Federal Reserve Bank
of San Francisco

Spring

1985

Economic Review

Number

2

The Federal Reserve Bank of San Francisco's Economic Review is published quarterly by the Bank's Research and Public Information Department under the supervision of Joseph Bisignano, Senior Vice President and Director of Research. The publication is edited by Gregory J. Tong, with the assistance of Karen Rusk (editorial) and William Rosenthal (graphics).

For free copies of this and other Federal Reserve publications, write or phone the Public Information Department, Federal Reserve Bank of San Francisco, P.O. Box 7702, San Francisco, California 94120. Phone (415) 974-3234.

Competition for Money Market Deposit Accounts

by
Michael C. Keeley
and
Gary C. Zimmerman*

The money market deposit account (MMDA) is the first liquid, short-term, small denomination deposit account in recent history to be free from interest-rate ceilings. Introduced in December 1982, it has become a very important source of deposits, with balances currently over the $450 billion level nationally, which represents about 15 percent of total deposits. In this paper, we analyze why the account has been so successful, where the funds in the account came from, how the shift in funds affected banks' deposit costs and how individual banks competed for and priced MMDAs.

The single most important step in the deregulation of interest rates on retail deposits at banks and thrifts was the authorization of the money market deposit account (MMDA) and the Super-NOW account, which were both free of interest ceilings.[1] The volume of funds that moved into MMDAs was staggering: MMDAs grew to over $300 billion (15 percent of total deposits) within three months after their introduction on December 14, 1982. Super-NOW accounts, however, attracted only some $30 billion during the three-month period after their introduction on January 5, 1983.

In this paper, we analyze why and how the MMDA so dramatically altered both depositors' and banks' holdings of deposits and the nature of competition for deposits. Our objectives are to explain why these accounts were so popular, where the funds came from, what determined individual banks' pattern of adoption of these accounts, and how sensitive the quantity of funds in these accounts was to variations in both their own interest rate as well as the rate on substitute assets. Our analysis focuses on the MMDA rather than the Super-NOW because the MMDA's impact on deposit holdings was so much larger.

Money market deposit accounts are an insured, short-term, ceiling-free account with limited transaction features that were, in the language of the authorizing Garn-St Germain Act of 1982, to be ". . . directly equivalent to and competitive with money market mutual funds. . .". However, prior to their introduction, there was considerable uncertainty about the sources and stability of MMDA deposits and how they would be priced. Some thought that MMDAs would attract large quantities of money from the money funds, and there even was speculation about the long-term viability of the money funds given that MMDAs were insured. Others thought that most MMDA deposits would come from other funds already on deposit at banks and thrifts. If large amounts

*Senior Economist and Economist, respectively. Research assistance from Joni Whitmore and Maureen O'Byrne and comments from Jack Beebe, Fred Furlong, Hang-Sheng Cheng, Randall Pozdena and Barbara Bennett are much appreciated.

of low-interest "core" deposits like passbook savings accounts were transferred into MMDAs, there were fears that banks' and thrifts' deposit costs would rise substantially. There was also uncertainty about how quickly funds would be attracted to MMDAs and whether they would be responsive to minor rate changes (as are large, $100,000 and over, Certificates of Deposit, CDs), allowing individual institutions to increase their market shares by slightly outbidding their competitors. Moreover, it was unclear whether the MMDA's rate-sensitivity would change over time—with balances more rate-sensitive during the initial introductory period than once depositors had shifted funds into the new accounts.

We address these issues by analyzing the competition for MMDAs, both among banks as well as between banks and the money funds. Our empirical analysis of these issues utilizes monthly data on the rates and quantities of deposits of MMDAs and other accounts for a sample of 59 banks in the 12th Federal Reserve District. (See Data Appendix.) Data for individual banks, unlike aggregate data, enable us to address questions of interbank competition, which most previous studies of deposit flows cannot.[2] We do, however, use nationwide aggregate data to estimate the flows of funds into MMDAs from the money funds and from other deposit accounts.

Our analysis uses both micro and macro data to estimate the parameters of the supply function of MMDA deposits to banks. Both short- and long-run, own- and cross-price elasticities are estimated. In addition, we analyze the process of adjustment in financial markets both to the introduction of a new account and to changes in rates once the initial adjustment is complete.

The organization of the remainder of this paper is as follows. In Section I, we provide a brief description of the MMDA account and a historical perspective on why it was introduced. Aggregate data for the nation are used to measure flows of funds into MMDAs, both from the money market funds and from funds already on deposit with banks and thrifts. Section II outlines a theory of banks' demand for MMDA deposits and the nature of deposit adjustment costs. Implications of this theory for the market acceptance and pricing of MMDAs are developed as well. In Section III, the competition for MMDAs is analyzed. First, the time pattern of adoption of MMDAs is modeled as a function of banks' characteristics and pricing strategies. Then, estimates of the short- and long-run, own- and cross-interest elasticities of the supply of deposits to banks are presented. Section IV presents our summary and conclusions.

