

FEDERAL RESERVE BANK OF KANSAS CITY

Economic Review



Second Quarter 1993

Can IRAs Cure the Low National Savings Rate?

Asymmetric Effects of Monetary Policy

*The Impact of Monetary Policy on Bank Lending:
The Role of Securities and Large CDs*

*The Changing U.S. Pork Industry: A Dilemma
for Public Policy*

Are There Too Many Governments in the Tenth District?

Economic Review

Federal Reserve Bank of Kansas City

Second Quarter 1993

Volume 78, Number 2

- Can IRAs Cure the Low National Savings Rate?* 5
By C. Alan Garner
- Asymmetric Effects of Monetary Policy* 21
By Donald P. Morgan
- The Impact of Monetary Policy on Bank Lending:
The Role of Securities and Large CDs* 35
By William R. Keeton
- The Changing U.S. Pork Industry: A Dilemma for Public Policy* 49
By Alan Barkema and Michael L. Cook
- Are There Too Many Governments in the Tenth District?* 67
By Glenn H. Miller, Jr.
-

Contents

Can IRAs Cure the Low National Savings Rate?

5

By C. Alan Garner

Many Americans believe the low national savings rate is a serious economic problem. Because savings, and in turn investment, are key determinants of real income growth and future living standards, economists and fiscal policymakers have proposed various policy changes as possible cures for the low national savings rate. A popular proposal has been to encourage greater participation in Individual Retirement Accounts (IRAs). Discussion of IRAs has temporarily waned as the Clinton administration focuses on such issues as long-term deficit reduction, health care costs, and infrastructure investment. Nevertheless, in coming years, proposals for expanding IRA participation are likely to reappear. There is disagreement, however, about whether increased IRA participation would actually raise national savings.

Garner examines whether changing the tax laws to encourage greater IRA participation would be a reliable way to boost the nation's savings. He identifies three basic problems that kept IRAs from being an effective savings incentive in the 1980s and shows why recent reform proposals would not solve these problems.

Asymmetric Effects of Monetary Policy

21

By Donald P. Morgan

Recent U.S. experience suggests that tight monetary policy slows the economy more than easy monetary policy accelerates it. When monetary policy was tight in 1988 and 1989, the economy seemed to slow in response. Yet when monetary policy was eased in 1990, the economy did not respond accordingly. The suggestion that monetary policy has such asymmetric effects is not altogether new or unorthodox. Indeed, mainstream economists adopted this view for several decades after the Great Depression, when easy monetary policy seemed powerless to revive the economy. Recent studies have revived interest in asymmetry. Theoretical research has suggested reasons why tight policy may have more impact than easy policy. And empirical studies have produced evidence of asymmetry using changes in monetary growth to identify the stance of policy.

Morgan finds evidence that monetary policy has asymmetric effects using two alternative measures of the stance of policy: the federal funds rate and a narrative index based on the statements of policymakers.

*The Impact of Monetary Policy on Bank Lending:
The Role of Securities and Large CDs*

35

By William R. Keeton

The seemingly small impact of easier monetary policy on bank lending during the recent recovery has added to the controversy over the lending view of monetary policy. According to the lending view, easier monetary policy stimulates the economy by increasing bank lending. But critics claim that bank lending may fail to rise because banks may use the deposits generated by the easier policy to buy securities or retire large CDs.

The controversy over the lending view has focused only on the *direct* effect of monetary policy—both sides have largely ignored the *indirect* effect. Easier monetary policy not only increases bank lending directly by increasing deposits, but also indirectly by lowering open-market rates.

Keeton analyzes the implications of bank security holdings and large CDs for the direct and indirect effects of monetary policy on bank lending. He argues that when both the direct and indirect effects are taken into account, banks' ability to fund loans by selling securities and issuing large CDs may strengthen monetary policy rather than weaken it.

The Changing U.S. Pork Industry: A Dilemma for Public Policy

49

By Alan Barkema and Michael L. Cook

The shape of the U.S. pork industry is changing dramatically, as pork production shifts into the hands of fewer, larger farmers with closer ties to processors and consumers. Such a change points to the loss of thousands of small hog farms, which has triggered a public policy debate in Iowa, Kansas, and other leading hog producing states. Primarily responsible for the changes under way are today's discriminating consumers, who challenge the industry to pack improved nutrition into more convenient products. The industry is responding with an arsenal of new technologies and is abandoning its traditional way of moving pork to the market.

Barkema and Cook consider the changes under way in the U.S. pork industry today and what the changes suggest in the years ahead. The authors conclude that the wave of structural change in the pork industry will continue, resulting in a more integrated industry of fewer, larger farms with closer market ties to pork processors. The industry's emerging structure poses a dilemma for public policy, which must balance the loss of traditional small farms against the economic benefits to consumers of higher quality, lower cost products.

Are There Too Many Governments in the Tenth District?

67

By Glenn H. Miller, Jr.

Many taxpayers in the Tenth District and elsewhere are concerned that excessive spending and taxation by the state and local government sector are due to too many governments. Some urge consolidation, aimed at eliminating duplication of effort, as the best way to increase efficiency. One way consolidation might be achieved is by centralizing the state-local sector, that is, by providing services from the state house rather than from courthouses and city halls. Another way might be to merge units of local government. But making governments bigger through consolidation does not necessarily make government more efficient or the public sector smaller. Indeed, some researchers suggest just the opposite—that a greater number of governments in a certain area will reduce the overall size of the public sector in that area.

Miller reviews the evidence on the consolidation argument, especially on how the number of governments and the structure of the state-local sector affect the efficiency and size of the public sector. He concludes that those interested in controlling tax burdens must look beyond the number of governments and not jump to the conclusion that consolidation inevitably makes government more efficient.

Can IRAs Cure the Low National Savings Rate?

By C. Alan Garner

Many Americans believe the low national savings rate is a serious economic problem. Because savings, and in turn investment, are key determinants of real income growth and future living standards, economists and fiscal policymakers have proposed various policy changes as possible cures for the low national savings rate. A popular proposal has been to encourage greater participation in Individual Retirement Accounts (IRAs), which provide a tax-advantaged account for retirement savings. Last year, for example, the Bush administration proposed a new "flexible" IRA, and Senators Bentsen and Roth introduced a bill liberalizing IRA eligibility and creating a new kind of IRA. Legislation based on the Bentsen-Roth plan was passed by Congress late in 1992, but was not signed into law.

Discussion of IRAs has temporarily waned as the Clinton administration focuses on such issues as long-term deficit reduction, health care costs, and infrastructure investment. Nevertheless, in coming years, proposals for expanding IRA participation are likely to reappear. IRA reform remains popular with many fiscal policymakers

anxious to raise the national savings rate. And IRAs are politically appealing as a form of middle-class tax relief.

There is disagreement, however, about whether increased IRA participation would actually raise national savings. National savings is the sum of government savings and private savings. Increased IRA participation would reduce government savings by decreasing tax revenues and raising the budget deficit. Nevertheless, higher IRA contributions could increase national savings if private savings were to rise by more than the decline in government savings. However, economic studies reach differing conclusions about whether, and how much, IRAs increase private savings.

This article argues that changing the tax laws to encourage greater IRA participation would not be a reliable way to boost the nation's savings. The first section explains why the low savings rate is a source of concern and briefly describes how IRAs work. The second section shows that IRAs were not successful in raising the national savings rate in 1982-86, the period of broadest IRA participation. Finally, the third section identifies three basic problems that kept IRAs from being an effective savings incentive in the 1980s and shows why recent reform proposals would not solve these problems.

C. Alan Garner is a senior economist at the Federal Reserve Bank of Kansas City. Carrie Ross, an assistant economist at the bank, helped prepare the article.

NATIONAL SAVINGS AND IRAs

Many Americans are concerned about future U.S. living standards because of the sluggish growth of productivity and real output over the last two decades.¹ Workers often feel uneasy about their own living standards in retirement and about the economic prospects for their children and grandchildren. Recent debate about future living standards has centered on the low U.S. savings rate and policy options, such as IRAs, for raising the savings rate.

The low savings rate

The national savings rate has been low in recent years compared with both past U.S. savings rates and savings rates in other industrial countries. For example, the national savings rate averaged 2.4 percent of net national product over the last five years, which was well below the average 8.8 percent rate in the 1960s. International statistics also suggest that the savings rate is far lower in the United States than in other industrial countries, such as Canada, Germany, and Japan.²

A low national savings rate may hurt future living standards by reducing domestic investment and productivity growth. If an economy is closed to international capital flows, domestic investment and savings are closely related because capital formation requires that real output be shifted away from consumer goods into new plant and equipment. A low savings rate would thus reduce the quantity of capital available for workers to use in the production process. A lower level of capital per worker would make workers less productive and cause firms to pay lower real wages than if the savings rate were higher.

But a low savings rate may hurt future living standards even if the economy is open to international capital flows. Because companies in an open economy can borrow abroad, they may be able to finance the same capital stock as if the

savings rate were higher, and thus worker productivity and real wages may also be the same. Nevertheless, future generations will have to consume a smaller share of the net national product because of higher interest and dividend payments to foreigners. Thus, a low national savings rate may hurt future living standards by increasing U.S. indebtedness to foreigners.³

How IRAs work

Because of such concerns about future living standards, expanded IRA programs have been advocated as a possible cure for the low savings rate. An IRA is a tax-advantaged account designed to encourage retirement savings. Many taxpayers can receive tax benefits by deducting part or all of their IRA contribution from their taxable income during the year in which the contribution is made. Under current tax laws, this deduction is gradually eliminated for high-income taxpayers with a pension plan.⁴ All IRA contributors can benefit, however, by deferring taxes on their earnings until the funds are withdrawn from the IRA. An IRA may have an additional tax benefit if the household's income tax rate drops in retirement, allowing withdrawals to be taxed at a lower rate than during the household's working years.

Nevertheless, many households do not contribute to an IRA because IRA balances are illiquid, meaning the funds cannot be converted into cash without a large loss in value. A taxpayer who withdraws funds from an IRA before reaching 59-1/2 years of age must, by law, pay a 10 percent withdrawal penalty in addition to the deferred income tax. The withdrawal penalty is designed to encourage taxpayers to use IRAs solely for retirement savings. A taxpayer may therefore be unwilling to put funds into an IRA if those funds might be needed in the near future. However, an IRA may not be illiquid for older households because the law allows a taxpayer over 59-1/2 years of age to withdraw funds from an IRA without penalty.

Table 1

Participation in the IRA Program, 1979-89

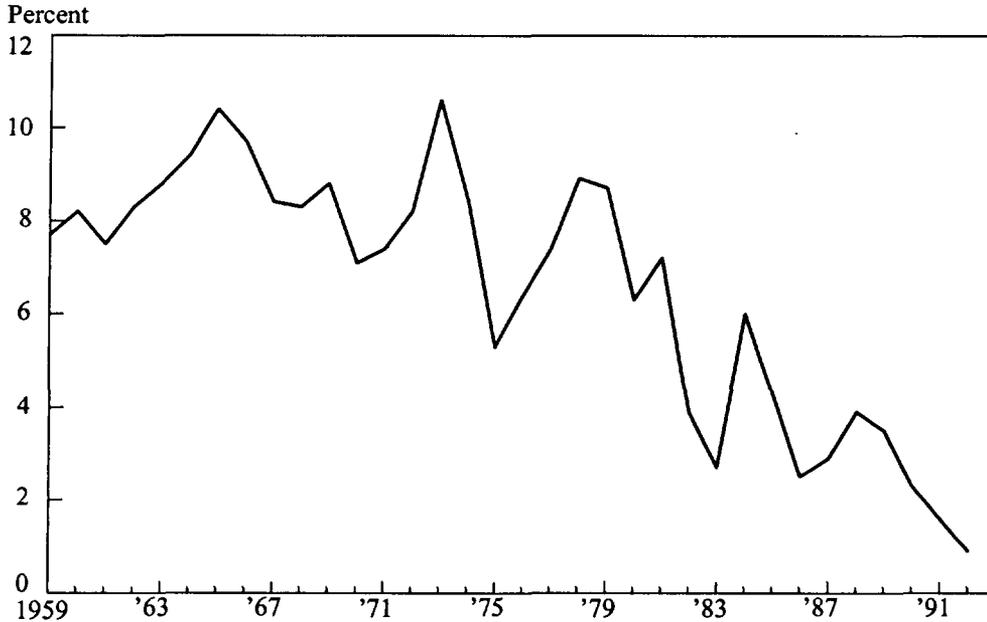
Year	Number of returns claiming IRA deduction (millions)	Amount of IRA deductions claimed (billions)
1979	2.5	\$3.2
1980	2.6	3.4
1981	3.4	4.8
1982	12.0	28.3
1983	13.6	32.1
1984	15.2	35.4
1985	16.2	38.2
1986	15.5	37.8
1987	7.3	14.1
1988	6.4	11.9
1989	5.8	10.8

Sources: Joint Tax Committee, U.S. Congress; and Internal Revenue Service, U.S. Department of the Treasury.

The amount of IRA contributions rose sharply in the first half of the 1980s in response to changing eligibility requirements. The IRA was originally created in 1974 to encourage retirement savings by workers without a pension plan. The eligibility requirements were liberalized by the Economic Recovery Tax Act of 1981, which increased the deduction limit for IRA contributions and opened the program to workers with a pension plan. As Table 1 shows, the number of tax returns claiming IRA deductions grew dramatically from 3.4 million in 1981 to 16.2 million in 1985. The dollar amount of IRA contributions also grew rapidly from \$4.8 billion in 1981 to \$38.2 billion in 1985.⁵

But contributions to IRAs dropped sharply after the passage of the Tax Reform Act in 1986. The requirements for making deductible IRA contributions were tightened as part of a general philosophy of broadening the tax base to permit lower income tax rates. In particular, tax reform reduced the appeal of IRAs by decreasing or eliminating the IRA deduction for higher income workers with a pension plan. The cut in personal income tax rates also lowered the tax benefits from contributing to an IRA. Table 1 shows that the number of tax returns claiming IRA deductions and the dollar amount of IRA contributions dropped sharply in 1987, and both have remained lower in recent years.

Chart 1

National Savings Rate

Source: Bureau of Economic Analysis.

DID IRAs RAISE NATIONAL SAVINGS IN THE 1980s?

Two kinds of empirical evidence are available to see whether expanded IRA eligibility in 1982-86 raised the national savings rate. Aggregate evidence shows changes over time in the savings rate for the entire economy. Cross-section evidence looks at the effects of IRA participation on saving by individual households at a given point in time. Neither kind of evidence provides much support for the view that IRAs raised the national savings rate.

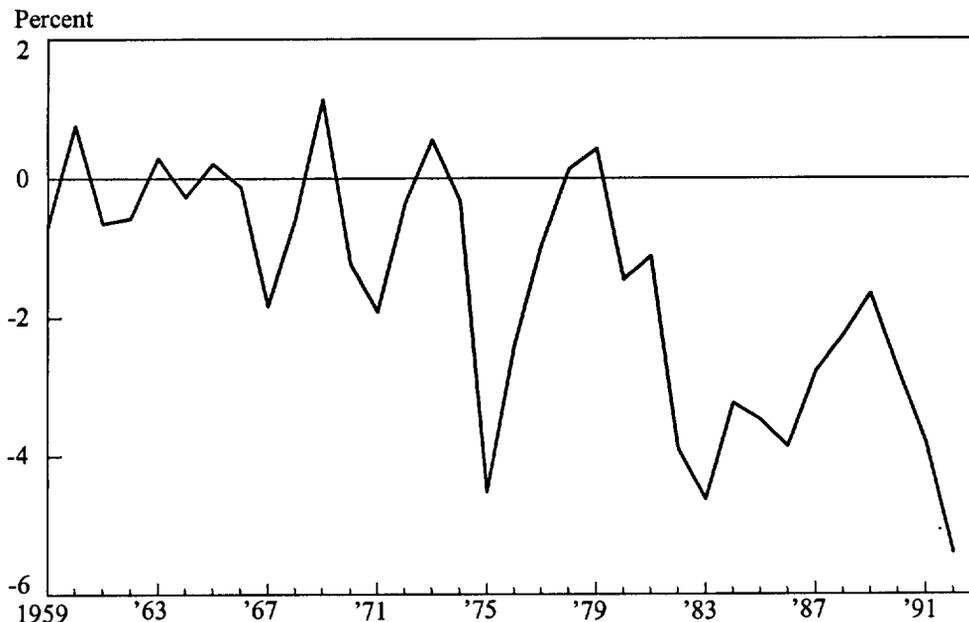
Aggregate evidence

Widespread availability of IRAs in 1982-86

did not halt a persistent downward trend in the national savings rate. Chart 1 shows the national savings rate, defined as national savings divided by the net national product.⁶ Although national savings fluctuated substantially from year to year, the national savings rate fell from 7.2 percent in 1981 to 2.5 percent in 1986, the final year of broad IRA participation. National savings has remained low since 1986, falling to 0.9 percent of the net national product in 1992.

The decline in the national savings rate in the 1980s partly reflected greater dissaving by the government sector, which includes the federal government and state and local governments. Although increased IRA participation in 1982-86 reduced tax revenues, the large decline of government savings was primarily caused by other factors.

Chart 2

Government Savings Rate

Source: Bureau of Economic Analysis.

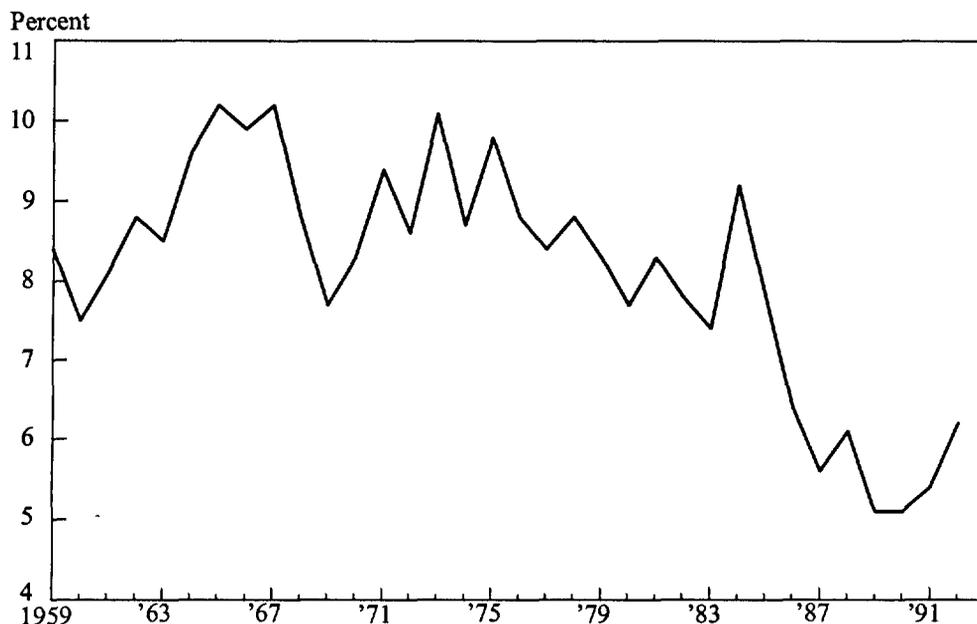
Chart 2 shows the government savings rate, defined as the government sector's budget surplus or deficit divided by the net national product. The government savings rate fell from 0.4 percent in 1979 to -4.6 percent in 1983 as government deficits mounted because of large federal tax cuts and the effects of back-to-back recessions on government revenues. The government savings rate gradually improved to -1.7 percent in 1989, but then dropped sharply to -5.4 percent in 1992, again reflecting the effects of sluggish economic growth on government revenues.

Increased IRA participation in 1982-86, however, might be expected to have more of an effect on private savings, savings by businesses and households.⁷ Yet Chart 3 shows the private savings rate declined from 8.3 percent of the net national product in 1981 to 6.4 percent in 1986. After the

passage of the Tax Reform Act, the private savings rate continued declining to 5.1 percent in 1990 before recovering slightly in 1991 and 1992. Even with this small recovery, the private savings rate remains low by postwar standards.

Aggregate savings rates, thus, do not provide any evidence that IRAs raised the national savings rate in 1982-86. But proponents of expanded IRA programs still believe that IRAs raised both national and private savings. The downward trend in the national savings rate does not, in their view, settle the issue of whether IRAs were an effective savings incentive because national savings might have been even lower without broad IRA eligibility. Indeed, proponents emphasize that the low point in the private savings rate occurred after IRA eligibility was restricted by tax reform.

Chart 3

Private Savings Rate

Source: Bureau of Economic Analysis.

Cross-section evidence

Because of the differing interpretations of aggregate savings trends, recent empirical research on IRAs relies heavily on cross-section evidence showing differences in savings behavior across a large number of households at a given point in time. Information on these households is typically drawn from federal tax returns or surveys of consumer spending and finances from the early to mid-1980s. The empirical results apply most directly to personal savings, a major component of private savings, but such results are then used with other assumptions to estimate the impact of IRAs on the national savings rate.

Cross-section studies by Venti and Wise (1990, 1992) suggest that raising the limit on IRA

contributions would substantially increase the national savings rate. Their 1990 study, for example, concludes that about one-third of the increase in IRA contributions would come from personal tax savings, which decrease the government savings rate, but the remaining two-thirds of the contributions would come from higher personal savings. The increase in personal savings would therefore be larger than the decrease in government savings, raising the national savings rate. Venti and Wise also investigate whether IRA contributors switched funds from existing non-IRA financial assets into IRAs to reduce their tax payments. Such asset switching would reduce government tax revenues without raising personal savings. But Venti and Wise find virtually no switching of savings from non-IRA financial assets into IRAs.

Several economists are critical of the empirical studies by Venti and Wise. For example, Deaton argues that the large savings effect of IRAs in the Venti and Wise studies may be due to statistical problems in working with cross-section data. And Gravelle argues that their results depend heavily on an arbitrary theoretical structure that is inconsistent with conventional economic theory.

Other critics argue that the statistical results of Venti and Wise cannot distinguish between their theory and an alternative in which IRAs have no effect on the savings rate (Gale and Scholz; Joines and Manegold). Venti and Wise's major finding is that IRA contributors had higher than average levels of personal savings. This finding is consistent with the view that IRAs stimulate savings, but it is also consistent with an alternative view that people with a strong desire to save are likely to save more in all forms, including IRAs. Critics contend that Venti and Wise do not control adequately for differing desires to save across households. As a result, the positive association between being an IRA contributor and having a high savings rate does not show whether being an IRA contributor causes a high savings rate or being a heavy saver causes IRA contributions.

Gale and Scholz conclude that increasing the IRA contribution limit would have a much less positive effect on the national savings rate. They find that many IRA contributors had reached the stage in life where they needed to save heavily for retirement. For such savers, IRAs were an attractive way to save but largely captured savings that would have occurred anyway. Gale and Scholz estimate that an increase of \$100 in the IRA contribution limit would raise national savings by only \$2, assuming the tax deduction for the new IRA contribution is entirely saved. But national savings would actually decrease by \$14 if half of the tax deduction were consumed. Thus, Gale and Scholz cannot rule out the possibility that expanded IRA contributions would lower the national savings rate by increasing private savings less than the decrease in government savings.

Joines and Manegold also find that expanding IRA eligibility in 1982-86 did not produce a large increase in the national savings rate. An important feature of this study is that Joines and Manegold follow the behavior of a group of households over time. Because IRA eligibility requirements for many households varied in response to the 1981 change in federal tax law, Joines and Manegold can infer the effect of IRAs on personal savings from actual behavior rather than relying on theoretical assumptions. This study provides evidence of substantial shifting by IRA contributors from their existing financial assets into IRAs. Joines and Manegold conclude that any increase in national savings is likely to be much smaller than estimated by Venti and Wise. Moreover, like Gale and Scholz, they cannot rule out the possibility that increased IRA contributions might lower the national savings rate.

Summary of 1980s evidence

The empirical evidence from the 1980s gives little reason to believe that expanding IRA programs would raise the national savings rate. Increased IRA participation in 1982-86 did not halt the downward trend in the aggregate savings rate. Moreover, although cross-section studies reach differing conclusions, the studies with the best research methods conclude that IRAs provided a weak incentive for higher savings in the 1980s, and may have even reduced national savings.

WOULD IRAs RAISE NATIONAL SAVINGS IN THE 1990s?

Recent proposals to expand IRA participation are somewhat different from the IRA programs of the 1980s. Thus, the past failure of IRAs to raise the national savings rate does not automatically imply that current proposals would be unsuccessful.

ful in the 1990s. To assess recent IRA reform proposals, this section begins by diagnosing why IRAs failed to raise the national savings rate in the past. Then, it argues that the proposed changes in the IRA program would not correct past problems that made IRAs an ineffective savings incentive.

Why were IRAs unsuccessful in the 1980s?

Economic theory suggests three major reasons why the IRA program did not increase the national savings rate in the 1980s.

Insensitivity to the rate of return. The first reason IRAs were unsuccessful in the 1980s is that private savings may be relatively insensitive to changes in the expected rate of return on the taxpayer's investments. Economic theory does not clearly predict whether the higher after-tax return on an IRA would raise or lower the private savings rate. A theoretical argument, called the substitution effect, implies the savings rate would increase because households would save more now to attain a higher future level of consumption. But another theoretical argument, the income effect, may work in the opposite direction. According to this effect, the tax savings from an IRA would increase the household's lifetime spendable income, allowing it to save less both now and in the future. The change in the private savings rate would therefore depend on which theoretical effect predominated.⁸

Target saving provides the clearest case where an increase in the after-tax return on savings could actually reduce the savings rate. A target-saving household wishes to accumulate a specific dollar amount by some future date. For example, suppose a household wishes to save a lump sum now that will grow to \$1,000 in ten years. If the after-tax rate of return were 3 percent, the household would have to save \$744 now to have \$1,000 in ten years. But if an IRA raised the after-tax rate of return to 5 percent, the household would only need to save about \$614. A target-saving house-

hold might therefore reduce its current savings rate if opening an IRA raised the after-tax rate of return.

Some empirical studies also find that the private savings rate is relatively insensitive to changes in the after-tax rate of return. A prominent study by Boskin suggests that an increase in the rate of return causes a large improvement in the savings rate. But a reexamination of the issue by Friend and Hasbrouck finds little support for the belief that higher after-tax rates of return stimulate savings. A recent study by Hall also detects little or no relationship for the U.S. economy between the expected return on savings and the total amount saved.⁹

Weak marginal incentives. The second reason IRAs were unsuccessful in the 1980s is that IRAs may not have increased the after-tax rate of return that many taxpayers earned on an additional dollar of savings. Economic theory applies to marginal spending and saving decisions—decisions to spend or save an additional dollar of income. But some households may have contributed as much as possible to an IRA and then saved even more in non-IRA financial assets. For such contributors, the tax benefits from the IRA were exhausted, and the decision to spend or save an additional dollar of income depended on the lower after-tax rate of return earned at the household's regular tax rate.

