order to increase market discipline. He does not think a case has been made for a federal regulatory role, and believes that federal supervision could actually increase total insolvency costs. Harrington would like to see the abandonment of state

rate-setting, but would not alter the industry's anti-trust exemption.

With respect to guaranty funds, Harrington argues that guarantees result in policyholders having reduced incentives to buy coverage from safe insurers; the market collectively has more information and knowledge than the regulators, and the spreading of insolvency losses through guaranty funds can reduce pressure on government to commit adequate resources to solvency monitoring. It would be desirable to require a large co-payment from the policyholders, especially those who are best able to monitor insolvency. Harrington also makes a case for post-insolvency assessments being superior to an accumulated

Harrington argues that rate regulation of property-liability insurance has little or no justification, and would limit the regulatory role to requiring appropriate information disclosure.

fund. The arguments presented against federal regulation of property-liability insurers draw heavily on the thrift experience, and particularly the role of Congress in condoning forbearance for insolvent institutions.

Harrington argues that rate regulation of property-liability insurance has little or no justification, and he would limit the regulatory role to requiring appropriate information disclosure. The industry is highly competitive, with ease of entry, and market forces can most efficiently determine rates. Harrington contrasts the industry to public utilities, where rate regulation is necessary. Rate regulation can result in insurers exiting certain lines or states, reducing net worth and thereby increasing insolvency risk; it can also result in insurers being less innovative. Regulation can directly increase expenses and distract management as a result of the rate hearing process.

Harrington sees the cooperative development of policy forms and sharing of loss data as entirely constructive, lowering costs, easing entry, and increasing forecast accuracy. He sees the forecasting of

future losses by advisory organizations as serving a useful function to the extent that they improve individual insurer forecasts. He is concerned that a substantial change in the industry's antitrust exemption could lead to higher prices and less stability, and result in a surge of costly litigation.

J. Robert Hunter vigorously challenges Harrington's characterization of the property-liability insurance market as highly competitive, as well as his proposal to remove rate regulation while preserving the industry's exemption from antitrust laws. Hunter presents evidence that the public does not have sufficient information to select insurance companies on the basis of cost or service quality. He also cites findings that collusion on rates has been the norm, not the exception, in the industry. Hunter reviews the mechanism by which the Insurance Services Office, an industry service organization, provides insurers with advisory rates. He argues that, even with plans to exclude expense factors from the rate data, some critical components of the rate formula will still be provided that instead should be calculated independently by individual insurers, if collusion is to be prevented.

Hunter could agree to easing or even phasing out rate regulation, but only if all anticompetitive forces were eliminated. Specifically, he mentions the antitrust exemption, the anti-rebate laws, the anti-group laws, the barriers to entry by banks, the information gap, and the underwriting selection problem.

With respect to solvency, Hunter challenges Harrington's proposal to decrease guaranty fund coverage in order to improve market discipline. He would expand coverage for personal lines and small businesses. Even with respect to large commercial customers, he notes that loss of insurance protection could have secondary effects on the public when the business, as well as the insurance company, fails. Hunter calls for federal minimum standards for solvency regulation, and direct federal regulation of alien reinsurance and alien surplus lines markets.

Robert Litan agrees with most of Harrington's points, but he would not reject a federal solvency role and would draw different lessons from the thrift crisis. Litan faults the state regulators for their performance in connection with the larger failures of property-liability insurance companies in recent years. He attributes recent efforts by the NAIC to improve state regulation to the threat of federal regulation. Litan proposes creating a federal regulatory program and a national guaranty fund system as an alternative to state regulation and guaranty funds.

Insurers that chose the federal system would no longer be subject to rate regulation. While Litan acknowledges some adverse selection problems with his proposal, he sees it as a way of forcing reform of the state systems, or having property-liability insurance regulation gravitate to the federal level.

Litan draws on his interpretation of the thrift crisis to support the idea that a pre-funded guaranty system would be superior to the usual post-insolvency assessment procedure. He points out that thrift regulators engaged in forbearance largely because of insufficient funds to resolve failed institutions.

Litan is concerned that major exogenous events pose a substantial threat to the industry, citing specifically a potential major earthquake and possible court rulings making insurance companies responsible for the cost of cleaning up hazardous waste sites. He suggests steps that could be taken in advance to protect the industry from being overwhelmed by such calamities.

Richard Stewart briefly outlines what he sees as the major issues in rate regulation and in dealing with the underwriting cycle. He then turns to the issue of solvency and argues that insolvency is a natural outcome for a property-liability insurer.

It is the liabilities of the insurer, not the assets, that are of most concern, and these liabilities extend far into the future. In Stewart's view, the future is not going to be like the past, and therefore it is nearly impossible to estimate the extent of these liabilities for pricing or reserving purposes. In the general liability line the threats are systemic, further adding to the industry's susceptibility to catastrophes on the liability side. Moreover, the industry is intensely competitive, and the incentives and rewards are concentrated on the front end of a transaction, with willingness and ability to pay claims coming much later.

If it is the duty of the regulator to prevent insolvencies, it is very hard to accomplish this by early detection and swift action because of the uncertainty about the extent of the liabilities. However, it is easy to forbear and avoid recognition of insolvency for several years, thereby escaping responsibility. In Stewart's view, this perverse incentive for the regulator increases the risk of even greater losses.

Our system of compensation for accidents functions through an insured civil liability procedure. In the event of insurance company insolvency, the victims include not only direct policyholders but large groups of individuals, whose only link may be the use of a common product or exposure to a form of

pollution, and who are terribly hurt by the insurance company insolvency. We should not think only of corporate America in considering guaranty fund protection surrounding the property-liability insurance system.

Stewart believes that state regulation, with improvements such as those currently in process, can do a satisfactory job of detecting and acting against emerging insolvencies. However, liquidation and guarantees for large-scale general liability insolvencies should be managed at the national level.

Conclusions

The ability of domestic insurance companies to meet their obligations is vital not only to the welfare of their customers but also to the economy and social fabric of the country. In recent years the structure of the life insurance industry has changed in a way that has increased the risk of major insurers becoming insolvent or illiquid. Capital ratios have not increased in response. At the same time the property-liability

insurance industry has become more leveraged and perhaps more vulnerable to large-scale losses.

Opinions differ widely as to the extent and duration of the current weaknesses in the asset quality of life insurers, but it is generally agreed that state regulation and the system of guaranty funds are being materially strengthened by various initiatives. Experts disagree, however, about the ability of even strengthened state systems to avert solvency problems or to safeguard policyholders and others in the event of failures of major insurers. Agreement on the desirability and extent of protection to be provided for policyholders, pensioners, and savers dependent on an insurance company's ability to pay, would facilitate determination of what, if any, federal role is desirable in regulation or in administering guaranty funds.

Congressional interest in examining the insurance industry, continuing downgrades in ratings of individual companies, and the prospects for a prolonged period of depressed commercial real estate values, all suggest that insurance industry solvency issues will be with us for some time.

Conference Series

The Federal Reserve Bank of Boston Conference Series

No. 1	Controlling Monetary Aggregates	June 1969
No. 2	The International Adjustment Mechanism	October 1969
No. 3	Financing State and Local Governments in the Seventies	June 1970
No. 4	Housing and Monetary Policy (out of print)	October 1970
No. 5	Consumer Spending and Monetary Policy: The Linkages	June 1971
No. 6	Canadian-United States Financial Relationships	September 1971
No. 7	Financing Public Schools (out of print)	January 1972
No. 8	Policies for a More Competitive Financial System	June 1972
No. 9	Controlling Monetary Aggregates II: The Implementation	September 1972
No. 10	Issues in Federal Debt Management	June 1973
No. 11	Credit Allocation Techniques and Monetary Policy	September 1973
No. 12	International Aspects of Stabilization Policies	June 1974
No. 13	The Economics of a National Electronic Funds Transfer System	October 1974
No. 14	New Mortgage Designs for Stable Housing in an Inflationary Environment	January 1975
No. 15	New England and the Energy Crisis (out of print)	October 1975
No. 16	Funding Pensions: Issues and Implications for Financial Markets	October 1976
No. 17	Minority Business Development	November 1976
No. 18	Key Issues in International Banking	October 1977
No. 19	After the Phillips Curve: Persistence of High Inflation and High Unemployment	June 1978
No. 20	Managed Exchange-Rate Flexibility: The Recent Experience	October 1978
No. 21	The Regulation of Financial Institutions	October 1979
No. 22	The Decline in Productivity Growth	June 1980
No. 23	Controlling Monetary Aggregates III	October 1980
No. 24	The Future of the Thrift Industry	October 1981
No. 25	Saving and Government Policy	October 1982
No. 26	The Political Economy of Monetary Policy: National and International Aspects	July 1983
No. 27	The Economics of Large Government Deficits	October 1983
No. 28	The International Monetary System: Forty Years After Bretton Woods	May 1984
No. 29	Economic Consequences of Tax Simplification	October 1985
No. 30	Lessons from the Income Maintenance Experiments	September 1986
No. 31	The Merger Boom	October 1987
No. 32	International Payments Imbalances in the 1980s	October 1988
No. 33	Are the Distinctions between Debt and Equity Disappearing?	October 1989
No. 34	Is There a Shortfall in Public Capital Investment?	June 1990
No. 35	The Financial Condition and Regulation of Insurance Companies	June 1991

Conference Series



The Financial Condition and Regulation of Insurance Companies, proceedings of a conference sponsored by the Federal Reserve Bank of Boston, is now available. Each year since 1969 the Bank has gathered a distinguished group from universities, government, business, and finance to discuss important issues affecting economic policy. The Bank's June 1991 economic conference examined recent trends in the insurance industry, considering both life and property-liability companies. A group of academics, economists, government regulators, and industry officials met to consider the competitive pressures, the dramatic changes in risk factors, and the decline in asset quality currently facing the industry. The proceedings include all the conference papers and discussions; the volume begins with an overview of the conference, prepared by the editors. Among the authors and discussants are Joseph M. Belth, Gerard M. Brannon, Jeffrey Cohen, J. David Cummins, Scott E. Harrington, J. Robert Hunter, Terence Lennon, Robert E. Litan, Earl R. Pomeroy, Richard E. Stewart, James M. Stone, Frederick S. Townsend, Jr., and Kenneth M. Wright.

