Federal Reserve Bank of New York

Quarterly Review

Winter 1991 Volume 15 No. 3-4

Table of Contents

Table of Contents

The Decline in U.S. Saving and Its Implications for Economic Growth

by Ethan S. Harris and Charles Steindel

By conventional measures the U.S. saving rate declined dramatically over the last ten years. Household saving averaged just 3.8 percent of GNP in the 1980s, down from a 5.0 percent average over the previous thirty years. Corporate saving has also fallen, and the government has become an increasing net borrower. Overall, the net national saving rate—domestic funds available for new investment—dropped to just 3.0 percent in the 1980s, less than half of its historical average of 7.5 percent.

Despite these dismal statistics, some would argue that the drop in saving rates is not a cause for concern. No apparent disaster has attended the low saving rate; instead, in the 1980s the United States enjoyed the longest peacetime expansion of the postwar period. Spurred by booming stock and real estate markets, the value of wealth rose dramatically during the decade. Broader measures of saving incorporating government capital investment and consumer durables show much more saving than conventional measures. Furthermore, even if saving has fallen, in a market economy the rate of saving is no more than an expression of people's "time preference"—if consumers have chosen to spend more today and to leave less for tomorrow, why should we question their choice?

This article examines the saving data and finds that concerns about the low saving rate are indeed well founded. The first half of the article documents the trends in a variety of measures of saving. We find that any measure of saving that focuses on the actual acquisition of productive assets shows a clear decline in the 1980s. Broader measures of saving do show higher *levels* of saving but also show the same downward trend as the conventional measures. Although capital gains from the stock market caused some wealth-based measures of saving to surge in the 1980s, empirical tests reveal that these measures do not capture the growth of productive capacity: stock price appreciation is a poor substitute for real asset accumulation.

The second half of the article explores the consequences of the saving decline. Low saving has not caused a sudden collapse in the economy, but it has caused a steady erosion of the nation's growth potential and it has been accompanied by a sharp increase in net indebtedness to foreigners. A simulation model of the economy suggests that low saving relative to past trends has already cost the economy about 15 percent of its capital stock, lowering the nation's potential output by 5 percent. This drag on growth comes at an inopportune time. In the next several decades declining growth in the working age population will increasingly constrain economic growth. At the same time, rising environmental costs, increasing payments to foreign owners of U.S. assets, and a growing retirement population will make an increasing claim on output. Continued low saving and investment reduces the nation's ability to respond to this squeeze on living standards.

On the international front, low saving has contributed significantly to the worsening of the nation's trade and investment position. This development in turn has fueled support for restrictions on international trade and investment. It may also have reduced investor confidence and increased vulnerability to shocks from abroad. The U.S. appetite for foreign capital is especially troubling in light of the growing capital needs of

the emerging market-oriented economies of Eastern Europe and the developing world.

It is not too late to undo the damage of the 1980s. As our simulation model shows, a recovery in the net saving rate to its pre-1980s level would gradually rebuild the capital stock and would tend to reverse the deterioration in the external debt position. This saving recovery could be accomplished by balancing the federal budget and raising the private saving rate by about 2 percentage points, or by pushing the government balance into surplus and buying down some of the debt accumulated in the 1980s. Deficit reduction efforts in the last several years, including the recent budget accord, are important steps in this direction, but further action would be needed to complete the process. In the short run, a higher saving rate would mean lower current spending; within a decade, however, consumption would recover to well above its current path. Under reasonable assumptions about people's time preference—how they value consumption today relative to consumption tomorrow—the delayed gratification would be well worth it.

The 1980s decline in U.S. saving

Saving is one of the most important but most widely misunderstood topics in economic analysis. Looking at the subject broadly, we can identify three overlapping concepts of saving. All are useful in certain contexts, but they are not equally useful measures of the long-run health of the economy. The three are 1) saving as the increase in net worth, 2) saving as unspent income, and 3) saving as the supply of capital.

These would be equivalent if all unconsumed income were used to purchase capital and if all assets remained fixed in price. But because these conditions are not satisfied in the real world, it is important to distinguish carefully between the concepts.

The relation between the three concepts

The most comprehensive way to gauge saving is to trace changes in wealth or net worth. Some commentators argue that the rapid increase in wealth in the 1980s reflected an equally rapid increase in the nation's productive potential. To demonstrate why the wealth-based measure of saving may be a deceptive indicator of changes in productive power, we trace the relationship of changes in wealth to saving and to growth in capacity. Wealth is affected not only by saving out of current income—the outright purchase of assets—but also by changes in the value of existing assets. Consider the basic definition of wealth, W, as the product of a real stock of assets, A, and a price per unit of assets, P. (Clearly, wealth is equal to assets less liabilities. For simplicity, we are using only the word "assets.")

$$(1)\quad W = PA.$$

Taking first differences, we can derive an expression for the change in wealth (suppressing subscripts on lagged values of variables for the sake of simplicity):

$$(2)\quad \Delta W = P\Delta A + A\Delta P.$$

The first term of equation 2, $P\Delta A$, is the product of asset prices and real asset accumulation. It is conceptually equal to the conventional definition of saving as unspent income, as exemplified in the U.S. National Income and Product Accounts (NIPA). The second term, $A\Delta P$, is the product of the stock of assets and the change in asset prices, and represents that portion of wealth accumulation due to revaluations, or capital gains.

Wealth accumulation clearly bears some causal relationship to the asset accumulation concept of saving. Not only does asset accumulation increase wealth, but growth in wealth may also affect decisions to acquire new assets. If people feel more wealthy because of capital gains on their assets, they may decide to spend more and put aside less resources for asset accumulation.

Neither wealth accumulation nor asset accumulation necessarily measures the growth of productive capital. The assets viewed as components of wealth by the residents of a nation may include items that are not necessarily part of the productive capital stock, such as government debt.[1] Suppose the asset list consists of two items, productive capital, K, and other assets, O. Further suppose that each asset has a price associated with it, P_K for productive capital and P_O for other assets. We can then rewrite the basic expression defining wealth as

$$(3)\quad W = P_K K + P_O O.$$

Taking first differences, we have

$$(4)\quad \Delta W = P_K\Delta K + P_O\Delta O + K\Delta P_K + O\Delta P_O.$$

The first two terms of this expression, $P_K\Delta K + P_O\Delta O$, sum to total asset accumulation (the conventional definition of saving). Note, however, that assets may be

[1]Considerable controversy exists in the economic literature over the issue of viewing government debt as part of aggregate wealth. Although an individual's holdings of government debt are clearly part of his or her wealth, it has been argued that some individuals in the population take account of the future taxes that will be levied to redeem the debt and feel poorer as a result. See Robert Barro, "Are Government Bonds Net Wealth?" *Journal of Political Economy*, vol. 82 (November-December 1974), pp. 1095-1118. For a nontechnical discussion of Barro's work and the literature it spawned, see "The Public Purse," *Economist*, November 24, 1990, pp. 77-78.

accumulated in both "productive" and "nonproductive" forms. The last two terms represent the capital gains on both types of assets.

Neither increases in wealth nor asset accumulation is a precise measure of productive investment. The growth of the productive capital stock is equal to purchases of new capital, $P_K\Delta K$, plus the portion of the capital gains on existing capital that can be associated with increased productivity. Thus, total wealth accumulation overstates productive investment because it includes all capital gains and purchases of unproductive assets. Asset accumulation alone also includes purchases of unproductive assets but fails to account for any increased productivity of existing assets.

In the actual data, asset accumulation—the conventional definition of saving—and wealth accumulation are fairly readily observed. The supply of productive capital, a primary focus of this article, is more difficult to measure. Any division of assets into productive and nonproductive categories must be somewhat arbitrary. Likewise, a further division of capital gains on productive assets into those reflecting additions to productivity and a residual category will also be arbitrary.

The next portion of the article examines the data. As we shall see, untangling the various measurements of saving reveals a consistent pattern: a downtrend in the supply of capital and in the growth of productive capacity.

Saving as unspent income

The NIPA compiles data on saving defined as unspent income. Income that is not spent is necessarily used to acquire assets or repay debts. Sectoral saving is merely sectoral income less the sum of transfers to other sectors and spending on currently consumed goods and services. Table 1 documents movements in saving, defined this way, over the last generation for the total economy and for the household, corporate, and government sectors. It breaks down the postwar period into four phases: the high-growth 1950s (1953-61), the boom years (1962-73), the productivity slowdown (1974-79), and the most recent period (1980-89, further divided into 1980-84 and 1985-89).

Household saving

Household or personal saving as a share of GNP is shown in the first column in Table 1. Personal saving is usually measured as a percentage of disposable personal income—the commonly reported personal saving rate—but to facilitate comparison with other measures, it is here shown relative to GNP. It is clear that personal saving was unusually low in the expansion of the 1980s.

Personal saving is arbitrarily defined in the NIPA. For example, increases in corporate profits add to personal income and saving only if they are distributed as dividends. But if corporate stock values reflect increases in undistributed profits, household shareholders benefit from retained earnings. Another anomaly arises in the area of employee benefit plans. Employer payments into the reserve funds of private retirement and insurance plans are counted in personal income and saving, while similar payments by government employers are not included. Given these anomalies in the construction of personal saving, it makes sense to focus on broader measures of saving.[2]

Corporate and private saving

Corporate saving consists of corporate profits after payment of taxes and dividends. Column 2 of Table 1 shows that corporate saving, like household saving, was

[2]Personal saving is still a useful indicator for other purposes. While the level of the personal saving rate at any one time may be a poor clue to overall saving (or even to saving by and in behalf of households), changes in the personal saving rate may give some indication as to the underlying strength of consumer demand. In general, although special factors such as government pay raises can distort monthly or quarterly data, declines in the personal saving rate are associated with strong growth in consumer demand, while increases are associated with weakness in consumer demand.

Table 1

U.S. National Saving

(Percent of GNP)

	(1) Household Saving	(2) Corporate Saving	(3) Private Saving	(4) Government Saving	(5) Net National Saving
1953-61	4.6	2.8	7.4	−0.6	6.9
1962-73	5.1	3.3	8.3	−0.4	7.9
1974-79	5.4	2.5	7.9	−1.2	6.8
1980-89	3.8	1.7	5.5	−2.5	3.0
1980-84	4.7	1.6	6.2	−2.5	3.8
1985-89	2.9	1.9	4.7	−2.5	2.2

unusually low in the expansion of the 1980s. In part, the weakness in corporate saving reflects the increasing share of corporate revenues going toward interest expense. However, rising corporate interest payments help increase household income and saving. Nonetheless, the sum of household and corporate saving, net private saving, has also been unusually low in recent years (column 3).

Government saving

By NIPA definition all government outlays are either spent on currently produced goods and services or transferred to other sectors, so government saving is simply tax receipts less spending—that is, the government surplus. As column 4 shows, the consolidated government budget position (federal plus state and local) went from approximate balance or small deficit for the bulk of the postwar period into deep deficit in the early 1980s. Somewhat surprisingly, the late 1980s saw no improvement in government saving: partial success in curbing the federal deficit was offset by a deterioration in the state and local surplus.[3]

Like the definitions of corporate and household saving, the definition of government saving is arbitrary. The U.S. NIPA differs from the national income accounts of some other countries in treating capital spending by the government as a current outlay rather than saving. However, including government capital spending in saving and investment will not change the downtrend: government spending on nonmilitary structures fell to 1¾ percent of GNP in the 1980s from about 2⅓ percent in the 1970s.

The NIPA data on government saving contains other distortions. Like corporate saving, government saving has been held down by increased interest payments to households. In addition, as mentioned above, government contributions to employee benefit plans are not viewed as compensation of government workers. Since both distortions are offset in the household sector, a less deceptive idea of trends in saving can be found by looking at total national saving—the sum of government and private saving (column 5). These figures confirm that national saving has reached exceptionally low levels in recent years.

Saving as the increase in wealth

The Federal Reserve Board compiles detailed sectoral data on wealth accumulation and holdings that can be used to calculate both parts of equation 2—the asset accumulation, or saving, portion and the revaluation portion. However, a number of adjustments are necessary to make these "Flow of Funds" data useful for our purposes.

A preliminary issue is the relationship of these data to the NIPA. Information on sectoral asset accumulation in the Flow of Funds differs from its NIPA counterpart for definitional reasons (some saving flows are allocated to different sectors in the two systems), although in principle national saving is defined identically.[4] There are also statistical differences between the systems. Table 2 uses Flow of Funds data to calculate saving flows as defined in the NIPA. This procedure suppresses the definitional differences between sectoral saving in the two systems and makes it possible to identify the pure statistical differences. A comparison of Tables 1 and 2 shows that the decline in household and private saving in the 1980s is less pronounced in the Flow of Funds data than in the NIPA. With the government sector included, however, the national saving

[3]The annual data from 1986 to 1989 do show some reductions in the overall government deficit.

[4]The Flow of Funds counts purchases of consumer durable goods in asset accumulation and saving. We do not follow this procedure and have removed consumer durables from the data on household asset accumulation and wealth. The Flow of Funds treatment of consumer durables does not add to the statistical discrepancies between the Flow of Funds and the NIPA because the Flow of Funds uses the NIPA data on durable goods spending.

Table 2

U.S. National Saving as Measured with Flow of Funds Data

(Percent of GNP)

	(1) Household Saving	(2) Corporate Saving	(3) Private Saving	(4) Government Saving	(5) Net National Saving
1953-61	5.6	1.9	7.5	−.9	6.7
1962-73	5.6	2.2	7.8	−.5	7.3
1974-79	5.6	2.7	8.3	−1.5	6.8
1980-89	4.7	1.5	6.1	−2.9	3.2
1980-84	5.5	1.6	7.0	−2.9	4.1
1985-89	3.9	1.4	5.2	−2.9	2.3

decline from the mid-1970s to the late 1980s is comparable in the two systems (although the Flow of Funds shows a decline of about 1 percentage point less if the comparison is made from the 1960s).

We will use the Flow of Funds sectoral asset accumulation data without any definitional or statistical adjustments.[5] We will, however, make some adjustments in the Flow of Funds wealth data to derive a comprehensive national wealth accumulation series. Some of these adjustments are conceptually simple and easy to make; others, however, are more complicated. We present alternative ways of handling the more difficult adjustments.

First, the data are adjusted for biases caused by inflation. In an inflationary period, the nominal value of wealth must grow at least at the rate of inflation to maintain its purchasing power. Accordingly, for all our measures of wealth accumulation we will deduct an estimate of the inflationary component to get a more relevant measure of wealth accumulation trends.[6]

Second, the data are carefully consolidated to avoid double counting. There is no ideal way to measure aggregate wealth accumulation, because one sector may own a claim to the wealth accumulated by another sector. (Household ownership of corporate stock is the most obvious example.) The most natural way to consolidate the nation's balance sheet is to assume that household wealth accumulation accurately represents economy-wide wealth accumulation, since households are the ultimate beneficiaries of the income generated by productive assets. For example, corporate accumulation of productive assets should at least indirectly increase the wealth of household shareowners.

To analyze the role of saving in economic growth, the data should ideally be adjusted so that only assets adding to the economy's productive potential are included in wealth. For example, household wealth includes holdings of government debt. If this government debt is used to finance capital assets (such as roads and bridges), then it should be included in economy-wide wealth, but if it is used to finance current spending (such as government salaries), then it does not add to economy-wide wealth. In practice, changes in government debt, even at the state and local level, do not appear closely related to changes in government capital and therefore may be better left out of our wealth calculations.

Another problem that may require adjustment of the data is the inclusion of corporate stock at market value in the standard household wealth measure. In the short run, increases in plant and equipment owned by corporations may not be reflected in stock market values. There may also be swings in stock market values that do not reflect changes in the productive potential of firms. In particular, changes in tax laws and shifts in investor sentiment can have as strong an impact on stock prices as changes in true productive capacity.

To eliminate some of these distortions, Table 3 presents a number of alternative measures of aggregate wealth accumulation. All these measures net out the increase in wealth necessary to maintain its purchasing power. Column 1 shows the inflation-corrected increase in the Flow of Funds measure of household wealth (excluding holdings of consumer durable goods). This measure includes corporate equity holdings in the form of both direct household ownership and indirect ownership through mutual funds and fiduciaries. Column 2 removes the acquisition of government debt, federal as well as state and local. Column 3 replaces the inflation-corrected increase in the market value of corporate equity holdings with the increase in corporate net worth, a measure which will more closely reflect corporate accumulation of productive capital. (The corporate net worth series values physical assets at their reproduction cost, so it is affected by changes in asset prices in relation to the general price level as well as by actual investment.)

Column 1 shows a marked resurgence in the conventional measure of wealth accumulation in recent years. In fact, wealth accumulation in the second half of the 1980s was stronger than in any period since the 1950s. Column 2 shows that removing private sector accumulation of government debt makes a modest difference to this result, but the resurgence in the late 1980s is still

[5]Except for our removal of consumer durables from Flow of Funds asset accumulation and household wealth. See footnote 4.

[6]The inflation measure used in the calculations was the increase in the implicit price deflator for personal consumption expenditures. Similar calculations were done in Carol Corrado and Charles Steindel, "Perspectives on Personal Saving," *Federal Reserve Bulletin*, vol. 66 (August 1980), pp. 613-26.

Table 3

Household Wealth Accumulation in Excess of General Price Inflation

(Percent of GNP)

	(1)	(2)	(3)
	Total	Excluding Government Debt	Total, with Corporate Equity Valued on Net Worth Basis
1953-61	13.2	12.9	8.9
1962-73	7.9	8.0	9.6
1974-79	8.6	8.3	15.8
1980-89	8.4	6.9	4.6
1980-84	5.2	4.2	3.7
1985-89	11.6	9.7	5.5

evident. Changing the treatment of corporate wealth accumulation, however, makes a crucial difference. Column 3 shows that when corporate net worth is used in place of stock market values, the 1980s as a whole emerges as a period of pronounced weakness in wealth accumulation, especially in comparison with the strength of the 1970s.[7]

The distinction between columns 1 and 3 comes essentially from radically different estimates of capital gains. Table 4 shows capital gains (in excess of general price inflation) on the market value of corporate equity and on corporate net worth. In the second half of the 1980s, capital gains on corporate equity averaged about 7 percent of GNP—compared with capital *losses* of about 2 percent of GNP in the late 1970s. Capital gains on corporate net worth were essentially zero in the 1980s, after amounting to about 2½ percent of GNP in the late 1970s. The capital gains and losses on corporate equity feed into the wealth accumulation series shown in column 1 of Table 3, while those on corporate net worth feed into the series shown in column 3. The sharp divergence in the movement of the two capital gains series in the 1970s and 1980s lies behind the divergent movement of the two wealth accumulation series.

Saving as the supply of capital

The decline in national saving in the 1980s did not necessarily result in a one-for-one drop in productive investment. First, the official data may misclassify some categories of spending. On the one hand, although both consumer durables and government capital expenditures are classified as current spending, they may be more akin to investment. On the other hand, some components of investment may add less to productive capacity than others. Plant and equipment investment in some ways is very different from inventory and residential investment. Second, foreigners are responsible for a portion of capital formation in the United States. If foreign capital inflows exceed outflows, then national investment will exceed national saving. Third, as noted earlier, increases in the value of existing assets may implicitly add to the supply of capital if these revaluations reflect increases in their productive potential.

The contribution of foreign saving

Net national saving is one possible measure of the supply of capital. This aggregate represents the resources Americans make potentially available for funding productive capital formation. However, capital formation in the United States need not be financed just from domestic sources. The first column of Table 5 repeats the data on trends in national saving shown in Table 1. The second column adds net foreign investment in the United States (excluding foreign purchases of government debt) to net saving. We see that despite the surge in foreign investment in the 1980s this measure of the supply of capital has fallen well below its pace in earlier decades.

The large net capital inflows of the 1980s were the financing counterpart of the U.S. current account deficit. In principle, current account deficits could stem from high levels of domestic investment that draw in foreign funds. In practice, however, the current account deficits of the 1980s mainly reflected high U.S. consumption. Thus, the foreign inflow simply offset part of the weakness in U.S. saving rather than contributing to a high investment rate. Furthermore, in any circumstances, foreign investment inflows are not perfect substitutes for domestic saving. A foreign-owned factory might employ just as many workers and produce as many goods as an American-owned plant. However, the profits from the factory's operations will be earned by the foreign own-

[7] Indeed, if government debt accumulation is removed from the column 3 data, the 1980s look even more anemic, with a decade-average ratio of wealth accumulation to GNP of about 3 percent.

Table 4

Household Capital Gains in Excess of General Price Inflation

(Percent of GNP)

	Corporate Equity†	Corporate Net Worth
1953-61	7.2	1.0
1962-73	0.5	0.0
1974-79	−2.1	2.6
1980-89	4.8	−0.4
1980-84	2.4	−0.7
1985-89	7.2	−0.2

†Including gains on holdings of insurance companies and pension funds.

Table 5

Net Capital Supplies from Saving

(Percent of GNP)

	(1) From National Saving	(2) From National and Foreign Saving
1953-61	6.9	6.4
1962-73	7.9	7.1
1974-79	6.8	6.0
1980-89	3.0	4.0
1980-84	3.8	3.9
1985-89	2.2	4.1

ers, not by Americans. Thus, foreign investments will not produce as much income for Americans—or U.S. GNP—as otherwise equal American investments.[8]

It follows that neither of the two aggregate measures in Table 5 precisely captures the saving available to generate GNP growth.[9] The column 1 measure (net national saving) does not include foreign productive investment in the United States and so understates the accumulation of productive capital. The column 2 measure overstates the accumulation of capital: all else equal, capital owned by foreigners will generate less GNP than capital owned by Americans. Nonetheless, the decline in both measures in the 1980s suggests strongly that the supply of capital to the United States out of U.S. and foreign saving has fallen.

Redefining productive investment

On the expenditure side of the NIPA, national saving equals the sum of net foreign investment, residential and inventory investment, and nonresidential fixed investment. Clearly, not all of these investment categories contribute equally to the growth of productive capacity—only the last is conventionally viewed unambiguously as capital accumulation. For example, it is plausible to argue that residential investment does not add to the productive capacity of the United States in the same way that other categories do.[10] Although this view may represent only a value judgment—homes surely do add to economic well-being—many important issues, such as the future external position of the United States, would hinge more directly on the growth of resources in the business sector of the economy than on the growth of the housing stock.

An opposite problem arises with government investment and spending on consumer durables. Government spending on infrastructure clearly adds to the productive capacity of the U.S. economy, but it is not counted as saving in the NIPA (although in the household wealth data, infrastructure spending financed through debt sales to the private sector would be indirectly counted). The national income accounts in most other countries address this problem by distinguishing between government consumption and investment. Consumer durable spending, like housing, creates a stream of future services to consumers, yet it is counted as consumption rather than investment. The failure to include these investment-like expenditures could distort the saving picture.

Although it is reasonable in principle to include infrastructure spending in the supply of capital, in practice actually counting such spending is difficult. Not only is it difficult to classify government capital outlays as productive or nonproductive, but it is also difficult to calculate service lives and depreciation schedules for such unique assets. (What is the true service life of an airport?) Consumer durables pose similar problems. Nevertheless, including these categories in saving does not change the overall picture. While it is true that government investment and consumer durable expenditures are fairly sizable, they also show the same downward trend as private saving in the 1980s, especially outside the military. The decline in government spending on structures was noted earlier; consumer durable spending fell from over 12 percent of GNP in the middle 1970s to less than 9 percent in the 1980s.[11]

Are capital gains a form of saving?

A more critical issue in analyzing the connection between saving trends and capital formation concerns the treatment of capital gains. From an individual perspective, capital gains can properly be included in saving: the fundamental purpose of saving is the accumulation of wealth so that consumption may be higher in the future, either for an individual or for his or her heirs. Anything that adds to wealth can, from an individual's viewpoint, be considered saving. From a policy viewpoint, the aggregate wealth accumulation data shown in column 1 of Table 3 are important because they give a sense of how rapidly or slowly U.S. consumers are reaching targeted levels of wealth. The acceleration of capital gains in the 1980s probably played an important role in the strong growth of consumer spending in the period and the weakness in

[8]These issues are discussed more fully in M.A. Akhtar, "Adjustment of U.S. External Balances," in Federal Reserve Bank of New York, *74th Annual Report, 1988*.

[9]Another issue is the treatment of depreciation. All saving measures used in this paper are net of depreciation. Net saving is available for use in increasing the net capital stock. If productive capital is defined in terms of the gross capital stock, then analysis might better focus on gross saving trends (gross saving is net saving plus depreciation). No definitive case can be made for the superiority of the net capital stock to the gross stock as a measure of productive U.S. capital. See A. Steven Englander and Charles Steindel, "Evaluating Recent Trends in Capital Formation," Federal Reserve Bank of New York *Quarterly Review*, vol. 14, no. 3 (Autumn 1989), pp. 7-19. However, the trends in gross saving are similar to those in net saving.

[10]Of course, on an ex ante basis the saving flows invested in residences could have been invested in other forms of capital.

[11]For a discussion of the importance of infrastructure spending, see Michael J. Boskin, Marc S. Robinson, and Alan M. Huber, "Government Saving, Capital Formation and Wealth in the United States, 1947-85," in Robert E. Lipsey and Helen Stone Tice, eds., *Measurement of Saving, Investment, and Wealth*, pp. 287-356 (Chicago: University of Chicago Press, 1988). It has been argued that the decline in infrastructure growth has played a critical role in the weakness of U.S. productivity growth since 1973. See David A. Aschauer, "Is Public Investment Productive?" *Journal of Monetary Economics*, vol. 23 (March 1989), pp. 177-200.

national saving. However, if rapid growth in household wealth is due mainly to revaluations of existing assets, or to government debt issuance, the increased wealth may not represent increased productive capacity. Hence, the supply of saving available for capital formation may be inadequate, and increases in consumption stemming from capital gains may not be sustainable.

Table 1 reflected this traditional view by omitting capital gains and losses. However, some portion of aggregate capital gains will reflect increases in the true productive power of assets, and we may legitimately include these gains when we compute saving as the supply of capital available for increases in productive capacity.

Table 6 attempts to construct comprehensive measures of the supply of capital. In Table 6 a portion of aggregate capital gains (as always in this article, over and above overall price inflation) is added to the Table 5 saving flows.[12] Gains and losses on residential real estate are excluded because realistically they are not part of and cannot be made available for productive investment,[13] but gains and losses on other assets are included. Essentially, columns 1 and 2 of Table 6 correspond to columns 1 and 3 of Table 3 but exclude changes in the value of government debt and capital gains on residential real estate. Columns 3 and 4 of Table 6 add to columns 1 and 2 changes in the value of foreign asset holdings. Columns 1 and 3 use stock-market-based valuation of corporations; columns 2 and 4 value corporations on a reproduction cost basis.

The key point in Table 6, as in Table 3, is that it makes a crucial difference whether corporate wealth is valued by the stock market or by reproduction cost. If we value corporations by the stock market, the increase in productive wealth as a share of GNP in the 1980s was sharply higher than in the 1970s and even somewhat higher than in the 1960s. If we value corporate assets at reproduction cost, the accumulation of wealth was quite low.

It seems natural to disregard the stock-market-based data since, by construction, the reproduction cost measures are more closely related to the increase in the officially measured stock of physical capital. Nevertheless, the stock market measures should be influenced by expected future streams of earnings, and increases in these measures may pick up expected future increases in the productivity of the existing capital stock (say from future improvements in technology). In this sense, at least some share of capital gains in the stock market may represent a form of "investment" and "saving."

A simple way to test whether all stock market capital gains actually reflect future increases in productive capacity and output is to use lagged changes in a capital input measure derived from stock market data in an aggregate production function. Details of this exercise can be found in the Box. In general, the results indicate that stock-market-based series have had little ability to predict future output. This finding means that we can reject the hypothesis that all past stock market capital gains and losses reflected changes in the future productive power of capital. Thus, in the past, not all changes in stock market wealth were a form of "investment." Our finding further suggests that a measure of the amount of saving actually available for capital formation might well exclude stock-market-based capital gains and losses, although it is certainly true that stock market fluctuations will in part reflect changes in the long-run potential of the economy. The problem is that we can neither readily differentiate this source of market fluctuations from others, nor assume that all market changes reflect changes in long-run potential.

Once we have recognized the limitations of the stock-

[12]The actual asset accumulation flows used in Table 6 were taken from the Flow of Funds. They differ from the NIPA saving series for reasons discussed earlier.

[13]Actual purchases of residential real estate are included in the aggregate supply of capital calculations because ex ante they were available to be invested in other forms.

Table 6

Estimates of the Supply of Capital

(Percent of GNP)

	Domestic Sources Only		Domestic and Foreign Sources	
	(1) Equity Market Valuation	(2) Reproduction Cost Valuation	(3) Equity Market Valuation	(4) Reproduction Cost Valuation
1953-61	13.0	8.7	12.6	8.2
1962-73	7.0	8.7	6.8	8.6
1974-79	5.7	12.9	5.2	12.8
1980-89	7.8	4.0	8.9	5.0
1980-84	6.1	4.7	6.5	5.0
1985-89	9.5	3.4	11.3	4.9

Box: The Stock Market as a Measure of Saving

Measures of wealth based on the stock market rose more rapidly in the 1980s than those based on the reproduction cost of capital. If the increase in the stock market represents a form of productive saving for the future, then past experience should show that a measure of the capital stock based on the stock market is a good indicator of future output growth.

We assume that output can be explained by a Cobb-Douglas production function, which in logarithmic terms can be written as

$$\ln Y = a + \alpha \ln L + (1-\alpha) \ln K + \lambda t,$$

where Y equals output; L, labor input; K, capital input; λ, the rate of total factor productivity growth; and t, time.†

The expression can be restated as

$$\Delta \ln(Y/L) = (1-\alpha)\Delta \ln(K/L) + \lambda.$$

We estimated this equation for the nonfinancial corporate sector, comparing a number of measures of the capital input. One measure is the standard net nonresidential capital stock; the others are derived from stock market pricing of the capital stock. We assumed that all the difference between nonfinancial corporate net worth and the market value of nonfinancial corporate equity can be assigned to different valuations of the capital stock. A constant-dollar valuation of the capital stock based on the stock market was derived by dividing this nominal value by the implicit capital stock deflator. If changes in the stock market are truly indicative of future increases in productivity and output (thus making these changes a form of saving), a lag on this measure should help explain output. Accordingly, simple three-year and five-year moving averages of the real stock market capital variable were used as proxies for K.

The estimated equations included a number of standard corrections for cyclical productivity changes and shifts in trend productivity growth. The table reports the coefficients of the capital input variables and the equations' residual standard errors. We see that the three-year and five-year stock market variables show little sign of being adequate proxies for the capital input. The estimated coefficients on these variables, which ideally should equal capital's share of income (about one-third), are barely positive. The coefficient on the net capital stock, though considerably higher than anticipated, is more plausible (it is possible that this term picks up the effect of omitted variables such as inventories and natural resources). Finally, the standard errors of the regressions with the stock market variables are quite high relative to the standard errors in the regression using the conventional capital stock variable.

†For further details, see Englander and Steindel, "Evaluating Recent Trends."

Performance of Capital Input Measures in Production Relationships

Capital Input Measure	Coefficient	Equation Residual Standard Error
Net capital stock	.770	.01
Stock-market-based measure (three-year average)	.005	.07
Stock-market-based measure (five-year average)	.043	.07

Note: The estimated equations are of the form:

$$\ln(\text{lprod}) = a_0 + a_1 \ln(\text{caphrs}) + a_2\, \text{cycl} + \sum_{i=3}^{7} a_i T_i,$$

where

lprod = nonfarm business sector labor productivity
caphrs = the ratio of capital to hours worked
cycl = a measure of capacity utilization (the ratio of actual to potential real GNP, as calculated by the Federal Reserve Board staff)
T_i = a set of time trends (allowing for breaks in 1952-61, 1962-68, 1969-73, 1974-79, and 1980-88).

market-based data, the evidence on U.S. saving trends becomes clearer: saving, in the sense of supplying capital for the expansion of the U.S. economy, has reached extremely low levels. This conclusion is true for all three concepts of saving; it is true whether we include or exclude foreign capital; and it is true regardless of how broadly we define investment.[14]

The case for higher U.S. saving

Clearly U.S. saving in the 1980s was low by just about any measure, but should this be a cause for concern? On the surface, the economy performed reasonably well despite the low saving rate. Indeed, the declining saving rate spurred consumption, contributing to the cyclical recovery following the 1982 recession. For the decade as a whole, annual growth in GNP per capita averaged only about 0.5 percentage point below the postwar average.[15] U.S. external debt grew, but the foreign investors brought in new capital, and net indebtedness to foreigners remained small as a share of GNP. If low saving has hurt the U.S. economy, the effects are well disguised.

The remainder of this article uses a simulation model of the U.S. economy to uncover, and quantify, the subtle costs of low saving. The model links the three basic components of growth—saving and investment, labor force growth, and technological advance—to economic growth and the U.S. external debt position. Three variations of the model are employed to accommodate the diverse views held by growth experts about the interaction between investment and technological change: the traditional model that considers technology to be independent of investment and two alternative forms of the model that regard new capital investment as a spur to technological change. Details of the model are presented in the Appendix.

With this model we ask: What has been the cost of low saving in terms of economic output, living standards, and external indebtedness? What would be required to rectify the situation? Would the sacrifice of current consumption be worth it? And finally, what are the limits of what higher saving rates can accomplish?

The legacy of the 1980s

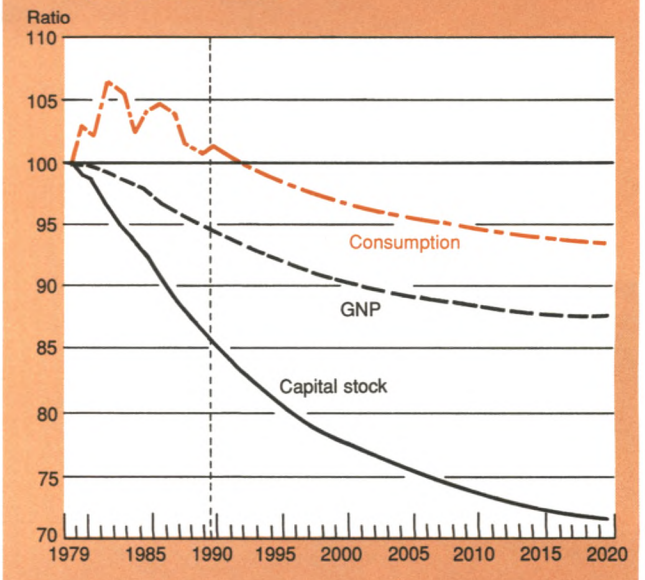
Low saving in the 1980s left the U.S. economy with a relatively low capital stock, low output, and a large foreign debt. A simple way to quantify the damage is to compare two saving scenarios (using the NIPA saving data). The first, a "status quo" scenario, assumes that the net saving rate followed its actual path in the 1980s, falling to about 2.0 percent in 1990, and will remain at that level into the future. The second, the "1950-79 trend" scenario, assumes that the net saving rate never fell from its 1950-79 average of about 7½ percent.

Chart 1, using traditional model estimates, shows that low saving made possible a surge in consumption in the 1980s but at considerable long-run cost. By 1989 low saving had cost the U.S. economy about 15 percent of its capital stock and about 5 percent of its potential GNP. Furthermore, by the end of the century, the accumulated loss could grow to 28 and 10 percent, respectively. In fact, even the gain to consumption should be short-lived: by the early 1990s weak economic growth should push consumption below the 1950-79 trend scenario.

The U.S. net external debt position suffered as well. In the 1970s, with U.S. saving and investment roughly in balance, the United States was a modest net capital exporter. As the saving rate declined in the 1980s, however, an increasing portion of investment was financed by net foreign capital inflows. Of course, the

Chart 1

Impact of Low Saving



Note: The chart shows the status quo scenario relative to the 1950-79 trend scenario.

[14]The United States was not the only nation to see a decline in saving in the 1980s. See Andrew Dean, Martine Durand, John Fallon, and Peter Hoeller, "Saving Trends and Behavior in OECD Countries," *OECD Economic Studies*, no. 14 (Spring 1990), pp. 7-58.

[15]Real GNP growth averaged 2.6 percent from 1980 to 1989, down from 3.6 percent over the previous thirty years. In per capita terms, the decline was more modest, reaching 1.6 percent in the 1980s from 2.2 percent in the previous period.

Table 7

U.S. External Position in Book and Market Value

(Percent of GNP)

	Book Value			Market Value		
	Assets	Liabilities	Net Assets	Assets	Liabilities	Net Assets
1979	20.4	16.6	3.8	26.0	17.9	8.1
1989	27.0	39.7	−12.7	42.2	44.7	−2.5
1999†	36.3	55.6	−19.3	48.5	63.5	−15.0
2009†	46.1	68.9	−22.8	58.3	81.2	−22.9

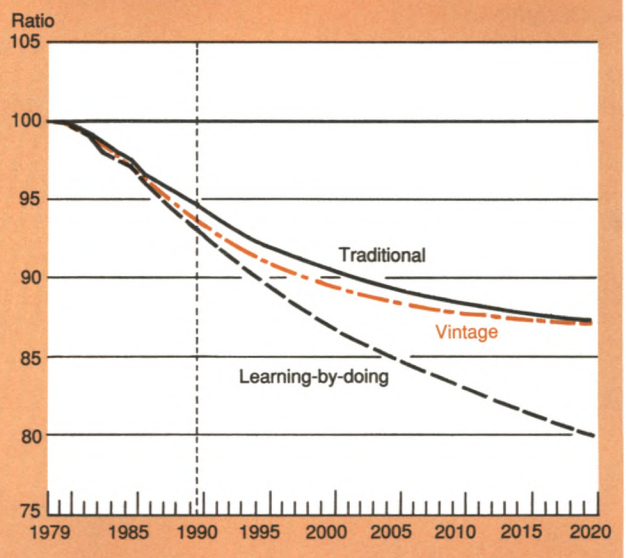
†Projections assume a status quo scenario with net capital inflows of 1.6 percent of GNP per year, GNP growth of 6.3 percent per year, and growth in total capital flows of 8 percent per year.

causes of the surge in external debt are complex, but it is reasonable to argue that low saving played a key role.[16] Foreign investments in the United States grew fivefold in the 1980s, while U.S. investments abroad grew less than threefold. As a result, the United States went from a net creditor position of 4 percent of GNP in 1979 to a net debtor position of 13 percent a decade later (Table 7). If net capital inflows continue at the current pace of 1½ percent of GNP, U.S. net indebtedness will reach 19 percent of GNP by the turn of the century and eventually grow to a peak of over 25 percent.[17]

The cost of low saving may be even greater than the traditional model suggests. Many recent studies of economic growth have emphasized the link between capital formation and technological change. If capital "embodies" new technology, a decline in saving may be a double blow to the economy: not only is there less capital, but existing capital also becomes increasingly outdated. Chart 2 shows the GNP path under three assumptions about technological change: the traditional model, in which technology is independent of investment; the "vintage" model, which assumes higher investment lowers the average age and adds to the productivity of capital; and the "learning-by-doing" model, which assumes that new investment not only adopts the latest technology, but actually encourages further innovations.

Chart 2

Impact of Low Saving on Real GNP under Alternative Models



Note: The chart shows the status quo scenario relative to the 1950-79 trend scenario.

The alternative models suggest much stronger impacts on GNP from lower saving than does the traditional model. In both alternative models the level of GNP is an additional 1 to 2 percentage points lower by 1989. In the vintage model these technology effects eventually peter out.[18] In the learning-by-doing model,

[16]The economic mechanism was as follows: low private saving and high government borrowing put upward pressure on U.S. interest rates; this made U.S. investments relatively more attractive, encouraging net capital inflows; these inflows in turn took some of the pressure off of interest rates. The overall result was that U.S. net investment fell about one-third less than net saving during the 1980s, with the difference accounted for by increased capital inflows.

[17]A number of measurement problems plague the official data. Adjusting the data, however, does not alter the general picture of a deteriorating trend. For example, in the official data, direct investment is measured at book values, a procedure which understates the value of the generally older U.S. investments abroad. When the data are adjusted to market values (as shown in Table 7), the level of net debt is much lower but the trend is nearly as bleak: from 8 percent in the black in 1979 to 3 percent in the red in 1989, and 15 percent in the red by the turn of the century. For details behind these calculations, see Appendix.

[18]As the economy approaches its long-run growth path, growth in the capital stock slows so that new capital is no longer a dispropor-

however, lower investment means a slower pace of "learning" and technological innovation, permanently reducing annual GNP *growth* by 0.3 percent.

The benefits of a saving recovery

It is not too late to undo the damage of the 1980s. Charts 3 and 4 present traditional model estimates of the impact of a saving recovery. Chart 3 shows the impact on economic growth of a "recovery" scenario, in which the net saving rate rebounds to its historical average of about 7½ percent over the next five years. Chart 4 puts this recovery in perspective, comparing it with the 1950-79 trend scenario (in which the saving rate never declines) and with the status quo scenario (in which the saving rate remains at its 1990 level of about 2 percent). All three scenarios assume that the labor force grows in line with the official projections of the Social Security Administration. The simulations also assume a 1 percent contribution to growth from technological advance.

In the status quo case the economy continues along its current low growth path until about 2010 (Chart 3). Saving is just sufficient to replace worn out capital and provide for net capital growth of about 2 percent. Real output, consumption, and investment also settle into an equilibrium growth path of 2 percent. Early in the next century, as the "baby boom" generation begins to retire, the labor force stops growing, pulling down GNP growth to about 1½ percent. With the saving rate unchanged, the United States continues to rely on foreign capital inflows equivalent to 1½ percent of GNP.

An increase in the saving rate upsets this equilibrium. Higher saving flows into investment, and net capital stock growth surges, pulling up GNP growth as well. The GNP growth acceleration is relatively modest but can extend over a long period of time. In the first five years of the saving recovery, GNP growth averages 0.5 percentage point higher than in the status quo; the differential then notches up to 0.8 percent in the second five years and declines thereafter. The growth expansion is self-limiting, however. In the years following the saving rebound, the capital stock rises relative to GNP but saving remains fixed as a share of GNP, so growth in the capital stock tends to slow over time. As a result, consumption, investment, and GNP all gradually settle into a new path at higher levels, but with the growth rate back at its original pace of about 2 percent (Chart 4).[19]

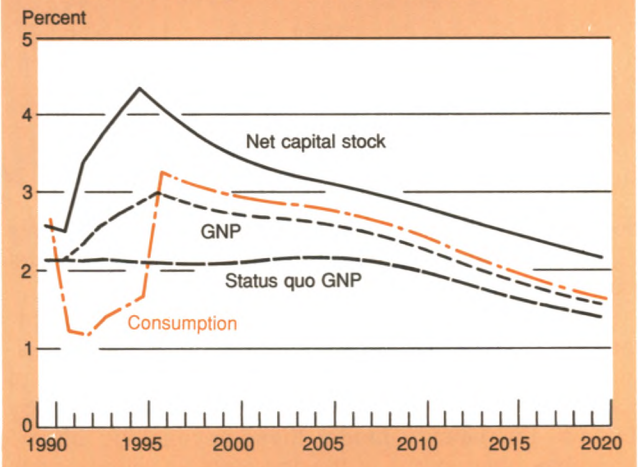
Eventually the temporary fall in saving is "forgiven,"

[19] In the traditional model, permanently raising the GNP growth rate requires an ever-increasing saving rate. For example, raising growth by 1 percent would require increasing the saving rate by about 0.4 percentage point each year *for as long as growth is to be kept higher.*

Footnote 18 continued
tionate part of the total. At this point the average age of the capital stock stops declining and technological advance returns to its long-run trend. See Appendix for details.

Chart 3

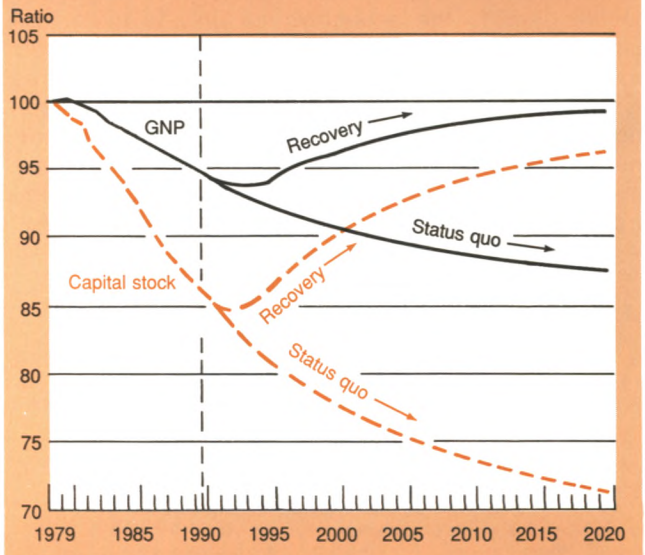
Impact of a Saving Recovery on Real Growth



Note: Chart shows growth in net capital stock, GNP, and consumption for the saving recovery scenario.

Chart 4

Comparison of the Recovery and Status Quo Scenarios



Note: Each series is plotted relative to the 1950-79 trend scenario.

in the sense that GNP returns to its 1950-79 trend path. As Chart 4 shows, however, this recovery can take a considerable period of time. By 2010, twenty years after the saving rate rebounds, real GNP and the capital stock are still 1.8 percent and 6.0 percent, respectively, below their 1950-79 trend scenario levels.

The benefits of higher saving are greater if the new investment also encourages technological change. Chart 5 shows the recovery and status quo paths for GNP under both the traditional model and the alternative learning-by-doing model.[20] The low saving of the status quo scenario is especially damaging in models with embodied technology. Unlike the traditional model, in which the loss of output eventually stabilizes, the learning-by-doing model shows GNP steadily dropping relative to trend. Clearly a saving recovery is preferable to this steady loss of output. Nevertheless, even with the saving recovery, the learning-by-doing model never fully forgives the saving shortfall of the 1980s: GNP settles below the trend scenario because of the permanent lost learning of the 1980s.

