New England Economic Review

Federal Reserve Bank of Boston

Sept./Oct. 1990

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.

Contents

Contents

Oil Shock III?

Norman S. Fieleke

Vice President and Economist, Federal Reserve Bank of Boston. Lisa O'Brien and Thomas J. Miles provided research assistance.

As if to affirm that "History repeats itself," the third oil shock in seventeen years has been threatening. But history never repeats itself exactly, and important differences exist between this incipient oil shock and the previous two, including differences in the likely magnitude of the shock, in the vulnerability of oil consumers, in the state of the world economy, and in the economic policy options open to governments in the industrial oil-importing countries. All of these differences are favorable, tending to weaken the force of the blow.

The Magnitude of the Shocks

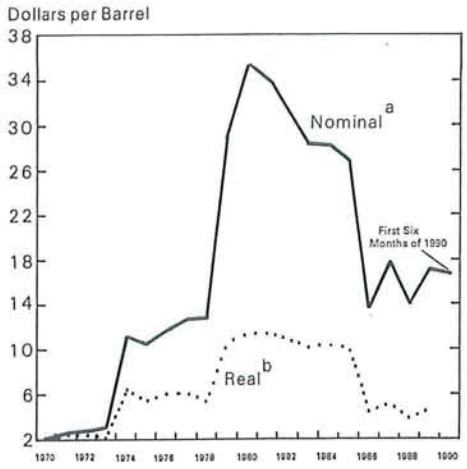
From 1973 to 1974 the average world price of crude petroleum rose by 261 percent, from $3.10 per barrel to $11.20, as the members of OPEC (Organization of Petroleum Exporting Countries)[1] restrained their output in the face of strong demand (Chart 1). In relation to the average price that the oil-exporting countries and others were paying for goods from the industrial countries—one measure of the "real" price of oil—the cost per barrel rose by a significantly lower, but still dramatic, 190 percent. This real price increase meant that the quantity of goods the industrial countries had to give up in exchange for a barrel of oil was nearly three times as great in 1974 as in 1973. Never before had the world witnessed an exercise of monopoly power on such a scale and with such success.

During the second oil shock the real price of crude went up by 110 percent, from 1978 to 1980. By 1986, however, this second price increase had been reversed.

The latest shock began to develop in July 1990, as the members of OPEC undertook negotiations over a reduction in their crude oil production. On July 27 they agreed upon a reduction designed to raise the average selling price of their crude to $21 a barrel, about 20 percent

Chart 1

World Crude Oil Price, 1970-90



[a]Average of Dubai, United Kingdom, and Alaskan North Slope spot crude prices, reflecting relatively equal consumption of light, medium, and heavy crudes worldwide.
[b]Nominal price deflated by the export unit value index for industrial countries, 1970=100.

Source: IMF, *International Financial Statistics Yearbook 1989* and *International Financial Statistics*, August 1990; and IMF staff.

above the level then prevailing. Some previous OPEC agreements to cut back on production had been undermined by countries that produced in excess of their assigned quotas, but this latest agreement was bolstered by an Iraqi threat of military action against such countries, a threat aimed chiefly at Kuwait and the United Arab Emirates. On August 2 Iraq followed up its threat with the invasion and conquest of Kuwait, and on August 6 the United Nations Security Council responded by calling for, among other things, an embargo against the acquisition of oil from either Iraq or Iraqi-occupied Kuwait. In consequence of all these developments oil prices soared.

Where prices will settle remains to be seen. Barring warfare that disrupts oil production or greatly increases demand, it seems unlikely that prices will average much more than $25 a barrel in the near future. Indeed, they might stabilize somewhat below that level, even if Iraq continues to occupy Kuwait and the embargo is fairly effective. The primary reason is that production in some countries within OPEC reportedly has been expanded to offset most of the reduction in the flow of oil from Iraq and Kuwait. Another consideration that should restrain the price increase, at least for the near term, is the unusually large stock of petroleum in storage, a matter that is discussed more fully below.

Consequences

For oil-importing countries, a sharp reduction in the supply of oil generates the loathsome combination of price increases and higher unemployment. It is not mere coincidence that each of the previous oil shocks was followed by both recession and a surge of inflation in the industrial countries. Any analysis of these developments should recognize that an abrupt reduction in the supply of oil should, in and of itself, generate only a one-time increase in the general price level, rather than a prolonged increase in the rate of inflation. Indeed, in principle it would be possible to prevent even a temporary bulge in the rate of inflation by precipitately tightening monetary policy.

The argument against such monetary policy shifts is that they are likely to cause recessions even in the absence of oil shocks. And the vulnerability to recession is even greater following an oil shock, because the oil price increase itself entails a contractionary, as well as an inflationary, influence. The contractionary influence is twofold: the residents of oil-importing countries are obliged to devote a larger share of their spending to foreign oil rather than to goods and services produced at home, since nothing can readily take the place of foreign oil; and certain lines of activity, such as the fuel-intensive airline industry and the manufacture of "gas-guzzling" automobiles, promptly suffer a decrease in sales and lay off employees who are not immediately absorbed by other industries.

Precise quantification of even the short-term effects of such oil price increases is far from simple, as we need to know—among other things—the importance of oil in both production and consumption, as well as the possibilities of substituting other factors for it, and the stance of monetary and fiscal policies. Nonetheless, a consensus seems to have emerged about the effects on output and inflation in the industrial countries. On average, a 10 percent rise in world oil prices is believed to lower real gross national product by about 0.2 percent and to raise consumer prices by perhaps 0.3 percent. A price of $25 per barrel would represent an increase of about 50 percent above the levels prevailing in 1989 and in 1990 just before the latest shock began.

Ten or fifteen years ago such a price increase

would have had consequences far more distressing than those foreseen today. As recently as 1986, many economists thought that a 10 percent rise in oil prices would lower real GNP in the typical industrial country by more than 0.5 percent and raise consumer prices by somewhat more, or roughly twice the magnitudes now accepted (Fieleke 1988, p. 5).

The reasons for this lessened vulnerability are straightforward. Most important is the progress made in using energy, especially oil, more efficiently. For the twenty-four OECD (Organisation for Economic Co-operation and Development) countries collectively, the quantity of oil consumed in generating each $1 billion (at 1985 prices) of gross domestic product was reduced from 254,000 metric tons in 1973 to 154,000 in 1988, or by nearly 40 percent (Table 1). For the United States, the comparable efficiency gain was 34 percent, while West Germany and Japan enjoyed more sizable gains of 41 percent and 48 percent, respectively. Over this same period the total amount of energy (including oil) required to produce a unit of GDP was also lowered substantially: by 24 percent for all OECD countries as a group, 23 percent for West Germany, 26 percent for the United States, and 30 percent for Japan. This progress was, of course, stimulated in good measure by the sharp increases during the oil shocks in the price of oil and of energy sources that could substitute for oil.

A second reason for the heightened immunity is that oil now supplies a smaller share of the world's energy usage than at the onset of the two earlier shocks. As reported in Table 2, the share of world primary energy production accounted for by crude oil was nearly one-half in 1973 and 1978, but had declined to slightly more than one-third by 1988.

Still another reason that an oil shock should produce less "bang for the buck" now in the industrial countries is that their economies have been less close to overheating than just before the earlier shocks—if rates of inflation and unemployment are any guide. Correspondingly less risk exists that inflationary expectations will be significantly heightened this time.

Of course, the international merchandise trade and current-account balances of the oil-importing countries will shift toward smaller surpluses or larger

Table 1
Measures of Energy Efficiency for Selected Countries, 1973–88
Thousands of Metric Tons of Oil or Oil-Equivalent per Billion Dollars of GDP

	OECD		United States		Japan		West Germany	
Year	Energy-to-GDP	Oil-to-GDP	Energy-to-GDP	Oil-to-GDP	Energy-to-GDP	Oil-to-GDP	Energy-to-GDP	Oil-to-GDP
1973	474	254	587	271	363	274	444	248
1974	465	241	580	261	372	271	430	222
1975	453	234	572	261	345	247	407	211
1976	458	237	577	265	357	258	421	222
1977	452	237	571	275	335	248	407	214
1978	447	233	562	266	324	238	411	215
1979	443	225	550	253	322	229	416	211
1980	424	205	530	230	302	199	393	188
1981	405	187	503	210	282	178	374	167
1982	394	179	497	206	268	165	363	161
1983	385	171	480	197	263	162	358	156
1984	382	166	466	189	268	158	362	154
1985	375	158	452	183	256	141	363	152
1986	367	159	439	184	253	142	357	156
1987	365	156	440	181	244	136	352	149
1988	361	154	436	179	253	143	343	146

Note: Gross domestic product (GDP) was valued at current purchasing power parities, then deflated by the U.S. GDP deflator with 1985 = 100. "Energy" is total energy requirement, and "oil" is total oil requirement.

Source: OECD, International Energy Agency, *Energy Balances of OECD Countries 1987–1988* (Paris: OECD, 1990), pp. 70, 122, 142, 186, 207, and *Energy Balances of OECD Countries 1970–1985* (Paris: OECD, 1987), pp. 10, 244, 334, 532; OECD, *National Accounts 1960–1988*, vol. 1 (Paris: OECD, 1990), pp. 139, 145.

Table 2
Components of World Primary Energy Production, 1973, 1978, and 1988

Component	Percent of Total		
	1973	1978	1988
Natural Gas Liquids	1.7	1.8	2.2
Nuclear	.9	2.3	5.7
Hydroelectric	5.5	6.2	7.0
Dry Natural Gas	17.6	17.9	20.2
Coal	26.4	26.2	27.9
Crude Oil	47.9	45.7	37.6
Total	100.0	100.0	100.0

Note: Detail may not add to totals shown because of rounding.
Source: U.S. Department of Energy, *International Energy Annual 1980*, pp. 6–12, *International Energy Annual 1985*, p. xi, and *International Energy Annual 1988*, p. 5.

deficits with the oil exporters. This result of the oil price increase should not pose a major problem, either, if previous experience is relevant. Following the two earlier, substantially greater oil price shocks, the large current-account surpluses that initially accrued to OPEC were virtually eliminated within a couple of years. In fact, OPEC has run current-account deficits half the time since the first oil shock.

U.S. Dependence on Foreign Oil

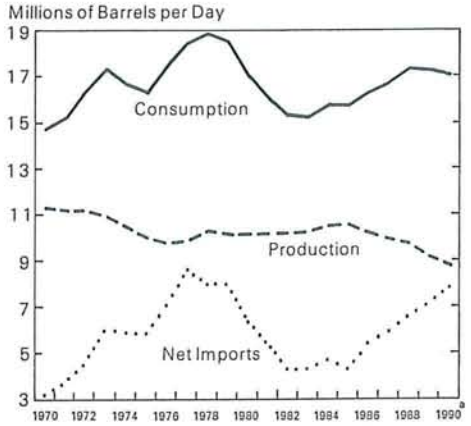
Although the United States did succeed in lowering the amount of oil used per unit of GDP, the nation has sharply increased its net imports of petroleum in recent years (Chart 2). This development was largely a response to the huge 1986 price decline, which occurred after Saudi Arabia stepped up its production in an attempt to enlarge its receipts and/or its market share in the face of relatively weak demand. The plunge in price discouraged U.S. production while enlarging consumption—and enlarging net imports, which have become almost as great as production.

Even so, the U.S. economy is less dependent on imported petroleum than are most "Big Seven" economies. As can be seen in Table 3, Japan and West Germany each imported nearly twice as much petroleum per billion dollars of GDP as the United States did in 1988. Canada and the United Kingdom, being net exporters of petroleum, are easily the best positioned of these countries.

The embargo against trade with Iraq and Kuwait adds a new dimension to this oil shock. As detailed in Table 4, in 1989 the United States carried on a sizable trade with the members of OPEC, including Iraq and Kuwait. The trade was dominated by U.S. imports of petroleum (included in industrial supplies and materials), but also incorporated $4.5 billion in U.S. exports of capital goods, along with substantial U.S. exports of industrial supplies and foodstuffs. Embargo of all commerce with Iraq and Kuwait not only would suspend some $3 billion of U.S. petroleum imports from those countries but would interrupt U.S. exports to them amounting to more than $2 billion (at 1989 rates of trade), including nearly $700 million in foods, feed, and beverages.

Chart 2

U.S. Production, Consumption, and Net Imports of Petroleum Products, 1970-90



Note: All data for 1989 and production data for 1990 are preliminary.
a) For the first 221 days of 1990.
Source: U.S. Energy Information Administration, *Annual Energy Review 1989*, p. 115, and *Weekly Petroleum Status Report*, August 10, 1990, p. 3.

The Position of OPEC

By this time OPEC has become almost a household word. Like many household words, however, its full significance is not widely understood.

OPEC is not a business entity and does not engage in commercial transactions; it is an intergov-

Table 3
Net Imports of Petroleum and Products per Billion Dollars of Gross Domestic Product in Seven Major Industrial Countries, 1988

Country	Thousands of Barrels per Billion Dollars of GDP
Japan	931.3
West Germany	929.4
France	812.3
Italy	772.1
United States	494.3
Canada	−341.4
United Kingdom	−350.6

Note: GDP is measured at 1988 prices and 1988 purchasing power parities. A minus sign signifies net exports.
Source: U.S. Central Intelligence Agency staff; OECD, *National Accounts 1960–88*, vol. 1 (Paris: OECD, 1990), p. 145.

ernmental organization registered with the United Nations Secretariat. It was founded in Baghdad, Iraq, in September 1960, by Iraq, Iran, Kuwait, Saudi Arabia, and Venezuela, countries that resented reductions in oil prices that had occurred during 1959 and 1960. (Unlike the present, at that time oil export prices were set by the international oil companies—which paid taxes to the countries from which they withdrew oil—and during 1959 and 1960 the companies had lowered prices markedly in response to diminishing demand.) Thus the stated principal purpose of OPEC is to coordinate the petroleum policies of the member countries and to determine the best means for safeguarding their interests.

To qualify for membership, a country must have substantial net exports of crude petroleum and must be accepted by three-fourths of the full members, including all of the founding five. By November 1973 the membership had expanded to the current thirteen, including Algeria, Ecuador, Gabon, Indonesia, Libya, Nigeria, Qatar, and the United Arab Emirates, in addition to the initial five. All decisions at OPEC conferences require a unanimous vote.

As can be seen in Table 5, the bulk of the world's oil reserves are controlled by the major reserve holders within OPEC. And OPEC is surely a cartel in the sense that it is a combination of producers designed to limit competition. Its members have striven to coordinate their sales of crude oil so as to influence the price.

While no doubt should exist about these goals, much doubt is warranted about the Organization's success. In particular, OPEC has been riven by disputes over both economic and political issues. Economic disputes have revolved around such matters as price and output levels, and such a dispute was one reason for Iraq's invasion of Kuwait. Political differences had previously led to another war, between Iraq and Iran, which endured for eight years. Still, public conflict is not conclusive proof of cartel failure.

Any attempt to assess the true quantitative impact of OPEC is handicapped by the lack of an empirically verified economic model to explain either

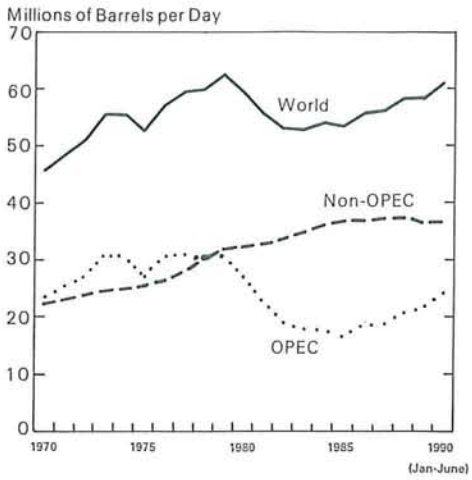
Table 4
U.S. Merchandise Trade with OPEC Countries, by Major End-Use Category, 1989
Millions of Dollars

	OPEC			Iraq and Kuwait		
End-Use Category	Exports	Imports	Balance	Exports	Imports	Balance
Foods, Feeds, Beverages	2,123.8	1,272.6	851.2	675.9	2.7	673.2
Industrial Supplies and Materials	3,197.4	28,073.3	−24,875.9	371.5	3,350.6	−2,979.1
Capital Goods	4,533.1	66.2	4,466.9	363.3	0.3	363.0
Automotive	1,347.5	40.1	1,307.4	403.0	0	403.0
Consumer Goods	1,276.8	971.0	305.8	126.8	1.6	125.2
Special Category (Military-type Goods)	427.6	167.2	260.4	56.1	26.2	29.9
Exports Not Elsewhere Classified and Reexports	286.8	0	286.8	25.5	0	25.5
Total	13,193.0	30,590.4	−17,397.4	2,022.1	3,381.4	−1,359.3

Note: Exports f.a.s.; general imports, f.a.s.
Source: U.S. Bureau of Economic Analysis, machine run.

Chart 3

Production of Crude Oil, 1970-90



Source: U.S. Energy Information Administration, *Annual Energy Review 1989*, p. 249; and *International Petroleum Statistics Report*, August 1990, p. 3.

cartel behavior or changes in the market prices of petroleum. At the least, it seems clear that OPEC has tried to manipulate the price of oil. Inspection of Charts 1 and 3 discloses that OPEC's total crude oil output was reduced somewhat from 1973 to 1974, when the price was soaring and non-OPEC output was increasing, and that OPEC's output diminished further in 1975. Likewise, in 1980, while the second major price rise was still underway, OPEC production began a major decline, and continued to decline for several years while the real price of oil remained at relatively very high levels. By contrast, non-OPEC production grew steadily throughout this period. Finally, as noted at the beginning of this article, the latest oil shock was initiated by a cutback in OPEC production designed to raise the selling price.

Not only has OPEC's output diminished during periods when prices were sharply rising and non-OPEC output was growing, but OPEC's output as a percentage of its capacity has fallen when prices were soaring. These findings suggest that OPEC was restricting its production at least in part to raise the price.

Moreover, these production shifts are not attrib-

Table 5

World Crude Oil and Natural Gas Reserves in the Ten Leading Reserve-Holding Countries, January 1, 1989

Crude Oil			Natural Gas		
Country	Reserves (billion barrels)	Percent of World Total	Country	Reserves (trillion cubic feet)	Percent of World Total
Saudi Arabia[a]	255.0	25.7	U.S.S.R.	1,500.0	38.1
Iraq	100.0	10.1	Iran	494.4	12.6
United Arab Emirates	98.1	9.9	United Arab Emirates	201.5	5.1
Kuwait[a]	94.5	9.5	United States	168.0	4.3
Iran	92.9	9.4	Qatar	156.7	4.0
U.S.S.R.	58.5	5.9	Saudi Arabia[a]	152.0	3.9
Venezuela	58.1	5.9	Algeria	104.2	2.6
Mexico	54.1	5.5	Venezuela	102.2	2.6
United States	26.8	2.7	Canada	95.1	2.4
China	23.6	2.4	Iraq	95.0	2.4
All Others	129.0	13.0	All Others	866.8	22.0
World Total	990.6	100.0	World Total	3,935.9	100.0

[a]Includes half the reserves in the Neutral Zone between Kuwait and Saudi Arabia.

Note: All reserve figures except those for the U.S.S.R. and natural gas reserves in Canada are proved reserves recoverable with present technology and prices. U.S.S.R. figures are "explored reserves," which include proved, probable, and some possible. The Canadian natural gas figure includes proved and some probable. Detail may not add to totals shown because of rounding.

Source: U.S. Energy Information Administration, *Annual Energy Review* 1989 (Washington, D.C.: Government Printing Office, 1990), p. 247.

Table 6
Coefficients of Correlation between Annual Crude Oil Production Levels in Selected Country Groupings, 1973–1990:QI

Country Groupings	Saudi Arabia	Non-OPEC
OPEC		−.85
Non-OPEC	−.69	
OPEC excluding Saudi Arabia	.59	
OPEC excluding Saudi Arabia, Iran, and Iraq	.71	−.79

Source: Underlying data from U.S. Department of Energy, Energy Information Administration, *Annual Energy Review, 1989* (Washington, D.C.: Government Printing Office, 1990), p. 249; and U.S. Central Intelligence Agency, *International Energy Statistical Review*, 31 July 1990, pp. 1, 2.

utable solely to Saudi Arabia, the largest producer, or to the war between Iraq and Iran that broke out in the fall of 1980. On the contrary, Saudi production has been much more positively correlated with production in the rest of OPEC (whether or not Iran and Iraq are included) than with production outside of OPEC (Table 6). In light of such behavior, OPEC may be classified as a "partial market-sharing cartel"—one in which the members usually raise and lower production jointly, although some may make larger percentage changes than others. By this interpretation, OPEC is not a "dominant firm cartel," in which Saudi Arabia (with perhaps a few other members) acts as the residual or swing producer while other members behave competitively.

The market-sharing cartel is the weakest form of cartel arrangement, so its behavior is difficult to predict. In the case of OPEC, prediction has become especially hazardous. With (at this writing) Saudi and Iraqi troops facing each other across the border of Iraqi-occupied Kuwait, not merely OPEC's policies but the future of OPEC itself have become more uncertain.

Some Policy Considerations

If "those who cannot remember the past are condemned to repeat it," what lessons from the past are being employed to reduce the damage from the third oil shock? As already noted, the industrial countries have significantly reduced the quantity of oil and of energy required to produce a unit of output, and the injury from this latest oil shock will be diminished accordingly. This development was not merely the predictable response of energy consumers to the previous oil shocks and the associated oil price increases; governmental taxation and other policies to promote energy conservation also played a role. If one can judge from Table 1, energy conservation policy in the United States has been less rigorous than in other OECD countries, as energy consumption in relation to GDP has consistently been substantially greater in the United States.

Another important action has involved widespread international cooperation. In November 1974, just thirteen months after the beginning of the first oil crisis, sixteen member countries of the OECD signed an agreement on an International Energy Program (IEP), and the OECD created the International Energy Agency (IEA) to oversee the program. This initiative recognized that energy security is a matter of collective security. Twenty-one OECD countries now belong to the IEA and participate in the IEP.

The chief goals of the IEA and the IEP are:

i) co-operation among IEA Participating Countries to reduce excessive dependence on oil through energy conservation, development of alternative energy sources and energy research and development;
ii) an information system on the international oil market as well as consultation with oil companies;
iii) co-operation with oil producing and other oil consuming countries with a view to developing a stable international energy trade as well as the rational management and use of world energy resources in the interest of all countries;
iv) a plan to prepare Participating Countries against the risk of a major disruption of oil supplies and to share available oil in the event of an emergency. (OECD International Energy Agency 1986, p. 2)

To attain the last of these goals, each IEA member is to maintain an emergency stock of oil sufficient to replace net imports for ninety days and is to be prepared to decrease oil consumption in the event of a supply disruption. Moreover, provision is made for the sharing of oil supplies during a crisis. As of mid-1990, total emergency stocks held by OECD countries were on the order of two billion barrels, with readily accessible stocks greater than ninety days' worth of net imports. As an active participant in the stock-building program, the U.S. government has some 587 million barrels of oil stored in its Strategic

Petroleum Reserve, mainly in underground caverns in Texas and Louisiana.

When, and how much, to draw upon these emergency stocks are judgments that should be based on the circumstances. In this latest oil shock, the imminent threat of armed conflict and still higher oil prices has tended to undermine the case for immediately tapping the reserves. Should the war threat ease, however, the case for using the reserves becomes very strong unless the supply of oil from OPEC is substantially maintained. To ameliorate oil shocks was, after all, the purpose for which the reserves were acquired, at considerable cost to the taxpayer. To deny the taxpayer the benefits would be a betrayal of trust. A danger exists that the oil reserves will become like national gold reserves—sacred, and inviolable until the economic damage becomes catastrophic.

OPEC, as well as the oil-importing countries, should have learned a lesson or two from the previous shocks. Large and abrupt reductions in the supply of oil tend to generate major recessions in the industrial countries that reduce the demand for OPEC's major export and lower the rates of return on the sizable investments that OPEC members have placed in the industrial countries. And oil consumers redouble their efforts to develop more reliable and economical sources of energy in the longer run. Surely it is in OPEC's best interest to resist the temptation to hold up its customers.

Summary

Barring a major armed conflict, the oil-importing nations should weather this third oil shock more easily than the previous two. For one thing, the percentage increase in the oil price is likely to be much smaller. For another, the major industrial economies now produce the typical unit of output with much less energy, and oil, than was the case during the 1970s. Also, these economies have not been so close to overheating as they were just before the earlier shocks, so that inflationary expectations should remain lower.

In spite of substantial progress, the United States uses considerably more energy, including oil, to produce a unit of output than do the other OECD countries as a group. Nonetheless, four of the Big Seven countries are more dependent on *foreign* oil than is the United States. The embargo against trade with Iraq and Kuwait will curtail U.S. exports to those countries, exports that amounted to more than $2 billion in 1989.

Although OPEC has tried to manipulate the price of oil, its success in the past over the long term is doubtful, especially in light of the huge price decline in 1986. OPEC is the kind of loose cartel whose price and output policies are exceptionally difficult to predict.

Not only the oil shocks, but also government policies in the industrial countries, have contributed to enhanced energy and oil efficiency. One important additional policy measure is the accumulation of sizable emergency stocks of oil to cope with supply disruptions. The risk exists, however, that the use of these reserves during a disruption will be delayed until an extreme emergency develops—the kind of emergency that, ironically, the reserves were intended to prevent. Another risk is that some influential members of OPEC may not yet have learned that large, abrupt oil price increases are inimical to their own interests.

[1] The 13 members of OPEC are Algeria, Ecuador, Gabon, Indonesia, Iran, Iraq, Kuwait, Libya, Nigeria, Qatar, Saudi Arabia, the United Arab Emirates, and Venezuela.

References

Fieleke, Norman S. 1988. *The International Economy under Stress.* Cambridge, Mass.: Ballinger Publishing Co.

Organisation for Economic Co-operation and Development, International Energy Agency. 1986. *Energy Policies and Programmes of IEA Countries: 1985 Review.* Paris: OECD.

How Does Public Infrastructure Affect Regional Economic Performance?

Alicia H. Munnell with the assistance of Leah M. Cook

Senior Vice President and Director of Research and Research Assistant, respectively, Federal Reserve Bank of Boston. *The author would like to thank colleagues at the Boston Fed for very valuable comments; Richard W. Kopcke deserves particular gratitude. Mary Chamberlain provided valuable programming assistance.*

Bridge collapses and water main explosions focus national attention on the crumbling condition of the nation's infrastructure. Catastrophic infrastructure failures are always a momentary spur to debate on the nation's capital investment policies. But increasingly these negative developments have been accompanied by economists' claims that public capital investment makes a significant contribution to national output, productivity, growth and international competitiveness.

These conclusions, which emerge from the work of Aschauer and others, have generally been based on observed patterns of national and international spending on public capital and various measures of economic performance. Reaction to these claims has been cautious; critics have charged that the empirical work overstates the impact on productivity by ignoring other factors, that the direction of causation between public investment and output growth is unclear, and that even if the historical empirical relationships were estimated correctly, they provide no clear indications for current policy.

This paper is not designed to answer all the criticisms but rather to offer one more brush stroke to the emerging picture of the relationship between public capital investment and private economic activity. It does this by exploring the impact of public capital on output, employment growth, and private investment at the state and regional level. The paper consists of four parts. Since no comprehensive measures of public or private capital are available at the state level, the first section explains the construction of such data and describes the distribution of these wealth measures by state and region. The second section uses these data to estimate an aggregate production function, in order to see whether the positive relationship between output and public capital, which has been documented at the national level, holds up for individual states and regions. The third section moves from the steady state to the adjustment process and explores the relationship between public invest-

ment and private investment, attempting to determine the direction and magnitude of the effect. Finally, the fourth section introduces the public capital data into a firm location model in order to see whether variations in public capital by state have had any impact on state-by-state employment growth.

The conclusion is that those states that have invested more in infrastructure tend to have greater output, more private investment, and more employment growth. This evidence supports results found in earlier studies. The empirical work also seems to indicate that public investment comes before the pickup in economic activity and serves as a base, but much more work is required to spell out the specifics of the link between public capital and economic performance.

I. Public and Private Wealth by State and Region

The U.S. Bureau of Economic Analysis (BEA) publishes annual data from 1925 to the present on the stock of private and public tangible wealth; these data include equipment and structures, but exclude land inventories and rental residential real estate. Despite the availability of public capital data, until recently this kind of input had been virtually ignored in the analysis of national production and growth. The oversight is difficult to explain, since the stock of public capital is not small. As shown in Table 1, in 1988 public capital amounted to almost $2.5 trillion, compared to $4.4 trillion in the private sector. Even ignoring investments devoted to military purposes, the stock of public capital amounted to $2.0 trillion, or 46 percent of the value of the stock of private capital.

Most of the $2.0 trillion of nonmilitary public capital consists of assets owned by state and local governments. Highways and streets account for 39 percent of the total state and local wealth, and water and sewer systems for another 16 percent; buildings (primarily schools and hospitals), other structures and equipment make up the rest (Table 2).

No data are available on the stock of private or public capital on a state-by-state basis. Hence, it was necessary to devise some way of dividing up the national totals published by the BEA. In the case of public capital, the approach taken was to create a state capital series based on annual state public investment data and BEA depreciation and discard schedules, and use this distribution of capital to apportion the BEA public capital totals. In the case of private capital, state investment data (except for manufacturing) were not available, so the approach followed was to apportion the BEA total on the basis of various measures of each state's activity in agriculture, manufacturing and nonmanufacturing (see Appendix A).

The results of this estimation procedure are presented in Table 3, which shows the per capita stocks of public and private capital by region for 1988 and

Table 1
Private and Public Nonresidential Net Capital Stock, 1988

Capital Stock[a]	Billions of Dollars	Percent of Total
Total	6846.4	100
Total Private	4364.8	64
Nonfarm business	4202.3	61
Farm	162.5	2
Total Public	2481.6	36
Military	490.9	7
Nonmilitary	1990.7	29
Federal	272.2	4
State and Local	1718.5	25

Note: Numbers may not add to totals because of rounding.
[a]Figures include equipment and structures only. Land, inventories, and rental residential capital are excluded.
Source: U.S. Bureau of Economic Analysis, unpublished data.

Table 2
State and Local Fixed Nonresidential Net Capital Stock by Type of Asset, 1988

Capital Stock	Billions of Dollars	Percent of Total
Highways and Streets	670.7	39.0
Water and Sewer Systems	265.7	15.5
Buildings and Other Structures		
Schools, Hospitals and Other Buildings	514.2	29.9
Conservation and Development Structures	29.3	1.7
Miscellaneous	126.7	7.4
Equipment	111.8	6.5
Total	1718.5	100.0

Note: Numbers may not add because of rounding.
Source: U.S. Bureau of Economic Analysis, unpublished data.

the ratio of private to public wealth. Table 4 presents information about the growth in public and private capital for the periods 1970–80 and 1980–88. The most striking aspect of the data is that while all regions invested in both private enterprises and public infrastructure during the 1970s, only the South and West continued to add to public capital in the 1980s.

This process of constructing state-by-state capital measures has produced 19 years of data for each of the 48 states in the continental United States; the question is whether it has produced any real information or whether, in effect, it has simply reproduced the relationships between aggregate inputs and outputs many times over. This is a particularly important question given that the procedure for constructing both private and public wealth involved apportioning national totals. Here the nature of the methodology is crucial; if the totals had been distributed to states, say, based on the national ratio of capital to labor, no new information would have been added.