Copies of conference volume no. 35, *The Financial Condition and Regulation of Insurance Companies*, may be obtained without charge on request to the Research Library—D, Federal Reserve Bank of Boston, P.O. Box 2076, Boston, Massachusetts 02106-2076.

Please type or print

Name _____ Title _____

Organization _____

Address _____

Please check

- Please send Conference Series No. 35 *The Financial Condition and Regulation of Insurance Companies*.
- Please send back-issue Conference Series No. _____

Note: Single volumes will be mailed without charge, but a \$10.00 payment (check drawn on a branch of a U.S. bank) will be required for more than 10 volumes or more than 10 copies of the same volume.

PLACE
STAMP
HERE

Research Library—D
Federal Reserve Bank of Boston
P.O. Box 2076
Boston, Massachusetts 02106-2076

The Municipal Bond Market, Part II: Problems and Policies

Fifty years ago Henry C. Simons challenged the concept of tax exemption when he remarked (1938):

The exemption of the interest payments on an enormous amount of government bonds . . . is a flaw of major importance. It opens the way to deliberate avoidance on a grand scale . . . the exemption not only undermines the program of progressive personal taxation but also introduces a large measure of differentiation in favor of those whose role in our economy is merely that of rentiers.

While the “program of progressive personal taxation” appears to have been left behind, Simons’ criticism of the exemption is still widely held. The purpose of this article is to identify the problems posed by tax exemption, and to assess some alternatives. The analysis goes well beyond the issue of equity, which is the heart of Simons’ complaint. This study asks whether the results of tax exemption represent an appropriate outcome, and questions whether tax exemption is really necessary to achieve the benefits stated in its favor.

This article is a companion to an earlier one (Fortune 1991) that examined the effects of the income tax code on the market for municipal bonds, concluding that the municipal bond market is a creature of tax policy. That article explored the history of the exemption, reviewed the relevant tax legislation, presented a theoretical model of the municipal bond market, and employed econometric methods to determine the role played by the tax code.

The first section of this article addresses three major problems of municipal bond market performance: market instability, vertical equity, and financial efficiency. These problems have driven the debate about reform of the market. The second section discusses several approaches to mitigating these problems. The third section focuses on an aspect of tax exemption that has received very little attention, its impact on

Peter Fortune

Professor of Economics at Tufts University. The author did the basic research for this article while he was a Visiting Scholar at the Federal Reserve Bank of Boston in 1990–91. He is grateful to the Federal Reserve Bank of Boston for research support.

resource allocation and economic efficiency. The section estimates the loss in economic output due to the exemption, concluding that while the loss is small relative to the size of the economy, it is, nevertheless, worthy of attention. The last section summarizes the article and its conclusions.

I. Municipal Bond Market Performance

Why does Congress allow municipal interest payments to be exempted from federal income taxes in the face of a very large chronic deficit in the federal budget, even though it is now clear that no constitutional provision requires that this tax policy continue? The rhetoric of tax exemption is philosophical, appealing to notions of appropriate intergovernmental relations and, in particular, to the doctrine of reciprocal immunity: no level of government should use its taxing authority to impose harm on another level. But the true force behind tax exemption is that it provides state and local governments with a valuable subsidy, which can be enjoyed at their discretion. Political support for the exemption is very strong, and it will continue unless a better way can be found to structure a subsidy to state and local governments.

An assessment of the economics of tax exemption, which is a subsidy of capital costs, suggests that the case for it is weak. The economic argument must

The true force behind tax exemption is that it provides state and local governments with a valuable subsidy, which can be enjoyed at their discretion.

rest on the view that, in the absence of a capital cost subsidy, state and local governments will produce an inadequate amount of public services with insufficient capital intensity. While the final word on this issue is not yet spoken, the debate continues in the current discussion about public infrastructure, such as highways, schools, and solid waste facilities. For

example, Munnell (1990) finds a high marginal productivity of infrastructure, suggesting that an inadequate amount is available, while Hulten and Schwab (1991) find no indication of inadequate infrastructure.

However, even if infrastructure is insufficient, it can be argued that better methods than tax exemption can be used to achieve these goals. Three fundamental criticisms of tax exemption have received the most attention. The first says that tax exemption induces unnecessary volatility into municipal bond yields. According to this "market instability" argument, tax exemption narrows the market for municipal bonds and makes that market more sensitive to changes in the distribution of investable funds between individuals and financial institutions, as well as to other factors that affect financial markets. The result is that municipal bond yields are more volatile than yields on comparable taxable bonds, introducing cyclical variations in the cost of capital for state and local governments. This also introduces variability into the value of the capital-cost subsidy enjoyed by municipalities.

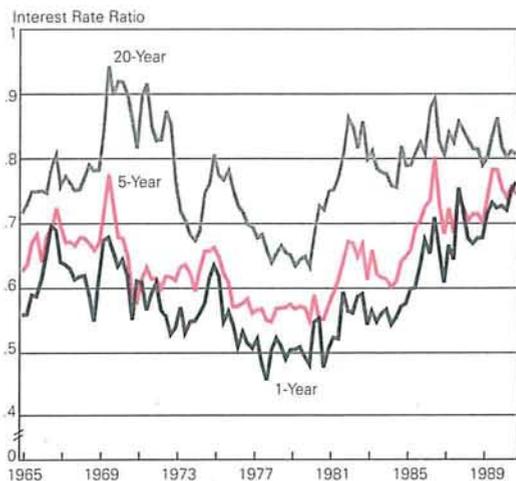
The second criticism, echoing Simons' complaint, is that tax exemption is inequitable; it confers upon the wealthy a valuable opportunity to increase their after-tax income, and it erodes the degree of vertical equity in our tax system by allowing the wealthy to avoid taxation in ways not available to the less affluent. This criticism is the most common in popular discussions of tax exemption.

The third criticism is that tax exemption is *financially* inefficient because it imposes greater costs on federal taxpayers than the benefits it confers upon state and local governments.¹ Still another criticism is that tax exemption fails to encourage *economic* efficiency. Instead, it is argued, tax exemption encourages overproduction of public services as well as overuse of capital by the public sector. A corollary is that the private sector has inadequate capital with which to produce goods and services. This view is based on the assumption that a competitive market economy, unfettered by government intervention in prices, will induce an appropriate allocation of resources. This issue will be discussed in the third section of the article.

¹ Note that the word "efficiency" in this context is used quite differently from the engineering context (getting the most for any given amount of inputs) or the economic context (Pareto-Optimality, or making each person as well off as possible given the positions of all other people). The focus of *financial* efficiency is on the very narrow question of how much benefit is received by lower levels of government per dollar of cost to the federal Treasury.

Figure 1

*Interest Rate Ratios for Selected Terms,
Prime Municipal Bonds vs.
U.S. Treasury Bonds*



Source: Salomon Brothers, Inc.

Market Instability

Figure 1 shows the interest rate ratio for municipal bonds of one-year, five-year, and 20-year maturities. For each maturity, this ratio is the yield to maturity on high-quality municipal bonds (R_m) (Salomon Brothers prime grade) over the yield on U.S. Treasury bonds (R_T) of the same maturity. Much of the movement in these interest rate ratios can be explained by changes in the income tax code (Fortune 1991).

It is clear that the interest rate ratio is highly variable for each maturity. From high ratios in the early 1970s, the ratios declined sharply until the early 1980s, after which they rose again. Thus, municipal bond yields are more volatile than are yields on U.S. Treasury bonds. It is interesting to note, however, that much of this volatility disappeared in the last half of the 1980s. The reduction in volatility in the 1980s was largely the result of the reduced progressivity of the tax system, as well as of tax policies that reduced commercial bank incentives to hold municipal bonds (Fortune 1991).

The interest rate ratio can be interpreted as determined by the tax rate of the marginal investor in

tax-exempt bonds; indeed, this implicit tax rate can be inferred from interest rate data as $t_m = 1 - (R_m/R_T)$, or equal to one minus the interest rate ratio for municipal bonds. The implicit tax rate (t_m) is also the rate of subsidy of state and local capital costs as a result of tax exemption. For example, if the marginal investor's tax rate is 30 percent, then state and local governments face a cost of capital that is only 70 percent of the cost associated with issuing taxable bonds. Thus, the variation in the interest rate ratio (R_m/R_T) translates into variation in the rate of subsidy.