Many IRA contributors in 1982-86 probably experienced weak marginal savings incentives because of their favorable financial situations. Table 2 shows that many IRA contributors were higher income households who had reached the stage in life where they needed to save for retirement. These households also had larger net worths and greater holdings of non-IRA financial assets than households without IRAs, suggesting that many IRA contributors were already saving heavily. Moreover, a large proportion of IRA participants exhausted the tax savings available from an IRA because they contributed up to the \$2,000 limit. Thus, participation in the IRA program did

Table 2

Characteristics of Households With and Without IRAs, 1986

Characteristics	Households Without IRAs	Households With IRAs
Median age (years)	49	50
Median three-year income	\$47,000	\$105,000
Median non-IRA financial assets	\$3,000	\$21,695
Median net worth	\$25,470	\$107,946

Source: William G. Gale and John Karl Scholz, "IRAs and Household Saving."

not increase the amount saved by many households because the IRA did not boost the expected return on an additional dollar of savings.

Asset switching. The third reason that IRAs were unsuccessful in the 1980s is that the IRA program reduced government savings by causing many households to switch some of their existing funds out of taxable non-IRA financial assets into IRAs to reduce their tax bills.¹⁰ But how much of their savings were households willing to switch into tax-advantaged IRAs?

From a theoretical standpoint, the amount of asset switching depended on whether contributors viewed IRAs and non-IRA financial assets as close substitutes. Proponents of IRAs often assume that such assets are not close substitutes because of the early withdrawal penalty on IRAs. In this view, households must keep some of their savings in non-IRA financial assets for short-term needs or to be prepared for emergencies. Households would not put such short-term savings into an IRA because of the withdrawal penalty. By this reasoning, the large IRA contributions in 1982-86

must have been new retirement savings rather than assets switched from existing non-IRA balances.

Critics of IRAs respond that a large amount of asset switching probably occurred in 1982-86 because IRAs and non-IRA financial assets were good substitutes for many contributors. The withdrawal penalty was irrelevant for IRA contributors over 59-1/2 years of age. The penalty would also not have been very important for contributors nearing 59-1/2 years of age because such contributors could get their funds back without penalty after a short wait. Moreover, Table 2 showed that many contributors had substantial non-IRA financial assets, some of which were probably being held for retirement. Regardless of age, many taxpayers were probably willing to substitute retirement savings in IRAs for retirement savings in non-IRA financial assets as long as they continued to hold enough non-IRA assets to meet unexpected needs. Such theoretical arguments are consistent with the findings of Joines and Manegold that a large amount of asset switching occurred in the 1980s.

Reform options

Various fiscal policymakers have recently proposed reforms to expand IRA participation and raise the national savings rate. The Bush administration, for example, proposed a new kind of IRA—the Flexible Individual Retirement Account (FIRA)—that would be less of a retirement savings program than the current IRA. In this proposal, taxpayers would receive no initial deduction for their FIRA contributions of up to \$2,500 per year. But regardless of the taxpayer's age, there would also be no tax or penalty on withdrawals after seven years.¹¹ As a result, FIRAs might be useful in saving for intermediate-term financial goals, such as a child's college education or the downpayment on a house.

In the other leading IRA proposal of recent years, Senators Bentsen and Roth advocated restoring the 1982-86 eligibility requirements for IRAs as well as creating a new kind of IRA. The Bentsen-Roth proposal would allow all taxpayers, regardless of income or pension coverage, to contribute up to \$2,000 to a conventional IRA. This proposal would also create a new kind of IRA, similar to the Bush administration's FIRA, with no initial tax deduction for contributions but also no tax on the investment earnings when the funds are withdrawn. Under either option, funds could be withdrawn without penalty for such specific purposes as buying a first home or paying college expenses.

Although efforts to broaden IRA participation have subsided recently, IRAs remain popular with many fiscal policymakers as well as many taxpayers. The chances are good therefore that bills seeking to reform the IRA program will be debated again in the future. But because future reform proposals may differ from the Bush administration and Bentsen-Roth proposals, this section identifies five options that could be combined in various ways in future reform proposals.

Option 1: Raise the income ceiling. A possible IRA reform option is to allow higher income tax-

payers with pension plans to make deductible IRA contributions, while keeping current dollar limits on the amount of the contribution. Such a reform would simply move the IRA program back to the eligibility requirements that existed before 1987. Raising the income ceiling for deductible contributions might be expected to increase the amount of IRA contributions because many contributors in 1982-86 were higher income taxpayers.

But even if the amount of IRA contributions were to increase sharply, this reform option would not necessarily raise the national savings rate. As the previous section showed, IRAs were apparently unsuccessful in raising the national savings rate in the 1980s despite a large increase in the number of contributors. And the problems of weak marginal savings incentives and asset switching are likely to be more severe for higher income contributors than for the average household. Thus, an increase in the income ceiling might cause additional asset switching and channel savings that would have occurred anyway into tax-advantaged IRAs, both of which would reduce government savings.

Raising the income ceiling for deductible IRAs also might not stimulate as large an increase in IRA contributions as occurred in 1982-86. The tax advantages from an IRA are now less than in 1982-86 because personal income tax rates are lower.¹² In addition, taxpayers make greater use of another tax-advantaged savings program, the 401(k) plan, available through many employers. Like an IRA, such plans allow the contributor to defer tax payments on some current income, as well as investment earnings, until the funds are withdrawn at retirement. But such plans may be more attractive than IRAs because of their higher contribution limits and, often, matching contributions by the employer. Any household that does not take full advantage of its tax-deferred 401(k) plan would not experience greater savings incentives under this reform option.¹³

Option 2: Raise the contribution limit. A sec-

ond reform option is to raise, or remove entirely, the dollar limit on IRA contributions. Many high-income households and households with a strong taste for saving experience weak marginal savings incentives when contributions are limited to \$2,000 because they are already saving above this level. Raising or removing the dollar limit would encourage such households to save more because they could earn a higher after-tax return on an additional dollar of savings. But this increase in the after-tax return might not stimulate much, if any, additional savings if the private savings rate is relatively insensitive to the rate of return. The increase in private savings might even be too small to offset the decline in government savings, reducing the national savings rate.

The first two IRA options also may have political disadvantages. Both options would give additional tax benefits primarily to households with above-average levels of income and wealth. Yet President Clinton proposes higher income tax rates for those at the very top of the income distribution. Providing new tax benefits for higher income taxpayers through either option might be seen as inconsistent with the new administration's tax objectives, and probably would provoke complaints from political groups that are concerned about the income distribution.

Option 3: Allow penalty-free withdrawals. A third reform option is to make IRAs more attractive to the typical saver by allowing penalty-free withdrawals for certain reasons, such as buying a first home or paying college expenses. Such a reform would make IRAs somewhat more liquid and thus more appealing to younger taxpayers and middle-income households that may feel uncomfortable about tying up their savings until retirement. Adopting this reform option might therefore result in a substantial new flow of funds into IRAs.

But this reform option might not raise the national savings rate. The movement of funds into such IRAs might largely represent asset switching from existing non-IRA financial assets, where households presumably are putting their current

savings for homebuying or a college education. This reform option also might increase the relevance of the target saving example because households probably have a better idea of the target sum needed for college tuition or the downpayment on a home than about the amount needed for retirement. By offering a higher after-tax rate for achieving these savings goals, this reform option might cause households to save less for such non-retirement objectives.

Option 4: Create a "back-loaded" IRA. A fourth reform option is to create a "back-loaded" IRA, meaning an IRA where the initial contributions are not deductible but IRA distributions are also not taxed. The back-loaded IRA might require the taxpayer to keep funds in the IRA for a certain number of years before gaining any tax benefits, as in the Bush administration's proposed FIRA. A back-loaded IRA might even require that the funds be left in the account until 59-1/2 years of age, similar to the current IRA.¹⁴

But a back-loaded IRA does not solve the economic problems of the conventional IRA. The back-loaded IRA does not, for example, cure the low sensitivity of national savings to changes in the after-tax rate of return, nor does it increase the marginal incentive to save for households that are already saving above the IRA contribution limit. Moreover, if the immediate tax deduction is an important savings incentive—as proponents of the current IRA maintain—the back-loaded IRA might provide less incentive to save than the present IRA.

The back-loaded IRA also creates a political danger of worsening the government budget deficit over the long run. Little government revenue would be lost in the short run because the back-loaded IRA gives no immediate deduction for IRA contributions. Such a feature might be attractive to politicians wishing to provide middle-income tax relief in a time of large government budget deficits. But if taxpayers were to shift funds from taxable assets into a new back-loaded IRA, government revenue losses would gradually increase

as investment earnings accrued without being taxed. Because such revenue losses are not immediately apparent, fiscal policymakers might be tempted to give away a large amount of future tax revenue, worsening the government savings rate in the years ahead.

Option 5: Create a Premium Savings Account.

A fifth reform option is the Premium Savings Account (PSA), a new approach designed to provide better marginal savings incentives than the current IRA. Bernheim and Scholz propose that each taxpayer would have to save some fixed "floor" amount, based on the taxpayer's income, before being allowed to contribute to a PSA. For each dollar of savings above the floor amount, the taxpayer could contribute one dollar to the PSA up to some "ceiling" amount, also based on the taxpayer's income. For example, a single taxpayer with an income of \$60,000 might have a floor of \$6,000 and a ceiling of \$8,000. By saving \$7,500 in a given year, the taxpayer could thus contribute \$1,500 to a PSA.¹⁵ Properly constructed, this system of floors and ceilings could maximize the number of households in each income class that would experience an increased marginal incentive to save.

But the effectiveness of a PSA program also remains uncertain. Although the PSA would provide stronger marginal savings incentives than the current IRA, the PSA still might have little effect if the private savings rate is insensitive to the after-tax rate of return. As a result, government savings still might decline by more than the increase in private savings and thus worsen the national savings rate. In addition, the PSA would pose administrative challenges because such a system would require a measure of savings that is not available from current tax returns. And it is unclear that economists understand consumer behavior well enough to design an effective schedule of ceilings and floors for the PSA.

The Premium Savings Account is therefore an intriguing proposal that—properly constructed—might provide better marginal savings incentives

than the other reform options. As such, the PSA clearly deserves further research and evaluation. But at this point, economists cannot be sure how such a system should be designed or whether it would really increase the national savings rate.

CONCLUSION

Because of the mixed results from past research, economists cannot say with certainty whether a broader IRA program could raise the national savings rate in the 1990s. But the national savings rate declined in the mid-1980s despite a substantial flow of funds into IRAs. Cross-section evidence also suggests that IRAs did not stimulate much new savings in 1982-86, the period of broadest IRA eligibility. IRAs were unsuccessful in the 1980s for three main reasons—the low sensitivity of savings to the after-tax rate of return, weak marginal savings incentives, and asset switching into IRAs from taxable financial assets. Although some of the reform options in the preceding section might stimulate more savings than the current IRA, none solves all of the economic problems that made IRAs ineffective in the 1980s. As a result, the reform options do not offer a reliable cure for the low national savings rate.

But this conclusion does not mean that fiscal policymakers can do nothing about the low national savings rate. Although policymakers may have weak tools for influencing the private savings rate, they have substantial influence over the government savings rate. As the second section showed, a decline in government savings is a major reason for the low national savings rate. Fiscal policymakers may not be able to eliminate fluctuations in government revenues caused by the business cycle, but they can improve the government balance sheet by reducing the budget deficit over time. The resulting increase in government savings should raise the national savings rate and ultimately improve U.S. living standards.

ENDNOTES

¹ Productivity growth slowed from an average of 2.4 percent annually in the 1960s to 1.3 percent in the 1970s and only 0.8 percent in the 1980s. Likewise, the growth of real net national product per person slowed from 2.7 percent annually in the 1960s to 1.7 percent in the 1970s and 1.3 percent in the 1980s.

² Net national product is gross national product minus a capital consumption allowance. The Joint Committee on Taxation of the U.S. Congress reports that the U.S. savings rate averaged 3.6 percent of gross domestic product in the 1980s. In contrast, the average savings rate in the 1980s was 8.4 percent in Canada, 10.2 percent in Germany, and 17.8 percent in Japan.

³ Even in an open economy, a low national savings rate may lower the domestic capital stock. For a large country such as the United States, a low savings rate may raise the real interest rate in world capital markets, thus depressing real investment spending. Moreover, some empirical evidence suggests that world capital markets are not fully integrated. Feldstein and Horioka find that, looking across a sample of industrial countries, a low domestic savings rate is associated with low domestic investment. Thus, even though the U.S. economy is open to international capital flows, the low savings rate may explain part of the slowing in productivity growth and real output growth in the last two decades.

⁴ The maximum individual contribution to an IRA is the lesser of \$2,000 or the individual's compensation. Taxpayers who are not covered by a pension plan may deduct their entire IRA contribution. For taxpayers covered by a pension plan, the deduction is gradually eliminated as adjusted gross income rises from \$40,000 to \$50,000 for families and from \$25,000 to \$35,000 for single persons. A taxpayer also may contribute up to \$250 to the IRA of a nonworking spouse. Deductible contributions are taxed when the funds are withdrawn from the IRA, but withdrawals of nondeductible contributions are not taxed.

⁵ Most taxpayers did not contribute to an IRA even during the period of broadest eligibility requirements. Only about 16 percent of all federal tax returns claimed an IRA deduction at the height of the program in 1985, and less than 6 percent of tax returns claimed an IRA deduction in 1989.

⁶ Although the measures of the savings rate in this section are widely used to discuss trends in savings, many economists believe these savings rates are subject to measurement errors. For example, these savings rates do not include capital gains or losses on household financial assets, even though such gains or losses could dramatically change house-

hold wealth. These savings rates also treat the purchase of a durable good, such as a car or refrigerator, as a form of consumption. Many economists argue that such purchases are really a form of savings because a durable good provides services to a household for several years rather than being consumed all at once. Bovenberg and Evans examine these issues in greater detail, however, and conclude that the decline in the savings rate cannot be attributed solely to measurement errors.

⁷ Private savings is personal savings plus undistributed corporate profits. Personal savings is simply disposable income, household income after tax payments, minus personal consumption expenditures. Undistributed corporate profits include inventory valuation and capital consumption adjustments. The private savings rate is private savings divided by the net national product.

⁸ Some proponents of IRAs argue that conventional economic theory does not capture the full savings incentive from IRAs because such theory ignores important behavioral effects (Shefrin and Thaler; Thaler). For example, Venti and Wise (1992) hypothesize that widespread promotion of IRAs by financial institutions in the early 1980s caused households to pay more attention to their retirement needs, stimulating a greater increase in savings than conventional theory suggests. But fiscal policymakers should not, at present, place much confidence in behavioral analyses of IRAs. Such theories are relatively new and untested, unlike conventional theory which has been applied usefully to many policy issues. Behavioral theories of savings do not yet provide a reliable basis for analyzing the effect of IRAs on private savings.

⁹ Bosworth and Bovenberg provide brief surveys of empirical research on the sensitivity of private savings to the after-tax rate of return. The wide range of estimates prompts both authors to conclude that no consensus exists on the magnitude of this effect. But the high degree of uncertainty is another reason why expanded IRA eligibility is not a reliable way to raise the national savings rate. Fiscal policymakers should concentrate on policy actions where the effects on private behavior can be anticipated with greater certainty.

¹⁰ Feldstein argues that previous analyses of IRAs overstated their adverse effect on government tax revenues by ignoring a positive effect on corporate tax payments. If IRAs raise the savings rate, a higher stock of corporate capital will lead to larger profits and higher corporate tax receipts. Thus, Feldstein asserts that the revenue loss from IRAs is much smaller than was previously estimated, and may even be a revenue gain over some time horizons. But Feldstein's arguments hold only if IRAs raise the national savings rate

because savings must first increase to produce a larger stock of corporate capital. As this article shows, such a positive effect of IRAs on national savings is very much open to dispute.

¹¹ In the Bush administration's proposal, the flexible IRA would be available to single persons earning up to \$60,000 per year and married couples earning up to \$120,000. Withdrawals of investment earnings within three years of the initial contribution would face both income taxation and a 10 percent penalty. Withdrawals of earnings between three and seven years after the contribution would be taxed at the regular income tax rate but would not face an additional penalty.

¹² The Clinton administration has, however, proposed an increase in the top personal income tax rate from 31 percent to 36 percent. In addition, people with taxable incomes over \$250,000 would face an additional 10 percent surtax, resulting in an effective income tax rate of about 40 percent. But personal income tax rates would generally remain lower than in the early 1980s.

¹³ For some higher income taxpayers with pension coverage, the availability of both deductible IRAs and 401(k) plans might increase the marginal incentive to save. Suppose the taxpayer currently has an income level too high to make deductible IRA contributions. Also suppose the taxpayer wants to save \$6,000 but can only contribute \$5,000 to a 401(k) plan because of rules restricting contributions by highly compensated employees. Earnings on the last \$1,000 of savings would, thus, be taxed at the regular income tax rate.

If the taxpayer becomes eligible for a deductible IRA with a maximum contribution of \$2,000, earnings on the last \$1,000 of savings could be sheltered from taxes until retirement. As a result, the after-tax return on an additional dollar of savings increases, which might induce the taxpayer to save somewhat more than \$1,000 outside the 401(k) plan.

Some higher income taxpayers may, therefore, experience a stronger incentive to save if both 401(k) plans and deductible IRAs are available. But there still might not be enough new private savings to offset the loss in government savings from this reform option. Moreover, higher income taxpayers without a pension plan would not experience any increase in their savings incentives because such taxpayers can already make deductible IRA contributions.

¹⁴ In this case, the back-loaded IRA and the current IRA would offer the same expected after-tax returns to the saver over the life of the investment. This equivalence requires that the marginal income tax rate be the same at the time of the contribution and at the time of withdrawal. Ozanne examines cases where the marginal tax rate varies over time and argues that a back-loaded IRA may provide stronger incentives than the conventional IRA for nonretirement saving, such as saving for the downpayment on a home.

¹⁵ The maximum contribution to the PSA for this taxpayer in any given year would be \$2,000, which could be made only if the taxpayer saved \$8,000 or more. A single taxpayer with a smaller income would face lower floor and ceiling amounts. For example, a taxpayer with an income of \$50,000 might have a floor of \$2,700 and a ceiling of \$4,700.

REFERENCES

- Bentsen, Lloyd, and Bill Roth. 1992. "The IRA: Saving Grace," *The Washington Post*, July 30.
- Bernheim, B. Douglas, and John Karl Scholz. 1992. "Private Saving and Public Policy," in James M. Poterba, ed., *Tax Policy and the Economy*. National Bureau of Economic Research Conference Report, November 17.
- Boskin, Michael J. 1978. "Taxation, Saving, and the Rate of Interest," *Journal of Political Economy*, part 2, April.
- Bosworth, Barry P. 1984. *Tax Incentives and Economic Growth*. Washington: Brookings Institution.
- Bovenberg, A. Lans. 1989. "Tax Policy and National Saving in the United States: A Survey," *National Tax Journal*, June.
- _____, and Owen Evans. 1990. "National and Personal Saving in the United States: Measurement and Analysis of Recent Trends," *IMF Staff Papers*, September.
- Deaton, Angus. 1987. "Comment," in Martin Feldstein, ed., *The Effects of Taxation on Capital Accumulation*. National Bureau of Economic Research. Chicago: University of Chicago Press.
- Feldstein, Martin. 1992. "The Effects of Tax-Based Saving Incentives on Government Revenue and National Saving," National Bureau of Economic Research working paper no. 4021, March.
- _____, and Charles Horioka. 1980. "Domestic Saving and International Capital Flows," *The Economic Journal*, June.
- Friend, Irwin, and Joel Hasbrouck. 1983. "Saving and After-Tax Rates of Return," *The Review of Economics and Statistics*, November.
- Gale, William G., and John Karl Scholz. 1992. "IRAs and Household Saving," unpublished manuscript, July.
- Gravelle, Jane G. 1991. "Do Individual Retirement Accounts Increase Savings?" *Journal of Economic Perspectives*, Spring.

- Hall, Robert E. 1988. "Intertemporal Substitution in Consumption," *Journal of Political Economy*, April.
- Joines, Douglas H., and James G. Manegold. 1991. "IRAs and Saving: Evidence from a Panel of Taxpayers," Federal Reserve Bank of Kansas City, Research Working Paper no. 91-05, October.
- Ozanne, Larry J. 1992. "Past Experience and Current Proposals for IRAs," in Marvin H. Koters, ed., *Personal Saving, Consumption, and Tax Policy*. Washington: AEI Press.
- Shefrin, Hersh M., and Richard H. Thaler. 1988. "The Behavioral Life-Cycle Hypothesis," *Economic Inquiry*, October.
- Thaler, Richard H. 1990. "Anomalies: Savings, Fungibility, and Mental Accounts," *Journal of Economic Perspectives*, Winter.
- U.S. Congress, Joint Committee on Taxation. 1991. *Description and Analysis of S. 612 (Savings and Investment Incentive Act of 1991)*, Staff Report. Washington: U.S. Government Printing Office.
- U.S. Department of the Treasury, Internal Revenue Service. 1992. *Individual Income Tax Returns 1989*. Washington: U.S. Government Printing Office.
- Venti, Steven F., and David A. Wise. 1992. "Government Policy and Personal Retirement Saving," in James M. Poterba, ed., *Tax Policy and the Economy*, vol. 6. National Bureau of Economic Research. Cambridge, Mass.: MIT Press.
- _____. 1990. "Have IRAs Increased U.S. Saving?: Evidence from Consumer Expenditure Surveys," *Quarterly Journal of Economics*, August.

Asymmetric Effects of Monetary Policy

By Donald P. Morgan

Does tight monetary policy slow the economy more than easy monetary policy accelerates the economy? Recent U.S. experience suggests that may be the case. When monetary policy was tight in 1988 and 1989, the economy seemed to slow in response. Yet when monetary policy was eased in 1990, the economy did not respond accordingly.

The suggestion that monetary policy has such asymmetric effects is not altogether new or unorthodox. Indeed, mainstream economists adopted this view for several decades after the Great Depression, when easy monetary policy seemed powerless to revive the economy. Recent studies have revived interest in the asymmetric effects of monetary policy. Theoretical research has suggested reasons why tight policy may have more impact than easy policy. And empirical studies, which use monetary growth to identify the stance of policy, have produced evidence of asymmetry.

Of course, not all changes in monetary growth may reflect changes in policy. Recent growth in M2, for example, has been very slow even though monetary policy, judging from the statements of policymakers themselves, has been decidedly

easy. Moreover, various measures of money can give contradictory indications of the stance of monetary policy. So studying only the monetary aggregates may lead to the wrong conclusions about the stance and the impact of policy.

This article looks for evidence that monetary policy has asymmetric effects using two alternative measures of the stance of policy: (1) the federal funds rate and (2) a narrative index based on the statements of policymakers. The article finds some evidence of asymmetry using both measures of policy. The first section of the article traces the history and possible causes of asymmetry. The second section presents some evidence that the impact of monetary policy is asymmetric.

ROOTS OF ASYMMETRY

The notion of asymmetry was born in the Great Depression. That event convinced many that easy policy was powerless against recessions, even if tight policy could check a boom. One reason heard then for asymmetry is still heard today: a loss of confidence by firms and consumers during recessions makes monetary policy less effective. Two other reasons given today are credit constraints that augment only tight policy, and prices that are less flexible downward than upward.

Donald P. Morgan is a senior economist at the Federal Reserve Bank of Kansas City. Steve Ravel, a research associate at the bank, helped prepare the article.

History of asymmetry

Economists have entertained the possibility of asymmetry off and on for decades. Until the end of the 1920s, most believed the impact of monetary policy was symmetric. It was thought that policymakers controlled a lever that could lower or raise the level of economic activity equally well. By raising interest rates, the Federal Reserve could slow the economy, and by lowering rates, the Fed could stimulate the economy. The apparent success of monetary policy in turning around mild recessions in 1924 and 1927 bolstered faith in the effectiveness of easy policy (Hansen).

This faith in easy policy was shaken by the Great Depression, which convinced many economists that only tight policy was effective. After the economy turned down in 1929, short-term nominal interest rates soon declined to less than 1 percent. The low level of interest rates convinced the Federal Reserve it was pursuing an easy monetary policy. Yet the Depression persisted until 1934, leading many to conclude that easy policy was futile, "like pushing on a string."¹

According to monetary historians, the notion of asymmetry was widely adopted in the 1940s and 1950s (Mayer, Johnson). This view was also held by some within the Federal Reserve. A vice president of the Federal Reserve Bank of New York, for example, began an article in 1951, stating, "Two decades ago it still bordered on heresy to suggest that central bank control over interest rates was useless . . . Today that heresy has become widely accepted as dogma" (Roosa, p. 1).

Belief in asymmetry diminished in the 1960s and 1970s, after Friedman and Schwartz reexamined the monetary history of the Great Depression. Their analysis showed that monetary policy was not easy in the early 1930s, but actually was tight. So rather than proving monetary policy was impotent, they argued, the Great Depression was "tragic testimonial" to its power. This observation seemed to effectively weaken the primary evidence that easy policy was ineffective. By 1969,

Mayer concluded: "Belief in asymmetry of monetary policy has lost much of its support" (p. 150).

The loss of support for asymmetry, however, may have been premature. After all, Friedman and Schwartz's observation about the Great Depression proved only that tight policy was effective, not that easy policy was equally effective. And the recent sluggish recovery has once again shaken some observers' faith in the power of easy policy.² Moreover, recent theoretical research suggests reasons why easy policy may be less effective than tight policy.

Reasons for asymmetry

Changing outlook. One reason monetary policy might have asymmetric effects is because business and consumer confidence changes over the business cycle. This changing outlook could lead to asymmetry if firms and consumers are more pessimistic during recessions than they are optimistic during booms, or if the outlook of firms and consumers simply matters more during recessions.

It is easy to see how pessimism may thwart easy policy. If firms have a dim outlook on their business prospects, lower interest rates may not stimulate borrowing and investment. Similarly, if employment prospects are bleak, lower interest rates will not boost consumer spending on durables.

Pessimism was, and still is, a commonly invoked reason why easy policy may be weak. During the Great Depression, pessimism was thought to discourage borrowing and lending, thus countering the supposedly easy policy. The saying then was "You can lead a horse to water . . ." And just a year ago, Federal Reserve Board Chairman Greenspan pointed to the low "state of consumer and business confidence" as a reason the economy had responded sluggishly to easy policy (p.1).

Yet pessimism by itself cannot explain why easy policy is less effective than tight policy. During booms, the outlook of firms and consumers

changes from pessimism to optimism. And optimism presumably weakens the impact of tight monetary policy, just as pessimism weakens the impact of easy policy.

For a changing outlook to explain asymmetry, pessimism must weaken easy policy more than optimism weakens tight policy. This requirement would be satisfied if the outlook itself changes asymmetrically over the business cycle, that is, if firms and consumers are more pessimistic during recessions than they are optimistic during booms. This requirement would also be satisfied if the outlook simply matters more during recessions, meaning that firms and consumers worry more about the outlook and pay less attention to interest rates during recessions than during booms.

While both of these conjectures about business psychology might be true, they are still weak reeds to rely on for explaining asymmetry. There are, however, more compelling reasons.