A rise in the saving rate could also cause a dramatic improvement in the U.S. external asset position (Chart 6). During the 1980s increased foreign capital inflows replaced roughly one-third of the drop in net national saving, preventing an even more dramatic drop in net national investment. As a reasonable first approximation our model assumes that this process reverses when net national saving rebounds. With saving restored to its pre-1980s average, the United States could again become a (modest) net capital exporter, causing a steady decline in the net debt position as a share of GNP.[21] Note, however, that this improvement in the nation's net debt position does not eliminate our dependence on foreign capital. Instead, continuing efforts by investors to diversify their portfolios internationally should mean continued rapid growth in both external liabilities and assets, even as the gap between the two narrows.[22]

The long-run consumption reward

If the sole goal of economic policy is to maximize output, a higher saving rate is always better, and the only policy question is how to get it higher. Presumably the principal goal of higher saving is not just higher GNP, but higher living standards as well. The ultimate test of saving policy, therefore, is whether it improves the time profile of consumption.

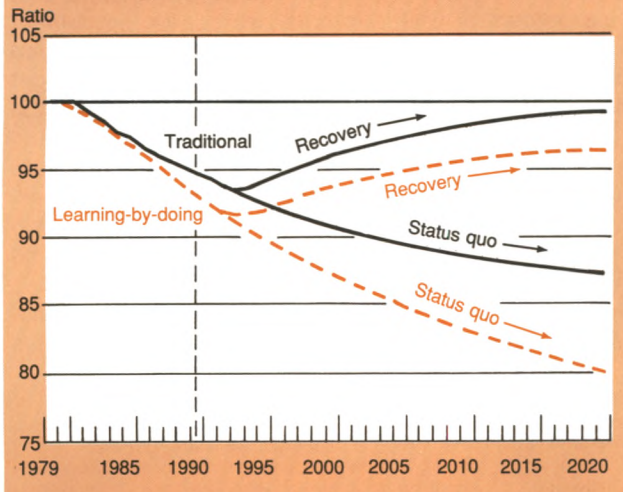
[20] To keep the chart uncluttered, the "vintage" model is excluded. As Chart 2 suggests, the vintage model behaves like the learning-by-doing model in the short run and the traditional model in the long run.

[21] As the Appendix shows, if net capital inflows do not decline in response to the saving recovery, the benefits to GNP from higher saving are even greater.

[22] In the 1980s, total capital flows, the sum of inflows and outflows, grew at an annual rate of 12.7 percent, almost twice as fast as nominal GNP growth of 7.6 percent. Our simulations assume that this diversification continues at a more modest pace in the future, with 8 percent growth in total capital flows and 6½ percent growth in GNP. See Appendix for further details.

Chart 5

Impact of a Saving Rate Recovery on Real GNP under Alternative Models



Note: Each series is plotted relative to the 1950-79 trend scenario.

Chart 6

International Capital Position in the Recovery Scenario

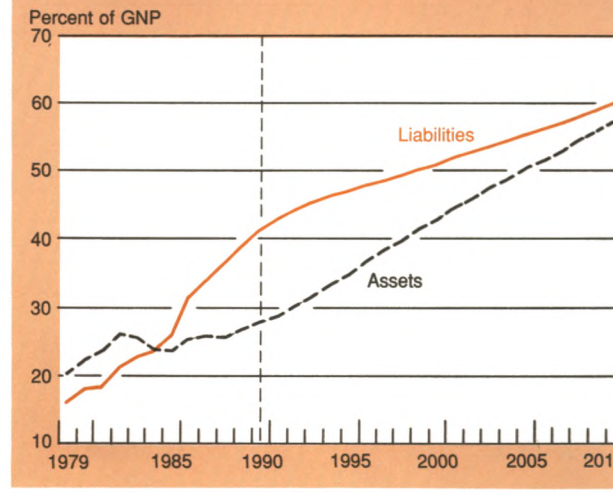


Chart 7 compares aggregate consumption for the status quo and recovery scenarios with aggregate consumption for the 1950-79 trend. A permanently lower saving rate can have considerable immediate benefit to consumption. Had the saving rate remained on trend in the 1980s, consumption would have averaged 3.3 percent less than it actually was. Yet the payback for this consumption binge is already being felt. By 1992, if the status quo continues, the cumulative income loss from low saving will already have pushed consumption below the trend level. Most of the costs of the consumption binge would be felt in the next century when consumption drops to 10 percentage points below trend and then stays there forever. (This is what some commentators mean when they say that low saving has "mortgaged" our future.)

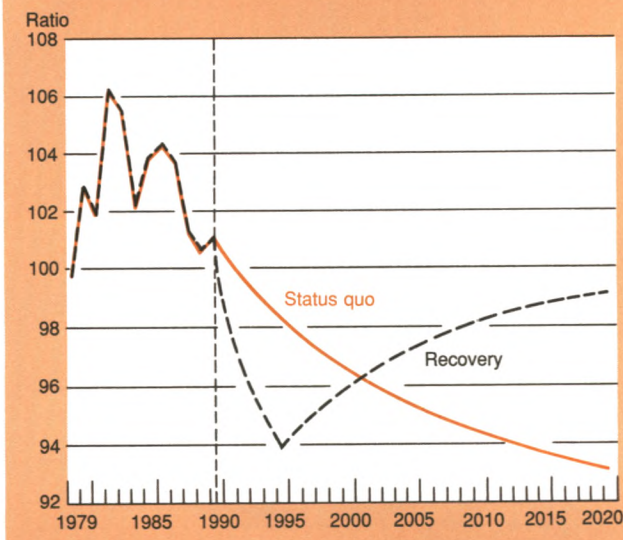
In the recovery scenario the excesses of the 1980s are reversed in the 1990s. Consumption drops sharply and then gradually converges back to the trend level. In fact, in the traditional model the highest possible consumption path is achieved with gross saving rates of close to 30 percent.[23] If saving rises above 30 percent, the gain to consumption from higher output is more than offset by the need for more resources to maintain the capital stock.

The consumption "reward" for higher saving can take many years to materialize. For example, consumption in the recovery scenario will not surpass consumption in the status quo scenario until the turn of the century (Chart 7). Determining whether this delayed gratification would be worthwhile requires some measure of people's time preference—the rate of discount that equates the utility of current and future consumption. The precise magnitude of this discount rate is a matter of considerable dispute, yet under any reasonable assumption the saving rate appears low. For example, even if the discount rate is as high as 5 percent and if people's time horizon is only thirty years (they essentially do not care about unborn generations), society is still better off under the recovery scenario than under the status quo. Indeed, the status quo saving rate would only be justified if the discount rate were as high as 10 percent.[24]

[23]The growth literature calls this the "golden rule" consumption path.

Chart 7

Real Consumption under Alternative Saving Scenarios



Note: Each series is plotted relative to the 1950-79 trend scenario.

Difficult, but not impossible

Restoring the capital stock to its pre-1980s path will require a sustained increase of more than 5 percentage points in the national saving rate. This task appears particularly daunting in view of the downward trend of the last decade. Furthermore, with most studies showing only a limited response of private saving to incentives, the main burden of adjustment must fall on the worst saver of all—the federal government.

Recent efforts to reduce the budget deficit suggest that significant progress can be made. The Congressional Budget Office estimates that in the absence of the fall 1990 budget accord the 1995 budget deficit (excluding revenues from the sale of thrift assets) would have been 2.8 percent of GNP. The accord reduced this to 0.8 percent of GNP—in other words, the agreement goes roughly two-thirds of the way toward a balanced budget. Continuing this process all the way to budget balance would increase the national saving rate a total of 3 percentage points from current values. By the end of this decade this partial rebound in the saving rate could add about 7½ percent to the U.S. capital stock, increase real GNP by about 2 percent, and slow growth in net external debt to the point where it would no longer be increasing as a share of GNP.

Not only should the budget deficit be eliminated, but a case can also be made for a budget surplus. With a

[24]If saving is "suboptimal," why don't people save more? In many ways government tax and spending policy tends to discourage saving and redirect income to less productive uses. But even without these distortions, private saving decisions are probably not "socially optimal." By ensuring a healthy economy, with a growing economic pie, saving contributes to social stability and confidence in the future. For these and other reasons, private decisions may yield less saving than what is collectively desirable.

budget surplus the Federal government could begin buying down debt accumulated in the 1980s, freeing up resources for private investment. One way to generate a surplus would be to balance the non-Social Security portion of the budget. The annual surplus of the Social Security System now offsets about one-fourth of the deficit in the rest of the budget. By the turn of the century, the Social Security surplus could reach 2 percent of GNP. Balancing the budget exclusive of Social Security trust funds and pushing the unified budget into surplus, therefore, could raise the saving rate almost 5 percentage points. This is close to the saving path suggested by the recovery scenario discussed earlier.

Conclusion

The 1980s saw net national saving fall to its lowest rate of the postwar period. All measures of saving that estimate the actual acquisition of productive assets confirm this finding. The costs of this poor performance have been subtle but quite real: temporarily higher consumption has been gained at the long-run expense of several years' worth of GNP growth and a complete reversal of the U.S. external debt position. In particular, our simulation results show the following:

- Traditional model estimates indicate that the drop in saving in the 1980s has already cost the U.S. economy about 15 percent of its capital stock, lowering potential output by about 5 percent. By the end of this century, if the status quo continues, the accumulated loss in capital and output will grow to 28 percent and 10 percent, respectively.
- The actual cost may be even greater. In an alternative, learning-by-doing model, which links capital formation to the pace of technological innovation, the estimated loss to potential output was over 7 percent in 1990 and could rise to about 15 percent by the year 2000.
- Foreign capital inflows in the 1980s prevented an even greater shortfall in the capital stock, but in the process the United States has gone deeply into debt. At current rates of net capital inflow, in ten years the United States will pay more than 1 percent of its annual income to service this foreign debt, an exact reversal of its position ten years ago.
- The U.S. net saving rate would have to climb 5½ percentage points as a share of GNP to offset the decline of the 1980s, restore the trend in capital growth, and end the deterioration of our external debt position.
- Most of this gap could be closed by balancing the federal budget excluding the Social Security surplus. The recent budget accord is a significant step in this direction.
- Raising the saving rate will require lower current consumption. The present saving rate can be justified only if people put a very low value on future consumption compared with present consumption. If we assume a reasonable "discount rate" of 2 percent per year—roughly the real return to government bonds in the postwar period—lifetime consumer satisfaction is maximized with a net saving rate four times the current pace.

A higher saving rate is not a cure-all for the nation's ills. Higher saving means a higher level of output, but it does not sweep away the inflation and unemployment problems of the business cycle. Although higher saving would probably reduce the nation's *net* foreign indebtedness, it will not mean an end to the gross inflow of foreign capital. Furthermore, not all saving is equally productive. The growth benefits of higher saving could be greatly increased by eliminating tax distortions favoring less productive investments. Finally, in the 1980s not only did private spending shift out of investment into consumption, but public investment lost out to current spending as well. A healthier economic outlook will require redirecting all kinds of spending toward investment—not only in plant and equipment, but also in infrastructure, education, environmental safeguards, and research.

Appendix: The Growth Model

All growth simulations in the text are based on a detailed neoclassical representation of the U.S. economy. This simple growth framework is a powerful tool for exploring alternative paths for the economy. Offering a clear connection between results and assumptions, the framework can be easily manipulated, and it has a long track record of use in previous research.† The model also has some disadvantages. It is highly simplified, lumping capital and output into very broad aggregates. It also ignores the short-run costs of changing the saving rate, focusing instead on the long run.‡ This Appendix reviews the main equations of the model and then tests the robustness of the results to changes in model parameters. Several notation conventions are followed: a "%" before a variable indicates a growth rate, "(-1)" means "lagged one period," "Δ" signifies the change from a year ago, and a "C" suffix means "measured in constant 1982 dollars."

Labor

Labor input is measured by aggregate hours worked. Growth in hours is assumed to equal the growth in working age population plus an add factor to account for increased participation rates:

$$(1)\quad \%LAB = \%POP + ADD.$$

Most simulations use the "middle" population projections of the Social Security Administration§ and assume an add factor of 0.1 percent. Variables measured in per capita terms also use the Social Security Administration's projections for total population.

Capital

The model treats the nominal gross saving rate as exogenous, distributes saving among its various uses, and then calculates the implied capital accumulation. The basic saving identity determines the share of GNP going to investment:

$$(2)\quad I/GNP = S/GNP + NFI/GNP,$$

where S is gross national saving, I is gross investment, and NFI is net foreign capital inflows. We assume that, in line with the 1980s experience, net foreign capital inflows decline by one-third of any improvement in the saving rate:‖

$$(3)\quad NFI/GNP = [.015 - .333*(S/GNP - .13)].$$

Investment is divided between residential (IR), nonresidential (IN), and inventories (IV). We assume that, consistent with the 1979-89 trend, a gradually declining share of investment goes into the residential sector and a fixed portion goes into inventories:

$$\begin{aligned}(4)\quad I &= IR + IN + IV \\ &= (.3 - .00178*T)*I + (.67 + .00178*T)*I \\ &\quad + .03*I,\end{aligned}$$

where T is a time trend equal to 1 in 1991.††

These nominal investment flows are converted to real investment by subtracting the assumed rate of inflation of 4 percent. A portion of real nonresidential investment (INC) is allocated to the farm (IFC) and "other" (IOC) sectors, and the remainder goes to nonfarm business (INFBC):

$$(5)\quad INFBC = INC - IFC - IOC.$$

These investment flows, along with assumed depreciation rates, determine capital accumulation. The key capi-

†The basic framework was developed in Robert M. Solow, "A Contribution to the Theory of Economic Growth," *Quarterly Journal of Economics*, vol. 70 (1956), pp. 65-94. Recent extensions of the model include Paul M. Romer, "Increasing Returns and Long-Run Growth," *Journal of Political Economy*, vol. 94, no. 5 (1986), pp. 1002-37; Robert E. Lucas, "On the Mechanics of Economic Development," *Journal of Monetary Economics*, vol. 22, pp. 3-42; and Maurice F. Scott, *A New View of Economic Growth* (Oxford: Clarendon Press, 1989). Two recent applications are a study of the Social Security system, Henry Aaron et al., *Can America Afford to Grow Old?* (Washington D.C.: Brookings Institution, 1989); and a study of demographic trends, Keith Carlson, "On Maintaining a Rising U.S. Standard of Living into the Mid-21st Century," Federal Reserve Bank of St. Louis *Review*, vol. 72, no. 2 (1990), pp. 3-16.

‡A sharp increase in the saving rate could cause the economy to weaken if slower consumption growth is not immediately offset by increased investment. Policy makers could mitigate some of the short-run impact. In any event, these initial effects will be unimportant in the long run.

§*1988 Annual Report of the Board of Trustees of the Federal Old-Age and Survivors Insurance and Disability Insurance Trust Funds* (Washington, D.C.: Government Printing Office).

‖From 1979 to 1989, net national saving declined 5.1 percentage points as a share of GNP; over the same period, capital inflows increased 2.0 percentage points relative to GNP, replacing 38 percent of the lost saving.

††To keep the model simple, we ruled out linking housing to population growth. This simplifying assumption has little bearing on the results.

Appendix: The Growth Model *(continued)*

tal equation is for nonfarm business (KNFBC):

$$(6)\ KNFBC = INFBC + (1 - \delta)*KNFBC(-1),$$

where "δ" is the depreciation rate and is assumed fixed at .111 for most of the simulation period. Residential capital is similarly determined by adding new investment (IRC) to lagged capital (KRC) and subtracting depreciation, but with a .025 depreciation rate.[‡‡]

Output

Real GNP is divided into six components, value added by nonfarm business (RNFBC), services from housing (RRC), farm (RFC), government (RGC), "other" (ROC), and the rest of the world (ROWC):

$$(7)\ GNPC = RNFBC + RRC + RFC + RGC + ROC + ROWC.$$

"Other" and government are assumed to grow at the same rate as aggregate hours plus a constant, and farm output is assumed to grow at a fixed rate. The rest-of-world component is a linear function of the accumulated external asset position, and it assumes a 7.5 percent return on new flows (see the next section). Residential output (RRC) is measured as a simple product of the return to housing services and the real stock of housing (KRC):

$$(8)\ RRC = .085*KRC.$$

Nonfarm business output is modeled using the traditional Cobb-Douglas formulation, as well as two variations with different assumptions about technological progress. The basic model is:

$$(9)\ RNFBC = [1 + .7*\%LAB + .3*\%KNFBC + \%TECH1]*RNFBC(-1),$$

where technological advance (%TECH1) is assumed to add a constant 1 percentage point to output growth. In the vintage model, %TECH2 depends on the average age of the capital stock:

$$(10)\ \%TECH2 = .01 - .02*\Delta AGE,$$

and AGE depends on the rate of gross investment:[§§]

$$(11)\ \Delta AGE = .84 - .92*(INFBC/KNFBC)*AGE(-1).$$

The vintage model implies that a one-year drop in the average age of capital adds 2 percentage points to output. In the learning-by-doing model, technological advance is a linear function of the rate of investment:

$$(12)\ \%TECH3 = .06*(I/GNP).$$

With an investment rate of, say, .18, this equation implies a contribution to growth from new technology of 1.1 percent per year. The parameters for both the vintage and learning-by-doing models are calibrated so that they explain roughly half of the contribution to growth from technology innovation in the postwar period. In other words, half of technology advance is assumed to be embodied in capital and the remainder is assumed to be independent of capital formation.

External asset position

The U.S. external asset position is equal to last period's net assets minus the current period's net capital inflows:

$$(13)\ NETA = NETA(-1) - NFI.$$

Net capital flows are determined as shown in equation 3 above. In addition, the net asset data are adjusted from book value to market value to account for the undervaluing of direct investments. This undervaluing of investment is particularly large for the older U.S. investments abroad. The methodology, which relies on stock market values, is drawn from a paper by Michael Ulan and William G. Dewald.[¶¶] For all future years we assume that average stock market values grow 6 percent per annum.

Model characteristics and sensitivity

With nonfarm business accounting for about 80 percent of output, the model behaves very much like a pure Cobb-Douglas model. Higher saving boosts growth, although some of the effect is mitigated by the leakage of capital abroad and the failure of some sectors to respond to the higher saving and capital formation. Once the saving rate stabilizes at a new higher level, the capital-output ratio and the growth in output, investment, and

[‡‡]We assume that, except for a transition period, inflation is the same for investment and noninvestment goods (4 percent) and that depreciation is constant. We also assume that during the first ten years, in line with recent experience, nonresidential investment inflation is only 3 percent and the depreciation rate rises .001 per year. These assumptions have roughly an offsetting impact on real capital formation.

[§§]The coefficients are derived from a regression for the period 1949-89.

[¶¶]"Deflating U.S. Twin Deficits and the Net International Investment Position," U.S. Department of State, Working Paper Series no. 12, 1989.

Appendix: The Growth Model *(continued)*

consumption all settle on a roughly constant long-run path. There is a minor tendency for growth to slow over time because of the expected slowdown in labor force growth as the "baby boom" generation reaches retirement.

By necessity the simulation model adopts a number of reasonable but somewhat arbitrary assumptions in order to produce usable results. Here we show the GNP response to the saving rate recovery under alternative model parameters. These sensitivity tests illustrate the robustness of the model results and give readers a chance to see how their own prior assumptions change the findings.

Most of the model parameters are not important to the basic findings of the model. As Table A1 shows, changing the labor force, technology growth, and depreciation assumptions does not materially affect the results. The most important assumptions in the model relate to the link between capital and output. The estimates for the alternative models in the table show that if technology is partly driven by investment, the output effects of higher saving can be considerably higher. Furthermore, the output effects are quite sensitive to the coefficient on capital. Our model assumes a coefficient of .3; estimates in the literature range from .2 to .33 and higher. Even with the lowest reasonable parameter value, however, the saving rebound has considerable output effects.

The international dimension

A second set of crucial parameters in the model relate to the role of foreign capital in the economy. In an economy open to foreign investment, such as the United States, a drop in saving need not result in an equal loss of investment and potential GNP. Instead, the GNP loss will be mitigated to the extent that foreigners fill the saving gap and that some of the output generated by their investment accrues to U.S. residents rather than to the foreign owners. Here we explore the sensitivity of our results to two key assumptions: the capital flow response to changes in saving and the return to foreign capital.

The capital flow assumption in the model (equation 3) is a compromise between two extremes. One extreme is to assume perfect capital mobility, with investors indifferent to the country and currency in which they invest. In this environment, a drop in the U.S. saving rate will only temporarily raise U.S. interest rates, causing a foreign capital inflow that fully offsets the decline in U.S. saving.[†††] The other extreme is to assume a closed economy response, in which changes in U.S. interest rates have no effect on net foreign capital flows.[‡‡‡] In this case, a drop in saving would raise U.S. interest rates and cause a one-for-one drop in U.S. investment. The assumption adopted in the article seems consistent with actual experience in the 1980s: lower saving was partially offset by foreign capital inflows but not enough to prevent a rise in U.S. interest rates relative to other countries and a drop in U.S. investment.

The assumed return to foreign capital is also a compromise between extremes. Foreign investment probably produces just as much output as domestic investment, but it produces less gross *national* product. Part of the income generated accrues to the owners, but part also

[†††]This result strictly holds only in the "small country" case. In practice, U.S. investment would tend to fall in proportion to the drop in the global pool of saving. For a discussion of these issues, see Martin Feldstein and Charles Horioka, "Domestic Saving and International Capital Flows," *Economic Journal*, vol. 90 (June 1980), pp. 314-29.

[‡‡‡]Here we assume that capital flows continue at historical rates and do not respond to any change in relative saving and interest rates.

Table A1

Impact of the Saving Recovery on GNP under Alternative Assumptions

(Percent Deviation from the Status Quo)

	After Ten Years	After Twenty Years	After Fifty Years
Traditional model	6.0	10.7	15.6
With 1 percent labor growth	6.0	10.7	14.9
With 2 percent technical advance	6.1	10.8	15.1
With 1 percent more depreciation	6.3	11.2	15.4
With 0.6 capital coefficient	12.4	25.3	44.5
Vintage model	7.5	12.1	15.9
Learning-by-doing model	7.6	14.9	29.8

Table A2

Long-Run Impact of the Saving Recovery on GNP under Alternative International Parameters

(Percent Deviation from the Status Quo)

Domestic Share* (In Percent)	Closed Economy	Compromise	Open Economy
37	17.9	17.0	14.0
57	17.9	15.6	9.8
76	17.9	14.2	5.8

*The portion of output that accrues to U.S. residents when foreigners invest in the United States.

Appendix: The Growth Model *(continued)*

accrues to the government in the form of higher taxes and to workers benefiting from the higher demand for labor. The model assumes a 7½ percent net return to foreign investors, implying that about half of the output gain from foreign investment goes to the investor and the remainder is diverted to U.S. residents.

Table A2 shows the results of varying both assumptions. Like the last column of Table A1, Table A2 shows the long-run (fifty-year) output gain under a saving recovery. The middle entry in the table shows the results for the traditional model with the standard assumptions. The greatest gain from saving is in the closed economy case, where changes in the saving rate have a one-for-one impact on domestic investment. The saving effect is weakest in the case of a pure open economy in which a very large domestic share of the output is generated by foreigners' investments (bottom right-hand corner of the table). In this case, higher U.S. saving simply displaces foreign investors, and the U.S. only gains to the extent that foreigners no longer earn their (small) profit share.

Comparing the Cost of Capital in the United States and Japan: A Survey of Methods

by James M. Poterba

Wide U.S. trade deficits in the early 1980s prompted policy analysts in government and industry to search for the sources of declining U.S. competitiveness. Many argued that U.S. managers failed to "take the long view," forgoing investment or market development projects with high future yields to maintain their current profits. Cultural factors, such as the weak implicit contracts between firms and workers, were often cited as the cause of falling competitiveness, even though these factors have evolved slowly while the U.S. trade position declined precipitously in the early 1980s.

During the last decade, a small but growing group of academics, policy makers, and businessmen has argued that the differential behavior of U.S. and foreign firms is a rational response to disparities in their economic environments. For example, George Hatsopoulos (1983) claimed that the cost of capital, or the pretax rate of return that firms must earn to generate the returns demanded by shareholders and creditors, was significantly higher in the United States than in Japan. He and others have argued that as a result, Japanese managers find it in their firms' best interest to undertake some long-horizon projects that U.S. managers would reject.

While the cost of capital is simple in concept, it is quite complex in practice. It depends on the rates of return demanded by shareholders and bondholders, the tax system confronting corporations, and a variety of auxiliary aspects of firm behavior. Any attempt to estimate the cost of capital must rely on a variety of assumptions about corporate financing and investment practices. Moreover, data for firms in different nations are rarely comparable, requiring further assumptions and approximations.

Given the central importance of the cost of capital in corporate investment decisions, it is no surprise that numerous studies have tried to compare the cost of capital facing U.S. firms with that of their international competitors. Given the estimation difficulties, however, it is also no surprise that these studies do not reach identical conclusions. Many but not all studies find that the cost of capital has placed U.S. firms at a competitive disadvantage relative to firms in other nations.

This article surveys the sizable literature comparing the cost of capital in different nations. It tries to isolate common conclusions and to highlight the methodological differences of previous investigations. The article does not attempt to compute "definitive" estimates of relative capital costs. Rather, it draws on earlier studies and emphasizes the underlying economic and institutional factors that may contribute to cost of capital disparities.

The article illustrates alternative cost of capital methodologies by focusing on the United States and Japan. Most previous studies have confined their analysis to these nations because of a worsening bilateral trade balance in the 1980s and the high visibility of Japanese import penetration in several high-technology U.S. markets.[1] Limiting the present analysis to the United States

James M. Poterba is Professor of Economics, Massachusetts Institute of Technology, and Associate Director of the Taxation Research Program, National Bureau of Economic Research. The paper was presented on December 6, 1990, at the Federal Reserve Bank of New York Colloquium on the Cost of Capital in the United States.

[1]The principal exceptions are the studies by the U.S. Department of Commerce (1983) and McCauley and Zimmer (1989). The former considers only the cost of funds and does not report complete cost of capital estimates. The latter analyzes capital costs in the United States, Japan, Germany, and the United Kingdom.

and Japan makes it unnecessary to discuss the institutional complexities of other nations, while still highlighting measurement issues concerning the cost of capital.

Even when only two nations are compared, relative capital costs can vary through time. This article concludes that Japanese firms have enjoyed a cost of capital advantage over their U.S. competitors throughout most of the last two decades, although the source of this advantage has shifted. At the beginning of the 1980s, for example, low costs of debt combined with debt-equity ratios substantially above those in the United States held down capital costs for Japanese firms. The increasing integration of world capital markets during the last decade has limited the differences in borrowing costs, however, and today any cost of capital advantage is due to lower costs of equity rather than to differential borrowing costs.

The article is divided into six sections. The first provides a brief overview of what the cost of capital is and how it affects managers' project evaluation. The analysis demonstrates that long-term projects are particularly affected by higher costs of capital.

The next two sections discuss the cost of funds, the required return that investors demand the firm earn after corporate taxes. The second section analyzes the cost of debt, while the third considers the more complicated problem of measuring the required return demanded by shareholders. Both sections present data on historical rates of return in the United States and Japan and briefly explain why required returns might differ across countries.

The fourth section discusses debt-equity ratios of firms in the two nations, noting shifting patterns through time and describing the institutions that have historically supported higher leverage in Japan than in the United States. The fifth section considers the influence of the corporate tax rate and the system of investment incentives on the cost of capital. Contrary to some prior claims, tax considerations do not appear to be central determinants of capital cost disparities between the United States and Japan. This section also reports the summary measures of capital costs presented in previous investigations.

The article's final section notes several policy options that would affect the cost of capital. These include changing the taxation of firms and shareholders as well as raising the national saving rate.

Table 1

Impact of Discount Rates on Long-Term Investments

Discount Rate	Current Value of $5 Cash Flow, Five Years in the Future	Economically Profitable "Waiting Time" for Hypothetical New Project†
4	$4.09	45.8 years
6	3.70	23.8
8	3.35	14.2
10	3.03	9.2
12	2.74	6.1

Source: Author's calculations.

†The second column reports the waiting time for a project with an up-front cost of $100,000 and annual profits of $25,000 once it begins yielding returns. The estimates in this column answer the question, How long could a firm wait until the profit flows began?

What is the cost of capital?

The cost of capital is the pretax real return that a firm must earn, gross of depreciation, to satisfy the demands of its shareholders and bondholders. If new projects do not earn a return at least as great as the cost of capital, the equity market will penalize managers for wasting corporate resources. The cost of capital therefore directly affects the optimal investment policy of corporations. As the cost of capital rises, firms will find fewer projects yielding returns high enough to warrant new investment. The cost of capital depends upon the required returns investors demand, on the tax treatment of investment, on the depreciation of the investment asset, and on the expected rate of appreciation for the productive asset.

To understand the link between the cost of capital, discount rates, and project choice, consider a simple example of a manager confronting a project requiring a onetime payment today that will return five dollars in today's prices five years in the future. How large an up-front payment will a manager be willing to make for this project? This depends on the discount rate that investors (and the manager) apply to the firm's cash flows. The first column of Table 1 presents the answer to this question for several different values of the discount rate.[2] When the discount rate is 4 percent per year, the manager is willing to give up $4.09 for each five dollars he will earn in five years. With a discount rate of 10 percent, however, the manager will only forgo $3.03 to earn $5.00 in five years.

A second example illustrates the same point. Consider a stylized project that costs $100,000 today but does not yield returns for several years. There is no uncertainty about the project's cash flows; once the project becomes productive, it yields $25,000 per year (in the prices of the first year) *forever*. Chart 1 sketches the cash flow pattern associated with this stylized project. The second column in Table 1 reports the number of years that a manager will agree to wait before receiving the project's positive cash flows. If the discount rate is 4 percent, the project will be profitable even if it takes

[2]Variations in the discount rate affect the cost of capital, although not all disparities across countries or firms in costs of capital are due to differential discount rates.

forty-five years before positive cash flows materialize. At a discount rate of 12 percent, however, any delay of more than six years renders the project unprofitable. These calculations illustrate that the discount rate is a particularly critical determinant of the attractiveness of long-term investments.

The cost of capital depends on the discount rate as well as many other considerations affecting the attractiveness of investment projects. It is a function of the returns demanded by bondholders and shareholders, the debt-equity mix used in financing new projects, the corporate tax rate, and the generosity of tax allowances on new investments. Formally, the expression for the cost of capital (c) is

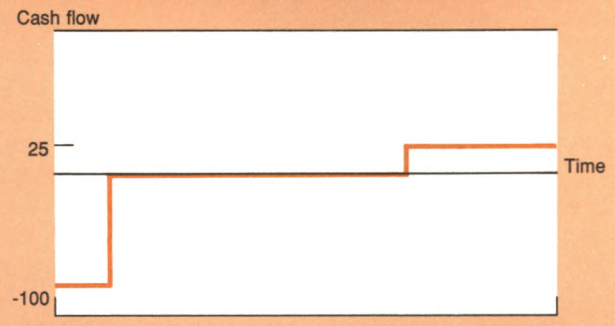
$$(1)\quad c = [r_{eq}(1-\beta) + \beta(1-\tau)r_b + \delta - \pi]^*[(1 - ITC - \tau z)/(1-\tau)],$$

where

r_{eq} = nominal rate of return demanded by equity holders
r_b = nominal rate of return demanded by bond holders
β = debt-to-total capital ratio
τ = marginal tax rate on corporate earnings
δ = economic depreciation rate of capital good
π = expected inflation rate
ITC = rate of investment tax credit
z = present discounted value of depreciation allowances on a new investment project.

This expression, though complex, is easy to understand. The first term in brackets is a weighted average of the required returns demanded by equityholders and bondholders, with weights β and (1-β) equal to the share of each type of financing in the firm's capital structure,[3] plus the cost of physical decay on the asset. The nominal cost of debt is multiplied by a $(1-\tau)$ term to reflect the tax deductibility of interest payments. Since expected inflation is subtracted from this term, it effectively depends on real debt and equity returns. The second term recognizes that investment incentives and depreciation allowances reduce the cost of purchasing capital goods. Thus, the cost of capital is lower as the investment credit (ITC) or benefits of depreciation allowances (z) are larger. The division by $(1-\tau)$ simply recognizes that profits are taxed, so that the post-tax return that the firm must deliver to its investors is "grossed up" by $1/(1-\tau)$. The next three sections focus on the components of this cost of capital formula.

Chart 1
Hypothetical Long-Term Investment Project



The cost of debt

The cost of funds is the rate of return that the firm must promise to its creditors or shareholders when it raises financial capital. Most firms use both debt and equity capital. This section considers the cost of debt, deferring the more controversial cost of equity until the next section.

The pretax cost of debt is the interest rate that a firm must promise on new corporate borrowings. There is no single borrowing rate for the corporate sector; different firms can borrow at significantly different rates, depending on their riskiness. Even a given firm does not borrow at a single interest rate; rather, it faces a spectrum of rates depending on the maturity of its debt issue and the proposed application of funds. Most studies ignore these sources of heterogeneity and use indexes of yields on high-grade corporate debt (BAA or better) to measure the cost of debt finance. This procedure is justifiable if the structure of risk and maturity premia is stable across countries and time. Such an assumption is not particularly plausible, but an alternative, empirically tractable methodology is difficult to find.

Nominal before-tax interest rates are *not* the key determinants of corporate borrowing costs. Rather, the cost of funds is affected by the *real, after-tax* cost of debt, defined by

$$(2)\quad r_{AT} = (1-\tau)i - \pi,$$

where π indicates the expected inflation rate. Variation in expected inflation rates across countries can lead to significant differences in nominal interest rates, even if real interest rates are similar. It is therefore important to

[3]Kester and Leuhrman (1990) emphasize that the marginal debt-equity mix in financing a given project may differ from the average debt-equity mix for the corporate structure. They correctly observe that the average debt-equity ratio of a firm or corporate sector may not reflect the appropriate debt-equity weighting on marginal projects.

correct nominal interest rates, even if crudely, for inflationary expectations. Equation 2 also emphasizes the link between the statutory tax rate and the after-tax borrowing rate. Since nominal interest payments are tax deductible, an increase in expected inflation that raises nominal interest rates by less than $1/(1-\tau)$ times as much as the inflation shock will reduce the after-tax cost of borrowing.

There are at least three different ways to measure expected inflation. The first assumes that actual inflation at any moment is a good proxy for what was expected. While obviously erroneous in some situations, this approach is simple and can also be interpreted as the ex post real interest rate paid by firms in a given period. A second strategy involves using either survey data or macroeconomic forecasts. While these data are somewhat arbitrary, especially if only one firm's forecast is being used, they are attractive precisely because they are statements of expectations. Finally, the most common approach is to compute a weighted average of past inflation rates and to argue that most individuals extrapolate the recent past to the future. Like the use of actual inflation rates, this approach will misstate expectations during periods when policy shocks or other factors lead to rapid revisions in inflationary prospects.

Three cost of capital studies indicate the varied approaches to measuring the real cost of debt. Hatsopoulos and Brooks (1987), who update and slightly modify Hatsopoulos' (1983) study, use Moody's BAA rate as the pretax interest rate for the United States, but for Japan, they construct their own estimate of long-term borrowing costs using the yield on heavily traded, low-risk Nippon Telephone and Telegraph bonds plus a "risk premium" equal to the yield spread between BAA bonds and Treasury bonds in the United States. This procedure assumes that the risk premium for corporate bonds is identical in the two nations. When paired with the assumption that actual inflation rates are reasonable proxies for expected inflation, it yields real after-tax interest rates in Japan that average more than 100 basis points below those in the United States during the 1970s and early 1980s.

Bernheim and Shoven (1987, 1989) focus on short-term borrowing costs, since their analysis argues that the capital market equates short-term risk-adjusted returns in the bond and stock markets. They explore several different measures of expected inflation and find that for the early 1980s, Japanese real interest rates were between 300 and 600 hundred basis points lower than their U.S. analogues. They also present evidence on long-term rates, finding disparities that, though smaller, again suggest lower Japanese real borrowing costs.

McCauley and Zimmer (1989) present the most systematic analysis of borrowing costs. They recognize the mix of long- and short-term borrowing in corporate capital structures and take an average of the interest rates on different maturity debt. They also correct observed interest rates for the presence of compensating balances, that is, requirements that borrowers hold some fraction of a loan in a low-interest account at the lending institution. These requirements effectively raise the cost of borrowing. McCauley and Zimmer (1989) follow Hatsopoulos and Brooks in subtracting the actual inflation rate from nominal interest rates when constructing the real after-tax cost of borrowing. Their results, for a more recent time period than either of the earlier studies, suggest no apparent differences in real after-tax borrowing costs in the United States and Japan. In part the difference in results is due to capital market integration beginning in the early 1980s.

While different costs of borrowing may have played an important part in historical differences between U.S. and Japanese capital costs, they are unlikely to be central today. Differences in real interest rates across nations are inconsistent with a perfectly functioning world capital market in which investors from a given nation earn the same rate of return regardless of where they invest their funds. Academic studies (surveyed, for example, by Mishkin 1984 and Frankel 1990) nevertheless suggest that there are differences in real interest rates between some countries. The size of the U.S. and Japanese markets and the active cross-border arbitrage in fixed income markets make large disparities in these markets unlikely.

A firm in either the United States or Japan could, in addition, try to exploit persistent differences in real interest rates by issuing bonds denominated in the other nation's currency and marketing them to foreign investors. This equilibrating force was absent in the years before 1980, when the Japanese capital market was relatively closed to outside investors or borrowers. Today, however, firms routinely make cross-border transactions of this type. This development reinforces the view that interest rate differences are unlikely to be a central component of the cost of capital differences between Japan and the United States.

The cost of equity

Estimating the cost of equity is the most difficult part of any cost of capital computation. The reason is that there is little evidence on the risk premium that equity investors require to hold stocks rather than less risky bonds. The risk premium is likely to vary through time, making it difficult to use historical data to assess this parameter. Consequently, researchers have differed more in their methods of measuring the cost of equity than in

their methods of measuring the cost of debt.

This section considers four approaches to measuring the cost of equity. The first subsection considers estimates that assume that past returns on corporate stock provide a good guide to required returns. The next three subsections discuss various measures of expected returns that are based on the ratio of actual earnings to share prices or assets values. A concluding subsection discusses the extent to which differences in equity cost can persist in a world capital market.

Estimates using historical data on equity returns

The simplest approach to measuring the required return on equity is to assume that the historical average differential in equity and debt returns indicates the extra return that investors demand for holding risky equity rather than riskless debt.[4] If required returns were constant through time, and if the data sample on equity and debt returns were long enough to measure the average returns precisely, then this procedure would yield reliable results. In practice, however, neither of these conditions is satisfied.

It is useful to begin with background information on the equity risk premium computed this way. Table 2 reports the average excess return on equities relative to government bills in the United States and Japan for several different time periods. The findings highlight the sensitivity of these results to the sample period.[5] The sharp rise in the Japanese equity market during the mid-1980s implies that any estimate of ex post returns that includes these years (and does not span a very long period) will show that Japanese investors demand *higher* equity returns than their U.S. counterparts. The 39 percent decline in the Japanese equity market during calendar 1990 has weakened, but not erased, the apparent differential in required returns.

The problems with using relatively short samples of historical returns are more fundamental than sensitivity to a few years of data. To understand the first problem, consider an economy in which institutional changes within a single year reduce by half an equity risk premium that has historically been constant. Share prices will *rise* in response to this news, and ex post measures of the equity risk premium will suggest that it has risen. In this case, however, the actual movement is just the opposite.

The second difficulty with ex post returns is that just as real interest rates appear to fluctuate, there is evidence that required returns vary over the business cycle and through time. Recent research in financial economics (for example, Fama and French 1988) suggests that a considerable share of the variation in equity returns, particularly over long horizons, can be forecast using the dividend-price ratio and related variables. Changes in financial markets and practices are also likely to affect the equity risk premium. The rise in the leverage of some U.S. firms during the 1980s, for example, probably raised their equity risk premia relative to what they would have been otherwise; the gradual reduction in the fear of deep and major depressions since the 1930s has probably lowered the relative cost of equity during the postwar period.

A third drawback to using historical data to calibrate required returns is the imprecision of the resulting estimates. During the last sixty years, the return on U.S. equities has exceeded that on Treasury bills by 7.5 percentage points per year. Given the significant annual variation in equity returns—the standard deviation of returns on the U.S. market is approximately 20 percent per year—the standard deviation of the *mean* return

[4]Bernheim and Shoven (1989) present some estimates based on this approach. Baldwin (1986) and Kester and Leuhrman (1990) also implicitly take this approach.

[5]Baldwin (1986) was among the first to bring equity returns data to bear on calibrating the required return; she concluded that, if anything, the risk premium was higher in Japan than in the United States. Kester and Luehrman (1990) perform a more sophisticated set of tests, asking whether the market pricing of particular categories of risk differs in the United States and Japan. They find no evidence of such differences, but their tests are restricted to only four years of data (1982-86).

Table 2

Excess Returns on Equities Relative to Bills: United States and Japan, 1926–1990

	U.S. Excess Return (In Percent)		Japanese Excess Return (In Percent)	
Sample Period	Mean	Standard Deviation	Mean	Standard Deviation
1926–89	7.5	20.0	—	—
1960–89	3.2	15.1	7.2	16.6
1960–79	1.5	14.3	3.8	16.9
1980–89	6.5	16.5	14.0	15.9

Source: Author's calculations, based on Ibbotson Associates (1990), and Morgan Stanley–Capital International Data.

estimated for the period since 1926 is approximately 2.5 percentage points. To specify a range with a 95 percent chance of including the actual mean differential, one would therefore need to admit possibilities from 2.5 to 12.5 percentage points. With such a range, convincing conclusions about the cost of equity are very difficult.

For Japan, the data problem is even more severe. Most analysts focus on returns in the Japanese equity market during the period since 1960 because the markets before the Second World War and in the early postwar years bore little resemblance to the sophisticated market of today. With only thirty years of data, however, the 95 percent confidence band for returns on the Japanese equity market ranges from 1.2 percent to 13.2 percent per year.

Estimates based on price-earning ratios

A second (and probably the most common) approach to measuring required equity returns relies on market-based measures of prospective equity returns. McCauley and Zimmer (1989a), Bernheim and Shoven (1989), and Ando and Auerbach (1988a, 1988b, 1990) all use some variant of this approach in studying cost of capital disparities. They use the earnings-price ratio, possibly corrected for international differences in accounting or other features, to measure investors' required returns.

Before considering the merits and difficulties of this approach, it is useful to summarize the trends through time in price-earnings ratios for the United States and Japan. These data are shown in Table 3 and Chart 2, without any adjustments. The rapid rise in Japanese share prices during the mid-1980s made the price-earnings ratio in Japan much higher than that in the United States. This is the basis for many findings that Japanese firms faced lower required returns on equity during this period.

There are both theoretical and empirical difficulties in using price-earnings ratios or, more accurately, their reciprocal (earnings-price ratios) to describe required returns. One theoretical objection is that rather strong assumptions are needed if the earnings-price ratio is to equal the current required return. For example, if required equity returns change through time, then the earnings-price ratio equals an average of current and future required returns, minus the expected growth rate of earnings. Today's required return is equal to the earnings-price ratio only if the required return is constant through time, or if by chance future variations offset each other and lead the average to equal the current value. A second difficulty is that observed earnings-price ratios reflect the stock market's expectation of future corporate growth. A low earnings-price ratio could therefore be the result of optimistic growth expectations rather than low costs of equity finance. In any

Table 3

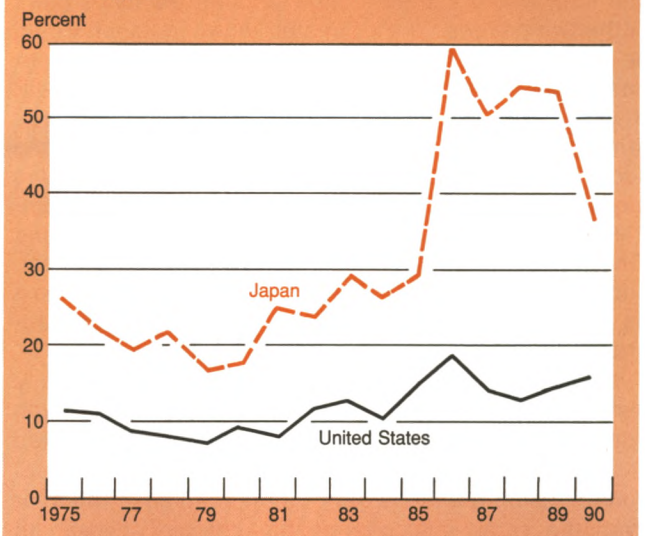
Price-Earnings Ratios for the United States and Japan, 1975–1990

Year	United States	Japan
1975	11.8	25.2
1976	11.2	22.0
1977	9.1	19.3
1978	8.2	21.5
1979	7.5	16.6
1980	9.6	17.9
1981	8.2	24.9
1982	11.9	23.7
1983	12.6	29.4
1984	10.4	26.3
1985	15.4	29.4
1986	18.7	58.6
1987	14.1	50.4
1988	12.9	54.3
1989	14.8	53.7
1990	15.9	36.6

Source: French and Poterba (1991a, Table 6). U.S. price-earnings ratios are taken from Standard & Poor's 500 index of actively traded stocks; Japanese ratios are from the Nomura Research Institute's 350 index of actively traded stocks.

Chart 2

Price-Earnings Ratios for the United States and Japan



Source: Kenneth R. French and James M. Poterba, "Were Japanese Stock Prices Too High?" Journal of Financial Economics, forthcoming.

case, the resulting earnings-price ratio must be corrected for expected growth differentials to compare required returns across countries.

A more practical objection to measuring equity returns with earnings-price ratios is that these ratios cannot be compared internationally because of accounting factors. Most studies relying on earnings-price information make some corrections to numbers reported by corporations; the United States–Japan comparison illustrates the type of corrections needed.