Financial Efficiency and Equity

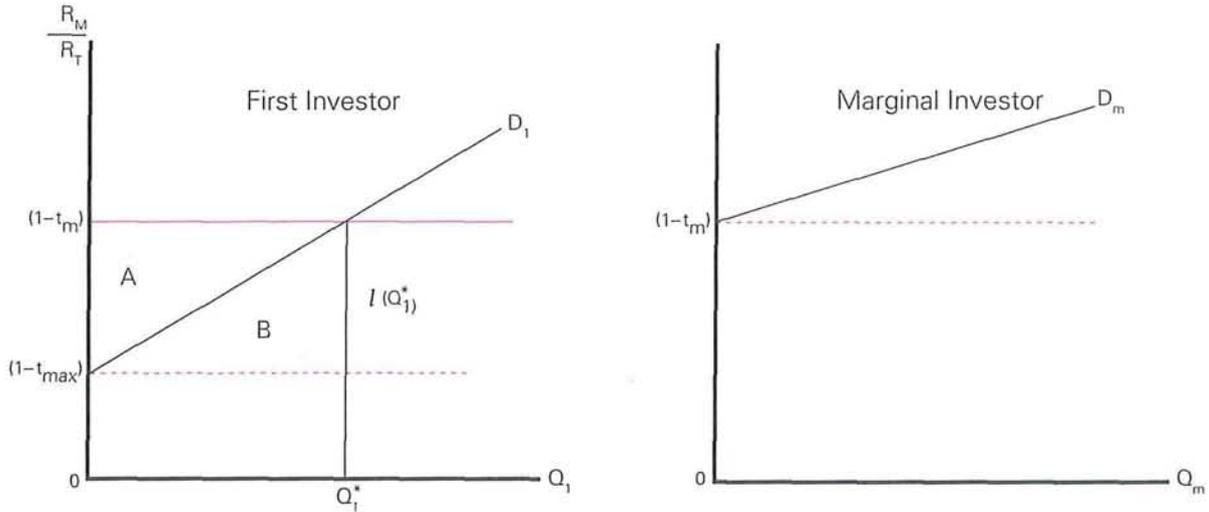
In order to assess the financial efficiency and equity problems, this study will use the model of the municipal bond market developed in Part I (Fortune 1991). Assuming that municipal bonds and taxable bonds are substitutes in investors' portfolios, each investor will choose an amount of municipal bonds based on her tax rate and on her assessments of the nonpecuniary advantages or disadvantages of municipal bonds. Among these nonpecuniary factors are differences in call features, tax rate uncertainty, duration, and liquidity. The optimal holding of municipal bonds will be that quantity for which $(R_m/R_T) = \ell + (1 - t)$, where t is her tax rate and ℓ is the "risk premium" required by the investor; the risk premium is the investor's compensation for nonpecuniary characteristics. While the tax rate is exogenous to the investor's decision, the risk premium is endogenous: as an investor contemplates increasing the amount she invests in municipal bonds, she will require a higher interest rate ratio to compensate for the increased risk of municipal bonds.

Assuming that the risk premium is zero for the first dollar of municipal bonds held by an investor, then if an investor holds no municipals, she considers the first dollar of municipals to be equivalent to a dollar of taxable bonds. This means that for infra-marginal investors, the interest rate ratio will exceed the value $(1 - t)$ by the risk premium required to induce them to hold municipal bonds. But for the marginal investor, who holds a small amount of municipal bonds, the interest rate ratio is $(1 - t_m)$, where t_m is the marginal investor's tax rate.

Figure 2 shows the demand functions for municipal bonds of two investors: the "first investor," whose tax rate, t_{max} , is the highest, and the "marginal investor," with tax rate t_m . The quantity of municipal bonds acquired is along the horizontal axis, and the

Figure 2

Individual Investors in the Market for Municipal Bonds



vertical axis shows the interest rate ratio. The broken horizontal lines at $(1 - t_{max})$ and $(1 - t_m)$, respectively, show each investor's demand function for municipal bonds if tax-exempt and taxable bonds are perfect substitutes. The upward-sloping solid lines labeled D_1 and D_m are the actual demand functions, with the vertical distance to the broken line representing the risk premium required to induce the investor to hold each quantity of municipal bonds.

Figure 2 assumes that the bond markets have settled into an equilibrium in which the interest rate ratio is just sufficient to induce a marginal investor with tax rate t_m to buy a small amount of tax-exempt bonds. The equilibrium interest rate ratio is $(1 - t_m)$, which is high enough to induce the first investor to hold Q_1^* in tax-exempt bonds. For each investor, the interest rate ratio has two parts. The first is the ratio required to give tax-exempts the same after-tax return as taxable bonds; for the first investor this is $(1 - t_{max})$. The second part is the risk premium required to induce the first investor to hold the quantity of tax-exempts he chooses. For the first investor the risk component is $\ell(Q_1^*)$, but for the marginal investor the risk component is (by assumption) zero.

Following an unfortunate convention, the term "windfall income" will be used to designate any

income from tax-exempts that is in excess of the income required to break even on an after-tax basis. Thus, for the first investor the amount of windfall income is given by the sum of areas A and B, multiplied by the taxable interest rate, or $(A + B) * R_T$. However, $(\text{area B}) * R_T$ is not really a windfall, for it is the amount of extra income required to induce the investor to hold Q_1^* . The only true excess income is measured by $(\text{area A}) * R_T$; this is the "investor's surplus," which exists because the investor earns interest on his infra-marginal investment in excess of the amount required. Note that in the case of a linear demand function, the investor's surplus will be 50 percent of the investor's windfall income.

Figure 3 shows the municipal bond market. The vertical line labeled SS is the supply function, showing the quantity of municipal bonds outstanding at each interest rate ratio. In order to focus attention solely on the demand function, it is assumed that this is not interest-elastic.² The upward-sloping schedule

² Considerable evidence suggests that, in the long run, the amount of debt issued to finance capital outlays is not interest-sensitive, though the timing of debt issue is influenced by the interest rate cycle. Recent evidence does suggest, however, that arbitrage activity does induce some interest sensitivity to the supply of municipal bonds (Metcalf 1990, 1991).

DD shows the demand for municipal bonds as a function of the interest rate ratio; this is the horizontal summation of each investor's demand function.

DD rises because, as the amount of bonds outstanding increases, the interest rate ratio must rise by enough to induce infra-marginal investors to switch some portion of their portfolios from taxable to tax-exempt bonds, as well as to induce new marginal investors to enter the market as the original marginal investors become infra-marginal investors. For each quantity of municipal bonds outstanding, the vertical distance to DD is $(1 - t')$, where t' is the tax rate of

willing to buy municipal bonds if the interest rate ratio were as low as $(1 - t_{max})$, but because Q_M^* of municipal bonds are sold, the interest rate ratio must be $(1 - t_m)$. The windfall income for the highest-bracket investors—per unit of taxable interest paid—is, therefore, $[R_M/R_T - (1 - t_{max})] * Q_M^* = (t_{max} - t_m) * Q_M^*$; the dollar amount of the windfall is this times the taxable interest rate, or $(t_{max} - t_m) * (Q_M^* R_T)$. Note again that this "windfall" is not all unearned: some portion of it (approximately half) is a necessary reward for risk.

If this analysis is extended to compute the total windfall income for investors with higher tax rates than the marginal investor, windfall income is then represented (per unit of R_T) in Figure 3 by area B; the dollar value of the total windfall is $R_T * (\text{area B})$. In practice, one can estimate the total windfall income using the following formula:

$$(1) \text{ Windfall Income} = (\bar{t} - t_m) R_T Q_M \\ = (\bar{t} - t_m) [R_M Q_M] / (R_M / R_T).$$

In this formula \bar{t} is the "average marginal tax rate," the average of tax rates paid by all investors in municipal bonds,³ and t_m is the marginal investor's tax rate, calculated from the observed interest rate ratio as $t_m = 1 - (R_M/R_T)$. Windfall income is the difference $(\bar{t} - t_m)$ multiplied by total interest paid on municipal bonds, $R_M Q_M$, and divided by the interest rate ratio; in Figure 3 this amount is shown as (area B) * R_T .

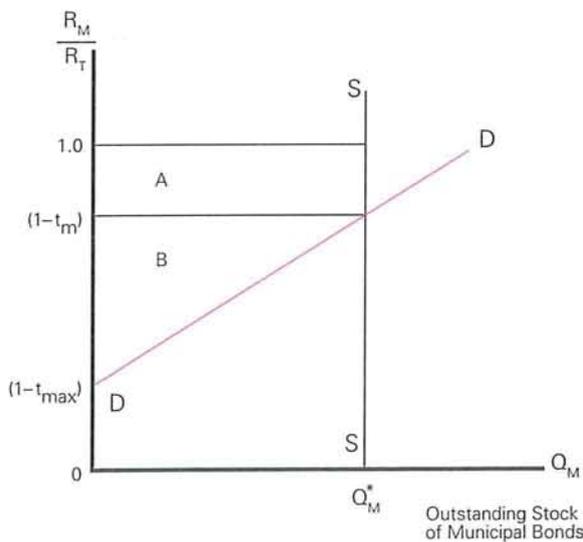
The equity problem is inextricably connected to the financial efficiency problem. In order to assess the degree of financial efficiency, the federal tax revenues lost because of tax exemption must be calculated and compared with the interest payments saved by state and local governments. Consider first the interest savings experienced by states and municipalities. In the absence of tax exemption, municipalities would pay an interest rate ratio of 1.0, but because of tax exemption they pay a rate ratio of $(1 - t_m)$, thereby reducing the rate ratio by $[1 - (1 - t_m)] = t_m$.⁴ Interest

³ The average marginal tax rate would be the sum of each investor's marginal tax rate weighted by the proportion of total municipal bonds outstanding that he holds, or $\bar{t} = \sum t_i s_i$ where i is an index over investors, s_i is the share of municipal bonds owned by the i th investor, and t_i is the i th investor's tax rate.

⁴ For expository convenience, it is assumed that $\ell = 0$ if tax exemption is not allowed; that is, that all nonpecuniary factors that lead to different pricing of municipal and private bonds are due to the exemption. This is clearly not true, and as a result this analysis tends to understate the interest savings of state and local governments.

Figure 3

Costs and Benefits of Tax Exemption



the investor who buys the last dollar of municipal bonds. Thus, for the quantity actually outstanding (Q_M^*) the tax rate of the marginal investor is t_m and, recalling the assumption that $\ell = 0$ for the marginal investor, the equilibrium interest rate ratio is $R_M/R_T = (1 - t_{max})$. The marginal investor is receiving exactly the interest rate ratio he requires to be induced to hold municipal bonds. But all infra-marginal investors are receiving windfall income, a portion of which is investor's surplus.

Consider the first few dollars of municipal bonds issued. These will be sold to the investors with the highest tax rate (t_{max}); these investors would be

savings is, therefore, measured by (area A) * R_T , which is

$$(2) \quad \text{Interest Savings} = t_m R_T Q_M.$$

The revenue cost to the U.S. Treasury is the sum of two components: the windfall income received by high-bracket investors plus the interest savings of municipalities. The dollar value of revenue cost is (area A + area B) * R_T . Thus,

$$(3) \quad \text{Revenue Cost} = R_T * [t_m + (\bar{t} - t_m)] Q_M \\ = \bar{t} R_T Q_M.$$

If, as has historically been true in the United States, taxation of income is progressive, then the *average* marginal tax rate exceeds the marginal tax rate ($\bar{t} > t_m$) and area (A + B) > area A. Therefore the revenue cost to the federal government must exceed the interest savings enjoyed by states and local governments by an amount known as "windfall income."