Credit constraints. Monetary policy may also be asymmetric due to the interaction of credit constraints and the demand for credit. Credit constraints arise if tight policy makes banks less willing to lend to some borrowers.

Tight monetary policy can lead to credit constraints if banks are unwilling to lend to riskier borrowers when market rates are high. As tight policy pushes up market rates, banks' cost of funds increases because banks must raise deposit rates along with market rates. All else equal, banks would simply pass on the higher cost of funds to borrowers by raising loan rates. But all else is not equal because higher loan rates can increase the risk of bankruptcy by increasing the borrower's obligation to the bank. If higher loan rates threaten to increase bankruptcy risk too much, banks may ration the quantity of credit available to riskier borrowers, leaving them credit constrained.

The interaction of such constraints and the demand for credit can skew the impact of monetary policy. By driving up market rates, tight monetary policy tightens the credit constraint on

some borrowers. Given a growing economy and strong demand for credit, the constraint effectively limits spending by these borrowers. This *binding* constraint augments the impact of tight monetary policy, leading to a larger decline in borrowing and spending than would result from higher market interest rates alone. On the other hand, easy policy relaxes credit constraints by lowering market rates. Relaxing the constraints, however, will not necessarily boost borrowing and spending if a slowing economy has reduced the demand for credit. In other words, if the credit constraint is no longer binding before policy is eased, relaxing the constraint will not augment easy policy.³

This reasoning suggests that if the credit constraints bind only when policy is tight and the demand for credit is strong, then tight policy will be more powerful than easy policy.⁴ This explanation of asymmetry seems more compelling than the changing outlook story because it has been sketched out in theoretical models and has some indirect empirical support.⁵

Prices less flexible downward. Another explanation for asymmetry involves the relative flexibility of prices. Monetary policy will have asymmetric effects on real output if prices are less flexible downward than upward. In that case, tight policy will cause output to fall with little change in prices, while easy policy will cause prices to rise with little change in output. Recent theoretical research suggests that prices may be inflexible downward because firms are already inclined to raise their prices to keep up with trend inflation regardless of whether policy is tight or easy.

These theories all begin with the assumption that it is costly for firms to adjust prices simply because firms must issue new price lists to their customers. Such "menu" costs make it expensive for firms to continuously adjust prices to their desired level. To economize, firms will set an initial price based on expected growth in spending for some period, say, a year. In the meantime, firms' desired price may change because of unexpected changes in spending. Firms may therefore

choose to adjust prices periodically when their initial price is too far from their desired price.

Trend inflation in this scenario will lead firms to adjust their prices asymmetrically in response to changes in monetary policy. Suppose policy was eased to a certain degree during the year, causing spending to rise. Although the higher spending would increase a firm's desired price above its initial price, the difference might be too small to justify the cost of adjusting prices. Trend inflation also increases the desired price, however, because firms wish to keep their individual prices in line with the general price level. The higher spending, together with inflation, might very well lead firms to increase prices rather than output.⁶

Reversing this logic reveals how trend inflation makes firms less likely to lower prices when monetary policy is tightened to the same degree. Although the resulting decline in spending inclines firms to lower prices, inflation inclines them to raise prices. The overall effect may wash, so that instead of cutting prices in response to lower spending, firms reduce output. Trend inflation can therefore cause prices to be less flexible downward than upward, which in turn leads to asymmetric effects of policy on output.

This explanation of asymmetry seems the most compelling both theoretically and empirically. While the discussion above simply traces the story, the explanation has been fully developed in several models (Tsiddon; Ball and Mankiw; Caballero and Engel). And empirical tests support the models. Caballero and Engel, for example, gathered data on 37 countries with low to moderate inflation rates. They found that as inflation rises across countries, the impact on output of unexpected increases in spending falls, just as the models predict.⁷ Caballero and Engel, however, studied the effect of spending changes in general, not just changes due to monetary policy.⁸ So the question remains whether monetary policy in particular has asymmetric effects.

EVIDENCE OF ASYMMETRY

Evidence that monetary policy has asymmetric effects was found by recent studies using the monetary aggregates to identify the stance of policy. Additional evidence is presented here when policy is identified using the federal funds rate and an index based on policymakers' statements.

Evidence using the aggregates

A recent study by Cover found evidence that monetary policy has asymmetric effects. Using quarterly data beginning in 1949, Cover estimated the effect of changes in M1 on output growth.⁹ He discovered that declines in money growth usually had a substantial and statistically significant effect on output. In contrast, he found that increases in money growth usually had a small and statistically insignificant effect on output.

Others have subsequently researched Cover's findings using different sample periods and different data. One researcher found the same results even if the sample excluded the Volcker era of 1979 to 1987, which included a period of severely tight policy.¹⁰ Delong and Summers found similar results using annual data back to the turn of the century. This author found very similar results using the broader monetary aggregates, M2 and M3, which sometimes—as is the case recently—diverge from the narrower M1 aggregate.¹¹ These various extensions suggest Cover's results do not merely reflect his choice of sample period, data frequency, or monetary aggregate.

A potential problem remains, however. All these extensions identify the stance of policy using the monetary aggregates, which can be misleading because not all changes in monetary growth reflect changes in policy. Current slow growth in M2 and M3, for example, may largely reflect portfolio shifts to higher yielding stock and bond funds, rather than tight policy. And over the longer sample periods covered in the studies just discussed,

variation in money growth could simply reflect changes in current output rather than changes in policy. If so, estimates of the impact of money growth on output will be inaccurate because the estimates will also measure the reverse impact of output on money growth. This identification problem has led researchers to consider alternative measures of the stance of policy.

Evidence using alternative policy measures

Some researchers now advocate using the federal funds rate to identify the stance of policy (Bernanke and Blinder). Others are returning to an older, narrative approach in which the stance of policy is identified by the statements of policymakers themselves (Romer and Romer).

The federal funds rate. The funds rate is a natural way to identify the stance of policy because policymakers have targeted this rate on and off for the last 30 years. Changes in the funds rate appear to have an asymmetric impact on output, just as changes in money growth do. The evidence is weaker, though, when the sample excludes the period in the early 1980s when policymakers dropped the funds rate target.

Measuring the stance of policy directly with the federal funds rate raises problems, however, because not all changes in the funds rate reflect changes in policy. For example, if the economy is growing rapidly and other market rates are on the rise, policymakers may let the funds rate drift up in step, even though they are not actively tightening policy. Similarly, policymakers may let the funds rate drift up with inflation, even though they are not tightening policy.

Identifying changes in the funds rate that reflect changes in policy calls for a two-stage regression procedure.¹² In the first stage, the level of the funds rate is regressed on its own lagged values, on current and lagged values of output growth and inflation, plus a constant and a trend variable.¹³ The variations in the funds rate not

explained by those variables, the residuals, are used to identify the stance of policy. The positive residuals represent tight policy because the residuals measure how much the current funds rate exceeds the level predicted by current and lagged values of output and inflation. The negative residuals represent easy policy, when the funds rate is lower than would be expected given current and lagged values of output and inflation.

In the second-stage regression, the real growth rate of output (the change in the log) is regressed on a constant and trend variable, lagged output growth, and lagged values of positive and negative residuals of the funds rate as shown in the equation below.

$$\Delta \log Y_t = a + b * Trend + \sum_{i=1}^8 C_i * \Delta \log Y_{t-i} + \sum_{i=1}^8 d_i * FF_{t-1} + \sum_{i=1}^8 e_i * FF_{t-i} + n_t.$$

The cumulative impact of policy on output is measured by the sum of the coefficients on the positive and negative residuals of the funds rate. Because changes in the funds rate in one direction tend to move output in the opposite direction, the estimated sums are expected to be negative.¹⁴ The equation was estimated with quarterly data over two sample periods. The full sample spanned 1963:2 through 1992:3. The second sample excluded 1979:4-1982:4, the period in which the Federal Reserve deemphasized the funds rate.

Changes in the funds rate clearly had an asymmetric impact on output over the full sample period (Table 1). The impact of increases in the funds rate is large and highly significant. The impact of decrease in the funds rate is small and insignificantly different from zero. And as the bottom line indicates, the difference in the impact is statistically significant.

The evidence of asymmetry is considerably weaker when the sample period excludes 1979:4-1982:4. Over this subsample, both increases and

Table 1

Impact of the Federal Funds Rate on Output

	Full sample period 1963:2 - 1992:1	Excluding 1979:4 - 1982:4
FF ⁺	-1.09*** (.33)	-1.75*** (.52)
FF ⁻	-.08 (.32)	-1.00** (.50)
FF ⁺ - FF ⁻	-1.01** (.48)	-.75 (.68)

Notes: Shown are sums of coefficients from regression of output growth on positive and negative residuals of federal funds rate. Residuals are calculated from a first-stage regression of the funds rate on constant, trend, eight lags of funds rate, and current and eight lags of output growth and inflation (see text for explanation). Standard errors are in parentheses. ** and *** indicate significance at 5 percent and 1 percent, respectively.

decreases in the funds affect output significantly and the impacts do not differ significantly. Arguably, however, there is some evidence of asymmetry even in this subsample. Increases in the funds rate are *more* significant than are decreases, and increases also appear to have a much larger impact, even if the difference is not statistically significant.¹⁵

Because asymmetry can be a visual concept, the empirical results are depicted in chart form. Chart 1 plots the cumulative impact on output of a change in the funds rate of one percentage point for one quarter.¹⁶ For the full sample period, an increase in the funds rate has a substantial and persistent effect on output (Panel A). One year after the increase, output is 0.6 percent below its initial level, and after two years output is more than 1 percent lower. A decrease in the funds rate, in comparison, has a small and temporary effect on output. One year after the decrease, output is less than 0.4 percent above its initial level, and

after two years output returns to its initial level.

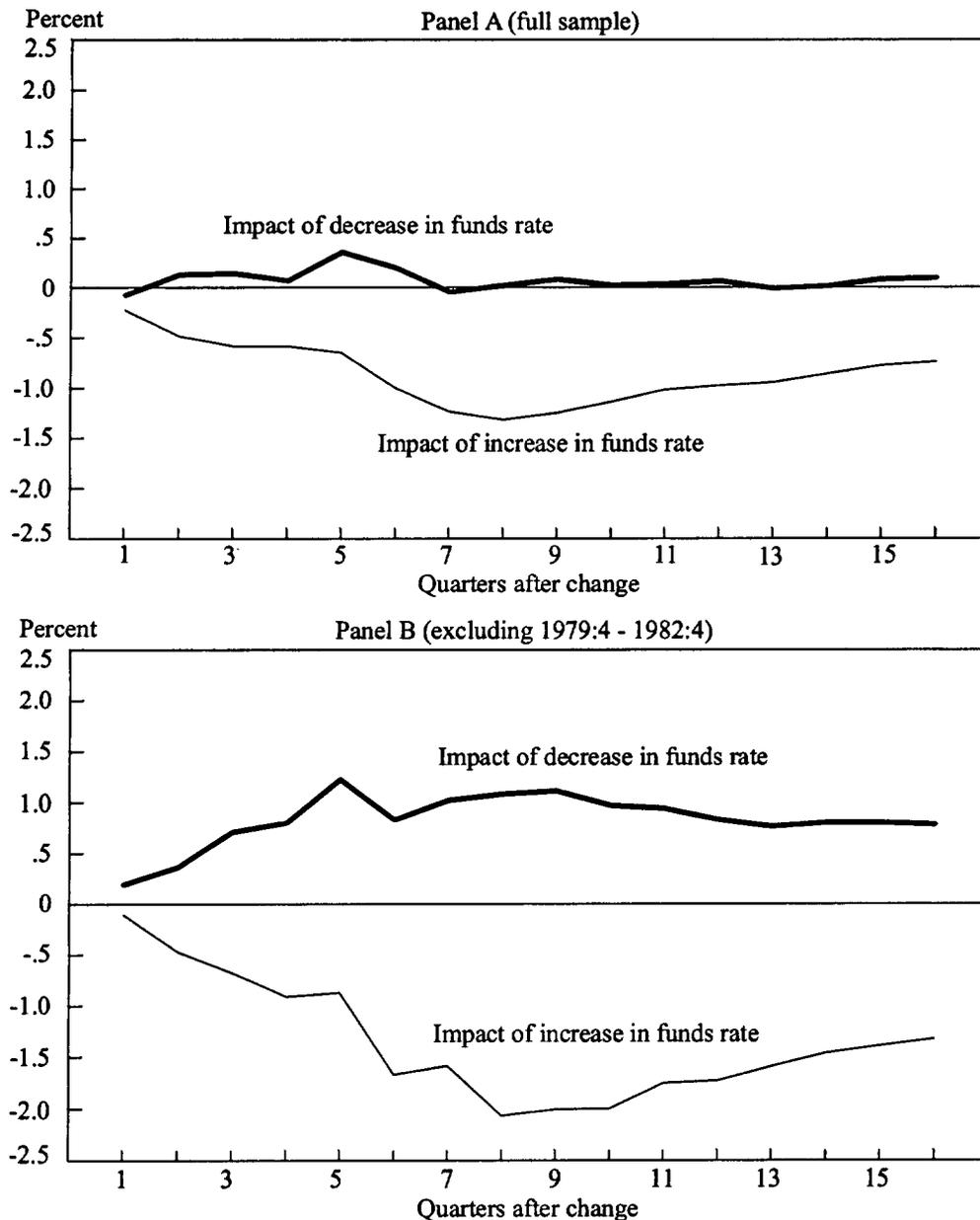
When the sample excludes 1979:4-1982:4, the asymmetry is less striking (Panel B). As with the full sample, increases in the funds rate seem to have a large impact, with output falling 2.0 percent below its initial level two years after the increase. Decreases, however, have a larger and more lasting impact on output compared to the full sample. The peak impact occurs a little more than a year later, when output is almost 1.5 percent below its initial level.

Thus, the pattern of results using the funds rate to identify policy, though not uniformly strong, tends to reinforce the evidence of asymmetry using the monetary aggregates.¹⁷ The following section uses an altogether different narrative approach to identify policy, based on policymakers' statements.

The narrative approach. This approach requires researchers to read the recorded statements of policymakers and to index the stance of policy according to those statements. Using such

Chart 1

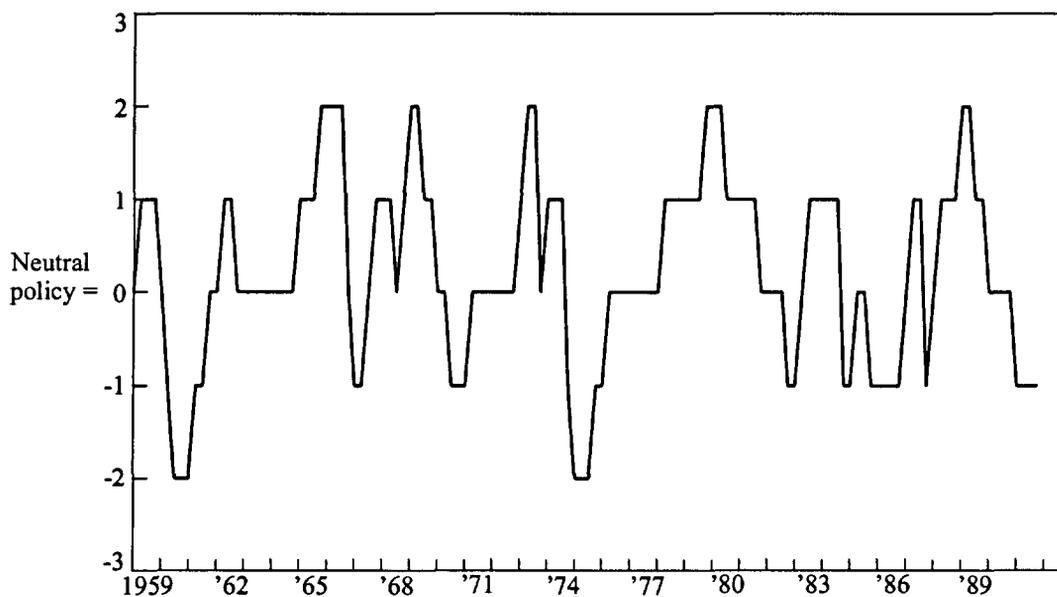
Asymmetric Impact of Changes in the Federal Funds Rate on Output



Note: Shown is the cumulative change in output from its initial level following a one percentage point change for one quarter in the federal funds rate. See text and endnote 16 for explanation.

Source: Author's calculations.

Chart 2

Boschen-Mill Index of Monetary Policy

Note: Value of index at end of quarter. 2 = very tight, 1 = tight, -1 = easy, -2 = very easy. Also see endnote 19.
Source: Boschen and Mill.

an index recently constructed by John Boschen and Leonard Mill, this section finds results similar to those above.

To construct their index, Boschen and Mill perused the policy records of the Federal Open Market Committee (FOMC) from 1953 through 1991.¹⁸ Based on their reading of the record, the authors classified the stance of policy into five categories: very tight, tight, neutral, easy, and very easy. They assigned each category a respective value of 2, 1, 0, -1, and -2. The Boschen-Mill index is plotted in Chart 2.¹⁹

The narrative approach is less subjective than it appears. Boschen and Mill found that their index was significantly correlated with five other indexes constructed by different researchers over

the last several years. This finding implies that different readers of the policy record form the same impression about the stance of monetary policy. So despite its apparent subjectivity, this approach should be a useful alternative to the funds rate or the aggregates in investigating the impact of monetary policy.

The investigation still requires the two-stage procedure used with the federal funds rate. The same procedure is necessary because monetary policy, even measured by the index, often responds to changes in output. So in order to accurately estimate the impact of policy on output, it is first necessary to identify changes in policy that are not due to changes in output.²⁰ As before, this identification is accomplished with a first-stage

Table 2

Impact of Boschen-Mill Index on Output

	Full sample period 1963:2 - 1992:1	Excluding 1979:4 - 1982:4
BM ⁺	-1.58* (.87)	-2.24*** (.74)
BM ⁻	-.21 (.93)	1.78 (1.12)
BM ⁺ - BM ⁻	-1.79 (1.55)	-4.02*** (1.49)

Notes: Shown are sums of coefficients from regression of output growth on positive and negative residuals of the Boschen-Mill index. Residuals are calculated from first-stage regression of the index on a constant, trend, eight lags of index, and current and eight lags of the output growth and inflation (see text for explanation). Standard errors are in parentheses. * and *** indicate significance at 10 percent and 1 percent.

regression of the index on its own lagged values, plus current and lagged values of output and inflation, which might be correlated with the index and with output.²¹

In the second stage, output growth is regressed on positive and negative residuals from the first-stage regression, which represent tight policy and easy policy. The second-stage regression, which took the same form as the equation above, was also estimated over two sample periods: the full sample, 1963:2-1992:2, and the subsample that excluded 1979:4-1982:4.²²

Policy also appears to have asymmetric effects when measured by the Boschen-Mill index (Table 2). Over the full sample, increases in the index had a large and marginally significant impact on output, while decreases had an insignificant impact. The bottom line, however, indicates that the impacts do not differ significantly.

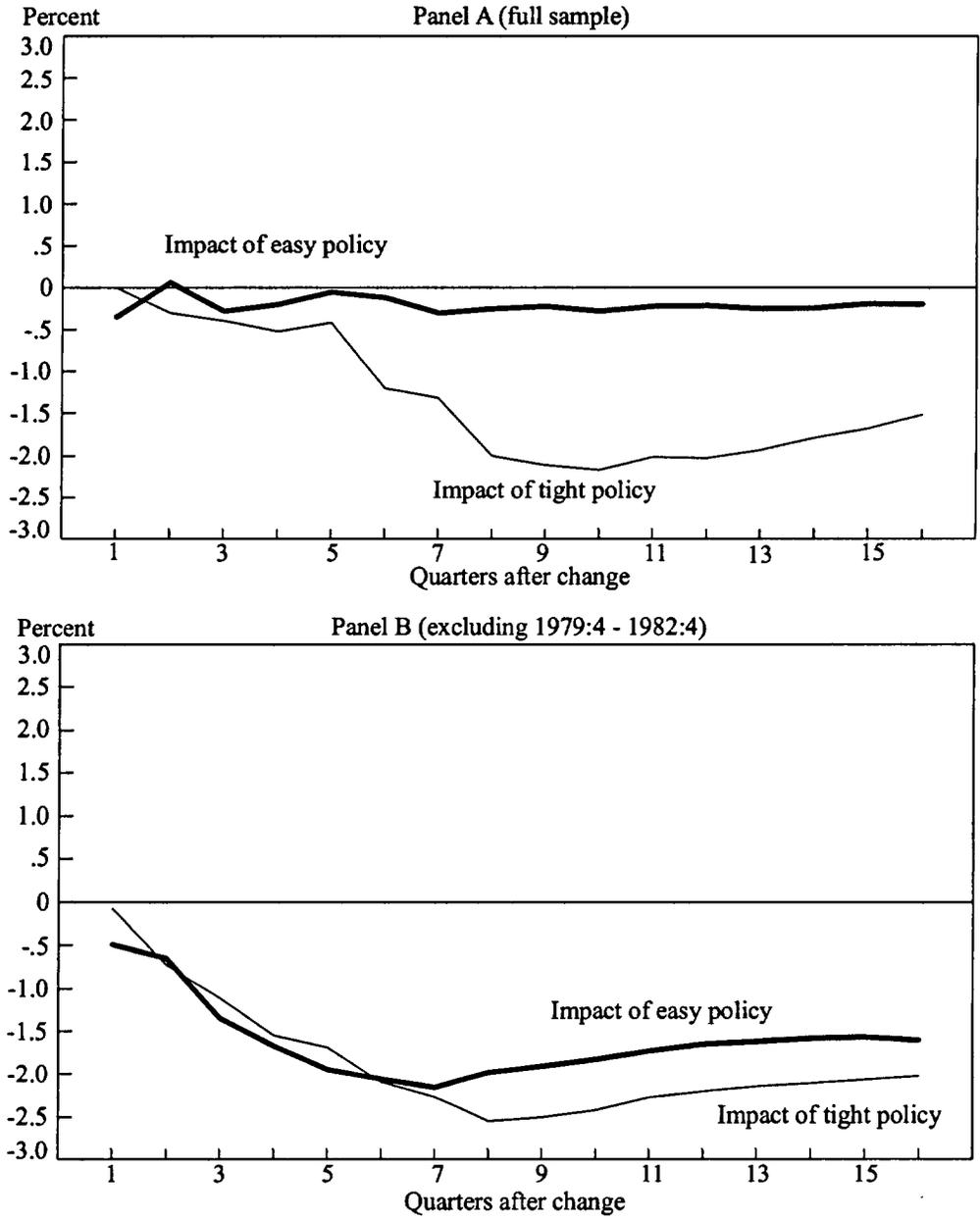
Unlike with the funds rate, the evidence of asymmetry is stronger when the sample excludes

1979:4-1982:4. Increases in the index had a very significant impact on output, while decreases in the index had an insignificant impact. Moreover, the difference in the impact is highly statistically significant. The stronger evidence with the index over the subsample is notable because the evidence with the funds rate was weaker over this sample period.

A picture illustrates the asymmetric impact of changes in the Boschen-Mill index on output (Chart 3). Over the full sample (Panel A), a one-unit increase in the index for one quarter, representing tight policy, has a dramatic and lasting effect on output. In contrast, a one-unit decrease, representing easy policy, has virtually no effect on output. When the sample excludes 1979:4-1982:4, an increase in the index reduces output dramatically (Panel B). Although a decrease in the index appears to reduce output, recall that the impact of a decrease in the index was insignificantly different from zero in Table 2.

Chart 3

Asymmetric Impact of Changes in the Boschen-Mill Index Output



Note: Shown is the cumulative change in output from its initial level following a one percentage point change for one quarter in the federal funds rate. See text and endnote 16 for explanation.

Source: Author's calculations.

SUMMARY

The view that monetary policy has asymmetric effects, once the mainstream view after the Depression, has undergone a revival. Recent experience in the United States suggests that easy policy may be weaker than tight policy, and recent theory suggests reasons why this may be so. Empirical research, using the monetary aggregates to identify policy, has produced evidence of asymmetry.

The results presented in this article provide some additional evidence of asymmetry using two

alternative measures of policy: (1) the federal funds rate and (2) a narrative index of policy. Although this evidence is not equally strong across measures and across sample periods, the same patterns recur. Tight monetary policy,¹ however measured, substantially and significantly reduced output in either sample period, while easy monetary policy usually had an insignificant effect on output. These results, together with recent theory and experience, make a case for further research on the possibility of asymmetric effects of monetary policy.

ENDNOTES

¹ Alfred Hansen, a prominent Harvard economist of that era wrote, "The monetary weapons can, indeed, be applied effectively to check an expansion" (p. 71). Later, in the same chapter on monetary policy in the Depression, he concluded, "But the decade of the thirties offers abundant evidence that cheap money alone is not adequate" (p. 82).

² Of course, the slow growth of M2 in the last several years leads monetarists to conclude that policy has been tight rather than easy during the current recovery.

³ For convenience, the text assumes that banks quantity ration credit at some point instead of raising loan rates. The equilibrium quantity of loans in that case is simply the quantity supplied at the point where the loan supply curve bends backward (Keeton). When policy is tightened, the equilibrium quantity of loans falls by the full amount of the shift in loan supply. Of course, easy policy presumably returns the loan supply curve to its original position. Nevertheless, the equilibrium quantity of loans will be less than before policy was tightened if loan demand has fallen in the meantime, and now intersects the original loan supply curve below the point where it bends backward, i.e., if there is no longer credit rationing at the new equilibrium.

⁴ Quantity rationing seems sufficient but not strictly necessary to explain asymmetry. The same information problems that lead to quantity rationing will lead banks to charge a steep risk premium over market interest rates for a given loan size. Even if firms can borrow at that rate, they may still be "credit constrained" because the premium will lead firms to forego profitable spending that cannot be financed internally (Fazzari, Hubbard, and Peterson). Moreover, the risk premium will increase with the size of the loan, implying a convex loan supply curve. As a convex loan supply curve

shifts back and forth (due to changes in policy), the equilibrium quantity of loans changes more when loan demand is high than when loan demand is low.

⁵ The discussion in the text relies on the interaction of credit constraints and the demand for credit to explain asymmetric policy effects. A closely related idea is developed in a model by Jackman and Sutton, in which a credit constraint causes asymmetry through permanent income effects. Higher interest rates force constrained consumers to reduce spending by the full amount that their loan payments increase. Lower rates relax the constraint, but spending increases less than proportionately because consumers spread out their spending across time. A somewhat related idea is suggested in Bernanke and Gertler's model, in which firms may be credit constrained because of low collateral. When firms are fully collateralized, and hence unconstrained, sharp declines in investment spending are more likely than sharp increases. Indirect evidence of such asymmetries comes from Kashyap, Lamont, and Stein, who find that firms' spending on inventory appears credit constrained only when policy is tight. More indirect evidence comes from Gertler and Gilchrist, who find that shocks to the federal funds rate have larger effects on spending (particularly by small firms) when output growth is below average.