Consolidation of subsidiary earnings. Until the mid-1980s, Japanese firms usually reported parent company earnings, excluding the profits of wholly and partly owned subsidiaries. Since more than half of the shares on the Tokyo Stock Exchange are owned by other traded corporations (see French and Poterba 1991a), omission of the retained earnings from partly owned firms can substantially affect reported earnings. This generates a downward bias in the earnings-price ratio as a measure of required returns, since the stock market will recognize the value of intercorporate equity holdings but earnings will not reflect the relevant cash flow. This problem can be corrected by inflating earnings (the approach in McCauley and Zimmer 1989a) or by removing the value of intercorporate holdings from the estimate of share value (French and Poterba 1991a).

Depreciation. In Japan, firms use the same depreciation lifetimes in computing tax and accounting earnings. In the United States, accounting depreciation is typically slower than that for tax purposes. The same project, if accounted for by a Japanese and an American firm, would therefore show different earnings flows in the two nations. The estimated return in Japan would be lower in the early years of the project, when Japanese depreciation would exceed that in the United States, and higher in later years, when the Japanese firm would have fully depreciated the asset. These accounting disparities need to be corrected in making any comparison of earnings-price ratios across countries. Ando and Auerbach (1990) and McCauley and Zimmer (1989a) convert depreciation for both Japan and the United States to an economic replacement-cost basis; French and Poterba (1991a) try to restate Japanese depreciation on U.S. accounting principles.

Inflationary effects on earnings. Inflation has many distorting effects on corporate earnings. It interacts with nominal accounting conventions to make reported accounting earnings a relatively poor proxy for economic profits. If nations have different inflation rates, or even the same inflation rate but different investment histories, then reported accounting earnings will be differentially biased.

Ando and Auerbach (1988a, 1988b, 1990) and McCauley and Zimmer (1989a) try to correct accounting earnings for inflationary errors. This involves restating depreciation allowances in terms of asset replacement cost rather than historical cost, subtracting spurious profits on goods in inventory sold at nominal prices that exceed the nominal acquisition price by much more than the real sales price exceeds the real purchase cost, and estimating real rather than nominal interest outlays. The relative importance and net effect of these corrections on U.S. and Japanese accounting earnings vary through time. The inflation rate in Japan was higher than that in the United States during the 1970s, but lower in the mid-1980s. In the 1970s, however, the greater leverage of Japanese firms made the inflationary misstatement due to nominal interest rates more important than that in the United States.

Other factors must be considered in correcting earnings-price ratios across nations, such as the treatment of reserve accounts in Japan and the disparate procedures for funding retirement plans in different countries. The factors discussed above, however, are the most important ones.

After correcting earnings-price ratios for the various considerations noted above, the ratios for Japan still appear lower than those for the United States. Lower earnings-price ratios in one country do not *necessarily* signal lower required returns, since the disparity could be due to differential growth expectations. One crude way to assess the importance of the latter effect relies on estimates of long-term real GNP growth published by macroeconomic forecasting firms. These show average Japanese long-term growth rates of approximately 4 percent per year, compared with values of approximately 2.5 percent per year for the United States. Even if discount rates were identical, one would therefore expect lower earnings-price ratios in Japan than in the United States. This is not a large enough growth disparity, however, to account for the differences in earnings-price ratios, nor are there any striking changes in the expected growth rate in the mid-1980s when the Japanese market soared.[6] This evidence consequently points toward lower required equity returns in Japan than in the United States, particularly in the late 1980s.

Estimates of required returns based on market earnings-price ratios can change substantially in relatively short time spans. This has occurred during the last year with the sharp decline in the value of the Japanese stock market. The earnings-price ratio in Japan has risen by more than one-third since December 1989, indicating a possible rise in required returns.

[6]This discussion draws on French and Poterba (1991a), which also presents data on macroeconomic forecasts.

Estimates based on historical profit rates

A third approach to measuring required returns, one which is related to the earnings-price calculations, involves measuring the rate of return on corporate assets—the profit rate. Rather than scale accounting earnings by a market-based measure of asset values such as the total value of outstanding equity, this approach divides by an estimate of the replacement value of the firm's capital stock. It suffers from all the difficulties of international comparisons that are associated with earnings-price ratios, with the additional difficulty that data on the replacement value of assets are not readily available and, when available, are often estimated in different ways for different nations. Nevertheless, computing the ex post profit rate can provide some evidence on the level of required returns.

Sustainable growth analysis

A fourth approach to estimating the cost of equity, used by Hatsopoulos (1983) and Hatsopoulos and Brooks (1987), involves estimating the sustainable growth rate for dividends that could be achieved by reinvesting current earnings without altering debt policy. By adding the sustainable growth rate to the current dividend yield, this approach provides another estimate of the required return on equities. Since this method is ultimately based on historical rates of return, not surprisingly it suggests that the cost of equity in Japan is lower than that in the United States.

Can the costs of equity differ?

This survey of previous work suggests that several different methodologies point to a similar conclusion: the cost of equity has been lower in Japan than in the United States for most of the last two decades.[7] Just as it was appropriate to ask if international differences in real interest rates could persist over long periods, one can ask whether arbitrage by investors and firms can eliminate disparities in expected equity returns. There are at least three reasons to suspect that differential equity returns *can* persist.

First, structural factors may lead to fundamental differences in the riskiness of U.S. and Japanese firms. Intercorporate share ownership in Japan and the significant role of banks in corporate finance affect firm behavior and may cushion investors from particularly adverse outcomes at a given firm.[8] This would suggest that even if the price of a particular type of security market risk were equated in Japan and the United States, the "real riskiness" of the Japanese corporate sector would be lower and therefore would command a lower total risk premium.

Second, the rapid increase in Japanese land prices during the 1980s may (until recently) have provided a ready source of collateral for Japanese corporate borrowing.[9] The value of land holdings for Japan's nonfinancial corporate enterprises rose from ¥256.3 trillion at the end of 1985 to ¥478.2 trillion at the end of 1988—an appreciation of between $1.5 trillion and $2 trillion, depending on which exchange rate is used.[10] This sharp rise in collateral value may have lowered equity costs in recent years; it would not provide an explanation for any lower equity costs in the 1970s and early 1980s.

Finally, it is possible that the strong assumptions of integrated world capital markets are inappropriate. Japanese investors are less well informed about U.S. than about Japanese equities and may therefore prefer holding domestic shares, even if the expected return on U.S. equities is somewhat higher (see French and Poterba 1991b). As for the arbitrage by corporate suppliers of equity, U.S. firms may face constraints on their ability to issue equity in Japanese markets. Japanese investors may convey low-cost capital to Japanese firms but not to U.S. firms traded in Tokyo (whose shares are primarily traded and priced in New York).

One explanation of the difference in equity costs that does *not* appear to explain the U.S.-Japanese case is high turnover in the U.S. stock market. Table 4 shows the turnover rates on the New York and Tokyo stock exchanges during the years 1985-89. Turnover rates in Tokyo exceed those in New York in some years. When one recalls that the Tokyo market includes very substantial blocks of cross-held shares that trade infrequently, the implied turnover rate for the "in play" shares is significantly higher than that in New York.

Weighting the costs of debt and equity: corporate leverage rates

The results in the last two sections suggest that the cost of debt may have been lower in Japan than in the

[7]The earlier discussion suggested that because of the limited data span, it is probably not possible to reject the hypothesis of similar required equity returns in the two nations. The test being applied in such cases, however, has extremely low power to detect deviations because there is so little data.

[8]Hoshi, Kashyap, and Scharfstein (1990) provide evidence of the resilience of investment and sales at "keiretsu" firms during

Footnote 8 continued
economic downturns, suggesting that financial practices affect real behavior.

[9]Kashyap, Scharfstein, and Weil (1990) provide some evidence that Japanese firms with greater land holdings have exhibited higher investment rates during recent years. This evidence is consistent with, although not definitive support for, the collateral explanation.

[10]Economic Planning Agency, *Annual Report on National Accounts, 1990* (Tokyo: Economic Planning Agency).

United States until the early 1980s. The cost of equity has been lower for most of the last two decades but particularly in the late 1980s. The net effect on the cost of capital depends on the relative weights placed on debt and equity in the two nations. These debt-equity ratios have not remained fixed over time but have changed significantly during the last two decades.

There are several measures of the debt-equity ratio of a firm or corporate sector. Although managers usually focus on the ratio of book debt to the book value of equity, this measure fails to capture the significant swings in the relative prices of debt and equity securities. The more natural measure is therefore the market value of debt divided by the market value of equity. Table 5 and Chart 3 show the recent history of an imperfect measure of debt-equity ratios for nonfinancial firms: the ratio of the book value of long- and short-term debt to the market value of corporate equity.

The central conclusion to be derived from Table 5 is that Japanese debt-equity ratios were significantly higher than their U.S. counterparts in the early 1980s, but that they have declined while U.S. leverage has remained stable or, if anything, increased.[11] In 1985, the Japanese debt-equity ratio was 1.3 to 1, compared with .67 to 1 in the United States. By March 1989, rising Japanese share prices had reduced the ratio of book debt to market equity to .55 in Japan, while the corresponding figure for the United States was still approximately .67.

The convergence of U.S. and Japanese leverage was due to two factors. First, U.S. nonfinancial corporations repurchased nearly $100 billion in equities *each year* between 1985 and 1990. Chart 4 shows the net secu-

[11]Ando and Auerbach (1988b) report estimates of the book debt-to-market equity ratio for the period since 1970, and their data suggest relatively little variation in either nation during the years before 1985.

Table 4

Turnover Rates for U.S. and Japanese Equity Markets

(In Percent)

Year	New York Stock Exchange	Tokyo Stock Exchange
1980	36	50
1981	33	50
1982	42	35
1983	51	44
1984	49	43
1985	54	48
1986	64	75
1987	73	96
1988	55	98
1989	52	73

Sources: Column 1 data are drawn from the *New York Stock Exchange Fact Book;* column 2 data are from the *Tokyo Stock Exchange Fact Book.*

Table 5

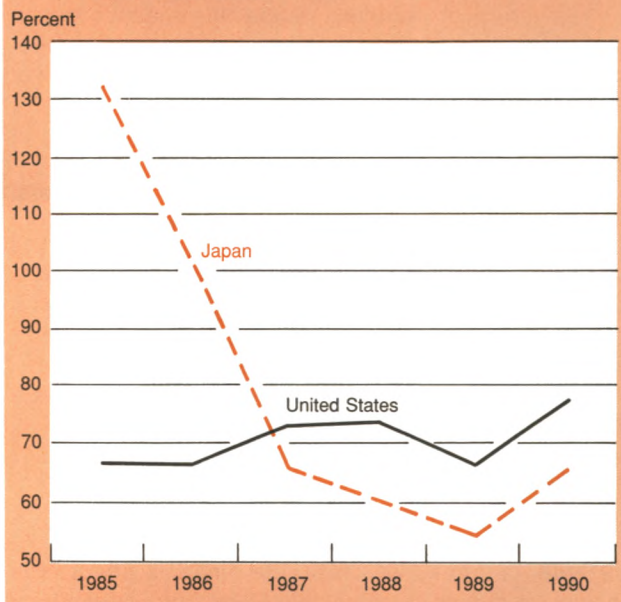
Debt-Equity Ratios (x100) for U.S. and Japanese Nonfinancial Corporations

Year	United States	Japan
1985	66.9	132.7
1986	66.6	100.7
1987	73.0	65.9
1988	73.7	60.1
1989	66.5	54.8
1990	77.1	65.8

Source: U.S. data are drawn from the Board of Governors of the Federal Reserve System, *Balance Sheets of the U.S. Economy,* 1990 edition. Japanese data are from *Daiwa Analysts Guide,* 1989 edition. Estimates show the ratio of the book value of corporate long- and short-term liabilities to the market value of corporate equity.

Chart 3

Debt/Market Equity Ratios for U.S. and Japan Corporations



Sources: Board of Governors of the Federal Reserve System, Balance Sheets of the U.S. Economy, 1990; Daiwa Analysts Guide, 1990.

Notes: Chart shows ratios of the book value of corporate long- and short-term liabilities to the market value of corporate equity. Japanese data are from end-March. U.S. data for 1985-89 are from end-year; for 1990, from end-September.

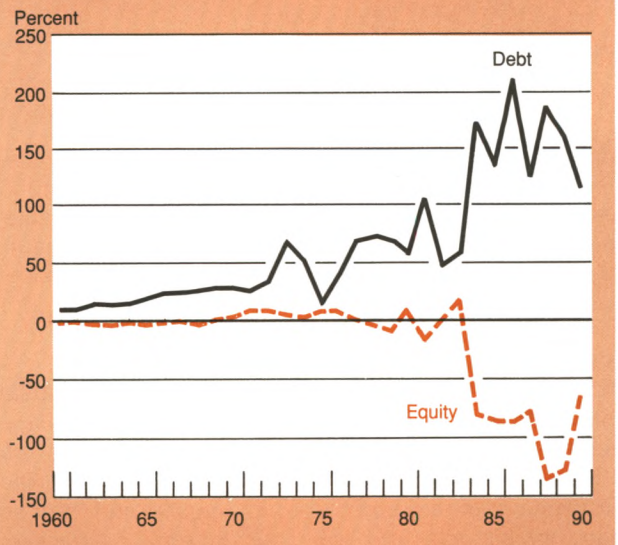
rity issues during this period, with large equity purchases, both direct repurchases and takeover acquisitions, matching substantial debt issues in recent years. Only rising equity values prevented the debt-equity ratio from rising sharply during this period. Second, the rapid increase in Japanese share values during the 1980s was not matched by escalating debt values or debt issue. Consequently, the debt-equity ratio of Japanese firms on a market value basis declined during the period.

Taxation and summary measures of the cost of capital

The least controversial part of most cost of capital studies is the treatment of tax incentives for new capital investment. There is broad consensus both on the approach to analyzing tax considerations and on the underlying tax code provisions that are important. Different studies have reached different conclusions, however, regarding the net effect of tax provisions on the relative costs of capital, primarily because of *different auxiliary assumptions*. This section sketches the relevant tax parameters—the statutory corporate tax rate and the net tax-induced reduction in the price of capital goods—then notes their values through time and explains how they affect the cost of capital. It concludes by presenting complete estimates of the cost of capital from several different studies.

Chart 4

Net Equity and Debt Issues by U.S. Nonfinancial Corporations



Source: Federal Reserve Board, Flow of Funds Accounts.

Note: 1990 data are first three quarters' flows at annual rate.

Tax parameters

The statutory tax rate affects the pretax returns that firms must earn, other things equal, to satisfy their owners. The magnitude of this effect depends on the fraction of the corporation's profits that are subject to corporate tax, that is, on the relative importance of debt and equity finance.

The generosity of tax depreciation schedules, including the availability of investment credits, is another key aspect of the tax code. To provide a unifying framework for comparing different depreciation schedules, most economic analyses focus on the present discounted value of tax depreciation benefits, given by

$$(3)\quad ITC + \tau z = ITC + \Sigma\ \tau_{t+k} * d_{t+k} / (1+\rho)^k,$$

where τ_{t+k} is the tax rate prevailing k years after the investment is made, d_{t+k} is the value of depreciation allowances (per dollar of initial investment) that the firm is allowed to claim, and ρ is the nominal discount rate applied by investors to cash flows with the risk characteristics of depreciation benefits. The value of investment allowances thus depends on the rate at which the future tax savings are discounted, as well as the statutory corporate rate. Higher tax rates make a given set of deductions more valuable.

The net effect of raising the corporate tax rate thus depends on the time path of depreciation allowances and the discount rate applied to these tax benefits. If these depreciation benefits were worth one dollar ($z = 1$ and $ITC = 0$), then changes in the corporate tax rate would have *no effect* on the cost of capital: the after-tax cost of a one-dollar project would be reduced, just as the after-tax return from the project would fall. Only when the value of depreciation allowances falls below one dollar does raising the corporate tax rate increase the cost of capital.

The tax parameters in both the United States and Japan have shifted significantly during the last decade. Table 6 presents the values of each tax component for the beginning and end of the decade. The first column shows the statutory corporate tax rate in each nation, with the U.S. rate falling from 50 percent, including federal as well as state taxes, at the beginning of the 1980s to 38 percent at the decade's end. By comparison, Japanese corporate tax rates are higher: the net tax rate was 53 percent in 1980 and remained at 50 percent at the decade's end.

The second column in Table 6 shows the depreciation benefits accruing to a firm that invests in general indus-

Table 6
Tax Parameters in Cost of Capital Calculations

Parameter	Japan		United States	
	1980	1988	1980	1988
Statutory corporate tax rate	.526	.499	.495	.380
Present value of tax reduction for new investment				
Autos	.465	.473	.534	.333
Industrial plant	.250	.355	.166	.142

Source: Bernheim and Shoven (1989, Table 5).

trial equipment, as well as industrial plant, in the two nations. Although the tax lifetimes in the two nations are similar, the tax benefits for the two examples of projects given here are greater in Japan. The reason is that in the late 1980s the discount rate applied to the cash flows is lower, and the statutory tax rate to which the deductions apply, higher, in Japan. The net effect of U.S. tax policies during the 1980s was to lower depreciation benefits by extending lifetimes, phasing out the investment tax credit, and reducing the statutory marginal tax rate. Consequently, these changes brought about an increase in the cost of capital.

Most studies of capital costs have argued that tax provisions in Japan are similar to those in the United States and that therefore relatively little of the cost of capital differential can be attributed to tax considerations. Bernheim and Shoven (1989) point out, however, that similar tax provisions operating in different economic environments can yield different tax incentives.

Summary costs of capital

Relatively few studies have made complete estimates of the cost of capital, although many have examined its components. Table 7 presents three sets of estimates from studies using different methodologies to assess U.S. and Japanese capital costs. The table shows the estimated capital cost in 1980 for each study, as well as the estimate for the most recent year available.

The studies compute somewhat different capital costs; Hatsopoulos and Brooks estimate an average cost of all capital services, McCauley and Zimmer the cost of capital for a plant investment with a twenty-year lifetime, and Bernheim and Shoven the cost of capital for an industrial plant. In addition, the strategies for estimating the cost of equity vary; Hatsopoulos and Brooks use the sustainable growth method, while Bernheim-Shoven and McCauley-Zimmer use estimates based on adjusted earnings-price ratios.

Despite these differences in approach, all of the studies conclude that the cost of capital is significantly higher in the United States than in Japan. The precise magnitudes differ, with Hatsopoulos and Brooks finding the largest differential (10 percentage points) in 1980, compared with only 2.7 percentage points in McCauley and Zimmer's study. In more recent years, the results suggest a cost of capital differential of approximately 5 percentage points between the two nations.

Table 7
Estimated Costs of Capital for the United States and Japan

Study	Year	United States (Percent)	Japan (Percent)
Hatsopoulos-Brooks			
	1980	14.1	4.0
	1985	9.7	3.8
McCauley-Zimmer			
	1980	11.5	8.8
	1988	11.2	7.2
Bernheim-Shoven			
	1980	18.7	11.0
	1988	11.1	4.1

Source: Hatsopoulos-Brooks values are estimated by the author from Figure 9 of Hatsopoulos-Brooks (1987) and correspond to the cost of fixed asset services (before depreciation). McCauley-Zimmer estimates are drawn from Table 2 of McCauley-Zimmer (1989b) and correspond to the cost of a twenty-year plant. Bernheim-Shoven estimates are drawn from Table 6 of Bernheim-Shoven (1989).

Conclusion and possible policy levers

Many different factors bear on a nation's cost of capital. This survey of previous work comparing the cost of capital in the United States and Japan suggests that differential costs of equity are the single most important explanation of apparent cost of capital differences. Many institutional and economic differences between the two nations may contribute to this disparity—in particular, Japan's higher saving rate, less burdensome taxation of equity returns, and greater flexibility in spreading corporate risk.

Because the cost of capital depends on many parameters, a wide range of policies can be used to affect it. Several possibilities are indicated below.

Changing investment incentives is probably the most direct way for policy makers to affect capital costs. An investment tax credit, for example, reduces the cost of capital and can be targeted to affect only some classes of assets. While much of the discussion leading up to the Tax Reform Act of 1986 stressed the need for a "level playing field," treating all assets equally for tax purposes, some have argued that particular asset classes should be subsidized because of their high social returns.[12] This is the basis for the research and development tax credit, as well as subsidies to low-income housing. The major disadvantage of more general investment incentives is their significant revenue cost. To remedy this problem, policy makers might consider more revenue-efficient subsidies, such as "incremental" investment tax credits on a firm's investment above some history-based target.

The *tax treatment of investors* is a second obvious source of policy leverage on the cost of capital. The analysis above treated the pretax returns demanded by debt and equity investors as fixed. These returns may vary, however, with the tax treatment accorded to different securities. The lower pretax required return on tax-exempt debt in contrast to taxable bonds rather clearly demonstrates that investor-level taxes affect required returns. In this regard, a change in the tax treatment of dividends—for example, by reducing shareholder taxes with an integration system—would reduce capital costs. Similarly, a capital gains tax reduction would lower the pretax return demanded on equities. A particularly cost-effective form of capital gains reduction, from the standpoint of reducing the cost of capital, would follow the Japanese experience and apply very low tax rates to capital gains on corporate equities while taxing gains on other assets at relatively high rates.[13]

Raising national saving is a third strategy for reducing the cost of capital. Higher national saving would expand the supply of saving relative to demand, lower required returns on both debt and equity, and ultimately reduce capital costs. While the direction of this effect is clear, the magnitude of the cost of capital reduction from a given saving increase is again controversial. With partially integrated world capital markets, part of any increase in domestic saving will flow abroad, thereby blunting the effects on domestic required returns. Although historical evidence suggests a rather strong association between domestic saving and domestic investment rates, international capital markets have become much better integrated during the last decade, and the leakage effects are therefore probably larger today than in the past.

[12]DeLong and Summers (1990) suggest that equipment investment yields particularly high social returns and therefore warrants subsidy beyond other classes of capital goods. They present international evidence showing that nations that encourage equipment investment by keeping the relative price of equipment low grow more rapidly than nations with higher equipment prices.

[13]The net effect of capital gains tax reduction on capital costs is controversial. While few doubt that lower tax rates would lower capital costs, there is less agreement on the size of the effect. Since many gains on corporate stock are realized long after they accrue, or never face tax because of basis step-up at death, the effective capital gains tax rate is significantly lower than the statutory rate.

References

Ando, Albert, and Alan J. Auerbach. 1988a. "The Cost of Capital in the United States and Japan: A Comparison." *Journal of the Japanese and International Economies* 2: 134-58.

_____.1988b. "The Corporate Cost of Capital in Japan and the United States: A Comparison." In John B. Shoven, ed., *Government Policy Toward Industry in the United States and Japan* (Cambridge: Cambridge University Press).

_____.1990. "The Cost of Capital in Japan: Recent Evidence and Further Results." *Journal of the Japanese and International Economies* 4: 323-50.

Aron, Paul. 1981. *Are Japanese P/E Multiples Too High?* Daiwa Securities America, New York.

_____.1988. *Japanese P/E Multiples: The Shaping of a Tradition.* Daiwa Securities America, New York.

Baldwin, Carliss. 1986. "The Capital Factor: Competing for Capital in the Global Environment." In Michael Porter, ed., *Competing in Global Industries* (Boston: Harvard University Press).

Bernheim, B. Douglas, and John B. Shoven. 1987. "Taxation and the Cost of Capital: An International Comparison." In Charles E. Walker and Mark A. Bloomfield, eds., *The Consumption Tax: A Better Alternative?* (Cambridge: Ballinger Publishing).

References *(continued)*

_____.1989. "Comparison of the Cost of Capital in the United States and Japan: The Role of Risk and Taxes." Stanford University, Center for Economic Policy Research, mimeo.

Boone, Peter, and Jeffrey Sachs. 1989. "Is Tokyo Worth Four Trillion Dollars? An Explanation for High Japanese Land Prices." Harvard University, unpublished paper.

Brazzell, D., A. Robbins, G. Robbins, and P.C. Roberts. 1986. *The Cost of Capital in the United States and Japan* (Washington: Institute for Political Economy).

DeLong, J. Bradford, and Lawrence H. Summers. 1990. "Equipment Investment and Economic Growth." Harvard University, mimeo.

Fama, Eugene, and Kenneth French. 1988. "Dividend Yields and Expected Stock Returns." *Journal of Financial Economics* 22: 3-27.

Frankel, Jeffrey. 1989. "Quantifying International Capital Mobility in the 1980s." National Bureau of Economic Research, Working Paper no. 2856.

French, Kenneth R., and James M. Poterba. 1991a. "Were Japanese Stock Prices too High?" *Journal of Financial Economics*. Forthcoming.

_____.1991b. "Investor Diversification and International Equity Markets." *American Economic Review*. Forthcoming.

Hatsopoulos, George N. 1983. *High Cost of Capital: Handicap of American Industry* (Washington: American Business Conference).

Hatsopoulos, George N., and Stephen H. Brooks. 1986. "The Gap in the Cost of Capital: Causes, Effects, and Remedies." In R. Landau and D. Jorgenson, eds., *Technology and Economic Policy* (Cambridge: Ballinger Publishing).

_____.1987. "The Cost of Capital in the United States and Japan." Paper presented at International Conference on the Cost of Capital, Harvard University, Kennedy School of Government, November 19-21, 1987.

Hoshi, Takeo, Anil Kashyap, and David Scharfstein. 1990a. "Corporate Structure, Liquidity, and Investment: Evidence from Japanese Panel Data." *Quarterly Journal of Economics*. Forthcoming.

_____.1990b. "The Role of Banks in Reducing the Costs of Financial Distress in Japan." National Bureau of Economic Research, Working Paper no. 3435.

Ibbotson Associates. 1990. *Stocks, Bills, Bonds, and Inflation: 1990 Yearbook*. (Chicago).

Kashyap, Anil, David Scharfstein, and David Weil. 1990. "Land Holdings and Investment by Japanese Firms." Massachusetts Institute of Technology, mimeo.

Kester, W. Carl, and Timothy A. Luehrman. 1989. "Real Interest Rates and the Cost of Capital: A Comparison of the United States and Japan." *Japan and the World Economy* 1: 279-301.

_____.1990. "The Price of Risk in the United States and Japan." Harvard Business School, mimeo.

King, Mervyn A., and Don Fullerton, eds. 1984. *The Taxation of Income from Capital* (Chicago: University of Chicago Press).

McCauley, Robert N., and Stephen Zimmer. 1989a. "Explaining International Differences in the Cost of Capital." Federal Reserve Bank of New York *Quarterly Review* 14.2: 7-28.

_____.1989b. "Explaining International Differences in the Cost of Capital: The United States and United Kingdom versus Japan and Germany." Federal Reserve Bank of New York, Research Paper no. 8913.

McKee, M.J., J.J.C. Visser, and P.G. Saunders. 1986. "Marginal Tax Rates on the Use of Labor and Capital in the OECD Countries." *OECD Economic Studies* 7: 45-101.

Mishkin, Frederic S. 1984. "Are Real Interest Rates Equal Across Countries?" *Journal of Finance* 39: 1345-57.

Shoven, John B., and T. Tachibanaki. 1988. "The Taxation of Income from Capital in Japan." In J.B. Shoven, ed., *Government Policy Toward Industry in the United States and Japan* (Cambridge: Cambridge University Press).

U.S. Department of Commerce, International Trade Administration. 1983. *A Historical Comparison of the Cost of Financial Capital in France, the Federal Republic of Germany, Japan, and the United States* (Washington, D.C.: Government Printing Office).

Bank Cost of Capital and International Competition

by Steven A. Zimmer and Robert N. McCauley

The rising share of U.S. corporate loans booked by foreign-owned banks and the withdrawal of U.S. banks from foreign lending (Chart 1) raise concerns about the competitiveness of U.S. banks. The increasing fraction of U.S. banking assets controlled by foreign banks is paralleled by the rising share of U.S. manufacturing assets or employment under the control of foreign firms.[1] Unlike U.S. banks, however, U.S. manufacturing and commercial firms are not retrenching their foreign operations.[2] U.S. banks, then, are lagging U.S. industrial firms in international competition as measured by asset growth.

This contrasting performance may reflect some unique features of U.S. banking law and the very different profitability of banks' and corporations' foreign operations. A fundamental economic force may also be at work. If a relatively high cost of capital burdens both U.S. banks and industrial firms but figures more critically in financial than in industrial competition, then it is understandable that U.S. banks might lag their industrial counterparts.

This article investigates how capital costs may have contributed to the declining competitiveness of U.S. banks. We compare capital costs facing commercial banks in different industrial countries to determine whether U.S. banks are in fact operating at a disadvantage. In addition, we seek to identify the factors that account for differences in the cost of capital and explore some of the implications of these disparities.

Our analysis reveals that Japanese banks enjoy a low cost of capital, German and Swiss banks face a moderate cost of capital, and U.S., U.K., and Canadian banks confront a high cost of capital. These differences can be traced to shareholders' valuations of bank earnings in different equity markets. In effect, shareholders allow banks from different countries to price their services at different levels. What appears, then, to a banker with demanding shareholders as razor-thin margins designed to win market share may appear to another banker with less exacting shareholders as a fully priced transaction. We illustrate this point by calculating the capital costs for three different financial products: a straight corporate loan, a commitment to lend, and an interest rate swap.

Differences in bank cost of capital may arise from differences in national saving behavior, macroeconomic stabilization policies, industrial organization, financial policies, and taxes. Taxes can exert a more important effect on bank cost of capital than on industrial cost of capital, but they do not account for the differences observed. Stronger official safety nets for foreign banks may serve to cheapen subordinated debt and equity costs.

The cost of capital differences assert themselves forcefully in wholesale lending. In addition to the broad shift of market share in lending to U.S. corporations from U.S. banks to foreign banks shown in Chart 1,

[1]Only in the rare industry such as chemicals, however, have foreign firms reached the one-third share achieved by foreign banks in U.S. corporate lending. See Ned G. Hownestein, "U.S. Affiliates of Foreign Companies: Operations in 1988," *Survey of Current Business*, vol. 70 (July 1990), pp. 127-43.

[2]Raymond J. Mataloni, "U.S. Multinational Companies: Operations in 1988," *Survey of Current Business*, vol. 70 (June 1990), pp. 31-53.

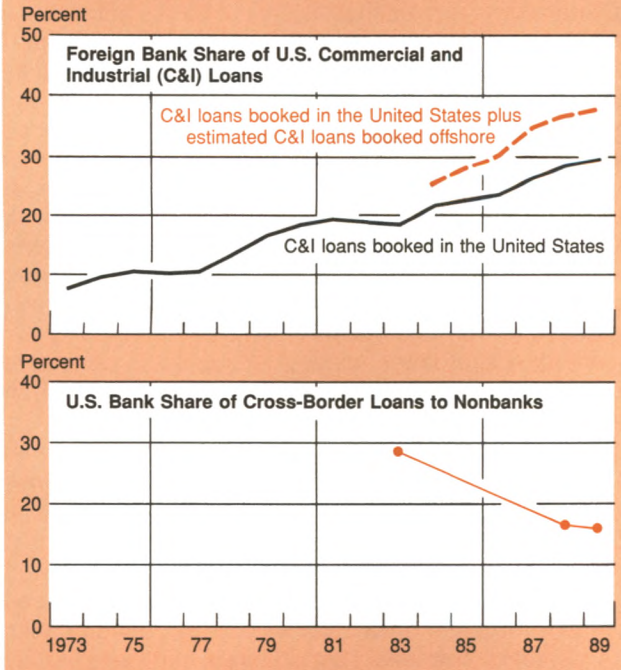
This paper was presented on December 6, 1990, at the Federal Reserve Bank of New York Colloquium on the Cost of Capital in the United States.

country-by-country gains by foreign banks argue for the importance of capital costs. The foreign banks with the greatest capital cost advantage have increased their market share most, while the foreign banks with little or no capital advantage have showed scant if any gains. Moreover, in the market for credit enhancement for commercial paper and municipal bonds, where equity capital costs are even more critical than in corporate lending, foreign banks have left even less of the market for U.S. banks. Banks with a cost of capital disadvantage experience particular difficulty competing in low-risk and low-value-added products since they have less potential to offset their disadvantage through lower labor and overhead costs, better production economies, and better risk management and assessment.

We begin by defining the cost of capital for banks as the fee or net spread between bank borrowing and lending rates that a financial product must generate in order to increase the market value of the bank. While usage of the term "cost of capital" in reference to banks varies, our focus on a required spread (or fee) has much to recommend it in a world of multicurrency lending and off-balance sheet banking. Since an international standard on bank capital limits a bank's ability to leverage up its equity, a bank's cost of capital is largely determined by the value that the stock market assigns to a bank's earnings and, to a lesser extent, by the risk premium paid on its subordinated debt.

Chart 1

Market Shares in Bank Lending



Sources: Bank for International Settlements; Federal Financial Institutions Council, Reports of Condition; Federal Reserve; Federal Reserve Bank of New York staff estimates.

Note: 1989 datum is estimated.

Defining the cost of capital

The cost of capital for banks differs from the cost of capital for industrial firms in two key respects. First, the cost of equity facing a bank assumes paramount importance, despite the fact that banks are more highly leveraged than commercial firms. Second, if the international accord on required capital ratios for banks binds at the margin, the equity required for a given project is readily quantified.

The primacy of equity

A bank's cost of capital for a financial product is the spread or fee that allows the required regulatory capital to earn the rate of return demanded by the market. If a bank prices a product below its cost of capital, the bank inflicts a loss on its shareholders.

Over two years ago, bank regulators from the major industrial countries established international regulatory constraints on bank pricing.[3] The so-called Basle Agreement requires banks from 1992 to hold 4 percent tier 1 capital, or shareholder equity, and 4 percent tier 2 capital, including subordinated debt, against risk-weighted assets. In the discussion below, we concentrate first on the more important tier 1 capital—henceforth referred to as equity.

While regulation sets a minimum on equity required for a bank product, the market determines the cost of that equity. A simple example illustrates how bank managers assess the cost of equity from the market. Suppose that bank managers are weighing a proposed corporate loan. For simplicity, assume that the credit committee judges the contemplated loan to pose much the same risks as the bank's other assets.[4] Bank management should issue new equity to support the loan if doing so promises at least to maintain the market value

[3]Committee on Banking Regulations and Supervisory Practices, Bank for International Settlements, "International Convergence of Capital Measurement and Capital Standards," July 1988, in Raj Bhala, *Perspectives on Risk-Based Capital* (Rolling Meadows, Ill.: Bank Administration Institute, 1989), Appendix, pp. 193-235.

[4]See Appendix C for the effects of relaxing this assumption.

[5]The bank may raise new equity through retained earnings or free up equity by asset sales. The cost of such equity is tantamount to the cost of new issuance if we interpret the bank's action as forgoing repurchase of its shares.

of its outstanding shares.[5] For the share issuance not to lower the bank's share price, the return on the new equity devoted to the contemplated loan must be at least as great as the profit rate on the bank's existing equity. The profit rate is defined as after-tax earnings divided by market value of equity.

Of course, making the loan runs up operating and other expenses. Thus the spread on the loan must be wide enough so that once labor costs, physical capital costs, expected loan default losses, and other expenses are accounted for, the after-tax earnings on the loan stand in the same relation to the allotted equity as overall bank earnings bear to the market value of the bank's outstanding shares. The net (after expenses) spread required to generate this required equity return is the bank's cost of capital for the loan.

Suppose the stock market consistently assigns a bank a share price twenty times earnings, and therefore bank management takes its required profit rate to be 5 percent. In considering a loan of average risk, the bank's managers would require that the loan return an annual profit of 5 percent on the 4 percent portion of equity allotted to it. This is equal to an after-tax return of 20 basis points on the value of the loan. If national and local taxes claim 50 percent of earnings, the bank needs a pretax net spread of 40 basis points just to cover its equity costs. Thus, the bank's cost of capital for a corporate loan is 40 basis points.

The role of debt

If banks may leverage every dollar of equity with $25 of assets, it may seem strange that the cost of capital as defined does not include the cost of debt financing—except the cost of subordinated debt in tier 2 capital. It should be recognized at the outset that off-balance sheet products such as letters of credit, commitments to lend, and interest rate swaps require equity to support them, in proportions set by the Basle Agreement, but no funding with debt. Interest rates have no direct bearing on the financial cost of these products to a selling bank.

We neglect deposit funding costs even for loans or other funded products because internationally active banks from different countries competing in the same market tend to fund themselves at similar interest rates at the margin. Under normal circumstances, one internationally active bank will pay much the same as another for deposits in London and other wholesale markets. Acquiring and holding even low-cost consumer deposits entail promotional and operating costs that tend to raise marginal costs to wholesale funding levels. Banks compete on the basis of the markup they charge over their cost of borrowed funds.

We readily concede that banks tend to enjoy a slight home-court advantage in selling deposits in their home currencies and in their home markets. Since the home-court advantage is often misunderstood, however, the following sections compare debt costs in the United States and Japan and consider whether low real interest rates at home actually help banks.

Debt costs and competition in U.S. commercial lending. Foreign banks ordinarily pay a 10-basis-point premium on their Yankee (foreign bank) certificates of deposit or purchased federal funds. Indeed, in late 1990 some foreign banks paid considerably more for end-year funding. If a foreign bank pays an extra 10 basis points—about 6 basis points after tax—on 24/25ths of its liabilities, it needs a 146-basis-point advantage in cost of equity on the last 1/25th (4 percent) just to pull even with its U.S. competitors.[6]

In the U.S. corporate loan market in particular, however, it is by no means clear that large U.S. banks have consistently enjoyed such a home-court advantage in wholesale funding. Although foreign banks have had to pay a premium to raise dollars in the United States, they may well have enjoyed a funding advantage in lending to U.S. firms in the late 1980s. Until December 1990, a U.S. bank had to hold a 3 percent Eurodollar reserve against funds raised in the Eurodollar market if the bank's home [illegible] had net obligations to the bank's foreign branches. A 3 percent reserve was also required on large certificates of deposit issued by U.S. offices. Consequently, once the bank had brought money into the United States from its foreign branches on a net basis, it faced no reserve incentive for raising funds offshore as opposed to onshore. A U.S.-chartered bank could not get around the reserve requirement by booking a loan to a U.S. corporation offshore because such loans were included in the computation of the required Eurodollar reserve. But since the consolidated reporting required to enforce this provision was not available for foreign banks' branches and agencies operating in the United States, they could avoid the Eurodollar reserve requirement by booking a loan to a U.S. firm offshore.

As long as the large U.S. certificate of deposit rate remained well below the London Interbank Offered Rate (LIBOR), the Eurodollar reserve requirement did not work to the disadvantage of U.S. banks. But as the United States drew in funds from the international banking system to finance its current account, domestic rates rose relative to offshore dollar rates. By mid-1987 a foreign bank able to borrow at LIBOR with no reserve to fund a loan to a U.S. corporation could regularly

[6] .96 {10 * (1—.39)} = .04*X*, where .39 represents combined federal, state, and local taxes. Solving for *X*, we have *X* = 146.4.

[7] See "Revision of Regulation D," *Federal Reserve Bulletin*, September 1980, pp. 758-73.

access cheaper funding than a U.S. bank choosing between issuing a reservable domestic certificate of deposit or bidding for a reservable Eurodollar deposit (Chart 2, shaded area). Data reported to the Bank for International Settlements by industrialized countries show that offshore claims on U.S. nonbanks reported by foreign banks grew from $48 billion at end-1984 to $189 billion at end-1990 and to $194 billion by mid-1990.[8] The Eurodollar reserve requirement may have been necessary for monetary control, but its uneven application surrendered some portion of the home-court advantage possessed by U.S. banks in borrowing dollars.

Debt costs and competition in Japanese commercial lending. The U.S. Treasury has argued that deposit regulation has put foreign banks operating in Tokyo at a disadvantage. Japanese banks have funded themselves with (a falling proportion of) consumer deposits carrying low, regulated interest rates, while foreign banks do not have access to such funds. The practice of average cost pricing on corporate loans has tended to pass through the benefit of regulated rates to corporate borrowers and to make foreign banks particularly uncompetitive.[9]

[8]Bank for International Settlements, *International Banking and Financial Market Developments*; and *Federal Reserve Bulletin*, various issues, Table 3.14.

[9]As the Treasury notes, "Loan charges are based in part on the blended cost of funds to domestic institutions, which continue to be below those of foreign banks." Such pricing casts doubt on the notion that Japanese banks reap extraordinary profits from cheap deposits at home that can be used to finance expansion abroad. U.S. Department of the Treasury, *National Treatment Study* (Washington: Government Printing Office, 1990), p. 218.

Do low real interest rates at home help? In competing to offer borrowers narrow spreads over the cost of borrowed funds, a bank from a country with lower interest rates, whether or not adjusted for inflation, has no direct advantage over banks from a country with higher real interest rates. In this respect, competition among banks differs from competition among industrial exporters, which gives the edge to those with access to cheap money at home. A bank that borrowed in a low-interest currency to lend in a high-interest currency could expose itself to enormous risks. Only indirectly can low real interest rates at home help by showing up as cost of equity differences, but such differences are captured in our cost of capital measures.

The cost of equity

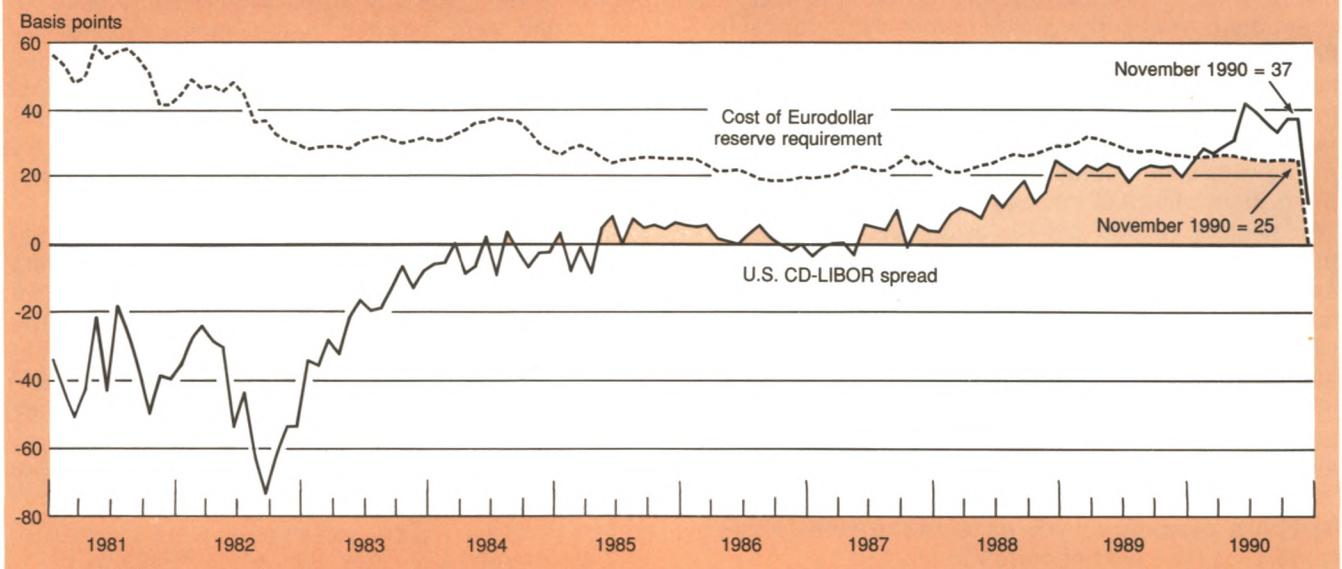
The return on a financial product has to be high enough to cover the required profit rate on the equity allotted to it. This required profit rate is best conceived of as the profit rate the bank can expect to sustain in the long term.

Conceptual problems

Although long-term sustainable profit rates cannot be

Chart 2

Cost of the Eurodollar Reserve Requirement and the Spread between Adjusted CD Cost and LIBOR



Note: CD cost is adjusted for FDIC insurance cost and reserve requirement of 3 percent.

observed, we can observe prevailing profit rates and adjust them to render them comparable across countries. Before explaining these adjustments, however, we consider three potential problems in using current profit rates as proxies for long-term profit rates—growth, cyclicality of profits, and undercapitalization.

Profitability. If investors expect a bank's profitability to rise, its current profit rate understates its true cost of equity because investors are paying up for earnings not yet in evidence. Thus, new firms perceived to have extraordinary growth potential and therefore priced at high multiples of current earnings are generally not thought to enjoy particularly low cost of equity. One must distinguish between growth in profits and growth in profitability, however. A bank's profits may grow simply because the bank reinvests a high proportion of its earnings or it issues new shares in volume. Growth in profitability requires more earnings from a given amount of capital and would have to result from a change in market structure, a change in cost structure, or other fundamental changes not easily achieved.

Japanese and U.K. banks' earnings prospects may raise the issue of profitability growth. In the United Kingdom, the entry of building societies into mainstream banking and the maturation of the consumer credit business point to lower bank earnings in the future; in Japan, deposit deregulation signals a reduction in profits for banks. In these instances, current profits overstate future profits and cost of equity is overstated. But if Japanese banks have, as some claim, pursued a market share strategy abroad with a view toward eventually raising spreads, then current profit rates may understate the cost of equity. Since these growth considerations are clearly difficult to quantify and, in the case of Japan, work in offsetting directions, no adjustment is made for growth in computing the cost of equity, although the possibility of the need for such an adjustment is recognized.

Cyclicality of profits. The second problem with using current profit rates as a measure of the cost of equity arises from the cyclicality of profits. If a firm is having a bad year, its profits do not proxy future profits well. The stock market may recognize a temporary downturn and price the shares of the firm in anticipation of higher future profits. In this case the current profit rate would understate the cost of equity.

The cyclicality of profits represents a significant theoretical shortcoming of the use of profit rates to measure the cost of equity. In practice, however, the cyclicality of profits is not problematic because of two factors: the behavior of bank managers and the behavior of the equity market.