This was not the approach; the share of public capital allocated to each state was based on actual state public investment data and the share of private capital was based on each state's involvement in specific types of economic activity. As a result, the data show significant variation; for example, the ratio of private to public capital, which averaged 2.5 for the nation, ranged in the 1988 state data from a low of 1.5 for New York to a high of 5.1 for Louisiana. Moreover, the rate of growth of public capital varied enormously by state both in the 1970s and particularly in the 1980s. For example, California, the state that ranked twelfth in the ratio of public capital to

Table 3
Stocks of Public and Private Capital by Region, 1988

Region	Public Capital		Private Capital		Ratio of Private to Public Capital
	Per Capita (Dollars)	Percent of Total	Per Capita (Dollars)	Percent of Total	
Northeast					
New England	5,953	4.9	13,748	4.4	2.3
Mid Atlantic	7,193	17.1	13,829	12.9	1.9
North Central					
East North Central	6,205	16.5	15,866	16.6	2.6
West North Central	7,501	8.4	18,455	8.1	2.5
South					
South Atlantic	5,788	15.3	14,520	15.1	2.5
East South Central	6,106	5.9	16,080	6.1	2.6
West South Central	6,330	10.7	25,165[a]	16.8	4.0
West					
Mountain	7,679	6.5	19,603	6.5	2.5
Pacific	6,573	14.8	15,256	13.5	2.3
Continental United States	6,509	100.0	16,551	100.0	2.5
Addendum					
Total Capital[b] (Billions of Dollars)	1,585.5		4,031.4		

[a]The high per capita private capital figure for the West South Central region is the result of a very large share of the nation's manufacturing and mining capital being allocated to Louisiana and Texas. The mining is understandable, since this sector consists largely of oil and gas production. Louisiana and Texas account for almost half of the nation's production of oil and gas, and oil and gas are extremely capital-intensive industries. The manufacturing capital is more difficult to explain, since the shares of manufacturing capital allocated to Louisiana and Texas are almost twice their shares of national value added by manufacturing industries. The main explanation appears to be the high ratio of capital to value added for the specific manufacturing industries located in these states. For example, both Louisiana and Texas are dominated by the petroleum and coal and the rubber and plastics industries; in 1985, these industries had a ratio of capital to value added of 1.37. This number was almost twice the ratio of capital to value added for the average of all the nation's manufacturing industries (.76). To ensure that these high private wealth figures were not distorting the results, separate equations were estimated for the remaining 46 states and the results were virtually unchanged.

[b]These totals differ from those shown in Table 1 for two reasons. First, they do not include Alaska, Hawaii, and District of Columbia. Second, the totals are beginning of year values, whereas the data on Table 1 represent end of year values.

Source: Author's calculations. See Appendix A.

Table 4
Average Annual Rates of Growth in Public and Private Capital by Region, 1970–80, 1980–88

	Public Capital									
	1970–80				1980–88				Private Capital	
Region	Highways	Water & Sewer	Other	Total	Highways	Water & Sewer	Other	Total	1970–80	1980–88
Northeast	1.0	4.8	2.7	2.4	0	1.0	−.2	.1	2.8	2.7
New England	.5	5.2	3.3	2.3	−.2	1.2	0	.1	3.0	4.1
Mid Atlantic	1.2	4.7	2.6	2.4	.1	1.0	−.3	0	2.8	2.2
North Central	1.2	2.3	2.5	1.9	.2	1.3	−.2	.2	3.3	.9
East North Central	1.0	2.1	2.1	1.6	0	1.2	−.5	0	3.3	.8
West North Central	1.6	2.9	3.7	2.5	.6	1.3	.3	.6	3.5	1.1
South	2.2	3.8	3.8	3.1	1.1	2.9	2.0	1.8	3.9	2.8
South Atlantic	2.8	4.4	4.4	3.7	1.0	3.0	2.4	1.9	4.6	3.8
East South Central	1.9	3.3	2.6	2.3	.6	1.1	−.3	.3	4.3	1.8
West South Central	1.6	3.2	3.7	2.6	1.7	3.6	2.9	2.5	3.2	2.3
West	1.2	2.9	1.9	1.8	.5	2.3	1.7	1.4	4.1	3.9
Mountain	1.9	3.1	4.7	3.1	1.9	4.7	4.1	3.2	4.3	2.7
Pacific	.9	2.9	1.3	1.4	−.2	1.5	.9	.7	4.1	4.5
Continental United States	1.5	3.4	2.8	2.3	.6	1.9	.9	.9	3.6	2.5

Source: Author's calculations. See Appendix A.

labor in 1970, had dropped to thirty-fourth place by 1986, and West Virginia, which ranked thirty-fifth in 1970, had risen to seventh place at the end of the period. In short, the individual observations appear to contain real information.

II. The Role of Public Capital in the Production Process

Several studies have examined public capital as an input in the production process. Aschauer (1989) introduced the obvious, but heretofore neglected, notion that the stock of public infrastructure as well as the stock of private capital may be a key to explaining the level of national output in the private sector. His results showed a strong relationship between output per unit of private capital and the stock of public capital; he also found a statistically significant relationship between the level of multifactor productivity and the stock of public capital. Munnell (1990), examining the labor productivity slowdown in the 1970s, found a similarly strong, statistically significant, relationship between the nation's stock of public capital and the level of labor productivity.

Studies at the subnational level have generally been constrained by the lack of wealth data. Nevertheless, several researchers have attempted to relate proxies for public capital to output. For example, Garcia-Mila and McGuire (1987) analyzed the effect of the stock of highways and educational expenditures (representing publicly provided human capital) on statewide production functions, and found that both had a significant positive effect on output.

Eberts (1986) has done similar work on a metropolitan area level. He created annual values of the public capital stock for each of 38 metropolitan areas and introduced them into a translog production function, with value added as output, hours of production and nonproduction workers as labor input, and private manufacturing capital stock as private capital. Eberts found that the public capital stock made a positive and statistically significant contribution to manufacturing output, but that its output elasticity was quite small (0.03).

A few researchers have examined the relationship between the growth, as opposed to the level, of output and public infrastructure; the results have

been mixed. For example, Hulten and Schwab (1984) explored whether the national productivity slowdown could be attributed to a decline in economic efficiency in the Snowbelt relative to the Sunbelt, due to aging infrastructure and a deteriorating capital stock. They disaggregated the growth in manufacturing value added for the nine Census regions into its components, and found that regional variation in output growth was not due to differences in productivity growth but rather to variations in the rate of growth of capital and labor. This evidence appeared to leave no role for variations in public infrastructure in determining regional differences in output growth.[1]

On the other hand, Aschauer (1990) recently completed a paper examining the relationship between income growth and highway capacity using state data. He found that highway capacity and pavement quality had significant positive effects on income growth and that these effects were relatively stable across regions.

The following analysis builds on this earlier work and treats public capital as an input whose services enhance the productivity of both capital and labor. Hence, public capital becomes another input in the production function and the equation looks as follows:

$$1) \qquad Q = (MFP) * f(K,L,G),$$

where Q is output, MFP is the level of technology, K is the private capital stock, L is labor and G is the stock of public capital. Assuming a generalized Cobb-Douglas form of technology yields a more specific relationship between inputs and outputs:

$$2) \qquad Q = MFP * K^aL^bG^c.$$

Translating this equation into logarithms produces a linear function that can be estimated:[2]

$$3) \quad \ln Q = \ln MFP + a\ln K + b\ln L + c\ln G.$$

The coefficients a, b, and c are the output elasticities of the factor inputs. In other words, the coefficients indicate the percentage change in output for a given percentage change in factor input. In production functions without public capital, making some further assumptions about factor markets and the nature of the production function allows the coefficients to be defined more precisely. Specifically, if factor markets are assumed to be perfectly competitive, so that factors are paid their marginal product, and if the production function exhibits constant returns to scale, so that a 10 percent increase in capital and labor leads to a 10 percent increase in output, then the coefficients equal the relative share of total income paid to capital and labor respectively. In the United States the relative shares of national income have been quite stable over many decades, with 35 percent of the total accruing to capital and 65 percent to labor.

While constant returns to scale over the private inputs has been the traditional assumption underlying most analysis of the Cobb-Douglas production function, the inclusion of public capital raises new questions about returns to scale. Given that increasing economies to scale play such an important role in determining the public provision of a good or service, one might be tempted to conclude that public capital in total may yield increasing returns to scale within the production function. Such a leap may be unwarranted, however. While a given highway may yield increasing returns to scale, the construction of an additional highway may not. Moreover, a doubling of the highway system would most certainly produce diminishing returns.

Given the uncertainty of the impact of public capital on returns to scale, several forms of the equation were estimated in addition to the original unconstrained equation. The first assumes that constant returns to scale holds only for the private inputs, but that the entire function shows increasing returns to scale. This assumption is captured by setting $a + b = 1$, so that the equation looks as follows:

$$4) \quad \ln Q = \ln MFP + a(\ln K - \ln L) + \ln L + c\ln G.$$

The alternative is that constant returns to scale applies to the entire production function, so that $a + b + c = 1$. Imposing the second constraint produces the third equation:

$$5) \quad \ln Q = \ln MFP + a(\ln K - \ln L) + \ln L + c(\ln G - \ln L).$$

The equations were estimated using pooled state output, capital and labor data for the period 1970 through 1986, the last year for which gross state product data were available. Labor is measured as total employment on nonagricultural payrolls from the Bureau of Labor Statistics. The public and private capital stocks are the data described in the first section. The unemployment rate is also included to reflect the cyclical nature of productivity. All dollar

Table 5

Regression Results: Output as a Function of Private Capital (K), Labor (L), and Public Capital (G), 48 States, 1970–86

Equation for Output (lnQ)										$\bar{R}^2$	SE	DW
			Private Capital Only									
1) No Constraint:	lnMFP	+	alnK	+	blnL	+	dU%			.992	.092	2.0
	6.75		.36		.69		.006					
	(69.2)		(38.0)		(82.4)		(4.0)					
2) a + b = 1:	lnMFP	+	a(lnK—lnL)	+	lnL	+	dU%			.990	.103	2.1
	7.32		.30		1.0*		−.002					
	(74.2)		(31.9)				(1.0)					
			Including Public Capital									
3) No Constraint:	lnMFP	+	alnK	+	blnL	+	clnG	+	dU%	.993	.088	1.9
	5.75		.31		.59		.15		−.007			
	(39.7)		(30.1)		(43.2)		(9.0)		(4.7)			
4) a + b = 1:	lnMFP	+	a(lnK—lnL)	+	lnL	+	clnG	+	dU%	.992	.090	2.0
	6.33		.34		1.0*		.06		−.007			
	(59.6)		(39.6)				(15.9)		(4.6)			
5) a + b + c = 1:	lnMFP	+	a(lnK—lnL)	+	lnL	+	c(lnG—lnL)	+	dU%	.990	.102	2.0
	6.82		.27		1.0*		.08		−.002			
	(45.8)		(23.3)				(4.4)		(1.0)			

Note: Q = gross state product; MFP = the level of technology; K = private capital stock; L = employment on nonagricultural payrolls; G = stock of state and local public capital; and U% = state unemployment rate; t-statistics in parentheses.

*Constrained to equal 1.

amounts used in the regressions are converted to 1982 dollars.

The regression results, which are summarized in Table 5, confirm, on the state level, that public capital has a significant positive impact on the level of output and does indeed belong in the production function. The first two equations show the estimated production functions without public capital; these equations look very sensible, with coefficients for capital and labor almost exactly in line with their shares of total income. When state and local public capital is added to the equation, it enters with a positive, statistically significant coefficient roughly half the size of that for private capital, and it reduces the standard error of the equation. The coefficient of 0.15 on public capital in equation 3 is noticeably smaller than the 0.35 estimated by Aschauer (1989) and Munnell (1990) in their analysis of national data. The number emerging from the state data implies that a 1 percent increase in public capital would raise national output by 0.15 percent.

The equations also provide some information about returns to scale. The coefficients of the factor inputs sum to 1.05 in the unconstrained equation, implying slightly increasing returns to scale. Constraining the equation either to have constant returns over the private inputs (a + b = 1) or over all inputs, both public and private, (a + b + c = 1) slightly increases the standard error.

Since public capital is an unpaid factor of production, the question arises as to how the benefits accruing from its contribution to output are distributed. It appears that capital and labor each receive a share roughly proportional to their output elasticities. In other words, the unconstrained elasticities for capital and labor in equation 3 are 0.31 and 0.59, respectively; if the 0.15 contribution from output from public capital is divided up proportionately, the result is very close to the traditional 35/65 division of income between capital and labor.

The coefficient of public capital is also sensible in that it implies a reasonable marginal productivity for public capital and equality between the productivity of public and private capital. That is, the elasticity of private sector output with respect to public capital is roughly half that with respect to private capital, and

the state and local public capital stock is approximately one-half the size of the private capital stock. With these proportions, the coefficients imply that a 1 percent increase in either public or private capital will increase output by 0.35 percent.[3] This result is important since the high values implied for the marginal productivity of public capital in Aschauer's results have been the target of criticism (Schultze 1990, p. 63).

Further support for the reasonableness of the results can be gleaned by examining the impact of various components of public capital on output. Table 6 summarizes the regression results with public capital broken into highways and streets, water and sewer systems, and other structures and equipment. Disaggregating in this fashion has almost no impact on the private labor and capital coefficients, yet yields coefficients for the components of public capital in line with expectations. Specifically, the major impact on output from public capital comes from highways and water and sewer systems, while other public capital, which consists primarily of buildings such as schools and hospitals, has virtually no measurable impact on private production.

The lack of effect from schools and hospitals does not mean that government-provided educational and health services have no effect on productivity. One would expect a well-educated and healthy labor force to be more productive than one without such advantages. Rather, the results suggest that the stock of buildings devoted to, say, education may not be the best indicator of the quality of educational services; teachers' salaries, for example, might be a measure. Moreover, even if physical capital were a good measure of service quality, in a highly mobile society the state that provides the educational or health services may not be the one that reaps the benefits.

Finally, separate production functions were estimated for each of the four major regions of the country to see if the relationships were stable across the states (Table 7). The relationship between inputs and outputs appears to vary significantly from one region to another. The question is whether any story can be told that explains the regional variations in the coefficients on labor, private capital and public capital.

One could argue that the large coefficient on labor for the Northeast, which indicates a high percentage change in output for a given percentage change in labor input, reflects the fact that the Northeast has a particularly well-educated, highly skilled labor force. At the same time, the relatively small coefficients on both the private and public capital in the Northeast may, in part, reflect the fact that this region has the lowest capital/labor ratio of any of the four; a relatively smaller amount of capital would imply a relatively smaller coefficient on capital in these equations, assuming the marginal productivity of capital is constant across the country. (These facts imply that the high wages earned by people in the Northeast are due to their intrinsic human capital rather than the amount of physical capital with which they have to work.)

The other surprising result pertains to the production functions for the South. This is the only region where the introduction of public capital significantly alters the coefficients on the private inputs. Once public capital is included in the equation, the coefficient on labor falls from 0.62 to 0.36; moreover, the coefficient on public capital itself is also very large (0.36). No obvious explanation leaps out; the only point that may be worth noting is that the South had

Table 6
Regression Results: Output as a Function of Private Capital (K), Labor (L), and Disaggregated Public Capital (H, WS, O), 48 States, 1970–86

Equation for Output (lnQ)												$\bar{R}^2$	SE	DW
				State-Local Capital										
lnMFP	+	alnk	+	blnL	+	clnH	+	dlnWS	+	elnO	+ fU%			
5.72		.31		.55		.06		.12		.01	−.007	.993	.085	1.9
(42.0)		(28.1)		(35.4)		(3.8)		(9.6)		(.7)	(5.2)			

Note: Q = gross state product; MFP = the level of technology; K = private capital stock; L = employment on nonagricultural payrolls; H = stock of highways; WS = stock of water and sewer systems; O = other state and local public capital, primarily buildings; and U% = state unemployment rate; t-statistics in parentheses.

Table 7

Regression Results: Output as a Function of Private Capital (K), Labor (L), and Public Capital (G), Four Regions, 1970–86

Equation for Output (lnQ)									$\bar{R}^2$	SE	DW
					Private Capital Only						
	lnMFP	+	alnK	+	blnL	+	dU%				
Northeast	9.31		.11		.95		−.01		.997	.068	1.5
	(28.2)		(3.3)		(28.9)		(3.2)				
North Central	6.90		.34		.72		−.003		.998	.048	2.0
	(27.9)		(14.2)		(41.2)		(1.8)				
South	6.03		.42		.62		−.01		.983	.098	1.7
	(31.1)		(22.4)		(30.3)		(4.7)				
West	4.92		.54		.58		−.02		.997	.058	1.7
	(31.6)		(36.9)		(51.4)		(7.9)				
					Including Public Capital						
	lnMFP	+	alnK	+	blnL	+	clnG	+ dU%			
Northeast	8.83		.09		.90		.07	−.01	.997	.067	1.5
	(22.7)		(2.7)		(22.2)		(2.3)	(3.7)			
North Central	5.68		.34		.62		.12	−.004	.998	.046	2.0
	(15.8)		(15.1)		(22.3)		(4.5)	(2.6)			
South	3.15		.38		.36		.36	−.02	.988	.082	1.7
	(10.1)		(22.8)		(12.0)		(10.8)	(6.8)			
West	4.53		.51		.53		.08	−.02	.997	.056	2.0
	(23.4)		(28.0)		(28.7)		(3.2)	(8.4)			

Note: Q = gross state product; MFP = the level of technology; K = private capital stock; L = employment on nonagricultural payrolls; G = stock of state and local public capital; and U% = state unemployment rate; t-statistics in parentheses.

the highest rate of public investment during the 1970s, and was virtually the only region that continued to increase its public capital stock in the 1980s.

In summary, estimates of production functions based on pooled cross-section state data for the period 1970–86 indicate that public capital contributes to private output. The coefficient on public capital implies that its marginal productivity is the same as that for private capital. The benefits of the contribution from public capital seem to be divided between private capital and labor in proportion to the elasticity of private sector output with respect to each input. Moreover, the components of public capital that one would expect to enhance private output—namely, highways and streets, and water and sewer systems—are the ones that have the statistically important relationship; public buildings, such as schools and hospitals, appear to have no direct measurable impact. Finally, the relationship between public capital and output holds up on a regional basis, although more work is needed to explain some of the variation in the coefficients.

III. Public Capital and Private Investment

Another aspect of the role of public capital in the production process is its impact on private investment. In other words, the discussion in this section shifts from documenting a steady-state relationship to exploring the adjustment process. In this process, two opposing forces may be at work. On the one hand, public capital appears to enhance the productivity of private capital, thereby raising the rate of return and encouraging more private sector investment. On the other hand, public capital may serve as a substitute for private capital; to the extent this occurs, more public capital will result in less private investment.

Eberts and Fogarty (1987), in an effort to determine the effectiveness of public infrastructure as a local investment policy, employed the Sims test of "Granger causality" for a sample of 40 metropolitan areas using investment data from 1904 to 1978. They found a statistically significant positive relationship between public outlays and private investment in all

but seven of the 40 cases. In those cities where a relationship existed, public capital investment appeared to influence private investment the majority of the time, but in a substantial number of cases the opposite was true and private investment appeared to precede public investment.

This section explores what can be learned from the state-by-state public and private capital data to supplement the scant existing evidence on the relationship between private investment and public capital. The investigation consists of three parts: the first involves restating the production function estimated earlier to demonstrate the significant positive impact of public capital on the marginal product of private capital; the second involves the estimation of a translog production function where interaction terms can indicate the extent to which public and private capital are complements or substitutes; and the third consists of an effort to estimate an investment function that summarizes the key relationships.

The simple Cobb-Douglas production function used earlier can be rewritten so that the productivity of private capital is the dependent variable. That is,

$$6) \qquad Q/K = MFP * K^{(a-1)}L^{b}G^{c}.$$

Again, translating this equation into logarithms produces a linear function that can be estimated.

$$7) \qquad \ln Q - \ln K = \ln MFP + (a-1)\ln K + b\ln L + c\ln G.$$

The results of estimating this equation are shown in Table 8. Not surprisingly, given that it is simply a rearrangement of the general equation, the relationships are the same as those already described. For the current discussion, the usefulness of the equation in this form is that it highlights the positive, statistically significant relationship between the productivity of private capital and the stock of public capital. Through this mechanism, the stock of public capital would be expected to encourage private investment.

The next step is to determine the nature of the relationship between public and private capital. Are they substitutes or complements in the production process? One way of addressing this issue is to estimate a translog production function; this nonlinear relationship between output and factor inputs includes cross-product terms, which indicate the substitutability or complementarity of the inputs. Variables are entered in the translog function as deviations from their means.

The results of the estimation process are presented in Table 9. The first set of coefficients for private capital, labor, and public capital are similar to those estimated in the simple Cobb-Douglas; as before, public capital has a positive impact on private sector output. The coefficients of the quadratic terms provide an indication of economies of scale for each of the factor inputs. The coefficients indicate slight increasing returns to scale for the private inputs, but constant returns to scale for public capital.

Information on substitutability or complementarity is provided by the coefficients of the cross-product terms. These estimates show a strong substitutability between private capital and labor, as expected, and a somewhat weaker degree of substitution between private capital and public capital. Labor and public capital appear to be complements, although the relationship is not statistically significant.

In an effort to gain more information about the nature of the substitutability between private and public capital, another translog production function was estimated with public capital disaggregated into highways and streets, water and sewer facilities, and other public capital. As before, the results indicate

Table 8
Regression Results: Productivity of Private Capital as a Function of Private Capital (K), Labor (L), and Public Capital (G), 48 States, 1970–86

Equation for Private Capital Productivity (lnQ-lnK)										$\bar{R}^2$	SE	DW
lnMFP	+	(a − 1) lnK	+	blnL	+	clnG	+	dU%				
5.75		−.69		.59		.15		−.007		.91	.088	1.9
(39.7)		(67.2)		(43.2)		(9.0)		(4.7)				

Note: Q = gross state product; MFP = the level of technology; K = private capital stock; L = employment on nonagricultural payrolls; G = stock of state and local public capital; and U% = state unemployment rate; t-statistics in parentheses.

Table 9
Regression Results: Translog Production Function, 48 States, 1970–86

Equations for Output (lnQ): Including Aggregate Public Capital	Coefficient (t-Statistic)	Disaggregating Public Capital	Coefficient (t-Statistic)
$\ln K - \ln \bar{K}$	.22	$\ln K - \ln \bar{K}$	.21
	(18.9)		(16.1)
$\ln L - \ln \bar{L}$	.69	$\ln L - \ln \bar{L}$	.67
	(37.5)		(35.7)
$\ln G - \ln \bar{G}$	.16	$\ln H - \ln \bar{H}$	.04
	(9.1)		(2.7)
		$\ln WS - \ln \bar{WS}$	.15
			(10.9)
		$\ln O - \ln \bar{O}$	−.02
			(1.1)
$(\ln K - \ln \bar{K})^2$	.27	$(\ln K - \ln \bar{K})^2$	.27
	(11.7)		(10.3)
$(\ln L - \ln \bar{L})^2$	.13	$(\ln L - \ln \bar{L})^2$	.17
	(3.2)		(3.1)
$(\ln G - \ln \bar{G})^2$	.03	$(1n H - \ln \bar{H})^2$	.02
	(0.5)		(0.3)
		$(\ln WS - \ln \bar{WS})^2$	.01
			(0.4)
		$(\ln O - \ln \bar{O})^2$	.09
			(3.9)
$(\ln K - \ln \bar{K})(\ln L - \ln \bar{L})$	−.39	$(\ln K - \ln \bar{K})(\ln L - \ln \bar{L})$	−.35
	(9.8)		(7.9)
$(\ln K - \ln \bar{K})(\ln G - \ln \bar{G})$	−.14	$(\ln K - \ln \bar{K})(\ln H - \ln \bar{H})$	−.10
	(2.1)		(1.6)
$(\ln L - \ln \bar{L})(\ln G - \ln \bar{G})$	.12	$(\ln K - \ln \bar{K})(\ln WS - \ln \bar{WS})$	.08
	(1.4)		(2.1)
		$(\ln K - \ln \bar{K})(\ln O - \ln \bar{O})$	−.20
			(4.4)
		$(\ln L - \ln \bar{L})(\ln H - \ln \bar{H})$	.11
			(2.0)
		$(\ln L - \ln \bar{L})(\ln WS - \ln \bar{WS})$	−.05
			(0.6)
		$(\ln L - \ln \bar{L})(\ln O - \ln \bar{O})$	−.04
			(0.8)
U%	−.006	U%	−.006
	(4.7)		(5.2)
intercept	11.0	intercept	11.0
	(1190.3)		(1168.1)
$\bar{R}^2$	.995	$\bar{R}^2$	.996
DW	1.7	DW	1.7

Note: Q = gross state product; K = private capital stock; L = employment on nonagricultural payrolls; G = stock of state and local public capital; H = stock of highways; WS = stock of water and sewer systems; O = other state and local capital, primarily buildings; and U% = state unemployment rate; t-statistics in parentheses.

that most of the impact of public capital on private production comes from water and sewer systems and, to a lesser extent, from highways; other public capital has no measurable impact. As in the equation with aggregate public capital, the quadratic terms indicate that none of the components of public capital exhibit increasing or decreasing returns to scale.

The coefficients of the cross-product terms of private capital and the components of public capital are completely in line with one's intuition. Highways and streets appear to be substitutes for private capital; this seems quite reasonable in that smooth, well-maintained roads will reduce the wear and tear on commercial vehicles. Moreover, private employers or developers may sometimes be required to build their own access roads. Water and sewer facilities are strong complements to private capital; these inputs are generally publicly provided and clearly augment private production. On the other hand, other public capital is a direct substitute. As noted before, this residual consists primarily of hospitals and schools, both of which have private sector counterparts; it also consists of power plants, which are definitely part of the private sector in some states.

Thus, public capital, as hypothesized, has the potential for either encouraging or discouraging private sector investment. One attempt was made to combine these two influences into the simplest possible model of investment. Specifically, the production function indicates that the desired stock of capital ($\bar{K}$) is positively related to the level of output (Q), the supply of labor (L) and the stock of public capital (G). At the same time, the desired stock is positively related to the marginal productivity of capital (MPK) relative to the cost of capital. Assuming the cost of capital is constant, the desired stock can be expressed as

$$8) \qquad \bar{K} = f(Q, L, G, MPK).$$

The simple Cobb-Douglas production function suggests that the marginal product of capital can be expressed as a function of the logarithms of private capital, labor, and public capital:

$$9) \quad MPK = \ln MFP + (a - 1)\ln K + b\ln L + c\ln G$$

This means that

$$10) \quad \bar{K} = \ln MFP + (a - 1)\ln K + b\ln L + c\ln G + dQ + eL + fG.$$

A stock adjustment approach was taken, whereby investment in a given year partially closes the gap between the desired and the existing stock of capital; that is,

$$11) \qquad K_t - K_{t-1} = \alpha(\overline{K} - K_{t-1}).$$

Introducing the described specification of the desired capital stock into the stock adjustment model yields

$$12) \quad K_t - K_{t-1} = \alpha(\ln MFP + (a-1)\ln K + b\ln L + c\ln G + dQ + eL + fG - K_{t-1}).$$

The results of estimating this equation are shown in Table 10.[4] (In addition to the traditional coefficients and t-statistics, Table 10 includes beta coefficients; these coefficients, which standardize for the magnitude of the individual variables, provide a better indication of the relative importance of the various factors in explaining employment growth.) The signs of the coefficients on public capital are as predicted. As one of the variables that determines the marginal productivity of private capital, public capital enters the equation with a positive coefficient. (Unfortunately, the signs on the other variables representing the marginal productivity of capital are reversed; the logarithm of private capital should be negative and the log of labor, positive.) Thus, public capital appears to stimulate private investment through its influence on the productivity of private capital. On the other hand, the stock of public capital has a negative, statistically significant effect on private investment. Given that private and public capital are substitutes, an increase in the stock of public capital, all else equal, will reduce the required level of private capital and private investment.

It may be pushing these results too far, but it is hard to resist estimating the net effect of public capital on private investment. On the one hand, a 0.1 increase in the log of public capital implies a $96 billion increase in private investment. In dollar terms, 0.1 increase in the log is roughly equivalent to a 10 percent increase in the public capital stock, or $172 billion. From these numbers, $1 of additional public capital appears to increase private investment by 56 cents. On the other hand, the coefficient on last period's capital stock indicates that an additional $1 of public capital reduces private investment by 11 cents in that year (more in subsequent years). On balance, the equation suggests that each additional dollar of public capital appears to increase private investment by 45 cents.

Table 10
Regression Results: Investment as a Function of the Marginal Productivity of Capital (MPK), Output (Q), Private Capital (K), Labor (L) and Public Capital (G), 48 States, 1975–86

Equation for $K_t - K_{t-1}$	Coefficient (t-Statistic)	Beta
Marginal Productivity of Capital		
lnK	199.7	.05
	(0.4)	
lnL	−853.1	−.23
	(1.2)	
lnG	959.9	.24
	(1.0)	
G	−.11	−.81
	(3.8)	
L	−861.6	−.44
	(1.1)	
$\sum_{i=0}^{4} Q$	.09	1.97
	(3.7)	
K_{t-1}	−.02	−.30
	(2.1)	
intercept	−10,641.0	
	(1.4)	
$\overline{R}^2$	.46	
DW	2.2	

Note: Q = gross state product; MFP = the level of technology; K = private capital stock; L = employment on nonagricultural payrolls; G = stock of state and local public capital; t-statistics in parentheses.

The simple investment equation, however, can certainly be improved, so the results should be interpreted only as an invitation for future researchers to pursue this topic. The more robust results in the investment area are 1) public capital positively affects the marginal productivity of private capital, and 2) public capital and private capital in the aggregate are substitutes. A careful estimation of the net effect of these two forces remains to be done.

IV. Infrastructure and Firm Location

The third strand in the literature pertaining to infrastructure and economic activity focuses on the relationship between public capital and new business formation or employment growth. For, after all, to

demonstrate a systematic relationship between public capital, output, and investment is only the first step; the challenge is to describe the mechanism through which public capital enters into the process.

Infrastructure could influence the location decisions of both firms and households. For example, high-quality roads, sewer systems, schools and hospitals would be expected to encourage people to move to a given area; similarly, firms requiring large amounts of water in their production process, such as fabric dyeing, would be attracted, all else equal, to areas with water supply facilities that can meet their needs.

Although an enormous literature explores the factors entering the firm location decision, relatively little work has been done focusing on the role of infrastructure in that process.[5] A notable exception is a recent study by Eberts (1989), who examined the relationship between changes in metropolitan area capital stock and firm openings. He found statistically significant positive effects in the case of small businesses, with lesser impact on large firms. He also looked at changes in the public capital stock, but did not find a significant relationship between public investment and openings.

This section uses the state-by-state public capital data to see whether public infrastructure is important in explaining state variations in private economic development. At the state level, the best indicators of economic development and growth are employment trends; hence, the empirical work examines the relationship between employment growth and public capital within the context of a firm location model.