⁶ This discussion is partial equilibrium in the sense that inflation is taken as given in analyzing an individual firm's pricing decisions. The studies being discussed, however, determine inflation endogenously as a function of individual firm's decisions.

⁷ More precisely, the model predicts that the degree of asymmetry increases with the average inflation rate across countries: the impact of negative spending shocks increases

with inflation *and* the impact of positive spending shocks decreases with inflation. This is precisely what Caballero and Engel found. Ball and Mankiw's model predicts that the distribution of real growth rates of output will be skewed to the left because output is more likely to fall after a decline in spending than it is likely to rise after an increase in spending. They cite evidence of such skewness from Sichel in support of this prediction.

⁸ And because the degree of asymmetry increases with the inflation rate, the question remains whether policy has asymmetric effects in the United States, a country with a relatively low inflation rate.

⁹ To be precise, Cover estimated the impact on output of *unexpected* changes in money: changes not predicted by lagged output, interest rates, and other variables. Using unexpected changes is a way of excluding changes in money growth due to changes in demand, rather than changes in supply.

¹⁰ The anonymous referee of Cover's article at the *Quarterly Journal of Economics* found the same result excluding the Volcker period.

¹¹ The results are available upon request.

¹² Controlling for feedback between policy and output is usually accomplished with a vector autoregression (VAR). The nonlinearity implicit in this model precludes straightforward use of a VAR. The two-stage procedure used here, however, is akin to a VAR in which output is ordered before policy. Cover also used a two-stage procedure. However, he excluded current output growth from the first stage and included current money shocks in the second stage. This is equivalent to ordering policy before output in a VAR, which is not the usual prior. Moreover, it could have biased his results because money demand very likely depends on current income.

¹³ The first-stage regression was:

$$FF_t = a + b * trend + \sum_{i=1}^8 b_i * FF_{t-i} + \sum_{i=0}^8 C_i * \Delta \log Y_{t-i} + \sum_{i=0}^8 d_i * inflation_{t-i} + e_t$$

The sample period was 1961:2-1992:3. Using the change in the funds rate instead of the level did not alter the results. Including the deficit in the first-stage regression tended to strengthen the results. That specification was not reported, however, in order to maintain symmetry with the first-stage specification of the Boschen-Mill regression (to follow), in which the deficit seemed not to belong.

¹⁴ Herein, the terms "increases" and "decreases" are used to refer to positive and negative residuals of the first-stage regression and not simply to positive and negative changes in the observed series.

¹⁵ That the evidence of asymmetry is weaker when the sample excludes 1979:4-82:4 may or may not be surprising, depending on one's priors. Given the evidence over the full sample, one might have expected even stronger evidence after excluding a period in which the Federal Reserve was clearly not targeting the funds rate (implying that funds rate shocks did not necessarily represent changes in policy). On the other hand, one may have expected weaker evidence after excluding a period in which the funds rate was allowed to fluctuate more than ever before.

¹⁶ For simplicity, the federal funds rate is assumed to change for only one quarter, after which it returns to its initial level. Even such a temporary change in the funds rate will have persistent effects on output growth because, according to the regression equation, the past level of the funds rate affects the current growth rate of output. The impact of a change in the funds rate is calculated using the estimated coefficients in the regression equation and includes the indirect effects due to lagged changes in output growth. This "dynamic multiplier" is not to be confused with an impulse response function.

¹⁷ Roughly the same pattern of results was obtained if the deficit was included in the second-stage regression. With 12 lags included, easy policy was sometimes significant after two years. It is hard to believe, however, that a decrease in the funds rate this quarter would directly affect output more than two years from now. And even if it did, such a long lag still seems to preclude a useful stabilization role. Moreover, the restriction that the regression equation had only eight lags versus 12 lags could not be rejected.

¹⁸ The records included the FOMC directives and the associated policy discussions in the minutes of the FOMC meetings.

¹⁹ The series constructed by Boschen and Mill was monthly. For compatibility with quarterly data, however, this article uses the value of the index on the last month of the quarter. Boschen and Mill also used negative values to indicate tight policy and positive values to indicate easy policy. For comparability with the results using the funds rate (which increases when policy is tight), this article uses the opposite signs: positive values represent tight policy and negative values represent easy policy.

²⁰ This issue arises in any study in which the data investigated do not result from controlled experiments. For

example, if psychologists wish to investigate whether alcohol causes depression, they must first determine that their subjects were not drinking because they were depressed in the first place. Psychologists can pre-screen their subjects and reject those who give depression as a reason for their drinking. This pre-screening is analogous to the first-stage regression here (and with the funds rate), which “rejects” changes in policy due to changes in output growth. But other than declining output growth, or a recession, why would policy ever be eased? Perhaps because of public pressure, or because of a shift by existing policymakers toward complementary goals, such as calming financial markets, or because new policymakers with different goals joined the FOMC. All of

these reasons could lead to a change in policy that was not directly in response to a change in output growth.

²¹ Inflation was included in the first-stage regression to maintain symmetry with the funds rate specification and because Boschen and Mill did so. The results do not change significantly, however, when inflation is excluded from the first stage.

²² Although the Boschen-Mill policy index runs only through 1991:4, the second-stage regression uses only lagged policy so the estimation runs through 1992:1.

REFERENCES

- Bernanke, Ben, and Alan Blinder. 1992. “The Federal Funds Rate and the Channels of Monetary Transmission,” *The American Economic Review*, September, pp. 901-21.
- Bernanke, Ben, and Mark Gertler. 1989. “Agency Costs, Net Worth and Business Fluctuations,” *The American Economic Review*, March, pp. 14-31.
- Ball, Laurence, and N. Gregory Mankiw. 1992. “Asymmetric Price Adjustment and Economic Fluctuations,” mimeo, Harvard University, May.
- Boschen, John F., and Leonard O. Mills. 1993. “The Narrative Approach to Evaluating Monetary Policy: Consistency of Interpretation and the Relation to Monetary Activity,” mimeo, Federal Reserve Bank of Richmond.
- Caballero, Richardo J., and Eduardo M.R.A. Engel. 1992. “Price Rigidities, Asymmetries, and Output Fluctuations,” National Bureau of Economic Research Working Paper No. 4091, Cambridge, Mass., June.
- Cover, James Peery. 1992. “Asymmetric Effects of Positive and Negative Money-Supply Shocks,” *The Quarterly Journal of Economics*, November, pp. 1261-82.
- DeLong, J. Bradford, and Lawrence H. Summers. 1988. “How Does Macroeconomic Policy Affect Output?” *Brookings Papers on Economic Activity*, No. 2, pp. 433-94.
- Fazzari, Stephen, R. Glen Hubbard, and Bruce Petersen. 1988. “Financing Constraints and Economic Activity,” *Brookings Papers on Economic Activity*, no. 1, pp. 141-95.
- Friedman, Milton, and Anna J. Schwartz. 1963. *A Monetary History of the United States, 1867-1960*. Princeton: Princeton University Press.
- Gertler, Mark, and Simon Gilchrist. 1992. “Monetary Policy, Business Cycles, and the Behavior of Small Manufacturing Firms,” mimeo, New York University, December.
- Gordon, Robert J. 1984. *Macroeconomics*, 3d ed. Boston: Little, Brown, and Company.
- Greenspan, Alan. 1992. Remarks before the Annual Meeting of the Independent Bankers Association of America, February 11, unpublished.
- Hansen, Alvin H. 1941. *Fiscal Policy and Business Cycles*. New York: W.W. Norton and Company.
- Jackman, Richard, and John Sutton. 1982. “Imperfect Capital Markets and the Monetariest Black Box: Liquidity Constraints, Inflation and the Asymmetric Effects of Interest Rate Policy,” *The Economic Journal*, 92, March, pp. 108-28.
- Johnson, Harry G. 1962. “Monetary Theory and Policy,” *The American Economic Review*, 52:3, June, pp. 335-84.
- Kashyap, Anil, Owen Lamont, and Jeremy Stein. 1992. “Credit Conditions and the Cyclical Behavior of Inventories: A Case Study of the 1981-82 Recession,” mimeo, University of Chicago School of Business, August.
- Keeton, William R. 1979. *Equilibrium Credit Rationing*. New York: Garland Press.
- Keynes, John Maynard. 1935. *The General Theory of Employment Interest and Money*. New York: Harcourt, Brace, and Company.
- Mayer, Thomas. 1968. *Monetary Policy in the United States*. New York: Random House.
- Romer, Christina D., and David H. Romer. 1989. “Does Monetary Policy Matter? A New Test in the Spirit of Friedman and Schwartz,” in Olivier J. Blanchard and Stanley Fischer, eds., *NBER Macroeconomics Annual*, pp. 121-80.
- Roosa, Robert V. 1951. “Interest Rates and the Central Bank,” from *Money, Trade, & Economic Growth*, Essays in Honor of John Henry Williams. New York: The Macmillan Company.
- Sichel, Daniel E. 1989. “Business Cycle Asymmetry: A Deeper Look,” Federal Reserve Board Working Paper no. 93, April.
- Tsiddon, Daniel. 1991. “On the Stubbornness of Sticky Prices,” *International Economic Review*, February, pp. 69-75.

The Impact of Monetary Policy on Bank Lending: The Role of Securities and Large CDs

By William R. Keeton

The seemingly small impact of easier monetary policy on bank lending during the recent recovery has added to the controversy over the lending view of monetary policy. According to the lending view, easier monetary policy stimulates the economy by increasing bank lending. But critics claim that bank lending may fail to rise because banks may use the deposits generated by the easier policy to buy securities or retire large CDs. The weak response of bank lending to the recent easing seems to support the critics. Proponents of the lending view dismiss this charge. Instead, they blame the weakness in bank lending on the balance sheet concerns of banks and borrowers.

The controversy over the lending view has focused only on the *direct* effect of monetary policy—both sides have largely ignored the *indirect* effect. Easier monetary policy not only increases bank lending directly by increasing deposits, but also indirectly by lowering open-market rates. Specifically, lower open-market rates stimulate lending by encouraging banks to fund new loans

through security sales or large CDs.

This article analyzes the implications of bank security holdings and large CDs for the direct and indirect effects of monetary policy on bank lending. The article argues that when both the direct and indirect effects are taken into account, banks' ability to fund loans by selling securities and issuing large CDs may strengthen monetary policy rather than weaken it. The first section describes the direct effect of monetary policy on bank lending and explains why critics believe banks' ability to fund loans from nondeposit sources has diminished this effect. The second section describes the indirect effect and argues that banks' access to nondeposit funds increases this effect. The last section shows that the increase in the indirect effect can outweigh the decrease in the direct effect, so that monetary policy is more effective on balance.

HOW NONDEPOSIT SOURCES OF FUNDS DECREASE THE DIRECT EFFECT

Critics of the lending view argue that bank security holdings and CD issuance have reduced the direct effect of monetary policy on bank lend-

William R. Keeton is a senior economist at the Federal Reserve Bank of Kansas City. Kenneth Heinecke, a research associate at the bank, helped prepare the article.

ing. This section describes the direct effect of monetary policy and explains why critics believe this effect has been diminished.

What is the direct effect?

A key tenet of the lending view is that monetary policy affects economic activity largely through the direct effect of changes in bank reserves on bank lending. Consider, for example, an open-market purchase by the Federal Reserve. Reserves and transactions deposits first increase by equal amounts. Banks, finding themselves with more reserves than they need to meet the reserve requirement on transactions deposits, then increase their lending and investment, leading to a further rise in transactions deposits and required reserves. This process continues until banks' required reserves have risen enough to eliminate the surplus of reserves.

According to the lending view, the rise in bank lending permitted by the increase in transactions deposits stimulates private spending, thereby boosting the economy. To make additional loans, banks must lower loan rates and relax credit standards. Large borrowers who can borrow on the open market may respond by substituting bank credit for open-market credit, leaving their total spending unchanged. However, small and medium-size borrowers who cannot easily borrow on the open market will use the additional credit to finance new spending.

Why banks' willingness to use nondeposit funds reduces the direct effect

The argument that bank security holdings and CD issuance have reduced the direct effect of monetary policy on bank lending is based on two claims. First, banks' increased willingness to fund loans from nondeposit sources has made the supply of bank loans highly elastic with respect to the return on loans. Second, when loan supply is highly elastic, changes in deposits have little effect

on bank lending.

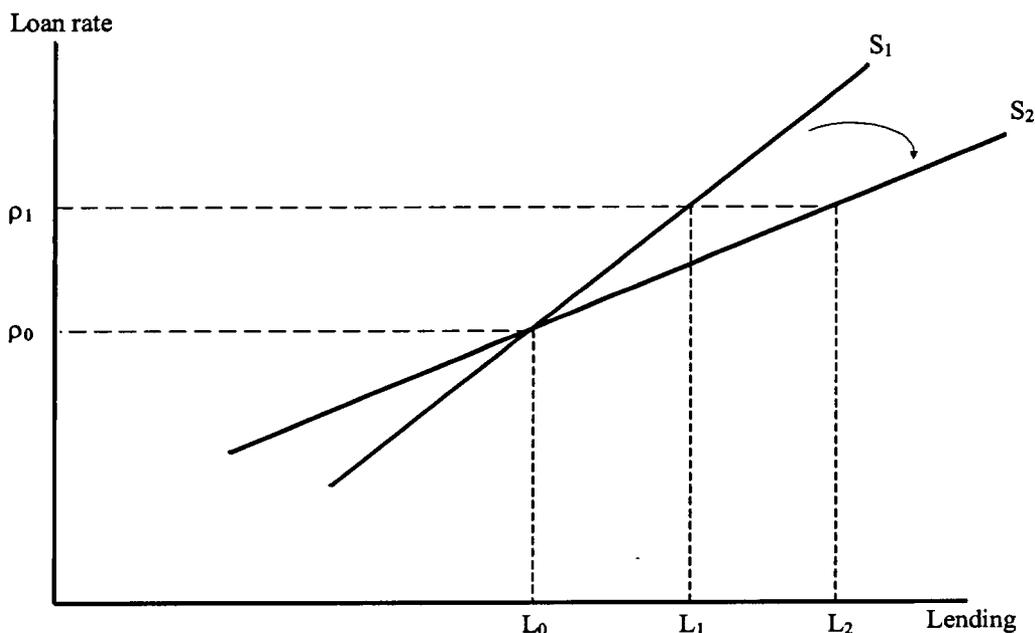
Willingness to use nondeposit funds makes loan supply more elastic. Critics of the lending view argue that banks have become more willing than in the past to fund loans by selling securities or issuing large CDs. The main factor that used to deter a bank from funding loans from such nondeposit sources was the risk of illiquidity. Loans could not be liquidated as quickly as securities to meet unanticipated deposit withdrawals. Thus, if a bank shifted too heavily from securities to loans, it might have to meet subsequent deposit withdrawals by borrowing on an emergency basis at above normal rates.¹ Similarly, liabilities like large CDs were viewed as a more volatile source of funds than core deposits because the liabilities were uninsured and held by large investors who were highly sensitive to relative yields. Thus, if a bank financed too many new loans with large CDs, the likelihood of a liquidity crisis would increase.

Recent developments may have reduced this risk of illiquidity, making banks more willing to use their securities and large CDs as a source of funds for loans. The growth of interstate banking and the trend toward consolidation have made deposits more predictable for many banks, reducing the risk of unanticipated withdrawals. Also, the removal of deposit rate ceilings, the growth of brokered deposits, and the development of secondary loan markets have made it easier for banks to meet deposit withdrawals or runoffs of CDs by attracting new funds or selling loans. Thus, banks may be more willing to sell securities and issue large CDs when lending becomes more profitable, and more willing to buy securities or retire large CDs when lending becomes less profitable.

The effect on loan supply of banks' increased willingness to use nondeposit sources of funds is shown in Figure 1. Total bank lending is measured on the horizontal axis and the loan rate on the vertical axis.²

The loan supply curve, *S*, shows how much banks would like to lend at each loan rate, given the rate of return on open-market assets and the

Figure 1
Loan Supply Becomes More Elastic



volume of transactions deposits. In deciding how much to lend for a given volume of transactions deposits, banks must weigh the benefit and cost of using nondeposit funds to finance additional lending. The benefit is the interest income on the new loans. The cost includes both the increased risk of illiquidity and the interest expense from using more nondeposit funds—the interest lost on the securities that are sold or the interest paid on the CDs that are issued. If the loan rate rises, the benefit from additional lending will increase. Thus, banks will expand their lending until the benefit and cost of additional lending are again equal.

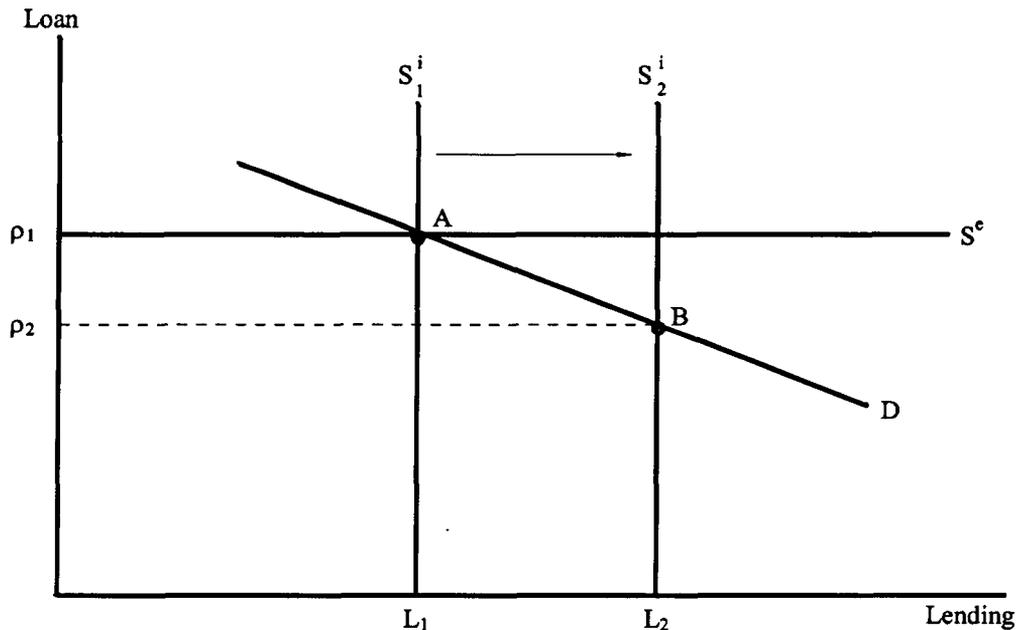
An increased willingness on the part of banks to fund loans from nondeposit sources flattens the

loan supply curve, causing it to rotate from S_1 to S_2 in Figure 1. Suppose, for example, that due to fundamental changes in the banking industry, increases in lending funded from nondeposit sources have less tendency to increase illiquidity risk. Then when the loan rate rises, banks will be able to expand their lending by a greater amount before the benefit of additional lending is outweighed by the cost—that is, before the extra interest income from loans is offset by the extra illiquidity risk and the extra interest cost of nondeposit funds. Thus, if the loan rate rises from ρ_0 to ρ_1 in Figure 1, the amount banks want to lend will increase from L_0 to L_2 rather than from L_0 to L_1 .

A more elastic loan supply reduces the direct effect. The simplest way to show that a more

Figure 2

Effect of a Change in Deposits
(with open-market rate held constant)



elastic loan supply curve reduces the direct effect of monetary policy on bank lending is to compare two extreme cases—a perfectly inelastic supply curve and a perfectly elastic one. Figure 2 shows the effects of an increase in reserves in these two cases. As before, open-market rates are held constant and each supply curve is drawn for a fixed volume of transactions deposits.

The vertical curve S^i represents the case of a perfectly inelastic loan supply. In this case, banks desire a fixed ratio of loans to transactions deposits no matter how high the loan rate. For example, if the costs of illiquidity are very high, banks may prefer a loan-deposit ratio just low enough to avoid any risk of illiquidity.³ Thus, when the loan rate rises and deposits remain unchanged, banks refuse

to supply more loans.

The horizontal curve S^e shows the opposite extreme of a perfectly elastic loan supply. In this case, there exists a minimum loan rate at which banks are willing to lend, given the rate of return on open-market assets. If the loan rate rises even slightly above this level, banks want to fund an unlimited amount of loans by selling securities or issuing large CDs—for example, because illiquidity costs are negligible.⁴ Conversely, if the loan rate falls even slightly below the critical level, banks refuse to supply any loans. At a loan rate this low, banks are unwilling to fund loans with large CDs and would rather invest their transactions deposits in securities than in loans.

Finally, the downward-sloping demand curve

D shows how much businesses and households wish to borrow from banks at each loan rate. As the loan rate rises, the demand for loans falls for two reasons. First, for large well-known businesses, higher loan rates make bank loans less attractive than other methods of financing spending, such as borrowing on the open market or drawing down liquid assets. Second, as the loan rate increases, borrowers reduce spending, decreasing their demand for bank loans.

Consider now the effects of an open-market purchase in the two cases, starting from the same initial equilibrium (point A) and holding open-market rates constant.⁵ As noted earlier, the open-market purchase will cause transaction deposits to rise. However, as shown in Figure 2, this increase in deposits will have very different effects on the two loan supply curves and thus on the volume of bank lending.

The increase in deposits causes the perfectly inelastic supply curve to shift from S_1^i to S_2^i . Because banks desire a constant loan-deposit ratio, they increase lending by the same proportion as deposits. At the initial loan rate, ρ_1 , banks now supply more credit than borrowers desire. As a result, borrowers bid down the loan rate and banks move down the new vertical supply curve until the excess supply of credit is eliminated. In Figure 2, the new equilibrium is at point B , with a higher level of lending, L_2 , and a lower loan rate, ρ_2 .

In contrast, the increase in deposits has no effect on the perfectly elastic supply curve, S^e . With open-market rates held constant, the loan rate at which banks are just willing to lend remains at ρ_1 . Thus, total lending remains at L_1 , with banks using all their new transactions deposits in excess of reserve requirements to buy securities or retire CDs. To see why this case differs from the previous one, suppose banks increase their lending above L_1 as their deposits rise. As before, borrowers will bid down the loan rate. Now, however, even the slightest decrease in the loan rate will cause banks to shift heavily out of loans into

securities or slash the amount of loans funded with large CDs. As a result, lending will contract until the initial equilibrium is restored.

The fact that lending does not increase when loan supply is perfectly elastic confirms that banks' willingness to fund loans by selling securities or issuing CDs decreases the direct effect of monetary policy on bank lending. Critics of the lending view argue that this decrease in the direct effect significantly reduces the importance of the lending channel in the monetary transmission mechanism.⁶

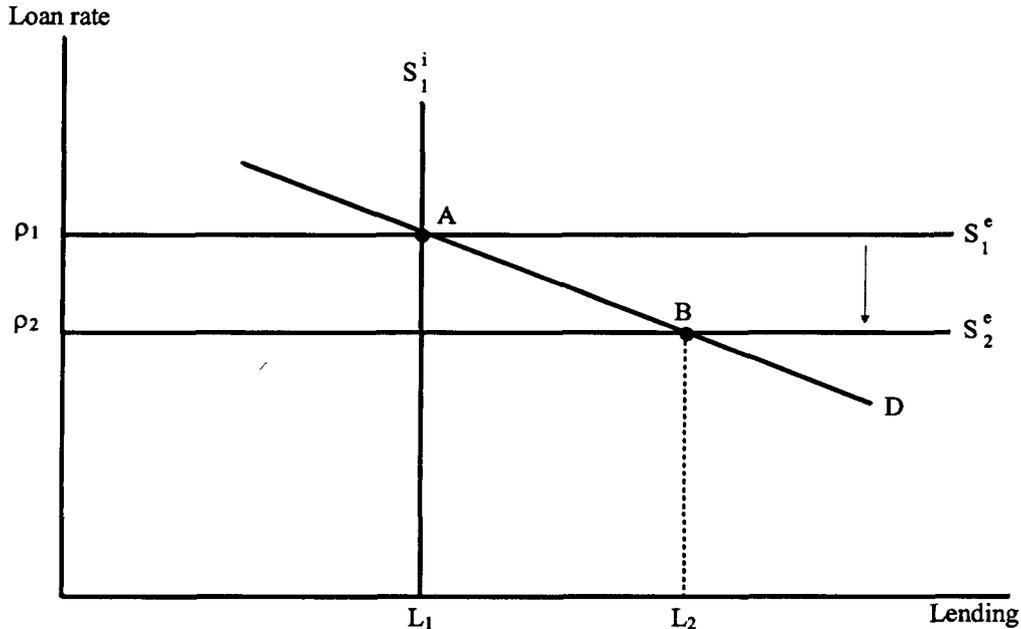
HOW NONDEPOSIT SOURCES OF FUNDS INCREASE THE INDIRECT EFFECT

To the extent bank security holdings and CD issuance have reduced the direct impact of monetary policy on bank lending, it is tempting to conclude that monetary policy will be less effective. Such a conclusion would be unwarranted, however, because it ignores the indirect effect of monetary policy on bank lending via changes in open-market rates.

What is the indirect effect?

The indirect effect consists of two steps—easier monetary policy lowers open-market rates, and lower open-market rates stimulate bank lending. To illustrate the first step, suppose the Fed increases the supply of bank reserves through an open-market purchase. As in the discussion of the direct effect, transactions deposits will rise until banks' required reserves just equal the new, higher supply of reserves. What the discussion of the direct effect ignored is that the public will use their excess transactions deposits to bid for open-market assets, driving the prices of such assets up and their rates of return down. At the same time, deposit rates will fall because banks will find it hard to invest the new transactions deposits prof-

Figure 3
Effect of a Change in Open-Market Rate
(with deposits held constant)



itably. Equilibrium will be restored when open-market rates fall enough relative to deposit rates to make the public content to hold the new, higher level of transactions deposits.

In the second step of the indirect effect, lower open-market rates make it more attractive for banks to use securities and large CDs to fund loans. As open-market rates fall, so will the rate of return on securities and the interest rate banks have to pay on large CDs. Thus, at the initial loan rate, banks will find it profitable to shift out of securities into loans and to finance new loans with large CDs. In other words, the fall in open-market rates will increase banks' desired ratio of loans to transactions deposits, reinforcing the direct effect of higher transactions deposits on bank lending.

Why banks' willingness to use nondeposit funds increases the indirect effect

As noted earlier, critics of the lending view believe banks' willingness to use nondeposit funds has made the supply of bank loans more responsive to changes in the loan rate. But if the supply of bank loans is more responsive to changes in the return on loans, it should also be more responsive to changes in the cost of using nondeposit funds to make loans. Thus, a decrease in open-market rates caused by easier monetary policy should lead to a bigger shift out of securities into loans and a bigger increase in the amount of new loans financed with CDs.