Thus, the financial inefficiency of tax exemption exists because of the equity problem, and reduction of the equity problem implies progress on the efficiency problem. The degree of financial efficiency can be measured by an "efficiency index," defined as the proportion of the revenue costs that accrues to states and local governments as interest savings. This efficiency index is the ratio of area A to area (A + B), or

$$(4) \quad \text{Efficiency Index} = t_m / \bar{t}.$$

Estimates of the Revenue Costs, Interest Savings, and Efficiency

Several studies have attempted to measure the revenue costs and efficiency of tax exemption. One approach, the Meltzer-Ott method (Ott and Meltzer 1963), is to estimate the marginal tax rates from the interest rate ratio, estimate the average marginal tax rate from data on ownership of municipal bonds and on the tax rates of each sector, and use U.S. Treasury or Federal Reserve Board flow-of-funds data on the outstanding stock of tax-exempt bonds. The second approach, called here the OMB method, is to use the Tax Expenditure Budget, reported annually by the U.S. Office of Management and Budget (1990).

The Meltzer-Ott method is used here to estimate revenue losses and interest savings for 1990. (See the Appendix, *Measuring the Cost of Tax Exemption*.) The year 1990 was chosen for two reasons: it is the

most recent year for which data are available, and it is sufficiently long after the Tax Reform Act of 1986 to allow a new equilibrium in the ownership of municipal bonds to be reached. As discussed in Part I of this study (Fortune 1991), the Tax Reform Act of 1986 created dramatic changes in the municipal bond market. First, the ownership of municipal bonds shifted sharply from financial institutions, particularly commercial banks, to households: while financial institutions and households each held about 50 percent of municipal bonds in 1985, the household share of outstanding tax-exempts rose to about 65 percent by the end of 1990. Second, the corporate income tax rate declined dramatically, from 46 percent to 34 percent, as did the maximum personal income tax rate, from 50 percent to 33 percent. Both acted to increase the interest rate ratio.

Poterba and Feenberg (1991) estimate that in 1988, after the Tax Reform Act was fully implemented, the average marginal income tax rate for households was 28 percent. For financial institutions, which held about 35 percent of outstanding municipals, the tax rate was 34 percent. The weighted average of those tax rates is 30.1 percent; this will be used to derive estimates of the average marginal tax rate for 1990.

The marginal tax rate for 1990 is assumed to be 23 percent, based on 1985-90 average interest rates of 8.77 percent for 10-year Treasury bonds and 6.78 percent for 10-year prime municipal bonds. At year end 1990, the outstanding stock of municipal bonds was \$837 billion. Combined with the previous assumptions, the Meltzer-Ott estimate of 1990 interest savings for state and local governments is \$16.9 billion, with a revenue cost to the Treasury of \$22.0 billion. The efficiency index is 77 percent.

The OMB method is based on the Tax Expenditure Budget, developed in 1968 by the Treasury Department under the direction of Stanley Surrey (1973). The Tax Expenditure Budget reports the estimated cost to federal taxpayers of the "loopholes" in the Internal Revenue Code during each fiscal year. Table 1 reports the revenue costs in the Tax Expenditure Budget for FY1990 at \$21.5 billion, very close to the \$22.0 billion derived from the Meltzer-Ott method. Thus it can be concluded that the costs to the federal taxpayer of tax exemption for state and local bonds were about \$22 billion. Applying the 77 percent efficiency index found by the Meltzer-Ott method, interest savings for states and local governments were \$16.9 billion.

It is important to note that in 1990 a large amount

Table 1
*Tax Expenditures in the Federal Income Tax:
 Revenue Losses from Exclusion of Interest
 on State and Local Debt, Fiscal Year 1990*
 Billions of Dollars

Total		\$21.515
Public Purpose Debt	\$10.730	
Private Purpose Debt	10.785	
IDBs for Businesses ^a	\$4.310	
IDBs for Authorities ^b	.720	
Mortgage Revenue Bonds	1.570	
Rental Housing	1.180	
Student Loans	.345	
Nonprofit Education	.235	
Nonprofit Health	2.190	
Veterans' Housing	.235	

^aIndustrial development bonds for energy facilities, pollution control, sewage and water facilities, small-issue IDBs.

^bIndustrial development bonds for airports, docks, sports and convention facilities, mass commuting.

Source: U.S. Office of Management and Budget (1990).

of private-purpose bonds received tax exemption, and only about 47 percent of these revenue losses were for public-purpose bonds. The use of tax-exempt bonds for private-activity purposes, particularly businesses, housing, and nonprofit hospitals, had been curtailed by the 1986 Tax Reform Act, but still involves significant revenue losses on bonds issued prior to August 1986.

II. Proposals for Municipal Bond Market Reform

Several reforms of the municipal bond market have been proposed, but as this section explains, none of them have been adopted. Instead, the market performance problems have been mitigated by a policy change that could not have been predicted 15 years ago: a dramatic reduction in the progressivity of personal income tax rates.

Elimination of Tax Exemption

One approach, which has little political support, would eliminate tax exemption and force municipalities to issue only taxable bonds. If this were done without grandfathering outstanding bonds, the U.S.

Treasury could recoup approximately \$22 billion to \$24 billion of tax revenues.

Because the efficiency, equity, and volatility problems all are due to the difference between yields of taxable and tax-exempt bonds, this approach would entirely eliminate those problems. It also would increase the cost of capital faced by states and local governments, as well as eliminate the human capital invested in the underwriting of tax-exempt bonds. The political power of the financial community and that of state and local government officials are reasons to doubt that this proposal will be implemented.

Substitution of a Direct Subsidy

A more moderate proposal would substitute a direct subsidy for tax exemption. In order to do this, Congress might eliminate tax exemption entirely, restricting states and local governments to issuing taxable bonds. Congress could then restore a capital cost subsidy by committing the U.S. Treasury to pay each state or local government a direct subsidy related to the size of its interest payments. If the Treasury wrote checks to states and local governments in amounts equal to the proportion σ of their interest payments on taxable bonds, the net interest cost of municipal borrowing would be $(1 - \sigma)R_T$.

Elimination of tax exemption cuts the connection between tax rates and the demand for municipal bonds. In effect, the demand schedule for municipal bonds becomes horizontal at an interest rate ratio of 1.0: the interest rate ratio will be unity or, stated differently, the municipal bond yield, R_M , will always equal the taxable bond rate. The total interest paid by municipalities will be $R_T Q_M$.

The payment of a direct subsidy equal to the proportion σ of interest payments reduces the net interest paid by state and local governments on taxable bonds from R_T to $R_T(1 - \sigma)$. Whether municipalities are better off under the direct subsidy plan than under tax exemption depends on the subsidy rate: if $\sigma > t_m$, the direct subsidy will reduce interest costs by more than the value of tax exemption. If, in addition, $\sigma < \bar{t}$, the direct subsidy will also reduce the costs to the Treasury. Thus, any value of the subsidy rate between \bar{t} and t_m will make both levels of government better off while also eliminating the equity and efficiency problems.

Why has this reform not received much support? This seems especially surprising since the subsidy rate could be set high enough to increase the capital cost subsidy to state and local governments and still

Table 2

Measurement of the Equity and Efficiency Effects of a Taxable Bond Option (TBO)

Item	Before TBO	After TBO	Increment
Cost to U.S. Treasury	$A_1 + A_2 + B_1 + B_2 + C_1$	$A_1 + A_2 + B_1 + B_2 + C_1 + C_2$	C_2
on Tax-Exempts	$A_1 + A_2 + B_1 + B_2 + C_1$	$A_1 + B_1 + C_1$	$-(A_2 + B_2)$
on Taxables	none	$A_2 + B_2 + C_2$	$A_2 + B_2 + C_2$
Interest Savings of State and Local Governments.	$A_1 + A_2$	$A_1 + A_2 + B_1 + B_2 + C_2$	$B_1 + B_2 + C_2$
on Tax-Exempts	$A_1 + A_2$	$A_1 + B_1$	$B_1 - A_2$
on Taxables	none	$A_2 + B_2 + C_2$	$A_2 + B_2 + C_2$
Windfall Income of Investors	$B_1 + B_2 + C_1$	C_1	$-(B_1 + B_2)$
on Tax-Exempts	$B_1 + B_2 + C_1$	C_1	$-(B_1 + B_2)$
on Taxables	none	none	none

Note: The areas in this table are defined in units of the taxable bond rate. To convert them to dollar values, each area should be multiplied by R_T . See Figure 4 for a visual presentation.

any rate ratio greater than $(1 - \sigma)$ no tax-exempts will be issued, so when $R_M > (1 - \sigma)R_T$ the supply schedule coincides with the vertical axis. Finally, at $R_M = (1 - \sigma)R_T$ the supply schedule is horizontal between the vertical axis and SS. Thus, with a taxable bond option the equilibrium interest rate ratio will be $(1 - \sigma)$, the amount of tax-exempt bonds outstanding will be Q_{TE}^* , and the amount of taxable bonds will be $(Q_M^* - Q_{TE}^*)$.

The taxable bond option will eliminate the volatility problem because the equilibrium ratio of tax-exempt to taxable interest rates will be set at $(1 - \sigma)$: any shifts in DD or SS will alter the composition of municipal debt, but will not affect the equilibrium interest rate ratio. For example, a rightward shift in SS in Figure 4 will lead to an increase in municipal bonds outstanding, all of which will be in the taxable form issued at the net cost of $R_T(1 - \sigma)$. Thus, the interest rate ratio is unaffected by either supply or demand shifts because taxable bonds are the marginal form of debt.