Bank managers, like most corporate managers, generally seek to smooth reported profits because equity markets tend to reward steady profitability. Bank managers command a considerable array of devices to smooth reported profits, including setting reserves and recognizing capital gains on, for instance, real estate. If managers approximate current profits to sustainable profits by realizing gains in bad years and not doing so in good years, then stated profits may represent long-term earnings better than one might otherwise think.

A second factor mitigating the measurement problems associated with the cyclicality of profits is the apparent short-term horizon of most equity markets. If a firm has an unusually poor year, then equity markets might be expected to recognize that profits will recover in the future and to price the equity accordingly. In reality, equity markets appear to price shares largely on current performance. Evidence for this observation can be found in the fact that price-earnings ratios tend to be either noncyclical or even procyclical.[10]

Undercapitalization. Cost of equity can easily be overstated for an undercapitalized bank. If asset losses reduce a bank's equity to levels below international capital standards, the bank must reduce assets or issue new equity. A bank with $100 billion in assets but only $2 billion in equity, for instance, faces the choice of raising $2 billion in equity to meet the Basle standard or selling off $50 billion in assets. If new equity is issued, the current shareholders will share current earnings with the new owners; if assets are reduced, the current shareholders will lose the income earned by the assets. In either case, earnings per share are set to decline. Investors for their part should recognize the impending dilution of their claim or asset shrinkage and value the share in anticipation of reduced earnings per share. As a result, the current earnings in relation to market capitalization of an undercapitalized bank will tend to overstate its cost of equity.

The market value of a poorly capitalized bank is reduced for two reasons other than anticipated asset shrinkage. First, a poorly capitalized bank is riskier because it is highly leveraged. Second, the loss of value on a portion of a bank's assets raises questions about management. The effect of anticipated asset reduction on the market value of the bank is almost certainly greater than these other two effects since it represents a first-order effect—reduction in future profits—as opposed to just an increase in volatility of profit, a second-order effect. Thus, while it is often said that an undercapitalized bank has a high cost of equity because of its riskiness, it is probably more correct to say that the appearance of a higher cost of equity will

[10]Robert N. McCauley and Steven A. Zimmer, "Explaining International Differences in the Cost of Capital: the United States and United Kingdom versus Japan and Germany," Federal Reserve Bank of New York, Research Paper no. 8913, August 1989.

change as such a bank adjusts assets to equity.

Measuring the cost of equity

Reported profits of banks cannot be directly compared across countries because of different accounting practices, different economic conditions, and the interaction of the two. We make four separate adjustments to stated profits: an adjustment for the differential treatment of developing country debt by banks across the different countries, an adjustment to impose equity accounting on shares held by Japanese and German banks, an adjustment for the interaction of growth and inflation with banks' net nominal asset position, and an adjustment for discrepancies between stated depreciation charges and economic depreciation.

The income data adjusted are taken largely from the annual reports of banks in the sample. The banks in the sample are listed in Table 1. Summary measures of the adjustments appear in Table 2.

Loan loss reserves for developing country exposures. Losses on assets introduce greater difficulty in measuring a sustainable profit rate for a bank than for an industrial firm. When an industrial firm writes down assets, appraisers can refer to next best use or even scrap values. Valuing a substandard loan, by contrast, requires a judgment of the borrower's capacity to pay as well as any collateral asset value. Bank stock analysts as a result make larger errors in forecasting earnings than industrial stock analysts.

One can adjust for asset quality problems by lowering earnings over time after a problem becomes evident, as banks generally do and as we do below. As an alternative, one may restate earnings between the booking of an ultimately problematic loan and the emergence of the problem, so that an ex post appropriate reserve is built up beforehand.

The banks in this study all have significant claims on developing countries that have restructured their debts, and all have reported lower profits as a result of making

Table 1
Banks in Sample (by Country)

United States
Bank America Corporation
Bankers Trust New York Corporation
Chase Manhattan Corporation
Chemical Banking Corporation
Citicorp
First Chicago Corporation
J.P. Morgan and Company
Manufacturers Hanover Corporation
Security Pacific Corporation

Canada
Bank of Montreal
Bank of Nova Scotia
Canadian Imperial Bank of Commerce
Royal Bank of Canada

Switzerland
Credit Suisse
Swiss Bank Corporation
Union Bank of Switzerland

Japan
Bank of Tokyo
Dai-Ichi Kangyo Bank
Fuji Bank
Industrial Bank of Japan
Long-Term Credit Bank of Japan
Mitsubishi Bank
Sanwa Bank
Sumitomo Bank

Mitsubishi Trust and Banking
Mitsui Trust and Banking
Sumitomo Trust and Banking

United Kingdom
Barclays
Lloyds Bank
Midland Bank
National Westminster Bank

Germany
Commerzbank
Deutsche Bank
Dresdner Bank

Table 2
Summary of Adjustments to Cost of Equity
(Percentage Point Adjustments)

Period Averages 1984-90	United States	Japan	United Kingdom	Canada	Germany	Switzerland
Developing country	4.95	−0.18	0.60	0.86	0	0
Cross-holding	0	1.50	0	0	1.53	0
Depreciation	−1.51	−0.05	−2.88	−1.39	−0.64	−0.39
Net nominal assets	−1.16	−0.15	−0.23	−1.10	−0.67	−0.41
Total	2.28	1.13	−2.51	−1.63	0.22	−0.80
Developing Country Year by Year						
1984	−12.93	−0.71	−6.46	−11.91		
1985	−7.63	−0.72	−3.24	−6.85		
1986	−5.59	−0.24	−1.47	−3.81		
1987	54.19	0.15	7.44	18.61		
1988	−12.79	0.17	−1.00	0.07		
1989	20.06	0.32	8.32	8.47		
1990	−0.73	−0.25				

provisions for possible losses. The extent of reserving or charging-off of such developing country loans varies by country and by bank—with the extremes ranging from less than 50 percent reserved to 100 percent reserved. Part of the difference in reserving against losses reflects differences in portfolio composition, and thus performance, and differences in bank commitment to the relevant markets. Quite apart from such differences in composition and bank strategies, however, bank managements have reserved in accordance with their own outlook for servicing and with national norms, regulation, and taxation.

Increases in loan loss reserves reduce profits: banks that have reserved heavily against restructured developing country loans report, all other things being equal, lower profits than banks that have not reserved as heavily. To render bank profits more comparable, we attempt to impose a uniform profile of reserving against such loans. Reported profits are lowered in years that banks have reserved less than is indicated by the uniform profile; conversely, reported profits are raised in years that banks have reserved more than is warranted.

Although only uniform treatment of restructuring country exposures permits a comparison of bank earnings and thus of cost of equity for banks in the late 1980s, no definitive benchmark for reserving is available. We set the uniform reserve ratio by end-1989 toward the high end of international usage to reflect our understanding of German and Swiss practice. This approach is the most practical, because standards of disclosure for German and Swiss banks make it very difficult to restate their accounts to any other standard.

In our adjustment we restate profit flows as if the reserves were taken on an after-tax basis; that is, we assume that the bank gets a full tax break on all reserves. In reality the tax authorities in some countries, notably the United States and Japan, have not generally allowed their banks to reserve on an after-tax basis (see equations, Appendix A). The effect of national differences in tax treatment of developing country reserves is often overstated: if losses are in the event realized on such portfolios, U.S. and Japanese banks share their losses with their respective governments; if the loans provided against ultimately perform, additional income will be recognized and taxes will be paid. Thus, differences in tax treatment will ultimately prove to be differences of timing. The time value of early deductibility of potential loan losses is not trivial, but it should be noted that transactions such as debt-equity swaps serve to bring forward tax benefits.

The decline in the dollar since 1985 has tended to shrink the developing country debt problem for foreign banks. Because most of the claims on restructuring countries were denominated in dollars, successive rounds of reserving have cost foreign banks less. For instance, reserving against 5 percent of their troubled country claims cost eight Japanese banks an average of 27 percent of their net income in the year to March 1985. After the yen appreciated from 250 to the dollar in March 1985 to 124 in March of 1988, however, reserving against 10 percent of the same banks' restructuring country exposure cost them only 31 percent of their net income. The importance of the dollar-yen exchange rate to Japanese banks was underscored by market reports of Japanese banks' bidding for $1.3 billion to cover the loss of dollar assets entailed in the exchange of Venezuelan debt in December 1990.[11]

Cross-held shares. Banks in Japan, Germany, and Switzerland hold substantial equity stakes in commercial and industrial firms. The dividends from these shareholdings contribute to a bank's income, but the retained earnings associated with the shares do not. Retained earnings show up in earnings only when capital gains are harvested through the sale of shares.

Haphazard realization of capital gains can misstate income quite seriously. A bank may go for extended periods without realizing capital gains on its shares in order to delay payment of taxes. In this case the profits of the bank would be understated. We hold that, for the correct statement of income, retained earnings associated with bank shareholdings should be consolidated with the bank's income. In other words, retained earnings associated with shareholdings, as well as dividends actually received, contribute income in the long term.

We correct for this problem in two steps. The first step is to subtract from after-tax income the after-tax capital gains from equity realizations. The second step is to add the retained earnings associated with the shares (less effective income tax on the earnings) to after-tax income. For Japanese banks, the retained earnings rates used are taken from our earlier estimates for nonfinancial firms and reflect all other adjustments to stated profits (see equation in Appendix A).[12]

Although such an adjustment serves consistently to narrow the differences between price-earnings ratios of Japanese and U.S. industrial firms, it bears on the comparison of Japanese and U.S. banks in a markedly different way. Japanese banks traditionally resembled other Japanese firms in leaving long-held shareholdings on their books at cost. As a result, their profit on sale of equity holdings fell well short of the retained earnings

[11]Konosuke Kuwabara, "Dealers See Dollar Staying Close to the Level of ¥132," *Japan Economic Journal*, December 22, 1990, p. 22.

[12]Robert N. McCauley and Steven A. Zimmer, "Explaining International Differences in the Cost of Capital," Federal Reserve Bank of New York *Quarterly Review*, vol. 14 (Summer 1989), pp. 7-28.

associated with their shareholdings. As the shape of the international agreement on capital adequacy became clearer, however, Japanese banks stepped up their realization of gains on long-held shares, reportedly by selling and repurchasing shares so as not to disturb the pattern of cross-shareholdings.[13] City banks in particular stepped up their realizations since they had farther to go to meet the Basle standard. By 1988 and 1989, capital gains on shareholdings exceeded the retained earnings associated with cross-held shares by a wide margin and also accounted for a large portion of pretax income (Chart 3).

Thus, the dormancy of Japanese city banks' shareholdings into 1987 understated their earnings, but heavy turnover in the three years to March 1990 overstated earnings. Churning of the equity portfolio in these years masked weak earnings and built up shareholders' equity.[14] We hypothesize that bank managers accepted the tax costs of realizing the gains in order to show enough earnings growth to market their equity in quantity.

German banks appear to have realized gains on their holdings of equities in the late 1980s much less often than Japanese banks. So, even though the fraction of German equities held by German banks represents only about half the fraction of Japanese equities held by Japanese banks,[15] adjusting for cross-held shares

[13]Robert Zielinski and Nigel Holloway, *Unequal Equities: Power and Risk in Japan's Stock Market* (Tokyo: Kodansha International, 1991), p. 187.

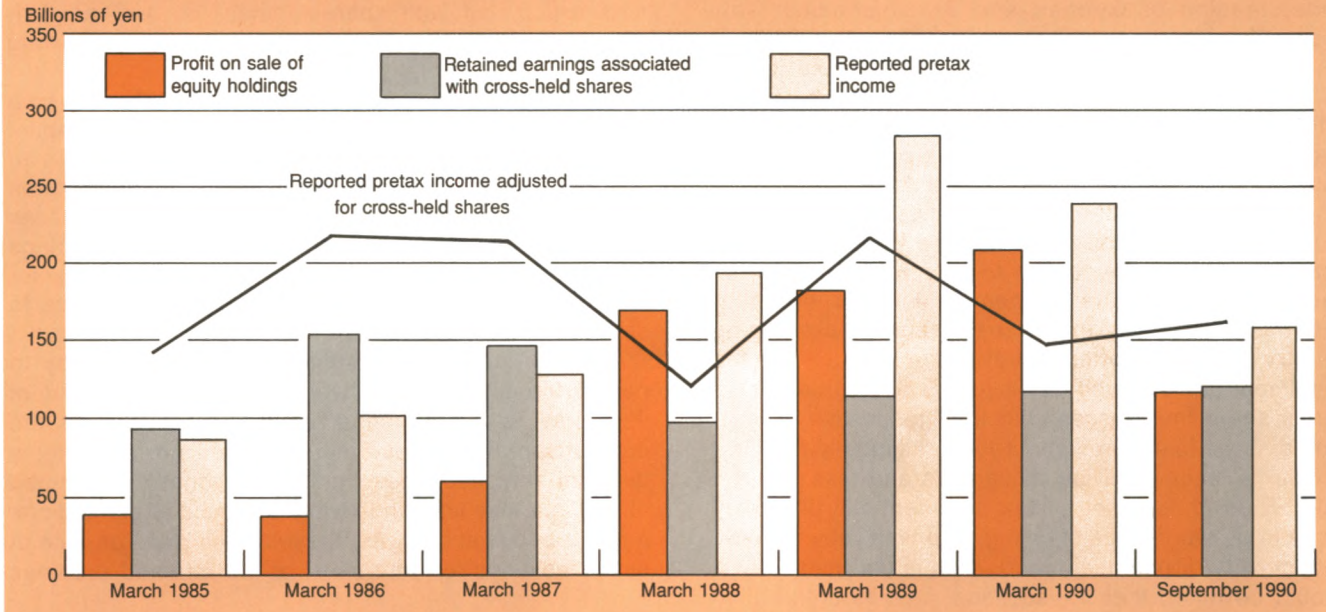
[14]Recognition of the real weakness of Japanese banks' earnings for five years is key to any assessment of the cost of equity through

Footnote 14 continued
the anticipated dividend yield. Any growth in dividends for the last five years does not proxy earnings growth but instead represents a rise in the effective payout ratio. For an application of the anticipated dividend method that concludes that the required return for Japanese banks is higher than that for U.S. banks, see Jack Glen and Richard Herring, "P/E Multiples: Comparative Evidence for Japan and the United States," University of Pennsylvania, August 1990 (processed).

[15]According to Claudio E.V. Borio, *Leverage and Financing of Non-Financial Companies: An International Perspective*, Bank for International Settlements, Economic Paper no. 27, May 1990, p. 30, German banks hold 8 percent German shares, although data on participations in unlisted firms are not available; Japanese banks hold 17 percent of Japanese shares. The shareholdings of the German banks may be more concentrated in the three sample banks.

Chart 3

Japanese Bank Earnings from Cross-held Shares



Sources: Profit on sale of equity holdings is from Nikkei Newsletter on Bond and Money for 1988, 1989, and 1990 and from Federal Reserve Bank of New York staff estimates for previous years; pretax income is from annual reports; retained earnings associated with cross-held shares are Federal Reserve Bank of New York staff estimates.

Notes: Chart plots averages of six city banks and two long-term credit banks. Results for the six months to September 1990 are doubled.

boosts the cost of equity as much for German banks as for Japanese banks in our sample period.

Swiss banks hold shares as well, but their standards of disclosure render it very difficult to know how much their reported incomes should be adjusted. Unlike Japanese banks, Swiss banks serve as market makers in their domestic stock markets, so they turn over at least some of their shares enough to realize capital gains and thereby to show retained earnings on shareholdings in their income statements. Moreover, because Swiss firms pay out more of their earnings than Japanese firms, the omission of retained earnings on crossheld shares produces less downward bias in earnings.

A first glimpse of the hidden wealth of a Swiss bank permits estimates to be made of the understatement of earnings and thus the cost of equity for Swiss banks. Union Bank of Switzerland disclosed hidden reserves of 2 billion Swiss francs at the end of 1989, as compared with reported shareholders' equity of 10.6 billion Swiss francs without the newly disclosed amount. The sources of income on the hidden reserves reported by the bank for 1989 connect the reserves to the bank's shareholdings: "premium income received from the exercise of conversion rights and warrants as well as the dividends on shares reserved for warrants."[16]

Three different ways of assessing the understatement of earnings associated with these heretofore unacknowledged post-tax retained earnings yield strikingly similar results.[17] Whether our estimates of the cost of equity are biased downward by undisclosed income of Swiss banks depends on the degree to which investors incorporate the hidden income in the price that they are willing to pay for the shares. If they have paid for hidden income, then our estimates understate the cost of equity by about 1 percentage point for Swiss banks.

Inflation-related adjustments. Adjusting stated profits to eliminate inflation effects will reveal a sustainable *real* rate of return. The first adjustment aligns the different time profiles of returns on real and nominal assets. A second adjustment removes the portion of a bank's earnings that simply maintains the real value of the shareholders' stake. Without the latter adjustment, Brazilian banks finish first in the bank profit league. In reality, income runs just to keep equity in place.

To see why the first adjustment is necessary, compare the income flows from a nominal and a real asset in inflation. A portion of the income flow from a nominal asset such as a bond merely compensates for the inflationary erosion of the real value of the principal. The income flow from a real asset such as real estate or equity represents a real return since inflation does not on average erode the value of the principal over time.

If a nominal liability funds a real asset of equal value held for two years in an environment of steady inflation and no asset price risk, then accounting flows at first understate income and then overstate income. In year one, the servicing cost of the nominal liability exceeds the receipt from the real asset by the inflation rate, and the position produces a loss. At the end of year two, however, the sale price of the real asset has increased with two years' worth of inflation and so exceeds the cost of retiring the nominal liability. This gain from inflation overwhelms the net servicing cost in the second year, and net income is reported. The investment shows no net profit over the two years, and indeed the theoretical net income each year is zero.

No problem arises with steady inflation and no asset growth. Inflationary gains on real assets find their way into the income statement either at sale or as the returns on the assets grow with inflation. As a bank grows, however, or as inflation varies, profits are misstated.

The effects of varying inflation and asset growth depend on the balance sheet of the bank. U.S., U.K., and Canadian banks tend to have more nominal assets than nominal liabilities; that is, they hold net nominal assets. German, Japanese, and Swiss banks, holding larger shares of real assets such as equities or leasable assets, have more nominal liabilities than nominal assets. This contrast reflects the different banking traditions: strictly commercial banking in the English-speaking countries as against a combination of commercial and industrial banking in the continental and Japanese economies, which industrialized later. Because of these differences, U.S., U.K., and Canadian banks tend to overstate profits if inflation rises or assets grow, while the other banks tend to understate profits under these circumstances.

The adjustment for the different time profiles of nominal and real assets takes two steps. To adjust for the misstatement of profit owing to shifting inflation, we subtract from stated earnings the product of the net nominal assets and the difference between the prevailing inflation rate and an average of the inflation rates over the life of the bank's real assets. To adjust for growth, we subtract from stated earnings the product of the inflation rate and the change in the net nominal position (see Appendix A for equation).

The second adjustment is both simpler and more important. With inflation, a bank's equity must grow with

[16]Union Bank of Switzerland, *Annual Report, 1989*, p. 22.

[17]UBS reported that the hidden reserves generated income in 1989 that amounted to 1.0 percent of the bank's market capitalization at end-1989 and 0.7 percent after taxes. If the heretofore undisclosed reserves grew at the 10 percent average annual growth rate shown by assets from 1979 to 1989, then undisclosed income would have ranged from 0.6 percent to 1.2 percent of market capitalization in 1984-89 and averaged 0.9 percent for the period. If the hidden reserves instead grew in line with reported shareholders' equity, then undisclosed income would have ranged from 0.4 percent to 1.3 percent in 1984-89 and averaged 0.9 percent of market capitalization.

the price level just to maintain its real value. The portion of a bank's retained earnings that suffices to hold constant the real value of shareholder equity does not contribute to real sustainable profits. Failing to reduce profits accordingly biases upward the apparent required profit rates of banks in high-inflation countries.

This adjustment is required only for that portion of shareholders' equity in excess of the bank's depreciable assets.[18] As described below, the adjustment for the difference between economic and stated depreciation accounts for the cost of maintaining the real value of assets subject to depreciation. So shareholders' equity less depreciable assets is multiplied by the inflation rate, and the product is subtracted from stated profits (see Appendix A for equation).

Depreciation adjustment. Stated depreciation charges differ from true depreciation for two reasons. Because depreciation is taken on the historical cost of the physical asset instead of its replacement cost, accounting depreciation charges understate true depreciation costs in an inflationary environment. In addition, the depreciation rates allowed for tax purposes may differ from physical depreciation rates.

To correct for the inflation bias, we first infer the age distribution of the bank's physical assets from past depreciation rates and then mark them to current prices by using cumulative gross national product price deflators. Depreciation on the repriced physical assets is substituted for the depreciation taken on the historically priced assets.

For U.K. banks, corrections were made to reflect the low depreciation rates on buildings in the 1980s that resulted from earlier tax provisions practically allowing the cost of buildings to be treated as an expense. For Japanese banks, corrections were also made to offset accelerated depreciation on some bank buildings.

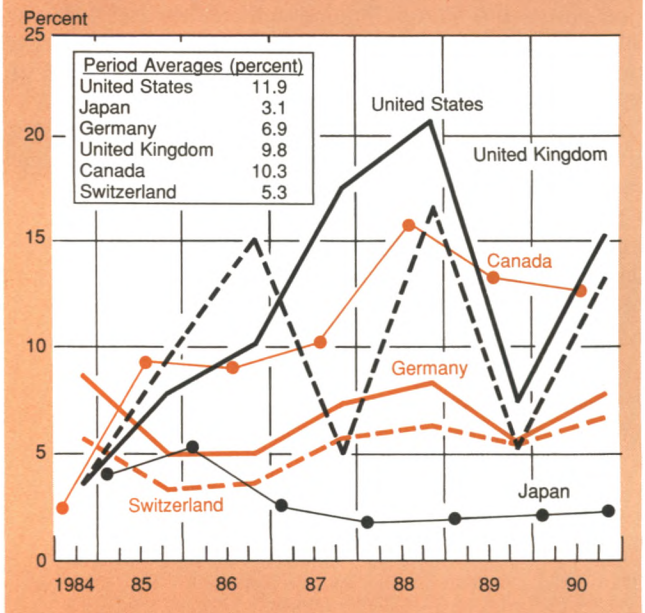
Results

The cost of equity for banks over the sample period 1984-90 varies markedly across countries (Chart 4).[19] In interpreting the figures, one should focus on period averages more than single years because the estimate can be biased for any given year in isolation. This is especially true for 1984 and 1985, when our adjustment for developing country reserves sharply reduces adjusted profits for U.S., U.K., and Canadian banks. For 1984-90 as a whole, equity markets in the United States, United Kingdom, and Canada burden banks with required returns of around 10 percent. German and Swiss banks face moderate equity costs in the 6 percent range, while Japanese banks enjoy very low equity costs of 3 percent.

Japanese bank share prices fell sharply in 1990, but Japanese banks' cost of equity remained low by international standards. The decline of Japanese bank share prices in the first quarter of 1990 only served to restore the cost of equity to its level of a year before. The further decline of Japanese share prices through September 1990, in conjunction with a half-year of flat reported earnings, raised the cost of equity by 1 percentage point in September to 3.1 percent. Even at this level, Japanese bank equity costs remained low in absolute or relative terms. The recovery of Japanese bank share prices in the fourth quarter brought down the cost of equity estimated for end-year. In retrospect, the decline in the cost of equity for Japanese banks in 1987-89 is somewhat surprising since bank stocks peaked in 1987 and have regularly underperformed the

Chart 4

Cost of Equity for Banks



Source: Federal Reserve Bank of New York staff estimates.

[18] For an approach that does not account for depreciable assets, see William M. Peterson, "The Effects of Inflation on Bank Profitability," in Richard G. Davis, ed., *Recent Trends in Commercial Bank Profitability* (New York: Federal Reserve Bank of New York, 1986), pp. 89-114.

[19] Data for U.S., U.K., Swiss, and German banks are from end-December, 1984-90; data for Japanese banks are end-March, 1985-90; data for Canadian banks are from end-October, 1984-90. Income data are partially estimated for 1990. Results for the United States and Britain do not include measures for Bank of America and Midland, respectively; both banks' real *and nominal* assets were shrinking over much of the sample period.

market since. The apparent inconsistency of falling share prices and falling cost of equity is resolved by noting both the weakness of earnings apart from capital gains and the substantial issuance of new shares.

The cost of lower tier 1 and tier 2 capital

The Basle Agreement permits internationally active banks to mix a variety of forms of capital with shareholders' equity in meeting the overall capital requirement of 8 percent against risk-weighted assets. Tier 1 capital can include not only common equity but also certain preferred shares. Qualifying as tier 2 capital are instruments and balance sheet items at the border of equity and debt: long-term preferred shares, revaluation reserves, general loan loss reserves, debt securities to be retired with equity, perpetual debt, and subordinated debt.

The following sections examine the cost of tier 1 preferred shares, tier 2 preferred shares, loan loss reserves, and subordinated debt. Tier 1 preferred shares appear quite cheap by comparison with common equity, although there may only be a market for these in New York. Over time, subordinated debt is likely to supplant loan loss reserves as tier 2 capital; national differences in subordinated debt costs suggest the importance of official support of banks to investors in bank capital instruments.

The cost of tier 1 preferred shares

Preferred shares are somewhat less costly than equity. This cost advantage reflects the lower risk and, in the United States, the exclusion of most preferred share dividend income from corporate taxation.

To qualify as tier 1, preferred shares must satisfy a stringent standard that has made them relatively rare in banking: they must be noncumulative. In other words, if a bank reaches the point where it eliminates its common share dividend and then eliminates its preferred share dividend as well, it does not promise to make good on the skipped preferred share dividend when and if it resumes paying dividends.

Barclays successfully marketed tier 1 preferred to U.S. investors in the summer of 1989 at a yield quite low by comparison with bank common equity costs. The yield to investors on the $180 million issue was 9.2 percent, and the cost to Barclays, payable out of post-tax income, was 8.2 percent.[20] This latter cost is comparable to subordinated debt costing 3.6 percent above LIBOR and implies a required spread on a corporate loan of roughly a quarter of that implied by British banks' cost of common equity.[21] It is little wonder that British banks are said to be looking to maximize the tax efficiency of such issues.[22]

An important advantage to a British bank of such equity is that it protects the bank's capital adequacy from exchange rate fluctuations.[23] When the pound sterling approached parity with the dollar at the end of 1984, British banks with very substantial dollar books watched their assets rise in relation to their sterling equity. With equity in dollars, dollar appreciation works to raise the sterling value of both assets and equity, so that the ratio of the two is more stable.

If noncumulative preferred shares are marketable in the United States but not abroad, foreign banks may enjoy an advantage over U.S. banks in hedging their capital adequacy. In the absence of a market for noncumulative preferred shares in nondollar currencies, U.S. banks could protect their capital adequacy from dollar depreciation only by taking a long position in the foreign currency. Such a position has the drawback, however, of introducing variability in earnings and capital even as it stabilizes the capital-asset ratio (see Box, p. 50).

The cost of tier 2 preferred shares

The development of the market for variable-rate preferred shares in New York provided banks with relatively inexpensive tier 2 preferred, at least under normal circumstances. For instance, on December 21, 1989, Morgan auctioned $250 million of cumulative preferred for periods of forty-nine to seventy-seven days with an interest rate of 6.75 percent.[24] This yield, payable out of after-tax net income, was equivalent to a deductible rate of about 2.4 percent above three-month LIBOR. The problem with auction-rate preferred, as some banks learned in 1990, is that under adverse circumstances, the bank can face the choice between watching the

[20] The wedge is introduced by the combination of the refundability under the U.S.-U.K. tax treaty of the Advance Corporation Tax, which integrates British corporate and investor income taxes, and the same treaty's withholding tax of 15 percent (*Barclays Bank PLC: Prospectus Supplement*, May 4, 1989, p. S3).

[21] The pretax, floating-rate equivalent cost to Barclays was 3.6 percent over LIBOR by the following calculation: $R/(1-t) - L$, where R (= 8.156 percent) is the post-tax cost of preferred shares, t (= .35) is the British corporate tax rate, and L (= 8.91) percent is the fixed-rate yield that can be swapped against LIBOR. Leveraged up twenty-five times, the required spread to cover this tier 1 equity would be about 17 basis points.

[22] Simon London, "Basle to Decide on Preference Capital," *Financial Times*, January 23, 1991, p. 15.

[23] Brian Pearse, Barclays' finance director, commented that "this issue as dollar-denominated capital will help insulate us—as a sterling-based bank—if the dollar suddenly appreciates further" (see John Evans, "Barclays to Issue Preferred Shares in US To Boost Capital, Reduce Currency Risk," *American Banker*, May 9, 1989, p. 3).

[24] J. P. Morgan, *Annual Report, 1989*, p. 47.

auction fail, paying a high rate, or retiring the preferred shares.

The cost of general loan loss reserves

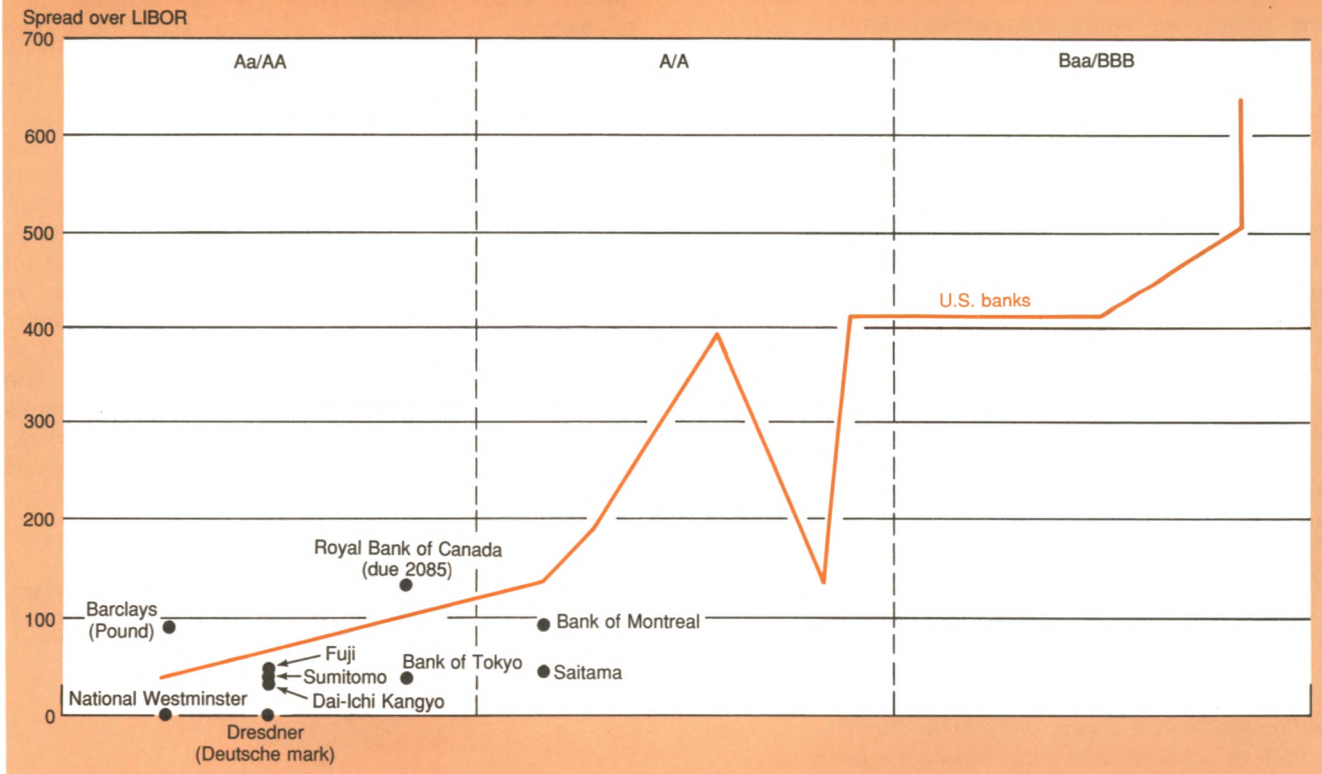
From some banks, general loan loss reserves will be very expensive; for others, they will be essentially costless. An addition to loan loss reserves usually comes at the expense of retained earnings, a component of shareholders' equity. As a result, the cost of general loan loss reserves may be taken to be as high as the cost of equity. (If a general loan loss reserve can be established out of pretax funds, then its cost is reduced by the time value of the early receipt of a tax reduction.) To the extent that a bank reserves for reasons other than meeting capital requirements (for instance, if regulators compel the bank to raise nonspecific reserves or if bank managers are simply underscoring their overcapitalization), then the reserves do not represent a marginal capital cost to the bank.

The cost of subordinated debt

The cost of subordinated debt varies systematically with the bond market's perception of the strength of a bank. Indicators of such strength include the ratings assigned by rating agencies. Plotting the yields (in relation to LIBOR) of fixed- and floating-rate notes in the Eurobond market and fixed-rate notes of U.S. banks trading in the U.S. corporate bond market highlights the premium that investors demand for accepting subordinated status (Chart 5). Nevertheless, four cautions are in order in interpreting these observations: ratings are not universally accepted as indications of bank strength, as evi-

Chart 5

Subordinated Debt Costs, November 1990



Sources: Association of International Bond Dealers; International Financing Review; Morgan Securities; Morgan Stanley.

Notes: For fixed rate bonds and notes, spreads over LIBOR equivalent to spreads over Treasury yields were calculated using appropriate midpoints of interest rate swap rates. Horizontal placement of observations within rating class reflects subdivisions of rating classes and any split in rating by agencies.

denced even by disagreements between rating agencies; yield information reflects indications by market-makers, rather than transaction prices; secondary market yields do not necessarily reflect rates obtaining on new issues; and these yields reflect a particular market environment, rather than an average of market environments.[25]

These qualifications aside, do banks from some countries enjoy an advantage in raising subordinated debt over banks from other countries? It appears that foreign banks' subordinated debt costs do not rise as quickly in response to lower ratings as the costs of U.S. banks. In Chart 5, foreign bank costs suggest a flatter curve than that described by U.S. banks' costs.

If foreign governments offer banks stronger support, then investors in bank bonds do not bear as much risk. (This interpretation is only strengthened by the rating agencies' practice of factoring into their ratings their perception of relations between banks and governments.) Since Continental Illinois, U.S. policy has regularly permitted bondholders of failed banks to take losses; foreign governments, by contrast, maintain ambiguous policies in this matter. In addition, banks with strong links to their industrial customers, including shareholdings, may enjoy a built-in market for subordinated debt that holds down debt costs at higher risk.

If these subordinated debt costs are taken to be representative of tier 2 capital costs, then tier 2 costs add anywhere from 1 basis point to 10 basis points to the equity cost of making a corporate loan. (Subordinated debt is capped at 50 percent of tier 1 capital but may account for 100 percent of tier 2 capital at the margin for some banks.) If a bank has to pay a pretax premium of 25 basis points for subordinated debt, then with a tier 2 requirement of 4 percent (1/25th) for a corporate loan, the bank needs to build another basis point into the loan spread before taxes. If a bank has to pay 200 basis points over LIBOR on its subordinated paper, it needs to charge an additional 8 basis points more before taxes to cover its tier 2 debt costs.

Cost of capital for financial products

Once bank managers have calculated their cost of equity, they can work out the spread or fee they must charge on individual products to cover their equity costs. If bank managers can identify the marginal source or relevant mix of tier 2 capital, then they can add its cost to the spread or fee needed to cover the cost of equity. Tier 2 capital costs are not included in capital costs as reported below.

[25]November 1990 was by no means a typical market environment for subordinated bank debt, but the sale by Japanese banks of subordinated debt in late summer 1990 permitted a wider set of observations.

The Basle Agreement has set the risk weight for each product as follows:

Product	Weight
Corporate loan	100 percent
Commitment to lend	50 percent
Interest rate swap	5 percent of notional amount plus 100 percent of positive mark to market value

These weights mean that a corporate loan absorbs shareholders' equity at the rate of 4 cents on the dollar, a commitment to lend absorbs at a rate of 2 cents per dollar, and an interest rate swaps absorbs 0.2 cents per dollar of contract initially. Given anticipated interest rate movements, the effective weight on the interest rate swap is more than double the initial weighting, so a bank should price a swap as if it absorbs 0.4 cents of equity.[26]

Cost of capital formulas for funded and unfunded products

Armed with an estimate of the cost of equity, the overall requirement of 4 percent equity, Basle risk weights, and corporate income tax rates, the bank manager can roughly calculate the required spread. The spread actually charged must be high enough so that, after the bite of income tax, the required return on the allotted equity is covered:

$$\text{(1) spread} * (1-\text{tax rate}) = \text{cost of equity} * \text{risk weighting} * 4 \text{ percent},$$

$$\text{or } S * (1-tc_t) = \text{COE} * \text{RW} * 0.04.$$

Consider a corporate loan. If the tax rate is 40 percent and the cost of equity is 10 percent, then the required spread for the full weight loan would be 0.67 percent, or 67 basis points: spread $*$ $(1-0.4) = 10 * 1.0 * 0.04$.

In reality, the determination of the cost of capital is slightly more complex because the bank is funding 4 percent of the loan with shareholder equity. The payment to equity is therefore the spread on the loan plus the real (net of inflation) after-tax return earned by investing the shareholder equity in a riskless asset. If the bank earns a positive (negative) real after-tax return from the investment of shareholder equity, then the required spread on its loans is narrowed (widened).

[26]The bank should anticipate an average amount of capital it will have to hold against the swaps. This average value reflects several different factors, including the volatility of interest rates, the length of the swap, and the term structure of interest rates. We estimate the average risk weighting of the ten-year interest rate swap to be 11 percent.

Although the more complete equation looks quite different,[27] the results are similar. If we assume the same parameters as earlier, and in addition assume an inflation rate of 5 percent and a riskless rate of 8 percent, then the required spread on the loan would be 71 basis points. A variation of this calculation for an unfunded product with a 100 percent risk weight, such as a standby letter of credit, shows the required fee to be 68 basis points.[28] Appendix B shows how tier 2 capital costs enter into the cost of capital for financial products, and Appendix C discusses the role of risk.

The role of taxes in bank cost of capital

Corporate income taxes play a rather different role in bank cost of capital than they do in cost of capital for nonfinancial firms. While corporate income taxes have an ambiguous effect on the cost of capital for nonfinancial firms, they clearly work to raise the cost of capital for financial firms. This difference results from the fact that financial assets are not physically depreciated and, to a lesser extent, from the way we define bank cost of capital.

Cost of capital for physical projects is defined as the real rate of return an asset must generate to cover the after-tax cost of the funds used to finance it. Corporate income taxes increase the pretax return that must be generated in order to meet a given after-tax return. At the same time, a high corporate income tax rate raises the value of the depreciation deduction and any investment tax credit that the firm can claim, and thereby reduces the required real rate of return. Moreover, the more leveraged a firm is, the less corporate income taxes work to raise required returns because of the tax deductibility of debt.

Generally, if the tax depreciation of an asset is slower than the physical depreciation of the asset, then the corporate income tax works to raise the cost of capital for the asset. If the tax depreciation of the asset is faster than the physical depreciation of the asset, then corporate income taxes may have no effect or may even reduce cost of capital, particularly if the firm is quite leveraged. In practice, the tendency of tax codes to permit physical assets to be depreciated faster than assets are losing economic value makes the cost of capital for physical projects less sensitive to corporate income tax rates.

Since financial assets do not depreciate, the corporate income tax raises the cost of capital by increasing the pretax rate a project must generate to meet a given required after-tax return. If the corporate income tax rises from 25 percent, a rate near the low, Swiss end of the spectrum, to 50 percent, a rate below the high, Japanese end of the spectrum, the cost of tier 1 capital rises by about 50 percent for a given cost of equity. By contrast, corporate income taxes do not directly affect the banks' cost of subordinated debt.

Since banks are highly leveraged, it would appear that their cost of capital should not be sensitive to corporate tax rates. Recall, however, that bank cost of capital is defined as a required spread or fee, and not as an overall interest rate. For example, consider a rise in the corporate tax rate that widens the required loan spread from 30 to 50 basis points. All other expenses aside, the bank needs to lend at a minimum of 8.5 percent instead of 8.3 percent at an interbank rate of 8 percent. Although this increase appears quite small, bankers compete in terms of spreads. Moreover, taxes show up directly in the required fee for products such as letters of credit and swaps.

Empirical results

Combining cost of equity estimates with equations 2 and 3 (see footnotes 27 and 28) produces the required spreads and fees on various financial products over the period 1984-90. The source of the various parameters is as follows: inflation is from the GNP deflator; the tax rate combines federal, regional, and local top-bracket corporate income tax rates; and the nominal interest rate is a riskless rate, approximated by an annual average of LIBOR less 1 percent.

We consider three different financial products: a standard corporate loan, a commitment to lend with a life greater than one year, and a ten-year dollar interest rate swap. The standard corporate loan—a funded

[27] (2) $0.96 * S * (1-tc_t) + 0.04 * \{\frac{1+[r_t * (1-tc_t]}{1+\pi_t} - 1\} = COE * RW * 0.04,$

where
S = required spread on loan (net of noninterest expenses)
tc_t = marginal corporate income tax rate at time t
r_t = riskless nominal interest rate at time t
π_t = inflation rate at time t
COE = cost of equity
RW = risk weighting.

[28] A variation of equation 2 applies to unfunded products such as letters of credit, commitments to lend, or swaps. The bank does not have to float debt to finance an unfunded product by definition, but equity is still required to underwrite the risk of the product. The proceeds from the required equity issue may be thought of as placed at the riskless interest rate. Unlike the spread on the loan, which is counted over the 96 percent portion of the loan financed by debt, the fee counts over the entire value of the project:

(3) $F * (1-tc_t) + 0.04 * \{\frac{1+[r * (1-tc_t)]}{1+\pi_t} - 1\} = COE * RW * 0.04$

where F = required fee on an unfunded project.

If we solve for equation 3 using the same parameters as in equation 2, we get 68 basis points as the required fee on a standby letter of credit, an unfunded project with a 100 percent risk weight.

financial product—is evaluated using equation 2, while the commitment to lend and the swap—both unfunded products—are evaluated using equation 3.

The required net spread on a corporate loan shows substantial variation across countries (Chart 6). A U.S., Canadian, or U.K. bank needs net spreads of 60 to 80 basis points while a Japanese bank needs only 10 basis points. It must be kept in mind that these spreads are net of all other expenses. If a U.S. bank has to allow 25 basis points for expected loan losses and another 25 basis points for providing and servicing the loan, then the bank will need a gross spread of about 130 basis points on the loan.

While the pattern of required fees on a commitment to lend in the different countries follows the pattern of spreads on the corporate loan, the results for the interest rate swap merit particular attention. The required annual net fee on this item is between 5 and 10 basis points for banks in the United States, United Kingdom, and Canada. Interest rate swap spreads can dip below 5 basis points. Thus, U.S., U.K., and Canadian banks cannot even earn enough on swaps to cover the cost of tier 1 equity. The problem is particularly acute when one recalls that cost of capital represents the net fee—after all expenses—required on the swap. Recent work by Zimmer indicates that a bank needs about 3 basis points to cover expected default losses on a ten-year interest rate swap.[29] Thus, even if we ignore compensation of the swap team and any overhead, a U.S. bank would need bid/ask spreads in excess of 15 basis points on interest rate swaps to cover the costs of capital and of expected defaults.

International taxation and bank cost of capital

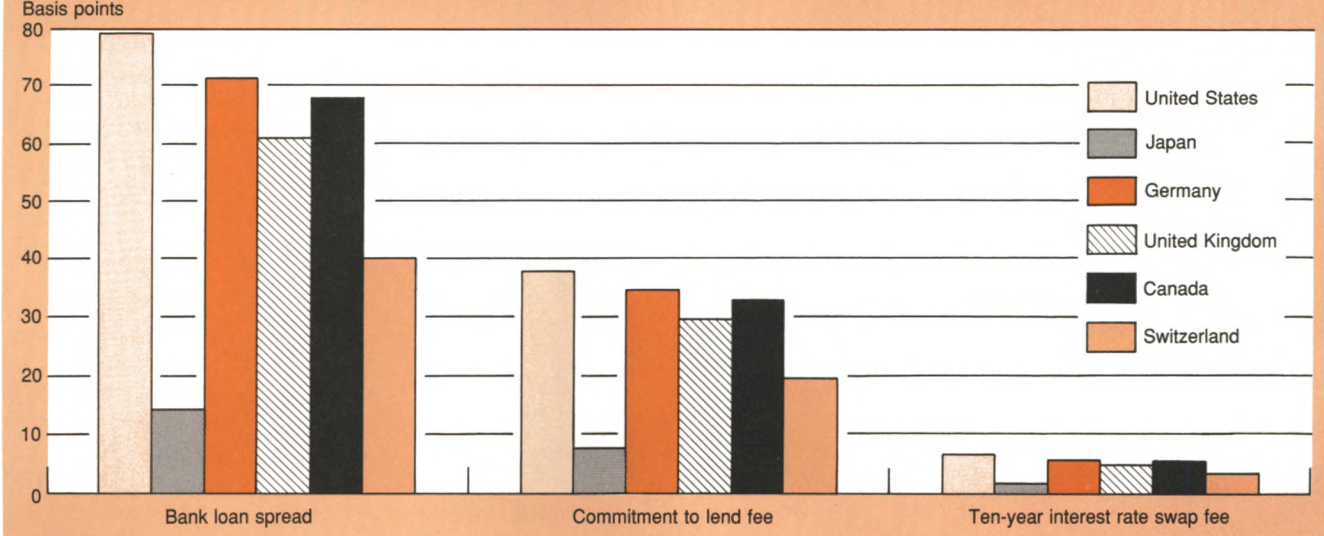
Internationally active banks generally go head to head in various markets, and such competition brings more than home country taxation into play. The cost of capital measures given in Chart 6 reflect U.S. taxes for U.S. banks, Japanese taxes for Japanese banks, and so on, and thus only apply to banks lending from their home country, whether engaged in home-court competition or in foreign lending from the head office. A multinational bank borrowing and lending in a number of markets must pay attention to a variety of tax codes. The interaction of tax codes, it turns out, tends to mitigate somewhat the benefits of hailing from a home country with low corporate taxes.