The theoretical literature and empirical studies of firm location are heavily oriented toward the locational decisions of individual manufacturing firms. The theory assumes that firms want to maximize their after-tax profit, so the location decision is driven by the firm's profitability at alternative locations. Profits depend on the difference between sales and the costs of production. Sales, in turn, depend on the nature of the market. For a company making intermediate products, useful data include the number and size of potential purchasers of the intermediate product and the number and size of competitors. If the firm produces for the consumer market, then the number

Table 11
Regional Data on Employment Growth (1970–80 and 1980–88) and Its Potential Determinants, 1970 and 1980

	Average Annual Rate of Private Employment Growth		Unemployment Rate		College Graduates		Urban Population		Tax Burden		Population Density[a]		Hourly Wage (Manufacturing)		Cost of Energy[b] (Per Million BTUs)	
Region	1970–80	1980–88	1970	1980	1970	1980	1970	1980	1970	1980	1970	1980	1970	1980	1970	1980
	Percent												1982 Dollars			
Northeast	.8	1.9	4.6	7.1	11.2	17.3	89.2	88.1	11.3	11.5	301	302	8.38	8.33	3.05	4.30
New England	1.9	2.6	4.9	5.9	12.2	19.3	82.9	81.2	10.5	10.4	189	196	7.92	7.61	3.81	4.52
Mid Atlantic	.5	1.7	4.5	7.5	10.9	16.6	91.2	90.5	11.6	11.8	372	369	8.53	8.60	2.91	4.26
North Central	1.7	1.3	4.7	8.2	9.6	14.8	71.5	70.5	10.3	9.6	75	78	9.20	9.66	2.96	3.91
East North Central	1.3	1.2	5.1	9.2	9.5	14.5	78.7	77.2	10.3	9.6	165	171	9.45	9.99	2.85	3.91
West North Central	2.7	1.5	3.8	5.7	9.9	15.4	53.8	54.0	10.5	9.7	32	34	8.51	8.85	3.35	3.93
South	3.7	2.6	4.5	6.4	9.7	15.0	66.8	67.8	9.3	8.7	71	86	7.26	7.65	1.86	4.20
South Atlantic	3.4	3.7	4.2	6.3	10.3	15.5	71.1	71.7	9.4	8.9	113	136	7.03	7.21	2.65	3.47
East South Central	2.9	2.1	4.8	7.9	7.7	12.1	53.5	53.4	9.4	8.7	72	82	7.08	7.48	2.08	3.76
West South Central	4.8	1.0	4.8	5.6	10.1	15.7	68.9	70.8	9.1	8.5	45	56	7.77	8.39	1.44	4.67
West	4.4	2.7	6.8	6.8	13.2	19.3	83.9	83.1	11.4	10.0	29	36	9.28	9.16	2.10	4.07
Mountain	5.9	2.6	5.1	6.2	12.9	18.9	60.7	62.4	10.8	10.1	10	13	8.42	8.60	2.22	3.32
Pacific	4.0	2.8	7.3	7.0	13.2	19.4	91.5	90.8	11.5	9.9	80	95	9.53	9.36	2.05	4.51

Note: See Appendix B for details on sources of data.
[a]Measured as number of persons per square mile of land area.
[b]Measured as the ratio of expenditures on fuel and purchased electricity to consumption of fuel and purchased electricity, for the industrial sector.

Table 12
Regression Results: The Role of Public Capital in Private Employment Growth, 1970–88, 1970–80, and 1980–88

	Employment Growth							
	1970–88 (1970 Levels)		1970–80 (1970 Levels)		1980–88 (1980 Levels)		1980–88 Growth (Based on 1970–80 Changes)	
Explanatory Variable	Coefficient (t-Statistic)	Beta	Coefficient (t-Statistic)	Beta	Coefficient (t-Statistic)	Beta	Coefficient (t-Statistic)	Beta
Cost of Labor								
WAGE	−1.4	−.52	−.8	−.20	−1.0	−.70	−.1	−.44
	(4.1)		(1.6)		(4.4)		(3.6)	
U%	.4	.39	.4	.28	.3	.36	.2	.20
	(3.3)		(2.3)		(2.2)		(1.4)	
COLL	.3	.46	.3	.33	.2	.39	.1	.09
	(3.8)		(2.7)		(2.5)		(.7)	
Cost of Land								
POP DENSITY	−.003	−.64	−.003	−.41	−.002	−.24	.06	.41
	(5.0)		(3.2)		(1.3)		(3.2)	
Cost of Energy								
ENERGY	2.8	.56	1.8	.24	−.1	−.05	−.003	−.10
	(4.2)		(1.7)		(.3)		(.7)	
TEMP	.08	.34	.1	.38	−.008	−.03		
	(3.0)		(3.4)		(.2)			
Potential Sales								
URBAN	.01	.31	−.006	−.09	.03	.50	−.01	−.01
	(2.0)		(.6)		(2.9)		(.1)	
TAXES	−.3	−.32	−.3	−.24	−.4	−.30	−.4	−.22
	(2.6)		(1.9)		(2.0)		(1.7)	
PUBLIC CAPITAL	.0001	.35	.0002	.45	.0002	.18	.03	.24
	(2.7)		(3.4)		(1.0)		(1.7)	
INTERCEPT	−5.0		−10.1		5.3		−.02	
	(1.7)		(2.2)		(1.1)		(.01)	
$\bar{R}^2$	.63		.62		.41		.45	
DW	1.9		2.1		1.9		1.8	

Note: For description of variables, see Appendix B.

and income of potential customers at each location would be relevant. On the cost side, the most important factors are probably wages and the skill of the labor force, although land and energy costs are also relevant.

The equations estimated here include variables to capture both revenue and cost components of profitability. The specific form of the equation is based on the disequilibrium adjustment model, which is commonly used in cross-sectional studies of regional economic growth. In this model, the change in the dependent variable, in this case private nonagricultural employment, is related to levels of the explanatory variables at the beginning of the period. For example, the growth in employment between 1980 and 1988 will be related to revenue and cost measures in 1980.

Three explanatory variables are included in the equations to represent the labor market: the average hourly wage in manufacturing (WAGE), the state unemployment rate (U%), and the percent of the state's population with at least four years of college (COLL). Two additional variables are designed to measure energy costs: the cost per million BTUs of purchased fuels and electricity (ENERGY) and the normal daily maximum temperature in July (TEMP). Finally, population density (POP DENSITY) is included to capture the cost of land. On the sales side,

the percent of the population residing in metropolitan areas (URBAN) was introduced to reflect the potential market. Since both firms and individuals are interested in after-tax income, a variable was included measuring total state and local taxes as a percent of state personal income (TAXES). Finally, the stock of public infrastructure (PUBLIC CAPITAL) was introduced to determine whether it had an independent direct effect, once these other economic determinants were taken into account. The regional values for most of these variables are summarized in Table 11, and the public capital data are shown in Table 3.

The regression results, which are shown in Table 12, are quite interesting and suggest that infrastructure does contribute towards a state's employment growth. Some general comments are required, however, before exploring the results in more detail. First, unlike the production function equations reported earlier, where the variables to be included are fairly well defined, the list of potential variables to explain state-by-state employment growth is limitless. For example, to estimate the effect of taxes on the growth in employment, one study employed five separate tax measures (Plaut and Pluta 1983). The goal of the exercise described below was to include only those independent variables whose presence would be viewed as essential by most observers.

Second, no matter how disciplined an investigator attempts to be, the temptation to try a number of different combinations or alternative measures is sometimes overwhelming. Since this part of the study involved some "fishing," the most useful way to proceed is to make all results available to the interested reader, report those that seem most persuasive, and then indicate what was learned from the process. One source of comfort is the fact that, while its statistical significance varies, the magnitude of the coefficient for public infrastructure remains virtually unchanged regardless of what modifications are made to the rest of the equation.

The first three equations in Table 12 are similar in approach; they vary only in the period spanned or the initial conditions. That is, the first equation explains employment growth over the 1970–88 period using 1970 values for wages, state unemployment rates, and so on; the second shortens the period of employment growth to 1970–80 but maintains the 1970 level for the independent variables; the third equation looks at employment growth over the 1980–88 period using 1980 levels of the independent variables. The fourth equation takes a somewhat different approach in that it attempts to explain employment growth for the 1980–88 period on the basis of what happened to the independent variables during the period 1970–80. For example, the independent variable becomes the change in the state's hourly wage level from 1970 to 1980 instead of the level of the wage in 1980.

The results are generally in line with what one would expect. The cost, availability and quality of labor in a given state appear to play a central role in that state's employment growth; the lower the wage level, the greater the level of unemployment, and the more highly educated the work force in the base period, the greater the growth in employment during the subsequent period. Similarly, to the extent that population density serves as an indication of the cost of land, the results show that states with relatively plentiful, inexpensive land in the initial periods experienced the higher rates of growth in the subsequent periods.

The results for energy costs are somewhat less consistent. The original notion was that higher energy costs, all else equal, would reduce profitability and therefore discourage the establishment of new firms and inhibit employment growth. The data support this hypothesis in two respects. First, all else equal, states with warmer climates tend to have greater employment growth. Second, energy costs have a negative effect on employment growth in the 1980–88 period.

The regression results suggest that infrastructure does contribute towards a state's employment growth.

The inconsistency arises in that energy costs appear to have been positively associated with employment growth over the entire 1970–88 period and during the 1970s. Although this result means that the variable is not playing its intended role, the perverse relationship is understandable. The major oil and gas producing states—Texas, Oklahoma, and Louisiana—began the 1970s with energy costs far below the national average. These states then enjoyed among the highest levels of employment growth from 1970 to 1980 as OPEC created a dramatic runup in energy

prices. Awash in money and easy access to energy, these states increased their consumption of energy and had the highest energy costs in the nation by 1980. The collapse of energy prices in the beginning of 1980s, however, meant that employment growth virtually ceased during the 1980–88 period. This boom/bust phenomenon probably explains the performance of the energy cost variable far more than its role as a factor of production.

The two remaining variables look fairly sensible. The percent of the population living in urban cities has a positive effect on employment growth, except during the 1970s when the energy phenomenon dominates. The tax burden, measured simply as the ratio of total taxes to personal income, has a consistently negative, statistically significant effect on employment growth. This finding is something of a coup, since researchers have traditionally gone to great lengths to find a relationship between taxes and economic development; they have frequently constructed complex measures of tax effects, and have just as frequently been unsuccessful.

Public infrastructure matters in firm locational decisions and thereby affects employment growth.

The purpose of constructing this whole model, however, was to determine whether the amount of public infrastructure has a direct measurable effect on employment growth. One would expect this to be the case; a state that goes to the trouble of building roads, sewers, water supply facilities, and power plants, as well as schools and hospitals, would be expected to attract more new firms and more households than a state that did not undertake such activity. Remember, this refers to the level of public capital for a *given* level of taxes, wages, land costs and other factors. The results are consistent with the notion that public capital contributes to economic growth; the coefficient of public capital is positive and relatively consistent for the entire period and the two subperiods. These numbers imply that $1,000 more of public infrastructure per capita in the initial period contributes roughly 0.2 percent to the average annual rate of employment growth.

One might wonder how much weight to put on these results. As indicated above, several regressions were run, adding and deleting variables for unionization and personal income and substituting heating degree days for the maximum temperature variable. No matter which variables were included in the regression equation, the coefficient for public capital never fell below 0.0001 or rose above 0.0003 for any of the time periods. In terms of the statistical significance, the t-statistics never fell below 1.2 for the subperiods or rose above 4.1. The reader must come to her or his own conclusion, but the author is convinced that public infrastructure matters in firm locational decisions and thereby affects employment growth.

Before leaving this topic, one further equation was estimated. It may be a little unorthodox, but it is based on the notion that investment and employment decisions are less related to the initial levels of the relevant variables than to how these variables have been changing in the recent past. The results of testing this hunch empirically are summarized in the last equation of Table 12. As noted earlier, this equation relates the growth in employment for the period 1980–88 to the changes in the variables over the period 1970–80. The $\overline{R}^2$ indicates that this approach explains more of the variation in state employment growth than including the initial levels. Almost all the variables have the expected sign and magnitude (except for population density[6]), and the growth of public capital appears to be considerably more important in this equation than its initial level was in the earlier equations. This should be interpreted as nothing more than one additional bit of evidence that public capital affects state-by-state levels of economic activity.

V. Conclusion

This paper consisted of three exercises exploring the relationship between public capital and economic activity. The first looked at the role of productivity in the production process and found that public capital had a positive, statistically significant impact on private sector output. These results were robust. The coefficient on public capital implied the same marginal productivity as for private capital. The benefits from public capital, an unpaid factor of production, seem to be divided between private capital and labor in proportion to the elasticity of private sector output with respect to each input. When public capital was

disaggregated into highways and streets, water and sewer systems, and other structures and equipment, the coefficient of each component was in line with expectations. Finally, the relationship between public capital and output held up on a regional basis, although more work is needed to understand the variation in the coefficients.

The second exercise involved investigating the role of public capital in private sector investment. Here two opposing forces were at work. On the one hand, the evidence clearly indicated that public capital enhances the productivity of private capital; through this mechanism public capital would be expected to stimulate private sector investment. On the other hand, the results of a translog production function indicated the bulk of state and local public capital is a substitute for private capital; this substitutability indicates that, for any given level of output, the more public capital on hand the less private investment required. A simple investment equation suggested that both these effects were evident, but these results were not robust and much more work should be done.

Public capital affects state-by-state levels of economic activity.

The third exercise explored the relationship between public capital and employment growth in order to see whether the stock of a state's physical infrastructure influenced firm location and subsequent growth. Although the specific model into which public capital should be introduced is much less precise than that specified by a production function, the empirical work provided convincing evidence, at least to the author, that a state's investment in public capital had a significant positive impact on that state's private employment growth.

The evidence seems overwhelming that public capital has a positive impact on private sector output, investment, and employment. But public capital is not just another form of private capital. These physical resources were produced by the public sector because they contribute additional benefits that cannot be captured by a private sector investor; the presumption is that inadequate quantities would have been produced if left to private sector initiatives. The fact that public capital has these externalities and that the marginal productivities of public and private capital appear to be the same in the private production process suggest that the United States has underinvested in public capital. But one does not really need equations to arrive at that conclusion.

The conclusion that this country has underinvested in public capital and that public capital has a positive impact on economic activity does not mean that the United States should blindly double the amount of money it spends on public capital; nor does it mean that careful cost-benefit analyses are no longer needed for individual projects. Rather the results indicate that more spending on public investment, which is clearly needed to remedy serious safety hazards and to improve the quality of life, may also produce greater productivity and growth.

Appendix A—Creation of State Estimates of Capital Stocks

No state-by-state data are available on the stock of public or private capital. Hence, it was necessary to devise ways of dividing up the national totals published by the Bureau of Economic Analysis (BEA). The capital stock series selected were the constant-cost or "physical-volume" estimates, where assets are valued at a base-year price. In the case of public capital, the approach taken was to create for each year, 1969 to 1988, a state capital stock series based on annual state investment data and BEA discard and depreciation schedules, and use the state-by-state distribution of these series to apportion the BEA public capital totals for the nation. In the case of private capital, state investment data (other than for manufacturing) were not available, so the approach followed was to apportion the BEA national total for private capital on the basis of various measures of each state's activity in the agricultural sector, the manufacturing sector, and the nonfarm, nonmanufacturing sector. These calculations are described below.

Public Capital Stocks

An estimate of public capital stock was made for each state, and each state's share of the sum of these estimates was used to apportion the BEA national estimate of state and local public capital. The capital outlay data used as a basis for the state estimates of stock were taken from *Governmental Finances*, a U.S. Bureau of the Census publication, for the years 1958 to 1988. Capital outlay was defined as direct expenditure for the construction of buildings, roads, and other improvements, including additions, replacements, and major alterations to fixed works and structures, whether contracted privately or built directly by the government. Purchases of equipment, land, and existing structures were also classified as capital outlays. (Repair expenditures, classified under current operations expenditure, were not included here.)

Governmental Finances lists, state by state, the capital outlays for certain functions as well as total capital outlays. Some functions were not reported separately for the full time period, so it was not possible to estimate stock measures for all types of capital. Consistent series were available for highways, sewerage, and water supply facilities. (Data on capital outlays on water supply facilities were not available separately from 1958 to 1960, but as this is only a brief period and because water supply facilities are an important piece of "core" infrastructure, the stocks were estimated based on data from 1961 to 1988.)

The BEA procedure outlined in *Fixed Reproducible Tangible Wealth, 1929–1985* was followed in order to calculate public capital stock estimates for 1969 to 1988. The first step in this process was to deflate annual data on nominal dollar investment in each state into constant dollar investment, with the same deflators used by the BEA in its calculations of national public capital stocks. Obtaining an estimate for the gross capital stock required calculating the value of each year's investment that would have been discarded over the years. Assets are not always discarded at the end of the average service life, but rather some assets are discarded earlier and others remain in service longer. The retirement pattern used by the BEA to calculate gross stocks is a modified Winfrey S-3, with retirements starting at 45 percent of the average service life and ending at 155 percent of average life. The service lives used here were again taken from the BEA. Highways, sewer systems, and water supply facilities were assumed to last 60 years, thus this figure was used in the discard and depreciation calculations for these assets. The average service life for total public capital had to be estimated and was calculated as a weighted average of the service lives of its components, with the weights representing the component's percent of total constant dollar investment over the full period, according to the following formula:

$$\left(15 \text{ yrs.} * \frac{\sum \text{State \& local equipment investment}}{\sum \text{Total state \& local investment}}\right)$$

$$+ \left(50 \text{ yrs.} * \frac{\sum \text{State \& local investment in buildings, "other" structures}}{\sum \text{Total state \& local investment}}\right)$$

$$+ \left(60 \text{ yrs.} * \frac{\sum \text{State \& local investment in highways, water supply facilities, sewer systems, and conservation and development structures}}{\sum \text{Total state \& local investment}}\right)$$

$$= 50.68 \text{ years.}$$

This calculation was based on BEA investment data. The value of discards was then subtracted from the annual real investments. Summing these investment figures over time gave the gross value of the capital stock. These estimates were then summed across states, with each state's share of this sum used to apportion the BEA national estimate of state and local gross public capital stock.

A similar procedure was used to derive net capital stock estimates. The value in the end year (that is, the year for which the stock is being estimated) of total depreciation on each year's original investment was calculated. The BEA assumption of straight-line depreciation over the average service life of the asset was used. (Service life estimates were the same as above.) Subtracting depreciation from the original annual investments left the net value in the end year of each year's investment. These values were summed to obtain the net value of the capital stock in that year. The stock estimates were then summed across states. Each state's share of this total stock was then used to apportion the BEA national total amount of state and local public capital stock for that year. Net capital stock estimates were used in estimating the production function; they better reflect the productive capacity of the stock because they are adjusted for wear and tear, accidental damage, and obsolescence.

The sum of estimates across states equaled approximately 75 percent of the BEA total state and local net stock measure in 1970. By 1980 the state stock estimates created here summed to 97 percent of the BEA total. The sum of

state estimates in 1986 was 108 percent of the BEA total. This number exceeds the BEA total because of coverage and timing differences between Census expenditure data and the NIPA data on state and local expenditures used by the BEA.

Because public assets have long lives and investment data begin only in 1958, the stock estimates in the earlier years have the potential to underestimate stocks in the older parts of the country, where much investment may have occurred prior to 1958. Similarly, it may overestimate capital stocks in the newer areas of the country. Looking at the results of the procedure, the bias does not seem too pronounced, since older industrial states like New York, Illinois, Pennsylvania, Ohio and Michigan are all ranked in the top ten in terms of total public capital stock in 1969. While these estimates could undoubtedly be improved by collecting data over a longer time period, given the complete dearth of information on public capital stocks at the state level, and the limitations of consistent, currently available data, they represent a reasonable first attempt.

Private Capital Stocks

Private capital stocks were calculated by apportioning BEA national stock estimates of various sectors among the states, using a procedure similar to the one outlined in Costa, Ellson, and Martin (1987). This approach was adopted because investment data by state are available only for the manufacturing sector, while the production function is to be estimated for the state economy as a whole. Thus data limitations prevented using the perpetual inventory method to calculate private capital stocks. The private capital stock in a state is given by the following formula:

$$K_i = (AGK_i / \sum AGK_i)AGK + (MFGK_i / \sum MFGK_i)MFGK + (NFNMFGK_i / \sum NFNMFGK_i)NFNMFGK$$

where: AGK = BEA constant-cost value of capital stock in the agricultural sector
MFGK = BEA constant-cost value of capital stock in the manufacturing sector
NFNMFGK = BEA constant-cost value of capital stock in the nonfarm, nonmanufacturing sector
AGK_i = proxy for capital stock in agriculture in state i
$MFGK_i$ = proxy for capital stock in manufacturing in state i
$NFNMFGK_i$ = proxy for capital stock in the nonfarm, nonmanufacturing sector in state i.

Much of the data used as proxies was taken from the economic censuses, which occur every fifth year: agriculture, manufacturing, and several nonfarm, nonmanufacturing sectors: construction, mining, services, and retail and wholesale trade. Several nonfarm, nonmanufacturing sectors were apportioned using data from sources other than the economic censuses: rail, air and water transportation, trucking, electric and gas services, telephone, and banking. A state's share of the proxy in the census year was used to distribute BEA assets for that year, preceding years and following years. Thus, data from the 1972 Census were used to apportion among the states the BEA national stock estimates for 1969 to 1974; 1977 shares were used for the 1975 to 1979 stock estimates; 1982 shares were the basis for the estimates from 1980 to 1984; and 1987 data were used to apportion national asset totals for 1985 and 1986. (In cases where data were not available for the census year, data for the closest year were used or another estimating procedure was employed. These exceptions are described below.)

The BEA estimate of capital in agriculture was distributed among states based on the value of land, buildings, and equipment in agriculture. The value of land, buildings, and equipment taken from the *1987 Census of Agriculture* was used as a proxy to calculate the stock for 1985 and 1986. Data from the 1982 *Census* was used to calculate shares for 1980 to 1984. Stocks for 1976 to 1979 were based on data from the 1978 *Census*. Data from the 1974 *Census* were used in estimating stocks for 1972 to 1975, while stocks for 1969 to 1971 were estimated using 1969 *Census* data.

The BEA estimate of capital in manufacturing was distributed among states based on their shares of the gross book value of depreciable assets in manufacturing. Asset data were taken from the 1977 and 1982 *Census of Manufactures*. State asset data were not yet available from the 1987 *Census* so the 1985 *Annual Survey of Manufactures* was used to estimate 1985 and 1986 stocks. The 1972 *Census* did not report asset data by state so the 1971 *Survey* was used as a proxy for stocks for 1970 to 1974, while the 1969 *Survey* was used to apportion the 1969 stock.

The BEA estimate of capital in the nonfarm, nonmanufacturing sector was divided among the states according to the sum of estimates for many subsectors: construction, mining, retail and wholesale trade, banking, railroad transportation, trucking and warehousing, water transportation, air transportation, electric services, gas services, telephone and telegraph, and services. The sum of asset estimates for all states, for all subsectors, represented nearly three-quarters of the BEA national total of nonfarm, nonmanufacturing assets. The following equation describes this estimating procedure:

$$NFNMFGK_i = (shCONSTR_i * CONSTRK) + (shMI_i * MIK) + (shR_i * RK) + (shW_i * WK) + (shBK_i * BK) + (shRAIL_i * RAILK) + (shTRUCK_i * TRUCKK) + (shBOAT_i * BOATK) + (shAIR_i * AIRK) + (shELEC_i * ELECK) + (shGAS_i * GASK) + (shTEL_i * TELK) + (shSVCS_i * SVCSK)$$

where sh = share.

The BEA estimate of assets in construction (CONSTRK) was distributed among states based on their share of the gross book value of depreciable assets taken from the *Census of Construction* for 1972, 1977 and 1982. No state data were yet available from the 1987 Census so 1982 shares were used to estimate stocks from 1980 to 1986.

Assets in mineral industries (MIK) were apportioned in

two parts: assets in oil and gas extraction, and assets in all other mineral industries. The BEA figure for assets in oil and gas extraction was apportioned among the states based on their shares of oil production in 1972, 1977, 1982 and 1986. Production values for 1972 and 1977 were taken from the *Minerals Yearbook* while values for 1982 and 1986 were taken from the Energy Information Administration's *Petroleum Supply Annual*. (Since 1982, when the Department of Energy was created, it has been responsible for publishing data on fuel production. Prior to that time these data were tracked in the Bureau of Mines' *Minerals Yearbook*.) Assets in all other mineral industries were distributed according to the following methodology. The *Census of Mineral Industries* for 1977 and for 1982 listed end of year gross book value of depreciable assets, by state. These same data were not calculated in 1972, and the 1987 data were not available yet. The proxy for 1986 shares (used to distribute total asset values for 1985 and 1986) was calculated by increasing each state's 1982 asset value by the ratio of each state's value of nonfuel mineral production in 1986 to the value of its nonfuel mineral production in 1982:

$$\text{assets}_{i86} = \text{assets}_{i82} * \frac{\text{Value of non-fuel mineral production}_{i86}}{\text{Value of non-fuel mineral production}_{i82}}$$

The 1972 proxy was calculated in a similar manner, with the 1977 asset value multiplied by the ratio of the value of 1972 production to the value of 1977 production. State asset values were summed, and then each state's share of this total value was calculated and used to apportion the BEA's total national value of assets in mineral industries (excluding oil and gas extraction).

The value of retail and wholesale trade assets (RK and WK) were apportioned according to each state's share of sales, taken from the *Census of Wholesale Trade* (1972, 1977, 1982, and 1987) and the *Census of Retail Trade* (1972, 1977, 1982, and 1987). According to Costa, Ellson and Martin (1987), the differing structure of retail and wholesale trade across states does not significantly affect the asset/sales ratio.

Assets in banking (BK) were distributed in a manner similar to wholesale and retail trade, using each state's share of deposits in 1972, 1977, 1982, and 1986. The source for deposit information was the *Statistical Abstract of the United States*, and the data reflect deposits of insured commercial banks.

The national estimate of assets in rail transportation (RAILK) was divided among states based on their proportion of track mileage in 1972, 1977, 1982, and 1986. Data on miles of track by state were taken from *Railroad Facts*.

Trucking and warehousing assets (TRUCKK) were distributed to states using the number of trucks in each state. Data on number of trucks by state were available from the *Census of Transportation* for 1972, 1977 and 1982, and from the *1987 Census of Transportation* for a limited number of states. The average growth rate in the number of trucks for states that had both 1982 and 1987 data points was used to extrapolate the number of trucks in 1987 for states without 1987 data.

The BEA national estimate of assets in water transportation (BOATK) was apportioned among states based on data from *Waterborne Commerce of the United States* (1972, 1977, 1982, and 1986) on the value of commerce in ports.

Each state's share of total civil aircraft was used to distribute the national value of assets in air transportation (AIRK). The Federal Aviation Administration's *Census of U.S. Civil Aircraft* (1972, 1977, 1982 and 1986) provided the data on the number of aircraft.

The proxy used to distribute assets in electric services (ELECK) was the generating capacity installed in each state, taken from the *Statistical Abstract* for 1972 and 1977, and the *Inventory of Power Plants in the United States* for 1982 and 1986.

The national estimate of gas services assets (GASK) was divided among states based on their share of miles of pipeline and main. *Gas Facts*, a publication of the American Gas Association, was the source for these data.

Assets in telephone and telegraph (TELK) were divided among states using their share of miles of wire in cable. These data came from the Federal Communication Commission's *Statistics of Communications Common Carriers* for 1972, 1977, 1982, and 1986.

The final categories of assets to be distributed among states are those in the services sector (SVCSK). BEA national asset estimates in six service categories were apportioned using each state's share of sales in that category. These six estimates were summed for each state to approximate assets in services. The six categories were hotels, personal services, business services, auto repair services, amusement services, and legal services. Sales data were taken from the *Census of Service Industries* for 1972, 1977, 1982 and 1987.

The next step was to sum the asset estimates of all these nonfarm, nonmanufacturing subsectors for each state to arrive at a proxy for nonfarm, nonmanufacturing assets. These values were then summed across all states and each state's share of this sum was used to apportion the BEA national estimate of capital stock in the nonfarm, nonmanufacturing sector.

Appendix A: Sources

American Gas Association. 1973, 1978, 1983, 1989. *Gas Facts*. Arlington, Virginia: American Gas Association.

Association of American Railroads. 1973, 1978, 1982, 1986. *Railroad Facts*. Washington, D.C.: Association of American Railroads.

Costa, Jose da Silva, Richard W. Ellson and Randolph C. Martin. 1987. "Public Capital, Regional Output, and Development: Some Empirical Evidence." *Journal of Regional Science*, vol. 27, no. 3, pp. 419–437.

Musgrave, John C. 1987. "Fixed Reproducible Tangible Wealth in the United States, 1983–86." *Survey of Current Business*, vol. 67, no. 8, pp. 100–103.

U.S. Bureau of the Census. 1969, 1971, 1985. *Annual Survey of Manufactures*. Washington, D.C.: Government Printing Office.

———. *Census of Agriculture*, 1969, 1974, 1978, 1982, 1987; *Census of Construction*, 1972, 1977, 1982; *Census of Manufactures*, 1977, 1982; *Census of Mineral Industries*, 1977, 1982; *Census of Retail Trade*, 1972, 1977, 1982, 1987; *Census of Service Industries*, 1972, 1977, 1982, 1987; *Census of Transportation*, 1972, 1977, 1982, 1987; *Census of Wholesale Trade*, 1972, 1977, 1982, 1987.

———. 1958 to 1988. *Governmental Finances*. Washington, D.C.:

Government Printing Office.
———. 1973, 1974, 1978, 1984, 1985, 1988, 1989. *Statistical Abstract of the United States*. Washington, D.C.: Government Printing Office.
U.S. Bureau of Economic Analysis. 1987. *Fixed Reproducible Tangible Wealth in the United States: 1925–1985*. Washington, D.C.: Government Printing Office.
U.S. Department of the Army, Corps of Engineers, 1972, 1977, 1982, 1986. *Waterborne Commerce of the United States, Vol. 5, National Summaries*. Washington, D.C.: Government Printing Office.
U.S. Department of Energy, Energy Information Administration. 1982, 1986. *Inventory of Power Plants*. Washington, D.C.: Government Printing Office.
———. 1982, 1986. *Petroleum Supply Annual*. Washington, D.C.: Government Printing Office.
U.S. Department of the Interior, Bureau of Mines. 1972, 1977. *Minerals Yearbook, Vol. II, Area Reports: Domestic*. Washington, D.C.: Government Printing Office.
U.S. Federal Aviation Administration. 1972, 1977, 1982, 1986. *Census of U.S. Civil Aircraft*. Washington, D.C.: Government Printing Office.
U.S. Federal Communications Commission. 1972, 1977, 1982, 1986. *Statistics of Communications Common Carriers*. Washington, D.C.: Government Printing Office.