The equity and efficiency problems are only partially eliminated by the taxable bond option: they are totally eliminated for all taxable bonds issued, but they continue (though at a lower level) for tax-exempt bonds. To show this, Figure 4 has been decomposed into six areas: A_1 , B_1 and C_1 apply to the tax-exempt bonds sold, while A_2 , B_2 and C_2 apply to taxable bonds. Table 2 shows the interpretation of each of these areas.

The taxable bond option must cost the federal taxpayer more than reliance on tax exemption alone.

This incremental cost arises because the taxable bond option has an effect only if $\sigma > t_m$, that is, if at the margin the direct subsidy exceeds the indirect subsidy of tax exemption. Because the federal costs of any tax-exempts issued will not change (being determined by the tax rates of the investors in tax-exempts), the total costs to the federal taxpayer must rise. The size of this additional cost is measured by $R_T * (\text{area } C_2)$.

The interest savings enjoyed by state and local governments will increase by $R_T * \text{area } (B_1 + B_2 + C_2)$: $R_T * \text{area } B_1$ is the value of increased interest savings on tax-exempt bonds that are still issued, while $R_T * \text{area } (B_2 + C_2)$ is the increased savings on the volume of debt that shifts from the tax-exempt to the taxable form. Thus, a taxable bond option will increase the interest savings enjoyed by state and local governments.

In summary, a taxable bond option will eliminate the volatility problem and mitigate the equity and efficiency problems. The magnitude of the reduction in the equity and efficiency problems will depend upon the subsidy rate on taxable bond interest: the higher the subsidy rate the greater will be the share of municipal bonds issued in the taxable form, and the lower will be the equity and efficiency problems. Indeed, if the subsidy rate were set at t_{max} , all municipal debt would be issued in the taxable form, and equity and efficiency problems would be eliminated. This case would replicate the results achieved by legislative elimination of tax exemption and a direct subsidy rate of t_{max} .

The taxable bond option is clearly a compromise, which maintains tax exemption but also induces municipalities to issue taxable bonds. It has been opposed by the same groups that have opposed the more extreme reform of completely eliminating tax exemption and replacing it with a direct subsidy on taxable municipal bonds. While the opposition has been a bit less monolithic—with, for example, less concerted opposition among municipal finance officials—it has been sufficiently vigorous to prevent adoption of the taxable bond option.

A Flat Income Tax

A fourth approach to reforming the municipal bond market adopts a flat rate schedule for personal income taxes. Recall that the upward slope of the demand schedule in Figure 2 occurs for two reasons. First, the progressivity of the income tax rate schedule means that additional bonds outstanding must induce a higher rate ratio to compensate investors with tax brackets lower than that of the initial marginal investor. Second, each investor requires a risk premium, which increases with his holding of municipal bonds. With a flat tax-rate schedule the progressivity component disappears, and the market demand function depends solely on the risk premium schedules of individual investors. The mar-

Another approach to reforming the municipal bond market would adopt a flat rate schedule for personal income taxes.

ket demand schedule will, therefore, be flatter. This will reduce the instability, efficiency, and equity problems.

The Tax Reform Act of 1986 introduced a two-bracket personal tax rate schedule (15% and 28%), initiating a major step toward a flat rate system.⁵ The 1990 Revenue Reconciliation Act, which adopted a 15 percent, 28 percent, 31 percent schedule, was an additional step in this direction. While the move

toward a flat-rate system was not due to any effort to mitigate the problems of tax exemption, it has had that effect. The major appeal of this approach is political. High-income investors are happy to trade the value of municipal bonds as a tax shelter for lower tax rates; state and local governments still receive a subsidy (though it is smaller) and do not face the uncertainty about continued payment of a direct subsidy; and municipal bond underwriters do not find the demand for their services dramatically threatened.

III. Resource Allocation and Economic Efficiency

The equity and financial efficiency problems of the municipal bond market are not "social costs." Rather, they are "zero sum" costs in the sense that one sector's gain is matched by another sector's loss. For example, the financial efficiency problem is zero sum because it affects the distribution of income, not the aggregate amount of income received: the gains enjoyed by state and local government taxpayers through lower interest costs, and by affluent investors through windfall income, are matched by costs to federal taxpayers.

This section focuses on the social costs of tax exemption. The problem of social costs, or *economic inefficiency*, is inherent in any capital-cost subsidy; it will occur even in the absence of market instability or efficiency and equity problems.

The core of the social cost problem is the resource allocation effect of tax exemption. Because the exemption reduces the net interest cost paid by municipalities, it alters the relative amounts of capital and labor that states and local governments use to produce public goods. In addition, by affecting the relative prices of public and private goods, it induces economic agents to demand more public goods and fewer private goods, thereby shifting the composition of aggregate production. The ultimate effect of a capital cost subsidy enjoyed by the public sector (but not by the private sector) is to increase the share of output produced by the public sector, and to increase the relative capital intensity of public sector production.

⁵ In effect, however, this two-bracket schedule became a four-bracket one (15%, 28%, 33%, 28%) as the phasing-out of the 15% bracket and personal exemptions took effect at certain income levels only.

The Microeconomics of Economic Efficiency

The effects of tax exemption on resource allocation can be examined using standard microeconomic analysis. Figure 5 shows an Edgeworth-Bowley Box designed to illustrate this problem. The economy has two sectors: the private sector, designated by the subscript "p", and the state and local government sector, designated by the subscript "g". The two factors of production are capital, designated by K, and labor, designated by L. The box assumes that the total amount of each factor is fixed: the width of the box shows the total amount of capital (\bar{K}) and the height of the box is the total amount of labor (\bar{L}). Eastward movements represent a shift in capital from the private to the government sector (a rise in K_g and an equal decline in K_p), while northward movements represent a shift in labor allocation from the private to the government sector (a rise in L_g and an equal decline in L_p).

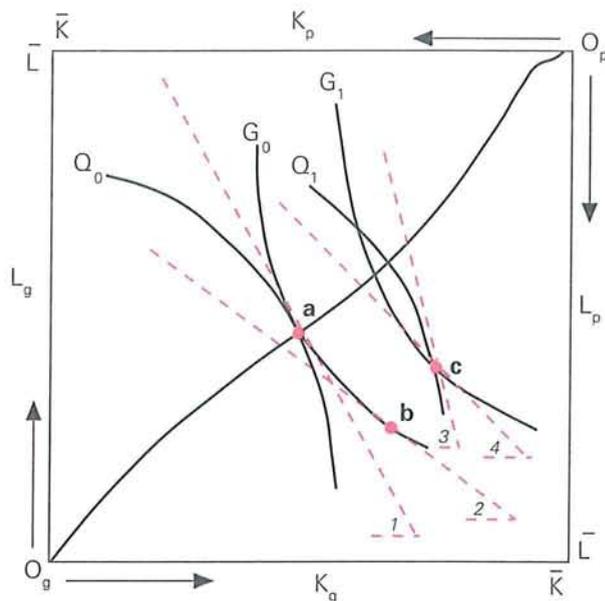
The economy's allocation problem is to determine how each factor will be allocated between the private and public sectors. This also determines how much of each good is produced. The Pareto-Optimal allocation of resources will place the economy on the curve connecting the southwest corner of the box, labeled O_g , to the northeast corner, labeled O_p . Any allocation of resources that moves the economy off this curve is an inefficient allocation, because it reduces the output of one sector without increasing the output of the other.

The southwest corner of the box is the origin from the vantage point of the government sector. At O_g the government sector uses no capital or labor and produces no output, while the private sector employs $K_p = \bar{K}$ and $L_p = \bar{L}$, producing the largest possible private output consistent with the economy's factor supplies. The number of government-sector "isoquants" is infinite; each isoquant is convex to this origin, and each shows the amounts of K_g and L_g that produce a given amount of public goods. For example, the curve labeled G_0 shows the combinations of government capital and labor that produce the amount G_0 of public goods, and the curve G_1 is the isoquant for a higher level of public goods. We know that $G_1 > G_0$ because some points on G_1 employ more of one factor while employing the same amount of the other factor, hence G_1 must represent higher output. Thus, the further northeast a government sector isoquant is, the higher the public good output that it represents.

The northeast corner, O_p , is the origin for the private sector, where no factors are employed by the

Figure 5

Tax Exemption and Resource Allocation



private sector and no private output is produced. At O_p all of the economy's capital and labor is employed by the government sector, and public good output is maximized. An infinite number of private sector isoquants, each convex to the origin O_p , represent the amounts of capital and labor necessary to produce a given level of private goods. Two of these isoquants are shown as Q_0 and Q_1 , with Q_0 representing the higher level of private output.

Each sector is assumed to minimize its production costs. It does this by equating the marginal value product of each factor to its price. If P_g and P_p are the prices of government and private goods, P_K and P_L are the prices of capital and labor, R is the interest rate, and δ is the depreciation rate on capital, cost minimization requires the satisfaction of the following conditions:

Government Sector: $P_g MP_{K,g} = P_K(R + \delta)$ and $P_g MP_{L,g} = P_L$ (5)

Private Sector: $P_p MP_{K,p} = P_K(R + \delta)$ and $P_p MP_{L,p} = P_L$

The equilibrium ratios of marginal products are

Government Sector: $MP_{K,g}/MP_{L,g} = [P_K(R + \delta)]/P_L$
(6)

Private Sector: $MP_{K,p}/MP_{L,p} = [P_K(R + \delta)]/P_L$

The marginal product ratios for each sector are represented by the slope of the isoquant for that sector. Because both sectors face the same factor prices, each sector will be induced to choose factor combinations that have the same marginal product ratios, that is, the same isoquant slopes. As noted above, the line connecting O_g and O_p is composed of all the points that represent an efficient allocation of resources. This line also turns out to be all the points at which the isoquants are tangent and, therefore, have equal slopes.