A multinational bank generally faces different effective corporate income tax rates in the different countries in which it does business. Since the cost of equity must be met after taxes, a single bank therefore faces differ-

[29]The calculation assumes that the swap is not netted. See Steven A. Zimmer, "Credit Risk in Interest Rate and Currency Swaps," Harvard University, 1988 (processed).

Chart 6

Spread or Fee Required to Cover Cost of Equity



Source: Federal Reserve Bank of New York staff estimates.

Note: Interest rate swap fee should be compared with one-half of swap spread as quoted in the market.

ent capital costs in different countries. As a result, the relation of capital costs for two banks from different countries varies across markets.

To assess the competitiveness of U.S. and German banks, for example, one cannot merely compare the cost of capital for a U.S. bank operating in the United States with the costs of a German bank operating in Germany. Consideration must also be given to the banks' capital costs in the same market, whether that market is the United States, Germany, Britain, or another country.

In calculating the effective tax rate for a bank in a foreign market we start with the corporate income tax payable at all levels of government in the foreign country. Next, for subsidiaries, we compute withholding taxes on dividends remitted, which tax treaties between pairs of countries lower from general rates of, for instance, 20 percent to 5 percent. Finally we factor in the home-country treatment of foreign taxes paid: some countries, such as the United States, give tax credits for a portion of the foreign tax paid (usually capped at comparable domestic taxes payable); other countries, such as Germany, exclude foreign source income from home-country taxation.[30] A bank may establish itself in a foreign market in the form of a branch or a subsidiary, and in general the tax rate differs by corporate form. Moreover, the taxation of a subsidiary's earnings depends on whether the subsidiary repatriates earnings or retains them.[31]

Cost of capital, or required spread, for a corporate loan thus depends not only on the cost of equity and taxes at home but also on taxes in the market served, the structure of tax treaties, and the corporate form of foreign operation (Chart 7 and Table 3). Each horizontal line in Table 3 represents a different market. For example, in the Japanese market (second line) a U.S. bank branch faces capital costs of 95 basis points as compared with 10 basis points for a Japanese bank, 55 basis points for a German branch, and 72 basis points for a Canadian branch.

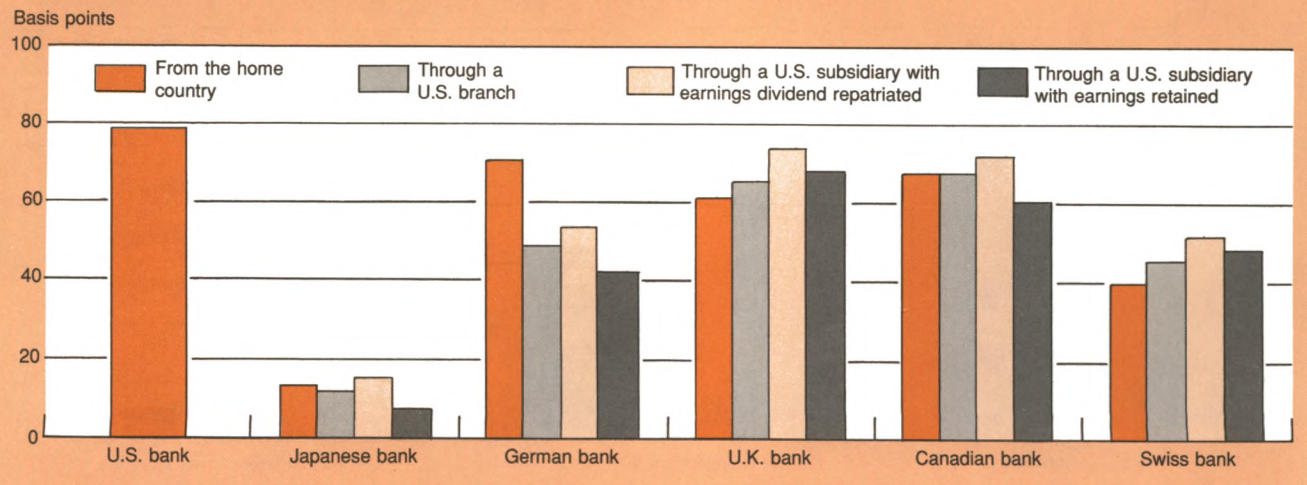
Same-market capital costs can tell quite a different story from home-market capital costs. Note that the table repeats our earlier finding that U.S. and German banks face capital costs of 76 and 65 basis points in their respective home countries. A German bank operating in the United States, however, faces capital costs of only 40 to 51 basis points; these figures bear compar-

[30]For general treatments of transnational taxation, see Julian S. Alworth, *The Finance, Investment and Taxation Decisions of Multinationals* (Oxford: Basil Blackworth, 1988); and Edmund Crooks, Michael Devereux, Mark Pearson, and Charles Wookey, "Transnational Tax Rates and Incentives to Invest," Institute for Fiscal Studies, Working Paper no. 89/9, October 1989.

[31]For instance, Bankers Trust reported in 1989 that it is saving U.S. taxes by not repatriating foreign earnings (about 25 percent of its shareholders' equity takes the form of undistributed earnings of certain foreign subsidiaries) and that federal taxes amounting to about 4 percent of shareholders' equity would have to be provided for, even after foreign tax credits, were the earnings not "permanently reinvested outside the United States" (*Annual Report, 1989*, p. 40).

Chart 7

Spread Required to Cover the Cost of Equity on a Loan to a U.S. Corporation



Source: Federal Reserve Bank of New York staff estimates.

ison with the 76-basis-point cost for a U.S bank operating at home (Chart 7). Conversely, a U.S. bank faces capital costs of 100 to 119 basis points in Germany, as compared with the 65-basis-point hurdle for a German bank operating at home.

The general result is that national differences in corporate tax rates tend to be flattened somewhat in cross-border operations by virtue of the interaction of national income tax codes and withholding taxes. In particular, banks from low-tax countries invariably face higher tax rates on their foreign operations. As a consequence, the usefulness of tax policy as a tool for reducing the U.S. disadvantage in cost of capital for banks competing abroad is limited. (See Box for related point.)

Note that the computed capital costs neither compel choice of corporate form nor exhaust relevant tax considerations. For example, British banks use both branches and subsidiaries in their U.S. operations despite the apparent advantage of branches. Further, it should be pointed out that effective tax rates do not reflect important aspects of certain tax systems. For example, the U.S. and Japanese tax systems both give tax credits for foreign income taxes paid up to the amount of domestic taxes payable on foreign income. For Japanese banks, however, the maximum credit for foreign taxes is based on worldwide income, while for U.S. banks foreign tax credits can only be applied to income earned in separate income "baskets"—a condition that restricts use of high withholding taxes levied by developing countries such as Brazil and Mexico.[32] In addition, Japanese banks can carry foreign tax credits forward and back three years, while no such provision exists for U.S. banks.

[32]A transition rule was provided for interest on loans to thirty-three troubled debtor countries. See Peat Marwick and Bank Administration Institute, *The Banker's Guide to the Tax Reform Act of 1986* (Rolling Meadows, Ill.: Bank Administration Institute, 1986), pp. 88-92.

Explaining international differences in bank cost of capital

These measures of the cost of equity are broadly in line with those of our earlier study of the cost of capital for industry in four countries.[33] The similarity suggests that the reasons adduced for differences in the cost of equity for nonfinancial firms may well carry over for banks. This section explores the relevance of the factors investigated empirically in the earlier study. In addition, it considers whether the particular risk features of banks introduce new sources of cross-country differences.

Macroeconomic and macrofinancial differences

Income tax differences do not explain international differences in bank cost of capital. Differences in household savings, macroeconomic stability, and relations among corporations, banks, and governments are consistent with observed differences in bank cost of equity.

Bank income taxes. Bank income taxes can exert a powerful influence on the cost of capital for banks, but international differences in bank income taxation offer very little help in making sense of international differences in bank cost of capital. The power of taxes is illustrated by Table 3, while the inadequacy of taxes as an explanation emerges from the broad similarity between country differences in the cost of capital (Table 3) and country differences in the cost of equity (Chart 4). In fact, if we consider average tax rates,

[33]See McCauley and Zimmer, "Explaining International Differences," *Quarterly Review*, for measures and empirical evidence supporting our explanations.

Table 3

Required Spreads on a Corporate Loan in Various International Markets

Branch / Subsidiary: Earnings Repatriated / Subsidiary: Earnings Retained

(In Basis Points)

Location of Market	Home Country of Parent Bank					
	United States	Japan	Germany	United Kingdom	Canada	Switzerland
United States	76	9 / 11 / 5	46 / 51 / 40	63 / 73 / 67	65 / 69 / 59	44 / 50 / 47
Japan	95 / 90 / 109	10	55 / 59 / 51	75 / 101 / 87	72 / 89 / 76	52 / 67 / 59
Germany	119 / 107 / 100	12 / 17 / 16	65	90 / 99 / 112	86 / 87 / 98	62 / 66 / 74
United Kingdom	68 / 76 / 76	9 / 10 / 3	43 / 36 / 36	59	65 / 65 / 51	41 / 38 / 42
Canada	82 / 95 / 102	10 / 13 / 7	58 / 56 / 44	75 / 87 / 74	65	55 / 63 / 51
Switzerland	62 / 76 / 76	9 / 10 / 2	40 / 41 / 32	59 / 59 / 52	65 / 65 / 46	38

Source: Federal Reserve Bank of New York staff estimates.
Note: Required spreads are computed on the basis of 1984-90 average cost of equity, interest rates, inflation, and 1990 tax rates.

Box: Implications of Currency Positioning of Shareholder Equity

In this Box we consider how the cost of capital for a bank's foreign affiliate depends on the net foreign exchange exposure of the affiliate. The calculations for Table 3 embody the assumption that the foreign affiliate has no foreign exchange exposure. In other words, the affiliate matches its foreign currency assets with funding in foreign currency liabilities, and the excess of assets over liabilities (shareholder equity) that supports the assets is held in the home currency. The bank invests shareholder equity in home-currency assets and therefore obtains a return on it that reflects home-country interest and inflation rates.

An argument can be made that the foreign affiliate should hold a long position in the foreign currency equal to 4 percent (required tier 1 capital ratio) of foreign assets in order to maintain the ratio of shareholder equity to assets in the face of foreign exchange movements. The drawback of such a position, of course, is that it ultimately represents a currency exposure to the parent, an additional source of volatility to earnings and shareholder wealth.

The question whether a foreign affiliate should take a long foreign exchange position in proportion to foreign assets is of more than theoretical interest. International capital standards requiring banks to hold capital against open foreign exchange positions are under negotiation. If international bank regulators decide that a long foreign exchange position in proportion to foreign currency assets represents an open foreign exchange position, then foreign affiliates will have an incentive to square foreign exchange positions and to hold shareholder equity in home-currency assets.

Since the question is still unresolved, we recalculate Table 3 on the assumption that the foreign affiliate holds a proportional long foreign position. The table below shows that the required spreads faced by many banks change because a bank that invests in a low-inflation currency tends to earn a higher *real after-tax* rate of interest† on riskless debt. Since the bank is assumed to invest shareholder equity (excess of assets over liabilities) in riskless debt, the higher real after-tax rate of return earned on the debt works to reduce the loan spread needed to cover the required return on equity. Thus, banks from high-inflation countries such as Canada can lower the cost of capital to affiliates in low-inflation countries by holding proportional shareholder equity in the low-inflation currency. Similarly, foreign affiliates of banks from low-inflation countries such as Japan and Switzerland encounter higher capital costs from holding proportional shareholder equity in high-inflation currencies.

In calculating the required spreads for the table, we implicitly assume that the parent bank defers realizations of foreign exchange gains and losses on the open position of the affiliate (typical for nonrepatriated earnings). Immediate realization would tend to shift the required spreads toward those in Table 3 in the long run. This shift arises from the tendency of a low-inflation currency to appreciate over time and thereby to increase the tax liability of the parent bank.

†The real after-tax rate of interest is roughly defined as {nominal interest rate · (1-tax rate)} – inflation rate. Since real interest rates across the countries are similar over the period considered, and since a bank's foreign affiliate faces the same tax rates in either currency, inflation is the prime determinant of the real after-tax rate of interest facing the bank.

Required Spreads on a Corporate Loan in Various International Markets: Proportional Long Position in Foreign Currency

Branch / Subsidiary: Earnings Repatriated / Subsidiary: Earnings Retained

(In Basis Points)

Location of Market	Home Country of Parent Bank					
	United States	Japan	Germany	United Kingdom	Canada	Switzerland
United States	76	21 / 25 / 15	50 / 54 / 41	59 / 66 / 62	71 / 76 / 65	26 / 33 / 30
Japan	82 / 79 / 94	10	45 / 48 / 42	60 / 75 / 66	68 / 80 / 70	21 / 34 / 28
Germany	114 / 106 / 97	24 / 30 / 28	65	83 / 85 / 94	93 / 90 / 100	37 / 44 / 50
United Kingdom	72 / 81 / 81	28 / 31 / 15	54 / 40 / 40	59	78 / 78 / 62	27 / 25 / 29
Canada	76 / 86 / 96	16 / 22 / 11	57 / 52 / 39	61 / 72 / 61	65	22 / 39 / 27
Switzerland	78 / 93 / 93	37 / 40 / 25	55 / 59 / 48	71 / 71 / 65	88 / 88 / 69	38

Source: Federal Reserve Bank of New York staff estimates.
Note: Required spreads are computed on the basis of 1990 tax rates and 1984-90 average cost of equity, interest rates, and inflation.

Japan and Germany tax banks quite heavily, while U.S. banks bear a tax burden more like that imposed by low-tax Switzerland.[34]

Household savings. Higher household savings rates in Japan, Germany, and Switzerland, perhaps reinforced by lower household access to bank credit, serve to lower equity costs for banks in these countries. However much the mobility of capital across borders has integrated debt markets, equity markets remain sufficiently distinct to allow differences in national savings behavior to assert themselves in the valuation of internationally comparable income streams.

Stability of growth. The particularly low equity costs for Japanese banks can be ascribed in part to the success of policy in smoothing growth. Japan's economy has grown markedly more steadily in the 1980s than the U.S., German, or British economies. Banks with heavy exposure to equity prices and high-leverage firms may particularly benefit from stabilization policy, which in the case of Japan works against a backdrop of adaptive corporate responses to economic challenges.

Relations between banks and corporations. Industrial organization in Germany and Japan tends to lower risk premia on the debt of industrial firms and to permit higher leverage at lower distress costs than would be entailed in the United States, Canada, or the United Kingdom. Close links between bankers and borrowers in Germany and Japan may also serve to lower capital costs for banks. The mixture of debt and equity claims on bank customers may spread risk, improve information flows, and facilitate bank influence over troubled debtor firms in ways that make German and Japanese bank shares more attractive to investors. These considerations are complementary to, but distinct from, the tax advantages of bank shareholding, especially in the presence of inflation. In addition, banks with close links to corporations may benefit from insensitivity of subordinated debt costs to bank risk.

Close links with corporations and other financial institutions served Japanese banks well in 1987-89, when the banks raised more equity than any other industry in Japan. By one count, the city banks raised the equivalent of $43 billion, including convertible issues, and each city bank in our sample raised between $3 billion and $6 billion equivalent.[35] Listings of the ten largest shareholders of, for instance, Sumitomo and Mitsubishi in March 1989 and March 1990 reveal the usefulness of reciprocal shareholdings. Over the fiscal year Sumitomo increased its outstanding shares by over 10 percent and Mitsubishi by over 7 percent. Yet the top ten shareholders remained in the same order for each bank, and indeed the stake of the top ten declined only 0.48 percent from 27.96 percent for Sumitomo and 0.72 percent from 29.00 percent for Mitsubishi.[36]

Capital constraints and financial deregulation may be straining the cross-holding pattern. Industrial firms among Sumitomo's and Mitsubishi's top ten shareholders did a bit better than the life and casualty insurers in taking up their share of new issues. Moreover, according to a news report in early April 1990, representatives of life insurers were letting it be known that they had had their fill of bank shares and could not be counted on to absorb any more.[37]

Relations among banks, corporations, and governments. Government policies to spread the costs of corporate distress beyond the immediately affected parties reduce the potential for losses by banks and their shareholders. If antitrust, trade, and industrial subsidy policies bolster distressed industries and firms in Germany and Japan more predictably than in Britain, Canada, and the United States, then investors in bank shares may face less risk and be willing to pay more for a given earnings stream.

Bank-specific factors

A puzzle arises from the comparison of the cost of equity for industry and banks in the United States, Japan, and Germany (Table 4). U.S. banks confront equity costs above those of U.S. industry, while Japanese and German banks enjoy a lower cost than those of Japanese and German industry, respectively. Why do Japanese and German banks seem to enjoy cheaper equity than their industrial customers?[38]

Risk, deposit insurance, and the cost of equity. The higher leverage characteristic of banks relative to industrial firms would seem at first glance to make banks more risky than their corporate customers. Higher leverage makes for more risky equity, and investors may be expected to demand a higher rather than a lower rate of return for bearing the extra risk. But deposit insurance

[34]Economic Advisory Committee of the American Bankers Association, *International Banking Competitiveness and Why It Matters* (Washington, D.C.: American Bankers Association, 1990), pp. 16 and 84.

[35]Zielinski and Holloway, *Unequal Equities*, pp. 184-6.

[36]Annual reports.

[37]"Major Life and Casualty Insurance Companies Selling Large Quantities of Stocks Formerly Considered 'Stable Long-Term Holdings,' Centering on Bank Stocks, in Response to Falling Markets," *Nikon Keizai Shimbun*. April 2, 1990, p. 1.

[38]Market misapprehension of the effect of inflation does not seem to resolve the puzzle. Inflation delivers unaccounted income to industry but gives a spurious boost to bank earnings. Investors' misunderstanding of the effects of inflation, therefore, would tend to raise industrial equity costs relative to bank equity costs. Yet industrial equity costs are higher relative to bank costs in the lower inflation economies of Germany and Japan.

in law and in practice allows a bank to operate at considerable risk without paying the price in the debt market that an industrial firm would pay.

Government support for bank debt funding reduces the downside risk of holding bank equity. A commercial or industrial firm that suffers losses large in relation to equity finds itself on a downward spiral as the heavier risk premia on debt put a greater burden on the firm's cash flows. A bank's debt costs rise much less quickly in response to loss of firm net worth. The value of government support for banks, in other words, rises precisely in bad times, and thereby reduces earnings volatility.

Differences in official safety nets. Differences in the strength and coverage of official safety nets may lower bank equity costs abroad in relation to industrial equity costs. As argued above, the cost of U.S. bank subordinated debt rises sharply as ratings are lowered, while foreign banks' subordinated debt costs seem much less responsive to ratings. The greater sensitivity of U.S. debt costs is understandable in light of losses that holders of U.S. bank holding company bonds have suffered in bank failures since the failure of Continental Illinois.

That the authorities in foreign countries rely less on market discipline on banks[39] finds expression not only in subordinated debt costs but also in equity costs. Lower sensitivity of subordinated debt costs to bad news itself limits earnings volatility and thereby lowers risk to equity holders. Moreover, less reliance on market discipline means that bank customers take their business elsewhere more slowly in response to bad news, so that once again bank earnings show less volatility. The recent experience of shareholders in large foreign banks has no parallel to the total loss of share value incurred by shareholders of Continental Illinois. Whether a government comes to the aid of an ailing bank itself or organizes a private sector rescue, shareholders in the rescued bank face less downside risk.

Are U.S. banks just riskier? Some observers would reject the explanations offered and contend that U.S. banks pay more for their equity because U.S. banks are simply more risky. At the extreme, this argument could posit a single international schedule relating bank cost of equity to risk, with the higher cost of equity faced by U.S. banks simply reflecting their greater risk. One could try to measure bank risk by measuring asset riskiness and leverage, but the difficulty of doing so is quite daunting. First, the relation of asset risk and leverage to bank risk is mediated by the official safety net. Second, as noted earlier, it is difficult in practice to separate the effects of asset risk and leverage from the conceptual problem of measuring cost of equity for an undercapitalized bank.

To understand the measurement problem, consider the claim that U.S. money center bank assets are more risky than their German and Japanese counterparts. This claim must be reconciled with the fact that German, Japanese, and even Swiss banks hold much larger equity positions—equities and participations—than do U.S. banks (Table 5). The decline in the Tokyo stock market in 1990 provided a reminder of how a sharp market downturn can reduce the value of bank assets and capital.

Direct measures of the market risk of bank equities do not suggest much link between the relative cost of equity and risk. Returns from U.S., Japanese, German, and British banks, at least, tend to match those of the Standard and Poor's 500, Nikkei, Commerzbank, and *Financial Times* indexes, respectively (Table 6).[40] U.S. banks' shares do seem quite a bit more risky than the

[39] Statement of E. Gerald Corrigan in *Deposit Insurance Reform and Financial Modernization,* Hearings before the Senate Banking Committee, 101st Cong. 2d sess. (Washington, D.C.: GPO, 1990) pp. 3-109.

Table 4
Comparison of Cost of Equity for Industry and Banking

(Percent)

	Period Averages		1984-88 Only	
	Industry	Banking	Industry	Banking
Germany	9.8†	6.9	7.8†	6.9
Japan	6.7	3.0	4.5	3.2
United States	10.5	11.9	11.2	12.0
United Kingdom	10.6	9.9	6.4	10.0

Source: Federal Reserve Bank of New York staff estimates.
Note: Period is 1977-88 for industry and 1984-1990 for banking.
†German industrial cost of equity includes estimated cross-holding adjustment not incorporated in McCauley and Zimmer, "Explaining International Differences," *Quarterly Review.*

[40] Use of the Tokyo Stock Price Index (TOPIX) rather than the Nikkei index did not materially affect the Japanese bank betas. Further, addition of a bond return variable did not materially affect either the U.S. or Japanese bank betas. The results for Japan are consistent with an analysis of daily returns in 1984-86 by Richard H. Pettway, T. Craig Tapley, and Takeshi Yamada, "The Impacts of Financial Deregulation upon Trading Efficiency and the Levels of Risk and Return of Japanese Banks," *Financial Review*, vol. 23 (August 1988), pp. 243-68; the results contradict the analysis of annual returns by Edward J. Kane, Haluk Unal, and Asli Demirguc-Kunt, "Capital Positions of Japanese Banks," in *Game Plans for the '90s, Proceedings of the 26th Annual Conference on Bank Structure and Competition* (Chicago: Federal Reserve Bank of Chicago, 1990), pp. 509-35. Our results for the U.S. money centers are consistent with Haluk Unal and Edward J. Kane, "Two Approaches to Assessing the Interest Rate Sensitivity of Deposit Institution Equity Returns," in Andrew Chen, ed., *Research in Finance*, vol. 7 (1988), pp. 113-37.

overall stock market only in 1990, a year in which the U.S. economy entered a recession. Indeed, U.S. bank share prices exaggerated the movement of general U.S. share prices most markedly in late 1990, as the economy contracted. By contrast, in 1986, a recession year for the Japanese economy, Japanese bank stocks seem, if anything, less risky than the market. This contrast is consistent with foreign banks' enjoying lower downside earnings volatility as a result of stronger official safety nets and, in some countries, industrial organization.

The limits of equity-market arbitrage

If equity costs are lower in Tokyo, why do firms from other countries not raise equity there? The experience of U.S. firms that have listed their shares in Tokyo shows that the mere exchange trading of shares in Tokyo does not result in share valuations different from the New York norm. Public offerings of shares in Tokyo by U.S. firms' Japanese subsidiaries or joint ventures, however, suggest that earnings streams *in Japan* as well as the earnings of Japanese firms were priced at

Table 5

Equity Shareholding by Banks in Japan, Germany, and the United States

(Percent)

	Book Value of Equity Securities Held as a Share of Bank's Book Shareholder Equity	Market Value of Equity Securities Held as a Share of Bank's Shareholder Equity Adjusted for Unrealized Gains on Equities	Market Value of Equity Securities Held as a Share of Total Assets
United States	11.7	11.7†	0.55
Japan	125.2	107.0	11.00
Germany	35.1	57.9	3.81
(including book value of equity participations)	(75.3)	(83.5)	(5.37)

Sources: Annual reports; Federal Financial Institutions Examination Council, Call Reports; International Bank Credit Analysis; *Nikkei Newsletter on Bond and Money*; Stephen Lewis, "German Banks' Ten-Month Results—A Solid Performance," *Salomon Brothers Germany Equity Research*, December 19, 1990; Federal Reserve Bank of New York staff estimates.

Notes: Data cover sample banks for Japan and Germany; data for U.S. average cover six sample banks in the second Federal Reserve district. Data for Japan are from March 1990.

†Book value of equity securities held is taken as a proxy for market value because of recent acquisition and high turnover.

Table 6

Relation of Bank Share Returns to Returns on Respective Market Indexes in Four Countries

	U.S. Money Center Banks			Japanese City Banks			German Gross Banks			U.K. Banks		
Period	Beta	Standard Error	R^2	Beta	Standard Error	R^2	Beta	Standard Error	R^2	Beta	Standard Error	R^2
1986-90	1.14†	0.067	.52	.92	.084	.32	1.03	.037	.75	1.01	.053	.57
1986	1.12	0.16	.52	.75	.15	.34	1.28†	.088	.81	1.10	.21	.35
1987	1.07	0.10	.70	1.40	.27	.35	.94	.052	.87	.97	.084	.73
1988	.92	0.13	.50	.68	.17	.25	1.20†	.093	.77	.67†	.12	.37
1989	1.16	0.20	.40	.37†	.18	.08	.89	.10	.60	1.11	.11	.66
1990	1.52†	0.23	.47	.82	.16	.39	.95	.094	.67	1.20	.13	.64

Sources: Standard and Poor's, Daiwa, Deutschebank, and *Financial Times*.

Notes: Data are weekly. Market indexes are Standard and Poor's 500 and Nikkei, Commerzbank, and Financial Times 100 indexes. Standard and Poor's money center bank index is a capitalization-weighted average of the share prices of Bankers Trust, Chase Manhattan, Chemical, Citicorp, Manufacturers, and Morgan. City bank index is a capitalization-weighted average of all thirteen city banks' share prices. Gross bank index is a capitalization-weighted average of the share prices of Commerzbank, Deutsche, and Dresdner. Financial Times bank index is a capitalization-weighted average of the share prices of the four sample banks plus Abbey National, Bank of Scotland, Royal Bank of Scotland, Standard Chartered, and Trustee Savings Bank.

†Betas are significantly different from 1 on a two-tailed test at 5 percent significance.

high multiples.[41] Whether U.S. financial firms with substantial presence in Tokyo could float equity in their Japanese subsidiaries at favorable prices is not clear.

Implications of the cost of capital disadvantage of U.S. banks

The measured cost of capital disadvantage of U.S. banks offers a simple account of U.S. banks' loss of market share at home and abroad as displayed in Chart 1. U.S. bankers would have had to offset relatively high capital costs with *much* better risk assessment or cost control to have maintained their share of corporate loans.

Bank strategies do not really offer an alternative explanation for U.S. banks' loss of market share. If U.S. bankers came to view wholesale lending as a commodity business not worth room on their balance sheets, then they would quite consistently shift their strategies toward consumer lending or reorient their corporate business to origination and risk management products. Such strategies, however, can be considered adaptations by banks burdened with high capital costs.

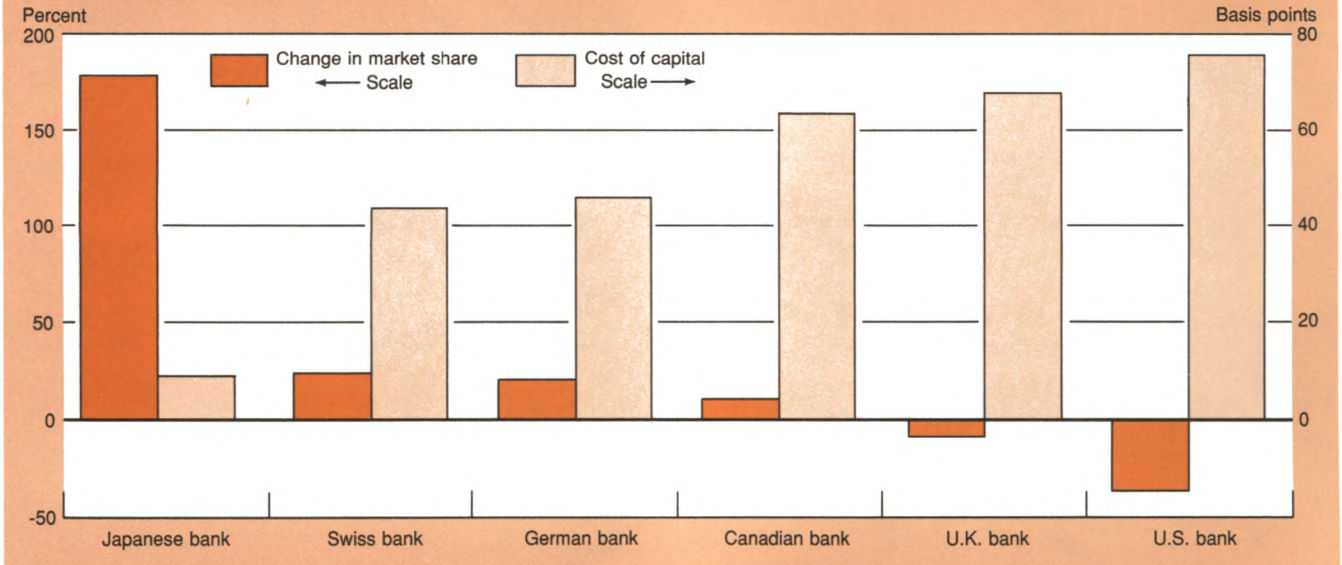
A more detailed look at commercial lending in the United States reinforces the connection between competitive outcomes and capital costs in the late 1980s. Here we consider only loans booked in the United States to commercial and industrial firms in the United States and take no notice of loans to U.S. businesses booked abroad, despite their importance, because it is not possible to decompose them by the nationality of foreign bank. The percent change in market share of sample banks from the six countries, that is, the share in March 1989 as a percentage of the share in March 1984, shows quite marked differences. Japanese banks almost tripled their share, Swiss and German banks showed substantial gains, Canadian banks slightly increased their shares, British banks lost market share somewhat, and U.S. banks suffered a 36 percent loss of market share (Chart 8). Banks with low capital costs gained market share at the expense of banks with high capital costs.

That foreign banks gained even more market share in the standby letter of credit (L/C) market than they did in commercial lending offers further evidence for the importance of capital costs in wholesale banking com-

[41] See Ted Fikre, "Equity Carve-outs in Tokyo," in this issue of the *Quarterly Review*.

Chart 8

Change in Market Share in Commercial Lending to Firms in the United States, 1984-89, and Cost of Capital for U.S. Loans



Source: Federal Financial Institutions Examination Council, Reports of Condition, as analyzed by George Budzeika, "Competitiveness in Commercial Lending in the United States," Federal Reserve Bank of New York working paper, appendix.

Note: Cost of capital bars show weighted average of the required spread for branch and subsidiary commercial loans, where weights reflect shares of commercial loans booked in branches and subsidiaries by each country's banks in 1989.

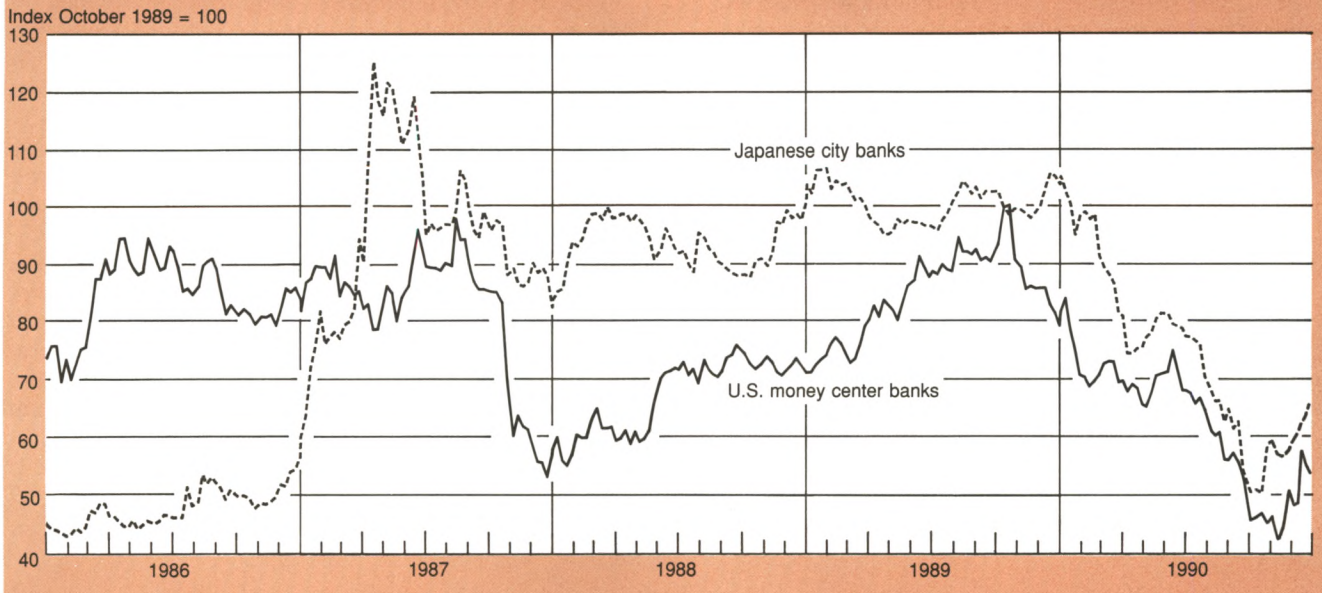
petition in the late 1980s. In the standby L/C market a bank sells a contract to pay a maturing obligation of a company or a municipality should the issuer fail to pay. Having paid for this contract, lower quality companies and state and local government agencies can raise money by selling commercial or municipal paper because buyers regard the writer of the standby L/C as the ultimate obligor. Unlike a commercial loan, a standby L/C is not funded under normal circumstances: the issuer usually retires the obligation that the bank

Table 7
Market Shares in Standby Letters of Credit

	Commercial Paper (Percent)				Industrial Revenue Bonds (Percent)	
	1985		1989		1985	1989
Sample Banks from	Number	Amount	Number	Amount	Number	Number
Canada	1	1	3	3	2	2
Germany	1	1	2	1	0	0
Japan	18	18	39	42	12	21
Switzerland	11	6	8	9	5	7
United Kingdom	11	10	7	6	13	10
United States	34	34	10	7	38	20
Memo:						
All U.S. Banks	51	57	18	14	59	43

Source: *Moody's Global Short-term Market Record*, as analyzed by Pat Wertman, "Letter of Credit Enhancement of Commercial Paper Issues: A Case Study of the Competitiveness of U.S. Banks," Federal Reserve Bank of New York working paper, January 1991.
Note: Data are for third-quarter 1985 and 1989.

Chart 9
U.S. and Japanese Bank Stock Indexes



Sources: Standard and Poor's index of share prices of Banker's Trust, Chase Manhattan, Chemical, Citicorp, First Chicago, Manufacturers Hanover, and J.P. Morgan; Daiwa Securities data on Japanese bank share prices.

Note: Japanese index is a capitalization-weighted index of share prices of twelve Japanese city banks.

has in effect guaranteed. As a result, any disadvantage faced by foreign banks in borrowing dollars in the United States, a disadvantage which may partially offset a cost of equity advantage in commercial lending, is largely irrelevant to competition in the L/C market.

In the event, foreign banks' cheaper cost of equity did find forceful expression in the standby L/C market in the 1980s (Table 7). U.S. banks in the sample wrote an estimated one-third by value of the $8 billion in identified standby L/Cs backing U.S. commercial paper in 1985 but only about 7 percent of the $19 billion in 1989.[42] Japanese banks' cost of equity advantage helped to raise their market share from 18 percent in 1985 to 42 percent in 1989. And much the same development is evident in the market for L/Cs backing industrial revenue bonds. It may be noted that some foreign banks in the sample, such as Deutsche Bank, wrote no L/Cs in 1989, although all sample banks had commercial loans outstanding then.

Should U.S. banks' loss of market share owing to a cost of capital disadvantage cause concern? Some would answer that if foreign banks can keep their shareholders happy while lending cheaply to U.S. corporations, then we should welcome the effect and not bemoan the cause. A year ago, some might have gone further to hold that foreign banks' lending to U.S. firms amounts to an insurance policy: should the capacity of U.S. banks to extend credit to the corporate sector become impaired, foreign banks could easily take up the slack. The greater the penetration of foreign banks, the more readily they could substitute for U.S. banks.

But events have called this view into question. Even as U.S. banks experience asset price declines that may dispose them to restrict corporate credit, the most prominent foreign banks operating in the United States have suffered their own asset price problems, in the form of lower equity prices in Tokyo. The extraordinary correlation of strains in the U.S. and Japanese banking systems is illustrated by the parallel movements of the share prices of money center and city banks (Chart 9). U.S. and Japanese bank share prices bore no relation to each other in the period 1987-89 but in 1990 showed strikingly high correlation.

As U.S. corporations that have borrowed from foreign banks or have backed their commercial paper with foreign bank L/Cs experience difficulty, how will foreign banks respond to private efforts and public policies designed to maintain the flow of corporate credit? The answer is important because a cost of capital disadvantage has shrunk U.S. bank market share.

[42]U.S. banks in total saw their market share shrink from 57 percent to 14 percent over the same period. See Pat Wertman, "Letter of Credit Enhancement of Commercial Paper Issues: A Case Study in the Competitiveness of U.S. Banks," Federal Reserve Bank of New York working paper, January 1991.

Conclusion

Equity markets in different countries imposed very different capital costs on banks in the late 1980s. New York, London, and Toronto burdened U.S., U.K., and Canadian banks with equity costs around 10 percent; Frankfurt and Zurich presented German and Swiss banks with equity costs in the 5 percent to 7 percent range; and Tokyo gave Japanese banks an edge with equity costs around 3 percent.

Subordinated debt costs appear to rise more quickly in response to lower ratings for U.S. than for foreign banks. For a U.S. average, higher cost of subordinated debt worsens the cost of equity disadvantage.

Taxes can exert a more powerful influence on the cost of capital for banks than on the cost of capital for industry. Nevertheless, differences in the cost of capital for banks appear to arise primarily from differences in household savings behavior and from differences in relations among banks, corporations, and governments. For Japan, the success of macroeconomic policy in smoothing economic growth may also help to cheapen bank capital.

In the wake of the Basle Agreement, cost of equity differences assert themselves as very different required spreads or fees on specific financial products. Banks facing a high cost of capital encounter substantial difficulty in competing in low-margin business lines. In the 1980s, banks with low capital costs gained market share in the U.S. wholesale market, while those with high capital costs suffered a loss of market share. Whether equity market valuations will converge and the current widening of spreads in banking persist long enough to reverse this trend remains to be seen.

Appendix A: Cost of Equity Adjustments

The first three adjustments to reported earnings are outlined in this appendix. For the depreciation adjustment, see our earlier article.†

Adjustment for developing country reserves

For British and Canadian banks, we add the following to stated profits:

(4) $R_t\ (1-tc_t) - WR_t\ (1-tc_t)$,

where

R_t = reported addition to developing country reserve
tc_t = effective corporate income tax rate
WR_t = warranted addition to developing country reserve
= (developing country exposure in or about 1985) * .14(1-.14)^n, n = year-1982.

For Japanese banks, no tax benefit is recognized for reserves, so the following is added to stated profits:

(5) $R_t - WR_t\ (1-tc_t) - (CO_t * tc_t) - NOL_t$,

where

CO_t = developing country charge-offs, including losses recognized on debt-equity swaps
NOL_t = net operating loss carry-forwards.

For U.S. banks, only a state and local tax benefit is recognized, so the following is added:

(6) $R_t\ (1-tsl_t) - WR_t\ (1-tc_t) - (CO_t * tc_t) - NOL_t$,

where tsl_t = combined state and local corporate income tax rate.

Cross-holding adjustment

For both German and Japanese banks, we add the following to stated profits:

(7) $\{[MVE_t * (ev_t - divd_t)] - CG_t\} * (1-tc_t)$,

where

MVE_t = market value of equity shares held at time t
ev_t = true profit rate on cross-held shares as calculated in McCauley and Zimmer, "Explaining International Differences."
$divd_t$ = dividend payout rate on market value of equity
CG_t = periodic realization of capital gains on cross-held shares

We calculate the market value of equity shares held by German banks as follows:

(8) $MVE_{1990} = BVE_{1990} + \dfrac{UCG_{1990}}{(1-tc_{1990})}$

where

BVE_t = value of cross-held shares as carried on books
UCG_t = value of accumulated unrealized capital gains, given on after-tax basis.

Since accumulated unrealized capital gains for German banks are available only for mid-December 1990,‡ we estimate the market value of equity for previous years as follows:

(9) $MVE_{t-1} = MVE_t * \dfrac{DAX_{t-1}}{DAX_t} - dMVE_t * \left(\dfrac{DAX_{t-1}}{DAX_t}\right)^{0.5}$,

where

DAX_t = German stock market index at time t
$dMVE_t$ = net additions to equity portfolio − $(divd_t * MVE_t)$.

Here we move backward from MVE_{1990}, iteratively estimating the earlier equity positions.

Inflation-related adjustments

To correct misstatements of profit due to differing time profiles of real and nominal returns, we make the following adjustments. First we adjust for the misstatement of flows owing to growth of assets by subtracting the following from profits:

(10) $(NOM_t - NOM_{t-1}) * [(1+\pi_t)^{0.5} - 1]$,

where

NOM_t = nominal assets less nominal liabilities at time t
π_t = inflation rate at time t.

To correct for the misstatement of profits owing to varying inflation rates, we subtract the following from stated profits:

(11) $NOM_t * \dfrac{(\pi_t - \pi^*_t)}{1+\pi_t}$,

where

$$\pi_t^* = \left[\sum_{i=t-k}^{t-1} w^{t-i} * \pi_i\right]$$

†McCauley and Zimmer, "Explaining International Differences," *Quarterly Review.*

‡Stephen Lewis, "German Banks' Ten-Month Results—A Solid Performance," *Salomon Brothers Germany Equity Research,* December 19, 1990.

Appendix A *(continued)*

where

$$\sum_{i=t-k}^{t-1} w^{t-i} = 1, \quad 1 > w > 0,$$

and k is the average life of real assets.

We then subtract from stated profits the amount that is needed to maintain the real value of shareholder equity:

$$(12)\ (SE_t - Ad_t) * \frac{\pi_t}{1+\pi_t},$$

where
SE_t = shareholder equity
Ad_t = depreciable assets.

Appendix B: Tier 2 capital and the cost of capital for financial products

The way in which cost of tier 2 capital enters into the cost of capital for a specific product varies with the source of tier 2 capital. If preferred shares are used, then the treatment is the same as tier 1 equity—the required return will have to be met out of after-tax earnings as shown for a funded asset in equation 3:

$$(13)\ 0.92 * S * (1-tc_t) + 0.08 * \{\frac{1+[r_t * (1-tc_t)]}{1 + (\pi_t * 0.5)} - 1\}$$

$$= (RW * 0.04) * (COE + CPS),$$

where CPS = cost of preferred shares.

Note that inflation is halved to reflect the fact that preference shares typically carry a nominal coupon.

The corresponding equation for an unfunded asset is similar except that the required fee is multiplied by 1 instead of 0.92. If the bank's marginal source of tier 2 equity is subordinated debt, then we can rewrite equation 13 as:

$$(14)\ 0.96 * S * (1-tc_t) + 0.04 * \{\frac{1+[r_t * (1-tc_t)]}{1 + \pi_t} - 1\}$$

$$= (RW * 0.04) * \{COE + [CSD * (1—tc_t)]\},$$

where CSD = cost of subordinated debt.

If the bank's marginal source of tier 2 funds is reserves, then the treatment is less clear. If the bank voluntarily takes reserves from shareholder equity, without an associated tax benefit, then the cost of tier 2 capital is the same as that of tier 1 capital. If there is an associated tax benefit, then the cost of capital can be written:

$$(15)\ 0.92 * S * (1-tc_t) + 0.08 * \{\frac{1+[r_t * (1-tc_t)]}{1 + \pi_t} - 1\}$$

$$= (RW * 0.04) * \{COE * (2—tc_t)\}.$$

The inclusion of tier 2 capital raises the calculated product-specific cost of capital. To the extent that banks use preferred shares or voluntary general loan reserves as tier 2 capital, the tier 2 capital costs will be a significant fraction of tier 1 capital costs. The cost of subordinated debt is low for banks in most countries—with the possible exception of the United States—and is unlikely to change the results significantly.

Our observation that bank cost of capital is highly sensitive to corporate income tax rates is not true for certain tier 2 capital. Subordinated debt and reserves with tax benefits are both tax deductible and consequently paid out of pretax income; the cost of these items is therefore insensitive to tax rates. Since tier 2 capital costs are likely to be much smaller than tier 1 costs, it will generally be the case that the overall product-specific cost of capital is still very sensitive to corporate tax rates.

Appendix C: Implications of Risk for Required Returns

If bank cost of capital is positively related to risk, then the required return on an individual product should be adjusted to reflect its effect on the overall risk of the bank. Specifically, we should determine the required return as:

$$(16)\quad S + \left\{ \frac{\text{dCOE}}{dloan} * \frac{E}{1-tc_t} \right\} + \frac{\text{dCSD}}{dloan} * SD.$$

The first term is the required spread or fee as calculated with equation 5. The second term is simply the change in the average cost of equity of the bank in response to the addition of the product and the capital allotted to it, multiplied by the value of outstanding equity. The third term is the sensitivity of the cost of subordinated debt to an additional unit of the loan, multiplied by the amount of outstanding subordinated debt.