Appendix B

Variables Used in the Firm Location Model of Employment Growth

Variable Name	Definition	Source
Dependent:		
CHPE	Average annual percent change in private nonagricultural employment	U.S. Bureau of Labor Statistics, *Handbook of Labor Statistics*, 1989
Independent:		
Cost of Labor		
WAGE	Hourly wage in manufacturing	U.S. Bureau of Labor Statistics, *Handbook of Labor Statistics*, 1989 and 1976
U%	Unemployment rate	U. S. Bureau of Labor Statistics, *Employment and Wages, Annual Averages 1980* and U.S. Department of Labor, *Employment and Training Report of the President, 1976*
COLL	Percent of the population aged 25 years or older that has completed at least four years of college	U.S. Bureau of the Census, *Census of Population, General Social and Economic Characteristics*, 1970 and 1980
Cost of Land		
POP DENSITY	Population density calculated as the ratio of total population to land area	U.S. Bureau of the Census, *Statistical Abstract of the United States*, 1979 and 1989
Cost of Energy		
ENERGY	Cost per million BTUs of purchased fuels and electricity in the industrial sector	U.S. Bureau of the Census, *Statistical Abstract of the United States*, 1983 and 1984 and U.S. Department of Energy, Energy Information Administration, *State Energy Price and Expenditure Report, 1987, State Energy Data Book, 1960–1979*
TEMP	Normal daily maximum temperature in July	U.S. Bureau of the Census, *Statistical Abstract of the United States*, 1979 and 1989
Potential Sales		
URBAN	Percent of the population living in metropolitan areas	U.S. Bureau of the Census, *Statistical Abstract of the United States*, 1984
TAXES	Total state and local taxes as a percent of personal income	U.S. Bureau of the Census, *Governmental Finances*, 1969–70 and 1979–80, and U.S. Bureau of Economic Analysis, *Survey of Current Business*, August 1987
PUBLIC CAPITAL	Per capita public capital stock	See Appendix A for a discussion of the creation of public capital stocks. Population data from U.S. Bureau of the Census, *Statistical Abstract of the United States*, 1979 and 1989

Note: All dollar values for equations employing 1970 levels were expressed in 1970 dollars, while dollar values for equations using 1980 levels were expressed in 1980 dollars. The variables in the equation employing changes in independent variables from 1970 to 1980 were calculated as the percent change in constant (1982) dollars for variables measured in dollars or the absolute change for those variables measured as percentages.

[1] The problem with this interpretation is that no measure of infrastructure is included in the equation and total factor productivity is calculated as a residual. If public capital is a legitimate input, then omitting it from the equation produces a biased estimated of multifactor productivity. See Munnell (1990).

[2] The productivity component can also be specified in a fashion that yields a time trend when the equation is translated into logarithms. Specifically, if $Q = MFPe^{\lambda t}K^{a}L^{b}G^{c}$, then $\ln Q = \ln MFP + gt + a\ln K + b\ln L + c\ln G$. Since equations with the time trend differed little from the simpler version described in the text, the results were not generally reported. This is confirmed by comparing Equation 3 from Table 5 and the same equation including the time trend.

	lnMFP +	λt +	alnK +	blnL +	clnG +	dU%
Eq. 3	5.75		.31	.59	.15	− .007
	(39.7)		(30.1)	(43.2)	(9.0)	(4.7)
Eq. 3′	5.70	.002	.30	.59	.17	− .008
	(39.3)	(2.7)	(28.9)	(42.6)	(9.4)	(5.4)

[3] In view of the importance of this number, it may be useful to report the calculation. The coefficient of each capital variable is the output elasticity, or the percentage change in output for a given percentage change in the input. In the case of public capital, this means that $0.15 = (\Delta Q/Q)/(\Delta G/G)$. Rewriting the equation in terms of marginal productivity produces $\Delta Q/\Delta G = 0.15(Q/G)$. In 1986, total gross state product (Q) was $3,680 billion and total state and local capital (G) was $1,595 billion. Substituting these values into the equation yields a marginal productivity of public capital of 0.35.

In the case of private capital, the relevant figures are 0.35 for the output elasticity and $3,670 billion for private capital. Introducing these figures into the equation yields $\Delta Q/\Delta K = 0.35 \times 3{,}680/3{,}670 = 0.35$.

[4] In estimating the equation, it is necessary to use lagged values of the determinants of marginal productivity of capital, since these determinants include this period's capital stock—the dependent variable.

[5] Several studies have attempted to examine the impact of publicly provided services on firm location decision. Investigators commonly include a measure of spending on welfare, which may be perceived by firms as an "unwanted" public expenditure, as well as measures of spending on "wanted" public expenditures, such as education or police and fire protection. See Wasylenko and McGuire (1985), Plaut and Pluta (1983), Bartik (1989) and Helms (1985).

[6] The change in population density appears to be playing the role of population growth rather than change in land cost in this equation. One would expect a close relationship between state population growth and the growth of nonmanufacturing employment, as local merchants expand to provide a wide array of services for the enlarged pool of consumers. Indeed, in an equation with manufacturing employment, rather than private nonagricultural employment, as the dependent variable, the change in population density is no longer statistically significant. This seems to confirm a strong positive relationship between the change in population density and the growth of nonmanufacturing employment.

References

Aschauer, David A. 1987. "Net Private Investment and Public Expenditure in the United States, 1953–1984." Chicago: Federal Reserve Bank of Chicago, Staff Memoranda, SM-87–10.

———. 1988. "Does Public Capital Crowd Out Private Capital?" Chicago: Federal Reserve Bank of Chicago, Staff Memoranda, SM-88–10.

———. 1989. "Is Public Expenditure Productive?" *Journal of Monetary Economics*, vol. 23, no. 2, pp. 177–200.

———. 1990. "Highway Capacity and Economic Growth: Concepts and Evidence." Unpublished.

Bartik, Timothy J. 1985. "Business Location Decisions in the United States: Estimates of the Effects of Unionization, Taxes, and Other Characteristics of States." *Journal of Business and Economic Statistics*, vol. 3, no. 1, pp. 14–22.

———. 1989. "Small Business Start-Ups in the United States: Estimates of the Effects of Characteristics of States." *Southern Economic Journal*, vol. 55, no. 4, pp. 1004–1018.

Barth, James R. and Joseph J. Cordes. 1980. "Substitutability, Complementarity, and the Impact of Government Spending on Economic Activity." *Journal of Economics and Business*, vol. 32, no. 3, pp. 235–242.

Bennett, James T. 1983. "The Impact of the Composition of Government Spending on Private Consumption and Investment: Some Empirical Evidence." *Journal of Economics and Business*, vol. 35, no. 2, pp. 213–220.

Carlton, Dennis W. 1979. "Why New Firms Locate Where They Do: An Econometric Model." In *Interregional Movements and Regional Growth*, William Wheaton, ed., pp. 13–50. Washington, D.C.: The Urban Institute, Coupe Papers on Public Economics.

———. 1983. "The Location and Employment Choices of New Firms: An Econometric Model with Discrete and Continuous Endogenous Variables." *The Review of Economics and Statistics*, vol. 65, no. 3, pp. 440–449.

Dalenberg, Douglas. 1987. "Estimates of Elasticities of Substitution Between Public and Private Inputs in the Manufacturing Sector of Metropolitan Areas." Eugene, Oregon: University of Oregon, unpublished Ph.D. dissertation.

Deno, Kevin T. 1988. "The Effect of Public Capital on U.S. Manufacturing Activity: 1970 to 1978." *Southern Economic Journal*, vol. 55, no. 2, pp. 400–411.

———. 1986. "The Short-Run Relationship Between Investment in Public Infrastructure and the Formation of Private Capital." Eugene, Oregon: University of Oregon, unpublished Ph.D. dissertation.

DeRooy, Jacob. 1978. "Productivity of Social Overhead Capital: North-South Comparisons." *Review of Business and Economic Research*, vol. 14, no. 1, pp. 44–60.

Dowrick, Steve and Duc-Tho Nguyen. 1989. "OECD Comparative Economic Growth 1950–85: Catch-Up and Convergence." *The American Economic Review*, vol. 79, no. 5, pp. 1010–1030.

Eberts, Randall W. 1986. "Estimating the Contribution of Urban Public Infrastructure to Regional Economic Growth." Cleveland: Federal Reserve Bank of Cleveland, Working Paper No. 8610.

———. 1989. "Some Empirical Evidence on the Linkage Between Public Infrastructure and Local Economic Development." Presented at the Western Economic Association Meetings, June 20.

———. 1990. "Public Infrastructure and Regional Economic Development." *Economic Review*, Federal Reserve Bank of Cleveland, vol. 26, no. 1, pp. 15–27.

Eberts, Randall W. and Kevin T. Duffy-Deno. 1989. "Public Infra-

structure and Regional Economic Development: A Simultaneous Equations Approach." Cleveland: Federal Reserve Bank of Cleveland, Working Paper No. 8909.

Eberts, Randall W. and Michael S. Fogarty. 1987. "Estimating the Relationship Between Local Public and Private Investment." Cleveland: Federal Reserve Bank of Cleveland, Working Paper No. 8703.

Fox, William F. and Matthew N. Murray. 1990. "Local Public Policies and Interregional Business Development." Forthcoming in *Southern Economic Journal*.

Garcia-Mila, Theresa and Therese McGuire. 1987. "The Contribution of Publicly Provided Inputs to States' Economies." Stony Brook, New York: State University of New York at Stony Brook, Research Paper No. 292.

Hansen, Niles M. 1965. "Unbalanced Growth and Regional Development." *Western Economic Journal*, vol. 4, no. 1, pp. 3–14.

Helms, L. Jay. 1985. "The Effect of State and Local Taxes on Economic Growth: A Time Series–Cross-Section Approach." *The Review of Economics and Statistics*, vol. 67, no. 4, pp. 574–582.

Hulten, Charles R. and Robert M. Schwab. 1984. "Regional Productivity Growth in U.S. Manufacturing: 1951–78." *The American Economic Review*, vol. 74, no. 1, pp. 152–162.

Kopcke, Richard W. 1982. "Forecasting Investment Spending: The Performance of Statistical Models." *New England Economic Review*, Federal Reserve Bank of Boston, November/December, pp. 13–32.

Looney, Robert and Peter Fredericksen. 1981. "The Regional Impact of Infrastructure Investment in Mexico." *Regional Studies*, vol. 15, no. 4, pp. 285–296.

McGuire, Therese J. 1986. "On the Relationship Between Infrastructure Investment and Economic Development." Report to the National Council on Public Works Improvement.

Mera, Koichi. 1973. "Regional Production Functions and Social Overhead Capital: An Analysis of the Japanese Case." *Regional and Urban Economics*, vol. 3, no. 2, pp. 157–185.

———. 1975. *Income Distribution and Regional Development*. Tokyo: University of Tokyo Press.

Moomaw, Ronald L. 1981. "Productive Efficiency and Region." *Southern Economic Journal*, vol. 48, no. 2, pp. 344–357.

Munnell, Alicia H. 1990. "Why Has Productivity Growth Declined? Productivity and Public Investment." *New England Economic Review*, Federal Reserve Bank of Boston, January/February, pp. 3–22.

National Council on Public Works Improvement. 1988. *Fragile Foundations: A Report on America's Public Works*. Washington, D.C.: Government Printing Office.

Newman, Robert J. 1983. "Industry Migration and Growth in the South." *The Review of Economics and Statistics*, vol. 65, no. 1, pp. 76–86.

Plaut, Thomas R. and Joseph E. Pluta. 1983. "Business Climate, Taxes and Expenditures, and State Industrial Growth in the United States." *Southern Economic Journal*, vol. 50, no. 1, pp. 99–119.

Schmenner, Roger W. 1982. *Making Business Location Decisions*. Englewood Cliffs, NJ: Prentice Hall.

Schultze, Charles L. 1990. "The Federal Budget and the Nation's Economic Health, Appendix B: The Productivity Payoff from Public Infrastructure Investment." In *Setting National Priorities Policy for the Nineties*, Henry J. Aaron, ed., p. 63. Washington, D.C.: The Brookings Institution.

U.S. Bureau of the Census. 1970, 1980. *Census of Population, General Social and Economic Characteristics*. Washington, D.C.: Government Printing Office, various issues.

———. 1979, 1984, 1989. *Statistical Abstract of the United States*, Washington, D.C.: Government Printing Office.

———. 1989. "State Population and Household Estimates with Age, Sex, and Components of Change: 1981–88." *Current Population Reports*, Series P-25. Washington, D.C.: Government Printing Office.

U.S. Bureau of Economic Analysis. 1990. Machine Readable Data on Gross State Product.

U.S. Bureau of Economic Analysis, Regional Economic Measurement Division. 1987. "State Personal Income, 1969–86: Revised Estimates." *Survey of Current Business*, vol. 67, no. 8, pp. 43–57.

U.S. Bureau of Labor Statistics. 1976, 1989. *Handbook of Labor Statistics*. Washington, D.C.: Government Printing Office.

U.S. Congress, Office of Technology Assessment. 1990. *Rebuilding the Foundations: State and Local Public Works Financing and Management*. Washington, D.C.: Government Printing Office.

U.S. Department of Energy, Energy Information Administration. 1981. *State Energy Data Report, 1960 through 1979*. Washington, D.C.: Government Printing Office.

———. 1989. *State Energy Price and Expenditure Report 1987*. Washington, D.C.: Government Printing Office.

U.S. Department of Labor. 1976. *Employment and Training Report of the President*. Washington, D.C.: Government Printing Office.

U.S. Department of Transportation. 1990. *Moving America: New Directions, New Opportunities*. Washington, D.C.: Government Printing Office.

Wasylenko, Michael and Therese McGuire. 1985. "Jobs and Taxes: The Effect of Business Climate on States' Employment Growth Rates." *National Tax Journal*, vol. 38, no. 4, pp. 497–511.

The Third Deficit: The Shortfall in Public Capital Investment

During the past few years, academic work, commission reports and natural disasters have highlighted the fact that we have been neglecting our stock of public capital and that this lack of attention can cause serious problems. At the Federal Reserve Bank of Boston's most recent economic conference, June 27, 28, and 29, a group of academics, economists, government officials and construction experts convened to delineate the magnitude of the shortfall in public investment, explain its potential economic consequences, and suggest mechanisms to help reverse the trend. The discussion was limited to public investment in physical capital, to make the topic manageable. The conference agenda is outlined below.

What Is the Current State of Our Infrastructure?
George E. Peterson, The Urban Institute
Discussants: Alan S. Blinder, Princeton University
Joel A. Tarr, Carnegie-Mellon University

Why Is Infrastructure Important?
David A. Aschauer, Federal Reserve Bank of Chicago
Discussants: Henry J. Aaron, The Brookings Institution
Richard A. Musgrave, University of California at Santa Cruz

How Does Public Infrastructure Affect Regional Economic Performance?
Alicia H. Munnell, Federal Reserve Bank of Boston
Discussants: Charles R. Hulten, University of Maryland
Ann F. Friedlaender, Massachusetts Institute of Technology

How Efficient Is Current Infrastructure Spending and Pricing?
Clifford M. Winston, The Brookings Institution
Discussants: Alan A. Altshuler, Harvard University
Michael E. Bell, The Johns Hopkins University

What Is the Role of the Private Sector in the Provision of Infrastructure?
John R. Meyer and Jose A. Gomez-Ibanez, Harvard University
Discussants: Sir Alan A. Walters, Putnam, Hayes & Bartlett, Inc.
Gail D. Fosler, The Conference Board

How Should Public Investment in Infrastructure Be Financed?
Edward M. Gramlich, University of Michigan
Discussants: Rudolph G. Penner, The Urban Institute
James M. Poterba, Massachusetts Institute of Technology

The proceedings of Conference Series No. 34 will be available early next year without charge on request to the Research Library—D, Federal Reserve Bank of Boston, P.O. Box 2076, Boston, Massachusetts 02106–2076.

The United States in Debt

Norman S. Fieleke

Vice President and Economist, Federal Reserve Bank of Boston. Lisa O'Brien provided research assistance. This article is drawn partly from the author's The International Economy under Stress *(Cambridge, Mass.: Ballinger Publishing Co., 1988).*

When the U.S. deficit on current international transactions soared to record levels during the mid-1980s, some observers perceived a grave loss of U.S. competitiveness that was "deindustrializing" America. Others warned of an imminent international financial crisis, predicting that the deficits would undermine confidence in the U.S. dollar (and in dollar-denominated assets) and induce a sharp drop in the dollar's foreign-exchange value and a sharp rise in U.S. interest rates. The heightened interest rates would precipitate a U.S. recession that would become worldwide—a "hard landing."[1]

Thus far, the landing has been far from hard. To be sure, the weighted average foreign-exchange value of the dollar did decline fairly steadily and significantly in real terms (adjusted for U.S. minus foreign inflation) during the years 1985 to 1987. But U.S. interest rates also generally declined, rather than rose, between the beginning and the end of this period, and both the U.S. and the world economies grew at a healthy pace. Between the end of 1987 and this writing, the foreign-exchange value of the dollar has changed relatively little, in spite of continued large U.S. current-account deficits; in fact, several central banks have on occasion sold large volumes of dollars in an effort to prevent the dollar from *rising* in the foreign-exchange markets! The economic expansion has continued, and the specter of a hard landing is invoked much less frequently in economic discourse.

The nonoccurrence to date of a hard landing does not prove that one will not take place. And even without a hard landing, the increasing U.S. indebtedness generated by the nation's current-account deficits will impose a growing burden on the U.S. economy. This article examines the growth of U.S. indebtedness to the rest of the world and the underlying causes, as well as the consequences and some proposed remedies.

The Magnitude of U.S. Indebtedness

As can be seen in Table 1, as recently as 1983 the United States was a net creditor in the community of nations, with assets abroad amounting to $89 billion more than foreign assets in this country. The transition from creditor to debtor status was swift and dramatic. The nation had attained its peak as a creditor in 1981, with a positive net international investment position of $141 billion. By the end of 1985 net indebtedness amounted to $117 billion, and by 1989 to $664 billion.

The change in the international investment position is attributable partly to asset purchases and sales (that is, capital flows) and partly to changes in the value of the assets that are held. To illustrate, Table 2 shows that capital flows comprised by far the largest component of the change in the U.S. position during 1989; foreigners loaned or invested about $215 billion in the United States, $88 billion more than U.S. residents invested in foreign countries. Aside from such capital flows, rising securities prices increased the value of stocks and bonds held both in the United States and abroad, with foreign assets in the United States increasing by $53 billion more than U.S. assets abroad on this count. Also, changes in the dollar exchange rates of other currencies somewhat altered the dollar value of foreign-currency-denominated stocks and bonds.

Table 1
International Investment Position of the United States at Year End, 1970–89
Billions of Dollars

Year	U.S. Assets Abroad (1)	Foreign Assets in the United States (2)	Net International Investment Position of the United States (column 1 less column 2) (3)
1970	165.4	106.9	58.5
1971	179.0	133.5	45.5
1972	198.7	161.7	37.0
1973	222.4	174.5	47.9
1974	255.7	197.0	58.7
1975	295.1	220.9	74.2
1976	347.2	263.6	83.6
1977	379.1	306.4	72.7
1978	447.8	371.7	76.1
1979	510.6	416.1	94.5
1980	607.1	500.8	106.3
1981	719.6	578.7	140.9
1982	824.8	688.1	136.7
1983	873.5	784.5	89.0
1984	895.9	898.1	−2.2
1985	949.7	1,066.9	−117.2
1986	1,073.4	1,347.1	−273.7
1987	1,175.9	1,554.0	−378.1
1988	1,265.6	1,796.7	−531.1
1989[p]	1,412.5	2,076.3	−663.7

Note: Detail may not add to totals shown because of rounding.
[p]Preliminary.
Source: *Survey of Current Business*, vol. 66, June 1986, p. 28; vol. 69, June 1989, p. 43; vol. 70, June 1990, p. 59.

This measurement of the U.S. position may be substantially in error, as the Commerce Department, the source of the data, points out. On the one hand, some U.S. claims on foreigners are understated because of certain measurement conventions or difficulties. For instance, U.S. official gold holdings—deemed, like U.S. holdings of foreign currency, to be a claim on foreigners—are valued at a most conservative $42.22 per ounce. Revaluing this gold stock at $400 per ounce—roughly the market price at this writing—would raise the reported value of U.S. assets abroad at year-end 1989 by nearly $94 billion. Similarly, U.S. direct investments abroad are carried at their original book value rather than at their higher current market value.

On the other hand, other measurement problems probably result in an understatement of the value of foreign assets in the United States. In particular, for years the United States has been receiving from abroad very large net receipts that cannot be traced to specific transactions—the so-called "statistical discrepancy" in the balance of payments. Some, perhaps most, of these net receipts—which totaled $22 billion in 1989—may well have been generated by "capital-account" transactions, particularly by foreign investment in the United States. Thus, some understatement of foreign assets in the United States seems likely. Such an understatement would, of course, tend to lower the reported net indebtedness of the United States below its true value, while the likely understatement of U.S. assets abroad would have the opposite effect.[2] On balance, it is hard to say whether the published measure of the U.S. net international investment position is significantly in error. Some considerations suggest an understatement, others an overstatement.

The transition of the United States from creditor to debtor status is not to be explained by transactions with a particular country or region. On the contrary, the U.S. position turned more negative (or less pos-

Table 2
The U.S. Net International Investment Position: Summary of Changes during 1989
Billions of Dollars

	U.S. Assets Abroad (1)	Foreign Assets in the United States (2)	Net International Investment Position of the United States (column 1 less column 2) (3)
Position at End of 1988	1,265.6	1,796.7	−531.1
Changes in 1989 Attributable to:			
Capital Flows	127.1	214.7	−87.6
Price Changes	13.3	66.7	−53.4
Exchange-Rate Changes	−2.3	−1.3	−1.0
Other Changes	8.9	−0.5	9.4
Total Changes	146.9	279.6	−132.7
Position at End of 1989[P]	1,412.5	2,076.3	−663.7

Note: Detail may not add to totals shown because of rounding.
[P]Preliminary.
Source: *Survey of Current Business*, vol. 70, June 1990, p. 55.

itive) with all major areas for which U.S. data are regularly published. As shown by Table 3, through 1988 the biggest swing was with Western Europe.

If the United States has become a sizable net debtor, which countries are the creditors? Unfortunately, data on net international investment position—or "net external assets," as the measure is generally called outside the United States—are officially published by only a few countries, and the comparability of these national measures is doubtful. Some data published by the International Monetary Fund (IMF) for the seven major industrial democracies suggest that Japan, Germany, and the United Kingdom have large net creditor positions (IMF 1988, p. 89). Other major net creditors probably include Switzerland and some members of the Organization of Petroleum Exporting Countries, especially Saudi Arabia and Iraq (Deutsche Bundesbank 1986, p. 30).

Once the United States became a net debtor, it became fashionable to compare its indebtedness with that of the less developed countries. U.S. indebtedness, it was widely reported, had come to exceed the indebtedness even of Brazil, the leading debtor among the developing nations. The comparison, however, was not valid. For one thing, the gross debt of the less developed countries was being compared with U.S. debt net of U.S. assets abroad. Such comparison is sometimes defended on the grounds that the external assets of developing countries typically are relatively small or, when privately owned, are beyond the control or influence of developing-country governments. The argument has merit, but to ignore all such assets is extreme.[3]

If measured gross, on roughly the same basis as less developed country debt is measured, the U.S. external debt came to $753 billion at the end of 1985, the year during which the nation became a net

Table 3
Net International Investment Position of the United States by Area at Year End, 1981 and 1988
Billions of Dollars

Area	1981	1988[P]	Change
Western Europe	−51.8	−436.9	−385.0
Japan	−1.7	−128.5	−126.8
Canada	66.9	53.5	−13.4
Latin American Republics and Other Western Hemisphere	99.3	−23.6	−122.9
Other	28.3	2.9	−25.4
Total	140.9	−532.5	−673.5

Note: At this writing data for the geographic areas listed are not available for 1989. The data shown were obtained in 1989, and the total in the column for 1988 differs somewhat from the total shown in Tables 1 and 2, which could be compiled from a 1990 source. In addition, detail may not add to totals shown because of rounding.
[P]Preliminary
Source: *Survey of Current Business*, vol. 69, June 1989, p. 42; U.S. Commerce Department staff.

debtor. This amount greatly exceeded the gross external debt of any less developed country. Indeed, the total external debt of all the capital-importing developing countries then amounted to only about 1¼ times the U.S. debt.[4]

In any event, by any conventional measure U.S. indebtedness increased dramatically. We shall examine some explanations for the U.S. external deficit and then consider the possible consequences of the deficit. Explanations can be classified into those that emphasize "supply-side" factors, "demand-side" factors, or both.

Supply-Side Explanations: Price Competitiveness

The large trade and current-account deficits that have ballooned U.S. net debt (Table 4) are often taken to signify a loss of U.S. "competitiveness." What is meant by competitiveness is seldom spelled out, but the concern is commonly with factors that underlie the aggregate supply of U.S. goods—factors such as technology, capital formation, research and development, and the quality of management and the labor force. Thus, to enhance U.S. competitiveness, action has often been proposed to upgrade the education of the work force (especially in math and science), to grant more favorable tax treatment to investment in capital equipment, to relax the antitrust laws so that firms could pool their research efforts, to provide better patent protection for new inventions, and so on.

Analyses of supply-side competition, or of competition among suppliers, commonly divide it into two broad categories: price competition and nonprice competition. Price competition is the subject of this section.

Arguably the best single index of a nation's changing overall price competitiveness is the change in its real exchange rate, that is, the change in its average price level relative to the average foreign price level after taking into account the change in the average foreign-currency price of its currency. Thus, a nation's price competitiveness will be impaired by a rise in its domestic prices relative to foreign prices, unless an offsetting decline occurs in the foreign-currency price of its currency.

Although analysts differ on precisely how to measure the real exchange rate, all widely used measures show big swings in U.S. price competitiveness during the period of deterioration in the U.S. trade and current-account balances. In general, the indexes suggest that the United States lost much price competitiveness between 1980 and 1985, but then rapidly regained the lost ground. For example, the index plotted as a solid line in Chart 1 shows a rise in U.S. relative prices of 37 percent (after incorporating nominal exchange-rate change) from 1980 to 1985, followed by a decline to approximately the 1980 level by the end of 1987. The "nominal" index plotted in the chart represents only the change in the foreign-currency price of the dollar. Clearly, it was this nominal exchange-rate change, rather than changes in domestic or foreign prices, that accounted for most of the large swings in U.S. overall price competitiveness over this period.

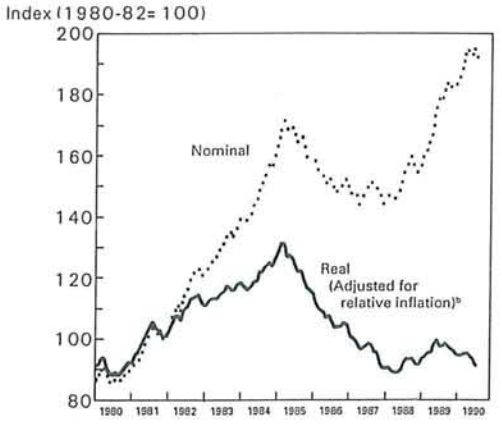
Table 4
U.S. Balances on Selected Components of International Current Account Transactions, 1970–89
Billions of Dollars; (−) signifies deficit

Year	Balance on Merchandise Trade (1)	Balance on Services and Income (2)	Balance on Unilateral Transfers (3)	Balance on Current Account (Column 1 + 2 + 3) (4)
1970	2.6	5.9	−6.2	2.3
1971	−2.3	8.2	−7.4	−1.4
1972	−6.4	9.2	−8.5	−5.8
1973	.9	13.1	−6.9	7.1
1974	−5.5	16.7	−9.2[a]	2.0
1975	8.9	16.3	−7.1	18.1
1976	−9.5	19.4	−5.7	4.2
1977	−31.1	21.8	−5.2	−14.5
1978	−33.9	24.3	−5.8	−15.4
1979	−27.5	33.1	−6.6	−1.0
1980	−25.5	34.9	−8.3	1.1
1981	−28.0	43.2	−8.3[b]	6.9
1982	−36.4	40.3	−9.8	−5.9
1983	−67.1	36.9	−10.0	−40.1
1984	−112.5	26.1	−12.6	−99.0
1985	−122.1	15.3	−15.5	−122.3
1986	−145.1	15.7	−16.0	−145.4
1987	−159.5	11.8	−14.6	−162.3
1988	−127.0	13.1	−15.0	−128.9
1989	−114.9	19.6	−14.7	−110.0

Note: Detail may not add to totals shown because of rounding.
[a]Includes extraordinary U.S. Government transactions with India.
[b]Break in series. Beginning with data for 1981, private remittances to foreign students in the U.S. are included.
Source: *Survey of Current Business*, vol. 70, June 1990, pp. 75–76.

Chart 1

Weighted Average Foreign Exchange Value[a] *of the U.S. Dollar, 1980-90*



[a] Measured in terms of 40 other currencies weighted according to maufactures trade; based on monthly averages of daily rates.
[b] Inflation measured in terms of wholesale prices in manufactures excluding food and fuel.
Source: Morgan Guaranty Trust Company.

It is widely agreed that the loss of U.S. price competitiveness between 1980 and 1985 contributed substantially to the increase in the U.S. trade deficit. But what caused the loss of price competitiveness? A number of factors could be responsible, not all of them supply-side in nature. Here a supply-side factor—productivity change—is considered; other factors are discussed in a following section.

Changes in the productiveness of a country's resources can have an important influence on the country's price competitiveness. If productivity rises, other things remaining equal, the money cost and price of a unit of output can fall. In evaluating overall productivity, one should consult a measure of the output yielded by a unit of all productive factors combined, including labor, land, and capital. But such measures of total factor productivity are extraordinarily difficult to construct. Consequently, international productivity comparisons are commonly based on indexes of output per input of labor in manufacturing, such as those in Table 5.

Do these indexes suggest that lagging productivity growth was responsible for the decline in U.S. price competitiveness between 1980 and 1985? Among the "Big Seven" countries listed individually, the United States ranked in the middle in productivity performance over this period, surpassing Canada, France, and West Germany but lagging behind Japan, Italy, and the United Kingdom. But how did the United States perform by comparison with its major industrial competitors as a group? As shown in the last column, eleven foreign industrial countries achieved an average increase of about 25 percent over the years 1980 to 1985, compared to an increase of nearly 21½ percent in the United States. This differential of some 3½ percent falls far short of accounting for the 37 percent deterioration in overall U.S. price competitiveness over this period.

In fact the U.S. lag in labor productivity growth was far greater during the 1970s than during the 1980s when the U.S. trade deficit increased so greatly. Between 1970 and 1975 labor productivity in manufacturing rose by 15 percent in the United States and by 26 percent in the eleven foreign industrial countries. And between 1975 and 1980 the increase was only 10 percent in the United States and 24 percent in the eleven other nations.

It is clear that the loss of U.S. price competitiveness during the period 1980–85 should not be attributed to the relative U.S. record on labor productivity in manufacturing over these years. To be sure, higher U.S. productivity growth could, in principle, have yielded lower U.S. inflation and, other things being equal, a smaller rise in the U.S. real exchange rate than that shown in Chart 1. The relative U.S. record on labor productivity in manufacturing, however, was extremely good during the 1980–85 period by comparison with the previous decade.