I. Background

Money market deposit accounts were in part a regulatory/legislative response to the success of the money market funds (MMFs) which, by late 1982, had attracted well over $240 billion, much of it from traditional bank and thrift depositors. Money funds offered significant advantages over the regulated accounts offered by depository institutions (hereafter referred to as banks). First, the funds paid returns to investors near the wholesale money market rate. Yields were determined by the return on the funds' portfolios of short-term money market instruments less a small administrative fee. Second, the funds were generally liquid, allowing investors transaction and check-writing privileges. These features were not available on banks' "money market certificates", a six-month time account with an indexed ceiling rate that yielded only slightly less than a six-month Treasury bill. Third, most funds' minimum opening balance was well below the $10,000 minimum required for banks' six-month "money market certificates" or the $20,000 minimum on 7- to 31-day time certificates. Fourth, the money funds were able to raise "deposits" on a nationwide basis, effectively skirting the prohibition that banks faced on interstate deposit-taking. Finally, MMFs could be linked with other mutual funds and security transactions.

Because of the money funds' dramatic growth, there was considerable political pressure to allow banks to offer comparable instruments so that they could compete on an equal footing. As it turned out, the MMDA and, to a much lesser extent, the Super-NOW accounts were just such instruments. Without them (or accounts similar to them), it appeared that depository institutions might continue to lose retail deposits to the money funds.

MMDA Terms

Although the MMDA was patterned after the MMFs, there are some differences, primarily regarding reserve requirements and regulatory limitations on transactions and minimum balances. The MMDAs are free of interest rate ceilings as long as a minimum balance of $1,000 is maintained ($2,500 prior to January 1, 1985), and are insured to $100,000. The MMDA is available to all depositors, including individuals, governments, nonprofit institutions and businesses—although non-personal deposits, unlike personal deposits, are subject to a 3 percent reserve requirement. In addition, unlike the money funds, MMDAs have transaction features that are restricted by regulation. Depositors are allowed up to six automatic, telephone or check transfers per month (with a maximum of three check transactions), although withdrawals made in person are unlimited. (See Appendix Table 1 for a detailed description of the characteristics of MMDAs, Super-NOWs and the money funds.)

Sources of MMDA Deposits

Where would the funds for the MMDA come from? In general, funds would be expected to come from other financial instruments that were close substitutes (from both banks' and depositors' perspectives) for regulated retail deposits. Since, at the time of the MMDA's introduction, open-market interest rates had been above the ceilings for a number of years, depositors who did not value the implicit interest in terms of added services would already have moved their funds out of the low-paying retail accounts. Thus, it seemed unlikely there would be a further shift of funds out of low-interest retail deposits, such as passbook savings, into MMDAs.

Rather, the funds for MMDAs, which have limited value as transaction accounts, would come from those sources where depositors had moved them in the first place, to avoid the ceilings. We thus expected particularly large inflows into MMDAs from the money funds. Depositors had shifted balances from regulated retail accounts into the money funds apparently because they viewed them as close substitutes. Moreover, since banks had used large CDs to replace lost retail deposits, we expected declines in these balances as well. Also, to the extent that the ceilings had induced depositors to move funds into other near ceiling-free accounts, such as the six-month money market certificate, some of those funds probably also would be moved into MMDAs.