Figure 3 illustrates this point by comparing the effect of a decrease in the open-market rate under

Table 1
Effects of Easier Monetary Policy on Bank Lending

	Direct effect (via changes in deposits)	Indirect effect (via changes in open-market rate)
Perfectly inelastic loan supply	Positive	Zero
Perfectly elastic loan supply	Zero	Positive

the two extremes of a perfectly inelastic loan supply curve and a perfectly elastic supply curve. The loan supply curves and the loan demand curve in this diagram are defined as in Figure 2. Recall that Figure 2 isolated the direct effect by showing the impact of an increase in transactions deposits with the open-market rate held constant. In contrast, Figure 3 focuses on the indirect effect by showing the impact of a decrease in the open-market rate with transactions deposits held constant. For convenience, Figure 3 also assumes that bank borrowers have no access to the open market, so that the loan demand curve is unaffected by changes in the open-market rate.⁷

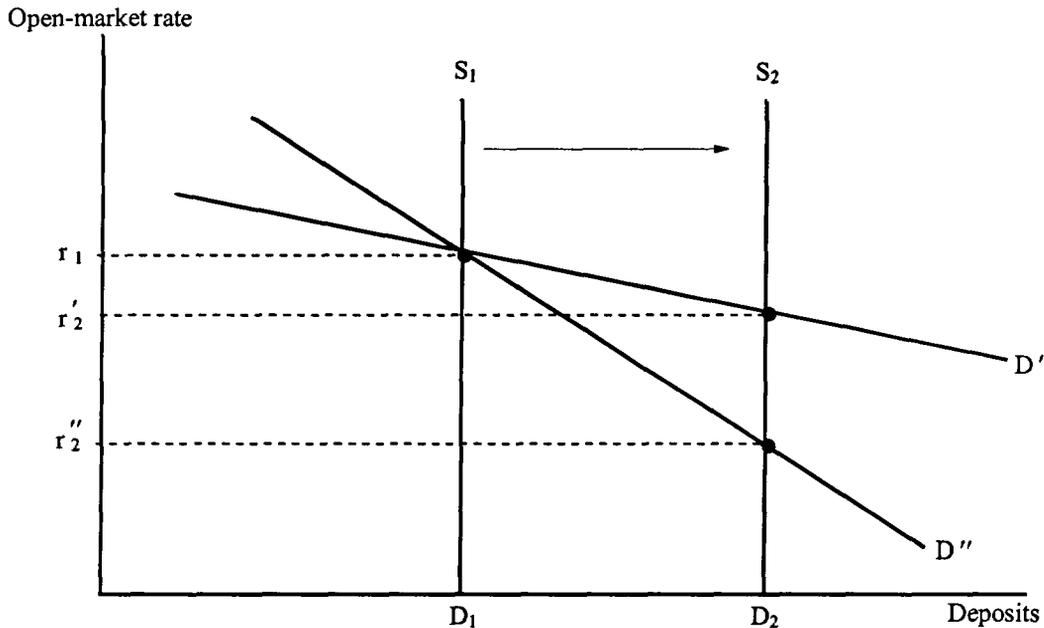
As shown in Figure 3, the fall in the open-market rate has no effect on the perfectly inelastic supply curve, S^i . The curve remains unchanged because banks desire a fixed ratio of loans to transactions deposits and transactions deposits are held constant. The decrease in the open-market rate reduces the cost to banks of funding loans by selling securities or issuing large CDs. By assumption, however, banks are unwilling to increase lending—for example, because doing so would increase their risk of illiquidity above zero. Thus, banks remain at point *A*, with the same total lending.

In contrast, the decrease in the open-market rate shifts the perfectly elastic supply curve, S^e ,

downward. As in the perfectly inelastic case, the fall in the open-market rate reduces the cost to banks of funding loans by selling securities or issuing large CDs. In this case, however, the loan supply curve shifts downward because the minimum loan rate at which banks are willing to lend falls from ρ_1 to ρ_2 . Suppose, for example, that banks require a fixed spread between the loan rate and the open-market rate to be willing to lend. Then the curve S^e will shift downward by the same amount that the open-market rate falls. At the initial loan rate, ρ_1 , banks now want to fund an unlimited amount of loans from nondeposit sources. As a result, they reduce the loan rate in an effort to attract more loan customers. This process continues until a new equilibrium is reached at point *B*, with a higher level of lending and a lower loan rate. Thus, while the fall in the open-market rate leaves lending unchanged in the perfectly inelastic case, it increases lending in the perfectly elastic case.

Table 1 summarizes the results of Figures 2 and 3 on the effects of easier monetary policy on bank lending. When loan supply is perfectly inelastic, the direct effect via changes in deposits is positive and the indirect effect via changes in open-market rates is zero (first row). On the other hand, when loan supply is perfectly elastic, the

Figure 4
Effect of Easier Monetary Policy on Open-Market Rate



direct effect is zero and the indirect effect is positive (second row). Thus, to the extent bank security holdings and CD issuance make loan supply highly elastic, they decrease the direct effect of monetary policy on bank lending but increase the indirect effect.

CAN NONDEPOSIT SOURCES OF FUNDS STRENGTHEN MONETARY POLICY?

Can the increase in the indirect effect outweigh the decrease in the direct effect, so that bank security holdings and large CDs actually increase the effectiveness of monetary policy? The answer

is yes, provided deposit demand is sufficiently insensitive to a change in interest rates or loan demand is sufficiently sensitive to such a change.

Interest sensitivity of deposit demand

Table 1 implies that monetary policy will have a bigger impact on lending with a perfectly elastic loan supply on one condition: the indirect effect in the perfectly elastic case exceeds the direct effect in the perfectly inelastic case. One way this can happen is if deposit demand is interest-insensitive, so that increases in reserves lead to large decreases in the open-market rate.

Figure 4 shows why an increase in reserves

will lead to a bigger fall in the open-market rate in the perfectly elastic case if deposit demand is interest-insensitive. This diagram measures the volume of transactions deposits on the horizontal axis and the open-market rate on the vertical axis. The vertical curves S_1 and S_2 show the supply of transactions deposits before and after the increase in reserves.⁸ The downward-sloping curves D' and D'' show two alternative demand curves for transactions deposits.

Each demand curve shows the amount of deposits the public desires to hold at different open-market rates, taking into account the response of the deposit rate to the open-market rate.⁹ When the open-market rate rises, banks will bid up the interest rate on transactions deposits. However, the deposit rate will rise less than the open-market rate, causing the public to shift out of transactions deposits into open-market assets. The flat curve D' assumes that an increase in the open-market rate leads to a large decline in demand for deposits, while the steep curve D'' assumes that an increase in the open-market rate leads to a small decline in demand for deposits. One reason the curve may be steep is that the public's demand for deposits may be unresponsive to changes in the relative returns on deposits and open-market assets. The other reason is that the relative returns on deposits and open-market assets may not change very much because deposit rates are highly flexible.

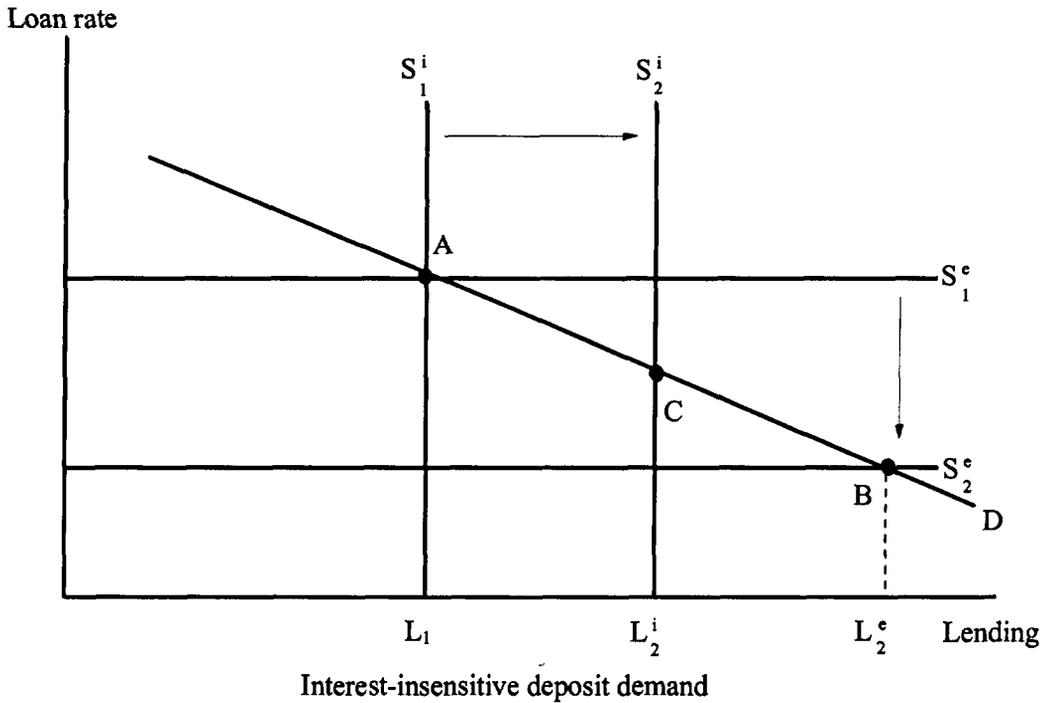
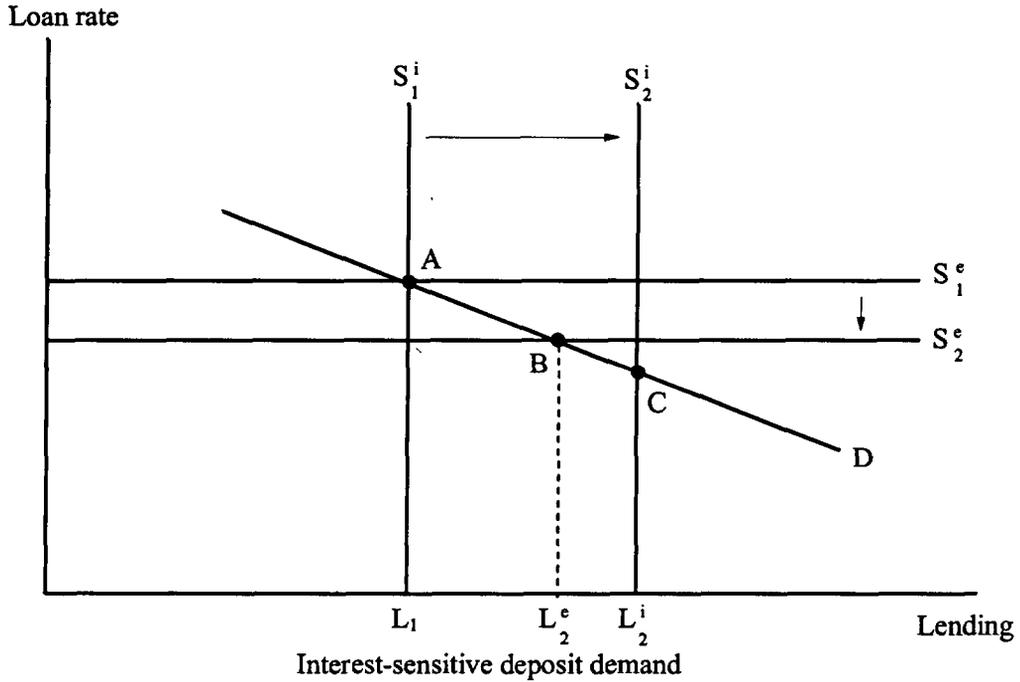
By shifting out the supply of transactions deposits, the easier policy creates an excess supply of deposits and drives down the open-market rate. If deposit demand is interest-sensitive (curve D'), only a small decline in the open-market rate is needed to induce the public to hold the higher volume of transactions deposits. As a result, the open-market rate falls only from r_1 to r_2' . However, if deposit demand is interest-insensitive (curve D''), a much larger decline in the open-market rate is needed to eliminate the public's excess supply of deposits. Thus, the open-market rate falls all the way from r_1 to r_2'' .

Figure 5 shows what this difference in the response of the open-market rate means for the bank loan market. The diagram shows the total effect of the increase in reserves on bank lending, taking into account both the direct effect via the change in deposits and the indirect effect via the change in the open-market rate. The top panel corresponds to the case of interest-sensitive deposit demand (the open-market rate falls only slightly), while the bottom panel represents the case of interest-insensitive deposit demand (the open-market rate falls sharply).¹⁰

Consider first the impact of the easier policy on bank lending with interest-sensitive deposit demand (top panel). Because the open-market rate falls only slightly, the perfectly elastic loan supply curve, S^e , shifts down by only a small amount.¹¹ Thus, in the perfectly elastic case, lending increases only from L_1 to L_2^e . In contrast, the increase in transactions deposits shifts out the perfectly inelastic supply curve, S^i , enough to increase lending all the way to L_2^i . Thus, when deposit demand is highly interest-sensitive, the critics of the lending view are correct—banks' willingness to fund loans from securities and large CDs reduces the impact of monetary policy on bank lending.

Consider next the impact of the same policy change when deposit demand is interest-insensitive (bottom panel). The increase in transactions deposits shifts the perfectly inelastic supply curve, S^i , outward by the same amount as in the top panel, leading to the same increase in lending as before. But because the open-market rate falls sharply when deposit demand is interest-insensitive, the perfectly elastic supply curve, S^e , now shifts down by a much greater amount.¹² As a result, the easier policy causes lending to increase even more in the perfectly elastic case than the perfectly inelastic case—that is, L_2^e now exceeds L_2^i instead of falling short of L_2^i . Thus, when deposit demand is interest-insensitive, banks' willingness to fund loans from securities and large CDs increases the impact of monetary policy on bank lending, the opposite of what the critics claim.¹³

Figure 5
Effect of Easier Monetary Policy on Lending



Interest sensitivity of loan demand

The reason interest-insensitive deposit demand leads to a large indirect effect in the perfectly elastic case is that increases in reserves cause large decreases in the open-market rate. Even if the decrease in the open-market rate is not large, however, the indirect effect can be large if small decreases in the open-market rate lead to large changes in lending. This will be the case if bank borrowers' demand for credit is relatively sensitive to the cost of borrowing.

In the top panel of Figure 5, for example, suppose the loan demand curve D becomes flatter, rotating counterclockwise around point A . The outward shift in the perfectly inelastic loan supply curve from S_1^i to S_2^i will lead to the same increase in lending as before. However, the downward shift in the perfectly elastic curve from S_1^e to S_2^e will now cause a bigger increase in lending than before. If the loan demand curve becomes flat enough, lending will increase more in the perfectly elastic case than in the perfectly inelastic case—that is, L_2^e will exceed L_2^i , just as it does in the bottom panel of Figure 5. Thus, if loan demand is relatively interest-sensitive, bank security holdings and CD issuance can increase the effectiveness of monetary policy.¹⁴

CONCLUSIONS

According to the lending view, monetary

policy works by increasing the supply of deposits to banks and thereby stimulating bank lending. Critics of this view sometimes argue that the willingness of banks to fund loans by selling securities or issuing large CDs insulates bank lending from changes in deposits, blocking the lending channel. It would be wrong to conclude, however, that monetary policy is less effective. Monetary policy not only affects bank lending directly, by changing deposits, but also indirectly, by changing the return on securities and the cost of CDs. Although the increased availability of nondeposit sources of funds diminishes the direct effect, it magnifies the indirect effect. If deposit demand is relatively interest-insensitive or loan demand is relatively interest-sensitive, the increase in the indirect effect may outweigh the decrease in the direct effect, rendering monetary policy more effective.

While nondeposit sources of funds can strengthen monetary policy in principle, this article has not proved that monetary policy has been strengthened in practice. In favor of the view that monetary policy has been strengthened, it could be argued that deposit rate deregulation has made deposit rates more flexible, reducing the interest sensitivity of demand for deposits. Based on theory alone, however, it is impossible to say whether deposit demand is now sufficiently interest-insensitive or loan demand sufficiently interest-sensitive for nondeposit sources of funds to strengthen monetary policy. Only direct empirical evidence on the slopes of the two demand curves can resolve the issue.

ENDNOTES

¹ A formal model of the liquidity motive for holding securities can be found in Tobin. King finds modest empirical evidence for the liquidity motive, but his study ends in 1979.

² Strictly speaking, what the vertical axis measures is not the loan rate but the expected rate of return on bank loans—the weighted sum of all possible rates of return to the bank, with each possible return weighted by its probability of occurrence. The expected rate of return on bank loans can

increase not only through higher loan rates but through stricter nonprice terms that reduce the risk of default. Whenever the term “loan rate” is used below, the reader should understand this to be shorthand for the expected rate of return on bank loans.

³ Suppose, for example, that a bank's funds consist entirely of transactions deposits and that the reserve requirement on transactions deposits is 10 percent. Suppose also that

the bank has some chance of losing half its deposits but no chance of losing more than that. Then the bank can keep its risk of illiquidity equal to zero by holding 10 percent of its deposits in the form of required reserves, 45 percent in the form of liquid securities, and 45 percent in the form of illiquid loans.

⁴ Note also that if the source of nondeposit funds is large CDs, a perfectly elastic supply curve requires that the expected rate of return demanded by investors on a bank's CDs not go up as the bank increases its lending.

⁵ As discussed at length below, changes in the supply of reserves generally lead to changes in open-market rates as well as changes in deposits, unless the public's demand for deposits is perfectly interest-elastic. The correct way to interpret Figure 2 is as a "thought experiment" to isolate the effect of changes in deposits from the effect of changes in open-market rates.

⁶ The most widely cited criticism along these lines is Romer and Romer. While disputing that the loan supply curve has become highly elastic, many proponents of the lending view appear to agree that this change would diminish the lending channel (Bernanke; Bernanke and Blinder). For further discussion of the controversy, see the surveys by Gertler and Gilchrist and by Morgan.

⁷ The only purpose of this assumption is to simplify the diagram. As shown in Keeton, the basic results require only that an equal increase in the loan rate and open-market rate reduce loan demand by decreasing borrowers' desired spending.

⁸ Figure 4 assumes for convenience that banks hold no excess reserves. In this case, the supply of transactions deposits equals the supply of reserves divided by the required reserve ratio.

⁹ Suppose, for example, that banks require a fixed spread between the loan rate and the open-market rate to cover the costs of making loans; that the cost of servicing each dollar of transactions deposits is c ; and that the required reserve ratio is k . Then, with perfect competition in the deposit market, the deposit rate will equal $(1-k)r - c$, where r is the open-market rate. Of course, market imperfections or deposit rate ceilings could prevent the deposit rate from adjusting to the perfectly competitive level.

¹⁰ As in Figure 3, bank borrowers are assumed to have no access to the open market, so that the demand curve is independent of the open-market rate. See footnote 7.

¹¹ Suppose, for example, that banks require a fixed spread between the loan rate and the open-market rate to be willing to lend. Then the curve S^e will shift down by the same amount that the open-market rate falls in Figure 4, $r_1 - r_2$.

¹² For example, if banks require a fixed spread between the loan rate and the open-market rate, the curve will shift down by the amount $r_1 - r_2$ shown in Figure 4.

¹³ The discussion has ignored two complications. The first is that the impact of an increase in reserves on economic activity depends not only on the change in bank lending but also on the change in open-market rates, since the latter influence spending by open-market borrowers. Taking this factor into account does not change the conclusions, however. The more the loan rate falls when policy is eased, the more the deposit rate will tend to fall and thus the more the open-market rate must fall for the public to hold the higher volume of transactions deposits. Thus, if the easier policy causes a bigger increase in lending and decrease in the loan rate in the perfectly elastic case than the perfectly inelastic case, it must also cause a bigger decrease in the open-market rate in the perfectly elastic case.

The second complication is that Figure 5 shows the effect of easier policy on bank lending at the initial level of GDP. If deposit demand rises significantly more than loan demand as GDP rises, the possibility cannot be ruled out that lending will end up increasing less in the perfectly elastic case than the perfectly inelastic case, even if lending initially increases more in the perfectly elastic case. However, the point remains that a perfectly elastic loan supply can strengthen monetary policy if deposit demand is sufficiently interest-insensitive. See Keeton for further details.

¹⁴ Recall that Figure 5 assumes bank borrowers have no access to the open market, so that loan demand is independent of the open-market rate. Under the more realistic assumption that loan demand depends on both the loan rate and the open-market rate, the condition for monetary policy to be more effective is that loan demand declines a large amount when the loan rate and open-market rate go up by equal amounts.

REFERENCES

- Bernanke, Ben S. 1988. "Monetary Policy Transmission: Through Money or Credit?" *Business Review*, Federal Reserve Bank of Philadelphia, November/December.
- _____ and Alan S. Blinder. 1988. "Credit, Money, and Aggregate Demand," *American Economic Review*, May.
- Gertler, Mark, and Simon Gilchrist. 1993. "The Role of Credit Market Imperfections in the Monetary Transmission Mechanism: Arguments and Evidence," *Scandinavian Journal of Economics*.
- Morgan, Donald P. 1992. "Are Bank Loans a Force in Monetary Policy?" *Economic Review*, Federal Reserve Bank of Kansas City, Second Quarter.
- Keeton, William R. 1992. "Monetary Policy Implications of Bank Security Holdings and Large CDs," Federal Reserve Bank of Kansas City, Research Working Paper RWP 92-8, December.
- King, Stephen R. 1986. "Monetary Transmission: Through Bank Loans or Bank Liabilities?" *Journal of Money, Credit, and Banking*, August.
- Romer, Christina D., and David H. Romer. 1990. "New Evidence on the Monetary Transmission Mechanism," *Brookings Papers*, 1.
- Tobin, James. 1982. "The Commercial Banking Firm: A Simple Model," *Scandinavian Journal of Economics*.

The Changing U.S. Pork Industry: A Dilemma for Public Policy

By Alan Barkema and Michael L. Cook

The shape of the U.S. pork industry is changing dramatically, as pork production shifts into the hands of fewer, larger farmers with closer ties to processors and consumers. The changing shape of the pork industry, the nation's second largest meat industry, points to the loss of thousands of small hog farms in the United States. The threat to traditional ways of farming has triggered a public policy debate in Iowa, Kansas, and other leading hog producing states.

Primarily responsible for the changes underway in the U.S. pork industry are today's discriminating consumers. Their more sophisticated tastes challenge the industry to pack improved nutrition into more convenient products. The industry is responding with an arsenal of new technologies. And to keep high-tech pork products on target for today's palates, the industry is abandoning its traditional way of moving pork from the producer's lot to the dinner table.

This article considers the changes underway

in the U.S. pork industry today and what they suggest in the years ahead. The first section reviews the structural changes currently taking place in the industry. The second section describes the causes underlying this structural evolution. The third section considers the industry's future structure. The final section contemplates the dilemma the changing pork industry poses for public policy.

The article concludes that the wave of structural change in the pork industry will continue, resulting in a more integrated industry of fewer, larger farms with closer market ties to pork processors. The industry's emerging structure, however, poses a dilemma for public policy, which must balance the loss of traditional small farms against the economic benefits to consumers of higher quality, lower cost products.

HOW IS THE U.S. PORK INDUSTRY CHANGING?

A wave of concentration and integration is sweeping the U.S. pork industry. Pork production is concentrating in the hands of fewer, larger producers and processors. Meanwhile, hog farmers and pork processors are developing closer ties, forming a more integrated industry from the hog farm to the supermarket.

Alan Barkema is a senior economist at the Federal Reserve Bank of Kansas City. Michael L. Cook is the Robert D. Partridge Endowed Professor of Cooperative Leadership at the University of Missouri—Columbia. Corey Waldinger, a research associate at the bank, helped prepare the article.

A more concentrated industry

In the past, hundreds of thousands of small, independent hog farms were the heart of the U.S. pork industry. Most small hog farms mixed hog production with a wide menu of other farm enterprises, especially the production of corn and soybeans, the main ingredients in hog feed. In the traditional production sequence, hog farmers purchased breeding stock from hundreds of small, independent producers and raised their offspring from birth to market weight on home-grown grains. The genetic makeup of hog herds varied widely from farm to farm, and a wide range of feeding and management systems boosted the variation in the nation's hog herd even more.

Today, the pork industry has a new makeup. The number of hog farms in the United States has plummeted, as the industry consolidates on fewer, larger, more specialized hog farms. During the past two decades, the number of hog farms has dropped from nearly 900,000 to only 250,000.¹ Despite that drop, the total volume of pork production has increased, underscoring the industry's consolidation on larger farms (Chart 1).²

The trend toward bigger hog farms has accelerated in the 1990s, as the latest production technologies have transformed the traditional hog shed into a specialized pork factory. Today, some modern hog farms produce and market more than half a million hogs a year. Multinational companies offer improved breeding stock with the most advanced genetics that enable pigs to grow faster and produce leaner pork with less feed. Climate-controlled buildings ensure top production regardless of weather. Computerized information systems enable well-trained managers to monitor herd performance and health. Diseases are kept in check with a steadily improving collection of health products, some of which are the creations of the latest advances in biotechnology. These specialized hog farms produce pigs that are nearly identical in size, shape, and quality—in sharp contrast to the variation found across the dwindling

ranks of smaller farms. Thus, consumers are provided pork products of more consistent quality.

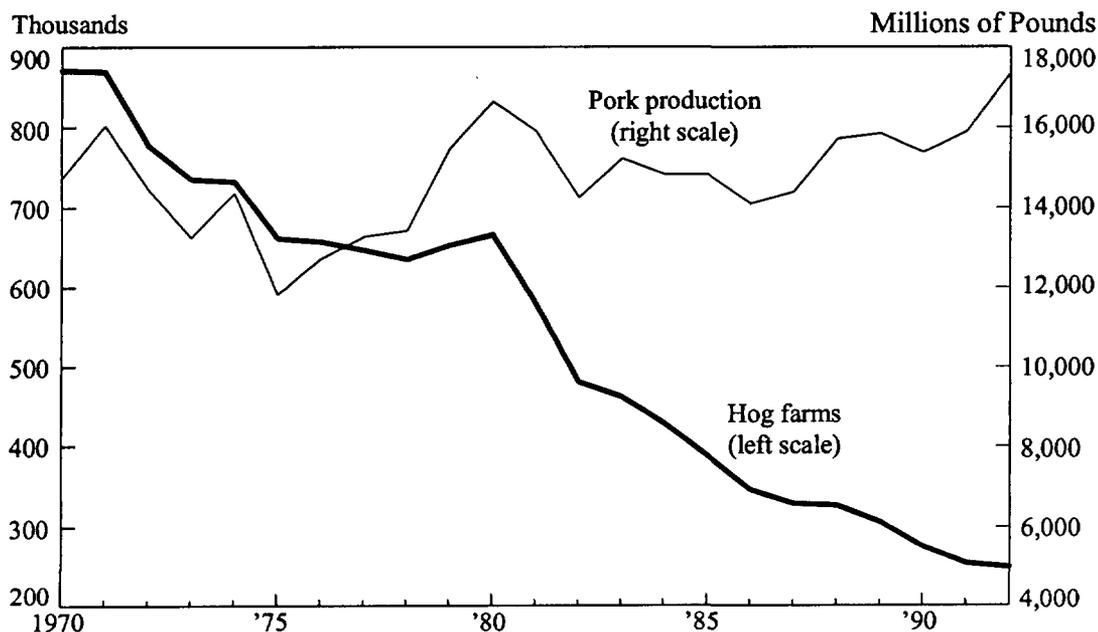
A more integrated industry

A change in marketing arrangements between hog farmers and pork processors has accompanied the industry's shift to fewer, larger farms. An increasing portion of the nation's hogs are produced under contract with pork processors or are owned outright by the processors, a system called vertical integration. As a result, marketing linkages between producers and processors are tighter. From 1980 to 1990, the percentage of the nation's hog production under contract or vertical integration doubled to about 10 percent (Manchester). Other data show that up to 16 percent of the nation's hogs were produced under contract or vertical integration in 1991 (Rhodes and Grimes). Thus, the trend toward a more integrated pork industry appears to be accelerating.