Supply-Side Explanations: Nonprice Considerations

Supply-side explanations of the U.S. trade and current-account deficits relate not only to the price competitiveness of U.S. suppliers but also to nonprice competition. U.S. firms were often said to have lost competitiveness because their products had become inferior in quality to foreign brands. Automobiles provide a good illustration. During the 1980s, surveys showed that U.S. consumers and engineers both considered foreign-brand cars generally to be of higher quality than U.S. cars. Consumers buying foreign cars were more likely to be satisfied with their purchase and to report a low frequency of repairs than were the buyers of U.S. cars.[5] In addition, U.S. firms were criticized for failing to tailor products to

Table 5
Output per Labor Hour in Manufacturing in Selected Industrial Countries, 1970–88
Indexes: 1980 = 100

Year	Country							
	United States	Canada	Japan	France	West Germany	Italy	United Kingdom	Eleven Countries[a]
1970	78.9	77.0	52.8	64.6	65.6	57.2	78.8	64.0
1971	83.2	82.4	55.9	68.1	68.1	58.2	82.6	67.5
1972	86.4	86.5	61.4	71.0	72.7	63.1	87.2	72.2
1973	91.1	92.0	67.7	75.8	77.3	68.3	93.6	78.0
1974	88.8	93.4	70.5	77.4	80.5	73.1	95.2	80.7
1975	91.1	90.2	71.5	80.2	83.0	70.6	93.1	80.8
1976	95.4	96.5	76.9	85.4	88.8	80.2	97.3	86.5
1977	98.3	101.8	81.5	90.4	92.1	81.9	98.1	90.3
1978	99.9	103.0	88.0	94.6	94.9	87.2	99.5	94.2
1979	99.9	103.9	93.6	99.2	99.6	95.5	100.6	98.3
1980	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
1981	102.3	104.8	103.7	103.0	102.2	102.7	105.1	103.7
1982	104.8	100.1	110.0	110.3	103.7	105.2	111.4	106.6
1983	110.3	107.3	116.0	113.1	109.8	110.8	120.8	113.2
1984	116.2	116.5	124.3	115.3	113.9	121.9	127.5	120.2
1985	121.4	119.5	131.3	120.0	118.2	128.4	131.6	125.1
1986	126.1	119.9	133.4	122.2	118.1	129.6	136.0	126.6
1987	130.7	122.7	143.8	123.7	119.6	132.9	144.8	131.9
1988	133.7	126.6	154.8	130.3	125.1	136.9	152.0	138.6

Note: The data relate to all employed persons, including the self-employed, in the United States and Canada, and to all employees (wage and salary earners) in the other countries. Although the indexes relate output to the hours of persons employed in manufacturing, they do not measure the specific contribution of labor as a single factor of production. Rather, they reflect the joint effects of many influences, including new technology, capital investment, capacity untilization, energy use, and managerial skills, as well as the skills and efforts of the work force.

[a] A trade-weighted average of Canada, Japan, France, West Germany, Italy, the United Kingdom, Belgium, Denmark, the Netherlands, Norway, and Sweden, but excluding in 1988 Belgium and the Netherlands, for which data are not available at this writing. The weights reflect the relative importance of each country as a U.S. manufacturing trade competitor as of 1980.

Source: *Monthly Labor Review*, vol. 13, April 1990, p. 98; staff of U.S. Bureau of Labor Statistics.

the preferences of foreign purchasers and for failing to mount aggressive, long-term marketing efforts in foreign lands.

Because such reports are so common, it may well be that U.S. competitiveness did lag in terms of quality and other nonprice considerations. How much weight to attach to these various nonprice dimensions is impossible to quantify with any precision. However, other evidence suggests that, whatever the shortcomings of U.S. firms, the worldwide performance of U.S. management did not degenerate as the U.S. trade deficit began to mushroom early in the 1980s.

Some of the most revealing evidence on the relative performance of U.S. management has to do with the record of multinational firms headquartered in the United States. By and large, management has more scope to influence the operations of a multinational firm than a national firm. With activities in more than one country, the multinational firm is not chained to local customs, regulations, or labor force. Over time management can shift activities of the firm from one nation to another as circumstances warrant, taking advantage of the best that each nation has to offer for the overall success of the firm. Thus, one may gain more insight into the international competitiveness of U.S. management by examining the performance of U.S.-based multinationals than by examining only the performance of U.S. firms.

Summary data on performance in export markets for manufactures are presented in Table 6. As shown by the bottom line, the overall share of U.S.-based multinationals—including exports by U.S. parents as well as by their majority-owned foreign affiliates (MOFAs)—increased somewhat between 1966 and 1984, the latest year for which data are available at this writing. From these numbers, one might argue that U.S. management was holding its ground in the

Table 6
U.S. Shares of the Value of World[a] Manufactures Exports
Percent

	1966	1977	1982	1983	1984
United States	17.5	13.3	14.3	13.7	14.0
U.S. Multinationals					
Parents	11.0	9.2	9.5	9.1	9.2
MOFAs[b]	8.2	9.7	9.7	9.9	10.3
Parents and MOFAs	17.7	17.6	17.7	17.7	18.1

[a] The "world" here is defined as all market economies.
[b]Exports by majority-owned foreign affiliates (MOFAs) as a percentage of exports by all countries except the United States.
Source: Robert E. Lipsey and Irving B. Kravis, "The Competitiveness and Comparative Advantage of U.S. Multinationals, 1957–1984," *Banca Nazionale del Lavoro Quarterly Review*, vol. 161, June 1987, p. 151.

international arena. By comparison, the first line shows that all U.S. manufactured exports, expressed as a share of the world total, declined from 1966 to 1977, rose from 1977 to 1982, and then declined slightly between 1982 and 1984 (a period when the U.S. trade deficit also increased sharply). Thus, the United States as a nation experienced some loss in competitiveness by this indicator between 1982 and 1984, but the rising market share of U.S. multinationals tends to exonerate U.S. management. It is especially interesting that the U.S. multinationals maintained their market share between 1982 and 1984 by raising the share of their MOFAs to compensate for a decline on the part of the U.S. parents.

Competitiveness and Aggregate Supply

Perhaps the best summary indicator of a nation's overall supply-side competitiveness is the share of world output that the nation supplies. Measures of this share are not precise; it is difficult to construct accurate comparisons of the outputs of different countries, partly because the composition and price structure of output vary from country to country. Nonetheless, such comparisons are regularly made by the Organization for Economic Cooperation and Development (OECD), whose membership includes twenty-four countries, nearly all of them industrialized.

In Table 7, which draws on the OECD data, it can be seen that the United States held its own with respect to gross output between 1981 and 1987, a period during which the U.S. trade balance registered a huge decline. Thus, at least by comparison with other countries, the United States did not display a serious aggregate "supply-side" problem during these years of deterioration in its trade balance.

If the supply side cannot be held responsible for the U.S. external deficits, what is to blame? After all, it is clear that the nation lost overall price competitiveness as the deficits began to increase. The answer may lie in the relationship between aggregate U.S. supply and demand and, more precisely, in the forces that influence that relationship.

Aggregate Supply and Demand

If the residents of a nation demand, or absorb, more output than the nation is supplying, the gap is filled by net imports from abroad. Although the output supplied by the nation may be growing rapidly, total demand within the nation may be growing even faster, so that the nation's trade and current-account deficits with the rest of the world expand

Table 7
U.S. and OECD Gross Domestic Product and U.S. Trade Balance, 1981–89

	U.S. GDP as Percent of OECD GDP		
Year	At Current Prices and Exchange Rates	At 1985 Prices and Exchange Rates	U.S. Merchandise Trade Balance (Billions of Dollars; Balance-of-Payments Basis)
1981	38	45	−28.0
1982	40	44	−36.4
1983	42	44	−67.1
1984	44	45	−112.5
1985	45	45	−122.1
1986	39	45	−145.1
1987	36	45	−159.5
1988	35	45	−127.0
1989	36	45	−114.9

Source: *Survey of Current Business*, vol. 70, June 1990, pp. 76–77; Organization for Economic Cooperation and Development, *National Accounts, 1960–88*, Vol. I pp. 123, 127 and 131 (Paris: OECD, 1990); and *Main Economic Indicators*, vol. 90, April 1990, p. 172.

Table 8
Real GNP and Real Domestic Demand in the United States and Other OECD Countries, 1981–89

	Real GNP				Real Domestic Demand			
	Level (1980 = 100)		Percent Change from Preceding Year		Level (1980 = 100)		Percent Change from Preceding Year	
Year	United States	Other OECD	United States	Other OECD	United States	Other OECD	United States	Other OECD
1981	101.9	101.5	1.9	1.5	102.2	100.0	2.2	.0
1982	99.3	102.8	−2.5	1.3	100.3	101.0	−1.9	1.0
1983	102.9	105.1	3.6	2.2	105.4	102.5	5.1	1.5
1984	109.9	108.9	6.8	3.6	114.5	105.5	8.7	2.9
1985	113.5	112.6	3.4	3.4	118.9	108.7	3.8	3.1
1986	116.7	115.6	2.7	2.6	122.9	112.9	3.3	3.8
1987	120.9	119.6	3.7	3.5	126.8	117.7	3.2	4.4
1988	126.3	124.9	4.4	4.4	131.0	124.3	3.3	5.4
1989	130.0	129.8	3.0	3.9	134.1	129.6	2.4	4.3

Source: *OECD Economic Outlook*, vol. 47, June 1990, pp. 181 and 188; *OECD Economic Outlook*, vol. 47, *Statistics on Microcomputer Diskette*, June 1990.

(unless there are offsetting price changes, such as import price reductions). To stem the growth of the deficit, the nation must retard the growth of its demand (that is, its absorption or expenditure) or accelerate the growth of its output.

The United States in the mid-1980s was such a nation. The data in Table 8 confirm that domestic demand grew faster than gross national product in the United States in every year from 1983 through 1986, a period during which dramatic increases occurred in the country's deficits on international trade and current account. Note that during most of this period U.S. output grew faster than output in other OECD countries as a group; however, U.S. demand grew even faster by comparison with demand in other OECD countries.

It seems, then, that the U.S. external deficits are not attributable to "supply-side" problems, certainly not supply-side problems alone. Demand, or more precisely, the changing relationship between demand and supply, seems a more promising subject for analysis. In what follows, some explanations involving both demand and supply are considered. Foreign as well as U.S. demand and supply are relevant, since some of what the United States supplies goes to satisfy foreign demand, while some of U.S. demand is satisfied by foreign supply.

Unfair Foreign Trading Practices

One explanation often advanced for the U.S. trade deficit is unfair foreign trading practices; the playing field is said to be "tilted" against the United States. This explanation involves references to both demand and supply. Although it is foreign rather than U.S. demand and supply that have allegedly been manipulated, the impact would have been to increase U.S. net imports. On the supply side, other nations have been charged with subsidizing or "dumping" their exports in world markets, thus lowering their supply prices and stealing both U.S. and foreign markets from U.S. suppliers. On the demand side, other nations are accused of imposing barriers against U.S. exports, thereby reducing demand for them.

To be sure, unfair trading actions do occur, and national governments, including the U.S. government, commonly undertake to shield firms within their borders against injury from such practices. In the United States, the law provides U.S. industries with remedies against import competition from dumped or subsidized merchandise, as well as against other practices deemed unfair. Dumping is defined as the sale of foreign merchandise at prices below those charged in the foreign producers' home market, or below the foreign cost of production. The antidumping statutes provide for the imposition of

antidumping duties to offset such price-cutting when a determination is made that a domestic industry is being materially injured—or threatened with such injury—by the dumped imports, or that the establishment of the industry is being materially retarded by such imports. Similarly, "countervailing" duties are imposed to offset foreign subsidies upon a determination by U.S. authorities that, because of subsidized import competition, a U.S. industry is being materially injured—or threatened with such injury—or that the establishment of the industry is being materially retarded.[6]

During 1987, when the United States incurred its largest trade deficit ever, the nation imposed new antidumping duties on fifteen products from twenty-six countries, and imposed new countervailing duties on seven products from twelve countries. Other actions were taken against practices that the United States deemed unfair on grounds other than those covered under the antidumping or countervailing duty laws (U.S. International Trade Commission 1988, pp. 5-4–5-11). Therefore, while unfair foreign trading practices may have operated to increase U.S. imports, it is plain that U.S. firms availed themselves of the provisions of U.S. law in order to stem such increases. The burden of proof rests with those who suggest that U.S. imports were bloated by unfair foreign trading practices in spite of the legal remedies that U.S. firms can invoke against such practices. Unfair foreign trading practices were to be found long before the U.S. trade deficit began to surge in the early 1980s, and it remains to be shown that those practices intensified so as to contribute substantially to the deficit.

Another difficulty with attributing the increased U.S. deficit to unfair foreign trading practices is that the increase was distributed widely across both commodity categories and geographic areas. This fact is documented in Tables 9 and 10. It seems most unlikely that virtually all major trading partners of the United States would simultaneously have intensified unfair practices in their trade with the nation.

Table 9 presents aggregate data on trade between the United States and each of its five leading trade partners, listed in order of magnitude of total U.S. trade with them in 1987. Similar data are shown for OPEC and for the rest of the world. Together, the five leading trade partners accounted for 53 percent of U.S. international trade (exports plus imports) in 1987; if OPEC is added, that share rises to 59 percent. Clearly, the U.S. trade balance deteriorated markedly from 1980 to 1987 with every listed area but OPEC, from which U.S. imports of petroleum declined dramatically.

To identify the areas with which the U.S. trade position deteriorated more than proportionately, the last column of the table shows what the value of U.S. exports and imports with each area would have been in 1987 if each area had retained the same percent-

Table 9
U.S. Merchandise Trade, by Major Trading Partners or Areas, 1980 and 1987
Billions of Dollars

		1987	
Country or Area	1980	Actual	Allocated on basis of 1980 shares[a]
Canada			
U.S. exports	41.6	62.0	46.5
U.S. imports	42.9	73.6	70.4
Balance	−1.3	−11.6	−23.9
Japan			
U.S. exports	20.8	27.6	23.2
U.S. imports	31.2	84.6	51.2
Balance	−10.4	−57.0	−28.0
West Germany			
U.S. exports	11.4	11.5	12.8
U.S. imports	11.7	26.9	19.2
Balance	−.2	−15.4	−6.4
Mexico			
U.S. exports	15.2	14.6	17.0
U.S. imports	12.6	20.3	20.6
Balance	2.7	−5.7	−3.6
United Kingdom			
U.S. exports	12.8	13.8	14.3
U.S. imports	9.8	17.2	I6.1
Balance	3.0	−3.5	−1.8
Organization of Petroleum Exporting Countries			
U.S. exports	17.4	10.7	19.4
U.S. imports	55.6	24.4	91.2
Balance	−38.2	−13.7	−71.8
Rest of the world			
U.S. exports	105.1	110.1	117.1
U.S. imports	86.0	162.7	141.0
Balance	19.1	−52.6	−23.8
Total, all areas			
U.S. exports	224.3	250.3	250.3
U.S. imports	249.8	409.8	409.8
Balance	−25.5	−159.5	−159.5

Note: Detail may not add to totals shown because of rounding.
[a]Each area is allocated the same fraction of total 1987 U.S. exports and imports as in 1980.
Source: *Survey of Current Business*, vol. 70, June 1990, pp. 86–88.

ages of total U.S. exports and imports as in 1980. Comparison of the last two columns reveals that the U.S. trade balance worsened not only actually, but disproportionately (the 1987 "actual" exceeds the "allocated"), with all listed areas except OPEC and Canada. While the greatest actual deterioration was with the "rest of the world," the greatest disproportionate deterioration, amounting to $29 billion, was with Japan, with the rest of the world a very close second.

The deterioration in the U.S. trade balance was distributed widely across commodity categories as well as across geographic areas. As indicated in Table 10, aside from the "all other" category, the balance worsened between 1980 and 1987 in every major commodity category except industrial supplies and materials, a category influenced by the decline in oil imports. More than proportionate deteriorations occurred in foods, feeds, and beverages, in capital goods, in automotive vehicles and parts, and in consumer goods, as the actual 1987 deficit was larger for each category (or the actual 1987 surplus was smaller) than it would have been if the category had accounted for the same percentage of total exports and imports as in 1980. (See last two columns of Table 10.)

Thus, the pervasiveness of the deterioration in the U.S. trade balance makes it unlikely that unfair foreign trading practices played a major role. Does this conclusion hold even for U.S. trade with Japan? The issue is raised most often with regard to Japan, partly because the U.S. deficit with that nation increased so sharply and amounted to more than two-fifths of the total U.S. deficit in 1989. While precise explanation of trade flows is very difficult, quantitative studies have concluded that the increase in the U.S. deficit with Japan was attributable mainly, or perhaps fully, to factors such as changes in prices, incomes, and the yen-dollar exchange rate. Any impact of unfair trading practices was adjudged to be decidedly secondary (Bergstrand 1986; Bergsten and Cline 1985, pp. 45–46).

Japan's record is not without blemish, however. In particular, evidence has been marshaled that Japan has offered some formidable "invisible" barriers to international trade. An invisible barrier is a system or regulation that applies to both domestic and foreign producers, but that works, perhaps unintentionally, to reduce the share of imports in domestic consumption. Government procurement policies, the wholesale and retail distribution systems, the setting of product standards, and the testing of products against these standards have commonly been alleged to constitute formidable invisible barriers in Japan. According to one investigation, if Japan's invisible barriers had been reduced to levels corresponding to those in the United States and the European Economic Community in the early 1980s, Japan's manufactured imports might have increased by 27 percent (equivalent to a rise of 7 percent in the country's total

Table 10
U.S. Merchandise Trade, by Major End-Use Category, 1980 and 1987
Billions of Dollars

End-Use Category	1980	1987 Actual	1987 Allocated on basis of 1980[a]
Food, feeds, and beverages			
Exports	36.4	25.3	40.6
Imports	18.5	24.8	30.4
Balance	17.9	.4	10.2
Industrial supplies and materials			
Exports	72.3	70.0	80.7
Imports	132.3	113.7	217.0
Balance	−60.0	−43.8	−136.3
Capital goods, except automotive			
Exports	76.3	92.4	85.1
Imports	31.4	85.1	51.6
Balance	44.8	7.2	33.5
Automotive vehicles, parts, and engines			
Exports	17.4	28.1	19.4
Imports	28.1	85.2	46.0
Balance	−10.7	−57.0	−26.7
Consumer goods (nonfood), except automotive			
Exports	17.7	20.3	19.7
Imports	34.2	88.8	56.1
Balance	−16.5	−68.5	−36.4
All other, including balance-of-payments adjustments			
Exports	4.2	14.3	4.7
Imports	5.2	12.1	8.6
Balance	−1.0	2.2	−3.9
All categories			
Exports	224.3	250.3	250.3
Imports	249.8	409.8	409.8
Balance	−25.5	−159.5	−159.5

Note: Detail may not add to totals shown because of rounding.
[a]Each category is allocated the same fraction of total 1987 U.S. exports and imports as in 1980.
Source: *Survey of Current Business*, vol. 70, June 1990, pp. 90–92.

imports), with at least half of the increased imports coming from the United States. At the same time, the investigation points out that such an increase would be far too small to eliminate the U.S. trade deficit with Japan. Thus, the conclusion remains that the deficit was generated mainly, if not totally, by causes other than unfair trading practices (Christelow 1985–86).

If unfair foreign trading practices are an improbable explanation of the U.S. trade and current-account deficits, what other explanations might be more convincing?

Probable Causes of the U.S. Trade and Current-Account Deficits

In its 1985 annual report (pp. 102–103), the Council of Economic Advisers identified three factors as the *immediate* causes of the U.S. trade deficit: (1) the appreciation of the dollar in the foreign-exchange markets after mid-1980; (2) the more rapid expansion of real income and demand in the United States than in the rest of the world after 1982; (3) the reduced demand for imports by the less developed countries that began to experience severe difficulty in servicing their debt and in obtaining new loans after mid-1982. Subsequent analyses have commonly cited the same factors. The weight of the evidence suggests that the first of these three factors accounted for more than half of the increased deficit, with the second factor accounting for perhaps one third, and other factors accounting for the balance (Hooper and Mann 1987, pp. 41–42, 95–96).

The roles played by dollar appreciation and by rapid U.S. demand growth have already been described. But what explains these factors themselves?

The dollar's value will rise in the foreign-exchange markets if the demand for dollars exceeds the supply at prevailing exchange rates. During the early 1980s, one important development that led to increased demand for dollars, relative to the supply, was an increase in net borrowing from abroad by U.S. residents. Foreign-currency balances were exchanged into dollar balances to accommodate this increase in U.S. borrowing, thereby bidding up the price of the dollar. The increased U.S. borrowing was caused largely by changes in federal fiscal policy, especially the shift toward deficit in the budget, which occurred at a time when U.S. monetary policy was relatively restrictive.

The key role played by government fiscal policy in inducing borrowing from abroad is suggested by an important accounting relationship: private domestic investment can be funded out of either the country's private saving or government saving, or out of funds loaned by foreigners. If government saving decreases without a compensating increase in private saving, private investors must tap foreign saving more heavily if they are to sustain their outlays.

The relative magnitudes involved in this accounting relationship for the United States are shown in Table 11, where private domestic investment in the fourth column is equal to the sum of its sources of financing, itemized in the first three columns. A negative number in one of the first three columns means that saving is being absorbed, on balance,

Table 11
Major Categories of Saving and Investment as a Percentage of GNP for the United States, 1970–90

Year	Gross Private Saving (1)	+	Government Saving (2)	+	Net Investment (Lending) by Foreigners (3)	=	Gross Private Domestic Investment (4)
1970	16.2		−1.0		−.5		14.7
1971	17.3		−1.8		−.1		15.6
1972	16.8		−.3		.2		16.7
1973	18.0		.6		−.6		17.6
1974	17.3		−.3		−.4		16.3
1975	19.0		−4.1		−1.4		13.7
1976	18.0		−2.2		−.5		15.6
1977	17.8		−1.0		.4		17.3
1978	18.2		.0		.5		18.5
1979	17.8		.5		−.1		18.1
1980	17.5		−1.3		−.5		16.0
1981	18.0		−1.0		−.3		16.9
1982	17.6		−3.5		.0		14.1
1983	17.4		−3.8		1.0		14.7
1984	17.9		−2.8		2.4		17.6
1985	16.6		−3.3		2.8		16.0
1986	15.8		−3.4		3.2		15.6
1987	14.7		−2.4		3.3		15.5
1988	15.1		−2.0		2.4		15.4
1989	15.4		−2.0		1.8		14.8
1990[a]	15.3		−2.5		1.5		13.9

Note: Detail may not add to totals shown because of statistical discrepancy.
[a]First quarter.
Source: Board of Governors of the Federal Reserve System, Fame Data Base.

rather than being made available for private domestic investment. Thus, in 1975 foreigners borrowed from current U.S. saving, rather than lending out of their own saving. Government in the United States also borrowed to finance a deficit in 1975; consequently, out of private saving amounting to 19 percent of GNP, only 13.7 percent was left for private investment within the United States (after adjustment for problems of measurement, known as the statistical discrepancy).

In 1982 the government deficit increased sharply in relation to GNP and then remained large by historical standards through the remainder of the 1980s (especially through 1986). Over the same period, private saving as a percentage of GNP declined, rather than rising to compensate for the greater government dissaving. Thus, from 1982 through 1987 private investment was increasingly financed by U.S. borrowing from abroad, as can be seen in Table 11. Such borrowing from abroad allowed total U.S. demand, or spending, to increase faster than U.S. output.

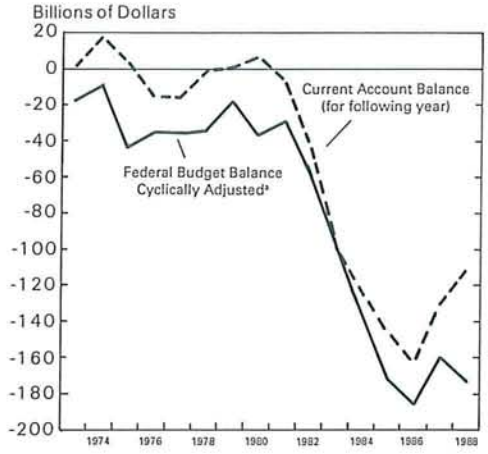
This net borrowing from abroad, it should be noted, is essentially the same as the current-account deficit in the U.S. balance of payments. Although the government deficit and the current-account deficit are thus related in an accounting sense, the sizes of the two deficits can still vary independently of each other, and on occasion inverse variation takes place. For example, from 1971 to 1972 the government deficit diminished while net borrowing from foreigners (the current-account deficit) expanded.

From 1982 through 1987 private investment was increasingly financed by U.S. borrowing from abroad, allowing U.S. demand to increase faster than U.S. output.

Such inverse variation is not likely when the government deficit changes dramatically in response to a change in government policy, as was the case in the United States in 1982, a year when a major federal tax reduction began to take effect even as spending on federal programs was accelerating. While views differ regarding the short-run impact of government

Chart 2

The Current Account and Federal Budget Deficits of the United States



a) Based on 6 percent unemployment GNP trend.

Source: *Survey of Current Business*, various issues and staff of the U.S. Bureau of Economic Analysis.

deficits, the dominant theory is that such a policy-induced surge in government borrowing in a country will put upward pressure on interest rates (adjusted for expected inflation) in that country, thereby attracting foreign investment. As foreign investors acquire the country's currency in order to invest there, they bid up the price of that currency in the foreign-exchange markets. The higher price of the country's currency will discourage foreigners from purchasing its goods but will encourage residents of the country to use their now more valuable currency to purchase foreign goods, so that the country's current account will move toward deficit (or toward a larger deficit). In addition, any increase in the country's total spending resulting from the enlarged government deficit will go partly for imports and for domestic goods that would otherwise be exported, also worsening the current-account balance. Again, to return to one of our central themes, we can see from this brief description that the deterioration of the current-account balance is associated with an increase in the country's total demand relative to the country's output.

Chart 2 supports the view that an increase in the government deficit tends to increase the current-account deficit at least over the medium run.[7] The

government deficit represented in this chart has been adjusted to exclude the effects of the business cycle; for example, any declines in tax revenues occurring because of recessions have been added back to the recorded level of government receipts, reducing the recorded deficit. Such adjustments are warranted because our interest is in deficits that tend to add to the preexisting level of borrowing and spending, rather than in deficits that merely offset a decline in aggregate borrowing and spending elsewhere in the economy. Since cyclically adjusted data are not available for state and local government deficits, Chart 2 uses data for the federal deficit, which has been the focus of concern. Also, the federal deficit for each year is matched with the current-account deficit for the following year, on the assumption that some time is required for an increase in the federal deficit to influence the current-account deficit.

As noted above, a change in the federal deficit is presumed to affect the current-account deficit partly through its impact on the dollar price of foreign exchange. Chart 3 suggests that the hypothesized relationship between the government deficit and the exchange rate did indeed prevail over the period 1973–85, although the relationship is rather loose. In this case, the government deficit for each year is paired with the exchange rate for the same year, with no lag, on the common assumption that exchange rates react promptly to stimuli, or even anticipate them (but then affect the current account with a lag). Also, the dollar price of foreign exchange, rather than the foreign-exchange price of the dollar, is plotted; therefore, a downward movement signifies appreciation of the dollar.

Although Charts 2 and 3 are suggestive, strong conclusions should not be drawn from them alone. The exchange rate and the current account are influenced not just by the government deficit but by other factors as well.[8] Other factors likely to have contributed significantly to the dollar's appreciation during the early 1980s—and thus to the current-account deficit—were an anti-inflationary U.S. monetary policy, U.S. tax law changes and deregulation that enhanced the after-tax profitability of investing in the United States, the easing of restrictions over capital outflows from Japan, and more restrictive government fiscal policies in some major foreign industrial countries (McCulloch and Richardson 1986, pp. 56–57; Feldstein 1985, p. 7; Helkie and Hooper 1988, Table 2-17). The net flow of capital into the United States was also fostered by the loss of investment appeal on the part of the less developed countries

Chart 3

The Real Exchange Rate and the Federal Budget Deficit of the United States, 1973-89

Source: *World Financial Markets*, various issues and staff at Morgan Guaranty Trust; *Survey of Current Business*, March 1990 p. 19 and staff at the U.S. Bureau of Economic Analysis.

that could not meet interest payments on their debt.

If net capital flows into the United States became so large during the 1980s, what form did these inflows take? As shown in Table 12, privately owned capital generally accounted for the great bulk of the inflows; foreign net purchases of U.S. securities, foreign direct investment, and inflows through U.S. banks all made substantial contributions. From 1986 through 1988, the private inflows were substantially augmented by inflows of officially owned capital, as 128 billion of dollar holdings in the United States were acquired by foreign monetary authorities, some of whom had sold their own currencies in exchange for dollars in an effort to limit their currencies' appreciation in the foreign-exchange markets.

If in this section we have correctly identified the leading causes of the U.S. external deficits, we are confronted with a puzzle: if the increase in the deficits was due primarily to the appreciation of the dollar and the relatively rapid growth of U.S. total demand (as immediate causes) during the first half of the 1980s, why in subsequent years did the deficits decrease so little (Table 4) as the dollar depreciated so greatly (Chart 1) and as U.S. demand grew more slowly (Table 8)?

Table 12
Capital Transactions in the U.S. Balance of Payments, 1980–89
Billions of Dollars

Type of transaction	1981	1982	1983	1984	1985	1986	1987	1988	1989[p]
Private capital, net	−22.6	−23.2	29.1	77.0	110.1	95.7	100.2	98.7	102.9
Securities, net	4.1	5.1	10.1	30.8	63.9	70.5	29.2	38.7	47.6
Foreign net purchases	9.8	13.1	16.9	35.6	71.4	74.8	34.5	46.6	69.5
U.S. Treasuries	2.9	7.0	8.7	23.0	20.4	3.8	−7.6	20.2	30.0
U.S. corporate bonds	2.1	2.8	2.2	13.9	46.6	53.8	26.5	26.8	33.0
U.S. corporate stocks	4.8	3.3	6.0	−1.3	4.3	17.2	15.6	−0.5	6.6
U.S. net purchases of foreign securities	−5.7	−8.0	−6.8	−4.8	−7.5	−4.3	−5.3	−7.8	−21.9
Direct investment, net	15.6	12.8	5.3	13.8	5.9	15.4	15.8	42.2	40.5
Foreign direct in U.S.	25.2	13.8	11.9	25.4	19.0	34.1	46.9	58.4	72.2
U.S. direct investment abroad	−9.6	−1.0[a]	−6.7	−11.6	−13.2	−18.7	−31.0	−16.2	−31.7
Net flows reported by U.S. banks	−42.0	−45.4	20.4	22.7	39.7	19.8	46.9	13.9	10.5
Other	−0.2	4.2	−6.6	9.7	0.6	−10.0	8.2	3.8	4.3
Official capital, net	−5.3	−7.5	−0.4	−5.5	−7.8	33.9	55.4	38.6	−15.3
Total reported capital flows, net	−27.9	−30.7	28.8	71.6	102.3	129.7	155.5	137.3	87.6
Statistical discrepancy	19.9	36.6	11.4	27.5	20.0	15.8	6.8	−8.4	22.4
Current-account balance	6.9	−5.9	−40.1	−99.0	−122.3	−145.4	−162.3	−128.9	−110.0

Note: Minus sign indicates an outflow.
[p]Preliminary.
[a]Break in series.
Source: *Survey of Current Business*, vol. 69, June 1989, p. 79; vol. 70, June 1990, pp. 72, 76, 77 and 97; and U.S. Department of Commerce staff.