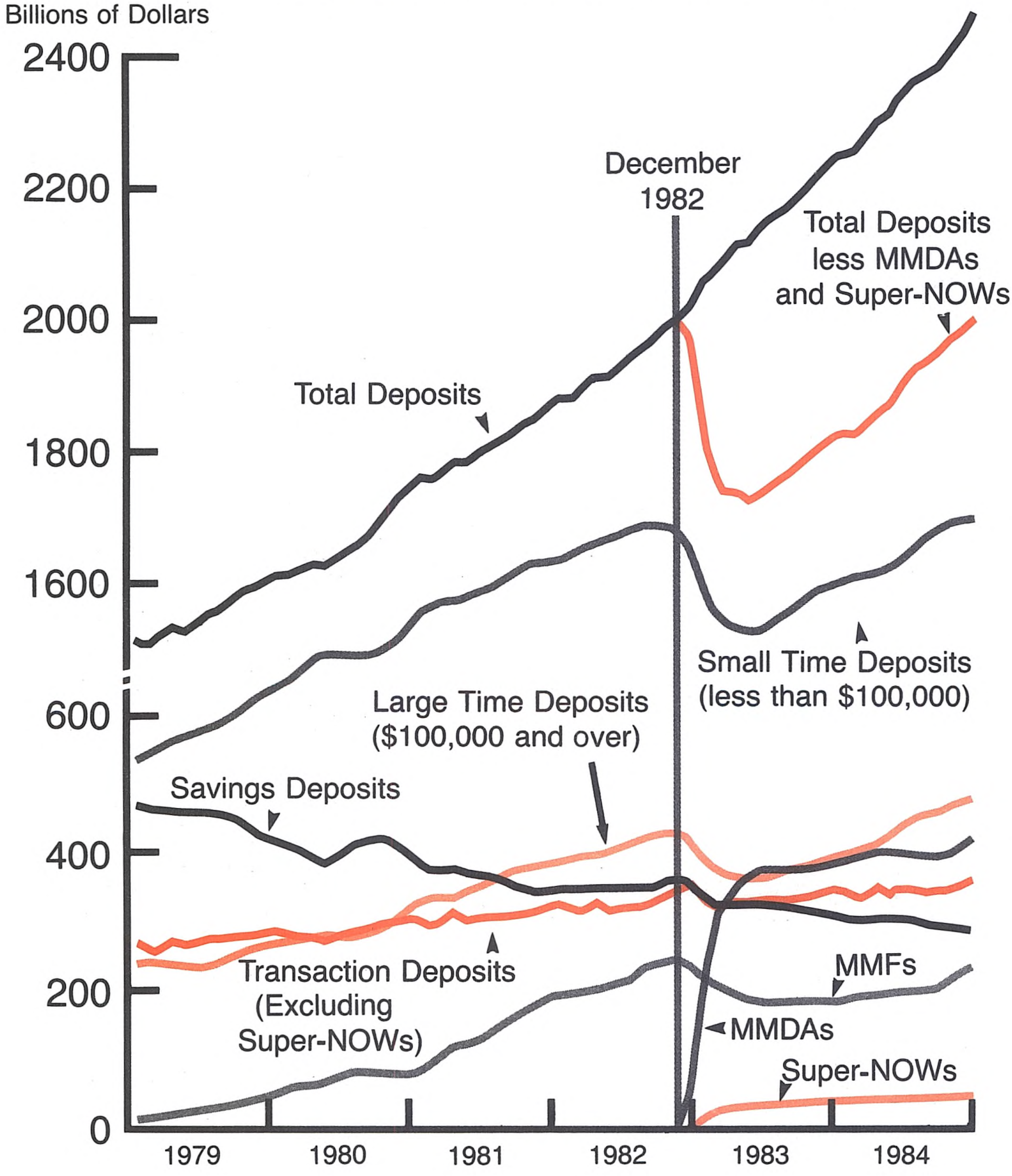
By analyzing the decline in various types of deposits that were contemporaneous with the MMDA's initial rapid deposit growth, it is possible to infer which types of deposits were probably the most important sources of MMDA funds. For banks and thrifts, aggregate deposit data (at the national level) for small and large time deposits, savings deposits and transaction deposits are analyzed along with data on money funds' assets. In Chart 1, these various deposit stocks are plotted along with MMDA and Super-NOW deposits to indicate which types of deposits fell as MMDAs grew. Also, monthly changes in various types of deposits were regressed on monthly changes in MMDAs (and other control variables) to provide quantitative estimates of the sources of MMDA deposits.[3]

Chart 1 shows that a substantial decline in the money market funds' assets coincided with the growth in MMDAs, suggesting that the money funds were substitutes for MMDAs and thus an important source of MMDA deposits. Similarly, our regression model shows a statistically significant decline of $.24 in money fund assets for each dollar that flowed into MMDAs.

Although the data suggest that the money

Chart 1

Money Market Funds vs. Components of Total Deposits (Combined Bank and Thrift Total)



funds were an important source of MMDA deposits, the money attracted from the MMFs did not lead to a comparable increase in the total deposits of the banking sector. As Chart 1 shows, there was only a slight increase in total deposits after December 14, 1982 despite the influx of funds into MMDAs from the money funds. Thus, the MMDA inflows must have been mostly offset by outflows from other types of deposits. In particular, we would expect some of the funds in banks' large CDs to leave the banking sector as money funds' assets ran off. In part, this is because of the direct effect of a reduction in the money funds' holdings of banks' CDs as the money funds contracted. Moreover, as banks experienced rapid inflows into MMDAs, they would reduce their purchases of relatively more expensive CDs, and some of the previous holders of large CDs would move their funds out of banking deposits. We find that there was indeed a statistically significant decline in large time deposits (CDs) of $.42 for each dollar increase in MMDAs. This drop was considerably larger than the decline in the money funds' assets.