In the traditional marketing system—called “open production”—farmers sell market-weight hogs to the highest bidder among local hog processing companies or their agents.³ Hogs arriving at market vary widely in size, shape, and quality characteristics. But the variation in the hogs makes relatively little difference in the price farmers receive because quality grades are broadly defined. In effect, farmers are generally paid for the total live weight of the hogs sent to market, even though quality varies widely.

Prior contractual arrangements between hog farmers and processors are rapidly taking the place of open production in the pork industry. In a typical contractual arrangement, the farmer is paid a flat fee, plus various performance incentives, to feed young pigs to market weight. The farmer provides land, labor, buildings, and equipment, while the contractor provides the young pigs, feed, veterinary supplies, and management advice. Thus, the contractor assumes much more control—and the farmer assumes less control—over

Chart 1

Pork Production and the Number of Hog Farms in the United States

Source: United States Department of Agriculture.

the production process than in traditional open production.

Alternatively, some pork processors now rely on vertical integration to raise their own hogs. The key distinction between vertical integration and contracting is that even more control over the production process is transferred to the processor. Under vertical integration, hog farmers relinquish their independence and simply become employees of the integrated pork firm.

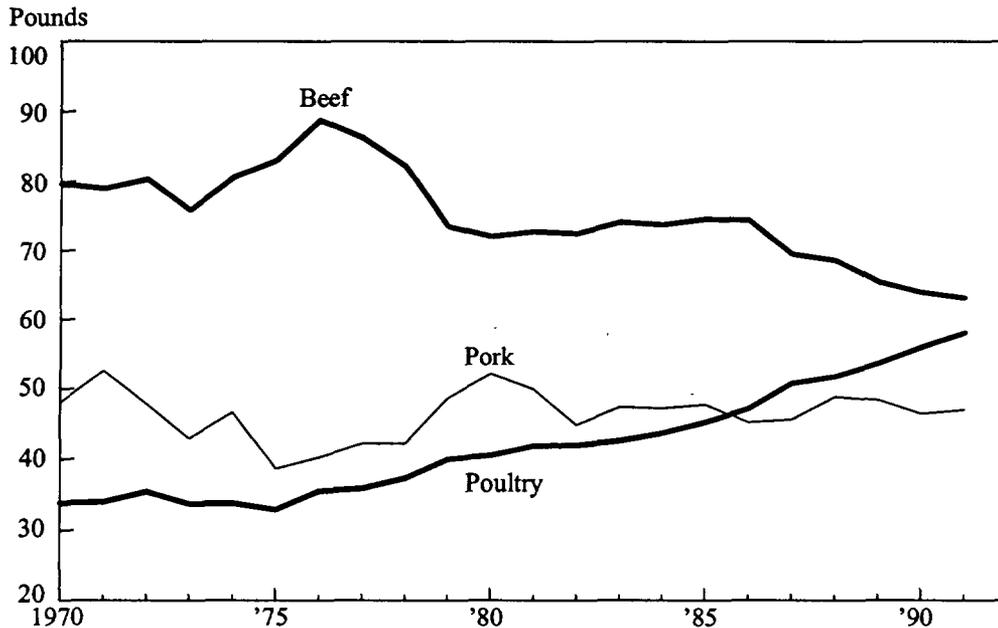
The shift toward a more integrated industry works hand-in-hand with the trend toward fewer and larger hog farms, much to the consternation of smaller producers. Processors who contract with farmers would prefer to manage a few contracts with large producers rather than many contracts with small producers. Integrated pork processing

firms are likely to produce enough hogs to account for a significant portion of their processing capacity. Thus, the trend toward a more concentrated, integrated pork industry has raised concerns that the nation's smaller, independent pork producers could lose access to markets and eventually be crowded out of the industry by much larger, integrated producers.

WHAT IS DRIVING THE PORK INDUSTRY TOWARD MORE CONCENTRATION AND INTEGRATION?

Three closely related factors appear to be driving the pork industry toward a more compact market structure. First, consumers have become

Chart 2

Per Capita Consumption of Beef, Pork, and Poultry

Source: United States Department of Agriculture, "Agricultural Outlook."

more discriminating, requiring the pork industry to design its products more carefully. Second, new technology is overhauling the pork industry, giving it the means to tailor its products for the consumer's more discriminating palate. Third, a more compact market structure has improved the flow of information between consumers and producers, ensuring that pork products are designed with consumer tastes in mind.

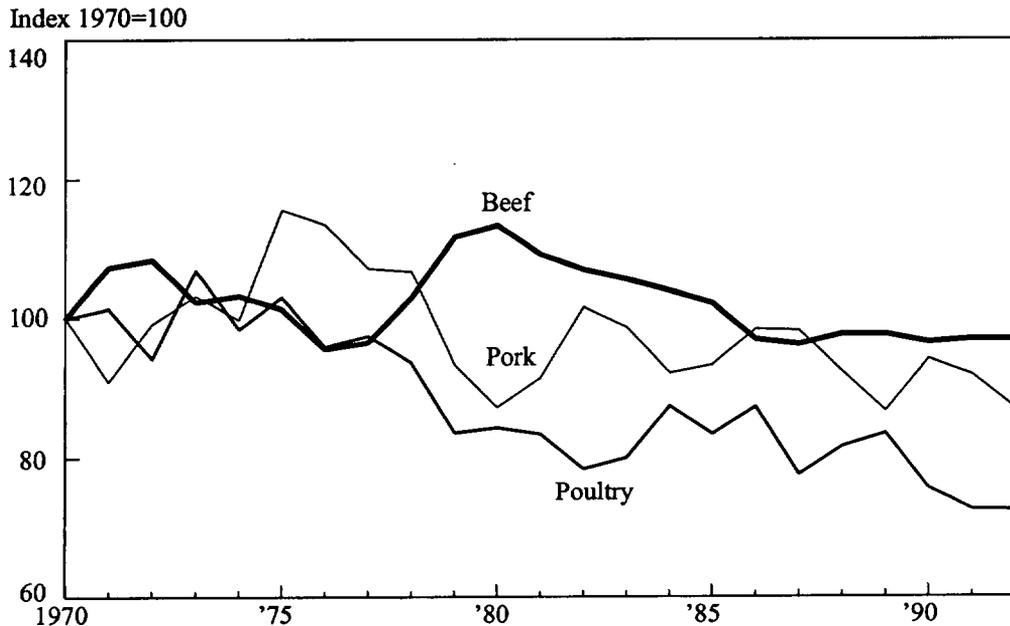
A more discriminating consumer

Consumers have become more discriminating buyers of pork and other food products. Faster paced lifestyles and new concerns about nutrition are dictating changes in American eating habits.

Pork consumption has been nearly flat for the past three decades, while poultry consumption has surged and beef consumption has fallen (Chart 2). The shift to poultry is partly explained by a sharp drop in the price of poultry products. The real price of poultry products has fallen more than a fourth during the past two decades, while beef and pork prices have been nearly flat (Chart 3). But another cause of the consumer's shift to poultry from red meat is the consumer's new penchant for convenience and healthier eating (Barkema and Drabenstott).

In the past, consumers were willing to do the lion's share of meal preparation themselves, purchasing relatively unprocessed food products at neighborhood grocery stores and butcher shops and transforming them into meals in their own kitchens. Modern consumers spend less time in the

Chart 3

Real Retail Prices of Beef, Pork, and Poultry

Source: United States Department of Agriculture.

kitchen, however, aiming to spend no more than 20 to 30 minutes preparing an average meal (Office of Technology Assessment). Strong demographic trends have shortened the consumer's tolerance for preparing meals. The proportion of women aged 25 to 54 in the work force has climbed steadily during the past two decades to about three-fourths, boosting sharply the number of single-individual and dual-income households. Both types of households are believed to spend less time preparing meals than traditional single-earner families. Thus, today's consumers are increasingly shopping for conveniently prepared food products that fit faster paced lifestyles.⁴

Another key element in the consumer's more discriminating demand for food is an increased concern about nutrition. Consumers increasingly

believe they are what they eat. Consumers are especially intent on reducing saturated fat, cholesterol, and sodium in the diet, perhaps in response to the health recommendations of groups like the American Heart Association (AHA).⁵ Concerns about heart disease and cancer have pushed up consumption of fresh fruits and vegetables and pushed down consumption of foods perceived to be fat-rich, like whole milk and red meat (Barkema and others 1991).

In sum, consumers are challenging the pork industry—and other segments of the food system—to tailor food products for more precisely defined market niches. To find success in today's food market, pork products must be conveniently prepared, while cutting back on saturated fat, cholesterol, sodium, and calories.⁶ The poultry indus-

try remains the leader in the race to win consumer favor with a wide range of new products that promise both convenient preparation and healthy eating. But recent marketing efforts by the pork industry, such as its "other white meat" campaign, appear to be gaining ground. The bottom line for the pork industry is that it must satisfy more discriminating tastes at competitive prices if its products are to win consumer acceptance at the supermarket.

A more capable producer

A happy coincidence for the pork industry is that advances in technology have enabled it to tailor its product for the consumer's increasingly discriminating palate. The product engineering process now begins on the hog farm itself, where the industry uses the latest technology to design hogs that produce the leaner pork consumers want. New measuring devices enable swine breeders to select the leanest, most productive animals. Then new reproductive technologies are used to rapidly multiply superior types of animals. Sophisticated computer software tracks and analyzes the performance of entire breeding herds. Thus, the industry is achieving much more rapid improvement in the industry's overall genetic pool. The result is leaner, faster growing, more efficient pigs than the old-fashioned animals of just a decade or two ago.

Meanwhile, advances in nutrition provide carefully formulated diets that enable the new super pigs to grow lean and fast with a minimum production of fat. One of the newest advances in swine nutrition is the development of the swine growth hormone pST, porcine somatotropin. A product of the latest biotechnology, pST promises to boost daily weight gains 10 to 20 percent, boost feed efficiency 15 to 35 percent, and cut fat by 50 to 80 percent. Thus, pST promises another quantum leap in the industry's ability to produce the leaner pork products that today's consumers demand.⁷

New market channels

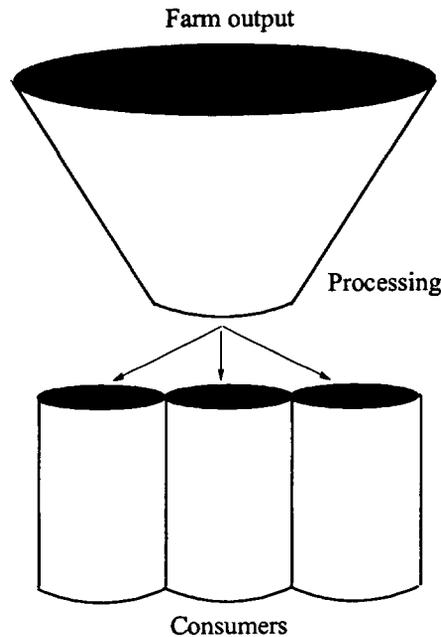
The pork industry has long relied on price signals from commodity markets to guide pork products from farm to grocery. Price signals link consumers to retailers, retailers to wholesalers, wholesalers to processors, and processors to hog farmers. But consumer preferences are becoming more specific than traditional price signals can handle.

Commodity price signals transmit general information well, but they are too fuzzy to transmit the more detailed information required in the modern pork market. For example, the traditional pricing system in the pork industry classifies market-ready hogs into only four quality grades, based on a visual examination of hog carcasses and the expected yield of lean meat. The grades were established decades ago to reward hog farmers for producing leaner hogs. But modern consumers want pork that is even leaner than the leanest of the old grades. Thus, the old grades provide little incentive for hog farmers to produce the kind of pork that today's consumers want.

Thus, new channels of communication are developing to ensure that pork products are properly engineered to meet the modern consumer's tighter specifications. Production contracts and vertical integration are especially effective ways to transmit the consumer's more demanding product specifications to hog farmers. Processors replace fuzzy price signals with crystal-clear contract provisions that specify the genetics, feeding program, and management system to be used on hog farms. Likewise, processors who integrate directly into hog production replace the traditional pricing system with the internal administrative commands of a single firm. Thus, both contracting and vertical integration in the pork industry tighten marketing linkages, ensuring that pork products remain on target for smaller consumer niches.

A sketch of the food system helps illustrate the change in market channels between producers and consumers in the new pork market. In the tradi-

Figure 1

The Traditional Pork Marketing System

tional food system, pork and other bulk farm commodities flowed into the processing sector through traditional commodity markets (Figure 1). The commodity hopper was wide because quality standards were wide. Consumers accepted pork products with broad quality characteristics and then transformed them into meals in their own kitchens.

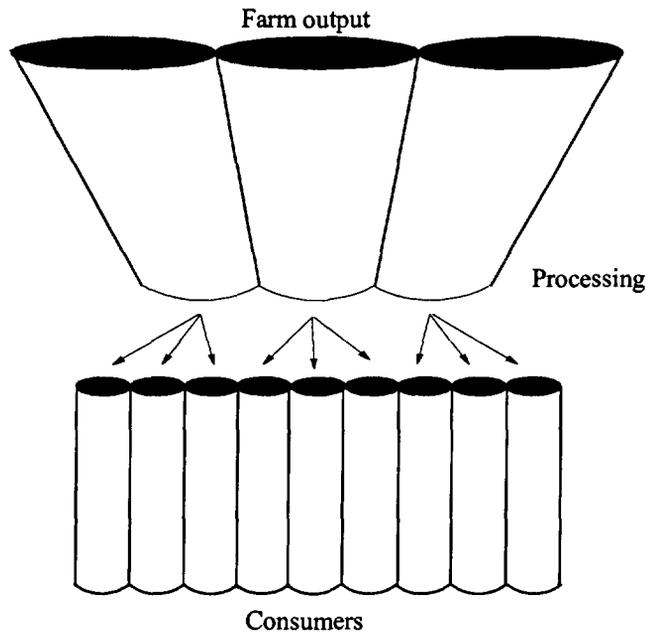
In the new food system, pork and other farm products flow into the processing sector through narrower market channels (Figure 2). The channels are narrower—that is, the diameter of the hoppers is smaller—because the consumers' food specifications are more detailed. Pork products must meet more stringent standards, because today's consumers are willing to spend less time in the kitchen transforming them into the foods they want. Targeting processed foods for smaller con-

sumer hoppers requires that farm products be targeted for smaller processing hoppers.⁸ Thus, the early steps of product development in the new pork market begin on the farm, rather than in the processing plant. For example, the pork industry is learning to trim fat with advances in genetics and nutrition on the hog farm, rather than with knives and cleavers in the butcher shop.

WHAT IS AHEAD FOR THE PORK INDUSTRY?

The U.S. pork industry is headed toward more consolidation and integration. The industry's structural realignment promises further erosion in the number of small hog farms as pork production

Figure 2

The New Pork Marketing System

concentrates in the hands of fewer, larger hog farms with closer ties to pork processors. Powerful economic forces, unleashed by changes in consumer demand and emerging technology, will drive the industry farther down the road toward a more integrated structure. Public policy, however, is geared toward preserving traditional small hog farmers. Thus, the changing pork industry and public policy seem headed for a collision.

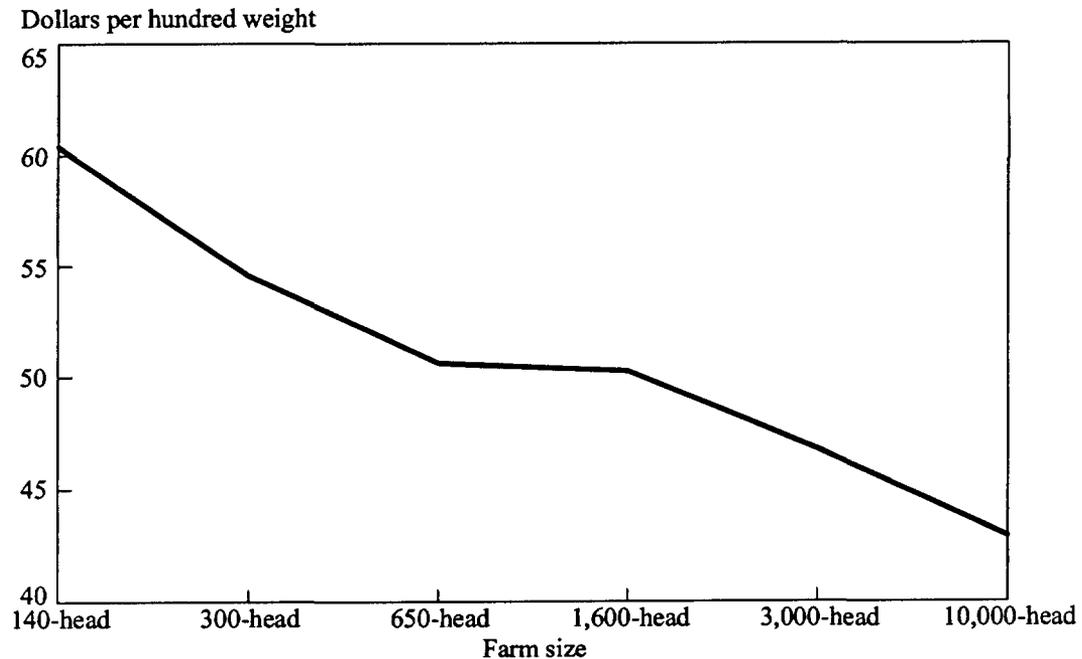
Economic factors behind more consolidation and integration

Two key economic factors will drive the U.S. pork industry toward more consolidation and integration in the years ahead: 1) a drive to cut costs

by capturing economies of scale, and 2) efforts to control the industry's increasing risks. Cutting production costs by shifting production to more efficient, larger farms will fuel further consolidation in the industry. Meanwhile, a further shift toward contracting and integration will reduce the industry's exposure to key sources of risk in the modern pork market.⁹

While both contracting and vertical integration are likely to increase in the years ahead, it is not clear which will be most common. Contract production is likely to be the norm for most large hog producers. But in some regions of scarce hog supplies, processors may ensure a steady supply of hogs for processing lines by integrating vertically into hog production. As more of the industry's processing capacity is met with contract

Chart 4

Average Hog Production Costs in the United States

Source: United States Department of Agriculture "Cost of Production-Livestock and Dairy, 1990."

production and processor-fed hogs, access to markets will shrink for independent hog farmers operating outside the umbrella of production contracts.

Cutting production costs on bigger farms. A steady stream of new production technologies will lead to further economies of scale in hog farming, spurring the current trend toward fewer, larger farms. Many industry observers believe the widening cost disadvantage of small farms points to a swift decline in the number of hog farms, from about 250,000 today to only 100,000 by the year 2000 (Hurt and others). The erosion in hog farms will be fastest among small farms, which are at the greatest cost disadvantage. As small farms leave the industry, hog production will concentrate further on bigger farms.¹⁰

Technology will drive the shift to fewer, larger hog farms. New production technologies are pushing down production costs on the nation's hog farms. But the industry's cost savings are achieved primarily by capturing economies of scale on big farms, where average production costs fall rapidly as the volume of production rises (Chart 4). For example, average production costs on farms producing 10,000 hogs annually, the largest size tracked by the U.S. Department of Agriculture, are nearly 30 percent lower than costs on small farms producing only 140 hogs a year. The steady, downward slope of the cost curve in Chart 4 suggests that even larger production units—like the new mammoth farms producing half a million or more hogs a year—have an even bigger cost advantage over smaller farms.

The newest technologies coming on stream are likely to push down production costs even more, especially on large farms. Some of the new technologies—like pST—are said to be “size neutral,” indicating they could lower production costs on both small and large farms. But these technologies are likely to be used more effectively on large farms where sophisticated production and management systems are already in place.¹¹

The economies of scale that are driving the industry’s consolidation should improve pork’s competitiveness with other meats—especially budget-priced poultry. The sharp drop in the price of poultry relative to beef and pork in the last two decades boosts its appeal to budget-conscious consumers. To remain competitive at the supermarket, the pork industry must hold down its costs—while simultaneously satisfying more discriminating tastes. Thus, further consolidation in the pork industry will be driven by the consumer’s demand for convenient and nutritious pork products in a highly competitive food market.

Reducing risk with contracting and integration. The pork industry is also driven toward more contracting and integration by its growing business risks. Ever larger capital investments expose the industry to bigger losses in risky markets. But hog producers and pork processors are learning to control market risks with contracting and vertical integration.

The industry’s further consolidation on bigger hog farms is making hog production a much riskier business than before. The highly specialized equipment used on big hog farms has little value in uses other than hog production. These modern pig factories have less flexibility than small farms to cut back production or liquidate assets altogether if hog prices fall or if a market cannot be found for market-ready hogs. Thus, uncertain marketing arrangements expose investors in modern hog farms to large losses.¹²

Uncertain marketing arrangements expose pork processors to similar risks. Operating costs in modern processing plants rise sharply if processing lines are not operated at optimal speed.

Thus, processors are exposed to large losses if a steady supply of market-ready hogs is unavailable. Moreover, pork processors must also ensure that the hogs entering their plants are of top quality. Otherwise they run the risk that the pork products they manufacture will be rejected by today’s more demanding consumers.¹³

Ensuring a steady supply of top-quality hogs for modern processing plants is likely to be of increasing importance to processors as pork production consolidates on fewer farms. Processors will have fewer sources of hogs, increasing the importance of continuing marketing relationships with large producers. Thus, the cost-driven consolidation in hog production will work hand-in-hand with the supply risks faced by processors to drive the industry toward a more integrated market structure.

Contracts between hog producers and pork processors are one way of reducing the risk of loss on large investments in specialized assets for both parties. Producers are ensured access to a market for their hogs at a guaranteed price, and processors are ensured a steady supply of hogs to keep processing lines running at optimal speed. Similarly, producers can expect compensation for hogs of exceptional quality—which the traditional pricing system might ignore—and processors can expect that hogs entering their plants will meet the more stringent quality specifications dictated by more demanding consumers.

Vertical integration goes a step farther than contracting in reducing the market risks associated with investments in specialized hog production and processing facilities. By locking together hog farms and processing plants under common management, the integrated firm can ensure a steady supply of high-quality hogs into its processing plants. As a result, production and processing facilities are used at optimum capacity, cutting the risk of loss on big capital investments.¹⁴

The trade-off for cutting marketing risks with contracting and vertical integration, however, is higher management costs. For example, proces-

sors who skirt traditional markets and contract directly with hog farmers face the cost of negotiating and managing numerous production contracts. Likewise, the vertically integrated pork firm faces the bigger burden of managing both hog production and pork processing instead of one or the other. But overall, an accelerating trend toward more contracting and integration in the U.S. pork industry suggests the economic advantages of contracting and integration outweigh any increase in management costs.

THE CLASH BETWEEN PUBLIC POLICY AND THE NEW PORK INDUSTRY

The continued structural realignment of the pork industry portends a new public policy challenge for rural America. Many rural communities will face a decline in local economic activity as the number of small, independent hog farmers erodes. But new, large-scale hog farms will boost economic activity in some other communities. The structural shift in the pork industry, therefore, promises the further concentration of rural economic activity in fewer pockets. Thus, policymakers are challenged to balance the impact of the industry's structural change on rural America with the economic benefits consumers derive from a more efficient industry that delivers higher quality products at lower cost.

Consumers have much to gain from the changes taking place in the pork industry. With stiff competition in the overall food market, the pork industry will pass along to consumers its new efficiency gains in the form of higher product quality and lower prices. Pork products compete directly with other protein foods including beef, poultry, fish, and even new products made from soybeans and other sources of vegetable protein. Thus, the risk is slim that a more concentrated pork industry could exercise monopoly power, which could constrain pork supplies and drive up prices for consumers.

While the structural evolution underway in the pork industry promises significant gains for consumers, it also promises to concentrate a significant source of rural economic activity in fewer pockets. Many rural communities will watch local economic activity dwindle as the number of small hog farmers erodes further. But large-scale production facilities will be a boon to economic activity in some other communities.¹⁵ Thus, the structural shift in the pork industry promises to create a new patchwork of a few rural winners and many losers.

Most public policy concerning farmers is made in Washington, D.C. Price supports for farm commodities, subsidies for farm exports, and other measures designed to boost farm incomes are legislated at the federal level. But the impact on farmers of a more compact pork industry is being debated in state capitals across the farm belt.

Public policy is generally designed to protect the interests of traditional small farmers. As the changing pork industry threatens the livelihood of smaller farmers, it is falling out of step with a host of public rules and regulations. Some of these rules restrict corporate farming in general, others restrict or regulate contracting and vertical integration in livestock production in particular, and still others seek to minimize the impact of concentrated livestock production on the rural environment. Thus, tension is building between public policy designed to protect the traditional pork industry and the new pork industry that is rapidly taking its place.

Pork production in the Tenth District. The tension between tradition-bound policy and a rapidly evolving pork industry is now evident in the Tenth Federal Reserve District.¹⁶ Pork production in the district has always taken a back seat to its northern neighbors in the Corn Belt, where pork production reigns supreme. But restrictions on the industry elsewhere are pushing pork production into parts of the district where it has seldom been seen.

Pork production has traditionally been an important part of agriculture in three district states,

Kansas, Missouri, and Nebraska. Nebraska ranks among the top five hog producing states in the nation, and Missouri and Kansas rank among the top ten. These three states account for the lion's share of pork production in the district and about a seventh of pork production in the nation.

But now pork production is beginning to migrate into other district states—Colorado, Oklahoma, and Wyoming—where the industry was almost nonexistent before. The share of the nation's pork production in these district states is still tiny, but a number of large pork firms, including Seaboard Corporation, DeKalb Swine, Cimarron Pork, Tyson, Pig Improvement Company, and National Farms, have recently chosen these non-traditional states for expansion.

Restrictions on the pork industry's activities in some traditional hog states are guiding the industry into these new production regions. For example, nine midwestern states (Iowa, Kansas, Minnesota, Missouri, Nebraska, North Dakota, Oklahoma, South Dakota, and Wisconsin), including six of the top ten pork producing states, have enacted some form of corporate farming law. The provisions of these laws vary widely, but generally they place restrictions on the farming or land holding activities of large, publicly traded corporations.¹⁷

A recent legislative battle in Iowa—by far the nation's leading hog producing state—and another currently raging in Kansas highlight the policy debate triggered by the pork industry's structural change. Since 1975, Iowa has prohibited processors from contract feeding of hogs. But in 1988, the state passed an exemption to the law that allows cooperatives who own processing facilities to engage in contract hog production with their members. Last year after an intensive debate, an attempt to repeal or add restrictions to the cooperatives' exemption was unsuccessful.¹⁸ Similarly, a 1988 Kansas law prohibits processors from contract feeding of hogs or from owning hogs directly. But a proposal to ease the law's restrictions is a topic of intense debate in the current session of the Kansas legislature.