Various explanations have been advanced. One of the more plausible is that responses to the dollar's depreciation are taking longer than did the responses to the appreciation. In particular, some foreign exporters, having just invested in gaining a larger share of the U.S. market in response to the appreciation, may have been loath to give up their market gains immediately and instead may have been sacrificing profits in order to retain most of their price competitiveness for the time being in spite of the dollar's depreciation. Another explanation is that factors other than changes in exchange rates and in total demand have an appreciable influence on the external deficit. For example, some studies have concluded that demand for imports grows relatively faster in response to income growth in the United States than it does in the rest of the world, so that the U.S. trade balance will deteriorate unless U.S. income grows much slower than income abroad, other things equal (Houthakker and Magee 1969). Some more recent research maintains that such a difference in the "income elasticity of demand for imports," if it exists at all, is not so important in sustaining the U.S. trade deficit as another factor, namely, the continuing introduction by foreign producers of new product lines that capture the fancy of American consumers (Helkie and Hooper 1988, pp. 20–23).

In any event, the U.S. current-account deficit remains large, and U.S. external indebtedness continues to mount. Having considered the causes of the U.S. deficit, we turn now to the consequences. These can be divided into two categories: past and future. Our chief concern is with consequences for the United States, rather than for the rest of the world.

The Consequences to Date

To some observers, the large U.S. external deficits connote something more alarming than reduced U.S. competitiveness; they connote the "deindustrialization" of America. According to this school, U.S. manufacturing not only has lost ground in export markets, but has been in retreat before a flood of competing imports. Strong action has, therefore, been recommended to preserve the viability of domestic industry.

To be sure, total U.S. output and employment would have been higher if exports had been greater, or competing imports smaller, other things being equal. But once this point is granted, how did the U.S. economy in fact perform under the intensified foreign competition?

The answer is perhaps best conveyed by aggre-

gative data on the production of goods and services. Because goods are generally more transportable than services, firms that produce goods usually are subject to more foreign competition than are firms that produce services. According to the data in Table 13, the production of U.S. goods grew faster than either U.S. GNP or the production of U.S. services after the 1981–82 recession, and the growth of goods production over this period compares favorably with that during previous recent expansions. By this measure, then, U.S. goods producers did well, even with the heightened foreign competition; evidently, the growth of total U.S. demand was rapid enough to accommodate a substantial rise in U.S. production as well as in U.S. imports. (During recessions, of course, goods output actually declines, while services output usually continues to grow.) While it would be a mistake to describe the performance of U.S. industry in superlatives, it is an even greater exaggeration to speak of the deindustrialization of America.

Although the U.S. external deficits have not been destroying American industry, they have exercised a moderating influence. As already noted, U.S. output would have grown even more rapidly in the absence of those deficits; the result might well have been an overheated economy, with appreciably higher inflation and interest rates.

This conclusion is supported by the data in Table 8. The growth rate of U.S. real domestic demand in the years 1983 through 1986 was high by historical standards—and extraordinarily high in 1983 and 1984, the two years of greatest increase in the U.S. trade and current-account deficits. Indeed, the 8.7 percent growth in U.S. domestic demand in 1984 was the highest since 1951. Had the United States been unable to acquire additional goods and services from abroad to help satisfy this surging demand, the nation could have experienced "bottlenecks"—if not more general shortages—as well as an acceleration of inflation. Even with the huge increase in its net imports, the U.S. economy expanded its output in 1984 by 6.8 percent, which was, again, the fastest rate of growth since 1951.

Partly because of the availability of imported goods, this rapid expansion took place without any rise in overall inflation as measured by the GNP deflator. To be sure, the nation's high rate of unemployment—7.4 percent in 1984—also militated against rising inflation. But the unemployment rate did fall steadily after 1983, and by end-1986, according to some authorities, was at or near the level at which it would no longer serve to restrain inflation (Wachter 1986, pp. 390–91). This level would have been reached much sooner without the increase in net imports.

It also seems clear that interest rates would have been higher in the United States had the nation been foreclosed from borrowing abroad. As reported in Table 11, gross private saving in the United States declined markedly as a share of GNP from 1981 to 1989 even though government dissaving during the 1980s was at unusually high levels; the nation stepped up its foreign borrowing to help offset these developments. Without the availability of foreign

Table 13
Percentage Changes in U.S. Real Output During Recessions and the Succeeding Expansions, 1969–90

Period (year and quarter)	Percentage Change			
	Goods	Services	Structures	Total GNP
1969:4 to 1970:4 (Recession)	−3.3	1.6	2.4	−.4
1970:4 to 1973:4 (Expansion)	17.6	12.0	12.7	14.5
1973:4 to 1975:1 (Recession)	−7.8	3.0	−18.9	−4.3
1975:1 to 1980:1 (Expansion)	25.2	18.7	28.1	22.4
1980:1 to 1980:3 (Recession)	−3.6	.8	−10.4	−2.3
1980:3 to 1981:3 (Expansion)	5.8	1.3	2.5	3.3
1981:3 to 1982:4 (Recession)	−7.3	1.2	−6.3	−3.2
1982:4 to 1990:1 (Expansion)	41.2	27.3	24.3	32.7

Source: Board of Governors of the Federal Reserve System, Fame Data Base.

financing, U.S. interest rates would have risen so as to choke back the level of private domestic investment to the lower level of financing provided from domestic sources alone. Even with the net inflow of foreign capital, U.S. interest rates, both short-term and long-term, reached record heights in the early 1980s (U.S. Bureau of the Census 1975, pp. 1001–1004; 1986, pp. 492–93).

Thus, the near-term consequences of its external deficits seem to have been rather beneficial for the United States. What about the longer term?

Must the Deficit Be Reduced?

Both common sense and experience testify that neither individuals nor nations can incur debt without regard to ability to repay. But it would be a gross exaggeration to suggest that the United States has been threatened with an imminent debt crisis. By no conventional statistical indicator is the nation in such desperate straits.

In this kind of analysis, it is common to distinguish between liquidity and solvency risk. Although countries do not declare bankruptcy, a country is insolvent if it is unable, either for economic or political reasons, to meet its debt obligations over the long term. By contrast, illiquidity means that a country cannot meet its obligations coming due in the near term, but can discharge those obligations, with accrued interest, in the longer run, along with the rest of its obligations.

To assist in evaluating such risk, analysts have developed various indicators of the burden of international indebtedness. Although these indicators are crude, they can help to signal emerging distress. Some widely used indicators focus on the share of a country's output or income that is owed to its creditors. Others focus on the share of export earnings that is absorbed by payments to creditors, recognizing that some significant fraction of those earnings must remain to pay for imports.

In Table 14 are data for some of these indicators that were readily available for a sample of diverse countries. At the end of 1989, gross external debt as a percentage of GNP was lower for the United States than for any of the other countries; no alarm was being sounded by this indicator. Nor was gross external debt as a percentage of exports high by comparison with the typical country listed. Somewhat less reassuring was the percentage of U.S. export receipts consumed by debt-service payments to foreign creditors. On balance, one could hardly make the case from such indicators that the United States was facing a debt crisis, especially since the indicators fail to take into account the relatively large foreign assets held by U.S. residents.

Table 14
Selected Debt Burden Indicators for the United States and Six Other Countries, 1989

	Gross External Debt as Percent of		Debt Service on Gross External Debt as Percent of Exports of Goods, Services and Private Transfers
	GNP or GDP	Exports of Goods, Services, and Private Transfers	
United States	14.1	123.3	72.0
Canada	39.8	144.7	55.5
West Germany	25.0	68.9	45.0
Denmark	72.0	175.0	117.5
Argentina	117.0	506.4	89.8
Brazil	38.9	304.3	103.2
Mexico	54.4	269.5	90.3

Note: Data are partly estimated.
Source: Morgan Guaranty Trust Co.

One key difference between the United States and the countries that have suffered debt-repayment problems in recent years is that the great bulk of U.S. external debt has been denominated in U.S. rather than foreign currency. Unlike debtors in these other countries, U.S. debtors generally have not had to acquire foreign exchange with which to service their external debts. Were this practice to continue, U.S. debtors would be unlikely to experience more difficulty in meeting their external obligations than in meeting their obligations to domestic creditors. In other words, any debt crisis encountered by the United States would be a general crisis, imperiling resident as well as foreign creditors, rather than an exclusively international crisis. Only a most unlikely development, such as systematic lending by foreigners to unsound U.S. businesses, or limitations by the U.S. government on U.S. payments to foreign creditors, would generate a peculiarly international problem. On the other hand, should U.S. external debt come to be denominated largely in foreign currencies, a depreciation of the dollar against those currencies

would, of course, increase the number of dollars that U.S. debtors were obliged to repay; and a sharp depreciation could provoke a debt crisis that was initially concentrated in the international sector.

The fact that U.S. external debt is owed overwhelmingly in dollars does not mean that the debt imposes no burden, nor does the improbability of an external debt crisis mean that the United States could continue along the path set upon a few years ago. Even if a nation were inclined to borrow without limit, others generally would not lend to it beyond its perceived capacity to service its debt. The moral for the United States is that it would be unable to continue incurring such relatively large current-account deficits in the long run, even if its government deficit remained large.

The long run, however, could be rather long. Some elementary computations are illuminating. In 1987 the U.S. current-account deficit attained a peak of 3.6 percent of U.S. GNP, with GNP at $4.5 trillion. Suppose that the current-account deficit were to run at 3.5 percent of GNP, and that nominal GNP were to increase by 6 percent each year, a fairly modest rate by recent historical standards. Also suppose that the average interest rate, or more generally, the average rate of return, earned by foreigners investing in the United States were 8 percent per annum. Finally, since the United States reportedly received nearly as much in interest and other income payments from foreigners as it made to them in 1989 (even though the data showed the nation then to be a sizable net debtor), we shall suppose that the United States did not in fact become a net debtor until this writing.

On these assumptions, the fourth column of Table 15 shows how net interest earned by foreigners would rise as a percentage of U.S. GNP over a fifty-year period. After a half-century, this interest burden would amount to about 4.7 percent of U.S. GNP. More likely, if U.S. net debt did rise markedly in relation to GNP, foreigners would demand higher interest rates to compensate for the reduced creditworthiness of the nation (for the seemingly greater risk of lending to it). Thus, the percentages in the fourth column would climb initially at a faster pace than shown, then at a slower pace as foreigners became reluctant to extend additional loans. Indeed, to contemplate a net foreign debt for the United States amounting to more than half its GNP, as this

Table 15
Net U.S. Interest Burden from External Debt under Differing Assumptions
Billions of Dollars, unless otherwise noted

		Annual Current Account Deficit Assumed to Be 3.5% of GNP			Annual Current Account Deficit Assumed to Be $150 Billion		
Year	Nominal GNP (1)	Cumulative Current Account Deficit (2)	Net Interest on Cumulative Deficit (3)	Net Interest as Percent of GNP (4)	Cumulative Current Account Deficit (5)	Net Interest on Cumulative Deficit (6)	Net Interest as Percent of GNP (7)
1	$ 4,500.0	$ 157.5	$ 12.6	.28	$ 150	$ 12	.27
2	4,770.0	324.5	26.0	.55	300	24	.50
3	5,056.2	501.4	40.1	.79	450	36	.71
4	5,359.6	689.0	55.1	1.03	600	48	.90
5	5,681.1	887.8	71.0	1.25	750	60	1.06
6	6,022.0	1,098.6	87.9	1.46	900	72	1.20
7	6,383.3	1,322.0	105.8	1.66	1,050	84	1.32
8	6,766.3	1,558.8	124.7	1.84	1,200	96	1.42
9	7,172.3	1,809.9	144.8	2.02	1,350	108	1.51
10	7,602.7	2,076.0	166.1	2.18	1,500	120	1.58
20	13,615.2	5,793.7	463.5	3.40	3,000	240	1.76
30	24,382.7	12,451.7	996.1	4.09	4,500	360	1.48
40	43,666.7	24,375.0	1,950.0	4.47	6,000	480	1.10
50	78,198.7	45,727.9	3,658.2	4.68	7,500	600	.77

Note: Nominal GNP is assumed to increase by 6 percent annually. Interest rate is assumed to be 8 percent and is applied to the net debt outstanding at the end of each period, which is taken to be zero prior to year one.

scenario does, might seem beyond the realm of reason. As the data in Table 14 suggest, such ratios did obtain for gross (and presumably net) external debt at the end of 1988 for some countries, but most were less developed, with much smaller economies, than the United States. Moreover, net interest ratios approaching the highest levels shown in column 4 of Table 15 would likely translate into something like two-fifths of U.S. exports of goods and services. In any event, our calculations are merely illustrations, not forecasts.

An interesting alternative is to assume that the U.S. current-account deficit did not rise with GNP but continued at an annual rate of about $150 billion (a rate somewhat exceeded during mid-1987), and to retain the other assumptions underlying the preceding computations. In this case, the net interest burden as a percentage of GNP would move upward for many years, as shown by the last column in Table 14, but would then decline (beginning with the eighteenth year, not shown in the table). This scenario seems much less threatening.

Variations in the underlying assumptions would, of course, yield different hypothetical outcomes. What seems clear from the calculations presented is that the U.S. current-account deficit could not remain so high *in relation to GNP* as it was in 1987. And in fact the deficit has decreased not only in relation to GNP but in absolute magnitude, amounting to $110 billion, or 2.1 percent of GNP, in 1989. It is not so immediately obvious that the deficit must be reduced still further. As illustrated in Table 15, even with a continuing annual deficit as great as $150 billion, U.S. GNP presumably would eventually increase more rapidly than U.S. net indebtedness (the accumulated deficit), so that the net interest burden would begin to decline in relation to GNP well before reaching the level of 2 percent.

It is not obvious that the U.S. current-account deficit must be reduced still further.

To suggest that the United States might continue to incur a sizable current-account deficit is not to imply that the nation can avoid any adjustment in its external accounts. As U.S. net interest payments to foreigners increase with U.S. net indebtedness, the nation will have to generate increasing net surpluses (or smaller net deficits) on other current-account transactions—essentially merchandise trade—in order to limit expansion of the overall current-account deficit. How this adjustment takes place is the topic of the next section.

The Nature of the Adjustment

The point has been made that total U.S. demand, or spending, increased faster than U.S. output during most of the 1980s, and that the nation is absorbing foreign saving to finance the gap. To reduce the imbalance, or to prevent it from rising as interest payments to foreigners go up, the United States must raise the growth rate of its output or reduce the growth rate of its spending. As can be seen in Table 8, the nation did succeed in lowering the growth rate of demand relative to that of GNP during the latter part of the 1980s. However, the rate of inflation, as measured by the GNP deflator, rose from 2½ percent in 1986 to 4¼ percent in 1989, implying that the nation's productive capacity was being strained beyond the point at which prices could be held relatively stable.

For the future, reducing the current-account deficit while restraining inflation will require that total demand grow more slowly than in recent years. Efforts to sustain output growth at the rate of the mid-to-late 1980s would court a marked rise in the rate of inflation. Of course, measures that raised output by raising productivity would not invite higher inflation. But raising the productivity of capital would tend to attract more investment from abroad, and as we have seen, investment from abroad works to enlarge rather than diminish the current-account deficit. Therefore, policies designed to raise the growth rate of output probably hold little promise for shrinking the external imbalances of the United States.

The alternative course, restraining total spending, is now underway. Slowing the growth of *consumption* spending, private or government, would, of course, be equivalent to accelerating the pace of saving, unless output growth slowed to the same degree. Alternatively, if the course of saving were left unchanged, the economy could cut back on the growth of its private *investment* spending. Cutting back on investment in plant and equipment, how-

ever, would reduce the nation's future output.

The government might well step up the rate of saving by contracting the budget deficit, either by cutting back its own spending programs or by raising taxes so that households would lower their consumption spending.[9] If reduction of the budget deficit—and of the economic stimulus the deficit provides—took place at a moderate pace, a recession need not ensue, since a goal of the deficit reduction would be to allow U.S. net exports, another stimulus, to expand more rapidly. One way that such deficit reduction could boost U.S. net exports would be by generating a depreciation of the dollar's foreign-exchange value, just as enlargement of the deficit had generated an appreciation.

Cutting the federal budget deficit seems a relatively appealing strategy for cutting the international trade deficit.

Thus, if U.S. spending must be constrained, cutting the federal budget deficit seems a relatively appealing strategy for cutting the international trade deficit. But are alternative or supplementary strategies available that do not rely on such direct attacks on spending? This economist is tempted to reply that there is no free lunch.

One alternative government strategy would be to do nothing at all—to take no action designed specifically to shrink the U.S. trade deficit, even in relation to GNP. As U.S. indebtedness mounted in relation to U.S. exports and GNP, investors would become more reluctant to lend to, or acquire net claims on, the United States, thus putting upward pressure on U.S. interest rates and downward pressure on the foreign-exchange value of the dollar. Indeed, this process seemed to be under way late in 1986 and at times during 1987, as some U.S. interest rates rose sharply in relation to rates in some other industrial countries even as the dollar dropped in value (Chart 1) against the currencies of those countries. The higher U.S. interest rates would discourage U.S. builders and other businesses from investing in new structures and equipment, and this reduced spending would help to improve the U.S. trade balance, albeit at the expense of future U.S. growth. Trade balance improvement would also be fostered by the depreciation of the dollar.

Just how dollar depreciation improves the U.S. trade balance is a matter of some debate. One conceivable route is via a reduction in the purchasing power of U.S. money balances. A rise in the dollar price of foreign currency (dollar depreciation) tends to raise the dollar prices of foreign goods imported into the United States, as well as the prices of substitute goods produced within the country. Thus, the purchasing power of U.S. residents could be somewhat diminished, discouraging spending and improving the balance of trade.

Dollar depreciation typically has another related price effect that also is helpful. The depreciation-induced rise in the dollar price of imports, and of exports, encourages U.S. businesses to shift resources into the production of export goods and of goods that can substitute for imports, and away from the production of goods that do not move in international trade. The same price movements encourage U.S. consumers to switch their purchases away from the goods that move in international trade and toward nontraded goods. Again, the tendency is to improve the trade balance. And if the prices of nontraded goods decline, or rise more slowly than before the depreciation, the nation need not experience a marked rise in its overall rate of inflation.

Still another government strategy to reduce the trade deficit would be protectionism. Now, the U.S. trade deficit has been very large, and any U.S. import tariffs or quotas severe enough to have a sizable initial impact on the deficit would certainly provoke foreign retaliation against U.S. exports. Even in the absence of retaliation, tariffs or quotas would not be very effective in decreasing the trade deficit unless they somehow reduced total U.S. spending. A tariff could reduce spending if the tariff revenue were used by the government to cut back on its budget deficit, but other taxes would offer the same opportunity without the cost of an international trade war. Protectionism, therefore, is not a promising approach to the problem.

Of the various strategies considered, then, the most desirable would be a combination of federal deficit reduction and tolerance of dollar depreciation. The adjustment process under way thus far has not been ideal. At this writing, significant federal deficit reduction is problematical and interest rates remain at levels relatively high by historical standards, tending to depress private investment. In this connection,

Table 11 indicates that the unusually high level of U.S. borrowing from abroad after 1982 was not accompanied by an unusually high level of private domestic investment. The implication is that the increased borrowing from abroad went mainly or entirely to finance increased consumption. Unlike sound investment, consumption generates no return with which to repay a loan. Thus to service its foreign debt, the United States will have to consume less than it otherwise would.

Summary

After 1982 the U.S. international investment position dramatically shifted from one of sizable net creditor to much more sizable net debtor, with further huge, debt-augmenting deficits in the offing. This transformation occurred even though the United States may have lost little or no competitiveness for "supply-side" reasons. In particular, during the years of rapid deterioration in the U.S. trade balance, U.S. labor productivity gains were virtually as great as those in other industrial nations; the performance of U.S.-based multinational firms suggests that U.S. management was maintaining its international competitiveness; and the United States did in fact maintain its share of world output.

Nor can unfair foreign trading practices explain much of the U.S. external deficits. The deterioration in the U.S. trade balance was distributed widely across commodity categories, as well as across geographic areas. It seems most unlikely that virtually all major trading partners of the United States would simultaneously have intensified unfair practices in trade with the United States in virtually all major commodity categories.

A more plausible explanation of the U.S. external deficit focuses on (1) the more rapid expansion of real income and demand in the United States than in the rest of the world after 1982, and (2) the appreciation of the dollar in the foreign-exchange markets after mid-1980, a development that reduced the price competitiveness of U.S. goods. Both of these factors stimulated greater growth in U.S. purchases of foreign goods than in foreign purchases of U.S. goods; both factors were themselves a result largely of the worldwide blend of monetary and fiscal policies, including the huge increase in the U.S. federal budget deficit. This increase in net federal spending boosted aggregate U.S. demand. Moreover, the increase in U.S. government borrowing associated with the budget deficit, coupled with an anti-inflationary U.S. monetary policy, tended to push up U.S. interest rates (adjusted for inflation), thus attracting investment by foreigners, whose purchases of dollar-denominated securities served to bid up the value of the dollar in the foreign-exchange markets.

Contrary to a widespread impression, the U.S. trade deficits were not accompanied by a "deindustrialization of America." Following the 1981–82 recession, the production of goods grew faster than the production of services within the United States, and the growth of goods production compared favorably with that during earlier economic expansions. Thus, U.S. goods producers fared relatively well despite the increased U.S. trade deficit. Rather than destroying large segments of American industry, imports from abroad helped to satisfy the swiftly growing U.S. demand, without the development of shortages and rising inflation.

Although the net foreign debt of the United States soared with the trade deficit, no crisis looms for the nation on indebtedness. Over the longer run, of course, foreigners would not be prepared to lend more and more to a nation whose indebtedness continued to rise in relation to its gross output and exports. Thus, the U.S. current-account deficit had to shrink in relation to the nation's output and exports; and the trade deficit in particular must diminish if the nation is to fund increasing net interest payments to its foreign creditors. The depreciation of the dollar that took place after February 1985 will contribute to this adjustment, as would further dollar depreciation and measures to reduce the federal budget deficit. The adjustment will not be painless for the United States, which will be obliged to consume less than it otherwise would.

[1] Perhaps the most articulate exponent of the hard landing scenario was Stephen Marris (1985 and 1987).

[2] While some part of the large unidentified receipts in the U.S. balance of payments has surely taken the form of foreign investment in the United States, it would almost certainly be a mistake to attribute all of these net receipts to such capital-account transactions. Much evidence exists that a significant portion of the receipts has been generated not by capital-account but by current-account transactions, such as the sale of U.S. goods and services abroad, or the charging of interest on U.S. loans to foreigners. Insofar as the unidentified receipts have resulted from current-account transactions, the reported value of U.S. net indebtedness requires no upward revision.

[3] It should also be noted that the measure of the U.S. position in Tables 1, 2, and 3 includes equity as well as debt claims, while the customary measures of the debt of less developed countries do not include equity held by foreigners. For a discussion of these matters, see Herman (1987, pp. 1–4).

[4] See Herman (1987, pp. 1–2); also see the IMF *World Economic Outlook* (latest issue) for more recent data than Herman supplies on developing-country debt.

[5] See, for example, U.S. Department of Commerce (1983, pp. 49–51), and *Consumer Reports*, various issues evaluating automobiles.

[6] For an outline of U.S. law and procedures relating to the imposition of antidumping duties and countervailing duties, see U.S. International Trade Commission (1989, pp. 140–42).

[7] After surveying the estimates yielded by a large number of multicountry models, Helliwell (1990, p. 17) concludes that "the U.S. fiscal policy of the first half of the 1980s was responsible for about half of the buildup in the external deficit. . . ."

[8] Among these factors are resource discoveries, changes in tastes and technology, and differences in national growth rates. Changes in tastes and in technology, however, as well as growth trend differentials, generally exert their influence gradually over long periods, and major resource discoveries are rare. From year to year, movements in the real exchange rate and current account are more powerfully influenced by business cycle fluctuations, by government controls, and by government monetary and fiscal policy, including changes in the government deficit such as those depicted in Charts 2 and 3.

[9] Considerable controversy exists within the economics profession over the impact of government revenues and spending on aggregate demand. For example, see Feldstein and Elmendorf (1990), Modigliani and Sterling (1990), and Kormendi and Meguire (1990).

References

Bergsten, C. Fred and William R. Cline. 1985. *The United States–Japan Economic Problem.* Policy Analyses in International Economics No. 13. Washington, D.C.: Institute for International Economics.

Bergstrand, Jeffrey H. 1986. "United States–Japanese Trade: Predictions Using Selected Economic Models." *New England Economic Review*, May/June, pp. 26–37.

Christelow, Dorothy. 1985–86. "Japan's Intangible Barriers to Trade in Manufactures." *Federal Reserve Bank of New York Quarterly Review*, Winter, pp. 11–18.

Deutsche Bundesbank. 1986. "External Assets of the Federal Republic of Germany in Mid-1986." *Monthly Report of the Deutsche Bundesbank*, vol. 38, October, pp. 30–46.

Feldstein, Martin S. 1985. "The View from North America." In *Global Economic Imbalances*, C. Fred Bergsten, ed. Special Reports No. 4. Washington, D.C.: Institute for International Economics.

Feldstein, Martin S. and Douglas W. Elmendorf. 1990. "Government Debt, Government Spending, and Private Sector Behavior Revisited: Comment." *The American Economic Review*, vol. 80, June, pp. 589–99.

Helkie, William L. and Peter Hooper. 1988. "An Empirical Analysis of the External Deficit, 1980–86." In *External Deficits and the Dollar: The Pit and the Pendulum*, Ralph C. Bryant, Gerald Holtham, and Peter Hooper, eds. Washington, D.C.: The Brookings Institution.

Helliwell, John F. 1990. "Fiscal Policy and the External Deficit: Siblings but Not Twins." NBER Working Paper No. 3313, April. Cambridge, MA: National Bureau of Economic Research.

Herman, Barry. 1987. "The United States as a Debtor Country: Indicators of Resource Transfer and Solvency." Working Paper No. 2, U.N. Department of International Economic and Social Affairs, February.

Hooper, Peter and Catherine L. Mann. 1987. "The U.S. External Deficit: Its Causes and Persistence." International Finance Discussion Papers No. 316. Washington, D.C.: Board of Governors of the Federal Reserve System.

Houthakker, H.S. and Stephen P. Magee. 1969. "Income and Price Elasticities in World Trade." *The Review of Economics and Statistics*, vol. 51, May, pp. 111–25.

International Monetary Fund. 1988. *World Economic Outlook*, April 1988. World Economic and Financial Surveys. Washington, D.C.: IMF.

Kormendi, Roger C. and Philip Meguire. 1990. "Government Debt, Government Spending, and Private Sector Behavior: Reply and Update." *The American Economic Review*, vol. 80, June, pp. 604–17.

Marris, Stephen. 1985. *Deficits and the Dollar: The World Economy at Risk.* Washington, D.C.: Institute for International Economics.

———. 1987. *Deficits and the Dollar Revisited.* Washington, D.C.: Institute for International Economics.

McCulloch, Rachel and J. David Richardson. 1986. "U.S. Trade and the Dollar: Evaluating Current Policy Options." In *Current U.S. Trade Policy: Analysis, Agendas, and Administration*, Robert E. Baldwin and J. David Richardson, eds. NBER Conference Report. Cambridge, MA: National Bureau of Economic Research.

Modigliani, Franco and Arlie G. Sterling. 1990. "Government Debt, Government Spending, and Private Sector Behavior: A Further Comment." *The American Economic Review*, vol. 80, June, pp. 600–603.

U.S. Bureau of the Census. 1975. *Historical Statistics of the United States, Colonial Times to 1970, Part II.* Washington, D.C.: Government Printing Office.

———. 1986. *Statistical Abstract of the United States: 1987.* Washington, D.C.: Government Printing Office.

U.S. Department of Commerce. 1983. *The U.S. Automobile Industry, 1982.* Washington, D.C.: USDC.

U.S. International Trade Commission. 1988. *Operation of the Trade Agreements Program, 1987.* Publication No. 2095. Washington, D.C.: USITC.

———. 1989. *Operation of the Trade Agreements Program, 1988.* Publication No. 2208. Washington, D.C.: USITC.

U.S. President. 1985. *Economic Report of the President.* February. Washington, D.C.: Government Printing Office.

Wachter, Michael L. 1986. "Comment on Lawrence H. Summers's 'Why Is the Unemployment Rate So Very High Near Full Employment?'" *Brookings Papers on Economic Activity 1986:2*, pp. 390–94.

New England Banks and the Texas Experience

Katerina Simons

Economist, Federal Reserve Bank of Boston. *The author thanks Thomas Cimeno, Lynn Browne, Richard Kopcke, Richard Randall, Eric Rosengren, and Steven Sass for helpful comments. Lawrence D. Herman provided valuable research assistance.*

Are New England banks going the way of Texas banks? The prospect that they might be is indeed alarming, given the magnitude of the collapse of the Texas banking industry. Of Texas's nine largest bank holding companies, all but one have either merged with an out-of-state organization or sought federal assistance. The loss absorbed by the Federal Deposit Insurance Corporation has thus far amounted to $8 billion, or about 44 percent of the 1987 balance of the FDIC fund.

New England banks are currently suffering from problems similar to those that caused the demise of many Texas banks. Nevertheless, this article concludes that the comparison of New England to Texas is unwarranted. Section I compares banking statistics of Texas and New England. Particular attention is paid to the influence of real-estate lending on bank performance. Section II relates the fortunes of the banking industry to the underlying health of the regional economy. It describes the leading sectors of the two economies and compares the real-estate cycles in Texas and New England. Section III concludes with a discussion of the future course of banking in New England.

I. Real Estate Lending and Bank Performance

The financial losses of Texas banks were largely due to the poor credit quality of their commercial real-estate portfolios. Texas banks lent heavily to developers constructing office buildings, strip shopping centers, shopping malls, industrial parks and warehouses—projects that suffered from serious overbuilding. As a result, in 1985, real-estate loans[1] accounted for 25 percent of nonperforming assets in Texas banks. Loans are classified as nonperforming when they are at least 90 days past due, nonaccruing, or renegotiated. Nonperforming assets are defined as nonperforming loans plus other real estate owned (OREO),

which consists primarily of foreclosed real estate. From 1985 to 1989, nonperforming real-estate assets in Texas banks had increased from $1 billion to $7 billion and accounted for almost 76 percent of nonperforming assets. As a result of these loan losses, from 1986 to 1989, Texas banks lost over $6 billion.

Nonperforming assets are an important indicator of the health of the banking system because they are closely related to the future level of loan losses; they also reduce interest receipts and equity capital. Nonperforming assets increased dramatically for New England banks in the first quarter of 1990, reflecting the deterioration in quality of banks' loan portfolios. As in Texas, the problems are mostly due to the declining value of real estate.

Texas banks reached a high point in their real-estate loan concentrations in the mid-1980s. By that time, the decline in real-estate values had already occurred. The Texas banks were slow, however, to recognize the problems of their borrowers and the declines in collateral values on real-estate loans they had already approved. Thus, the ratio of nonperforming assets to total loans did not reach its peak until 1988. At the same time, heavy charge-offs of construction and commercial real-estate loans were taken on the balance sheets of Texas banks.

In New England, on the other hand, banks quickly reported increases in nonperforming loans as real estate declined in value. Therefore, current conditions in New England are more reminiscent of Texas in the mid 1980s in terms of the point reached on the real-estate cycle, and a proper comparison of nonperforming loans and equity levels of New England banks today is better made to those of Texas banks in the late 1980s. Accordingly, charts and tables in this section compare current concentrations of real-estate loans in New England to those of Texas in 1985, and compare current New England nonperforming and equity ratios to those in Texas in 1988.