For example, consider point a, assumed to be the point at which the economy rests before introduction of tax exemption. At point a the isoquant G_0 is tangent to the isoquant Q_0 . Any other point on G_0 will, because of the shapes of the isoquants, be on a lower (more northeasterly) private-sector isoquant than Q_0 . Thus, any movement away from a gives lower private output for the same level of government output. The result is economic inefficiency, because the level of private output is lower than necessary to produce G_0 of public output.

In order to investigate the effects of tax exemption, assume that the economy is initially in a general equilibrium at point a, and that both sectors pay the same user cost of capital and wage rate. At this initial general equilibrium, the economy is Pareto-Efficient. If tax exemption is introduced, and the interest rate paid by the government sector, R_M , is below the rate paid by the private sector, R_T , then the relative factor costs for governments will be $P_K(R_M + \delta)/W$, measured on the box by the angle $\angle 2$. The private sector still faces the same factor price ratio, measured by $\angle 1$, so it wishes to remain at point a. But the government sector would want to move to point b, which minimizes the cost of producing G_0 of output under the new factor cost ratio.

Tax exemption has thrown the economy into disequilibrium: the private sector wants to use the amount of capital and labor represented by point a, leaving the government sector only $\bar{K} - K_p$ of capital and $\bar{L} - L_g$ of labor. In the initial equilibrium that was precisely the amount of capital and labor that the government sector wanted to use. But now the government wants to use more capital and less labor. In short, the introduction of tax exemption creates an excess demand for capital and an excess supply of labor. Furthermore, tax exemption has driven a per-

manent wedge between the factor prices faced by the private and public sectors: as long as R_M is less than R_T , the private sector faces a higher cost of capital relative to the cost of labor than does the government sector. Because of this wedge, the economy can never come to an equilibrium on the line O_gO_p : it can never be Pareto-Efficient.

Where is the new general equilibrium? Clearly the excess demand for capital must lead to a rise in the user cost of capital in the private sector; $P_K(R_T + \delta)$ must rise. Also, the excess supply of labor must lead to a fall in the wage rate, W , as labor becomes unemployed in the government sector and seeks employment in the private sector. The migration of capital to the government sector and of labor to the private sector, and the rise in the cost of capital combined with a decline in the cost of labor, will continue until the economy reaches a new point, like point c.

At point c the factor choices of the two sectors are consistent: the private sector wants to employ factors in exactly the amounts necessary to maintain full employment. Also, each sector is once again minimizing its production costs because it is once again equating the relative marginal products (slope of isoquant) to the relative factor costs. However, the relative factor costs, which were equal at a, are not equal at c: at point c $\angle 3$ is the factor price ratio for the

Tax exemption can drive a permanent wedge between the factor prices faced by the private and public sectors.

private sector, while $\angle 4$ is the price ratio for the public sector. Because $\angle 4 < \angle 3$, the government sector has a marginal product of capital less than that in the private sector and a marginal product of labor greater than that in the private sector.

The public sector is now producing with a higher level of capital intensity, while the private sector is producing at a lower capital-labor ratio. Clearly, point c is not Pareto-Efficient because a Pareto improvement would occur if resources were reallocated to reach a point on O_gO_p : this would allow production of more of one good with no sacrifice in the produc-

tion of the other good. But the price system will not induce that movement; the government has a permanent incentive to produce with too much capital and too little labor.

How far apart will be points a and c? Will c be to the southeast of a (more capital employed but less labor in the public sector) or to the northeast of a (more capital and more labor in the government sector)? The answers depend on two important considerations: technology, which fixes the substitutability between factors and thereby affects the curvature of the isoquants, and preferences, which determine the consumers' willingness to substitute private goods for public goods. So far as technology is concerned, the higher is the "elasticity of substitution" between capital and labor in each sector, the more each sector will alter its capital-labor ratio in response to the change in relative factor prices. For each sector, the elasticity of substitution has a minimum of zero, which corresponds to a fixed-coefficients technology. If both sectors have a zero elasticity, the curve $O_g O_p$ would represent the only possible points of equilibrium. Thus, if no factor substitutions can be made, no misallocation of factors between sectors can occur.

At the other extreme, the elasticities can be extremely high, approaching straight-line isoquants. In that case, very small changes in relative factor prices will induce extremely large changes in factor proportions, and the resource allocation effects of tax exemption will be large.

The final equilibrium will also be affected by preferences, which affect the substitutability between government and private goods. This is measured by the price elasticity of demand for government goods. Tax exemption will induce a fall in the relative price of government goods. If relative product prices have a very small effect on demand, tax exemption will have little effect on the relative quantities of each good; point c will be very close to point a. If, on the other hand, private and public goods are close substitutes, larger shifts in the mix of products will occur.

Except in the extreme case of zero substitution between factors and zero price elasticity of demand, it will always be the case that a capital-cost subsidy for the government sector will induce capital to move from the private sector to the public sector. However, the direction of labor movements will depend upon the price elasticity of demand for public goods. If this is sufficiently high, the capital-cost subsidy will induce consumers to switch from private to public goods so much that the public sector increases its employment of both capital and labor.

Measuring the Resource Allocation Effects of Tax Exemption

Arnold Harberger (1962) developed a simple general equilibrium analysis of the effects of taxation on the allocation of resources. In the intervening 30 years, a number of extensions and refinements of the basic model have been developed, but the Harberger model has become the standard for analyzing the resource allocation effects of a wide range of taxes. This section outlines the Harberger model. In the next section, the model is employed to derive estimates of the resource allocation effects of tax exemption.

The Harberger model is a formalization of the insights in Figure 5. The model assumes two producing sectors in the economy, a public sector, producing a "public" good in the quantity G , and a private sector, producing a "private" good in quantity Q . Each sector employs two factors of production, capital (K) and labor (L). The total amount of each factor is fixed in quantity, so that the factor allocation problem is restricted to the allocation of the total quantity of each factor between the two sectors. It is assumed that full employment of both factors prevails, so that no factor units fail to be allocated to production in the economy. Thus, if K_g and L_g are the capital and labor employed in the untaxed (government) sector that produces G , and \bar{K} and \bar{L} are the total amounts of capital and labor, then $K_p = \bar{K} - K_g$ and $L_p = \bar{L} - L_g$ are the capital and labor employed by the taxed (private) sector to produce.

Each sector has a production function, designated $Q = F(K_p, L_p)$ and $G = G(K_g, L_g)$, respectively. Each sector employs each factor up to the point where the marginal product value is equal to the factor price. The factor price of capital in the untaxed (government) sector is C_K , while the factor price of capital in the taxed (private) sector is $C_K + \Theta$, where Θ is the capital-income tax per unit of capital. The model assumes competitive factor markets, so that each factor is paid its marginal product value. Also, production functions exhibit constant returns to scale.

Three primary parameters affect the size of resource allocations resulting from a tax on capital in one sector. The first two are the elasticity of substitution between capital and labor in the two sectors, denoted by σ_p and σ_g : the greater is either σ_p or σ_g , the larger will be the changes in the capital labor ratios in the associated sector when factor prices change, and the smaller will be the changes in the

relative factor prices associated with changes in factor composition. This follows the general principle that the closer the substitutability between any two commodities, the larger will be the response in the ratio of the quantities used to any relative price change. Thus, a given change in relative quantities can be achieved by a smaller relative price change when two commodities are close substitutes.

The third primary parameter is the price elasticity of demand for the public good, E_g . The higher this price elasticity, the larger will be the shift in the allocation of the consumers' consumption bundle in response to any change in relative prices; for any given change in relative prices, the shift in demand between the taxed and untaxed sectors is greater when the goods are closer substitutes.

Estimates of the Effects of Tax Exemption

To estimate the resource allocation effects of tax exemption, it is necessary to assume values for the primary parameters, discussed above, which describe the response of economic agents to changes in relative prices. In addition, values must be assigned to several secondary parameters, which describe the allocation of resources in the economy. Among these are the capital income shares in each sector (f_K and g_K), the initial ratio of government sector capital to private sector capital (λ_K) and the initial ratio of government labor to private labor (λ_L).

The appropriate values of these secondary parameters will depend upon the definition of the private sector. Is it defined as nonfinancial corporations, all corporations, or all businesses including unincorporated enterprises? Does it include production of housing services? of farm output? The private sector has no single definition; here it has been defined to include all private nonagricultural production of goods and services except housing.

The U.S. Bureau of Labor Statistics establishment surveys of nonagricultural payrolls show that in the 1980-85 period there were 17.4 state and local sector employees for every 100 private sector employees; hence $\lambda_L = 0.174$. The U.S. Commerce Department's capital stock estimates (Musgrave 1990) indicate that in the 1982-89 period there was an average of \$40.50 of state and local sector capital for every \$100 of fixed nonresidential capital stock; hence, $\lambda_K = 0.405$.

According to Hulten and Schwab (1987), in the 1980-85 period about 24 percent of the value added in the state and local government sector was due to the services of the capital stock, hence $g_K = 0.24$. The

National Income Accounts indicate that over the same period about 60 percent of private sector value added was attributable to labor compensation, thus $f_K = 0.40$.

A great deal of work has been done on the elasticity of substitution between capital and labor in the private sector. The consensus appears to put this at somewhat less than unity; this study has chosen $\sigma_p = 0.90$ (Beckmann and Sato 1969). Considerably less agreement can be found about the elasticity of substitution in the state and local sector. Fortune (1983) reports results consistent with a Cobb-Douglas technology, implying $\sigma_g = 1.0$, a result supported by several studies cited in Blackley and DeBoer (1991).

However, one long-standing position argues that public sector activities are labor-driven and that the public sector does not have the same flexibility in the capital-labor ratio that the private sector enjoys (Baumol 1967; Baumol, Blackman, and Wolff 1985). This, it is argued, means that new capital-intensive technologies are not easily introduced and that the ability to substitute between capital and labor when relative prices change is weak. The result is low productivity growth and rising production costs in the state and local sector. A recent paper by Blackley and DeBoer (1991) supports the Baumol hypothesis, finding that capital and labor are weak complements. In order to allow for a wide range of estimates, this study has assumed two possible values of the state and local elasticity of substitution: $\sigma_g = 1.0$ and $\sigma_g = 0.25$.