Meanwhile, the pork industry is skirting the policy debate by moving into new production regions in nontraditional district states. The industry's geographic shift suggests the dividends paid by a new industry structure outweigh the possible drawbacks from moving into uncharted territory. How much of the industry the district gains will depend at least in part on the outcome of the public policy debate. Some states may attempt to preserve the pork industry's traditional structure, only to see the industry migrate to states where the regulatory environment is more accepting of its new structure. Overall, the Tenth District could capture a bigger share of the nation's pork industry, if the industry continues its shift into new production areas.

SUMMARY

The structural realignment underway in the U.S. pork industry will ultimately result in fewer, larger hog farms with closer marketing ties to pork processors and consumers. The industry's new structure promises consumers higher quality pork products at lower cost. But the industry's structural shift will also create winners and losers among traditional hog farms and rural communities, opening a public policy dilemma. Hog farms with a big enough scale of operations and the technical know-how to meet rigorous product requirements will thrive in the new pork market. Many smaller farms will be crowded out. Economic activity will rise in some rural communities and fall in others, as the industry concentrates in fewer pockets. Thus, a clear challenge emerges for public policy. Policymakers must balance the costs of a changing pork industry in rural America with the benefits to consumers of a more efficient industry that promises higher quality products at lower cost.

APPENDIX

HOW DOES THE MARKETING SCHEME AFFECT INDUSTRY COSTS?

The shift toward contracting and vertical integration in the U.S. pork market points to a key question: Under what conditions are these marketing arrangements favored over the traditional system of open production? Coase suggests that the fundamental reason for performing a variety of tasks in a single firm is that entrepreneurs find it cheaper to manage those tasks or services internally than to purchase some or all of them from others.¹⁹

The infant industry

Stigler's description of the industry life cycle is a good starting point for considering how the choice of a marketing structure can affect firm costs. In Stigler's view, the firm is an agglomeration of various processes, such as purchasing inputs, transforming inputs to outputs, and marketing final products. Panel A of Figure A1 depicts the cost structure for a pioneering firm (Firm 1) in a new or infant industry. The firm performs two functions or processes, process A and process B, described by average cost curves AC_A and AC_B , which sum to the firm's average total costs AC_F . The firm has no choice but to manage processes A and B internally, because the firm is virtually the only firm in the new industry.

As the industry grows, however, additional firms enter and the industrywide volume of both process A and process B increases. Eventually, industry volume is large enough to support a firm (Firm 2) which specializes in process B, exploiting economies of scale unavailable to the original pioneering firm. Thus, Firm 1 can lower its total production costs by relying on Firm 2 for process B, which it can buy at a price lower than its own

production costs regardless of volume (Panel B, Figure A1). This simple example suggests open production should become more common and vertical integration less common as a new industry grows, a trend which is opposite that occurring in the U.S. pork industry.

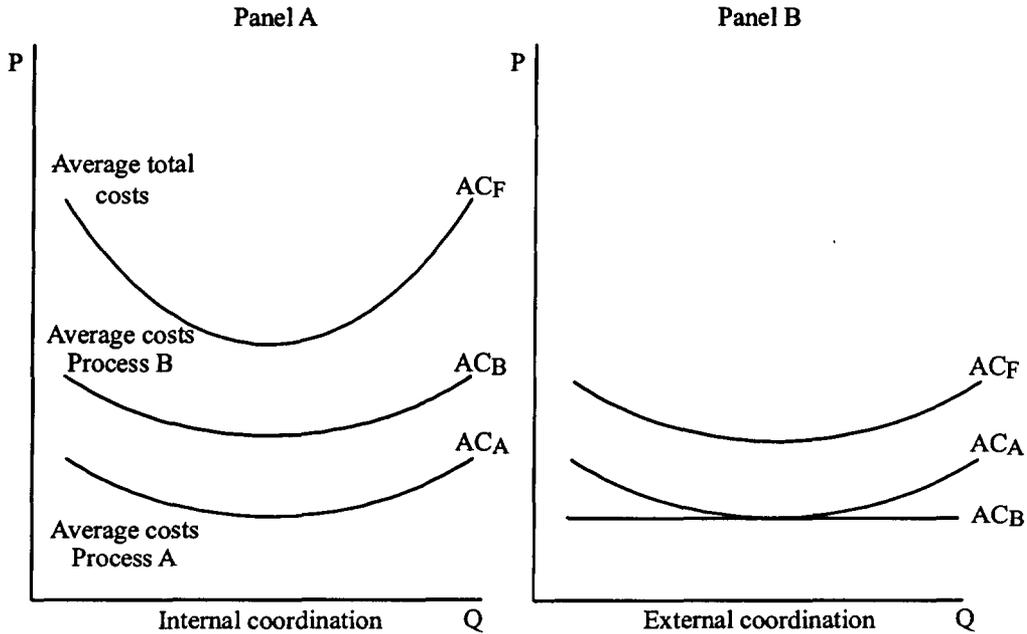
Adding transaction costs

The discussion so far has focused on production costs and ignored transaction costs, the costs of managing marketing relationships, which can change the picture markedly. Williamson's extensive work (1979, 1986) in the area has extended and refined Coase's original argument, attributing the choice of a marketing structure to the firm's cost minimizing decision. Williamson suggests that firms weigh the effects of different marketing schemes on production and transaction costs, with an eye to minimizing their sum.

A few minor adjustments to the sketch of Stigler's growing infant industry illustrate Williamson's ideas. Panel A of Figure A2 again shows the cost structure of the original pioneering firm (Firm 1) after its decision to rely on Firm 2 for process B. But Figure A2 also accounts for Firm 1's transaction costs (ACT)—the costs of managing its relationship with Firm 2.

Figure A2 also provides a useful framework for understanding the structural change underway in the U.S. pork industry by assuming that Firm 1 is a pork processor and Firm 2 is a hog producer. The pork processor's transactions costs might reflect both the cost of searching the countryside for a large enough supply of hogs to keep its processing lines running and the risk that market-weight hogs could be in short supply or of inferior quality. The addi-

Figure A1

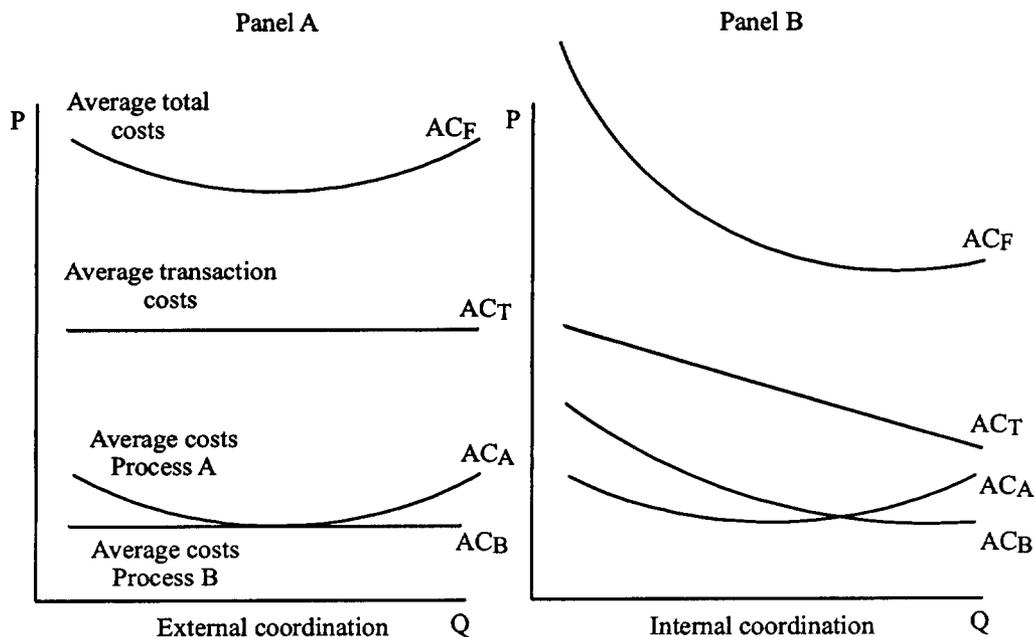
Growth of the Infant Industry

tion of these transaction costs significantly pushes up the pork processor's total average costs (ACF).

As an alternative to using the traditional marketplace, the pork processor and the hog producer may consider a contractual agreement, or the processor may integrate directly into hog production (Panel B, Figure A2). The production contract guarantees the hog producer a ready market at a sure price. And both contracting and integration ensure the processor a steady supply of top-quality hogs to keep processing lines running at optimal speed. Thus, market risks are reduced for both producer and processor. In the absence of contracting or vertical integration, on the other hand, market uncertainties could constrain both the producer and the processor from making large fixed investments in their businesses.

Lower marketing risks could embolden the vertically integrated pork processor to make large investments in hog production, driving down production costs by capturing economies of scale. The processor would probably have to add additional staff to manage the firm's new hog production activities. But the unit cost of the larger management load (ACT) would probably decline as hog production volume rose. Thus, the hog processor could find that transaction costs are smaller under vertical integration than under open production, when production volume is large enough to support the bigger management burden. As a result, at large production volumes, vertical integration enables the processor to maintain low production costs (the sum of ACA and ACB), reduce transaction costs (ACT), and thereby reduce total costs ACF .

Figure A2

Growth of the Infant Industry with Transactions Costs**ENDNOTES**

¹ In 1992, less than a third of the nation's hogs were found in small farm herds of fewer than 500 hogs; in 1970, roughly half the nation's hogs resided in such small herds.

² A trend toward more concentration is also evident in pork processing, although not as pronounced as in pork production. In 1992, the largest four pork slaughter firms accounted for about 42 percent of the nation's pork slaughter, up from 32 percent in 1972 (Hayenga and Kimle).

³ This classification scheme for different kinds of market structure is drawn from Mighell and Jones. See the appendix to this article for a more complete discussion of market structure.

⁴ According to Senaur and others (p. 310), "The history of food and agriculture is a story of gradually shifting roles. First food production, then processing, and now, increasingly,

food preparation have shifted out of the household."

⁵ The AHA advises consumers to limit total intake of meat, seafood, and poultry to no more than 6 ounces per day, use chicken or turkey (without the skin) or fish in most main meals, and to substitute meatless main dishes for regular entrees (American Heart Association 1985).

⁶ Senaur and others (p. 311) summarize the basic nature of food demand stating, "Over time, the fundamental human concerns regarding food remain largely unchanged. People desire a food supply that is reliable and affordable, furnishes the nourishment to sustain life and health, and provides satisfaction and pleasure when consumed."

⁷ While pST promises significant gains in pork quality and production efficiency, the jury is still out on its commercial use. The hormone is currently under review by the

Food and Drug Administration. Consumer acceptance of pork produced with pST is another important question. But survey data suggest consumers' desire to reduce fat outweighs consumers' concerns about any adverse impact of pST on food safety (U.S. Congress, Office of Technology Assessment).

⁸ Another way of describing the fragmentation of the pork market and other segments of the U.S. food market into smaller "hoppers" is to say that the number of unique transactions has increased. The characteristics and frequency of market transactions are key factors determining how markets are structured. See the appendix for a more detailed discussion of factors affecting market structure.

⁹ See the appendix for a more complete discussion of the role of economies of scale and risk in determining the structure of the pork industry.

¹⁰ In 1992, for example, less than 5 percent of the nation's hog farms held inventories of more than 1,000 head, the biggest size tracked by the U.S. Department of Agriculture. But those relatively few large farms accounted for nearly half of the nation's hog inventory.

¹¹ One recent study showed that pST could boost average net income per sow by \$110 to \$134 a year, with the biggest gain attained on the biggest farm in the study. Total net income increased about \$80,000 a year when pST was adopted on the biggest farm, compared with an increase of only \$8,000 a year on the smallest farm. Thus, pST is likely to widen the economic advantage of big farms over little farms (Office of Technology Assessment).

¹² A relatively high capital-labor ratio points to the critical role of large investments in specialized technology on big hog farms. Average production expenses on the 10,000-head farms shown in Chart 3 are lower than on the 140-head farms in every expense category. But the biggest difference is in labor expense, which is nearly two-thirds less on the big farms than on the small farms. As a result, the capital-labor

ratio on the larger farms averages more than half again as large as on the small farms.

¹³ To the authors' knowledge no data on the operating costs of pork processing plants are available. However, most industry observers believe that the shape of cost curves in pork processing plants would be similar to that of beef processing plants. See Barkema and Drabenstott for a more thorough discussion of operating costs in beef processing plants.

¹⁴ Cost reductions gained by locking together previously separate functions like hog production and pork processing are usually called economies of scope.

¹⁵ One Iowa study, for example, notes that the production of 40,000 hogs adds more than \$3 million to the value of the state's home-grown grains and boosts economic activity in nearby communities by more than \$6 million (Kliebenstein and Ryan).

¹⁶ The Tenth Federal Reserve District includes all or part of the states of Colorado, Kansas, Missouri, Nebraska, New Mexico, Oklahoma, and Wyoming.

¹⁷ Some corporate farming laws can also be interpreted as prohibiting the contract production of livestock "either as 'indirectly' engaging in farming or as the control of agricultural land" (Hamilton and Andrews, p. 2).

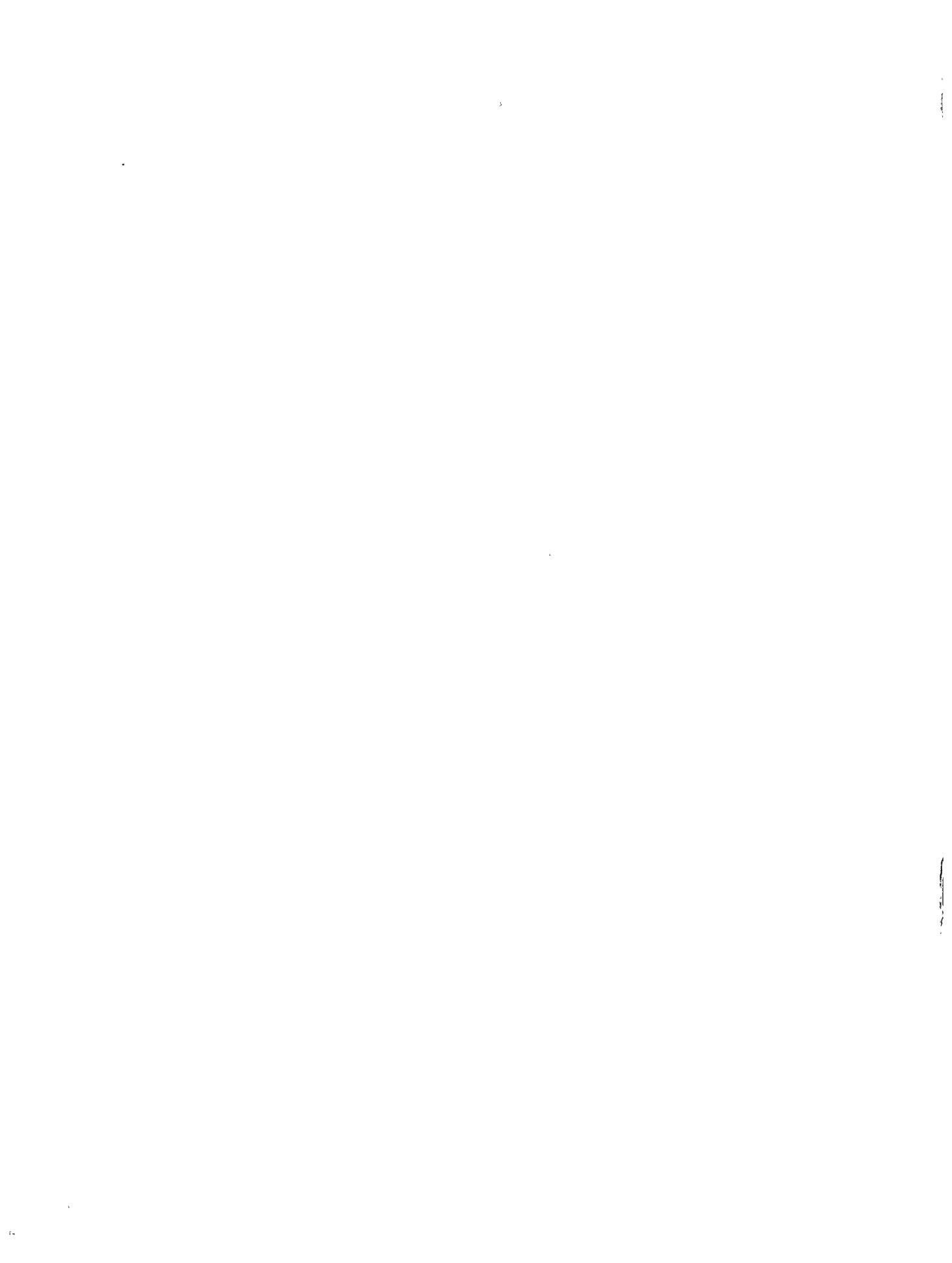
¹⁸ The Iowa Senate passed a bill that would have required cooperatives to receive the approval of 60 percent of its membership before it could engage in contract feeding of hogs. But the bill was amended in the Iowa House to establish a committee to study the issue further (Hamilton and Andrews).

¹⁹ Coase succinctly states the issue, "It is surely important to enquire why co-ordination is the work of the price mechanism in one case and of the entrepreneur in another" (p. 335).

REFERENCES

- American Heart Association. 1985. "The American Heart Diet, An Eating Plan for Healthy Americans."
- Barkema, Alan, and Mark Drabenstott. 1990. "A Crossroads for the Cattle Industry," Federal Reserve Bank of Kansas City *Economic Review*, November/December, pp. 47-64.
- Barkema, Alan, Mark Drabenstott, and K. Welch. 1991. "The Quiet Revolution in the U.S. Food Market," Federal Reserve Bank of Kansas City *Economic Review*, May/June, pp. 25-39.
- Barkema, Alan, Mark Drabenstott, and M. L. Cook. 1993. "The Industrialization of the U.S. Food System," Proceedings of the First Annual Meeting of the Agricultural and Food Marketing Consortium, Orlando, Fla.
- Coase, R.H. 1937. "The Nature of the Firm," *Economica*, vol. 4, pp. 386-405. Reprinted in Stigler and Boulding, eds. *Readings in Price Theory*. Chicago: Richard D. Irwin, Inc., 1952.

- Hamilton, Neil D., and Greg Andrews. 1992. "State Regulation of Contract Feeding and Packer Integration in the Swine Industry," Des Moines, Iowa: Agricultural Law Center, Drake University Law School.
- Hayenga, Marvin, and Kevin Kimle. 1992. "Packing Industry Competitive Structure and Pork Slaughter Capacity," Department of Economics, Iowa State University, Ames.
- Hurt, Chris, Kenneth A. Foster, John E. Kadlec, and George F. Patrick. 1992. "Industry Evolution." *Feedstuffs*, August 24, pp. 1, 18-19.
- Kliebenstein, J.B., and V.D. Ryan. 1991. "Integrating Livestock Industry and Community Development Strategies," conference proceedings of the Livestock Industry and the Environment, sponsored by Iowa State University, Ames, Ia., pp. 113-30.
- Manchester, Alden C. 1992. *Transition in the Farm and Food System*, U. S. Department of Agriculture, Economic Research Service.
- Mighell, Ronald L., and Lawrence A. Jones. 1963. "Vertical Coordination in Agriculture," U.S. Department of Agriculture, Economic Research Service, Farm Economics Division, Agricultural Economic Report No. 19.
- Office of Technology Assessment. 1992. "A New Technological Era for American Agriculture," Washington: U.S. Government Printing Office.
- Rhodes, V. James, and Glenn Grimes. 1992. "U. S. Contract Production of Hogs: A 1992 Survey," Report 1992-2, Department of Agricultural Economics, University of Missouri—Columbia.
- Senaur, Ben, Elaine Asp, and Jean Kinsey. 1991. *Food Trends and the Changing Consumer*, St. Paul: Eagan Press.
- Stigler, George J. 1951. "The Division of Labor Is Limited by the Extent of the Market," *Journal of Political Economy*, vol. 59, pp. 185-93.
- Williamson, O. E. 1979. "Transaction-Cost Economics: The Governance of Contractual Relations," *The Journal of Law and Economics*, vol. 22, pp. 233-61.
- Williamson, O. E. 1986. "Vertical Integration and Related Variations on a Transaction-Cost Economics Theme," in Joseph E. Stiglitz and G. Frank Mathewson, eds., *New Developments in the Analysis of Market Structure*, Cambridge, Mass.: MIT Press.



Are There Too Many Governments in the Tenth District?

By Glenn H. Miller, Jr.

Many taxpayers in the Tenth District and elsewhere have become concerned about their state and local government tax burdens. Yet, few citizens are willing to accept fewer public services to ease their tax burdens. As a result, keeping the lid on public spending requires that state and local governments provide services more efficiently.

It is believed by many taxpayers that excessive spending and taxation by the state and local government sector are due to too many governments. They urge consolidation, aimed at eliminating duplication of effort, as the best means to increase efficiency. One way consolidation might be achieved is by centralizing the state-local sector, that is, by providing services from the state house rather than from courthouses and city halls. Another way consolidation might be achieved is by merging units of local government. But making governments bigger through consolidation does not necessarily make government more efficient or the public sector smaller. Indeed, some researchers suggest just the opposite—that a greater number of governments in a certain area will reduce the overall size of the public sector in that area.

Consolidation of governments is a live issue in the Tenth District, where local governments are abundant and state and local taxes are rising. In Kansas, for example, a legislative committee was charged in 1991 with studying whether efficiency could be improved and taxloads reduced by combining some of the state's 3,800 local governments. The state's economic development agency argued for consolidation, but the Kansas Association of Counties denied that making governmental units bigger would make the public sector run more efficiently. Resolving this debate, in Kansas and elsewhere, will depend partly on how the consolidation of governments will affect the efficiency and overall size of the public sector.

This article reviews the evidence on the consolidation argument, especially on how the number of governments and the structure of the state-local sector affect the efficiency and size of the public sector. First, the article describes the state-local sector in the Tenth District. Next, the article examines whether the public sector is likely to be larger overall if consolidation is achieved by increasing the state's share of state-local activity. The article then examines whether consolidating local governments tends to increase their efficiency and reduce the size of government overall—and if so, what types of local government are more likely to gain from consolidation. The article concludes that those interested in controlling tax

Glenn H. Miller, Jr. is a vice president and economic advisor at the Federal Reserve Bank of Kansas City. Greg Shaffer, a research associate at the bank, helped prepare the article.

burdens must look beyond the number of governments and not jump to the conclusion that consolidation inevitably makes government more efficient.

THE STATE-LOCAL SECTOR IN THE TENTH DISTRICT

Kansas is not alone among district states in having a large number of local governments, nor in its interest in consolidating governments to help curb overall state and local government spending. Indeed, the district has a disproportionate share of the nation's local governments.

Types of local governments

In the Tenth District as elsewhere, local governments can be divided into two types of jurisdictions—general purpose governments and limited purpose governments. General purpose governments comprise counties, municipalities, and townships. Limited purpose governments comprise independent school districts and special districts, such as fire protection or water supply districts.

General purpose governments perform an array of functions. They build and maintain streets, highways, parks, and recreation areas. They also provide health and hospital services, social services, and police and fire protection. And they administer these and other activities. About 40 percent of all local governments in the district are general purpose governments, compared with 47 percent nationwide.¹

Limited purpose governments perform specific functions, such as providing fire protection or operating public schools. Most limited purpose governments perform only a single function, though a few provide several types of services. About 20 percent of all local governments in the district are independent school districts, compared

with 18 percent nationwide. About 40 percent are special district governments, compared with 35 percent nationwide.²

Number of local governments

Tenth District states have a total of 14,254 local governments, or about 17 percent of all local governments in the nation. In comparison, the district has only about 7 percent of the nation's population. Three district states—Kansas, Nebraska, and Missouri—rank among the ten states with the most local governments. Oklahoma and Colorado rank among the top 20. The number of local governments in district states ranges from 3,803 in Kansas to 331 in New Mexico (Table 1).

Despite the large number of local governments in the district, the trend since World War II has been toward fewer local governments. From the early 1950s to the late 1980s the number of governments in the district fell almost by half. The number of general purpose governments fell only modestly during this period, as a growing number of municipal governments more than offset a falling number of townships. Most of the huge drop in the number of local governments came from a decline in limited purpose governments. While the number of special districts doubled, that increase was swamped by an 85 percent fall in the number of school districts, as district states joined in the national wave of school district consolidation and reorganization that swept the country in the 1950s and 1960s.

Some district states have both a large number and a relatively high density of local governments. For example, Kansas and Nebraska have substantially more governments per county than the national average (Table 2). But the average number of local governments per county is lower than the national average in four other district states—considerably lower in New Mexico. Other measures of local government density, such as the

Table 1

Number of Local Governments by Type
Tenth District states, 1987

	Total	County	Municipal	Township	School district	Special district
U.S.	83,186	3,042	19,200	16,691	14,721	29,532
Colorado	1,594	63	266	—	180	1,085
Kansas	3,803	105	627	1,360	324	1,387
Missouri	3,148	115	930	325	561	1,217
Nebraska	3,152	93	534	454	952	1,119
New Mexico	331	33	98	—	88	112
Oklahoma	1,802	77	591	—	636	498
Wyoming	424	23	95	—	56	250
Tenth District	14,254	509	3,141	2,139	2,797	5,668

Source: U.S. Bureau of the Census.

number of governments in relation to land area and population, show a similar pattern. Thus by several measures, Kansas and Nebraska, and to a lesser extent Missouri, have relatively large numbers of local governments, while New Mexico has relatively few local governments.

In a region with such an abundance of governments, district citizens and policymakers naturally wonder if consolidation can offer relief from rising state and local taxes. The answer is not immediately apparent. Just counting governments is not enough to answer questions about whether there are too many governments or whether consolidation would lessen the tax burden. Some economists argue that combining governments increases efficiency in the provision of public services by yielding economies of scale, and that consolidation therefore results in a smaller public sector.

Others argue that when a government is the sole supplier of public services, it becomes bloated and inefficient. A smaller public sector results, they say, when many governments compete with one another.

Resolving this debate requires considering two questions: Can centralizing the state-local sector by increasing the state's share reduce the size of government overall? And can combining local governments into one or more larger jurisdictions reduce the size of government overall?

CENTRALIZING THE STATE-LOCAL SECTOR

Shifting a larger share of total state-local sector activity from courthouses and city halls to the

Table 2

Density of Local Governments*Tenth District states, 1987*

	<u>Square miles per government</u>	<u>Population per government</u>	<u>Governments per county</u>
U.S.	43	2,898	27
Colorado	65	2,051	25
Kansas	22	647	36
Missouri	22	1,610	27
Nebraska	24	507	34
New Mexico	367	4,471	10
Oklahoma	38	1,834	23
Wyoming	229	1,197	18

Source: U.S. Bureau of the Census.

state house is one way of consolidating to gain control of overall spending and taxation. Would such centralizing action increase efficiency and reduce the size of government, as local government activity became a smaller share of the total? Some researchers suggest the answer is probably no.