If the new product is particularly risky relative to the capital allotted to it, then the third term will be positive and the second term will probably be positive; if the product is particularly safe, then the third term will be negative and the second term will probably be negative. Our earlier calculations implicitly apply to products with average risk relative to allotted capital in the sense that the second and third terms are zero.

Two important implications follow from the assumption that bank cost of capital is sensitive to risk. First, banks facing a high cost of capital cannot mitigate their disadvantage by concentrating on products that have low regulatory capital weight relative to their risk. Second, it may be economical for some banks to carry excess capital.

In Brief
Economic Capsules

Equity Carve-outs in Tokyo
by Ted Fikre

The gap between price-earnings ratios of Japanese and U.S. stocks in the late 1980s has puzzled market analysts and other observers. Since 1987, the value assigned to a stream of corporate earnings has been two to four times higher in Tokyo than in New York. Although various accounting and economic factors have been cited as explanations for this difference, little effort has been made to determine which characteristics of "Japanese" equity attract pricing at such a high multiple of earnings.

Some insight into this problem, however, can be gained by examining the records of initial public offerings in Tokyo of shares of U.S. subsidiaries and joint ventures between Japanese and foreign firms. These so-called carve-outs reveal how Japanese investors valued the stock of a company operating in Japan but owned, at least to some degree, by foreigners. The pricing of seven such public offerings suggests that the Tokyo market assigned a high pricing multiple to earnings *generated in Japan*. All seven of these Tokyo carve-outs, which raised an aggregate of $0.7 billion equivalent over four years, received a price-earnings multiple significantly higher than the contemporaneous price-earnings ratio of their U.S. parents. Although any conclusion based on a very small sample must remain somewhat tentative, the puzzle of Japanese stock prices in the late 1980s gives these cases a particular claim on our attention.

Background

Tokyo stocks had higher price-earnings multiples than New York stocks throughout the 1980s, but pricing in Tokyo yielded a particularly large cost of equity advantage after 1986 (see chart). While price-earnings multiples for U.S. stock indexes ranged between 10 and 20 during the latter half of the decade, comparable Japanese multiples skyrocketed from around 25 at the end of 1985 to over 60 in 1987 and 1988. According to the Morgan Stanley Capital International indexes, the current level of price-earnings multiples in Japan, around 31 at the end of December 1990, remains well above the U.S. level of about 14, although adjustments for cross-shareholdings bring the price-earnings multiples significantly closer. Other broad indexes show similar results.[1]

Given the divergence of U.S. and Japanese price-earnings ratios in recent years, it is no surprise that the number of U.S. firms with listings on the Tokyo Stock Exchange grew from fifteen at the end of 1985 to seventy-two at the end of 1990. But if large, well-known U.S. corporations obtained listings in Tokyo in order to raise their price-earnings ratios, they were disappointed. A random sampling of twenty U.S. companies listed on the Tokyo Stock Exchange yielded an average price-earnings ratio of 12 as of December 31, 1990, roughly equivalent to aggregate U.S. levels. It is true that some of the benefits of listing shares in Tokyo, such as greater investor access and heightened prestige,

[1]On December 31, 1990, the price-earnings ratio for the Standard and Poor's 500 was 15.2, while the corresponding ratio for the Tokyo Stock Price Index was 38.3. Among narrower indexes, the Dow Jones Industrial Average showed a price-earnings ratio of 13.6 and the Nikkei 225, a ratio of 38.5.

may not be fully captured in price-earnings multiples. But merely listing one's stock in Tokyo does not suffice to achieve a higher capitalization for a given stream of earnings.

Japanese subsidiaries of U.S. companies may be in a better position to capitalize on the relative strength of Tokyo equity prices because their earnings, unlike those of U.S. companies listing shares on the Tokyo Stock Exchange, are generated in Japan. Several U.S. companies have in fact attempted to exploit the lower cost of equity in Japan by floating stock in their Japanese subsidiaries on the Tokyo market. In some cases the subsidiaries were wholly owned by the U.S. parent, and in others they were jointly owned by a Japanese and a U.S. company.[2]

Carve-outs of subsidiaries wholly owned by U.S. companies

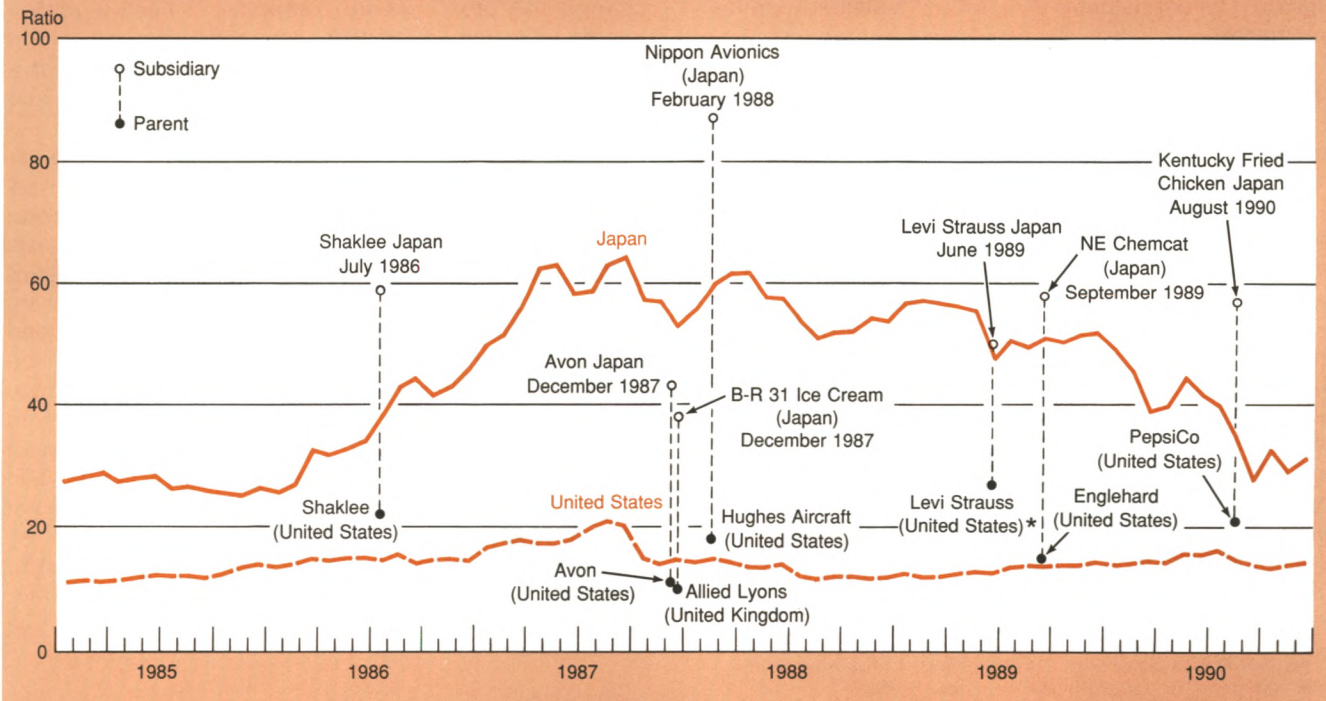
The first U.S. corporation to carve out and publicly offer equity in its Japanese subsidiary was Shaklee Corporation, a U.S. company involved in the direct marketing of food products. In July 1986 Shaklee Japan K.K. issued over 3 million shares on the Tokyo over-the-counter market at a price of $25.38 (¥3,900). These shares, representing a 22 percent stake in the Shaklee subsidiary, were valued at a price-earnings multiple of 59, quite a contrast to the multiple of 22 assigned by investors on the New York Stock Exchange to the parent company's stock.[3] The Shaklee deal, underwritten by Goldman Sachs and Nikko Securities, marked the first time that a non-Japanese investment bank took the lead in a Japanese issue.

More than a year passed before another U.S. company, Avon Products, took its Japanese subsidiary public in Tokyo. At the end of 1987, Avon floated a 40 percent stake in its Japanese unit on the over-the-counter market in Tokyo. The total value of the deal, underwritten by Morgan Stanley, was $218 million, making it the largest new issue in the history of the Jap-

[2]One of the Tokyo carve-outs described below involved a U.K.-Japanese joint venture.

[3]In February 1989 Shaklee sold its remaining 78 percent stake in the Japanese subsidiary to Yamanouchi Pharmaceutical for about $350 million, roughly 28 times earnings.

Price-Earnings Multiples of Equity Carve-outs Relative to U.S. and Japanese Markets



Note: Price-earnings ratios for Japan and the United States are taken from Morgan Stanley Capital International; price-earnings ratios for the subsidiaries, based on initial public offering prices, are computed using one-year lagged earnings.

* The parent company price-earnings ratio used for Levi Strauss is from September 1985, when management took the company private.

anese over-the-counter market. The shares were offered in December 1987 at a price-earnings multiple of 43, far above the multiple of 11 for the parent company's stock at that time.[4]

Another American company that took its wholly owned Japanese subsidiary public in Tokyo was Levi Strauss. In June 1989 the company offered a 15 percent stake ($80.6 million) in its Levi Strauss Japan subsidiary on the Tokyo over-the-counter market. Of the 4.1 million shares floated in Japan at a price of $19.66 (¥2,830), 1.1 million were newly issued and the remaining 3 million were existing shares previously held by the parent company. Like the Shaklee issue, the deal was underwritten by Goldman Sachs and Nikko Securities.

In the case of Levi Strauss a direct comparison of price-earnings multiples for the parent company and its subsidiary is impossible because the Levi Strauss parent had been taken private by management in September 1985. Nevertheless, the price-earnings ratio of 50 that Levi Strauss Japan achieved in 1989 was well above the multiple of 27 at which its parent exited the U.S. market in 1985. The behavior of Levi Strauss, going private in the United States and then taking a subsidiary public in Tokyo several years later, has elements of a long-term arbitrage strategy, designed to capitalize on persistent pricing differentials between the two markets.

Joint venture carve-outs

In addition to the wholly owned subsidiaries of U.S. companies that issued stock in Japan, several joint ventures between Japanese companies and U.S. or U.K. companies have floated shares in Tokyo. In December 1987 Allied Lyons, a U.K.-based food and beverage company, made an initial public offering of 14 percent of B-R 31 Ice Cream on the Tokyo over-the-counter market. B-R 31 Ice Cream is a fifty-fifty joint venture between Baskin-Robbins International (Allied's wholly owned subsidiary) and Fujiya of Japan. The 1.4 million shares were offered at $29.82 (¥3,610) each, yielding a price-earnings ratio of 38, well above the multiple of 10 at which Allied Lyon's stock traded in London at the time of the issue and below Fujiya's multiple of 83.[5] The $36 million of net proceeds from the deal, underwritten by Nomura, Daiwa, Nikko, and Goldman Sachs, was split evenly between the two partners.

All of the carve-outs through 1987 were listed on the over-the-counter market, not surprisingly given the strict listing requirements of the Tokyo Stock Exchange. However, in February 1988 Nippon Avionics, a joint venture between Hughes Aircraft and NEC, floated a $78 million issue on the second section of the Tokyo Stock Exchange. At the initial offering price of $13.85 (¥1,760), the stock secured a price-earnings ratio of 87, making it the most expensive initial public offering of all the equity carve-outs in Tokyo. The Nippon Avionics issue, underwritten by Goldman Sachs, was met so enthusiastically that within a week the shares were trading at over 200 times earnings. At the time of the Nippon offering, the stock of General Motors Hughes Electronic (the U.S. parent company of Hughes Aircraft) was trading at 18 times earnings in New York while NEC's stock was trading at a multiple of 102 in Tokyo.

In September 1989, just three months before the Nikkei peaked at 38,915, another U.S.-Japanese joint venture, NE Chemcat, issued shares on the Japanese over-the-counter market. Before the issue, Chemcat was 46 percent owned by Englehard Corporation and 54 percent owned by Sumitomo Metal Mining. The $90 million flotation, equivalent to a 15.6 percent share of Chemcat, reduced the stake of each partner but did not change the proportionate ownership. The 3.9 million shares, of which 1.35 million were newly issued, sold at a price of $23.13 (¥3,373) each, yielding a price-earnings ratio of 58, in between Englehard's ratio of 15 and Sumitomo's ratio of 72 at the time.

A more recent equity carve-out is particularly impressive because of the market climate in which it was executed. On August 21, 1990, with the Tokyo Nikkei plummeting below 24,000 as the Middle East crisis unfolded, Kentucky Fried Chicken Japan (KFCJ), a joint venture between PepsiCo and Mitsubishi Corporation, persisted in a previously planned listing on the second section of the Tokyo Stock Exchange. The flotation received an extremely strong reception by Japanese investors, who bought all the shares at the maximum value of ¥8,470 ($57.71) and then boosted the price up to ¥11,000 ($74.83) during the first day of trading. The resulting price-earnings ratio of 57 for KFCJ (based upon the initial public offering price) was far above the contemporaneous price-earnings ratio of 21 of PepsiCo and even higher than the multiple of 49 at which Mitsubishi's stock was trading in Tokyo.

The role of selection bias

It can be argued that the high price-earnings ratios obtained by the carved-out subsidiaries in Tokyo, more than double the parent price-earnings ratios on average, are partly attributable to a selection bias. Accord-

[4]By the end of December, less than two weeks after the issue, the stock was down 40 percent below the offering price.

[5]The B-R 31 issue demonstrated more strength in the ensuing period than did Avon Japan. Several months after the offering, the stock was trading at ¥3,950, almost 10 percent above the initial price. Inspired by the success of this first issue, Allied-Lyons and Fujiya sold another 16 percent of B-R 31 Ice Cream on the Tokyo market several months later.

ing to this view, a company would choose to carve out a subsidiary only if it expected the subsidiary's stock to be valued more highly than its own equity. Indeed, a sampling of U.S. subsidiaries carved out in the U.S. market appears to offer some support for this interpretation: the purely domestic carve-outs, like their counterparts in Tokyo, on average tend to be priced at higher price-earnings multiples than their parents.

Nevertheless, the size and regularity of the carve-out premium is much less striking in the domestic than in the Japanese carve-outs. A sample of eight U.S. carve-outs during 1988 and 1989 yielded carve-out and parent price-earnings ratios of 15.5 and 14.1, respectively (see table). This 10 percent premium is far smaller than the 220 percent premium received by the Tokyo carve-outs. Furthermore, unlike the Tokyo carve-outs, which all had higher price-earnings ratios than their parents, three of the eight U.S. carve-outs actually had lower price-earnings ratios than their parents. These results are consistent with evidence drawn from a broader sample by Schipper and Smith.[6] So, although carve-outs are generally priced at higher multiples than are their parents, the magnitude and consistency of the premium obtained by the Tokyo subsidiaries suggests a fundamental difference between the pricing of the U.S. and Japanese carve-outs. Selection bias, in other words, does not appear to be an important determinant of the differential between U.S. and Japanese carve-outs.[7]

[6]For seventy carve-outs undertaken between 1963 and 1983, Schipper and Smith reported a median subsidiary price-earnings ratio of 21.7 and a median parent ratio of 15. Furthermore, twelve of the seventy carve-outs received price-earnings multiples below those of their parents. Katherine Schipper and Abbie Smith, "A Comparison of Equity Carve-Outs and Seasoned Equity Offerings," *Journal of Financial Economics*, vol. 15 (1986), pp. 153-86.

Average Price-Earnings Ratios for Carve-out Parents and Subsidiaries

Carve-out Type	Market	Average Price-Earnings Ratios	Number of Cases Sampled
U.S. subsidiary of U.S. company			
U.S. parent	New York	14.1	
U.S. subsidiary	New York	15.5	8
Japanese subsidiary of U.S. company			
U.S. parent	New York	20.0	
Japanese subsidiary	Tokyo	50.7	3
Joint venture between Japanese and foreign firms			
U.S. or U.K. partner	New York or London	16.0	
Japanese partner	Tokyo	76.5	
Japanese subsidiary	Tokyo	60.0	4

Conclusion

We have seen that U.S. and U.K. firms that own subsidiaries or participate in joint ventures with Japanese companies were able to float equity in those operations on the Tokyo market at pricing in line with the broader Japanese market. The price-earnings ratio differentials between Japanese subsidiaries and their U.S. parents depicted in the chart roughly match the overall market discrepancies. While a simple comparison of price-earnings ratios between Japan and the United States can be misleading, factors known to explain a portion of the apparent pricing differential between the two markets probably do not explain the huge carve-out premium. The understatement of Japanese earnings associated with extensive cross-holdings of shares and low dividend payouts is known to boost reported Japanese price-earnings ratios somewhat misleadingly relative to those in the United States. However, carved-out firms, particularly the wholly owned subsidiaries of U.S. companies, are likely to hold fewer shares than Japanese firms in general.

Beyond accounting differences, why are the equity carve-outs of Japanese subsidiaries valued so highly? The large pricing advantage achieved by these Tokyo carve-outs over their U.S. parents does not appear to stem from the tendency of corporations to select for carve-outs those subsidiaries likely to be priced at a substantially higher price-earnings multiple than their own.

One interpretation points to general economic factors, such as the potential for higher and more stable growth and less business risk, that might lead to a favorable capitalization for earnings generated in Japan.[8] But one would expect that Japanese investors,

[7]Selection bias may, however, play a somewhat larger role in the case of Japanese joint ventures. Even if the U.S. parent of a jointly owned company wished to realize a carve-out premium, the Japanese parent might be reluctant to take its subsidiary public unless it was confident that it could achieve a price-earnings ratio considered good by the standards of the Tokyo market.

Therefore, it is not surprising that the joint venture carve-outs earned larger premiums than the carve-outs wholly owned by U.S. companies. On average, the price-earnings multiples of the joint venture subsidiaries were 280 percent higher than those of their U.S. parents, whereas the wholly owned subsidiaries commanded price-earnings ratios 180 percent higher than their parent companies. In addition, although only one of the four joint ventures carved out in Tokyo was priced above its Japanese parent, three of the joint venture carve-outs were priced above the overall Tokyo market.

[8]For a fuller discussion of the explanations behind this cost of capital advantage, see Robert N. McCauley and Steven A. Zimmer, "Explaining International Differences in the Cost of Capital," Federal Reserve Bank of New York *Quarterly Review*, Summer 1989, pp. 29-43.

recognizing this potential, would bid up the parent's stock price before the carve-out took place. The parent's stock price would then represent a weighted average of the Japanese and U.S. multiples, with the weights determined by the respective revenue shares of the Japanese and U.S. components of the firm. If this were the case, carve-outs would not raise total firm value: a high price-earnings multiple for the subsidiary in Tokyo would come at the expense of a lower multiple for the parent in the United States.

However, the Tokyo carve-outs provide no evidence of such a trade-off in pricing. Despite the high price-earnings ratios attained by the carve-outs, the parent stocks did not drop in price. Various reasons, such as imperfect or incomplete information, could be offered for this anomaly, but the best explanation would appear to be a preference on the part of Japanese investors for companies with exclusively Japanese operations. In buying the stock of these carved out subsidiaries and joint ventures, Japanese investors are making a "pure play" for Japanese operations and perhaps for Japanese management as well.[9]

The evidence from the carve-outs suggests that Japanese ownership is not a prerequisite for attaining the high multiples associated with the Tokyo market in recent years, although it does not preclude the possibility that such ownership could be a factor in the pricing of equity in this market. The valuation of these carve-outs demonstrates that an earnings stream generated entirely in Japan is sufficient to attract multiples comparable to those of the broader Tokyo market.

[9]By virtue of the residual parent company ownership, significant control of these subsidiaries and joint ventures remains in the hands of foreigners. However, even if these carved-out operations are not ultimately under Japanese control, they are probably under Japanese management at some level. For example, KFCJ's current president and chief executive is Japanese.

Optimal Monetary Policy Design: Rules versus Discretion Again

by A. Steven Englander

Over the last fifteen years, the entire direction of the debate on optimal monetary policy has been reversed. Earlier literature held that an optimal monetary policy could be devised by solving a "dynamic optimization" problem. It assumed that given a set of policy objectives and a model of the economy, the optimal path of inflation and GNP could be obtained.[1] The current literature argues that it is precisely the public's recognition that policy makers engage in such optimization that leads the public to expect positive inflation in equilibrium, even when both the public and the government view any inflation as undesirable. According to this argument, the public will perceive that the authorities are willing to exploit a trade-off between inflation and output, and it will adjust its expectations accordingly.

This article reviews and evaluates this newer literature on optimal monetary policy design.[2] It identifies the key arguments of the recent studies and assesses the realism of the author's assumptions. Particular attention is given to the intuitive underpinnings of the models advanced in these studies. In addition, the article analyzes the suggested mechanisms for achieving credible policies and considers whether the models' empirical implications are borne out in practice.

The main thrust of the current academic literature is to explain how a positive inflation rate can emerge on average, even when all parties view this as an inferior outcome that produces no extra output. The phenomenon that the authors are trying to explain is readily apparent: in the postwar period, inflation has averaged above zero in all countries belonging to the Organization for Economic Cooperation and Development (OECD), but few policy makers or economists believe that these inflation rates have contributed to economic well-being. The persistence of inflation at above desired levels in most OECD countries has led analysts to propose economic models yielding inflation as an equilibrium phenomenon explained by optimizing behavior on the part of the public and policy makers rather than happenstance.

[1] For example, J.H. Kalchbrenner and Peter A. Tinsley, "On the Use of Feedback Control in the Design of Aggregate Monetary Policy," *American Economic Review*, May 1976 (Papers and Proceedings of the 88th Annual Meeting of the American Economic Association, December 1975), pp. 349-55.

[2] Several other reviews of this literature have appeared in recent years. Nontechnical discussions can be found in Robert J. Barro, "Recent Developments in the Theory of Rules versus Discretion," *The Economic Journal*, vol. 95 (1985), Supplement, pp. 23-37; Alan S. Blinder, "The Rules-versus-Discretion Debate in the Light of Recent Experience," *Welwirtschafliches Archiv*, Band 123, Heft 3 (1987), pp. 399-414; and Stanley Fischer, "Rules Versus Discretion in Monetary Policy," National Bureau of Economic Research, Working Paper no. 2518, 1987, forthcoming in Benjamin Friedman and Frank Hahn, eds., *Handbook of Monetary Economics*. Somewhat more technical but largely readable are the surveys in Keith Blackburn and Michael Christensen, "Monetary Policy and Policy Credibility: Theories and Evidence," *Journal of Economic Literature*, March 1989, pp. 1-45; Torsten Persson, "Credibility of

Footnote 2 continued
Macroeconomic Policy: An Introduction and a Broad Survey," *European Economic Review*, vol. 32 (1988), pp. 519-32; Bennett T. McCallum, "Credibility and Monetary Policy," in *Price Stability and Public Policy*, pp. 105-28, Federal Reserve Bank of Kansas City, 1984; and Kenneth Rogoff, "Reputational Constraints on Monetary Policy," in Karl Brunner and Allan Meltzer, eds., *Bubbles and Other Essays*, Carnegie-Rochester Conference Series, no. 26, 1987. Of these authors, Barro and McCallum are most sympathetic to the policy thrust of the literature, and Blinder the least.

The proposed explanation is that the public views policy making as opportunistic: policy makers are willing to exploit a short-run inflation/output trade-off even if a long run trade-off neither exists nor is thought to exist. This explanation also yields a strong policy conclusion. If the public expects positive inflation because it believes that policy makers are trying to exploit this trade-off, the key to lowering actual and expected inflation is to guarantee that no such exploitation will occur. The mechanism by which this can be accomplished is to propose a readily visible rule that eliminates policy makers' discretion to inflate opportunistically. In large part this literature argues that the mere ability of policy makers to use discretion, even if the discretion is not actually exercised, will lead the public to expect positive inflation. Hence, the new monetary policy literature examines the old question of "rules versus discretion" from a new perspective.

The argument that the *structure* of the monetary policy making process (that is, the presence of discretion) rather than the *conduct* of monetary policy is the source of inflationary bias also points, as the authors see it, to the solution. Changing the structure of policy making to one guided by formal rules, they contend, might yield lower inflation on average with a relatively small loss of output. Implicit in such a recommendation is the assumption that preventing policy makers from responding to shocks or disturbances will yield only small losses. Advocates of such rules generally argue that feedback mechanisms can be incorporated in the rules to offset shocks and that the magnitude of such shocks might be lower if a consistent noninflationary policy rule were installed.[3]

An alternative view attributes the prevalence of inflation in recent decades to a combination of mistaken policies and adverse shocks, subsequently compounded by the unwillingness of policy makers to accept the output costs of disinflation through much of the 1970s. In this view, changing the structure of policy making would not accomplish much if the public and policy makers were unwilling to accept the costs of policies aimed at lowering inflation.

Such considerations have a direct connection with the issues surrounding the use of "intermediate targets" for monetary policy. It can be argued that if policy makers do not have a reputation for maintaining low inflation, they may find it necessary to pursue an intermediate target rule that can be monitored easily and on a timely basis by the public. This course may involve some loss of output or inflation control if the intermediate target is imperfectly linked to the final objectives. Nevertheless, the visible pursuit of a nominal intermediate target may provide sufficient offsetting benefits in the form of improved credibility and lowered inflation expectations to offset the imperfect linkage. In one sense, the rules-versus-discretion question involves comparing the losses from the imperfect linkages of intermediate targets to final objectives under a rules mechanism with the losses due to the inflation bias alleged to arise from discretion.

The focus of this article, however, is the interaction between the policy makers' goals and the public's expectations and behavior in response to these goals. Thus, intermediate targets will be discussed again only as a potential means for improving credibility. Much of the discussion below will assume that policy can successfully hit not only intermediate targets but also ultimate goal variables, such as inflation or nominal income growth. More specifically, the discussion will assume that policy makers can achieve their long-run inflation target *and* hit an output target temporarily by exploiting a short-term inflation/output trade-off. Over the long term it is assumed that output growth is at trend and is independent of policy.[4]

Recent literature has also examined the question of optimal monetary policy under conditions of considerable uncertainty about the structure of the economy and the policy makers' ability to hit targets on a period-by-period basis. In the face of such uncertainty, some results are weakened because the public, as might be expected, finds it more difficult to distinguish policy moves from random shocks—and to distinguish policy makers who are inflation prone from those who are not. Uncertainty about the structure of the economy also generally makes strict adherence to rules undesirable because it is difficult to design rules suitable under a broad range of conditions; in forming inflation expectations during periods of uncertainty, the public will usually place more weight on the policy makers' past inflation record.

After examining the policy implications of the recent literature, this article concludes that the policy relevance of this literature has been overstated. The the-

[3]Most of the literature is theoretical and does not make any effort to calculate the benefits or losses of adhering to a rule. Bennett T. McCallum, however, has written a series of articles proposing a specific feedback rule and attempting to estimate the losses from it. See, for example, McCallum, "The Case For Rules in the Conduct of Monetary Policy: A Concrete Example," Federal Reserve Bank of Richmond *Economic Review*, September-October 1987, pp. 10-17.

[4]This assumption can be identified with the Lucas supply curve, which is common in the literature. See Robert E. Lucas, Jr., "Expectations and the Neutrality of Money," *Journal of Economic Theory*, vol. 4, no. 2 (April 1972), pp. 103-24; and Robert J. Barro and David Gordon, "A Positive Theory of Monetary Policy in a Natural Rate Model," *Journal of Political Economy*, vol. 91, no. 4 (July 1983), pp. 589-610.

oretical insights emerging from this literature differ little from those of the earlier literature and are achieved only at the cost of analytical assumptions that are difficult to sustain empirically. Moreover, outside of a few extraordinary episodes, it is very difficult to find any concrete illustrations of the recent literature's key policy prediction that a credible disinflation can be relatively costless.

Some terminology

The recent literature on optimal monetary policy is difficult for nonspecialists to read, in part because the terminology is difficult. This section reviews the terminology and puts it in the context of the issues to be discussed in greater detail subsequently.[5]

In common language, a "consistent" policy is one that follows a well-defined set of rules over time. It would normally be viewed as superior to an "inconsistent" policy. Because the new literature on monetary policy emerged out of the earlier optimal control literature, the common usage has been altered. A *time consistent* policy is one that results from solving a long-term dynamic optimization problem *without* incorporating the effect of current policy actions on the public's expectations of the future.[6] The "consistency" that emerges in solving such problems is that the optimal policy in all future periods conforms to the policy determined in the initial period, provided that there are no unexpected occurrences or shocks to the economy. Put yet another way, in the absence of shocks, the optimal policy path laid out in time period 0 continues to appear optimal in period 1, period 2, and so on.[7] In no future period do policy makers have any reason to alter the policy path that they devised in period 0, again assuming that no shocks have occurred. If such shocks do occur, the time consistent policy path has the property that no currently anticipated developments would lead policy makers to contemplate changing their program in the future.

This type of consistency does not necessarily mean that the resultant policy path is desirable, only that policy makers see themselves as unable to do better. Whether the outcome is desirable in fact depends on how the public formulates its expectations. The assumption made by policy makers following a time consistent policy is that the public's behavior in each period depends on past policy decisions only. If the public's expectations are rational, however, so that on average the public correctly anticipates and reacts to future policy actions, the policy makers' decisions and the public's actions may be based on different views of the impact of the policy decisions. The public may correctly (on average) anticipate future policy moves because it recognizes the incentives faced by policy makers and incorporates these expectations into its current behavior, while policy makers assume that the public's decisions are independent of their future actions. In this situation, policy makers are aware of the public's current expectations but ignorant of how those expectations respond to policy actions. In this respect, the assumption of rational expectations on the part of the public provides it with an informational advantage over the policy makers.

The equilibrium that emerges is the outcome consistent with *both* views of the public's inflation expectations; it maximizes the policy makers' objective function, contingent on the public's current expectations. It is not necessarily the best outcome by any means. The public may base its expectations on worst case assumptions, and policy makers may find that the "optimal" policy in this case has the effect of validating these assumptions.

An example may illustrate this point. Assume that the public correctly believes that policy makers wish to lower the unemployment rate as much as possible provided that inflation does not exceed some critical threshold. For policy makers, the time consistent policy is to remain expansionary as long as inflation is below this critical value. The public, knowing that this is the policy makers' rule, will quickly adjust inflation expectation to the critical level, since it recognizes that government policy will quickly bring inflation there. Hence, the time consistent outcome is that inflation expectations and actual inflation adjust upward to the critical level, leaving the authorities little room in fact to implement the expansionary policy—that is, to lower the unemployment rate below some "natural rate."

In this example, the time consistent outcome has the following properties:

[5]The terminology and literature begin with Finn E. Kydland and Edward C. Prescott, "Rules Rather Than Discretion: The Inconsistency of Optimal Plans," *Journal of Political Economy*, vol. 85, no. 3 (June 1977), pp. 473-91; and Guillermo A. Calvo, "On the Time Consistency of Optimal Policy in a Monetary Economy," *Econometrica*, vol. 46, no. 6 (November 1978), pp. 1411-28. Willem H. Buiter, "The Superiority of Contingent Rules over Fixed Rules in Models with Rational Expectations," *Economic Journal*, vol. 91, no. 363 (September 1981), pp. 647-70, discusses the relationship between the new literature on monetary policy design and the older optimal control literature.

[6]Such models are often referred to as "causal" models since behavior can be traced directly to past events. By contrast, "noncausal" models allow expectations of future events to affect current behavior.

[7]Buiter, "The Superiority of Contingent Rules," citing Kydland and Prescott, "Rules Rather Than Discretion," and R. Bellman, *Dynamic Programming* (Princeton, N.J.: Princeton University Press, 1957), states that "a sequence of policy actions is time consistent if, for each time period, the policy action in that period maximizes the objective function, taking as given all previous policy actions and private agents' decisions and as given that all future policy actions will be similarly determined.

a) Policy makers always follow their perceived optimal rule of expanding output until inflation hits a critical level.
b) The public is not fooled; it correctly predicts the policy makers' action.
c) The outcome, characterized by a rapid jump in inflation expectations to the equilibrium level (the policy makers' threshold level), is unlikely to produce the output gains sought by policy makers.
d) At the equilibrium inflation rate, policy makers have no incentive to alter their policy.

By contrast, a *time inconsistent* policy path, which may in fact represent the optimal long-term policy path, does not necessarily appear optimal to policy authorities on a period-by-period basis.[8] As a result, each period policy makers would be tempted to renounce the initial time inconsistent policy path and substitute a new one. In the example above, the time inconsistent policy is to resist the temptation to lower the unemployment rate below the natural rate, even when the public's expectation is for zero inflation and expansionary policy would appear desirable from the policy makers' viewpoint.

The distinction between time consistent and time inconsistent policies can be illustrated further in the context of game theory. Consider the policy makers' payoff matrix, which specifies the value of a given outcome under a variety of circumstances and which is assumed to be known by the public (see the table).[9] From the policy makers' viewpoint, the best option is to inflate when the public expects no inflation, thereby gaining the benefits of faster growth (outcome C).[10] The worst option is to disinflate when inflation expectations are high, thereby producing a loss of output (outcome B).

In between these extremes are equilibrium outcomes. When the policy and the public's expectations are noninflationary (outcome D), the outcome is slightly worse than when the inflation takes the public by surprise, but better than when both public expectations and policy are inflationary (outcome A).

The key point is that the public recognizes that the government has an incentive to generate inflation whether the public expects high or low inflation. Outcome A is preferred to outcome B, and outcome C is preferred to outcome D—that is, the high-inflation strategy dominates.

Payoff Matrix from the Policy Makers' Viewpoint

Public's Expectation	Policy Decision: High Inflation	Policy Decision: No Inflation
High inflation	2 (A)	1 (B)
No inflation	4 (C)	3 (D)

Note: Higher numbers represent preferred outcomes.

The time consistent outcome is A: policy and expectations match, creating an equilibrium, and the authorities can do no better given the public's expectations. Outcome D, however, is the time inconsistent equilibrium and is clearly superior to A, but this outcome may be unsustainable. Once the public's expectations are decided, policy makers can do better by inflating. The public will also recognize that if inflation expectations are low, the authorities will choose C. Hence, the public will never expect the low-inflation equilibrium because policy makers' optimizing behavior consistent with that expectation yields high inflation.

Thus, in some instances adhering to a time inconsistent policy path is superior to following a time consistent path, provided that the public can be made to believe that policy makers are sincere in their pursuit of a policy that forgoes short-run optimization. In the game theory example, the superior time inconsistent outcome D could be achieved if policy makers could guarantee that they would not try to achieve C, their true optimum. Much of the policy makers' problem consists of convincing the public of their resolve to follow the time inconsistent path, when the public realizes the temptation to reoptimize. The problem resembles that of the Prisoner's Dilemma in that the outcome without cooperation between the players (in this case, the policy makers and the public) is likely to be inferior to that with cooperation. The absence of a mechanism to guarantee the cooperative solution rules out the superior outcome when cheating promises a better result for policy makers acting on their own.

A key element of the coordination problem is that the public is assumed to arrive at its expectation of current period policy before policy makers reach their decision. If the authorities moved first, the coordination problem

[8] In *Lectures in Macroeconomics* (Cambridge: MIT Press, 1989), Olivier J. Blanchard and Stanley Fischer offer the following definition: "A policy is dynamically inconsistent when a future policy decision that forms part of an optimal plan formulated at an initial date is no longer optimal from the viewpoint of a later date, even though no relevant new information has appeared in the meantime."

[9] Keith Blackburn, "Macroeconomic Policy Evaluation and Optimal Control Theory: A Critical Review of Some Recent Developments," *Journal of Economic Surveys*, vol. 1 (1987), pp. 111-48, provides a comprehensive review of the game theoretic aspects of this literature.

[10] The policy makers' objective function will be discussed below in greater detail.

would be mitigated because there would be no opportunity to fool the public. Paradoxically, the time consistency problem would be resolved because there would be no incentive to deviate from preannounced plans. Many of the proposed solutions to the time consistency problem amount to removing "surprise" as a policy tool. In the context of the table, they amount to forcing policy makers to choose between the no-inflation equilibrium (D) and high-inflation equilibrium (A).

To resolve the coordination problems that arise if the time inconsistent policy path is superior, policy makers may wish to commit or precommit themselves to the time inconsistent policy, which they know to be superior in the long term, and renounce the possibility of reoptimization. By committing themselves to the time inconsistent policy, they may hope to convince the public that they will not inflate, even when it would be advantageous to do so. A further difficulty may arise, however. If policy makers face no sanctions for violating their commitment or if the public cannot monitor on a timely basis policy makers' commitment, any commitment may lack credibility. Both the public and the policy makers may agree that the committed policy is best, but the public will not believe that the policy makers will follow through because of the period-by-period temptation to renege.

In practice, it may often be difficult to determine whether policy makers are adhering to the precommitted policy. Targets can be missed either because of random shocks to the economy or because policy makers are reneging on their commitments. Because of this ambiguity, advocates of precommitted policies often argue that following fixed rules makes it easier for the public to observe adherence to announced policies.[11] The rules can be very simple (for example, constant money growth rules) or more complicated, but they have to be understandable, and compliance has to be readily visible.

The requirement of ready visibility may make rules with *no feedback (open loop rules)* at times superior to rules in which policy actions are contingent on actual events. The public may lose confidence in its ability to monitor adherence to a rule if the rule permits action in response to events not readily observable. For example, assume that a particular monetary aggregate deviates from its precommitted path. The central bank may claim that it is merely accommodating a money demand shock. But the public, having no way to ascertain that such a shock has occurred, may assume that the deviation represents a policy easing and may therefore adjust inflation expectations upward.

[11]See the papers in Federal Reserve Bank of New York, *Intermediate Targets and Indicators for Monetary Policy: A Critical Survey*, 1990, for extensive references to the literature on monetary policy rules.

To sum up, the long-run optimal policy may be time inconsistent if the public can understand and predict future policy responses (that is, if the public has rational expectations). It may be preferable for policy makers *not* to optimize on the basis of expectations that they view as fixed, but rather to anticipate the negative effect that such optimization will have on expectations of future policy actions. More concretely, in the monetary policy case, policy makers who are expected to take advantage of low inflation expectations in order to pursue expansionary (and inflationary) policies may find that expectations are extremely sticky at undesirable levels in subsequent periods. Recognizing this, the policy makers may wish to commit themselves to a series of policy actions that may not be optimal on a period-by-period basis, but that are consistent with low inflation expectations in the long run. To succeed in the long run, such a commitment must be credible, and credibility in turn may depend on adherence to readily visible fixed rules. Fixed rules with no feedback make it easiest for the public to observe that policy is following its precommitted path.

Is there an inflationary bias to monetary policy?

This section considers the conditions under which positive inflation may emerge as an equilibrium, even when both the public and the policy makers view the outcome as inferior to one of zero inflation.[12] It examines the circumstances under which dynamic optimization by policy makers will produce an inferior result to a policy following relatively fixed rules. After the presentation of the basic model, a critical discussion of the assumptions needed to yield the equilibrium inflation result is presented. The section concludes with possible approaches to mitigating the alleged inflationary bias of policy.

How do inflationary biases emerge?

The basic structure of the models under discussion is very simple.[13] Policy makers try to achieve inflation and output goals that are inconsistent. The desired output level is greater than could be achieved at stable infla-

[12]For convenience of exposition and in common with the rest of the literature, this article will treat zero as the inflation target. In practice, measurement problems or nominal wage and price stickiness may make a positive but low level of inflation preferable. What is essential for the analysis is that the public view policy makers as willing to inflate above the target, whether it is zero or positive.

[13]Robert J. Barro and David Gordon, "Rules, Discretion and Reputation in a Model of Monetary Policy," *Journal of Monetary Economics*, vol. 12, no. 1 (July 1983), pp. 101-21; Barro, "Recent Developments in the Theory of Rules versus Discretion"; and Blanchard and Fischer, *Lectures in Macroeconomics*, provide clear descriptions of the analytical model underlying this section.

tion.[14] Policy makers face the choice between maintaining stable inflation at an output level lower than they would otherwise try to achieve or achieving desired output levels at the cost of ever-increasing inflation. Higher inflation emerges in the second case because a positive inflation surprise is the only mechanism bywhich policy makers can increase output to desired levels. In a multiperiod context, inflation surprises would be needed each period to maintain desired output; hence, spiraling inflation would emerge.

It is assumed that policy makers are less willing to tolerate additional inflation when inflation rates are already high. For example, going from 0 to 2 percent inflation will cause policy makers some discomfort, which may be offset by the temporary output gain. Each successive increment of inflation causes additional discomfiture, until inflation ultimately reaches a point at which policy makers are unwilling to accept the higher levels, even if output can thereby be maintained above the level corresponding to the natural rate. Thus, under these assumptions, there is a strict upper limit to the inflation rate policy makers would engineer, even if the public's inflation expectations were set naively. In many cases this upper limit will also be the public's equilibrium inflation expectation since the public knows that policy makers would not intentionally raise inflation any further.

It is also assumed that the public cannot be systematically fooled or surprised by inflation. The public knows policy makers' preferences and the structure of the economy, and knows that policy makers have an incentive to try to produce surprise inflation. The public also knows the increasing discomfiture of policy makers at high inflation rates. (The public's preference function is often assumed to be the same as the policy makers'—that is, to eliminate conflicting preferences as an underlying cause of equilibrium inflation.)

The public tries to predict the inflation rate by evaluating how policy makers are likely to act. The public recognizes that if policy makers observe low inflation expectations, they will have an incentive to create surprise inflation in order to reap output gains from the surprise. But the public also knows that at sufficiently high expected inflation levels, policy makers, by their own choice, will never inflate further, even by surprise, and might even choose to disinflate because of the perceived costs of a high level of inflation.

Using this knowledge, the public forms its expectations. It will never expect inflation to be so low that policy makers will have an incentive to create surprise inflation. Nor will the public expect the government to produce an inflation rate that is so high that the government would subsequently be tempted to engineer a recession (that is, create surprise disinflation) to reduce inflation to more acceptable levels. The rational expectation is thus for an inflation rate just high enough to eliminate the incentives for policy makers to surprise inflate and low enough to remove the incentive to surprise deflate. From the viewpoint of policy makers, losses from additional surprise inflation at this inflation rate just balance the perceived benefit of the additional output. The government, facing this expectation, has no incentive to produce any surprise. This outcome is characterized by inflation that is positive in equilibrium and output that is at the natural rate (but below the government's desired level). Nothing is gained on the output side from the additional inflation, but a welfare loss is incurred because of inflation. Thus, the outcome that emerges is inferior to the one that could be attained at zero inflation.

Although the extent of the knowledge attributed to the public by these models strains credulity, many of the specific assumptions are analytically convenient without being essential. What is essential is the public's assumption that policy makers are willing to use surprise inflation as a tool to generate higher employment. It is not necessary that the public know the exact form of the policy makers' preferences or the exact structure of the economy.

Underlying assumptions

The basic ingredients creating a conflict between short- and long-term policy making are (i) irreconcilable output and inflation goals, (ii) forward-looking or rational expectations on the part of the public (but not the policy makers), and (iii) a perceived ability on the part of policy makers to "surprise" the public with unexpected inflation.

Although these assumptions seem technical in nature, assessing their realism will clarify the realism of the entire analysis. In particular, the sensitivity of the analysis and the results to changes in the assumptions will help us to evaluate the claim that the structure of policy making is the source of persistent inflation in recent times. Indeed, one of the major contentions of

[14]Economists usually attribute this to some distortion that lowers output below its potential. The most common example is the distortionary tax that lowers supplies of labor and capital. Alternatively, political considerations may lead to a greater emphasis on short-term output gains as elections approach. See Alberto Alesina, "Macroeconomic Policy in a Two-Party System as a Repeated Game," *Quarterly Journal of Economics*, vol. 102, no. 3 (1987), pp. 651-78; Alberto Alesina and Jeffrey Sachs, "Political Parties and the Business Cycle in the United States, 1948-1984," *Journal of Money, Credit, and Banking*, vol. 20 (1988), pp. 63-82; and William D. Nordhaus, "Alternative Approaches to the Political Business Cycle," *Brookings Papers on Economic Activity*, 2:1989. In theory, the government's objective function could be strictly rising with output, but this would imply a willingness to trade leisure for output that would not be consistent with utility maximization by the public.

this article is that the implications of the time consistency literature are virtually indistinguishable from those of a standard backward-looking adaptive expectations framework. The additional theoretical elegance of the time consistency models is achieved only at the cost of assumptions whose empirical robustness is dubious.

Incompatible targets. The assumption of incompatible goals is essential.[15] In particular, policy makers are assumed to strive for an unemployment rate that is inconsistent with the natural rate. By assumption, the natural rate is the only unemployment rate at which inflation is stabilized; hence, policy makers must balance approaching their targeted unemployment rate against the extra inflation generated in getting there. There is no conflict between time consistent and time inconsistent policies if policy makers have only a single goal or multiple goals that are mutually supportive. If policy makers aim only at stabilizing inflation (at zero or any other value) or at stabilizing the unemployment rate at the natural rate (that is, the rate consistent with stable inflation), the time consistent policy path produced by dynamic optimization is fully consistent with the time inconsistent policy path toward the equilibrium of zero inflation (or any desired rate). Hence, the structure of policy making is irrelevant if policy makers are perceived as pursuing only a zero inflation target or a sustainable output target. It is only when the public views policy makers as regarding favorably the prospect of trading additional inflation for additional output that the inflationary bias emerges.

The reason that time consistency problems do not emerge when the output target is the natural rate is that the public has no reason to question the willingness or ability of policy makers to achieve their inflation and output goals. Because there is no conflict among goals, there is no question of commitment or credibility and no policy trade-off to exploit.[16]

As to the policy makers' objective function, the theoretical elegance of time consistency models appears greatly oversold. Undesirably high inflation as an equilibrium is derived at the cost of assuming that policy makers pursue targets that they know to be inconsistent. Often the pursuit is justified as a necessary consequence of the political process or as a way of offsetting other output-reducing distortions in the economy. In general, however, scant attention is paid to motivating the policy makers' assumed objective function empirically or theoretically.