This rather massive substitution of retail MMDAs for wholesale CDs has important implications for bank costs. Since MMDAs, like other retail deposits, are generally less costly than large CDs, this shift alone lowered banks' deposit interest costs. The inflow from money funds also has implications for their long-run viability. Although we estimate that MMDAs attracted nearly $90 billion from the money funds, this appears to have been a one-time shift since the money funds have continued to grow despite competition from the MMDA. Indeed, MMF assets have rebounded to near their pre-MMDA peak.

There was also a dramatic, statistically significant decline of $.52 in small-denomination (less than $100,000) time deposits for each dollar increase in MMDAs. After an actual decline of nearly $150 billion over a six-month period, small time deposits resumed their trend growth rate as the growth in MMDAs tapered off. This pattern suggests that there was a one-time shift of funds. The largest decline in the small time category took place in the popular six-month money market certificate, which already paid a near (wholesale) market rate of interest, but tied funds up for six months.

The impact of the switch from small time accounts to MMDAs on the cost of funds was probably not too large for many institutions. Nearly all of the funds in the small-time category were already paying near open-market rates or were tied to those rates by the end of 1982. Still, even though this switch did not directly alter the cost of these funds substantially, it changed the overall composition of deposits and shortened the (stated) maturity distribution for retail deposits.

Both savings (passbook accounts, at the time paying 5¼ percent at banks and 5½ percent at thrifts) and transactions balances—including demand deposits, Negotiable Orders of Withdrawal (NOW) and Automatic Transfer Service (ATS) accounts, and savings deposits authorized for telephone and preauthorized transfers, but excluding Super-NOWs—also appeared to fall slightly during the months following the authorization of the MMDA. However, the regression analysis does not provide evidence of a statistically significant shift. It is likely that the declines did not represent actual shifts because of build-ups (evident in Chart 1) in both savings and transactions balances in the weeks preceding the authorization of the MMDA. Knowing that banks would be authorized to offer these short-term market rate accounts as of December 14, many depositors with maturing instruments likely held funds temporarily in transaction or savings accounts until the new accounts were available.

The lack of a significant shift from savings accounts into MMDAs suggests that (passbook) savings accounts must be offering a large service benefit that offsets the binding effect of the interest ceiling. (The effect of binding interest ceilings is discussed in the Box.) This confirms our hypothesis that the gradual erosion of these accounts had left mainly depositors that prefer implicit interest in the form of non-taxable services rather than taxable interest. Our findings are consistent with evidence provided by Furlong (1984) that savings accounts—through

The Effect of Interest Rate Ceilings on Deposit Costs

To see why interest ceilings actually led to higher average and marginal deposit costs for banks operating in competitive deposit markets, consider the following analysis. A bank can "produce" or attract a given quantity of deposits, Q_O, by varying the relative amounts of unpriced service and interest payments it makes to its depositors. Such a tradeoff between interest and service payments is depicted in Figure 1.

At very low service levels, consumers are willing to give up a dollar of interest for less than a dollar of services, that is, services that cost the bank less than a dollar to provide. This is because of tax considerations (interest payments are taxable and service payments are not), the joint production of interest and services, and because of the transactions costs associated with buying or paying for each service separately. However, because of diminishing marginal utility, as service levels increase, consumers are willing to forego less and less interest per dollar of services. Thus, at high service levels, consumers are willing to give up a dollar of interest only in return for more than a dollar of extra service.

A bank seeks to minimize the cost of attracting a given quantity of deposits. The bank's cost of a given quantity of deposits is simply the sum of its interest and service payments, which is depicted as the 45° line in the figure. The minimum-cost combination of interest and service payments to attract Q_O deposits is given by the point of tangency in Figure 1, where the consumer is willing to forego a dollar of interest for a dollar of services. This is equal to the tradeoff the bank faces in providing interest and services.