Fiscal centralization in the state and local sector

Proponents of consolidating more state-local sector activity into state government can point to a nationwide twentieth-century trend in that direction. As state governments expanded into several new services, the state shares both of state and local spending and of total sector revenues

increased substantially from 1900 to the middle of the century.³ From 1900 to 1952, the state share of state and local spending rose from 12 to 35 percent, and state government's share of revenue rose from 18 to 50 percent. After the 1950s, the trend continued but at a much slower pace. By 1989, the state share of state-local direct general expenditures averaged 40 percent, and the state share of total revenue averaged about 56 percent (Table 3).

Fiscal centralization in the Tenth District differs little on average from centralization nationwide. But district states differ widely among themselves in the centralization of their state-local sectors, ranging from relatively low degrees of centralization in Colorado and Kansas to more highly centralized sectors in Oklahoma and New Mexico.⁴

Table 3

State Government Share of Total State and Local Revenue and Expenditures
Tenth District states, 1989

	Revenue * (Percent)	Expenditures + (Percent)
U.S.	55.6	40.0
Colorado	47.0	34.5
Kansas	50.9	38.7
Missouri	56.1	40.4
Nebraska	49.7	41.7
New Mexico	75.1	49.5
Oklahoma	62.4	47.1
Wyoming	58.8	40.6
Tenth District	55.9	41.1

* General revenue from own sources.

+ Direct general expenditures.

Source: U.S. Bureau of the Census.

Centralization and the size of the public sector

Some economists have brought a market-structure perspective to the question of how the state government's share of total state-local activity affects the overall size of the state and local government sector. Long interested in issues of monopoly and competition in the private sector, they have developed an analogous case in the public sector. They believe that large governments may seek to exploit their citizens through excessive taxation, and that competition among many smaller governments is an effective means to prevent such exploitation (Forbes and Zampelli, Oates 1985). As a leading researcher has put it for the state and local government sector in the United States, "other things

equal, those states with a more decentralized fiscal structure should have a smaller state-local sector" (Oates 1985, p. 750).

Several studies covering all states in the United States support the view that state-local sectors tend to be smaller in states with more decentralized fiscal structures (Giertz 1981, Wallis and Oates, and Oates 1989).⁵ That is, the public sector overall is likely to be smaller where the state's share of total state and local government activity is smaller. This finding suggests that more fiscal decentralization of the state-local sector might help hold down the overall size of the public sector and the tax burden.

Fiscal centralization and the size of the public sector also appear to be positively related in the Tenth District. The size of the public sector can be measured by the ratio of state and local govern-

ment spending to the size of the state economy. Using that measure, the size of the public sector in district states, on average, is not far from the national average (Table 4). But the range from the smallest (Missouri) to the largest (Wyoming and New Mexico) is sizable. In district states as nationwide, the more centralized is the state-local sector, the larger is the public sector in relation to the size of the state's economy.⁶ For example, New Mexico has a relatively centralized state-local sector and also a relatively large public sector. In short, centralization does not appear to be the answer to keeping the lid on public spending.

CONSOLIDATING LOCAL GOVERNMENTS

If consolidation by increasing the state's share of state-local activity is not likely to reduce the size of the public sector, would consolidating local governments be more effective? Discussions of this issue have gone on for decades, generally pitting those favoring increased consolidation against those supporting greater decentralization, or fragmentation, as the proper way to minimize local government expenditure for a given level of public services. The debate on consolidating local governments has been waged all across the nation, and the arguments are highly relevant for the Tenth District with its large number of governments.

Consolidation and economies of scale

Supporters of consolidation claim that having too many local governments causes "inefficiencies leading to less than effective methods of providing services, higher per-unit costs, [and] larger government outlays" (Dolan, p. 30).⁷ Consolidation, they point out, permits larger jurisdictions to capture economies of scale that smaller units cannot. Economies of scale exist when the cost per person served for some government activ-

Table 4

Ratio of State and Local Government Spending to Gross State Product, 1989

	Spending as a Ratio to GSP *
U.S.	.147
Colorado	.149
Kansas	.141
Missouri	.118
Nebraska	.144
New Mexico	.183
Oklahoma	.153
Wyoming	.187
Tenth District	.143

* Direct general expenditures.

Source: U.S. Bureau of the Census; Bureau of Economic Analysis.

ity falls as the number of persons served rises. Thus, a certain size and concentration of population are necessary for scale economies to be achieved. "Even though each sub-district may have different desired levels of a public service, small scale provision of these services may result in such a high per unit cost that it outweighs the advantages of diversity" (Giertz 1976, p. 202).⁸

Fragmentation and competition

Supporters of more decentralization of local

government believe that a greater number of governments can better tailor services to meet a broader range of public tastes (Giertz 1976). Moreover, they maintain that decentralization of local government enhances efficiency in providing public services. It does so by increasing the choices of governments available, thereby increasing competition between jurisdictions. Competition, in turn, disciplines inefficient jurisdictions through threatened or actual emigration of taxpayers and tax bases.

A key feature of the argument for decentralization, therefore, is the mobility of citizens and other resources among jurisdictions that offer choices between various sets of public services and tax burdens. Households and firms must be able to move fairly easily between jurisdictions if they are to be able to choose the jurisdiction that best fits their tax and public services preferences, and thereby to exert an influence on the cost of government. This condition may be met more readily in areas of high population density, such as metropolitan areas, than in low density rural areas.

Competitive pressures thus lead local governments to provide the preferred services and to provide them efficiently. Such competition is greater the larger is the number of governments in the specific geographic area. And, because the competition between jurisdictions associated with decentralization increases efficiency, the overall size of the local public sector is reduced (Nelson).

The evidence

Is the overall size of the public sector likely to be reduced by consolidating local governments? Or, is the overall size of the public sector likely to be reduced by fragmentation and competition? The weight of the evidence appears to be on the side of the fragmentation position, especially for general purpose governments.

Several studies support the fragmentation hypothesis (Eberts and Gronberg, Nelson, Zax 1989). These studies typically examine virtually

all counties nationwide and generally support the conclusion that greater decentralization of local governments is likely to yield a smaller public sector overall.

But these studies also reveal a dichotomy: the relationship between fragmentation and the size of the public sector appears to be different for general purpose governments than for single purpose governments. Increased fragmentation of general purpose local governments is likely to result in a smaller public sector. But increased fragmentation of single purpose governments is associated with a larger public sector.⁹ These results arise because of the different effects of fragmentation in different situations. Competition between jurisdictions is a significant factor for general purpose governments. Economies of scale are a significant factor for single purpose governments.

General purpose governments, which have territorial boundaries that prevent them from overlapping with others of the same type, are most likely to have their services shaped by competition. Within an area small enough to minimize changes in residence, job location, and social life, they are typically the jurisdictions among which households and firms choose in selecting their preferred local tax burdens and sets of public services. With enough choice among jurisdictions, citizens and firms can "vote with their feet" so as to influence local governments toward more efficient provision of public services. For these kinds of governments the studies cited here report a clear and strong finding that fragmentation yields a smaller overall local public sector, a relationship "consistent with increased efficiency through competition among local governments" (Zax 1989, p. 564).

A single purpose government is ideally suited to take advantage of economies of scale (Zax 1989). Typically the sole provider of just one service over a large area, a single purpose government may overlap one or more other jurisdictions. Decentralization, which implies smaller sized governmental units, could cause the loss of scale

economies. Moreover, a single purpose government is less likely to be competing with other jurisdictions in its area than would a general purpose government. Consequently, a large number of single purpose governments in an area does not necessarily increase the choices available there and thus does not bring competitive pressures to bear on the provision of public services.¹⁰ Fragmentation among single purpose governments therefore might sacrifice economies of scale without bringing the pressure for greater efficiency that comes from competition between jurisdictions. Public sector size and the number of single-function governments thus might be expected to be positively related (Nelson). The studies cited earlier find that fragmentation among single purpose governments tends to increase the size of the local public sector.¹¹

Not all studies support the fragmentation position. Take, for example, a study of local governments in Illinois (Dolan). This study found a strong positive relationship between fragmentation and increases in the cost of government. That is, the more fragmented were local governments within a specified geographic area, the higher the costs of government were likely to be in that area. Another study, which examined the effects of competition between county governments within metropolitan areas, also concluded that decentralization is positively associated with the size of the public sector (Forbes and Zampelli).¹² While the evidence from empirical studies is somewhat mixed, the results overall appear to favor fragmentation.

The evidence on the consolidation vs. fragmentation discussion may be summarized as follows. Studies of the local government sector generally support the view that a larger number of local governments in a specific area tends to result in a smaller overall size for the public sector in that area. But the relationship between the number of governments and public sector size appears to be different for general purpose governments than for single purpose governments. Increased fragmentation of general purpose governments results in a

smaller public sector, due to the discipline exerted by competition between jurisdictions. But increased fragmentation of single purpose governments results in a larger public sector, due to a loss of those economies of scale for which such governments are well-suited.

The Tenth District case

The research findings just discussed are from studies of the United States as a whole. Do the same kinds of relationships hold for states in the Tenth District? To answer this question, the relationship between the size of the public sector and the number of local governments in district states was studied using regression analysis.¹³ The aim of the analysis was to determine whether fragmentation reduces the size of the public sector in district states.

The results of the regression analysis suggest that decentralization leads to a smaller public sector in Tenth District states. When all of the more than 500 counties in district states were included in the analysis, the results show a small but statistically significant negative relationship between the total number of governments and the size of the public sector. That is, in district states as nationwide, fragmentation of government tends to reduce the size of the local public sector, apparently by encouraging competition between public service providers. Moreover, just as was true nationwide, there is a statistically significant negative relationship between the number of general purpose governments and public sector size in the district. The relationship between the number of limited purpose governments and public sector size is not statistically significant for the district.¹⁴

The relationship between decentralization of local government and the size of the public sector might differ across district states because of substantial differences in population per county.¹⁵ For example, the potential for competition between jurisdictions to reduce the size of the public sector

might be less in counties with smaller populations. To examine a possible difference, all counties in the district were divided into those with populations less than 10,000 (42 percent of the total) and those with more than 10,000 (58 percent). Each group was analyzed separately. The analysis shows a negative, statistically significant relationship between the total number of governments and the size of the public sector in both the smaller county group and the larger county group. The same was true for general purpose governments. And, as was true for all counties together, the relationship between the number of limited purpose governments and the size of the public sector was not statistically significant, either for the smaller or for the larger counties.

In summary, this analysis of district local governments shows that fragmenting local government tends to reduce the size of the public sector in district states as in the nation as a whole. The relationship also holds for general purpose governments in the district. But the district analysis does not show the positive relationship between the number of limited purpose governments and public sector size that has been found nationwide. Finally, there appears to be little difference between these relationships for district counties of less than, and more than, 10,000 population.

SUMMARY

Most citizens of the Tenth District live in states with a large number of local governments. Among their concerns are whether there are too many governments, and whether consolidation would improve efficiency in the provision of public services and lessen the burden of taxation. This article has reviewed a body of research that addresses these issues by studying the relationship

between consolidation and the size of the public sector. While the research findings are complex and not always strictly comparable, some broad conclusions may help citizens and policymakers in district states make informed judgments on issues of government consolidation and decentralization.

First, there is clearly more to deciding questions of consolidation vs. decentralization than just the number of governmental units involved.

Second, research suggests that decreasing the state share of the total state-local sector can help control the size of the public sector. However, consolidation issues may be more fruitfully considered at the local government level alone. Indeed, the mobility of citizens and resources is so important in the analysis that state-level inquiries may be inappropriate and therefore investigations at the local government level may be required (Zax 1989).

Third, while decentralization at the local government level generally results in a smaller public sector overall, the relationship appears to be different for general purpose governments than for single purpose governments. Fragmentation of general purpose governments may be appropriate where jurisdictional competition is likely to be beneficial. But fragmentation of single purpose governments should be approached carefully where economies of scale may exist. Thus, citizens and policymakers should not lump the two types of government together when making decisions about local government structure (Eberts and Gronberg).

Consolidation of governments, then, may not always be the answer for citizens concerned about their tax burdens. In some cases, the route to controlling tax burdens may actually be through decentralization, not consolidation, of governments. In the end, Tenth District citizens who wonder if they have too many governments may find that fewer is not inevitably better.

ENDNOTES

¹ Information in this section is from U. S. Bureau of the Census, *1987 Census of Governments*.

² The local public sector is more heavily weighted toward special districts in the region than in the rest of the country. Four district states—Colorado, Kansas, Missouri, and Nebraska—rank among the ten states with the most special district governments in the nation. Nearly all of the special districts in the region are single function districts, with water supply and fire protection districts among the largest in number. Most special district governments in the region conduct relatively small scale operations, and many do not have property taxing power.

³ This discussion draws heavily on Wallis and Oates. The extent of centralization of the state and local sector is given by the share of state government in the sector as a whole, measured either by the share of revenues raised or by the share of expenditures made. Choosing whether to use revenues or expenditures to measure the relative importance of a level of government is primarily a matter of deciding how to treat intergovernmental transfers. Using a revenue measure attributes the funds involved to the grantor, which initially collects the revenues. Using an expenditures measure attributes such funds to the grant recipient, which eventually spends the money. Researchers have used both measures of centralization, although the spending measure may have received more attention recently.

⁴ Studies of the United States as a whole suggest several geographic and demographic factors related to diversity and population concentration that help explain the observed differences in fiscal centralization from one state to another. For example, size in terms of land area and population, as well as population density and the degree of urbanization, are related to fiscal centralization (Giertz 1976, Wallis and Oates). Fiscal centralization in district states seems to respond to many of the same factors that influence centralization nationally. Colorado, for example, has the largest range in population density between its counties, the largest share of its population in metropolitan areas, and the lowest degree of fiscal centralization among district states. New Mexico, on the other hand, with a small population and low overall population density, has the highest degree of fiscal centralization among district states. Other factors, including unique patterns of historical development and explicit political choices, also contribute to the fiscal centralization that now exists.

⁵ Other studies find that the degree of centralization has little to do with the size of the government sector (Eberts and Gronberg, Oates 1985). Also see the review of the earlier

work in Oates 1989.

⁶ The conclusion for the district states is based on the rank correlation between size of the public sector as measured by total state-local spending as a share of gross state product and centralization as measured by the state share of state-local expenditures. The value of the rank correlation coefficient is +0.5, indicating that a more centralized state-local sector is associated with a larger public sector.

⁷ Those favoring consolidation also cite a number of political and social disadvantages of decentralization, including "confusion in responsibility for service provisions, reductions in political scrutiny and control, political unresponsiveness, [and] units of government concerned only about their own problems," as well as negative impacts on policy issues such as fair housing and school desegregation (Dolan, pp. 30, 43).

⁸ Economies of scale alone may not be enough to tip the balance in favor of consolidation, however. Small jurisdictions may be able to take advantage of scale economies by contracting for services with another government or with the private sector. Doing so could permit them to retain choices about the quantity and quality of services they make available, while taking advantage of the benefits of large size achieved by the contracting agency (Giertz 1976).

⁹ One of these studies reports little or no effect of the number of single purpose governments on the size of the public sector (Nelson, pp. 201, 203).

¹⁰ "In addition, since many special districts provide only minor services and since nearly half of them lack the authority to levy taxes . . . there may be little incentive for individuals to choose between these districts" (Eberts and Gronberg, p 4).

¹¹ Eberts and Gronberg caution that "it may be the case that part of the observed increase in spending associated with greater numbers of units simply indicates that additional special districts are providing additional services" (p. 8).

¹² It has been suggested that the Forbes and Zampelli study may not be directly comparable to the other studies discussed here (Oates 1989).

¹³ The size of the public sector—the dependent variable in the regression analysis—is measured as total local government spending as a proportion of personal income in each county. The fragmentation variable is the number of local governments in the county. The analysis was completed both with the total number of governments, and with local govern-

ments divided into general purpose and limited purpose jurisdictions. The other explanatory variables are per capita income, population, a dummy variable indicating whether or not a county is located in a metropolitan area, and intergovernmental revenue as a share of total revenue. The first three variables represent demand for local public services. The fourth shows the extent to which local governments receive support from higher levels of government.

14 Results for all governments:

<u>Fragmentation variable</u>	<u>Coefficient</u>	<u>T-statistic</u>
Total number of governments:	-.0007	5.59
Results by type of government:		
General purpose governments	-.0014	5.69
Limited purpose governments	-.0003	1.37

15 Some research suggests that fragmentation does not reduce the size of the public sector in counties with populations of less than 10,000 (Zax 1988).

REFERENCES

- Dolan, Drew A. 1990. "Local Government Fragmentation: Does It Drive Up the Cost of Government?" *Urban Affairs Quarterly*, September.
- Eberts, Randall W., and Timothy J. Gronberg. 1988. "Can Competition Among Local Governments Constrain Government Spending?" Federal Reserve Bank of Cleveland, *Economic Review*, Quarter 1.
- Forbes, Kevin F., and Ernest M. Zampelli. 1989. "Is Leviathan a Mythical Beast?" *The American Economic Review*, June.
- Giertz, J. Fred. 1976. "Decentralization at the State and Local Level: An Empirical Analysis," *National Tax Journal*, June.
- _____. 1981. "Centralization and Government Budget Size," *Publius*, Winter.
- Nelson, Michael A. 1987. "Searching for Leviathan: Comment and Extension," *The American Economic Review*, March.
- Oates, Wallace E. 1985. "Searching for Leviathan: An Empirical Study," *The American Economic Review*, September.
- _____. 1989. "Searching for Leviathan: A Reply and Some Further Reflections," *The American Economic Review*, June.
- U.S. Bureau of the Census, *1987 Census of Governments*.
- Wallis, John Joseph, and Wallace E. Oates. 1988. "Decentralization in the Public Sector: An Empirical Study of State and Local Government," in *Fiscal Federalism: Quantitative Studies*. Chicago: The University of Chicago Press.
- Zax, Jeffrey S. 1988. "The Effects of Jurisdiction Types and Numbers on Local Public Finance," in *Fiscal Federalism: Quantitative Studies*. Chicago: The University of Chicago Press.
- _____. 1989. "Is There a Leviathan in Your Neighborhood?" *The American Economic Review*, June.

*Federal Reserve Bank of Kansas City
Economic Policy Symposium Series*

Policies for Long-Run Economic Growth (1992)

Policy Implications of Trade and Currency Zones (1991)

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

Monetary Policy Issues in the 1990s (1989)

Financial Market Volatility (1988)

Restructuring the Financial System (1987)

Debt, Financial Stability, and Public Policy (1986)

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

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

Price Stability and Public Policy (1984)

Industrial Change and Public Policy (1983)

To order copies of any of these symposium proceedings, please write: Public Affairs Department,
Federal Reserve Bank of Kansas City, 925 Grand Boulevard, Kansas City, Missouri 64198.

*Federal Reserve Bank of Kansas City
Research Working Papers*

To order copies of the Research Working Papers published by the Federal Reserve Bank of Kansas City since 1989, please send the number of any paper you wish to receive along with your name and address to: Public Affairs, Federal Reserve Bank of Kansas City, 925 Grand Boulevard, Kansas City, Missouri 64198-0001.

<p>“Tax Analysis in a Dynamic Stochastic Model: On Measuring Harberger Triangles and Okun Gaps” Gregory Huffman and Jeremy Greenwood</p>	<p>RWP 89-01</p>	<p>“Money Supply Announcements and Real Economic Activity” Douglas Joines</p>	<p>RWP 89-10</p>
<p>“Price Flexibility and Macroeconomic Stability: An Empirical Simulation Analysis” John Caskey and Steven Fazzari</p>	<p>RWP 89-02</p>	<p>“An Empirical Study of Comparable Worth” E. Jane Arnault, Louis Gordon, Douglas Joines, and G. Michael Phillips</p>	<p>RWP 89-11</p>
<p>“Discretion, Wage Indexation, and Inflation” David D. VanHoose and Christopher J. Waller</p>	<p>RWP 89-03</p>	<p>“How Bad Is the Federal Budget Deficit?” Douglas Joines</p>	<p>RWP 89-12</p>
<p>“Exchange Rates in the 1980s” Craig S. Hakkio</p>	<p>RWP 89-04</p>	<p>“Discretionary Monetary Policy and Socially Efficient Wage Indexation” David D. VanHoose and Christopher J. Waller</p>	<p>RWP 89-13</p>
<p>“The Output and Inflation Effects of Dollar Depreciation” George A. Kahn</p>	<p>RWP 89-05</p>	<p>“Optimal Monetary Policy in a Multisector Economy with an Economy-wide Money Market” David D. VanHoose and John Duca</p>	<p>RWP 89-14</p>
<p>“Bank Credit Commitments and Credit Rationing” Donald Morgan</p>	<p>RWP 89-06</p>	<p>“Borrowed Reserve Targeting and Nominal Income Smoothing” David D. VanHoose</p>	<p>RWP 89-15</p>
<p>“Optimal Variance Ratio Tests for Serial Dependence and a Test for Mean Reversion” Jon Faust</p>	<p>RWP 89-07</p>	<p>“Bank Profitability, Deregulation, and the Production of Financial Services” Diana Hancock</p>	<p>RWP 89-16</p>
<p>“Market Indicators for Monetary Policy” C. Alan Garner</p>	<p>RWP 89-08</p>	<p>“Pawnbroking in America: The Economics of a Forgotten Credit Market” John P. Caskey</p>	<p>RWP 89-17</p>
<p>“Marginal Income Tax Rates and Real Economic Growth” Craig S. Hakkio and Mark Rush</p>	<p>RWP 89-09</p>		

"The Prepayment Experience of FNMA Mortgage-backed Securities"	Sean Beckett and Charles S. Morris	RWP 90-01	"Is Real Exchange Rate Variability Caused by Relative Price Changes? An Empirical Investigation"	Charles Engel	RWP 91-02
"The Foreign Exchange Risk Premium: Is It Real?"	Craig S. Hakkio and Anne Sibert	RWP 90-02	"A Comment on William Poole's Paper 'Exchange Rate Management and Monetary Policy Mismanagement: A Study of Germany, Japan, United Kingdom, and United States After Plaza'"	Craig S. Hakkio	RWP 91-03
"Real and Nominal Exchange Rates Since 1919"	Craig S. Hakkio and Douglas H. Joines	RWP 90-03	"Can the Markov Switching Model Forecast Exchange Rates?"	Charles Engel	RWP 91-04
"Stock Prices, News, and Business Conditions"	Grant McQueen and V. Vance Roley	RWP 90-04	"IRAs and Saving: Evidence from a Panel of Taxpayers"	Douglas H. Joines and James G. Manegold	RWP 91-05
"Tests of Mean-Variance Efficiency of International Equity Markets"	Charles Engel and Anthony Rodrigues	RWP 90-05	"Public Finance and Coordination Problems in a Common Currency Area"	Anne Sibert	RWP 91-06
"On the Foreign Exchange Risk Premium in a General Equilibrium Model"	Charles Engel	RWP 90-06	"Imperfect Information and Financial Constraints: New Evidence Using Bank Loan Commitments"	Donald P. Morgan	RWP 91-07
"The Risk Premium and the Liquidity Premium in Foreign Exchange Markets"	Charles Engel	RWP 90-07	"Strategic Capital Taxation in Large, Open Economies with Mobile Capital"	Jiming Ha and Anne Sibert	RWP 92-01
"Cointegration: How Short Is the Long Run?"	Craig S. Hakkio and Mark Rush	RWP 90-08	"Can Unconventional Preferences Explain Risk Premia in the Foreign Exchange Market?"	Anne Sibert	RWP 92-02
"Government Finance in a Common Currency Area"	Anne Sibert	RWP 90-09	"How Important Are Monetary and Fiscal Policy in Explaining Postwar Aggregate U.S. Data?"	Peter E. Kretzmer	RWP 92-03
"Bank Credit Commitments"	Donald P. Morgan	RWP 90-10	"Allais Theory Suggests a Solution and an Explanation for the Equity Premium Puzzle"	John E. Golob	RWP 92-04
"Realignment in Farm Lending: Strategic Issues for the 1990s"	Alan Barkema and Mark Drabenstott	RWP 91-01			

“Business Cycle Fluctuations in U.S. Regions and Industries: The Roles of National, Region-Specific, and Industry-Specific Shocks”
Todd Clark RWP 92-05

“A Regime Shift in Measured Per Capita Consumption, with Implications for Asset Prices and Returns”
John E. Golob RWP 92-06

“Does Money Matter Anymore? A Comment on Friedman and Kuttner”
Sean Beckett and Charles Morris RWP 92-07

“Monetary Policy Implications of Bank Security Holdings and Large CDs”
William R. Keeton RWP 92-08

“Bank Loan Commitments and the Lending View of Monetary Policy”
Donald P. Morgan RWP 92-09

“Bank Representation in Low-Income and Minority Urban Communities”
John P. Caskey RWP 92-10

“Determinants of Rural Growth: Winners and Losers in the 1980s”
Tim R. Smith RWP 92-11

“Monetary Transmission Channels in Major Foreign Industrial Countries: A Comment”
Craig S. Hakkio RWP 92-12

“Business Fixed Investment Spending: A Critical Survey of Modeling Strategies, Empirical Results, and Policy Implications”
Robert S. Chirinko RWP 93-01

“Determinants of the Decline in Union COLA’s”
Stuart E. Weiner RWP 93-02

“Bubbles, Fundamentals, and Investment: A Multiple Equation Testing Strategy”
Robert S. Chirinko and
Huntley Schaller RWP 93-03

Economic Review

President

Thomas M. Hoenig

Senior Vice President and Director of Research

Thomas E. Davis

Economic Research Department

Mark Drabentstott, Vice President and Economist

Bryon Higgins, Vice President and Economist

Glenn H. Miller, Jr., Vice President and Economic Advisor

Craig S. Hakkio, Assistant Vice President and Economist

George A. Kahn, Assistant Vice President and Economist

Charles Morris, Assistant Vice President and Economist

Gordon H. Sellon, Jr., Assistant Vice President and Economist

Stuart E. Weiner, Assistant Vice President and Economist

Alan Barkema, Senior Economist

Sean Beckett, Senior Economist

C. Alan Garner, Senior Economist

William R. Keeton, Senior Economist

Donald P. Morgan, Senior Economist

Tim R. Smith, Senior Economist

Todd Clark, Economist

Andrew J. Filardo, Economist

John E. Golob, Economist

Visiting Scholars

Robert S. Chirinko, University of Illinois (Urbana/Champaign)

Charles Engel, University of Washington

Keith E. Maskus, University of Colorado

Anne C. Sibert, University of Kansas

Second Quarter 1993, Vol. 78, No.2

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

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

The views expressed are those of the authors and do not necessarily reflect the positions of the Federal Reserve Bank of Kansas City or the Federal Reserve System. If any material is reproduced from this publication, please credit the source.



Economic Review
Federal Reserve Bank of Kansas City
Kansas City, Missouri 64198-0001
Second Quarter 1993, Vol. 78, No. 2