Rational expectations. The second requirement for time consistency problems to emerge is rational expectations by the public. That is, the public knows enough about the preferences of policy makers and the structure of the economy to forecast policy accurately on average.[17] Under rational expectations, policy makers cannot systematically fool the public and so cannot gain the extra output that is sought, even temporarily. There is an asymmetry here in that while both the public and the policy makers know the structure of the economy and the policy makers' preference, only the public optimizes on the basis of future events. Indeed, in the earliest models that developed the time consistency problem, it was explicit that the public reacted to both past and future policies, while policy makers optimized only on the basis of past events.[18] Such myopia on the part of policy makers is often attributed to their susceptibility to short-term political influences. Policy makers do not recognize that the public discerns and reacts to their incentives. If policy makers recognized that the public cannot be fooled, they would not make the effort to do so. Furthermore, in many cases, even if policy makers assumed (incorrectly) that the public had backward-looking expectations, they would nevertheless be deterred from inflating opportunistically as long as their discount rate was not too high and they viewed the public's expectations as responding reasonably promptly to actual inflation.[19] By implication, in those

[15]This assumption dates back to Kydland and Prescott, in "Rules Rather Than Discretion." It is used in Barro and Gordon, "Rules, Discretion and Reputation"; Barro and Gordon, "A Positive Theory of Monetary Policy"; Alex Cukierman and Allan H. Meltzer, "A Theory of Ambiguity, Credibility and Inflation Under Discretion and Asymmetric Information," *Econometrica*, vol. 54, no. 5 (September 1986), pp. 1099-1128; and virtually every other paper on the subject.

[16]Brian Hillier and James M. Malcomson, in "Dynamic Inconsistency, Rational Expectations and Optimal Government Policy," *Econometrica*, November 1984, pp. 1437-51, argue that the essence of the time consistency problem is that policy makers have two targets, inflation and output, but only one instrument. Surprise inflation becomes a second instrument that the policy makers are attempting to utilize.

[17]Rational expectations are not strictly required. As long as the public's behavior responds somewhat to its expectation of future policy, a time consistency problem can emerge. However, virtually all of the literature assumes rational expectations.

[18]For example, see Kydland and Prescott, "Rules Rather Than Discretion." In equilibrium, expectations are fulfilled, so both the policy makers' and the public's expectations are rational ex post.

[19]For example, the low-inflation equilibria discussed in Barro and Gordon, "Rules, Discretion and Reputation," can be interpreted as emerging because policy makers recognize that inflation expectations respond quickly to actual policies. Also see V.V. Chari and Patrick J. Kehoe, "Sustainable Plans," Federal Reserve Bank of Minneapolis, Research Department Working Paper no. 377, 1988; and Herschel I. Grossman, "Inflation and Reputation with Generic Policy Preferences," *Journal of Money, Credit, and Banking*, May 1990, pp. 165-77. In "Credible Disinflation in Closed and Open Economies," Queens University Discussion Paper no. 660, 1986, David Backus and John Driffill find that the response of expectations even with Fischer-Taylor-type overlapping wage contracts is sufficiently quick to avoid the bulk of the costs associated with time inconsistent policies.

models where high inflation equilibria emerge, policy makers believe that they can fool most of the people for a long time.

The assumption that the public holds rational expectations can also be challenged on empirical grounds. Most empirical tests of the rational expectations hypothesis reject it. In particular, inflation expectations appear to be more backward- than forward-looking and inflation "surprises" can last for a long time.[20] If such is the case, the premise that adherence to a credible policy rule will produce costless disinflation may prove to be far off the mark. In practice, policy makers may find it risky to adopt a policy path whose success depends crucially on the assumption that the public will both anticipate correctly and react immediately to the effects of future policies.

In considering the robustness of policy conclusions to be drawn from the models under review, it is important to recognize that backward-looking (for example, adaptive) expectations on the part of the public can yield many of the same results produced by rational expectations in these models. Adaptive or backward-looking expectations in a multiperiod context would not be strictly "rational," but in regimes of moderate or low inflation the results would not diverge greatly from rational expectations. As long as expectations eventually catch up to actual inflation, any systematic inflation surprise can only be transitory. During this transition, policy makers could temporarily generate higher output (a course not open to policy makers if strictly rational expectations are assumed), but long-run output growth would be unaffected as long as policy makers were unwilling to accept ever-increasing inflation. Equilibrium inflation would be higher and output temporarily higher.

With such backward-looking expectations, however, it makes no difference if policy makers are credible, and there is no conflict between the time consistent and time inconsistent solutions. From the policy makers' point of view, they are obtaining the best solution given their preferences and the structure of the economy. That is, they may feel that if inflation is running at very low levels, the short-run increase in output can justify a small, but long-run, increase in inflation. In practice, however, if inflation expectations react quickly to increases in inflation, the willingness to inflate is likely to be extremely curtailed.

The key point is that in the absence of rational expectations, policy makers, perhaps reflecting the tastes of the public, have preferences that lead them to exploit the inflation/output trade-off and that make them unwilling to accept the output losses required for a return to zero inflation. In this case, it is probably better to choose better policies or better policy makers than to impose a structure of rules that may respond inflexibly, and thus suboptimally, to economic shocks or changes in priorities.

Surprise inflation as a policy tool. The final critical assumption of these models is that policy makers can generate surprise inflation to exploit an inflation/output trade-off temporarily. While this assumption is commonplace in the literature, the process by which the inflation/output trade-off is exploited in practice is not clearly described. Indeed, it seems to rest on two assumptions: 1) that anticipated policy moves (such as an expected easing in monetary policy) should have little or no effect on output, and 2) that policy makers can manipulate the surprise component of inflation to alter the path of output temporarily.

Surprise inflation is not a tool directly at the disposal of policy makers. Some other instrument—interest rates, money growth, reserve requirements—must be used to implement policy. By common consensus, however, long and variable lags separate movements in these potential instruments from changes in inflation or output. It is doubtful whether the degree of surprise experienced by the public when inflation rates change is any greater than that experienced by policy makers or whether economic behavior is greatly affected because of ignorance of the price level.[21] Hence, it is unlikely that mistaken beliefs about the level of real wages or relative prices can generate significant output fluctuations.

How surprise inflation affects aggregate output may appear to be an arcane question. But it can help us to determine whether the structure of policy making is the key factor inducing persistently high inflation expectations. If policies that have been previously announced, or for some other reason are already expected, nevertheless can have an effect on real output, the structure of the problem assumed in the time consistency literature is altered fundamentally.[22] The reason is that policy makers can achieve output gains, at least in the

[20]See A. Steven Englander and Gary Stone, "Inflation Expectations Surveys as Predictors of Inflation and Behavior in Financial and Labor Markets," Federal Reserve Bank of New York *Quarterly Review*, Autumn 1989.

[21]As discussed below, Barro and Gordon, "A Positive Theory of Monetary Policy," and Finn E. Kydland, "Monetary Policy in Models with Capital," in Frederick van der Ploeg and Aart de Zeeuw, eds., *Dynamic Policy Games in Economics*, pp. 267-87 (Amsterdam: North Holland, 1989), argue that the effects of surprise inflation on nominal asset values and capital accumulation are of greater empirical significance than the effects of wage or relative price surprises on output.

[22]See Frederic S. Mishkin, *A Rational Expectations Approach to Macroeconometrics* (Chicago: University of Chicago Press, 1983), for example. Both his original work and his reworking of Robert J. Barro and M. Rush, "Unanticipated Money and Economic Activity," in Stanley Fischer, ed., *Rational Expectations and Economic Policy*

short run, without resorting to policy moves that fool the public. Policy makers would optimize subject to their knowledge that unsustainable expansionary policies lead to inflation. Depending on the policy makers' objective function, they might tend to choose inflationary or noninflationary policies, but the source of the inflation would be the policy makers' actions rather than the structure of policy making or expectations conditioned on future policies.

The public might revise its inflation expectation upward when it observed expansionary policy being implemented, but it would not do so in the absence of such policy. Again, the conduct of policy making, rather than its structure, appears to be the underlying determinant of inflation.

Recognizing that ignorance of the level of prices or real wages is unlikely to produce major output effects, some analysts have argued that the effects of inflationary policy moves are seen immediately in asset values and capital accumulation decisions (but before the inflationary effects show up in actual prices). Hence, the policy surprise operates through wealth rather than inflation. The empirical consequences of such redistributions of wealth, however, are difficult to pin down. Some authors contend that inflation leads to higher output because the lower real value of government debt allows the government to engage in further spending. In contrast, others argue that price inflation may actually lead to a reduction in output by lowering the incentives to accumulate capital.[23]

While the issue appears abstract, the considerable uncertainty attending the effects of surprise asset inflation makes it unlikely that such surprises are the mechanism by which an inflation/output trade-off is consciously exploited by policy makers. Yet the structure of such models and the policy conclusions that they yield presuppose that surprise inflation is the only means by which policy can affect outcomes. If this assumption is false, it is hard to make the argument that the mere presence of discretionary policy making yields an inflationary bias. Again the time consistency problem seems less important than systematic policy errors or preferences in generating inflation.

Credibility

If the zero inflation outcome is preferable to the equilibrium outcome in the eyes of both parties, why do they not agree to maintain the preferred alternative? The time consistency literature argues that the answer to this question involves the credibility issue. The public recognizes that policy makers have every incentive to assert that they will maintain low inflation, but it also recognizes that policy makers have a greater incentive to renege if the public accepts the assertion at face value. According to this view, the public in general will not believe that low inflation will be maintained unless policy makers are viewed as strong adherents of low inflation or policy makers can provide evidence that they are following a policy rule that will yield low inflation. It is in this latter context that adhering to an intermediate target path believed consistent with low inflation, for example, may reduce inflation expectations.

This is where credibility issues become important. A commitment can be credible either because policy makers have a reputation for backing their commitments or because a way of enforcing the commitment exists. Among the suggested strategies for achieving commitment are

i) requiring commitment through legislation
ii) ensuring that any breaches are obvious
iii) choosing policy makers whose sole objective is low inflation.

The mechanisms by which these proposals provide credibility are discussed below. This analysis concludes that the strategies, while possessing some attractive features, are extremely difficult to implement and may carry concomitant disadvantages that could greatly outweigh their potential benefits. Moreover, if inflation expectations are essentially backward-looking, such policies may be redundant and potentially damaging if they tie policy makers' hands unnecessarily. The discussion concludes with an analysis of a fourth consideration that may encourage commitment:

iv) The adverse consequences of a reputation for opportunism may encourage policy makers to adopt low inflation policies even in the absence of a specific policy rule.

Legislation. By mandating a specific inflation goal or an intermediate target, legislation has the appearance of eliminating discretion by policy makers and substituting prescribed behavior.[24] In this way, the authorities'

Footnote 22 continued
(Chicago: University of Chicago Press, 1980), suggest that, if anything, anticipated policy moves have more impact on output than unanticipated policy.

[23]In "Rules, Discretion and Reputation," Barro and Gordon emphasize the revenue-generating function of inflation; in "Monetary Policy in Models with Capital," Kydland emphasizes the effect of inflation on capital accumulation decisions.

[24]Legislation can be viewed as imposing a severe penalty on policy makers for pursuing inflationary policies. Mats Persson, Torsten Persson, and Lars E.O. Svensson, "Time Consistency of Monetary and Fiscal Policy," *Econometrica*, vol. 55, no. 6 (November 1987), pp. 1249-73; and Mats Persson, Torsten Persson, and Lars E.O. Svensson, "Time Consistency and Monetary Policy," *Econometrica*, vol. 55, no. 6 (November 1987), pp. 1419-31, suggest an alternative,

conduct of policy may gain credibility in the eyes of the public.

One problem with legislated solutions, however, is the difficulty of ensuring an adequate degree of flexibility. Legislation can permit deviations from the rule under certain specified circumstances, such as war or deep depression, but there may be other circumstances, more difficult to identify or foresee, that would also justify a deviation, even at the risk of higher inflation. If the set of exceptions is made too general, however, the entire legislation may lose its credibility. Moreover, if the legislation is predicated on the assumption that disinflation can be achieved costlessly, a conflict between the explicit inflation goals and implicit output targets may well emerge. The public may discount legislation that does not state explicitly whether output losses are an acceptable cost of disinflation. In much of the time consistency literature this problem is "eliminated" by the assumption that a sufficiently "credible" disinflation will be costless, but the literature offers no set of criteria by which to predict in advance whether the costs of disinflation have in fact been lowered.

A second role for legislation might be to reduce or eliminate the conflict among final goals. A definite statement that price stability is the primary goal for monetary policy and that any output target ought to be consistent with this goal on average might mitigate the time consistency problem because it might reduce any temptation to exploit the inflation/output trade-off.

Making dissonant behavior obvious—intermediate targets. A second possible way of ensuring adherence to the announced path is to remove the possibility of surprise inflation from the hands of policy makers. In practice this could be achieved by tying policy to a particular nominal aggregate. Deviations from target would, at least in theory, be readily visible and viewed as reneging on the commitment. Policy makers would be able to comply with the rule and benefit from the low inflation equilibrium. Once they deviated from the rule, the public would recognize their lack of commitment, and expectations would immediately shift upward. Faced with these two possible outcomes, policy makers would adhere to the rule.

Obviously this strategy requires that the aggregate in question be controllable and predictably related to the final objectives. If the first condition does not hold, it is impossible to determine whether deviations from target represent a willful effort by policy makers to create inflation, that is, to renege.[25] If the second condition does not hold, the credibility will be achieved at the cost of being unable to respond to shifts in the velocity of the aggregate in question. Unless an intermediate target satisfying both these criteria can be found, it will be impossible to have both credibility and control over final objectives. These trade-offs are crucial to determining the desirability of an intermediate target rule. The controllability criterion points to a narrow aggregate—if control is limited, then the observation that an intermediate target is conforming to, or deviating from, a desired path brings little information. With a narrow aggregate, however, the link to final targets may be long and uncertain, and adherence to the intermediate target may lead to shocks to the final target.[26]

By and large, there appears to be scant evidence that strict observance of an intermediate target would yield better control over final targets.[27] This raises an important practical question about the use of such intermediate targets. Would a poorly selected intermediate target itself lack credibility because the public would recognize

Footnote 24 continued
but not very practical, way of penalizing inflationary behavior. They argue that if the government is a net creditor (and bound by some restrictions on the term structure of its holdings), the reduction in the real value of its assets from inflation would provide a disincentive to inflate opportunistically.

[25] In fact, Torben M. Andersen, "Rules Versus Discretion in Monetary Policy: The Case of Asymmetric Information," *Journal of Economic Dynamic Control*, vol. 10 (1986), pp. 169-74, argues that if policy makers have better information than the public about the source of money demand shocks, they would have an incentive to dissemble even under a constant growth rate rule.

[26] See, for example, Bennett T. McCallum, "Targets, Indicators and Instruments of Monetary Policy," in William S. Haraf and Philip Cagan, eds., *Monetary Policy in a Changing Financial Environment* (Washington, D.C.: American Enterprise Institute Press, 1990); and, in the same volume, Benjamin J. Friedman, "Is the Monetary Base Related to Income in a Robust Way? A Commentary." These authors come to opposite conclusions about the suitability of the monetary base as an intermediate target. David Currie, "Macroeconomic Policy Design and Control Theory—A Failed Partnership," *Economic Journal*, vol. 95 (June 1985), pp. 285-306, provides a discussion of the ill effects of what he perceives to be a poorly chosen intermediate targeting strategy in the United Kingdom in the early 1980s. Some analysts argue that strict control of monetary aggregate growth over long periods would reduce the drift in velocity of the monetary aggregates; they claim that many of the velocity changes seen in the last generation were themselves induced by the high inflation rates of the 1970s and early 1980s. See Barro, "Recent Developments in the Theory of Rules versus Discretion"; and John J. Judd and John L. Scadding, "The Search for a Stable Money Demand Function: A Survey of the Post-1973 Literature," *Journal of Economic Literature*, September 1982, pp. 993-1023. The alternative view is that much of the shift in velocity was exogenous to inflation and caused by improved technology, which allowed much greater control by firms and individuals of assets, and by financial deregulation.

[27] An extensive survey of intermediate targets is found in Federal Reserve Bank of New York, *Intermediate Targets and Indicators*. A summary of the findings appears in Richard G. Davis' introduction to the volume and is reprinted in the summer 1990 issue of the Federal Reserve Bank of New York *Quarterly Review* under the title "Intermediate Targets and Indicators for Monetary Policy: An Introduction to the Issues."

that adherence requires compromising the final targets for long periods of time? Knowing that the relationships between intermediate and final targets are by no means tight and unchanging, the public may well discount adherence to such targets as being unsustainable, just as legislation predicated on a costless disinflation is likely to be discounted.

To get around the problem of achieving credibility under shifting relationships between intermediate and final targets, it has been proposed that there be some feedback from final targets to policy instrument settings or that final objectives themselves (inflation or nominal income growth) be targeted.[28] Various contingent rules have been proposed to increase the stability of real output. As more contingencies are built into the rules, the performance in historical simulations appears to improve, but the public may view adherence to a complicated rule as being too difficult to monitor and hence little better than discretion.

A second mechanism that has been proposed to make reneging obvious is to release the record of policy deliberations and decisions immediately after they are made. The argument is that the public could then promptly recognize the inflationary consequences of policy changes, rendering surprise inflation unfeasible. However, such proposals depend critically on the assumption that the lags between policy deliberations and their public release are used by policy makers to generate surprise inflation or disinflation. In fact, lags between monetary policy decisions and their public release are currently so short—about six weeks—that it is hard to believe that such lags could be a source of inflation surprises. Moreover, a plausible argument could be made that immediate release would be counterproductive. If immediate release of deliberations made them more subject to political pressures, inflation expectations might rise rather than fall.[29]

Choosing conservative policy makers. By choosing policy makers of impeccably noninflationary tastes, the public is relieved of the need to monitor policy makers' decisions. The literature conventionally describes these policy makers as "conservative." In other words, the public can choose policy makers who attach far greater weight to low inflation than high output and who are thus more likely to err on the low inflation side.[30] Indeed, the assumption is that they are more averse to inflation than is the public.

In an economy subject to random shocks, this approach is likely to be inferior to a policy combining discretion with output targets that are consistent with low inflation. For example, if there is a supply shock, policy makers who pursue both output and price targets will wish to distribute the shock between the two, while policy makers who focus only on inflation will allow output to take the complete shock in order to attain inflation targets. Choosing conservative policy makers is equivalent to selecting the latter. In doing so, society forgoes the flexibility embodied in the former. It is not possible to determine in advance whether the gain in credibility from choosing conservative policy makers offsets the resulting loss in flexibility. In general, the gain from flexibility is higher when policy makers use their discretion to smooth output and inflation in an economy subject to large shocks. By contrast, the gains from discretion could be small in a relatively placid economy, and strongly noninflationary policy makers might be preferable to more flexible ones in that setting.

If society prefers stability in both inflation and real output, it is preferable to allow policy makers discretion in spreading shocks between prices and output. An inflation bias would not emerge if policy makers were aiming on average at consistent inflation and output targets. Again, a trade-off between discretion and rules emerges only if the public knows that the ultimate output target is not feasible without inflation.

Reputation. Although formal rules seem most direct in constraining inflationary proclivities, the need to maintain a noninflationary reputation can be almost as effective in constraining opportunistic policy makers. If policy makers have a long time horizon and do not discount the future too heavily, they may be reluctant to exploit an inflation/output trade-off opportunistically because this will raise inflation expectations in subsequent periods. A long time horizon is necessary because it increases the period during which policy makers would be "punished" by higher inflation expec-

[28] Implicit in some of these rules is the assumption that inflation expectations and actual inflation rates will be more responsive to policy under a rule than under discretion. If factors other than the structure of policy lead to sluggish adjustment of actual and expected inflation, this presumption would not be justified. The inflation and real output growth engendered by such rules might then not be desired by either the public or the policy makers. R. Spence Hilton and Vivek Moorthy review a variety of such rules in "Targeting Nominal GNP," in Federal Reserve Bank of New York, *Intermediate Targets and Indicators*.

[29] See William Poole, "Central Control of Interest Rates: A Commentary," in Haraf and Cagan, eds., *Monetary Policy in a Changing Financial Environment*. The political business cycle literature, which treats political influences on economic policy making, is beyond the scope of this survey. For recent discussions and further references, see Nordhaus, "Alternative Approaches to the Political Business Cycle"; and Alesina and Sachs, "Political Parties and the Business Cycle in the United States."

[30] Kenneth Rogoff discusses the implication of selecting policy makers with an unusually strong aversion to inflation in "Reputational Constraints"; "The Optimal Degree of Commitment to an Intermediate Monetary Target," *Quarterly Journal of Economics*, vol. 100, no. 4 (1985), pp. 1169-90; and "Reputation, Coordination and Monetary Policy," in Robert Barro, ed., *Handbook of Modern Business Cycle Theory* (Cambridge University Press, forthcoming).

tations. The moderate time discount rate is necessary because the policy makers would otherwise put much more emphasis on short-run optimization, an approach which might lead to opportunistic behavior. Analysts have pointed to the long and overlapping terms of central bankers as a way of promoting an institutional long horizon.

The precise degree of restraint that reputational factors impose on policy makers depends in large part on how the public forms its expectations, how fast expectations respond to a change in policy, and whether, once policy has been opportunistic, expectations revert back to low inflation without a loss of output. However, the following general conclusion is robust: unless policy makers are extremely short-sighted, valuing short-term output gains very heavily, their own willingness to inflate may be greatly constrained by the prospect of a long period of high inflation and inflation expectations. Knowing that the penalties from a loss of reputation are severe, policy makers may even choose zero inflation. Indeed, in the context of these models the public may lower its inflation expectations because it knows that policy makers view these penalties as a deterrent. Hence, even where there is a willingness to behave opportunistically, discretionary time consistent optimization may not produce significantly higher inflation than time inconsistent policies aimed at zero inflation.

Reputation may be important in a different way even when policy makers do not have full credibility. Much of the literature compares results when policy makers have full credibility at zero inflation with results when there is no credibility at all—that is, when policy makers are expected to inflate to their maximum tolerable inflation level. Under such circumstances, zero inflation is not a credible result because policy makers have too much incentive to renege. However, there may be an inflation level that is above zero but below that of the no credibility level to which policy makers could make a credible commitment.[31] While the policy makers may wish to behave opportunistically, they may be deterred by the possibility that the public's inflation expectations would revert as a result to the fully noncredible level. Hence, policy makers may find it preferable to adopt policies consistent with this intermediate level of expectations rather than try to achieve additional output gains.

Such considerations may help explain why announcements of near-term zero inflation targets often carry little credibility. The public may feel that policy makers will too readily jettison the zero inflation target if there is the opportunity to obtain extra output. While this logic would appear to argue in favor of announcing more credible gradualist disinflation policies, there does not appear to be much evidence that such announcements themselves produce more credible and less painful disinflation. The reason may be that the short-run policy moves are often too small to be convincing. The public may also doubt the medium-term political sustainability of the gradualist policy if it could imply persistent restraint.

[31]Barro and Gordon, "Rules, Discretion and Reputation"; and Barro and Gordon, "A Positive Theory of Monetary Policy." See also John B. Taylor's comments on Barro and Gordon in "Rules, Discretion and Reputation in a Model of Monetary Policy: Comments," *Journal of Monetary Economics*, vol. 12, no. 1 (July 1983), pp. 123-25.

Uncertainty

Much of the previous discussion has been deterministic. The public knows with precision the aims of policy makers and the structure of the economy (including the linkages of intermediate targets to final outcomes). Loosening these assumptions of precise knowledge affects the results as intuition would suggest: in the short run, the public is less categorical in its interpretation of apparent policy moves; in the long run, it will interpret a string of positive inflation results as an indication that policy makers are willing to exploit an inflation/output trade-off.

Two types of uncertainty are discussed below. The emphasis is less on modeling than on exploring intuitively how uncertainty affects the results discussed earlier.

Uncertainty about the structure of the economy

When there are structural changes in the economy, rigid adherence to a policy rule may be less desirable even than the time consistent (discretionary) outcome. Simply stated, the benefits from allowing policy makers to offset shocks may well outweigh the losses from higher inflation expectations due to time consistency problems.[32]

In some cases the optimal strategy may be mixed: follow a rule during normal times when shocks are relatively small, but switch to discretion in the presence of large shocks.[33] The reasoning is that it is expensive for policy makers to specify behavior under important but relatively rare events, just as it is difficult for the public to specify behavior under all possible contingencies in its private contracts. In the event of a crisis, such as a war or major downturn, both the public and the policy makers are likely to view a rule as inferior to

[32]In general, it has to be assumed that only policy makers are able to recognize the shock. Otherwise the public would be able to incorporate the shock into its expectations. Buiter, "The Superiority of Contingent Rules," has a comprehensive discussion of how informational advantages may tip the scale in favor of discretion.

[33]See Robert P. Flood and Peter Isard, "Monetary Policy Strategies," *IMF Staff Papers*, vol. 36 no. 3 (September 1989), pp. 612-32.

discretion. Rather than adhere to imperfect rules or attempt to determine rules for all contingencies it may be preferable to allow discretion but require some ex post justification for invoking it.[34] According to this reasoning, however, the benefits from adhering to a rule may outweigh the benefits of discretion during normal periods, provided that an adequate rule can be formulated.

When the economy is subject to shocks, however, the public may be more inclined to expect higher inflation because it knows that in the short run policy makers could disguise policy moves by claiming that they are actually random shocks to the economy. As a consequence, policy makers may find it more difficult to acquire a reputation for noninflationary policies if they are not adhering to a verifiable rule. Over the medium term, however, discretion may remain compatible with noninflationary behavior by policy makers. For example, while uncertainty may mean that a given positive inflation shock cannot be interpreted as opportunistic behavior, negative and positive shocks should, on average, offset each other over the medium term. By considering whether an observed sequence of inflation rates is more consistent with stable inflation than with opportunistic behavior, the public may be able to establish with fair precision the true objectives of the policy makers. In one such model, policy makers maintain credibility as long as inflation remains within a certain range but lose credibility if inflation rates stray outside.[35] Again, the proof of the pudding emerges in the eating—in the presence of uncertainty, the past record of inflation performance is more useful than an imperfect proxy for policy as an indicator of policy makers' targets.

It has also been suggested that an explicit trade-off can be made between the loss of flexibility due to strict intermediate targeting and the risk that policy makers may turn out to be more opportunistic than expected. One such model argues that in an economy with weak ties between intermediate and final targets, the inflation record of policy makers with a strong noninflationary record ought to be judged over a longer period than in an economy where intermediate targets are closely tied to final goals and the policy makers' reputations are not as well established.[36] That is, policy makers of good reputation should be given more medium-term discretion when intermediate targets are unreliable indicators of the stance of policy.

Uncertainty about policy makers' preferences

A large literature analyzing how the public forms its expectations when it is uncertain of policy makers' preferences has emerged in recent years. Time consistency problems are replaced in this literature by the problem of identifying policy makers who are more (or less) willing to inflate opportunistically.[37] Once policy makers are found to be weak on inflation, they lose credibility and inflation expectations move up to the time consistent level.

This literature focuses on the incentives prompting opportunistic policy makers to look like inflation fighters. Revealing themselves to be opportunistic carries a permanent cost of higher expected inflation, so they have an incentive to look tough on inflation for some period of time. (If their time horizon is infinite, the effect may be absolute.) By and large, an incentive for opportunistic policy makers to adopt noninflationary policies emerges under a wide variety of conditions.

If there is uncertainty about the state and structure of the economy, as well as about policy makers' preferences, the advantages accruing to a noninflationary reputation diminish, however. The reason is that when the public sees an apparently inflationary outcome, it may be uncertain whether the outcome results from a policy action or from a random shock.[38] Policy makers can do little in the short term to convince the public of their noninflationary intent. Because of this ambiguity, opportunistic policy makers may inflate early because it may take some time for the public to catch on. While the formation of the public's inflation expectations would obviously be influenced by such ambiguities, the use of

[34]Flood and Isard point to the requirement in many countries that central bankers testify periodically before elected officials as an example of a mechanism that will limit abuse of discretion. As part of this testimony, the bankers are closely questioned about their policies.

[35]See Matthew B. Canzoneri, "Monetary Policy Games and the Role of Private Information," *American Economic Review*, vol. 75 (1985), pp. 1056-70. The decision rule is analogous to the rule used in quality sampling. If an unusual number of defectives emerges in a small sample of a larger lot, the entire lot is rejected. There is a finite chance that a few atypical defectives will lead to rejection of a basically good lot; similarly, there exists the chance that random shocks beyond the policy makers' control will lead to their acquiring a reputation as inflation-prone.

[36]See Michelle R. Garfinkel and Seonghwan Oh, "Strategic Discipline in Monetary Policy with Private Information: Optimal Targeting Periods," Federal Reserve Bank of St. Louis, mimeo.

[37]The seminal papers are David Backus and John Driffill, "Rational Expectations and Policy Credibility Following a Change in Regime," *Review of Economic Studies*, vol. 52, no. 2 (April 1985), 211-22; and Backus and Driffill, "Inflation and Reputation," *American Economic Review*, vol. 75, no. 3 (June 1985), pp. 530-38. See also Robert J. Barro, "Reputation in a Model of Monetary Policy with Incomplete Information," *Journal of Monetary Economics*, vol. 17, no. 1 (January 1986), pp. 3-20; and John Driffill, "Macroeconomic Policy Games with Incomplete Information: Some Extensions," in *Dynamic Policy Games in Economics*, pp. 289-323 (1989).

[38]See Driffill, "Macroeconomic Policy Games"; and Kazuo Mino and Shunichi Tsutsui, "Reputational Constraints and Signalling Effects in a Monetary Policy Game," New York University, mimeo, 1989.

a policy rule in this case has the same problem as in the situations described earlier—that is, where economic shocks are large, discretion plus noninflationary policy makers dominates rules.

Empirical evidence and conclusions

There is little firm empirical evidence on many of the issues discussed in this article. The optimal structure of monetary policy clearly depends on many factors whose importance is difficult to measure. These factors include a) the public's ability to predict policy makers' actions, b) the policy makers' goals, c) the predictability of linkages between policy tools and final goals, d) the extent of shocks to the economy, and e) the perceived credibility of policy makers. Although much of the literature has been written by authors who hold strong views on the qualitative importance of these factors, actual measurement is so difficult that theoretical analysis has been far more common than empirical work.

As a result, most of the empirical work has focused on measuring whether the output costs of disinflation respond to the perceived credibility of policy.[39] Credible disinflationary policies, supported by verifiable rules, should carry a lower output cost than less credible discretionary disinflationary policies. Yet efforts to distinguish credible from noncredible disinflations have not met with great success. Most empirical work has not found any significant decline in the output costs of disinflation either in the United States or in the rest of the OECD through the early 1980s, and, indeed, these relations appear to have been stable in most OECD countries since the 1960s.[40] This apparent stability has persisted despite the view of many that anti-inflationary policies became more "credible" in the early 1980s.

It is difficult to state with any confidence that a particular set of policies will generate a credible disinflation with low output costs. The countries that disinflated in conjunction with a "rule"—which took the form of tying their currencies to stronger currencies in the European Monetary System—generally experienced high unemployment in the process. A possible interpretation of these 1980s disinflations, one that would be in line with the time consistency literature, is that the policies put in place were not in fact credible. Thus, the public may have questioned the commitments of the policy makers to low inflation and hence may have refrained from altering behavior and expectations in response to the announced policies.

A problem with this interpretation is that it is difficult, if not impossible, to find alternative independent tests of the presence or absence of credibility. There are few examples of countries adhering to monetary targeting rules that might provide a baseline test of whether such rules produce credibility and lower the cost of disinflation. In the view of many authors, the low-inflation OECD countries do not appear to follow an explicit rule.[41] To the extent that low inflation is built into expectations in these countries, it is because of the countries' recent success in maintaining low inflation, rather than their adherence to an explicit rule.

A second possible interpretation of the 1980s experience is that disinflation is expensive because expectations are largely backward-looking and do not readily incorporate the effects of policy changes. Most studies have found this characterization to be broadly accurate—at least as it applies to labor markets—as long as inflation is low or moderate.[42] If this is so, policy makers and society have to accept the output costs if they wish to disinflate to very low inflation rates. With this backward-looking, rather than rational, view of expectations formation, the kind of time consistency problem described in the literature under discussion does not exist in reality.

Although most analyses have not found any empirically significant credibility effects, there are a few exceptions, primarily in cases of disinflating from hyperinflation. Disinflations in Central Europe in the 1920s and in Chile and Denmark more recently appear more successful, although economists still debate whether these disinflations were indeed painless.[43] What char-

[39]William Fellner, "The Credibility Effect and Rational Expectations: Implications of the Gramlich Study," *Brookings Papers on Economic Activity*, 1:1979, pp. 167-89, first suggested this approach.

[40]For the United States, see A. Steven Englander and Cornelis A. Los, "The Stability of the Phillips Curve and Its Implications for the 1980s," Federal Reserve Bank of New York, Research Paper no. 8303, 1983; Olivier J. Blanchard, "The Lucas Critique and the Volcker Deflation," *American Economic Review*, vol. 74, no. 2 (May 1984), pp. 211-15; and Robert J. Gordon and Stephen King, "The Output Cost of Disinflation in Traditional and Vector Autoregressive Models," *Brookings Papers on Economic Activity*, 1:1982, pp. 205-43. For other OECD countries, see James Chan-Lee, David T. Coe, and Menahem Prywes, "Microeconomic Changes and Macroeconomic Wage Disinflation in the 1980s," *OECD Economic Studies*, no. 8 (Spring 1987), pp. 121-57.

[41]See, for example, Michael M. Hutchinson, "Japan's 'Money Focussed' Monetary Policy," Federal Reserve Bank of San Francisco *Economic Review*, Summer 1986, pp. 33-46; and Bharat Trehan, "The Practice of Monetary Targeting: A Case Study of the West German Experience," Federal Reserve Bank of San Francisco *Economic Review*, Spring 1988, pp. 30-44.

[42]See Englander and Stone, "Inflation Expectations Surveys."

[43]See Thomas J. Sargent, "The Ends of Four Big Inflations," in Robert E. Hall, ed., *Inflation: Causes and Effects* (Chicago: University of Chicago Press, 1982); Michael Christensen, "Disinflation, Credibility and Price Inertia," *Applied Economics*, vol. 19, no. 10 (1987), pp. 1353-66; Michael Christensen, "On Interest Rate Determination, Testing for Policy Credibility and the Relevance of the Lucas Critique," *European Journal of Political Economy*, vol. 3 (1987), pp. 369-88; and Marianne Baxter, "The Role of Expectations in Stabilization Policy," *Journal of Monetary Economics*, vol. 15, no. 3 (1985), pp. 343-62. Keith Blackburn and Michael Christensen

acterizes these credible disinflations is that monetary, fiscal, and, in some cases, exchange rate policies were all subordinated to the disinflationary goal. In particular, it has been argued that fiscal tightening, which would make future monetization of government debt less tempting, was a key factor in convincing the public that the low-inflation path was sustainable.

Two other characteristics of these disinflations are noteworthy, however, and cast doubt on the relevance of these examples to the task of disinflating from moderate inflation. First, it may be easier to move from high to moderate inflation rates because both the policy makers and the public clearly desire to lower inflation. There is a high real output cost of hyperinflation in terms of time and energy spent exchanging "money" whose value drops daily into assets with more stable value. Hence, the ambiguity whether the policy objectives are in fact consistent is not as profound as at lower inflation rates. Also, hyperinflation in many cases greatly reduced the real value of government debt. As a result, fiscal policy could start de novo with little or no debt service burden. Whether disinflation from moderate to low levels of inflation can occur with so little cost is not clear. Other instances of disinflation from more moderate inflation levels have generally resulted in substantial output costs.[44] Indeed, even the cases of successful disinflation from hyperinflation involved some apparent output cost. What makes them seem painless is the low output cost per percentage point of inflation reduction.

Second, the key reform in each of these cases was generally not an explicit attachment to a monetary policy rule but rather the creation of a set of mutually consistent monetary and fiscal policies. The consistency of policies may also have served to convince the public that lower inflation was the preeminent goal. Moreover, in several cases of disinflation with relatively small output costs, a coordinated structure of wage bargaining may have been important in unwinding a wage/price spiral.[45] Although credibility may have been important, these considerations suggest that it is not rules per se that create credibility but policies that will lead to disinflation irrespective of the underlying economic model.

A final consideration is that, in practice, policy may be more credible in one market than another. In the case of Ireland's disinflation in the 1980s, it has been argued that a reduction in long-term interest rates reflected a policy credibility in financial markets that did not exist in labor markets, as reflected in the sharp rise in unemployment rates.[46] As long as labor market expectations are slow to adjust, it is unlikely that the output cost of disinflation can be eliminated.

Footnote 43 continued
provide a concise survey of this literature in "Monetary Policy and Policy Credibility: Theories and Evidence," *Journal of Economic Literature* (March 1989), pp. 1-45.

[44]Robert J. Gordon reviews several such instances of disinflation in the United States and abroad in "Why Stopping Inflation May be Costly: Evidence from Fourteen Historical Episodes in Inflation," in Robert E. Hall, ed., *Inflation: Causes and Effects* (Chicago: University of Chicago Press, 1982).

[45]Robert J. Gordon, "Why Stopping Inflation May be Costly," argues strongly for this interpretation in several OECD disinflations in the 1960s and 1970s.

[46]See Jeroen J.M. Kremers, "Gaining Policy Credibility for a Disinflation," *IMF Staff Papers*, vol. 37, no. 1 (March 1990); and Rudiger Dornbush, "Credibility, Debt and Unemployment: Ireland's Failed Stabilization," *Economic Policy*, no. 8 (April 1989). Christensen, "Disinflation, Credibility and Price Inertia," also provides evidence on the sluggishness of price expectations in labor markets.

Treasury and Federal Reserve Foreign Exchange Operations

November 1990—January 1991

The dollar was subjected to conflicting forces during the November-January period. Sentiment toward dollar investments continued to deteriorate as the U.S. economy weakened and as interest rate differentials moved further in favor of foreign currencies. But at times, political developments abroad—particularly the Persian Gulf conflict—encouraged greater demand for dollars and limited the extent to which negative sentiment toward the currency was reflected in exchange rates. With these offsetting factors helping to maintain a sense of two-way risk to dollar exchange rates, the dollar ended the period mixed against major foreign currencies, and the U.S. monetary authorities conducted no intervention operations in the foreign exchange market. The dollar closed the period down slightly against the German mark and up slightly against the Japanese yen (Chart 1). On a trade-weighted basis as measured by the staff of the Federal Reserve Board of Governors, the dollar ended the period 1 percent below its level at the beginning of the period.

The first part of the period: early to mid-November

In the early part of the period, market attention centered on evidence of diverging growth and interest rate trends in the major industrial economies. Ever since the Iraqi invasion of Kuwait in August and the associated rise in oil prices and decline in consumer confidence, analysts had been progressively revising downward their forecasts for U.S. economic growth. The release of October payroll employment data in the first week of November revealed an unexpectedly large drop which, together with subsequent data, reinforced the view that the U.S. economy was slowing down (Chart 2). At the same time, preliminary indications suggested inflationary pressures were subsiding. Under these circumstances, market participants widely expected the Federal Reserve to continue to ease money market conditions and possibly to reduce its discount rate.

In contrast, market forecasts for the German and Japanese economies remained relatively upbeat. The need to rebuild eastern Germany was seen as providing ongoing stimulus to the German economy. Japanese economic data provided little evidence that the economy or price pressures were slowing in response to the central bank's tight policy stance. Mindful of these economic trends, market participants expected German and Japanese interest rates either to rise further or to remain at existing levels. Indeed, on the first day of the period, the Bundesbank announced a one-half percentage point increase in its official Lombard rate, and many market participants expected further tightening after German national elections in early December. The Bank of Japan was considered less likely than the Bundesbank to tighten monetary policy, but was nonetheless seen as unwilling to ease monetary conditions given high oil prices and Japan's tight labor market conditions.

The divergent outlook for interest rates weighed on the dollar in early to mid-November. Short-term interest rate differentials had been steadily moving against the dollar since the spring of 1989, when dollar investments

A report presented by Sam Y. Cross, Executive Vice President in charge of the Foreign Group at the Federal Reserve Bank of New York and Manager of Foreign Operations for the System Open Market Account. Daniel H. Brotman was primarily responsible for preparation of the report.

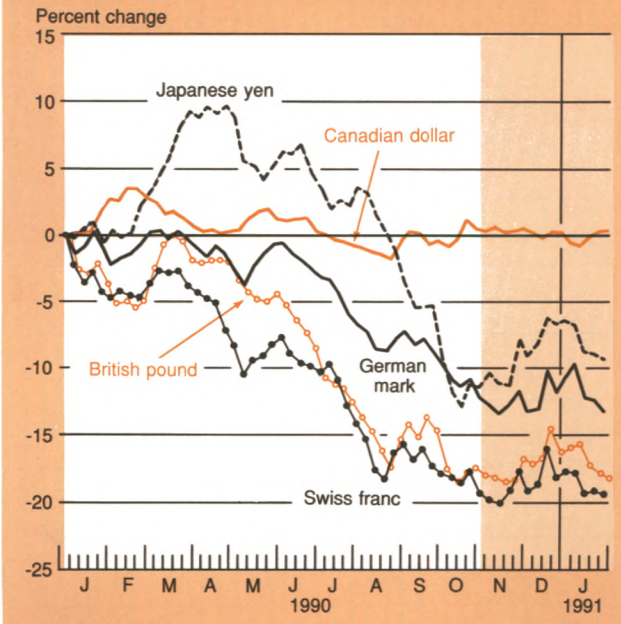
held a 4 to 6 percentage point interest rate advantage relative to the mark and yen (Chart 3). By late summer of 1990, the dollar's short-term interest rate advantage had been entirely eliminated. Thus, in early November, expected further declines in dollar interest rates, coupled with steady to higher rates abroad, threatened to push short-term U.S. interest rates well below mark and yen rates for the first time since 1980. Under these circumstances, the dollar declined 3¼ percent against the mark from its opening level of DM1.5170 to its November low of DM1.4660 on November 16. Its decline against the yen measured 2½ percent from ¥130.07 at the opening of the period to ¥126.70 on November 22.

The dollar was not the only currency affected by the divergent performance of major national economies. Pressures also developed among the European currencies during early November as the pace of German expansion contrasted with slowing growth or actual declines in the United Kingdom, Italy, France, and certain other European countries. The market conditions that had allowed several European central banks to lower domestic interest rates earlier in the year dissipated with the November increase in German interest rates. As the mark moved up from its relatively low position in the exchange rate mechanism of the European Monetary System (EMS), a number of participating central banks responded to the softening of their currencies relative to the mark by raising interest rates at a time when their economies were weakening or by intervening against marks to support their currencies. The Italian lira, the French franc, and the British pound were the currencies to come under the strongest downward pressures in November.

Dollar selling in response to the diverging economic trends was tempered somewhat by developments in the Persian Gulf. The Gulf conflict, while not the dominant market force that it later became, served as a background factor supporting the dollar at times during early and mid-November. Developments in the military and diplomatic arena at that time suggested that the probability of a war in the near term was rising. Many market participants interpreted the U.S. Administration's

Chart 1

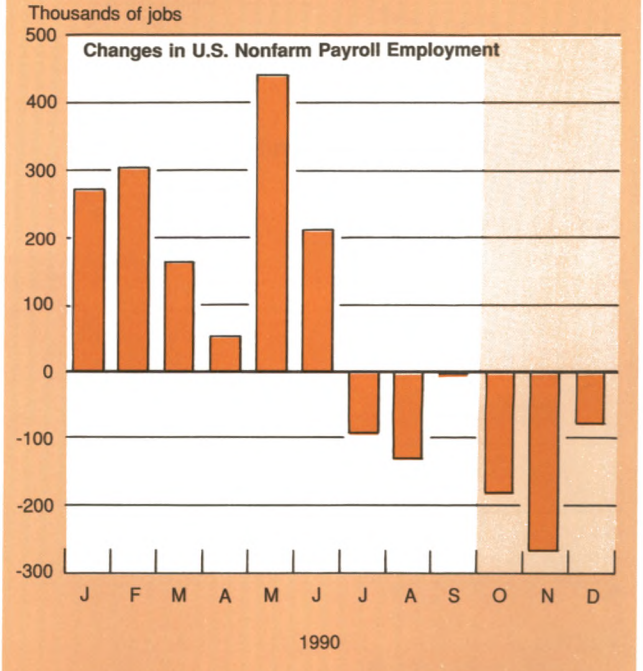
Concerns over the Persian Gulf conflict limited the extent to which negative sentiment toward the dollar was reflected in exchange rates.



Note: The above chart shows the percentage change of weekly average rates for the dollar from January 1990 through January 1991.

Chart 2

Declines in employment pointed to a weakening U.S. economy.