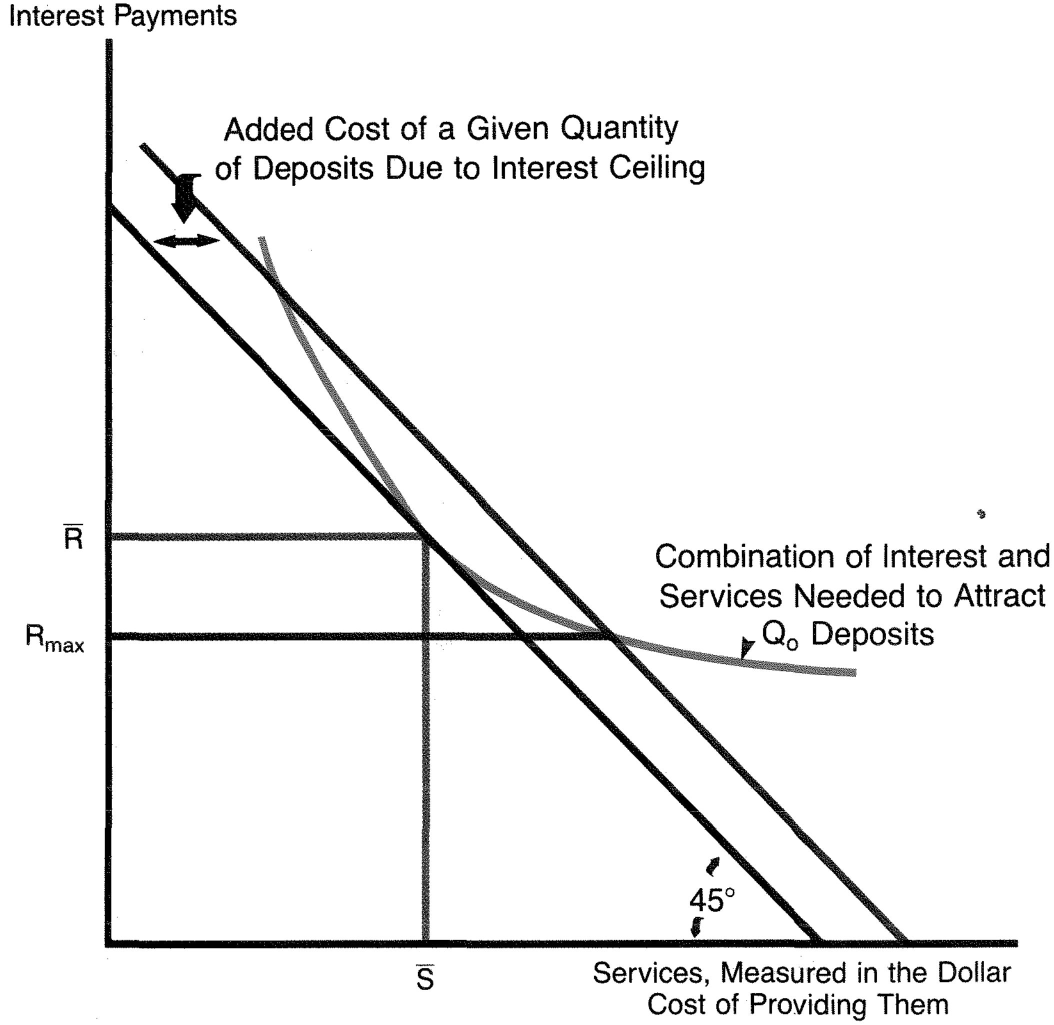
However, if an interest ceiling is imposed that limits interest payments to be less than $\bar{R}$—the interest the bank would have paid in the absence of ceilings—then the bank will be able to attract the same quantity of deposits only by offering services in excess of $\bar{S}$. Since a unit of these added services will be valued at less than a dollar even though they cost a dollar to provide, the bank will face higher total costs of attracting the same quantity of deposits. That is, average deposit costs will rise. It can be shown that the marginal cost of deposits will also rise. Thus, this analysis shows that ceilings that cause banks to pay less interest than they otherwise would leads to an inefficient substitution of service for interest payments and causes the average cost of deposits to rise.

A similar type of analysis can be performed for banks with access to wholesale markets for funds (the large CD market). The tradeoff between wholesale and retail deposits is similar to that depicted in the figure between interest and service payments. This is because there are costs of size intermediation that prevent wholesale deposits from being perfect substitutes for retail deposits. (Perfect substitutability would mean a linear tradeoff and also that banks would not take any retail deposits when ceilings are imposed.) Thus, interest ceilings force banks, which have access to wholesale markets, to rely more heavily on wholesale deposits than they otherwise would. Thus, the ceilings drive up the average and marginal costs of banks' deposits.

Since MMDAs have no interest ceilings, they reduce the average and marginal costs of deposits for banks constrained by ceilings in competitive markets. Thus, both banks and depositors have substantial incentives to switch funds into MMDAs.

Figure 1

Tradeoff Between Services and Interest Needed to Attract a Given Quantity of Deposits



high turnover—provide a substantial transaction service component different from that offered with MMDAs. Thus, the nearly $290 billion in savings deposits currently on the books are not likely to be shifted suddenly out of savings in response to higher returns on MMDAs.

Surveys of the sources of MMDAs made during the rapid growth period provide results similar to those discussed above, although most indicate a higher proportion of new funds.[4] As previously discussed, our evidence suggests that although a significant fraction of MMDA deposits were attracted from the money funds, these inflows were offset by outflows from other deposits. Although MMDAs apparently had only a minor impact on total deposit growth, they significantly altered institutions' mix of deposits—increasing retail deposits at the expense of wholesale deposits—and they were successful in allowing banks to compete with the money market funds for retail deposits.