Nonperforming Assets and Equity Capital in New England and Texas Banks

Chart 1 compares the ratio of nonperforming assets to total loans of Texas commercial banks, at its highest point in 1988, to similar ratios of New England and U.S. commercial banks in the first quarter of 1990, the most recent data available at the time of writing this article. The chart shows that while the ratio of nonperforming assets to total loans is now substantially higher for New England banks than for banks nationwide, the ratio for Texas banks in 1988

Chart 1

Nonperforming Assets as a Percentage of Total Loans and Leases

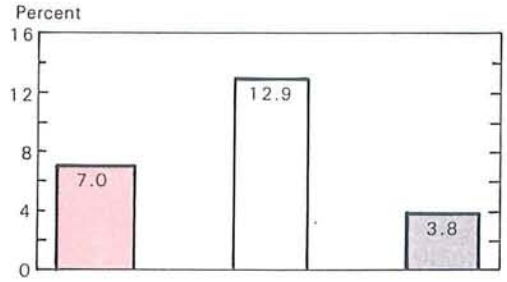


Chart 2

Equity Capital as a Percentage of Total Assets

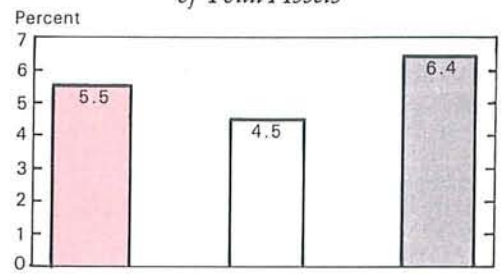
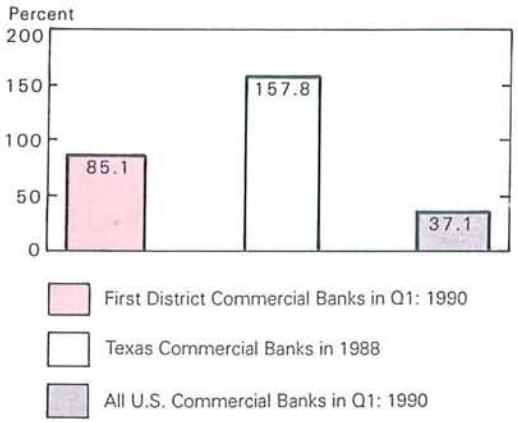


Chart 3

Nonperforming Assets as a Percentage of Equity Capital



Source: Board of Governors of the Federal Reserve System, call report data. Nonperforming assets include nonaccrual loans, loans more than 89 days past due but still accruing and other real estate owned.

was almost twice the current New England ratio. This suggests that the future loan losses of New England banks may not be expected to approach the levels suffered by Texas banks.

A substantial capital cushion is the primary line of defense against the threat of bank failure. Chart 2 compares the ratio of equity capital to total assets for New England banks with those of banks nationwide for the first quarter of 1990 and Texas banks in 1988. While New England banks have a smaller capital cushion than the current national average, it is larger than that of Texas banks in 1988. Chart 3 compares nonperforming assets to equity capital. In New England nonperforming assets are currently 85 percent of equity capital, a much higher level than the national average. In Texas, though, nonperforming assets exceeded equity capital in 1988.

The statistics in the charts describe the condition of New England banks at an isolated moment. They do not necessarily indicate the conditions that will prevail in the future. However, the numbers alone do not suggest that New England banks are headed for a Texas-style collapse. Moreover, the asset composition of banks in New England differs markedly from that of Texas banks, and, more importantly, the composition of industry in New England is more diversified than that of Texas.

Concentrations of Real-Estate Loans in Texas and New England

Table 1 compares the share of real-estate loans in the loan portfolio of Texas commercial banks at the end of 1985 with that of New England banks and U.S. banks in the first quarter of 1990. It shows that New England banks have a substantially higher concentration of real-estate loans (47.7 percent) than all banks in the nation (37.8 percent) and Texas banks in 1985 (34.7 percent). The heavy commitment of New England banks to real-estate lending is often cited as a reason for concern for their future performance.

The risks inherent in real-estate loans vary greatly. Construction and land development loans are commonly considered to be the most risky kind of real-estate lending, while residential mortgages are regarded as the safest. Commercial real-estate lending occupies an intermediate position.

Unfortunately, no national data record the losses associated with different categories of real-estate loans. A study by First Boston Corporation (1990), however, reports estimated losses on real-estate loans in Texas, based on a survey of a number of chief financial officers of Texas banks. According to these estimates, 60 percent of construction loans became nonperforming and one-half of these (that is, 30 percent of total construction loans) were eventually charged off; 20 percent of commercial mortgage loans would become nonperforming, and again, one-half of these would ultimately be charged off. The study did not report on losses on residential mortgages, because these were deemed too minor to be of importance.

Table 1
Real Estate Loans as a Percentage of Total Loans and Leases

Loan Category[a]	First District Commercial Banks, Q1:1990	Texas Commercial Banks, 1985	All U.S. Commercial Banks, Q1:1990
All Real Estate Loans	47.7	34.7	37.8
Construction and Land Development	8.2	14.1	6.6
Commercial Real-Estate Loans	14.4	10.5	10.8
Residential Real-Estate Loans	24.8	9.3	18.4

[a]Categories will not add to total, which includes farm loans.
Source: Board of Governors of the Federal Reserve System, call report data.

Table 1 also compares the composition of current real-estate lending of New England banks to the national average and to Texas banks in 1985. It shows that Texas banks concentrated their real-estate lending in construction and development. In contrast, New England banks have concentrated their real-estate lending in the safest category of real estate lending, residential mortgage loans. New England banks hold 25 percent of their loans in residential mortgages, while Texas banks had only 9 percent of loans in that category. New England banks also made a heavier commitment than Texas banks to commercial real-estate loans.

Although the proportions of nonperforming loans in different categories of real-estate lending are not available, they can be estimated by regression analysis. Table 2 presents estimates of ratios of nonperforming loans for New England banks for construction, commercial and residential real-estate loans in the past two years and the first quarter of 1990. (See the Appendix for a more complete descrip-

Table 2
Nonperforming Real-Estate Loans as a Percentage of All Real-Estate Loans at New England Commercial Banks[a]

	1988	1989	1st Quarter 1990
Construction	4.4	10.2	18.8
Commercial	1.1	2.1	3.3
Residential	.1	.6	.4
Total Nonperforming Real-Estate Loans	1.0	2.2	3.2

[a]Estimated for subcategories by regression analysis.
Source: See the Appendix.

tion of the regressions.) The results show that construction loans have the highest ratio of nonperforming loans among the categories, thus confirming the common wisdom that they are the riskiest type of real-estate loan. The results also show that during this period the proportion of nonperforming loans increased for all categories of real-estate lending. The only exception is residential loans, for which the proportion of nonperforming loans fell in the first quarter of 1990.

Although the emphasis on residential rather than construction lending among New England banks is encouraging, it is not, in and of itself, sufficient to avoid a Texas-style collapse. If the future losses on construction and commercial real-estate loans in New England eventually equal those of Texas, these losses would eliminate nearly one-half of the current equity capital of New England banks—a proportion similar to that in Texas. Because of differences in the regional economies of Texas and New England, however, New England banks are less likely to experience losses on their construction and commercial real-estate loans mirroring the magnitude in Texas.

II. The Regional Economies in Texas and New England

It is generally agreed that the Texas banking crisis was brought on by two distinct but related developments—the decline of oil prices and the subsequent collapse of the real-estate market. New England now appears to suffer from similar problems: a decline in the leading sectors of its economy, namely defense and high technology, and a weakness in real estate. Analogies between the current business conditions in New England and those of Texas in the mid-1980s are drawn with increasing frequency.[2] Nevertheless, important differences can also be found. First this section will compare the leading industries of Texas with those of New England, and then it will discuss potential consequences for the real-estate cycles in the two regions.

Leading Industries in Texas and New England

Both areas experienced substantial and prolonged weakness in industries that are commonly seen as economic drivers—mining and manufacturing in Texas and manufacturing in New England. Table 3 gives the composition of employment in Texas in 1985 and New England in 1988. Mining and manufacturing in Texas and manufacturing in New England both represent approximately 17 percent of employment. (Mining is insignificant in New England.) Mining and manufacturing employment in Texas was at its highest level in 1981. Between 1981 and 1984, the years that could be considered a peak for the Texas economy, employment in these two industries fell 10 percent. In New England, manufacturing employment has fallen roughly 16 percent since its highest point in 1984.

Traditionally, the performance of manufacturing and resource-based industries such as mining has

Table 3
Employment Shares in New England and Texas
Percent

	1988 New England	1985 Texas
Mining	.1	4.1
Manufacturing	17.4	12.1
Construction	5.9	7.1
Transportation and Public Utilities	3.9	5.1
Trade	21.6	22.4
Finance, Insurance	5.5	4.5
Real Estate	2.6	3.6
Services	29.0	23.0
Government	12.3	15.2
Agriculture	1.6	3.8

Note: Shares may not add to 100.0 because of rounding.
Source: U.S. Bureau of Economic Analysis, BE-55 release.

been regarded a critical determinant of regional economic growth. These industries, commonly referred to as "export" industries, serve national and international markets and their expansion is driven by developments outside the region. An expanding manufacturing sector creates a demand for industrial and R&D space; and indirectly through the jobs and income created, the expansion in manufacturing provides a stimulus to housing.

Texas. The Texas experience is consistent with the traditional view of economic growth. Although overall employment continued to grow for several years after employment in manufacturing and mining began to fall, eventually the weakness in these sectors spilled over to the rest of the economy.

The engine of growth in Texas was the booming gas and oil industry.

Several features make the Texas experience unique. The engine of growth in Texas was the booming oil and gas industry, which benefited from the surge in oil prices stemming from the Arab oil embargo in 1973 and the growing power of oil-producing countries to influence world oil prices. The price of Texas crude oil increased from $3 per barrel in 1971 to $35 per barrel in 1981. Such a spectacular rise in the price of oil created an atmosphere of heady optimism in Texas. New wealth was created at an unprecedented rate. Between 1978 and 1981, for example, the value of daily oil production in the state increased from $10 million to $33 million.

The oil boom accelerated the pace of Texas's economic growth and spurred the creation of new jobs in all sectors of the Texas economy. The growth of the oil and gas extraction industries increased the demand for professional, financial, and business services. It also stimulated the development of oil field machinery manufacturing. This industry, in turn, increased the demand for primary and fabricated metal products. The expansion of the energy sector increased personal income in the state and stimulated the demand for health care, education, and other consumer products. Thus, although the mining sector accounted for only 4 percent of the Texas employment in 1984, a study by Hill (1986), using an input-output model of the Texas economy, has estimated that 45 percent of all the new jobs created in Texas between 1972 and 1982 were the result, direct and indirect, of oil and gas exploration and development.

The sharp rise in oil prices carried with it the seeds of the subsequent oil price correction. After 1982 Texas crude oil prices weakened, falling to $15 per barrel by 1986. This pushed the Texas economy into a deep contraction; gross state product declined by more than 10 percent between the end of 1984 and the start of 1988.

New England. Table 4 shows the DRI/McGraw-Hill August 1990 forecast of the unemployment rate and the change in disposable income for New England and the United States between 1990 and 1992. According to the forecast, the downturn in New England will continue through 1990 and 1991, but the economy is expected to improve somewhat in 1992. All the while, the unemployment rate is expected to remain within 0.6 percentage points of the national average. By contrast, in 1986 and 1987, Texas unemployment rates were 9.3 and 8.3 percent, respectively, compared to rates of 6.9 and 6.7 percent for the nation as a whole.

For the New England economy to become as weak as Texas, its unemployment rate would have to rise 2 to 3 percentage points above the national average. There are two reasons why this is unlikely to happen. One is regional diversification, which may be measured by comparing the composition of employment in the region with that of the entire nation. Using this approach, Rosengren (1990) has concluded that the New England economy is diversified sufficiently to benefit from favorable national develop-

Table 4
Forecast of General Economic Indicators for New England and the United States

	1990	1991	1992
Unemployment Rate (%)			
New England	5.6	6.6	6.3
United States	5.5	6.0	5.9
Real Disposable Personal Income per Capita (% change from previous year)			
New England	+.9	−1.0	+.7
United States	+.5	−.4	+.9

Source: DRI/McGraw-Hill Forecast, August 1990.

ments. The second reason why the New England economy has been expected to experience a relatively limited contraction is the support that a growing national economy would provide to the New England manufacturing sector. Specifically, the DRI forecast is based on the assumption of 0.5 percent real GNP growth in the second half of 1990. Should a nationwide recession occur, however, New England may experience an even more significant contraction.

New England is, indeed, more dependent on high technology and defense than other parts of the country. Henderson (1990) estimates that the share of economic activity in New England that can be attributed to purchases by the Department of Defense is 6.2 percent. This figure includes both the direct and indirect purchases generated by defense, but does not include spillovers into the finance and service industries. For the New England states, this share ranges from a high of 7.5 percent for Connecticut to a low of 4.2 percent for Vermont. The comparable figure for the United States is 4.7 percent. Computer-related employment in New England accounts for 5 percent of the total employment in the region. All high-technology employment in New England, including computers, electronics and biotechnology, amounts to 10 percent of the total employment in the region. The comparative figures for the United States are 3 percent and 6 percent, respectively.

New England is more dependent on high technology and defense than other parts of the country.

On the other hand, firms in the high technology and defense industries employ a highly skilled labor force and provide a diverse industrial base that can adapt its skills, capital, and technology to other products. This gives New England the advantage of having a clustering of diverse but related skills and technologies. The research conducted at New England's major universities creates new technologies, attracts new ventures, and supplies highly skilled labor. The region's medical institutions also play a role as catalysts of new business opportunities, especially in the medical instruments industry. Computer and software industries are another important ingredient in the development of medically related firms. Thus the climate in New England is highly conducive to innovation, adaptation, and development.

In summary, New England has not yet experienced anything comparable to the Texas oil boom and bust. Accordingly, the current outlook for the region is that employment will decline, but that the unemployment rate will not exceed the national average by 2 or 3 percentage points, the way it did in Texas. As with all projections, this outlook will change with shifting developments in the national economy, particularly as the Gulf crisis brings added uncertainty.

The Real-Estate Cycle

The problems in the energy sector in Texas were compounded by speculation in real estate and overbuilding. For a time, the expansion of the finance, insurance, and real-estate industry and the sustained high levels of construction activity helped to offset the job losses in mining and manufacturing. In view of the subsequent collapse of the construction industry and the failure of many financial institutions in Texas, however, such rapid growth in the real-estate and financial sector can be seen as a sign of trouble. In the end, such speculation worsened the contraction caused by falling oil prices.

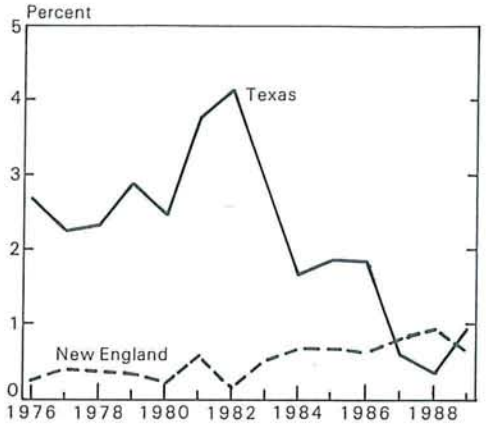
Growth in the finance, insurance, and real-estate industry in New England has been almost as rapid as that in Texas. Construction employment rose more rapidly in New England in the mid-1980s than it did in Texas. Of course, construction had been growing in Texas for a much longer period, and the construction industry was somewhat larger in Texas than in New England.

New England did not experience the significant overbuilding that contributed to the collapse of the real-estate market and the troubles of financial institutions in Texas. Office vacancy rates were low in New England before the downturn. The metropolitan office vacancy rate in Boston was 13.7 percent and that in Hartford was 17.3 percent in December of 1988. Nationwide, the metropolitan vacancy rate was 19.7 percent. In contrast, vacancy rates in Texas were high even before the slump; in 1984 the rate in Dallas was 20.7 percent, in Houston 27.6 percent, and in San Antonio 22.9 percent. The national average was 17.2 percent in that year. Current office vacancy rates are higher for nearly all these areas, but Boston remains well below the national average.

Vacancy rates alone do not measure the degree of construction "overhang," the amount of extra

Chart 4

Percentage Change in Population, New England vs. Texas



Source: U.S. Bureau of the Census, *Statistical Abstract of the United States*, 1982-3, 1988.

space that must be absorbed before excess supply ceases to be a problem. Both residential and office space is built in anticipation of future demand, which is sustained by continued low unemployment and the influx of new workers.

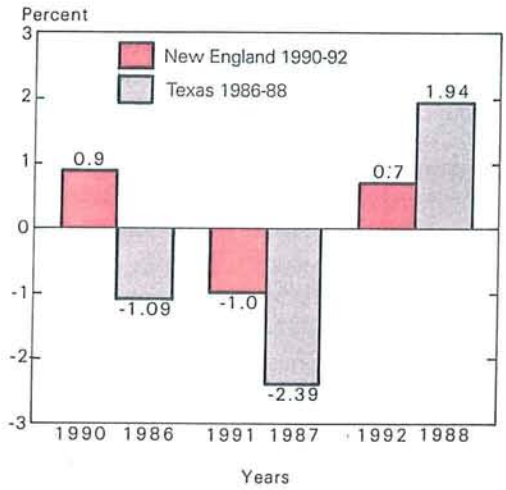
In Texas, the pattern of population growth tended to exacerbate construction "overhang," which is not the case in New England. The Texas population grew very fast before the downturn, as a result of immigration into the state. This fed the boom in construction of both office space and residential housing. When oil prices collapsed, the increase in unemployment halted immigration. Consequently, population grew more slowly than had been anticipated, with two results: first, a further slowdown of the service economy, thus reducing demand for office space; second, a reduced rate of household formation, thus eliminating the need for much new residential housing. As a result, construction overhang was exacerbated in Texas, further depressing real-estate prices.

In New England, on the other hand, expectations of future demand for office space and residential housing were not as great as in Texas, partly because most projections did not rely on immigration. Chart 4 compares annual percentage changes in population in New England and Texas between 1976 and 1989. As the chart shows, population growth in New England has been slow and without abrupt changes. As a result, the expected demand for both commercial and residential housing has been relatively stable. Thus New England does not suffer from construction overhang to the same extent as Texas.

The prospects of the residential housing market in New England depend to a large extent on the length and severity of the regional recession and the speed of recovery. Real per capita disposable income is a particularly relevant measure, since it mainly determines the prices people can afford to pay for housing. Table 4 shows a forecast of the change in this measure. After declining in 1991, real per capita disposable income in New England is expected to increase, beginning in 1992, at a rate close to the national average. This is significant in view of the Texas experience, where banks had losses even on residential mortgages, although their main problems were in commercial development. Chart 5 shows that Texas experienced sizable declines in real per capita disposable income in 1986 and 1987, contrasted with the much smaller decline in New England income

Chart 5

Rate of Change in Real Disposable Personal Income, Per Capita: New England Forecast 1990-92 vs. Texas 1986-88



Source: DRI/McGraw Hill August 1990 Forecast and U.S. Bureau of the Census.

expected in 1991. These numbers do not foretell a serious decline in the value of residential housing. This is good news for the residential real-estate lending prospects of New England banks, whose assets are more concentrated in this category than in commercial and construction real-estate loans.

III. Conclusion

The current New England banking situation and the Texas experience have a number of similarities. In both cases, a boom in the real-estate sector was followed by a sharp contraction caused by weakness in the leading sectors of the economy. In both cases, banks had greatly expanded their real-estate lending and the declining real-estate prices produced substantial loan losses. The similarities, however, do not imply that New England will go on to repeat the Texas experience. The expansion in Texas was driven primarily by increases in the price of oil. In contrast to Texas, employment in New England does not depend to a large extent on the price of a single commodity.

Rather than comparing New England to Texas, it may be more useful to regard the New England experience as an omen of financial stresses that could appear nationwide as the current economic expansion matures. Banks throughout the country have been increasing their commercial real-estate lending relative to their total assets. Should high office vacancy rates and low absorption rates depress real-estate markets, rising loan losses may depress the profits and equity capital of banks throughout the nation.

[1] Real-estate loans are defined as all loans secured by real estate and they include business and personal loans that are so secured.

[2] For example, see Fox Butterfield, "In New England, Worst Recession in U.S. Takes Hold," *The New York Times*, July 23, 1990; and Franklin Tucker, "Massachusetts May Not Be Texas, But It's Doing An Awfully Good Imitation," *Banker and Tradesman*, April 4, 1990.

Appendix

The following equation was estimated for all insured commercial banks in New England.

$$\frac{\text{NON PERF RE LOANS}}{\text{RE LOANS}} = b_1 \frac{\text{CONSTR}}{\text{RE LOANS}} + b_2 \frac{\text{COMM}}{\text{RE LOANS}} + b_3 \frac{\text{RESID}}{\text{RE LOANS}} + b_4 \frac{\text{FARM}}{\text{RE LOANS}}$$

where:

NON PERF RE LOANS — nonperforming real-estate loans
RE LOANS — real-estate loans
CONSTR — construction and land development loans
COMM — commercial loans
RESID — residential loans
FARM — farm loans

The share of nonperforming farm loans was estimated but not reported in the body of this article, because the share of farm loans is less than 0.5 percent of real-estate loans of New England banks. The equation was estimated separately for 1988, 1989, and the first quarter of 1990.

	Nonperforming Real-Estate Loans		
Independent Variables	1988	1989	Q1:1990
Construction	.044	.102	.188
	(6.68)	(7.59)	(8.15)
Commercial	.011	.021	.033
	(2.86)	(2.99)	(3.25)
Residential	.0007	.006	.004
	(.33)	(1.47)	(.75)
Farm	.085	.043	−.011
	(2.00)	(.47)	(−.09)
Mean of dependent variable	.01	.02	.03
Standard Error of Regression	.02	.03	.04
R-squared	.13	.14	.21

Note: t-statistics in parentheses.

References

First Boston Corporation. 1990. "Real Estate and Banks. The Impact of Nonperforming Loans." *Bank Market Service*, vol. 16, no. 19, March 30.

Henderson, Yolanda K. 1990. "Defense Cutbacks and the New England Economy." *New England Economic Review*, July/August, pp. 3–24.

Hill, John K. 1986. "Energy's Contribution to the Growth of Employment in Texas, 1972–1982." *Economic Review*, Federal Reserve Bank of Dallas, May, pp. 11–18.

Rosengren, Eric S. Forthcoming 1990. "How Diversified Is New England?" *New England Economic Review*, November/December.

The Need to Protect Depositors of Large Banks, and the Implications for Bank Powers and Ownership

Richard E. Randall

Vice President, Research Department, Federal Reserve Bank of Boston. The views expressed do not necessarily represent those of the Federal Reserve Bank of Boston or of the Federal Reserve System. *An earlier version of this article was presented at the Western Economic Association International annual conference on July 2, 1990. The author is grateful to Richard Kopcke, Stephen McCarthy and Eric Rosengren for helpful comments.*

The issues that will be addressed in this article on the U.S. banking system are highly controversial and very much the subject of intense debate at this time. The positions presented here are solely those of the author, drawn from thirty-five years of experience in bank supervision and discount window administration.

Three fundamental issues should be carefully considered before making any decisions on altering the federal safety net or the structure of the U.S. banking system. The first is whether or not bank depositors and other creditors can exercise timely and meaningful restraint on excessive risk-taking by bank managements. Most of the proposals for "reform" of the deposit insurance system rely on the premise that market forces can selectively alter bank lending practices so as to avoid major or widespread bank failures. The counterargument is that the market recognizes serious credit problems only after severe or even fatal damage has been done, and belated market reaction often increases the exposure of the deposit insurance fund.

The second issue is whether the government should handle the orderly resolution of large bank failures in such a way that uninsured depositors and other bank creditors are protected, popularly if inaccurately known as the "too big to fail" policy. Here the argument hinges on the significance of systemic effects where, in the absence of an expectation of such protection, a perceived problem in one large bank would trigger a deposit flight, not only from that bank, but from other banks believed to have serious weaknesses. At issue is the ability of depositors, other creditors and supervisory authorities to assess the risks in the various banks on a timely basis. Another question concerns the special role of banking in the economy, and the critical elements of banking that give rise to the government's interest in controlling the resolution of banking problems. Still another aspect is the possibility of nonbanking firms or institutions requiring similar federal intervention under certain circumstances.

The third fundamental issue is the degree to which banking should continue to be insulated from other financial and nonfinancial activities. The debate here should center on the implications for the federal safety net of a broadening of bank activities or of the movement of nonbanks into banking functions. One question to be resolved is the reliance to be placed on artificial barriers, known as firewalls, in protecting the bank safety net from nonbanking risks. Also to be considered is whether U.S. banks can be sufficiently profitable to survive in world competition in the absence of a broadening of powers.

These three issues are so interconnected that a discussion of any one would be incomplete without the other two. For example, an argument in favor of market discipline through reduced deposit insurance coverage must presuppose that uninsured depositors of large banks need no longer be protected. Otherwise, lower deposit insurance coverage would serve only to drive small and mid-size banks out of business in favor of large banks. Conversely, an argument that the effects of systemic runs on large banks believed to be in hazardous condition will not be destabilizing rests in part on an assumption that the market can distinguish banks with dangerous risk concentrations from those with only superficial or transitory weaknesses. Decisions on broader banking powers have implications for the risks assumed by the safety net. More importantly, any decision to allow nonbanks into essential banking activities may broaden the range of firms that must be included under the federal safety net. Clearly, questions relative to the need for and nature of the safety net, and the dependability of firewalls, must be resolved before undertaking a major restructuring of the financial services sector.

A number of additional issues of importance bear on any restructuring of the banking system and bank safety net. Among these are: (1) the proper interpretation of the recent disastrous experience of the thrift industry and its supervisory and deposit insurance backstops; (2) the question of what to do about the use of brokered funds; (3) the value of proposals for mark-to-market accounting, risk-based deposit insurance premiums, private deposit insurance, early resolution of failing banks, narrow banks, required issuance of subordinated debt, much higher bank capital requirements, and the like; (4) the value of firewalls versus the importance of synergies; (5) the bank holding company as a source of strength to its banks; and (6) the potential for reducing the banking system's propensity for major loss concentrations through more forward-looking supervision. In the author's view, all of these issues can be considered as subsets of one or more of the three fundamental issues identified above.

This article will focus primarily on the second of the three fundamental issues identified, the too big to fail policy, and its implications for the third fundamental issue, banking structure. It will be necessary, however, to briefly explain at the outset the position the author has developed in previous works on the first issue, the ability of the market to recognize credit problems on a timely basis.

I. Timely Recognition of Bank Credit Problems

While several types of risk can lead to the failure of banks, the overwhelming majority of commercial bank failures can be attributed to credit risks in the loan portfolio. Indeed, with respect to large commercial banks it is hard to think of cases of failures or near failures primarily attributable to other causes. Centran (1985) and First Pennsylvania (1980) are the only recent examples that come readily to mind. Thus, the focus here will be on recognition of credit problems in bank loan portfolios.

The largest losses that the Federal Deposit Insurance Corporation fund has had to absorb to date have come from relatively large banks, say those ranking in the top one hundred banking institutions by size.

The overwhelming majority of commercial bank failures can be attributed to credit risks in the loan portfolio.

The level of concentration of total banking assets in large institutions as a group, and the propensity for large banks to develop similar risk concentrations (as recently illustrated by loans to less developed countries, energy and shipping loans, highly leveraged transactions and real estate development lending), suggest that the relevant focus of our analysis should be on large banking organizations.

Banks exist as financial intermediaries largely

because a broad class of borrowers cannot readily obtain credit directly in the market as a result of informational and collateral control factors. Banks must follow business borrowers closely in terms of changing loan balances, frequent credit analysis, customer contact and collateral handling—functions that the market cannot perform easily, except for the largest and strongest concerns. Bank examiners periodically review bank loan portfolios and also ensure that large banking organizations have internal loan review functions. Yet there has been ample evidence recently, particularly in the Southwest and New England, that even specialists with full access to loan records sometimes have difficulty in evaluating the potential losses in a bank's loan portfolio before a high degree of vulnerability has developed.

Market analysts, whether they represent bank stock investors or creditors, have relatively little to go on in forming a judgment on the potential for major losses in a bank's loan portfolio. They can follow nonperforming loans, provisions for bad debts and charge-offs, but these indicators lag the actual risk-taking by months or even years and are trailing indicators of the credit problems that eventually emerge. Furthermore, a sudden deterioration in such indicators has little predictive value since it is seldom clear whether it is the result of a housecleaning, or the tip of an iceberg.

It has been the author's experience that the best evidence of a potential credit problem is a rapid growth in a particular loan category with high inherent risk characteristics. An example would be a rapid growth in construction lending, resulting in a high concentration. If the concentration can be further narrowed, say to condominiums in a particular geographic area that appears to have the potential to be overbuilt, the negative implications become even stronger.

While greater emphasis on this type of analysis should help in timely evaluation of risk, standardized data pertinent to concentrations are limited. It is usually only in the later stages of risk-taking that the sophisticated market can clearly distinguish irresponsible overconcentrations from reasonable specialization. The typical depositor, and even the large depositor with analytical resources, has little potential for making timely judgments on bank risk-taking in loan portfolios.

To demonstrate that, in fact, the various market forces have not been able to identify serious credit problems on a timely basis, the author in an earlier study examined each of the forty large bank holding companies that developed a problem, as defined in the study, in the 1980s through mid-1987.[1] The study related the timing of stock price movements, stock analysts' warnings, and bond rating changes to the period when most of the damage was being built in by lending practices and rapid loan growth. Typically, the potential for serious credit problems developed over a period of three to five years before the actual problems became externally obvious through alarming increases in nonperforming loans or provisions for loan losses.

The evidence of the study clearly shows that market forces were not able to identify the emerging credit problems until substantial, sometimes fatal, damage had been done. The market never reflected the problem before the approximate time of the first clear external signal from the bank itself, through an announcement of high nonperforming assets or loan loss provisions. The study also showed that analysts usually were unable to determine whether such an announcement represented the full recognition of an isolated problem, or the first revolution of a death spiral.

The best evidence of a potential credit problem is a rapid growth in a particular loan category with high inherent risk characteristics.

The issue here is not whether the market can see emerging problems in time to permit uninsured depositors to flee before the bank is closed, or to ensure that supervisors do not leave failing banks in operation after they have been fatally damaged. The point of market discipline is to prevent, or at least mitigate, serious problems, and to accomplish this it must be applied at the stage where unacceptable risks are being taken, so that management will be persuaded to diversify away from those risks before banking or economic factors can seriously damage the bank. Belated imposition of market pressure only complicates the efforts of the supervisory authorities to resolve failing banks so as to avoid systemic effects on other banks.

The author has, in the previous work cited, presented evidence that the market is incapable of

fulfilling this role, and the recent experiences with various large bank holding companies in the Northeast strongly buttress the earlier findings. To the author's knowledge, no findings to the contrary have been published.

The argument is sometimes made that the ability of depositors to identify unsafe banks has not been tested because our current system has made it unnecessary for them to do so. It is contended that in a world without deposit insurance, or other elements of the federal safety net, depositors would have the incentive to determine bank risk either themselves, or by hiring professional services.

Once banks are recognized as problems, it would seem to be the supervisor's role to control further risk-taking through close interaction and frequent on-site review.