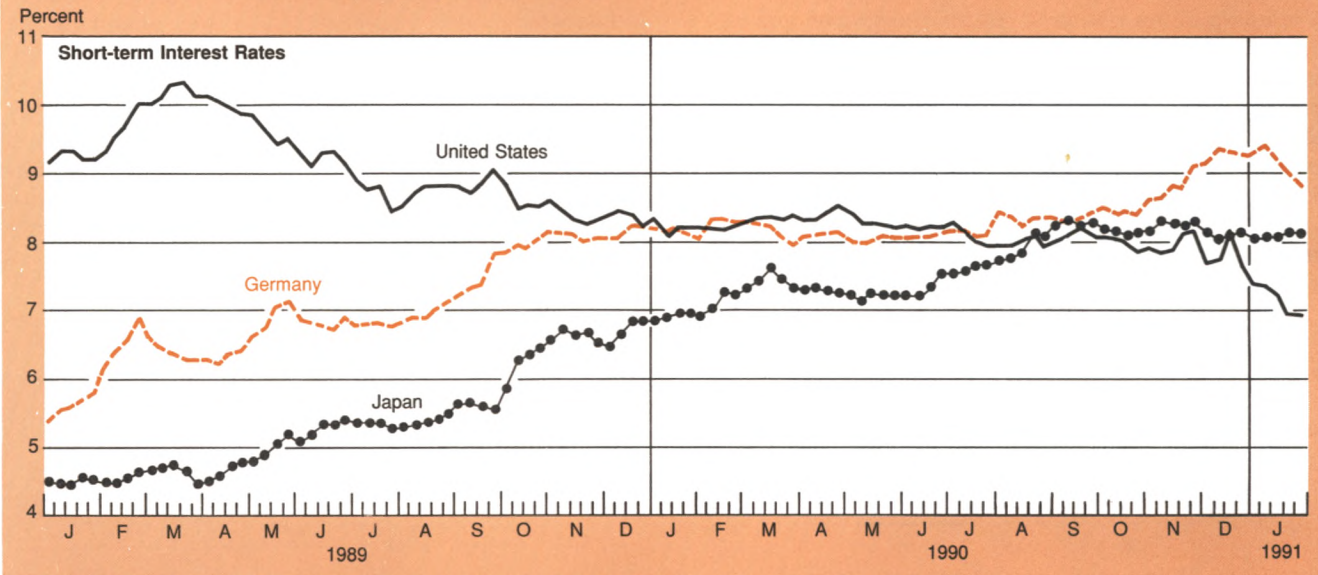


Note: The above chart shows changes in monthly U.S. nonfarm payroll employment from January through December 1990. The shaded area represents data released during the period. The December figure is preliminary.

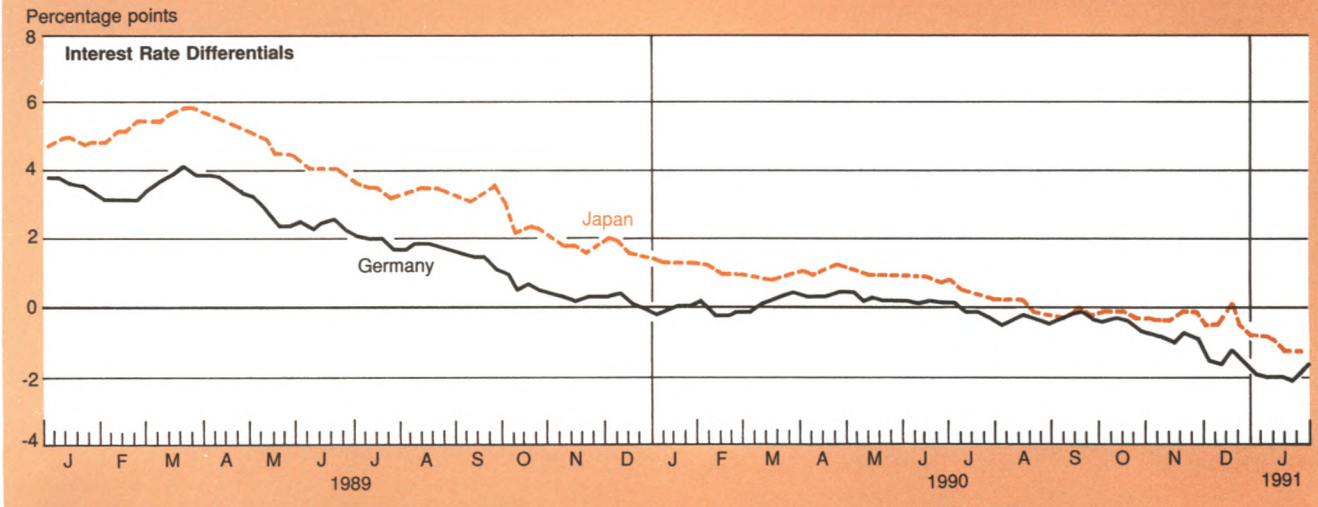
announcement on November 8 of a large reinforcement of U.S. forces in the Gulf as indicating that the United States was preparing for an outbreak of hostilities. Past experience had demonstrated a tendency for the U.S. dollar exchange rate to benefit from "safe haven" inflows during periods of political instability or military conflict abroad, and market participants increasingly came to build in a safe haven premium for the dollar. Inthat environment, dealers became increasingly reluctant to take on large short dollar positions. Thus, notwithstanding negative sentiment about the U.S. economy and the belief that interest rate differentials against the dollar would increase, the prospect of a safe haven effect associated with the outbreak of war helped

Chart 3

U.S. short-term interest rates continued to decline . . .



and interest rate differentials moved further against short-term dollar investments.



Note: The top chart shows weekly average U.S., German, and Japanese three-month Eurocurrency interest rates from January 1989. The bottom chart shows the dollar rate less the foreign rate.

to cushion the dollar's decline.

The middle of the period: late November to mid-January

Beginning in late November, the dollar came under several waves of upward pressure that pushed the currency above its opening levels and to its highs of the period. These pressures primarily reflected heightening expectations that the Gulf conflict would result in an early war. But the dollar's rise was aided by other factors, including a perceived deterioration of the political situation in the Soviet Union and two episodes of acute upward pressure on U.S. money market rates.

From November 29, when the U.N. Security Council set a deadline for Iraq to withdraw from Kuwait, until January 16, when Operation Desert Storm began, market attention focused almost entirely on the Gulf crisis. As the threat of war hung over the market during this month and a half, market participants of all types showed an increased reluctance to take on new risks or to respond fully to changes in underlying economic conditions. With interbank dealing in any case about to wind down as the year-end approached, many dealing institutions took the opportunity to impose an early halt to or reduction in their market-making activities. Many commercial and institutional participants decided to move to the sidelines and, to the extent possible, to postpone further transactions until the Persian Gulf situation was clarified. In this environment, markets became unusually thin and illiquid, and managers of interbank trading rooms at many institutions took steps to reduce the position-taking latitude of their trading staff.

Meanwhile, pressures in the federal funds and other short-term money markets began to appear in late November as banks bid aggressively to secure money to cover year-end accounting statements. These pressures, coming earlier and with much greater intensity than in past years, occurred against a background of heightened concerns over bank credit quality. At the same time, the efforts of many institutions to improve capital ratios, trim balance sheet size, and enhance internal liquidity reduced the availability of and increased the demand for short-term interbank funds, thereby pushing rates upward. Some market participants unable to secure funds in the interbank market bought dollars in the foreign exchange market to meet their year-end requirements. In response, the dollar moved up in late November and early December. When these pressures temporarily subsided in early December, the dollar retraced most of its rise and, in fact, edged down to touch a new post-World War II low against the mark of DM1.4625. But year-end pressures reemerged late in December and again helped support the dollar at that time.

Another reason for the dollar's rise starting in lateNovember was the growing expectation that the finance ministers and central bank governors of the Group of Seven (G-7) would soon meet and discuss exchange rate issues. With strains appearing in the exchange market involving the dollar and other currencies, some market participants believed the G-7 might take steps to stabilize exchange rates. This notion gained credence as several G-7 officials indicated that a meeting would occur in January.

Around mid-December, market unease over the political situation in the Soviet Union also contributed to the dollar's resilience. The December 20 resignation of Soviet Foreign Minister Shevardnadze raised concerns about the outlook for the success of the Soviet leadership's policies of political openness and economic restructuring. Because Germany was viewed as most vulnerable to the spillover effects of negative developments in the Soviet Union, the mark eased. The mark moved lower not only against the dollar and the yen but also against its partner currencies in the EMS. The mark's softer tone helped reduce, albeit temporarily, pressures that had been building throughout December within the EMS exchange rate mechanism.

In these circumstances, the dollar reacted only modestly to a series of actions by the Federal Reserve to ease monetary conditions. These included three moves in December and early January which led to declines in the federal funds rate totaling 75 basis points and one move to reduce the Federal Reserve's discount rate by

Table 1

Federal Reserve Reciprocal Currency Arrangements

In Millions of Dollars

Institution	Amount of Facility January 31, 1991
Austrian National Bank	250
National Bank of Belgium	1,000
Bank of Canada	2,000
National Bank of Denmark	250
Bank of England	3,000
Bank of France	2,000
Deutsche Bundesbank	6,000
Bank of Italy	3,000
Bank of Japan	5,000
Bank of Mexico	700
Netherlands Bank	500
Bank of Norway	250
Bank of Sweden	300
Swiss National Bank	4,000
Bank for International Settlements:	
Dollars against Swiss francs	600
Dollars against other authorized European currencies	1,250
Total	30,100

Table 2

Drawings and Repayments by Foreign Central Banks under Special Swap Arrangements with the U.S. Treasury

In Millions of Dollars; Drawings (+) or Repayments (−)

Central Bank Drawing on the U.S. Treasury	Amount of Facility	Outstanding as of November 1, 1990	November	December	January	Outstanding as of January 31, 1991
Central Bank of Honduras†	82.3	34.8	−34.8	—	—	—

Note: Data are on a value-date basis. Components may not add to totals due to rounding.

†Represents the ESF portion of a $147.3 million short-term credit facility established on June 28, 1990.

50 basis points on December 18. In addition, the Federal Reserve on December 2 announced plans to eliminate reserve requirements on nonpersonal time deposits and on net Eurocurrency liabilities in two stages during December.

Trading in the foreign exchange market remained listless even after the usual year-end holiday lull. During the early weeks of January, as participants awaited the January 15 U.N. deadline for Iraq to withdraw from Kuwait, the dollar tended to move during the day in response to the latest statements or signals regarding diplomatic efforts to avert war. Thus, the dollar eased following announcements that U.S. Secretary of State Baker would meet his Iraqi counterpart in Geneva and that the U.N. Secretary General would visit Iraqi leader Saddam Hussein in Baghdad, only to rebound later when these approaches proved fruitless. Against this background, however, the dollar edged up intermittently. The dollar's movements around this time were greatest against the Japanese yen, which was seen as having the most to lose from any disruption in oil supplies as a result of war and the most to gain from an expected oil price decline in the event of a peaceful settlement. But the dollar also rose against the mark. By mid-January, the dollar was trading up to levels as high as ¥137 against the yen and DM 1.55 against the mark, or roughly 5 percent and 2 percent, respectively, above its early November levels against those two currencies.

The end of the period: mid- to late January

The dollar's response to the outbreak of war on January 16 took many market participants by surprise. Having anticipated a wave of sustained dollar buying upon the outbreak of war, many interbank dealers had quietly been building up long dollar positions as the January 15 deadline approached. In the event, the dollar did move up on the first reports of bombing over Baghdad to highs of DM1.5525 and ¥138.00. However, the currency quickly gave way to selling pressures as market participants took profits on these long positions. Within a few hours after Operation Desert Storm began, the dollar had declined from its highs by about 3 to 4 percent. Oil prices fell back sharply while bond and stock markets around the world rallied.

Table 3

Net Profits (+) or Losses (−) on United States Treasury and Federal Reserve Foreign Exchange Operations

In Millions of Dollars

	Federal Reserve	U.S. Treasury Exchange Stabilization Fund
Valuation profits and losses on outstanding assets and liabilities as of October 31, 1990	+5,363.3	+2,876.3
Realized November 1, 1990–January 31, 1991	0	0
Valuation profits and losses on outstanding assets and liabilities as of January 31, 1991	+5,688.0	+3,027.2

Note: Data are on a value-date basis.

Thereafter, the dollar edged lower through the remainder of January. From time to time, dollar demand increased in response to concerns over the severity and scope of the Gulf conflict. This was the case, for instance, when missile attacks on Israel raised fears that the war might widen. But the dollar's tendency to firm on negative reports out of the Gulf began to wane as market participants appeared to grow more confident that the war would be relatively short and that the United States and its allies would be victorious.

As the exchange market grew accustomed to news from the Gulf and liquidity returned to more normal levels, market participants directed more attention to the economic developments and interest rate changes

that had gone almost unnoticed in December and early January. Against this background, the dollar began to decline again. Statements by Federal Reserve Chairman Greenspan on the potential for further monetary easing if growth of the monetary aggregates remained sluggish, and on the risks of a long and deep recession if the Gulf war were to drag on, were noted. These comments, coupled with President Bush's call for lower interest rates in his State of the Union speech, heightened expectations of further near-term cuts in dollar interest rates.

In a statement issued after their January 21 meeting, G-7 finance ministers and central bank governors "agreed to strengthen cooperation and to monitor developments in exchange markets" and stated they were "prepared to respond as appropriate to maintain stability in international financial markets." Market participants did not conclude at the time, however, that officials were prepared to take immediate and concrete action to stem further dollar declines.

Some market participants came to interpret the January 21 G-7 statement as suggesting that further interest rate increases abroad might be avoided as U.S. rates declined. Indeed, the expectation that Germany would postpone further tightening became so widespread during the last two weeks in January that pressures within the EMS eased, and European authorities were reportedly able to scale back their intervention mark sales. On the last day of the period, however, the Bundesbank increased its official discount and Lombard rates by 50 basis points, an action whose timing took the market by surprise. However, the Bundesbank characterized its move as technical and subsequently took steps to keep money market rates from rising.

Thus, as the period closed, sentiment toward the dollar remained negative as market participants, believing that dollar interest rates would decline further, expected interest rate differentials to continue to move against the dollar. The dollar closed the period at DM1.4768 against the mark, down 2½ percent from its November opening levels and only slightly above its post-World War II low against that currency. Against the yen, the dollar closed the period 1 percent above its opening levels at ¥131.25.

* * * *

As noted in our report for the August-September 1990 period, the U.S. Treasury Exchange Stabilization Fund (ESF) repurchased $2,500 million of foreign currencies from the Federal Reserve on November 1 to reverse certain previous warehousing operations. From that date through the close of the period, outstanding warehousing of foreign currencies with the Federal Reserve remained at $4,500 million, down from the peak of $9,000 million reached in March 1990.

The Treasury also continued to provide SDRs in exchange for dollars to foreign monetary authorities requiring SDRs for the payment of IMF charges and for repurchases. These exchanges totaled $204.3 million equivalent of SDRs over the three-month period.

The ESF's share of a multilateral credit facility established in June 1990 for Honduras was repaid in full during the period, with payments of $34.0 million on November 15 and $0.8 million on November 20. The ESF portion of this special facility expired at the end of November, and as of the end of January 1991, the Treasury had no special swap arrangements outstanding.

As of the end of January, cumulative bookkeeping or valuation gains on outstanding foreign currency balances amounted to $5,688.0 million for the Federal Reserve and $3,027.2 million for the ESF. The latter figure includes valuation gains on warehoused funds. These valuation gains represent the increase in dollar value of outstanding currency assets valued at end-of-period exchange rates, compared with rates prevailing at the time the foreign currencies were acquired.

The Federal Reserve and the ESF regularly invest their foreign currency balances in a variety of instruments that yield market-related rates of return and that have a high degree of quality and liquidity. A portion of the balances is invested in securities issued by foreign governments. As of the end of January, holdings of such securities by the Federal Reserve amounted to $8,114.8 million equivalent, and holdings by the Treasury amounted to the equivalent of $8,000.6 million valued at end-of-period exchange rates.

Treasury and Federal Reserve Foreign Exchange Operations

August-October 1990

During the August-October period, sentiment towards the dollar was generally negative. Exchange market participants continued to focus on signs of sluggish economic activity in the United States and on the movement of interest rates against the dollar. The growth prospects of the U.S. economy were widely perceived as weak, and the adverse trend in interest rate differentials, which had narrowed by several hundred basis points since early 1989, was expected to continue.

The crisis in the Persian Gulf had both positive and negative effects on the dollar. Immediately following the Iraqi seizure of Kuwait on August 2, the dollar rose to its highs of the period amid expectations that the conflict would trigger heavy flows into the dollar. Thereafter, although market participants were attracted to U.S. assets at times when fears of war intensified, the dollar was undermined by concerns that the U.S. economy was more vulnerable than other major economies to the steep rise in oil prices caused by the conflict.

In this environment, the dollar moved generally lower during the period, declining almost 5 percent on a trade-weighted basis as measured by the index of the staff of the Federal Reserve Board of Governors. Against individual currencies, the dollar declined between 4 percent and 4½ percent on balance against the major European currencies, reaching record lows against the German mark and Swiss franc. It declined against the Japanese yen by almost 11 percent to trade at its lowest levels against that currency since January 1989. The dollar was relatively unchanged against the Canadian dollar. The U.S. authorities did not intervene in the foreign exchange market during the period.

A report presented by Sam Y. Cross, Executive Vice President in charge of the Foreign Group at the Federal Reserve Bank of New York and Manager of Foreign Operations for the System Open Market Account. Thaddeus D. Russell was primarily responsible for preparation of the report.

* * * *

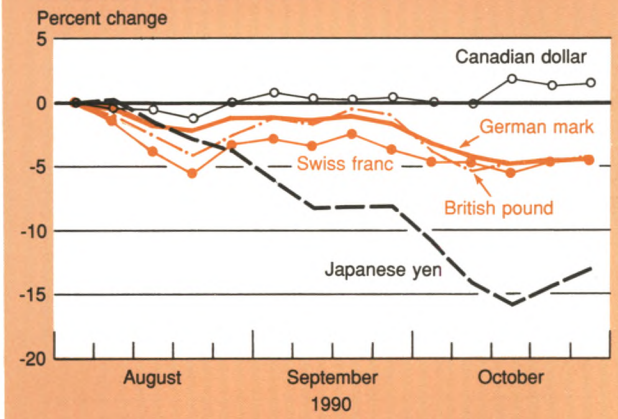
The outlook for the U.S. economy was a focus of attention in the exchange market throughout the period under review as market participants looked to each new economic statistic for signs of how significantly the U.S. economy was slowing. A report released just before the period had shown second-quarter GNP growth to be less rapid than had been expected at an annual rate of 1.2 percent. In early August, a number of data releases and reports reinforced impressions of slowing economic activity, including data on employment, industrial production, and capacity utilization as well as the Federal Reserve's "beige book" survey of economic conditions around the country.

As the period progressed, subsequent data releases provided mixed and hard-to-interpret signals about the U.S. economy. But the view of the economy in the exchange market and among observers more generally became increasingly negative, in large part because of concern over the economic impact of the sharp increases in oil prices resulting from the Persian Gulf crisis. Market participants believed that the U.S. economy was less able to cope than some of the other industrial economies with the potential effects of sharply higher oil prices on business activity and prices. A September 25 report revising second-quarter economic growth downwards to a 0.4 percent rate suggested to market participants that the U.S. economy

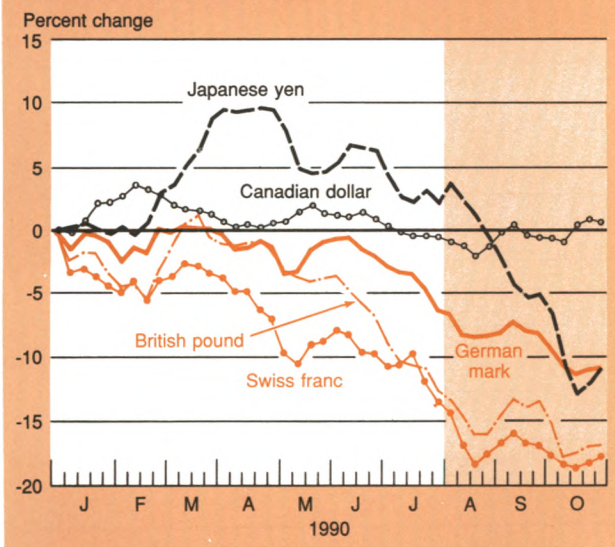
was weakening markedly, even before the economic effects associated with the Persian Gulf crisis had begun to affect it. Other economic data released over the period provided a more mixed impression, including preliminary U.S. GNP data released on October 30 estimating growth of 1.8 percent at an annual rate for the third quarter.

Chart 1

The dollar moved generally lower through the period. The downward movement was particularly sharp against the Japanese yen, . . .



making the dollar's decline against the yen over the course of the year comparable to its decline against the German mark.



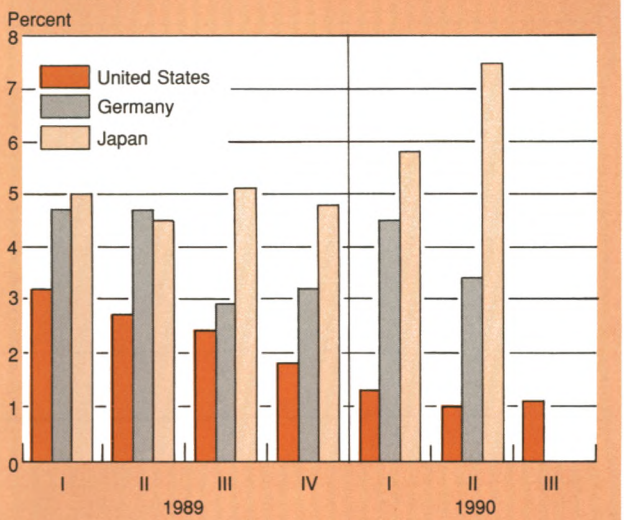
Notes: The top chart shows the percent change of weekly average rates for the dollar from July 31, 1990. The bottom chart shows the percent change of weekly average rates for the dollar from December 31, 1989. All rates are calculated from New York closing quotations.

Spreading perceptions of slowing U.S. economic activity added to the view that interest rates in the United States would continue to go down and that interest rate differentials would move further against the dollar. Expectations of lower interest rates were reinforced by the prospect that some form of compromise would be reached to reduce the U.S. fiscal deficit. After a major U.S. money center bank announced large staff cuts and increased provisions for problem loans late in September, U.S. banks also became a focus of discussion in the exchange market, with some market participants believing that the condition of U.S. banks added to the likelihood that the Federal Reserve would ease.

On September 30, news of a budget accord between negotiators from the White House and Congress also increased expectations that the Federal Reserve would soon allow an easing in the federal funds rate. After that initial budget package failed to pass Congress on October 5, however, the focus of market attention shifted away from interest rates. As the budget negotiations became protracted, the market grew preoccupied with the stalemate itself, which was widely viewed as evidence of the unmanageability of the budget process and of serious disarray within the U.S. government over economic management generally. Thus, concern over the impasse continued to weigh on the dollar until the

Chart 2

During the past two years, the pace of economic activity measured on a year-on-year basis has slowed in the United States while remaining strong in Japan and Germany.



Note: This chart shows quarterly changes in real GNP on a year-on-year basis.

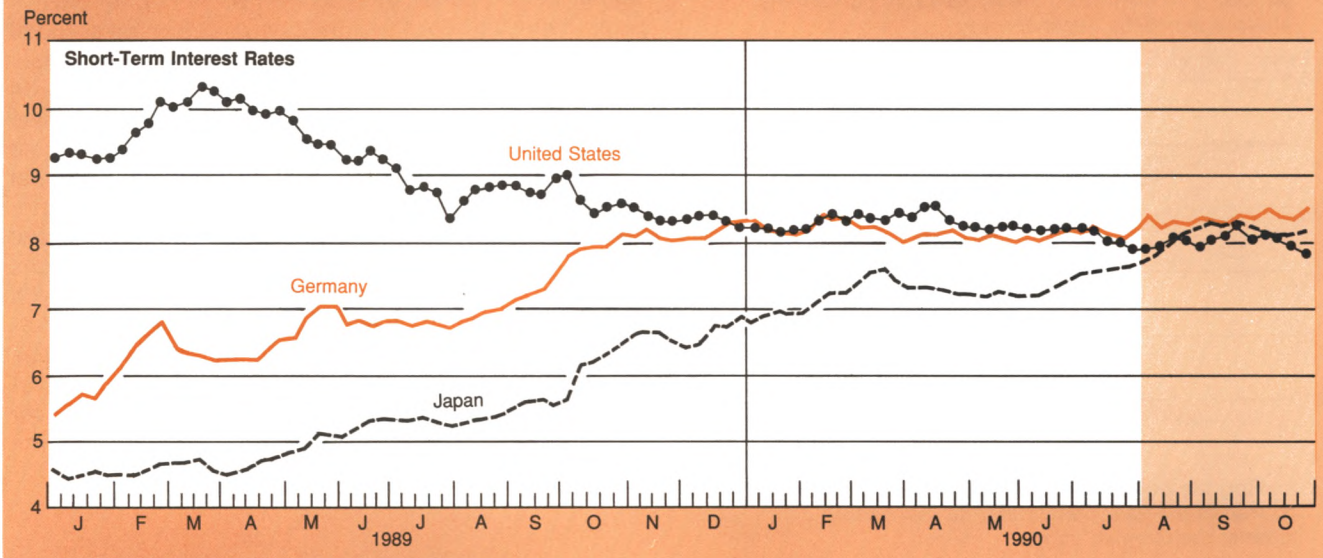
closing days of the period. Even when a new budget acceptable to the President was finally approved by Congress on October 27, it gave little lift to market sentiment toward the dollar.

The decline in the dollar during the period occurred principally during three waves of selling pressure.

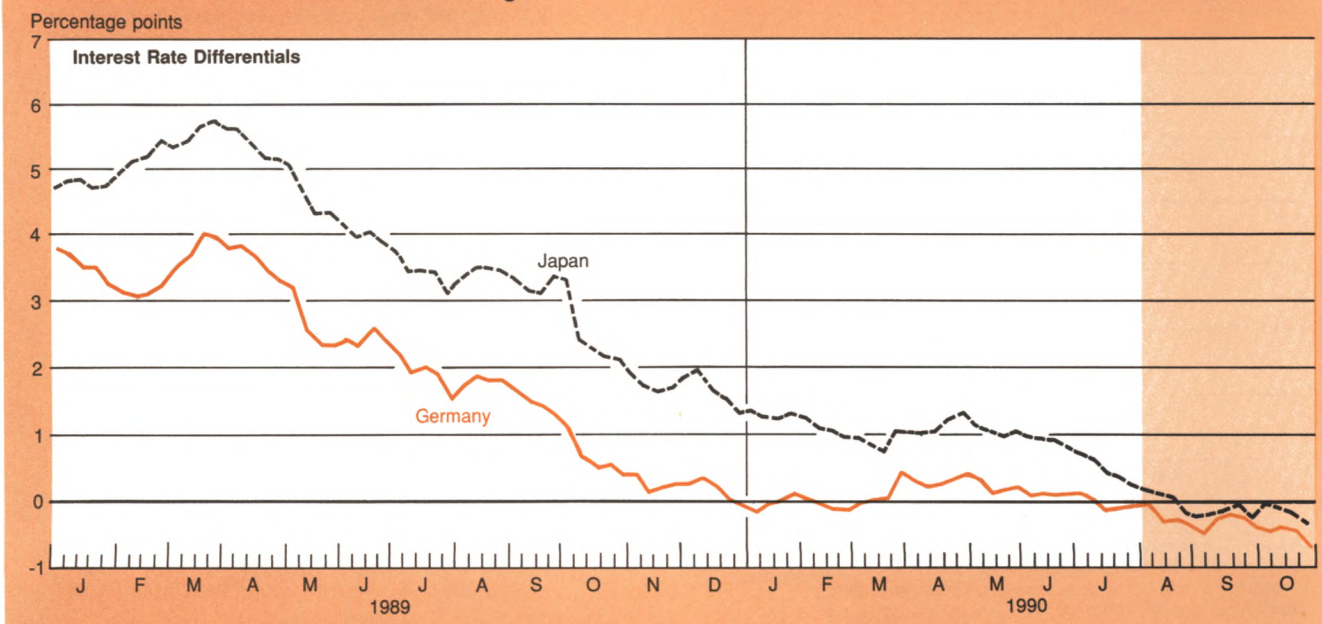
The first occurred during the first three weeks of August. Although the dollar initially firmed on news of Iraq's invasion of Kuwait, reaching its period highs on August 2 of DM 1.6215 against the mark and ¥151.60 against the yen, it quickly started to decline against the European currencies as market participants became

Chart 3

Since early 1989, increases in Japanese and German interest rates, together with declines in U.S. rates, . . .



eliminated the dollar's interest rate advantage.



Notes: The top chart shows weekly average three-month Euro interest rates. The bottom chart shows the dollar rate less the foreign rate.

more concerned over U.S. economic prospects. At this time, the dollar showed little net movement against the Japanese yen, the currency initially the most negatively affected by fears of a disruption of oil flow from the Middle East.

The second wave took place around mid-September when the dollar declined against the yen but traded relatively steadily against other major currencies. The dollar's decline against the yen stalled for a time around the September 22 meeting in Washington of the Group of Seven Finance Ministers and Central Bank Governors. The communiqué released after the meeting stated that the officials had noted the yen's appreciation since their last meeting and that they had "concluded that exchange rates were now broadly in line with continued adjustment of external imbalances."

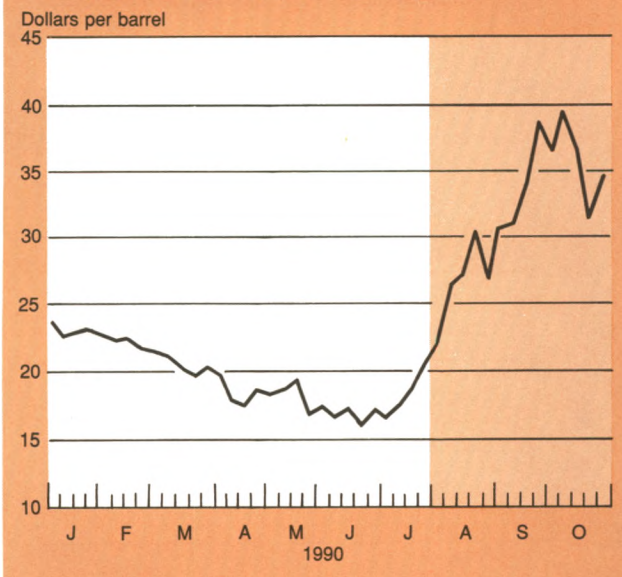
From late September through mid-October, the third wave occurred, with the yen leading a generalized rise of foreign currencies against the dollar. At that time, market participants became increasingly concerned about the impasse over the U.S. government budget, and perceptions developed in the market that officials, both in the United States and abroad, were not concerned about the dollar's decline. The dollar traded as low as ¥123.75 against the yen on October 18 and DM 1.4910 against the mark the next day, its lows for the period.

Late in October, steps were taken toward dispelling the impression of a lack of official concern. Treasury officials made clear in statements to the press that the Administration was concerned about the dollar, and rejected suggestions that the decline was welcomed. At about the same time, market rumors of U.S. intervention served as a reminder to market participants of the possibility of official action to support the dollar. In fact, the U.S. monetary authorities did not intervene during the three months under review.

The extent to which the dollar moved against individual currencies was further influenced by developments in their respective countries. With the formal unification of Germany on October 3, the pressures and anticipated costs associated with the integration of the East German economy into that of West Germany were a matter for reevaluation in the exchange market. The German mark continued to benefit from the perception that a large fiscal deficit and the fast pace of domestic economic expansion underway in the western part of the country, driven in part by demand from the east, would keep German interest rates firm or rising. Market participants noted repeated assurances from the Bundesbank that it would adhere to a strict, anti-inflationary policy stance, as well as the call for a strong mark to keep inflation in check and to help attract capital to

Chart 4

Oil prices rose sharply during the period.



Note: The chart shows weekly average prices per barrel of West Texas oil.

Table 1

Federal Reserve Reciprocal Currency Arrangements

In Millions of Dollars

Institution	Amount of Facility October 31, 1990
Austrian National Bank	250
National Bank of Belgium	1,000
Bank of Canada	2,000
National Bank of Denmark	250
Bank of England	3,000
Bank of France	2,000
Deutsche Bundesbank	6,000
Bank of Italy	3,000
Bank of Japan	5,000
Bank of Mexico	700
Netherlands Bank	500
Bank of Norway	250
Bank of Sweden	300
Swiss National Bank	4,000
Bank for International Settlements:	
Dollars against Swiss francs	600
Dollars against other authorized European currencies	1,250
Total	30,100

finance economic integration. The mark's strength was dampened periodically during the period as large upward revisions in estimates of the expenses associated with unification suggested that the costs and difficulties had been misgauged. Concerns about these problems and the upward trend in German interest rates also contributed to the sharp declines in German stock prices during the period.

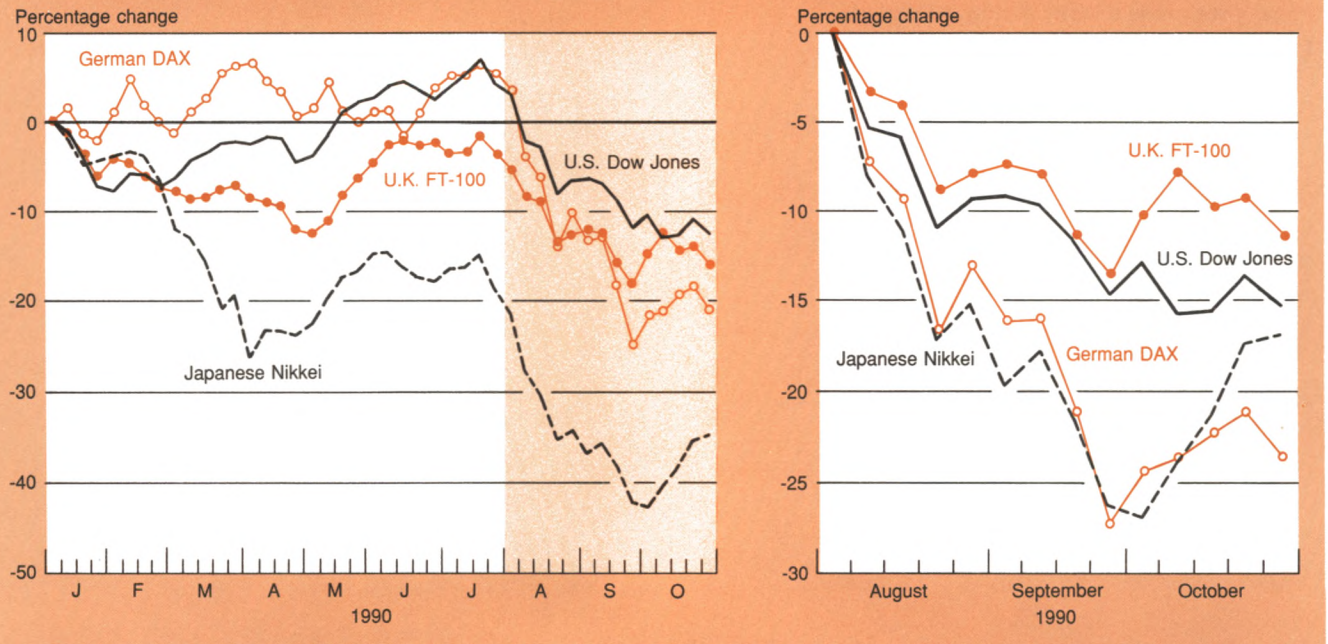
Among other European currencies, the pound sterling moved higher against the dollar during the period. It thereby moved broadly in line with the rise of the mark, despite signs of a weakening in economic activity, rising unemployment, and declining output and retail sales. The pound gained some support from safe-haven flows and from the perception that sterling would benefit from the United Kingdom's North Sea oil fields. Also, through much of the period, sterling was buoyed by expectations that the currency would soon join the Exchange Rate Mechanism (ERM) of the European Monetary System. On October 5, these expectations were borne out when it was announced that the pound was entering the ERM with a 6 percent margin of fluctuation. During the rest of October, the pound declined, moving below its ERM parity rate against the mark of DM 2.95.

Like the mark, the Swiss franc closed the three-month period almost 4½ percent higher on balance against the dollar. Early in the period, the Swiss franc led the rise against the dollar and strengthened against all major currencies. At that time, the franc appeared to benefit to some extent from the nervousness and uncertainties surrounding the situation in the Middle East. Its strength was also based on the Swiss National Bank's tight, anti-inflationary policy stance. After moving up to an all-time high of SF 1.2525 against the dollar on August 23, the franc fluctuated below this level through the end of October, while other foreign currencies subsequently moved higher. The franc's rise stalled after the Swiss central bank took advantage of the leeway provided by the currency's strength to moderate its tight monetary policy slightly, a move acknowledged in public comments towards the end of August.

The Japanese yen appreciated significantly against other major currencies during all but the initial days of the three-month period. The first effect of the invasion of Kuwait was to push the yen down against other currencies as the exchange market initially reacted to Japan's heavy dependence on imported oil and fears of a complete disruption of Persian Gulf oil shipments.

Chart 5

The world's major stock markets declined during August and September.



Notes: The panel on left shows the percent change in weekly averages of daily closing levels. Panel on right shows the percent change in the weekly averages since the beginning of the quarter under review.

Table 2

Drawings and Repayments by Foreign Central Banks under Special Swap Arrangements with the U.S. Treasury

In Millions of Dollars; Drawings (+) or Repayments (−)

Central Bank Drawing on the U.S. Treasury	Amount of Facility	Outstanding as of July 31, 1990	August	September	October	Outstanding as of October 31, 1990
Bank of Guyana†	31.8	13.4	0.0	−13.4	—	—
National Bank of Hungary‡	20.0	20.0	−12.1	−7.9	—	—
Central Bank of Honduras§	82.3	57.3	−22.6	0.0	0.0	34.8

Note: Data are on a value-date basis. Components may not add to totals because of rounding.
The ESF's special facility with the Bank of Mexico, inactive since July 31, 1990, expired on September 14, 1990.
†Represents the ESF portion of a $178 million short-term credit facility that expired on September 20, 1990.
‡Represents the ESF portion of a $280 million short-term credit facility that expired on September 14, 1990.
§Represents the ESF portion of a $147.3 million short-term credit facility established on June 28, 1990.

However, the yen soon began moving higher against both the dollar and other currencies as these concerns receded and market participants came to focus more on the rising cost of oil—a cost which the Japanese economy seemed better able to absorb than other countries. Furthermore, market participants expected that movements in interest rate differentials would continue to favor the yen. Market participants believed that the Bank of Japan, already concerned about the fast pace of Japan's economic expansion and inflationary pressures, would be quick to raise interest rates in response to the increase in energy costs resulting from the Persian Gulf crisis. In fact, the Japanese central bank did raise its discount rate by ¾ percentage point on August 30.

In response to rising market interest rates that both preceded and followed the discount rate hike, talk spread that Japanese investors were finding the returns they were getting at home to be adequate and would no longer be investing abroad as much as before, especially in the United States. Meanwhile, the decline in Japanese equity prices resumed, with the Nikkei index of the Tokyo Stock Exchange down 48 percent at the beginning of October from its levels at the start of the year. Accordingly, a number of Japanese banks, in response to the sharp falls in values of their domestic stock investments as well as their bond holdings, repatriated funds to shore up their domestic capital positions ahead of the end of the fiscal half year September 30. The yen's rise gained more momentum as Japanese companies and investors also moved to raise their hedge ratios on foreign holdings from below-average to above-average levels.

As the yen rose, Japanese officials were increasingly questioned about their attitudes towards exchange rates as some small- and medium-sized Japanese firms began to report that they were losing export competitiveness. Official comments at first left questions in the market as to whether either the Japanese or the U.S. authorities cared if the yen continued to rise. But in late October a large customer purchase of dollars against yen carried out by this Bank was seen in the market. Then, various remarks by U.S., Japanese, and French officials renewed market participants' wariness that the authorities might intervene to support the dollar.

The U.S. dollar rose slightly on balance against the Canadian dollar during the three months. In the early part of August, the Canadian currency firmed to its highest levels in twelve years against the U.S. dollar. At that time, market concerns over a possible disruption of Persian Gulf oil shipments helped buoy the currency

Table 3

Net Profits (+) or Losses (−) on United States Treasury and Federal Reserve Foreign Exchange Operations

In Millions of Dollars

	Federal Reserve	U.S. Treasury Exchange Stabilization Fund
Valuation profits and losses on outstanding assets and liabilities as of July 31, 1990	+3,547.5	+1,519.5
Realized August 1, 1990–October 31, 1990	0	+415.6
Valuation profits and losses on outstanding assets and liabilities as of October 31, 1990	+5,363.3	+2,876.3

Note: Data are on a value-date basis.

because of Canada's position as a net exporter of oil. However, the currency subsequently began to move lower, particularly after Canadian officials confirmed that the economy had entered a recession and that they were prepared to lower interest rates.

* * * *

The Exchange Stabilization Fund (ESF) renewed warehousing arrangements with the Federal Reserve that fell due within the period. These transactions resulted in realized profits of $415.6 million for the ESF, reflecting the difference between the rates at which the Treasury had acquired the funds and the rates at which the warehousing agreements were renewed. As of October 31, the last day of the period under review, the ESF's outstanding warehousing of foreign currencies with the Federal Reserve totaled $7,000 million, unchanged for the period under review.

The U.S. Treasury, however, had initiated steps prior to the end of the period that resulted in the reversal of $2,500 million of the warehousing of foreign currencies effective November 1, the day after the period's close. The reversal of warehousing of foreign currencies finalized on November 1 was financed in part by the Treasury's issue on October 31 of an additional $1,500 million of SDR certificates to Federal Reserve Banks. The remainder was financed from ESF cash balances. As of November 1, outstanding warehousing of foreign currencies with the Federal Reserve totaled $4,500 million, half the level outstanding earlier in the year.

The Treasury also continued to exchange SDRs for dollars with foreign monetary authorities that needed SDRs for payment of IMF charges and for repurchases, exchanging a total of $558.4 million equivalent of SDRs during the period.

Multilateral credit facilities previously established for Guyana and Hungary, in which the ESF participated, were repaid in full during this period, while a similar facility for Honduras was partially repaid. On September 14, a special Mexican short-term credit facility established in March by the U.S. monetary authorities expired. All drawings on the facility had been repaid prior to the period under review.

Guyana. At the beginning of the period, Guyana's outstanding commitment to the Treasury on its multilateral financing facility totaled $13.4 million. Guyana made four payments in September, including final repayment on September 20, the facility's expiration date.

Hungary. The Treasury's $20 million share of the first two drawings by Hungary was outstanding at the start of the period. Hungary reduced the amount outstanding on its second drawing by $4.8 million on August 1 and the amount outstanding on its first drawing by $7.3 million on August 20. The drawings were fully repaid on September 5. Hungary also completed repayments to the BIS (representing certain member central banks) before the September 14 expiration date of the facility.

Honduras. On August 1, Honduras made a partial repayment of $22.6 million to the Treasury, leaving an outstanding balance of $34.8 million on the Treasury's part of a multilateral facility.

As of the end of October, cumulative bookkeeping or valuation gains on outstanding foreign currency balances were $5,363.3 million for the Federal Reserve and $2,876.3 million for the ESF (the latter figure includes valuation gains on warehoused funds). These valuation gains represent the increase in dollar value of outstanding currency assets valued at end-of-period exchange rates, compared with rates prevailing at the time the foreign currencies were acquired.

The Federal Reserve and the ESF invest their foreign currency balances in a variety of instruments that yield market-related rates of return and that have a high degree of quality and liquidity. A portion of the balances is invested in securities issued by foreign governments. As of the end of October, holdings of such securities by the Federal Reserve amounted to $8,238.7 million equivalent, and holdings by the Treasury amounted to the equivalent of $8,331.6 million valued at end-of-period exchange rates.

Recent FRBNY Unpublished Research Papers†

9033. Steindel, Charles. "Recent Trends in Capital Formation." November 1990.

9034. Radecki, Lawrence J. "A Survey of the Origins and Purposes of Deposit Protection Programs." November 1990.

9035. Chan, Anthony. "Are the Preliminary Announcements of Some Macroeconomic Variables Rational?" November 1990. With Hareesh Dhawale.

9036. Hickok, Susan. "Factors behind the Shifting Composition of U.S. Manufactured Goods Trade." December 1990.

9037. Boldin, Michael D. "Characterizing Business Cycles with a Markov Switching Model: Evidence of Multiple Equilibria." December 1990.

9101. Seth, Rama, and Alicia M. Quijano. "Growth in Japanese Lending and Direct Investment in the United States: Are They Related?" January 1991.

9102. Park, Sangkyun. "A Triggering Mechanism of Economy-wide Bank Runs." February 1991.

9103. Park, Sangkyun. "Bank Failure Contagion in Historical Perspective." February 1991.

9104. Yuengert, Andrew M. "Estimating Immigrant Assimilation Rates with Synthetic Panel Data." February 1991.

9105. Yuengert, Andrew M. "Self-Employment and the Earnings of Male Immigrants in the U.S." February 1991.

9106. Uctum, Merih. "Foreign Interest Rate Disturbance, Financial Flows, and Physical Capital Accumulation in a Small Open Economy." February 1991.

9107. Seth, Rama. "Patterns of Corporate Leverage in Selected Industrialized Countries." February 1991.

†Single copies of these papers are available upon request. Write Research Papers, Room 901, Research Function, Federal Reserve Bank of New York, 33 Liberty Street, New York, N.Y. 10045-0001.

Single-copy subscriptions to the *Quarterly Review* (ISSN 0147-6580) are free. Multiple copies are available for an annual cost of $12 for each additional subscription. Checks should be made payable in U.S. dollars to the Federal Reserve Bank of New York and sent to the Public Information Department, 33 Liberty Street, New York, N.Y. 10045-0001 (212-720-6134). Single and multiple copies for U.S. and for other Western Hemisphere subscribers are sent via third- and fourth-class mail, respectively. All copies for Eastern Hemisphere subscribers are airlifted to Amsterdam and then forwarded via surface mail. Multiple-copy subscriptions are packaged in envelopes containing no more than ten copies each.

Quarterly Review subscribers also receive the Bank's Annual Report.

Library of Congress Catalog Card Number: 77-646559

Federal Reserve Bank of New York
33 Liberty Street
New York, N.Y. 10045-0001

Return Postage Guaranteed