II. Market Adjustment

In theory, the growth of funds in MMDAs would be determined jointly by households' and banks' portfolio decisions. Banks, within the limits imposed by competition, would set the rates (and other terms) on these accounts, and consumers could reallocate their portfolios in response to those rates subject only to the costs of such reallocations. Banks, of course, might try to take into account households' responsiveness to interest rates when setting rates, but households' reallocations might also depend on their expectations about how banks would price these accounts in the future. Furthermore, information and other unique transaction costs associated with moving funds into MMDAs might significantly affect the flow of MMDA funds.

Banks' Demand for MMDAs

Interest ceilings had led banks that were in competitive markets to substitute nonpriced services for interest payments and wholesale deposits for retail deposits. By allowing banks to attract funds directly through price competition, MMDAs were a lower cost means of attracting funds than had existed previously.[5] (See Box.) This regulatory innovation lowered the average and marginal costs[6] of attracting deposits and simultaneously increased the effective rate paid to depositors, and thereby provided a strong incentive for banks to attract funds into these accounts as well as a strong incentive for depositors to shift funds into these accounts.[7]

MMDAs' very rapid acceptance by the marketplace confirms that there were substantial cost savings for banks in offering such accounts and that depositors preferred the combination of explicit interest, maturity and services offered by MMDAs to those available on at least some pre-existing accounts. However, because both banks and depositors would probably experience adjustment costs associated with opening such accounts and shifting funds into them, the adjustment of actual to desired stocks of funds in these accounts would not be instantaneous. These adjustment costs have a number of important implications discussed below.

Adjustment Costs

As Flannery (1982) and others have shown, the existence of bank-specific transaction and information costs mean that retail deposits have a specific capital component making them a "quasi-fixed" factor of production.[8] That is, retail deposits are somewhat like specific human capital in that the transaction and informational costs involved in opening an account are largely specific to the bank in question. For example, a consumer must invest time to learn about a bank's rate, location and procedures, and must fill out various forms to open an account. Most of this investment, however, becomes worthless if the consumer switches to another bank or investment alternative.

As Becker (1962, 1964) has shown in the human capital context, the cost of specific investments will be shared. If banks paid the full costs (and received the returns) of the specific investment associated with opening accounts (by compensating depositors for the time and money costs of opening accounts), customers could switch accounts with no cost to themselves and therefore would not take into account these bank-specific investment costs. Banks, in turn, would earn no return on their investment in set-up expenses. Conversely, if depositors bore the entire (time and money) cost of setting up an account, banks could lower the rates paid without taking into account the lost investment the depositor would incur in switching accounts. The theory of specific capital investment predicts, therefore, that the costs of specific capital investments will be shared by both parties so that they both at least partially take into account the effects of their behavior on these specific investments. Thus, when an investment has a strong specific capital component, as does opening an MMDA account, the trading parties share the costs of the specific investment. This sharing, in turn, provides both parties with an incentive to continue their relationship to protect their investment.

These shared costs of setting up new accounts mean that both banks and depositors face adjustment costs when shifting deposits into MMDAs. For example, if a bank wishes to attract more deposits (at least partially through deposits into new accounts), it must bear part of the initial set-up costs as well as pay explicit interest. Depositors also bear part of the initial set-up costs. Adjustment costs for depositors lead to differences in the short- and long-run interest elasticities of the supply of deposits to banks, and imply it takes time for actual stocks to adjust to changes in desired deposit stocks. Similarly, adjustment costs for banks imply differences in bank's short-run and long-run interest elasticities of the demand for deposits.

As Flannery (1982) points out, these adjustment costs can lead banks to pay rates of interest on deposits in excess of their marginal revenue products in the short-run when demand temporarily declines, and to pay less than the marginal revenue products when demand temporarily increases, in order to minimize adjustment costs. Over a long-run period, however, deposit costs are equated (on average) to marginal revenue products. Flannery also notes that if non-retail deposits, such as CDs, have no or very small specific capital costs, then banks will use these instruments to meet temporary fluctuations in demand. Thus, it will be occasionally worthwhile for banks to pay higher rates on retail deposits than on wholesale funds such as large denomination CDs to avoid the adjustment costs associated with changing the level of retail deposits.

Adjustment costs are likely to be important for all types of retail deposits regardless of whether they are newly authorized accounts. However, unlike existing accounts in which one can simply deposit a check, virtually all MMDA deposits in the first few months were new deposits with the associated bank-specific set-up costs. Because of this, banks could partly compensate depositors for the costs of opening a new account by paying high interest rates initially. However, for a short-term account like the MMDA, this strategy is not cost-effective once a substantial number of new accounts are opened since high rates would have to be paid to both new and existing accounts. Thus, one strategy for banks would be to pay high rates initially to partly compensate depositors for the initial bank-specific set-up costs, but to compensate depositors in some other way for the costs of opening a new account once the rate of new account formation slowed.