The final parameter whose value must be assumed is the price elasticity of demand for state-local goods, E_g . A survey of the literature by Inman (1979) reported an average value of 0.50 for the uncompensated elasticity. DeBartolo and Fortune (1982) estimated the compensated elasticity at 0.15.⁶ Both values will be used here.

The Harberger model calculates the effect of a tax imposed on each unit of capital in the private sector. The value of Θ must be derived from an analysis of the impact of tax exemption on the cost of capital for the private sector. The optimality condition for the capital stock is given below in Equation (7), where $MP_{K,p}$ is the marginal physical product of capital, τ is the tax rate, Z is the present value of depreciation allowances, C_K is the nominal after-tax rate of return required on capital goods, π is the anticipated inflation rate, δ is the depreciation rate for private capital

⁶ The uncompensated price elasticity of demand includes the income effect of a relative price change, and is typically larger than the compensated elasticity, which is the substitution effect.

Table 3
Parameter Values Used in the Economic Efficiency Model

Parameter	Definition		Values
f_K	Capital Share of Value Added, Private	1980-85	.40
g_K	Capital Share of Value Added, Public	1980-85	.24
λ_K	Ratio of Public/Private Employment	1980-85	.174
λ_L	Ratio of Public/Private Capital Stock	1982-89	.405
σ_p	Elasticity of Substitution, Private	—	.90
σ_g	Elasticity of Substitution, Public	—	.25, 1.00
E_g	Price Elasticity of Demand, Public Goods	—	.15, .50
θ	Added User Cost of Private Capital	1980-85	.03

Note: The private sector is non-farm. The public sector is all state and local governments.

and γ is the rate of change in the relative price of capital goods.

$$(7) \quad (1 - \tau)P_p MP_{K,p} = P_K(1 - \tau Z)[C_K - \pi + \delta + \gamma].$$

This can be converted to the following condition for the marginal product of capital:

(8)

$$MP_{K,p} = (P_K/P_p)(1 - \tau Z)\{[C_K - (\pi - \delta - \gamma)]/(1 - \tau)\}.$$

The right-hand side of Equation (8) is the appropriate definition of the user cost of capital for the purposes of this study. Following Miller (1977) this study adopts the view that, in security market equilibrium, the after-tax required return on capital, C_K , is $R_T(1 - \tau)$, so the "grossed up" pre-tax return required on capital is simply R_T .⁷ If issuance of tax-exempt bonds were extended to the private sector, the before-tax interest rate would be R_M rather than the higher rate R_T .⁸ Thus, the additional cost of capital paid by private businesses because they are not allowed to issue tax-exempt debt, assuming that π , δ and γ are independent of the existence of tax exemption for private debt, is:

$$(9) \quad \Theta = (P_K/P_p)(1 - \tau Z)[R_T - R_M].$$

During the period 1980-85, the corporate tax rate was 0.46 and the present value of depreciation for \$100 of investment was about \$40 for equipment and \$28 for structures (Kopcke 1981).⁹ Because the private fixed nonresidential capital stock was split almost equally between equipment and structures, the average value of Z for the 1980s was 0.34. Assuming

$(P_K/P_p) = 1.0$ and employing the 1980-85 average for Aaa corporate bond yields and Standard & Poor's high-grade municipal bond yields ($R_T = 0.1267$ and $R_M = 0.1002$), then $\Theta = 0.0265$: tax exemption is equivalent to imposing a tax of 2.65 cents per unit of private sector capital. This analysis uses $\Theta = 0.03$.

The assumed parameter values are reported in Table 3. The results of the analysis are reported in Table 4. As expected, tax exemption reduces the net cost of capital for the public sector and raises the private sector cost of capital. It also creates a rise in the public sector's capital-labor ratio ranging from 0.66 percent (if $\sigma_g = 0.25$) to 2.13 percent (if $\sigma_g = 1.00$). Furthermore, aggregate output is reduced by 0.07 to 0.23 percent, with private output falling and public output increasing. The magnitude of the decline in aggregate output depends upon the elasticity of substitution between capital and labor in the public sector: the greater is σ_g , the larger the reallocation of

⁷ Miller argues that firms will use the lowest-cost form of finance. Let $R_T(1 - \tau)$ be the interest rate paid by businesses on long-term debt (with deductibility of corporate interest), and R_E be the return required by investors on equity. Furthermore, apart from risk differentials, investors will require that equity generate a yield equal to the yield on municipal bonds, so $R_E = R_M$. In this case, $C_K = \min[R_T(1 - \tau), R_M]$. Because equity and debt are very close substitutes, firms will issue one or the other until a security market equilibrium is achieved in which $R_T(1 - \tau) = R_M$. In this case, one can use either R_T or $R_M/(1 - \tau)$ as the required pre-corporate tax return on capital.

⁸ Following the logic of the previous footnote, extension of tax-exempt bonds to private businesses would mean that the interest cost of capital is $C_K = \min[R_M(1 - \tau), R_M]$. In this case, debt is always cheaper than equity, so $C_K = R_M(1 - \tau)$. As a result, the pre-tax interest cost is R_M .

⁹ This assumes: a pre-tax discount rate of 10 percent, service lives of five and 20 years for equipment and structures, respectively, and sum-of-years digits depreciation.

Table 4
Estimated Social Costs of Tax Exemption:
Selected Values of σ_g and E_g
 Percent except where noted

	$E_g = .15$		$E_g = .50$	
	$\sigma_g = .25$	$\sigma_g = 1.00$	$\sigma_g = .25$	$\sigma_g = 1.00$
<u>User Cost of Capital</u>				
Private Sector	+ .29%	+ .86%	+ .36%	+ .92%
Public Sector	-2.71	-2.14	-2.64	-2.08
<u>Product Prices</u>				
Private Sector	+ .12	+ .35	+ .14	+ .37
Public Sector	-.65	-.51	-.63	-.50
Aggregate	+ .03	+ .25	+ .05	+ .27
<u>Output</u>				
Private Sector	-.10	-.24	-.17	-.32
Public Sector	+ .12	+ .13	+ .39	+ .43
Aggregate	-.07	-.20	-.11	-.23
<u>Capital Employed</u>				
Private Sector	-.26	-.71	-.36	-.82
Public Sector	+ .63	+ 1.75	+ .89	+ 2.01
<u>Labor Employed</u>				
Private Sector	+ .00	+ .07	-.04	+ .01
Public Sector	-.05	-.38	+ .23	-.07
<u>Capital per Employee</u>				
Private Sector	-.26	-.78	-.32	-.83
Public Sector	+ .68	+ 2.13	+ .66	+ 2.08
<u>Aggregate Output</u>				
Aggregate				
(\$ billions)	-\$2.38b	-\$6.56b	-\$3.43b	-\$7.63b
Per Capita (\$)	-\$10.18	-\$28.15	-\$14.66	-\$32.67

Note: The calculations are based on 1980-85 data. The aggregate output index is a Divisia index of the proportional changes in public and private output, with a 12 percent share of value added in the public sector.

capital and labor and the greater the decline in aggregate output.

The last two rows of Table 4 translate the proportional change in aggregate output to dollar values, using the 1980-85 average level of nonfarm, nonfederal value added. The decrease in aggregate output ranges from \$2.38 billion, for low values of E_g and σ_g , to \$7.63 billion for high values of those parameters. Translated to per capita values, tax exemption costs from \$10.18 to \$32.67 per person. The estimate using the preferred parameter values ($E_g = 0.50$, $\sigma_g = 0.25$) is a total of \$3.43 billion or \$14.66 per capita.

These estimates indicate that tax exemption creates mild social costs in the form of output forgone. The amounts are not so dramatic as to make the exemption a serious problem on this score, but they

are also not so small as to make the analysis of social costs irrelevant. The reader should be aware that these costs are measured against the alternative of a perfectly competitive economy without subsidies; for the imperfect world we face, the introduction of tax exemption could, in fact, improve resource allocation. Indeed, those who believe that public infrastructure is insufficient in the U.S. economy argue that market outcomes do not efficiently allocate resources and that government should intervene to increase public infrastructure.

IV. Summary and Conclusions

This article addresses the policy issues created by the exemption of municipal coupon payments from the federal income tax base. The four fundamental problems are the trio of market instability, equity, and financial efficiency, which are "zero-sum" problems that have their primary effect on the distribution of income; and the problem of economic efficiency, which addresses social costs.

The market instability problem arises because tax exemption creates greater variability in the interest rates paid by state and local governments, hence it provides a capital-cost subsidy that varies according to financial market conditions. The equity problem is that tax exemption provides an opportunity for affluent taxpayers to increase their after-tax income, thereby forcing less affluent taxpayers to take on a greater tax burden. The financial efficiency problem is the mirror image of the equity problem: the cost to the federal taxpayer exceeds the interest savings of state and local governments by the amount of windfall income enjoyed by affluent investors.

This study has calculated the magnitude of these three problems in 1990, concluding that tax exemption cost the U.S. Treasury about \$22 billion to \$24 billion in that year, in exchange for which states and local governments reduced their interest payments by \$16.9 billion to \$18.5 billion. The "missing" amount, \$5.1 billion to \$5.7 billion, was received by affluent investors as "windfall income." Only a portion of the windfall income enjoyed by affluent investors is really excess income, or "investor's surplus," however; the remainder is required by investors to compensate them for the risk and illiquidity of municipal bonds. If, as is speculated here, about 50 percent of windfall income is investor's surplus, the excess income of affluent investors from tax exemption was about \$2.5 billion in 1990.

This article also discusses several possible reforms of the municipal bond market that would eliminate or mitigate the zero-sum problems. The first is elimination of the exemption. The second is a direct subsidy, which would eliminate the exemption but replace it with federal payment of a portion of state and local government interest costs. The third is a taxable bond option, a combination of tax exemption and a direct subsidy. None of these reforms have received sufficient political support, but the problems have been mitigated by the major changes in the tax code under the Tax Reform Act of 1986. The move toward a flat income tax system has reduced the

magnitude of these problems.