This argument presents problems. First, the issue is less a question of motivation than of useful information. As already noted, the best information on actual risk-taking is loan growth data by broad categories. While this information might stimulate analysts to question management about lending terms and to think about the bank's area of expansion relative to pertinent economic trends, it is hard to imagine analysts recommending deposit flight from non-problem banks based on such external risk evaluations.

On the other hand, it is just such evaluations by supervisory surveillance units that should prompt recommendations for on-site examinations to ascertain if risk-taking is reasonable. External analysis of risk-taking is also important to stockholders, who are immediately damaged when a supposedly strong bank announces a jump in its loan loss provision or level of nonperforming assets.

Once banks are recognized as problems, it would seem to be the supervisor's role to control further risk-taking through close interaction and frequent on-site review. Nonetheless, much of the argument for the need to control "moral hazard" through greater market discipline seems to be in terms of banks that are publicly recognized to have problems. It is particularly difficult for even the most experienced external analyst to judge the potential for further material losses in a bank's loan portfolio after one or more announcements concerning loan problems have been made. After original judgments on loan concentrations and economic environment have been made, the acknowledgment of the problem by the bank seldom leaves the analyst with much basis for evaluating the extent of the damage. It often comes down to the credibility that management has with the analyst, and the banking history of the 1980s includes numerous instances where analysts' confidence in management's veracity was misplaced. In several cases management itself was blind to the real extent of the damage.

The conclusion that the actions of depositors and other creditors cannot prevent major credit problems in banks, but can only complicate the orderly resolution of failing banks, has several important implications. It follows that the various proposals to increase depositor/creditor-imposed market discipline will not reduce the potential drains on the deposit insurance fund, and some may have the opposite effect. A recognition that the loss potential in a bank's loan portfolio is not so measurable as to be readily expressed in an accounting sense should limit interest in marking such assets to market. Similarly, it should be recognized that no preset formula for risk-based premiums can capture risk-taking with sufficient discernment that a pricing scheme can be relied on to deter excesses. While the author supports risk-based premiums as a logical step, expectations of its impact should be realistic.

II. Protection of All Depositors in Large Banks

The phrase "too big to fail" is a useful shorthand for an informal policy that has been in effect for some time in this country, and in one form or another in all industrial nations. The phrase itself is inaccurate, so it is useful to explore just what the policy is and why it exists.

Too Big to Fail Concept

In the United States, large banks whose capital has been depleted are not prevented from failing. Some form of legal reorganization takes place so that

the owners of the banks lose their entire investment before the Federal Deposit Insurance Corporation and the uninsured creditors of the banks suffer a loss. If the deposit insurance system works as intended, losses to the FDIC are not losses to the federal government or the taxpayers, but are met by the insured banks as their assessments replenish the insurance fund. The only way that the federal government (the taxpayers) would be required to absorb losses would be if the losses were so great, and the banking system so damaged, that the industry as a whole could not absorb these losses.[2]

The rhetoric on the issue often refers to the taxpayers "bailing out" the banks that get themselves in trouble. In fact, the policy under discussion is not designed to prevent nonviable banks from failing, and certainly not to protect either the stockholders or managements of failing banks. Taxpayer funds would only be called upon in the event of a "meltdown" of the entire commercial banking system, not to resolve individual banks.

The concept of too big to fail effectively extends insurance coverage to all depositors and creditors of large banks (or bank subsidiaries of large bank holding companies) as part of a supervisory reorganization. This is done to avoid runs on these banks that could lead to similar runs on other large banks perceived by the public to be in questionable condition. Such a policy is necessary because of a combination of several factors:

1. Uninsured bank depositors are not willing to risk losses, even relatively small ones, and in the absence of implied protection by the insurance fund, will withdraw funds as fast as possible from banks in questionable condition.
2. Bank creditors are unable to distinguish failing banks from damaged but viable banks, and will tend to shift funds to banks believed to be in safer condition on the receipt of bad news.
3. Supervisors, despite their superior access to information, will often require time to evaluate the viability of a large bank once a run begins.
4. Even after the authorities determine that a large banking organization is failing, it requires time, at least several weeks if not months, to arrange for its disposition. This is aside from any legal impediments to a determination of insolvency.

Thus, in the absence of implied protection, large-scale deposit runs would be possible in any large bank whenever the market was surprised by bad news. In a deteriorating environment with several large banks reporting increasing credit problems, market concerns for possible failures would probably be heightened, lowering depositors' tolerance for bank weakness. For banks that ultimately do fail, the elapsed period between a major deposit run and eventual resolution by the authorities will probably be several months. This would be true even if there are no legal constraints on the disposition of nonviable banks before they are insolvent in terms of book net worth. These banks, and other large banks that are troubled but are capable of surviving, will have to be supported through the discount window, or some other mechanism such as FDIC assistance, for a prolonged period, and potentially in very large amounts.

Myth or Necessity?

Some economists and representatives of large banks have argued that the need for protecting creditors of large banks is a myth.[3] They contend that very large failing banks could be disposed of, with depositors and creditors absorbing losses, without significant repercussions for the banking system or the economy. They note that, while deposit runs may occur in particular large banks, large-scale conversion of deposit balances to currency or to foreign currency deposits is unlikely, leaving monetary aggregates and the level of economic activity essentially unchanged. They argue that any systemic bank runs could be handled by the Federal Reserve as the lender of last resort.

In the absence of implied protection, large-scale deposit runs would be possible in any large bank whenever the market was surprised by bad news.

If one imagines the isolated closure and liquidation of a single very large bank with well-known problems at a time when the domestic and international banking systems are in unquestioned good health, it could be argued that the disruption that would result, while considerable, would be transitory with limited systemic effects and no prolonged neg-

ative impact from a macroeconomic point of view. (A counterargument would be that even if such action causes no immediate difficulties, it could be taken as precedent and result in instability at a later time when several large banks were in trouble.) This scenario of an isolated failing large bank cannot be assumed, however, and is not the model to use in considering this major policy issue, since the most serious banking troubles are apt to stem from economic events that affect a number of large banks at about the same time. Recent examples include the Southwestern energy recession, various regional real estate gluts, the economic failures of heavily indebted developing countries, and bank exposure from loans to highly leveraged borrowers.

A more appropriate scenario to consider would involve several large banks in danger of failing at about the same time, including some money center banks and perhaps a few major foreign banks. Problems would likely stem from the impact of some economic event on several banks, and banks could be adversely affected by more than one economic event because of a coincidence in timing. There would likely be a high degree of public uncertainty as to the depth of the underlying economic problems and the timing of recoveries. Most uninsured depositors and other bank creditors would be concerned about the possible failure of particular banks, and would be prone to hasty reaction to rumors and misinterpretation of information. Adverse developments in one bank could cause instability in other banks perceived to have similar problems.

The most serious banking troubles are apt to stem from economic events that affect a large number of banks at about the same time.

Supervisors would face similar uncertainties, even though they had much more information on the weaknesses of specific banks. While the depositor need only decide that the situation warrants pulling funds from one bank and putting them into another, the supervisor must determine if a particular bank is likely to fail, quantify the degree of any potential insolvency, and devise and execute a strategy for resolving the institution. A careful evaluation of the credit exposure of a troubled major bank involves a significant portion of the available examiner resources, and evaluations must be updated frequently as conditions change. When a number of large banks are in trouble at once, the supervisors will not necessarily be in a position to know the viability of a particular major bank when a deposit run develops. In a chaotic situation where depositors are rapidly shifting deposits from bank to bank, and creditors of banking concerns are refusing to roll over notes, the authorities must decide whether to seize particular institutions or support them, in some cases without a current evaluation.

The consequences of seizing an institution that is damaged, but still viable, are fairly serious, so the temptation will be to support banks in questionable condition until a reassessment can be made. Such support may involve heavy discount window lending on increasingly uncertain collateral. This problem should be mitigated, but will not be eliminated, by prompt resolution techniques.

In the payments area, sudden runs on a number of major institutions could place great pressure on banks and the Federal Reserve System to limit daylight and overnight exposure to other banks and customers. It is not hard to visualize scenarios in which the payments system would cease to function efficiently for an extended period while multiple runs on large banks continued. This could produce a snowballing of defaults and delinquencies, and lead to failures of weak firms and disruption of business generally. The effect could be to depress economic activity for a number of months.

Numerous borrowers would abruptly be forced to try to find other lenders as their usual banks experienced major deposit runs and were forced to suspend lending activities. Defaults could occur on bank and bank holding company debt as well as that of other firms, leading to a flight to quality and likely disruption in various markets. Some funds could flow to foreign banks in search of safety, disrupting normal intermediation patterns even where the funds continued to be denominated in dollars.

The contagion of uncertainty could cause runs on any major foreign banks that were believed to be in difficulty, further adding to the general confusion. Bank supervisory, deposit insurance, and discount window personnel could become overwhelmed by the combination of failures of nonviable banks and liquidity crises in viable banks. This could result in

delays and misjudgments that increased the costs to the insurance fund, the banking industry and the public, and prolonged the period of disruption.

It is probably true that, even in a chaotic situation such as that described above, the total volume of deposits of the banking system would not be substantially reduced by direct conversion to currency or foreign-denominated balances. The amount of funds available for loans, however, could be substantially reduced. As deposits run from weak banks to stronger banks, the banks receiving the sudden influx of deposits cannot be expected to increase loans quickly, taking on customers squeezed out of other banks. Much of the influx would be considered temporary funds and invested accordingly. Capital adequacy considerations and the time necessary for information gathering, credit analysis, and loan approval would also limit the ability of healthy banks to absorb the lending activity of the weak and failing banks. Thus, a period of significantly reduced bank lending would result, with negative implications for the level of economic activity.

The supervisory authorities have good reason not to commit to payouts of only insured depositors of major banks. The potential for unleashing forces beyond their control is simply too great.

The banking system is central to the payments mechanism and the provision of short-term credit, and also affects the financial markets and the transmission of Federal Reserve open market operations.[4] The discussion above suggests that the level of disruption to the banking system and bank customers and creditors that could result from a crisis of confidence in the major banks could significantly depress the level of economic activity. It could also increase the losses to be absorbed by the banks, increasing the risk that the banking system itself could be overwhelmed and unable to support the deposit insurance fund.

Thus, the supervisory authorities have good reason not to commit to payouts of only insured depositors of major banks. The potential for unleashing forces beyond the control of the supervisors is simply too great. The authorities do not necessarily have to make an explicit commitment to safeguard the uninsured depositors and other creditors of large banks, although 100 percent insurance is a viable alternative. The current level of uncertainty about implied support seems to have produced no great problems. It should be recognized, however, that even if deposit insurance were to be eliminated entirely, the government would probably still decide that it could not allow the largest banks to be closed and liquidated. Support for creditors of large banks is necessary because of the special role of banks in the provision of short-term credit, in the payments system, and in the transmission of liquidity through the economy.

Co-Insurance

Various proponents are recommending co-insurance, under which the bulk of deposits would be protected by the insurance fund, but all deposit balances over a specified amount would share in any loss.[5] For example, all deposit balances over $100,000 might be given a 10 percent haircut, subject to further disbursements later if losses (in excess of those absorbed by capital) prove less than this. If the haircut proves to be insufficient to cover the net loss, the balance would be absorbed by the insurance fund.

The major problem with such a proposal is that most large depositors are likely to run from a bank about as fast when their exposure is 10 percent as when it is 100 percent. Thus the scenario for instability envisioned above is just as much a threat with co-insurance as without it. While one cannot be certain that this is what will happen, corporate and institutional treasurers have a reputation for being risk-averse in their cash operations and have little motive to gamble on a distressed depository. Since the proposal involves an explicit assurance that the authorities will resolve large banks in such a way as to impose the fractional loss on depositors, the potential would be created for extreme sensitivity to adverse developments. It would seem to be incumbent on the proponents of co-insurance to demonstrate that their proposal will not leave the banking system subject to dangerous deposit volatility. This is quite aside from evidence that the market discipline aspects of the proposal are unlikely to come in time to influence risk-taking, as discussed in the first part of this paper.

Policy Ambiguity versus Full Insurance

The practice in the United States, and apparently in the other industrial countries, is to be ambiguous about special protections given to creditors of large banks. This country has no written policy, no clear demarcation of how large a bank must be to qualify for protection, and no certainty that the authorities will not some day experiment with a payout of only insured deposits in a relatively large bank. Despite these deliberate ambiguities, deposit instability has not been excessive to date as the market has assumed that the too big to fail policy is operative. While the United States has not experienced excessive deposit volatility when large banks have announced problems, it might take only one misguided attempt to make banks more vulnerable to market pressure to create a much less stable environment.

An alternative approach to the current ambiguity would be to fully insure all bank creditors.[6] Since the creditors in large banks appear to be de facto insured now, and payouts of only insured depositors in small banks are rare, 100 percent insurance would not materially increase the real burden on the insurance fund. Furthermore, an argument can be made that it would have no significant impact on so-called moral hazard, as long as supervisors control growth and asset quality in known problem banks. Since uninsured depositors in large banks (1) are not able to exercise timely market discipline and (2) currently tend to behave as though they are protected, at least until the risk of failure becomes quite visible, little constructive depositor pressure would be lost with 100 percent insurance.

In the author's view, there is no good reason not to go to 100 percent insurance, and doing so would eliminate unnecessary ambiguity and strengthen the competitive position of small and mid-size banks. On the other hand, there is probably no compelling reason to push for such a change at this time and the current wisdom probably would have to be altered significantly to gain wide support for this concept.

Protection of Nonbank Entities

If the government must concern itself with the orderly disposition of banks, must other entities be given comparable treatment? Other industries have no industry-supported insurance funds comparable to the bank insurance fund to protect their creditors, but this is not the key factor.[7] Bank creditors would have to be protected in order to prevent systemic runs, even if deposit insurance did not exist. Are there other industries where systemic risk factors create the potential for broad-based and prolonged disruption of the real economy, financial markets, or the medium of exchange? Should a firm be considered for government intervention simply because it is a very large employer, or because many individuals would lose insurance coverage or invested funds?

These questions have no easy answers and it appears that this is the appropriate area for policy ambiguity. In order to minimize potential exposure to the taxpayer, the safety net should be limited to banks. On the other hand, authorities should be alert to any situations that have the potential for raising questions about appropriate federal intervention.

For example, it might be unwise to allow the privatization of the federal secondary market-makers in mortgages, the Federal National Mortgage Association and the Federal Home Loan Mortgage Corporation, if a likelihood existed that they would have to be recapitalized someday by the taxpayers. Conceivably this could become necessary in order to prevent a nationwide crisis, perhaps in the form of a shutdown in the origination of new mortgage-backed securities and widespread defaults on existing securities.[8] Another example might be a concentration in a particularly vulnerable asset category that could threaten reserves of several of the nation's largest life insurance companies, for instance, junk bonds or particular categories of commercial real estate loans.

Every effort should be made to minimize any exposure of the government, and to remove public expectations of government support, keeping open the mechanisms to act if a situation should ever develop where governmental support was absolutely essential. The implications of this desire to limit the federal safety net to essential functions should be a key factor in considering changes in the interrelationships between banking and other industries.

III. Implications for Limitations on Bank Activities and Ownership

Banking organizations in this country have been limited to certain activities for many years, and the type of entity that can own a bank has been similarly restricted. The recent trend has been to gradually expand the range of permissible activities, but this has generally been done through administrative rulemaking rather than by congressional action.

Pressure is increasing, mostly from the large domestic and foreign banks, for this country to adopt some form of universal banking system. Some would like to see the legal recognition of a financial services industry, with banks, investment bankers, insurance companies, and various other types of financial institutions free to engage in one another's traditional activities and to acquire firms so engaged. A more extreme position espoused by several of the largest banks calls for a complete breakdown of distinctions between banking and commerce by allowing combinations of banks and any other type of business.

Some proponents acknowledge that banking is special and that the federal safety net must be maintained for banks, but would not like to see it extended to other activities in which banks may engage. Some argue that this can be done through administrative firewalls. Others complain that firewalls would nullify most of the advantages of the synergies that make interindustry combinations attractive. They tend to downplay the need for a federal safety net and seek to reduce the role of deposit insurance. They see no need to protect the creditors of large banks.[9]

The federal safety net should be maintained with respect to banking institutions and every effort made to avoid extending it to other types of firms.

Most proponents of these various degrees of deregulation of bank powers and ownership make one or more of the following arguments:

1. Large U.S. banks cannot compete internationally if they remain restricted and will continue to shrink in relative size.
2. Large U.S. banks are relatively unprofitable because of activity restrictions.
3. Ownership restrictions increase the cost of capital for U.S. banks by limiting the range of potential purchasers of bank ownership or reducing the franchise value of a bank charter.
4. The risk of failure in a firm diversified across several industries is substantially reduced, lessening concerns about risk to the federal safety net.

Thus, proponents attempt to make a case that it is imperative and urgent that the United States open up the banking industry or face a continuing decline of the industry and this country's economic well-being.

The thrust of this article is that the federal safety net should be maintained with respect to banking institutions and that every effort should be made to avoid situations that would require extending the safety net to other types of firms.[10] It is beyond the scope of this article to consider the full implications of these conclusions for bank powers and ownership, but three questions will be explored briefly:

1. Can the safety net's exposure be limited within a multi-industry firm through the use of firewalls, and will firewalls or Chinese walls nullify synergies?
2. Is it feasible to develop a compromise banking structure that limits risk exposure to the safety net while permitting banks to take advantage of natural synergies in nonbanking activities?
3. How valid are arguments that banks must diversify to compete?

The Effectiveness of Firewalls and the Restraints of Chinese Walls

Firewalls are administrative segregations of functions designed to prevent losses in one area of a firm from being transmitted to another area. Typical firewall devices include corporate separation of functions, separate work forces, separate funding sources, prohibitions on certain transactions between units, and the like. Firewalls may be distinguished from so-called "Chinese walls" by the purpose they serve. Chinese walls, as the term is typically used, are to prevent improprieties, such as the use of privileged information by another branch of the firm for corporate gain, or violation of a fiduciary responsibility in the process of avoiding a loss in another unit.

From the standpoint of protecting the bank safety net from having to cover losses in nonbanking activities, the fundamental question is whether firewalls can be depended on in extremis. When a multi-industry firm is struggling to survive, the senior management and owners will be under great pressure to avoid failure, and the temptation to shift assets from a bank subsidiary to a troubled nonbank subsidiary may prove irresistible. The risk of subsequent detection and punishment may count for little compared to surviving a crisis and saving the enterprise. Thus we should treat firewalls with considerable skepticism as a basis for allowing activities that

would otherwise be denied.

Furthermore, efforts to protect the safety net with strong firewalls may nullify the value of synergies that made the combination of bank and nonbank functions seem desirable in the first place. For example, large banks want to have the same people who meet with commercial customers to discuss a potential loan be able to arrange a security issue or a combination of the two if that fits a customer's needs. Strict segregation of lending and investment banking personnel for firewall purposes would make that impossible.

A further complication is whether Chinese-wall considerations should dictate such a segregation anyway. It is noted that Switzerland's banks appear to be moving toward the American position of separation of commercial and investment banking largely because of Chinese-wall-type considerations.[11]

Alternative Banking Structures

If one does not accept the reliability of firewalls, but does believe that it is necessary to protect the creditors of large banks, it is difficult to see how banking organizations and other financial services companies can be allowed to combine without broadening the coverage of the federal safety net. Perhaps that would be an acceptable consequence to those who are convinced that it is important to move toward universal banking with respect to financial services. However, the issue of extending the safety net to nonbank financial firms has not been raised by proponents of expanded powers or made part of the debate on financial restructuring.

Some argue that we must choose between complete separation of bank and nonbank financial firms on the one hand, and unrestricted entry on the other, with no middle ground. Since banks have some investment banking powers now, the former position would mean a rollback from current practices. Neither this alternative, nor an expansion of the federal safety net to cover much of the financial services area, seems very attractive.

It might be worthwhile to search for a compromise solution that would avoid the least desirable aspects of the two extreme positions. For instance, if it were possible to devise a structure whereby commercial and investment banking could coexist in an institution without destroying the advantage of synergies, consideration might be given to allowing certain investment banking functions to coexist with commercial banking without firewalls, even though the effect would be to broaden the activities covered by the bank safety net. This might be seen as a reasonable compromise that would not nullify synergies or place unrealistic faith in firewall safeguards, and have the advantage of not forcing banks to retreat from the level of investment banking involvement they have already attained. The different risk characteristics of commercial and investment banking would argue for maintaining separate accounting and applying separate capital adequacy standards, even though personnel separation was not required.

While permitting banks to engage in a financial activity without protective firewalls would broaden the safety net coverage, this might be considered acceptable if the inherent risks were not much greater than various bank lending activities, and if it were felt that these risks could be evaluated and controlled through the supervisory process. However, if nonbank financial firms were free to enter banking, then the safety net might soon be broadened substantially as large nonbank financial firms acquired banks, and it became necessary to protect creditors of the entire entity. Since this would seem to be undesirable, it suggests that consideration might be given to an asymmetrical structure in which banks would be permitted limited involvement in those nonbank financial activities where natural synergies are strong and conflicts and risks containable. However, banks would not be permitted to acquire major nonbank firms in certain industries (such as investment banking). Nonbank financial (and nonfinancial) firms would be excluded from owning banks or having direct access to the payments mechanism, and would continue to be outside the protection of the safety net.

Such a compromise structure would be vulnerable to criticism that it is unfair to some parties, and that it employs arbitrary limits on activities. Generally such solutions are to be avoided. Furthermore, the suggestion is made solely on conceptual grounds without careful consideration of how Chinese wall aspects might nullify synergies, or of the relative risk factors in various financial activities. It also does not attempt to deal with the complex question of how the line should be drawn between permissible and impermissible activities of banking institutions. Nonetheless, some compromise along the lines suggested above may be the only way to permit broader banking activities without greatly increasing the exposure of the taxpayer implied in an expansion of the safety net to other industries.

Evaluation of the Compulsion for Universal Banking

It is beyond the scope of this article to present a detailed analysis of the relative profitability or competitiveness of American banks. However, a few observations can be made. Earnings of large American banks were depressed during much of the 1980s by unusually high loan loss provisions and loan nonperformance. While large foreign banks suffered from some of the same problems, particularly exposure to less developed countries, the American banks suffered heavily from the effects of energy price changes on the economy of the Southwest and a succession of severe regional real estate problems. We should adjust bank earnings for this period of extraordinary credit problems and reexamine competitiveness before concluding that American banks are at a disadvantage due to restrictions on activities.

No evidence has been offered that earnings would have been better had our banks been allowed unrestricted entry into investment banking, insurance or other proscribed financial activities.[12] In fact, investment bankers in this country are currently experiencing a period of severe unprofitability and retrenchment, and concerns have been expressed recently about the credit risk in the loan and investment portfolios of some insurance companies.[13] It is by no means clear that our larger banks would be stronger today had they not been constrained from full participation in these industries.

No evidence has been offered that earnings would have been better had our banks been allowed unrestricted entry into investment banking, insurance or other proscribed financial activities.

The push for universal banking in the United States today is reminiscent of the movement in American industry toward conglomerates in the 1960s and early 1970s. By the mid 1970s the markets had become convinced that managements are not generally successful at managing widespread, multi-industry firms, and divestitures became commonplace.[14] In fact, much of the leveraged buyout lending and junk bond issuance of the 1980s related to continuing divestitures by multi-industry firms. One might well question the economics of universal banking in the United States in view of this history. Indeed, the markets could reflect the negative experience with conglomerates in such a way that the cost of capital would actually be higher for universal banks than banks that adhere to more traditional activities.

IV. Conclusions

This article started with the premise, based on the author's earlier work, that market discipline cannot be effective in deterring excessive credit risks in banks. This is because the market cannot recognize and properly evaluate such risks on a timely basis, and belated market reaction is counterproductive.

The main thrust of this article is that the authorities must continue to handle large bank failures in such a way that they do not trigger a systemic flight to perceived quality. This requires that all depositors of such banks be given at least implicit assurance that their funds will be protected. This protection is provided by an industry-supported insurance mechanism, and the burden would fall to the taxpayer only in the event that the banking industry as a whole was unable to absorb the level of losses generated.

While situations could develop where it would be necessary to use taxpayer funds to absorb losses in nonbanks, efforts should be made to avoid or minimize the potential for such needs. We should be very concerned with the potential for expansion of the safety net to large nonbanks as a consequence of a broadening of bank powers and entry into banking. These safety net concerns cannot be alleviated by administrative firewalls, and attempts to limit safety net exposure by building firewalls are apt to nullify whatever synergies may exist between particular banking and nonbanking activities. Chinese walls to prevent improprieties may also nullify synergies, further casting doubt on the real value of broadened powers.

Various arguments have been presented by proponents of universal banking as to the necessity and urgency for creating a financial services industry or breaking down barriers between banking and commerce. These arguments seem vulnerable to challenge with respect to complaints about the inherent uncompetitiveness of American banks and the presumption that multi-industry firms are more profit-

able and safer than single-industry firms.

In sum, the United States should not draw back from the current implicit backing given creditors of large banks. Bank involvement in investment banking and other financial activities should be limited to areas where synergies need not be nullified by Chinese walls and where risks are acceptable without imposing prudential firewalls. We should not hesitate to continue to restrict entry into banking by nonbank firms in order to avoid broadening the safety net, even if that means asymmetrical treatment of banks relative to nonbanks.

---

[1] See Randall (1989).

[2] This is what happened in the 1980s with respect to the thrift industry and its insurance fund, and it happened on such a massive scale that the cost to the taxpayer will be very painful. However, the thrift industry, its supervisors, and the administrators of its insurance fund all operated under different rules, in a different industry culture and a different political environment, from the banking industry.

This is not to deny that a major disaster could befall the banking industry someday, but the banking industry is less vulnerable to many of the particular problems that beset the thrift industry, while having its own set of stress points. It will be more productive to focus on areas of real vulnerability in the banking industry and its insurance fund rather than to revisit the series of calamities that essentially wiped out the thrift industry.

In order to minimize the possibility that the deposit insurance fund could be exhausted while banks remain able to replenish it, assessments should be raised promptly to reflect, and even anticipate, any abnormal losses that must be absorbed by the fund.

[3] See in particular American Bankers Association (1990) and Kaufman (1989).

[4] See Corrigan (1987) for a discussion of why banks are special.

[5] See American Bankers Association (1990) and Federal Reserve Bank of Minneapolis (1990).

[6] See Independent Bankers Association of America (1990).

[7] The Pension Benefit Guaranty Corporation and the Securities Investor Protection Corporation have some similarities to deposit insurance but are not considered comparable.

[8] Presumably the current market assumption of government backing stems from both the status of these entities as government-sponsored corporations and their dominant role in the secondary mortgage market.

[9] Representative of the variety of views on this subject are American Bankers Association (1990), Corrigan (1990), Reed (1989), Rideout (1990), *The Economist* (1990), and Weatherstone (1989).

[10] The safety net essentially consists of the following elements: deposit insurance, full protection of depositors in large banks (implicit), discount window access, and Federal Reserve backup of Fedwire settlement.

[11] Kraus (1990).

[12] See Boyd and Graham (1988) and Kwast (1988) for research in this area.

[13] See Garcia (1990) and Power (1990) for recent examples of news stories on the troubles of nonbank financial institutions.

[14] See Clark (1990) for a summary of this phenomenon.

References

American Bankers Association. 1990. *Federal Deposit Insurance: A Program for Reform.* Washington, D.C., March.

Boyd, John H. and Stanley L. Graham. 1988. "The Profitability and Risk Effects of Allowing Bank Holding Companies to Merge with Other Financial Firms." Bank Structure and Competition Conference, Federal Reserve Bank of Chicago, May 11–13.

Burstein, Melvin L. 1990. "Banking Safeguards Should Not Be Dismantled." *American Banker*, June 7, p. 4.

Cabinet Council on Economic Affairs. 1985. *Recommendations for Change in the Federal Deposit Insurance System.* U.S. Treasury, January.

Cates, David C. 1990. "Liquidity Lessons for the '90s." *Bank Management*, April, pp. 20–25.

Clark, Lindley H. Jr. 1990. "De-Diversification: Aid to Productivity." *The Wall Street Journal*, May 11, p. 1.

Corrigan, E. Gerald. 1987. *Financial Market Structure: A Longer View.* Federal Reserve Bank of New York, January.

———. 1990. *Statement (with Appendices).* U.S. Senate, Committee on Banking, Housing and Urban Affairs, May 3.

The Economist. 1990. "Survey: International Banking." April 7, pp. 5–68.

Federal Reserve Bank of Minneapolis. 1990. *Federal Deposit Insurance System Issues.* Comment to the Department of the Treasury, March.

Garcia, Beatrice E. 1990. "Equitable Life Makes Painful Decisions, But Firm Felt It Had Few Other Options." *The Wall Street Journal*, June 1, p. B9C.

Gilbert, R. Alton. 1990. "Market Discipline of Bank Risk: Theory and Evidence." *Review*, Federal Reserve Bank of St. Louis, January/February, pp. 3–18.

Greenspan, Alan. 1990. *Testimony.* U.S. House of Representatives, Financial Institutions Supervision, Regulation and Insurance Subcommittee, May 14.

Independent Bankers Association of America. 1990. *Protecting the Federal Deposit Insurance System.* Submission to the U.S. Treasury Department, February.

Kaufman, George G. and Larry R. Mote. 1988. "Bank Securities Activities: Current Position and Future Prospects." Bank Structure and Competition Conference, Federal Reserve Bank of Chicago, May 11–13.

Kaufman, George G. 1989. *Are Some Banks Too Large to Fail? Myth and Reality.* Working Paper Series, Federal Reserve Bank of Chicago, August.

Keehn, Silas. 1989. *Banking on the Balance; Powers and the Safety Net—A Proposal.* Federal Reserve Bank of Chicago.

Kraus, James R. 1990. "Swiss Separate Lines of Business." *American Banker*, May 11, p. 10.

Kwast, Myron L. 1988. "Bank Securities Powers: Are There Diversification Gains?" Bank Structure and Competition Conference, Federal Reserve Bank of Chicago, May 11–13.

Power, William. 1990. "Will Merrill Lynch Have to Take Another Earnings Hit?" *The Wall Street Journal*, May 29, p. C1.

Randall, Richard E. 1989. "Can the Market Evaluate Asset Quality Exposure in Banks?" *New England Economic Review*, July/August, pp. 3–24.

———. 1990. *Federal Deposit Insurance System Study, Comments to the Department of the Treasury.* Federal Reserve Bank of Boston, February 20.

Reed, John S. 1989. "Moving Financial Services into the 21st Century: Challenges and Prospects." *Issues in Bank Regulation*, Bank Administration Institute, Summer.

Rideout, Thomas P. 1990. "U.S. Banks Need New Rules to Win on Global Ballfield." *American Banker*, March 15, p. 12.

Seidman, L. William. 1988. "Deposit Insurance for the Nineties—Remarks Before the National Press Club." Washington, D.C., Federal Deposit Insurance Corporation, November 30.

Sprague, Irvine. 1989. "Will Top Banks Get a Break on Deposit Insurance?" *American Banker*, November 16, p. 4.

Weatherstone, Dennis. 1989. "Firewalls and the Structure of the Future." Bank Structure and Competition Conference, Federal Reserve Bank of Chicago, May 3–5.

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

Address Correction Requested
Return Postage Guaranteed

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