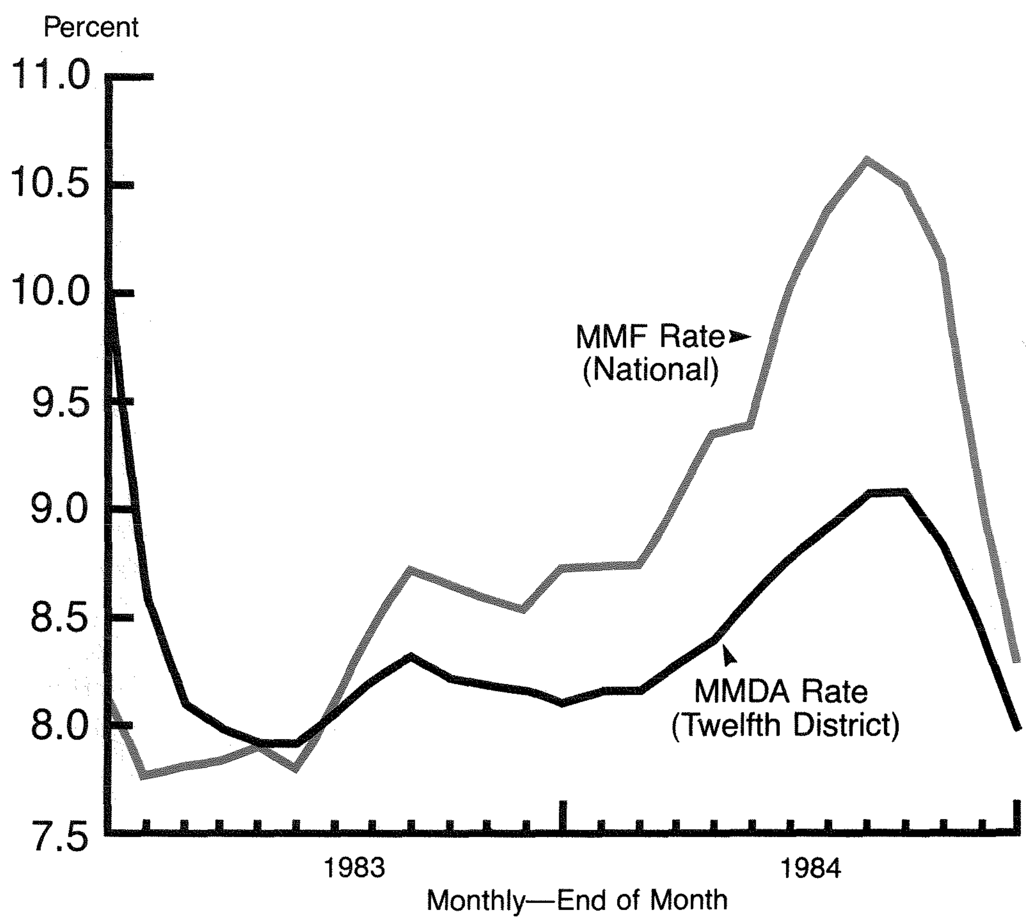
This specific-capital theory of retail deposit flows implies that the cost and quantities of such deposits will respond sluggishly or incompletely to changes in wholesale market interest rates. As Flannery and James (1984) point out, this means that the effective maturity of a bank's retail liabilities need not equal their stated maturity (or time to repricing). Thus, retail deposits will not be supplied perfectly elastically, and the short-run interest elasticity of the supply of deposits to banks will be considerably less than the long-run elasticity.

The Market Acceptance of MMDAs

When a new cost-saving technology such as the MMDA is introduced, the adoption of that technology is not instantaneous because there are costs in learning about the technology as well as costs involved in actually adopting the new technology. The rate at which a new technology is adopted depends on the cost-savings it affords compared to the information and other adjustment costs of adopting it. Although it appeared to take MMDA deposits 3 to 4 months to reach an equilibrium (see Chart 1), approximately 80 percent of banks nationwide offered MMDA accounts starting on December 14, when they were first authorized (and which was only two months after the enabling Garn-St Germain Act was passed). Thus, the costs associated with offering them must have been greatly exceeded by the expected returns. However, the costs associated with opening an MMDA account must have been significant since it took at least three months for deposits to reach an "equilibrium" level.