Finally, the article discusses the social costs of tax exemption, which arise from the loss of output as resources are reallocated from the private sector to the public sector in response to lower public sector capital costs. Using the period 1980–85 as the basis for estimates, this study concludes that in the 1980–85 period the tax exemption reduced the annual aggregate output (value added) of the nonfarm, non-federal government sector by \$2.4 billion to \$7.6 billion, depending on the assumptions. The preferred estimate is \$3.4 billion, which translated to per capita amounts equals \$14.66 per person.

Appendix: Measuring the Cost of Tax Exemption

The U.S. Treasury Department used the Meltzer-Ott method in 1965 (Joint Economic Committee 1966) to calculate the interest savings and revenue costs on state and local bonds sold in 1965, over the lifetime of those bonds. The Treasury Department estimated an average marginal tax rate of 42 percent and a marginal tax rate of 28 percent. The interest savings over the lifetime of gross state and local bonds newly issued in 1965 were \$1.9 billion, with a revenue cost of \$2.9 billion. Using formula (4) above in the text, these estimates imply an efficiency index of about 65 percent.

These early Treasury estimates are incorrect because they rest on a confusion between average and marginal analysis. The bonds sold in 1965 were incremental to the stock of outstanding municipal bonds, and the likely purchasers were the near-marginal investors in tax-exempts, whose windfall income would be very small. But the 1965 application of the Meltzer-Ott method assumes that the incremental supply of bonds is bought by the *average* investor, whose tax rate is measured by the average marginal tax rate. The result is a potentially serious exaggeration of the costs of new bond issues. The method employed is, therefore, more suitable to estimation of the costs of eliminating tax exemption for all outstanding bonds; in this case the average marginal tax rate is relevant.

The Meltzer-Ott method also makes some strong assumptions about market adjustments that occur in response to tax exemption. First, the method infers tax rates from the existing pattern of ownership of municipal bonds, and assumes that in the absence of tax exemption those owners would simply have bought taxable bonds (including, of course, taxable municipals) to replace the no-longer-available tax-exempt bonds. Second, it assumes that the general level of interest rates on taxable securities is not affected by the existence of tax exemption. However, the adjustments that would occur if tax exemption did not exist are far more complex than these assumptions suggest.

Consider the second point first. The effect of tax exemption on the taxable bond rate depends on the elasticity of the supply of both taxable and tax-exempt bonds. The Meltzer-Ott method assumes that either the outstanding stock of municipal debt is independent of interest rates (as, for convenience, is assumed in the text) or the private

sector supply of debt is infinitely interest-elastic. In the first case, the introduction of tax exemption would induce governments to switch their issues from taxable to tax-exempt form, but investors would switch exactly that amount of their portfolios to tax-exempts and out of taxable bonds. Because the shift in demand for taxable bonds (as investors switch from taxables to tax-exempts) is exactly matched by the shift in the supply function (as governments issue tax-exempts rather than taxable bonds), the net result is no change in the taxable bond yield. In the second case, increased issues of municipal bonds in response to tax exemption "crowd out" an equal amount of taxable bonds, leaving the taxable bond yield unchanged.

If, in contrast to the assumption of the previous section, state and local governments respond to lower interest costs by issuing more bonds, the introduction of tax exemption will increase the quantity of loanable funds demanded and push up the general level of interest rates. As this happens, private borrowers will reduce their bond issues in response to the higher costs. Only if the supply of private taxable bonds is infinitely interest-sensitive will the taxable bond rate remain unchanged; if not, the taxable bond rate must go up.

Now consider the first point. The Meltzer-Ott method assumes that investors simply switch from tax-exempts to taxable bonds, so that the pattern of ownership of outstanding tax-exempt bonds indicates the relevant tax rates of those who would otherwise invest in taxable bonds. However, this need not be true. For example, suppose that tax exemption were eliminated for all outstanding municipal bonds and that current holders of tax-exempt bonds try to shift into the next best tax shelter—common stocks. In this case, portfolio changes might create no additional taxes apart from temporary capital gains tax revenues. The net effect on tax revenues will depend not on the tax rates of investors who switch from tax-exempts to equities, but upon the tax rates of those who sold the equities and switched into taxable bonds. Presumably these tax rates are lower than the rates of the former tax-exempt bondholders because the equity sellers gave up the tax shelter of municipal bonds. Thus, the method tends to overstate the relevant average marginal tax rate.

References

- Baumol, William J. 1967. "Macroeconomics of Unbalanced Growth: The Anatomy of Urban Crisis." *The American Economic Review*, vol. 57 (June), pp. 415-26.
- Baumol, William J., Sue Anne Batey Blackman and Edward N. Wolff. 1985. "Unbalanced Growth Revisited: Asymptotic Stagnancy and New Evidence." *The American Economic Review*, vol. 75 (September), pp. 806-17.
- Beckmann, Martin J. and Ryuzo Sato. 1969. "Aggregate Production Functions and Types of Technical Progress: A Statistical Analysis." *The American Economic Review*, vol. 59 (March), pp. 88-101.
- Blackley, P. R. and L. DeBoer. 1991. "The Structure of State and Local Government Production." *Public Finance Quarterly*, vol. 19, no. 2 (April), pp. 147-65.
- DeBartolo, Gil and Peter Fortune. 1982. "The Demand for Public Services: Inferences from Municipal Bond Referenda." *National Tax Journal* (March), pp. 55-67.
- Fortune, Peter. 1973a. "Tax Exemption of State and Local Interest Payments: an Economic Analysis of the Issues and an Alternative." *New England Economic Review* (May/June), pp. 3-31.
- . 1973b. "The Impact of Taxable Municipal Bonds: Policy Simulations with a Large Econometric Model." *National Tax Journal* (March), pp. 3-31.
- . 1983. "A Test of the Cobb-Douglas Production Function for Local Governments." *National Tax Journal* (June), pp. 233-36.
- . 1984. "Tax Exemption and Resource Allocation: Implications for Prices, Production and Factor Choice." *Public Finance Quarterly* (July), pp. 347-54.
- . 1988. "Municipal Bond Yields: Whose Taxes Matter?" *National Tax Journal* (June), pp. 219-33.
- . 1991. "The Municipal Bond Market, Part I: Politics, Taxes and Yields." *New England Economic Review* (September/October), pp. 13-36.
- Galper, Harvey and John E. Petersen. 1971. "An Analysis of Subsidy Plans to Support State and Local Borrowing." *National Tax Journal* (June), pp. 205-34.
- Harberger, Arnold. 1962. "The Incidence of the Corporation Income Tax." *Journal of Political Economy* (June), pp. 215-40.
- Huefner, Robert. 1971. *Taxable Alternatives to Municipal Bonds: An Analysis of the Issues*. Research Report No. 53, Federal Reserve Bank of Boston, Boston, MA.
- Hulten, Charles R. and Robert M. Schwab. 1987. "Income Originating in the State-Local Sector." NBER Working Paper No. 2314 (July).
- . 1991. "Public Capital Formation and the Growth of Regional Manufacturing Industries." NBER Conference on Tax Exemption, Annapolis, MD, April 1991.
- Inman, Robert P. 1979. "The Fiscal Performance of Local Governments: An Interpretive Review." In P. Mieszkowski and M. Straszheim, eds., *Current Issues in Urban Economics*, pp. 270-321. Baltimore MD: Johns Hopkins University Press.
- Joint Economic Committee. 1966. "Comparison of the Interest Cost Saving and Revenue Loss on Tax-Exempt Securities." In *State and Local Public Facility Needs and Financing*, 89th Congress, 2nd Session, Volume 2, pp. 327-33. Washington, D.C.: Government Printing Office.
- Kopcke, Richard W. 1981. "The Efficiency of Traditional Investment Tax Incentives." In Board of Governors of the Federal Reserve System, *Public Policy and Capital Formation*, pp. 163-73. Washington, D.C. (April).
- Metcalf, Gilbert. 1990. "Arbitrage and the Savings Behavior of State Governments." *Review of Economics and Statistics*, vol. 72, no. 3 (August), pp. 390-96.
- . 1991. "The Role of Federal Taxation in the Supply of Municipal Bonds: Evidence from Municipal Governments." *National Tax Journal*, vol. 44, no. 4, Part I (December), pp. 57-70.
- Miller, Merton. 1977. "Debt and Taxes." *Journal of Finance* (May), pp. 261-75.
- Munnell, Alicia. 1990. "How Does Public Infrastructure Affect Regional Economic Performance?" *New England Economic Review* (September/October), pp. 11-33.
- Musgrave, John. 1990. "Fixed Reproducible Tangible Wealth in the United States: Revised Estimates." *Survey of Current Business*, (September), pp. 51-75.
- Ott, David J. and Allan H. Meltzer. 1963. *Federal Tax Treatment of State and Local Securities*. Washington, D.C.: The Brookings Institution.
- Poterba, James M. and Daniel R. Feenberg. 1991. "Which Households Own Municipal Bonds? Evidence From Tax Returns." NBER Working Paper No. 3900 (November).
- Seltzer, Lawrence. 1941. "Possibilities of Speeding the Elimination of Tax-Exempt Securities." *National Tax Journal Proceedings, 1941*, pp. 189-99.
- Simons, Henry C. 1938. *Personal Income Taxation*. Chicago, IL: University of Chicago Press.
- Surrey, Stanley. 1973. *Pathways to Tax Reform: The Concept of Tax Expenditures*. Cambridge, MA: Harvard University Press.
- U.S. Office of Management and Budget. 1985, 1990. *Special Analyses, Budget of the United States Government*. Washington, D.C.: Government Printing Office.

Federal Reserve Bank of Boston
P.O. Box 2076
Boston, Massachusetts 02106-2076

Address Correction Requested

Bulk Rate
U.S. Postage
PAID
Richmond, VA
Permit No. 930