Banks apparently expected substantial long-run cost savings by attracting funds into MMDAs because many institutions waged aggressive promotional campaigns. Many developed concerted marketing efforts to attract funds via television, radio and print media and direct mail. Institutions also employed bonuses

Chart 2
MMF Rate vs. Bank MMDA Rate



for depositors opening new accounts (an incentive to move funds between institutions) and, perhaps most importantly, many offered premium interest rates considerably above money market fund yields. This pricing behavior is consistent with the specific capital model which predicts that the specific capital investment required to open a new account will be shared by banks through high initial direct interest payments, through cash bonuses for opening accounts, or both.

In Chart 2, we plot the average rate paid on MMDAs (in our sample of western banks) and money funds from the end of 1982 to the end of 1984. This chart shows that the rates paid on MMDAs were considerably higher than the money fund rate in December 1982 and January 1983, but that the rates were close by March of 1983. This is the exact pattern—high initial rates followed by declines—predicted by the specific capital model because high interest rates cannot continue to be used to partly compensate depositors for the cost of opening new accounts once a substantial number of new accounts have been opened. After May 1983, the rate on the money funds exceeded that on MMDAs and was considerably more variable. This pattern also is consistent with the specific capital model which predicts that rates on retail accounts will behave more sluggishly than wholesale rates. Also, since MMDAs offer both federal deposit insurance (not available on MMFs) and more convenience features (for example, access through automated teller machines) than money funds, we would expect MMDA rates generally to be lower than money fund rates except during periods when interest rates temporarily decline.

III. Competition for MMDAs

In this section, we analyze the competition among banks for MMDAs. Two broad issues are considered: the determinants of banks' initial patterns of adopting MMDAs and the interest-sensitivity of MMDA deposits. To analyze differences in banks' patterns of adoption of MMDAs, we employ a logistic model of the percent of each bank's total deposits in MMDAs as a function of time. Differences in the banks' parameters of the logistic are then modeled as functions of various bank characteristics and pricing strategies. To analyze the interest-sensitivity of MMDAs, a standard stock adjustment model of MMDA deposits is employed. Short-run and long-run, own- and cross-interest elasticities of the supply of MMDA deposits to banks are estimated.

The Pattern of Adoption of MMDAs

As theory suggests, different banks would adopt different strategies to attract MMDA deposits depending on the expected benefits and costs. Those banks (and depositors) in the most competitive deposit markets would have the greatest incentives to switch to MMDAs because these banks would have experienced the greatest added costs in attracting deposits due to the inefficiencies of non-price competition. They would have, in effect, been paying high rates of interest (explicit plus implicit) on deposits while depositors were receiving low rates because depositors valued the "free" services and convenience at less than their cost. Similarly, banks that had substituted wholesale deposits for retail deposits might have different incentives to shift to MMDA funding than banks that had substituted nonprice competition for price competition. Also, banks may be in different loan markets and thus have different demands for deposits. Finally, banks may be in different markets for other factor inputs, such as labor and real capital, and thus face different prices for other factors of production. These different input prices also would lead to different demands for MMDAs.

To determine whether banks did differ in their strategies to attract MMDAs and what effects these strategies had on the time-pattern of adoption of MMDAs, we employed an empirical strategy first suggested by Griliches (1957) for analyzing the adoption of new technologies.

If, in fact, the costs and benefits of attracting

MMDAs differ among banks, then we should expect differences among banks in the time-pattern of adoption of MMDAs. For banks in the aggregate, there was an S-shaped pattern of adoption (See Chart 1). Although we would expect individual banks to have the same general S-shaped pattern of adoption of MMDAs, the parameters of this function are expected to differ among banks because of differing costs and benefits to them of attracting MMDA deposits.

To determine empirically whether the parameters of this general function do differ among banks, we fit separate logistic functions to each bank's percentage of total deposits in MMDAs over time. The logistic is an S-shaped function that captures the evolution over time in the share of total deposits in MMDAs in each bank. If the parameters of the logistic—which determine its origin, rate of growth and equilibrium percentage—differ substantially across banks, they then can be analyzed as functions of the banks' strategies for attracting deposits to determine how various strategies affected banks' time-pattern of adoption of MMDAs.

The logistic[9] is defined as:

$$P(t) = \frac{K}{1 + e^{-(a+bt)}} \qquad (1)$$

where:

P(t) = the percentage of deposits at time t in MMDAs

K = the equilibrium percentage of MMDAs of total deposits (when $t = \infty$)

b = the rate of growth of the percent of deposits in MMDAs

a = a parameter that positions the logistic on the time scale.

The logistic function, equation 1, is estimated for each of the 59 banks in our sample using the first 12 months of data on deposits. The method of estimation is non-linear least squares (with Gauss-Newton iterative optimization) that enables us to estimate the parameters K, a and b simultaneously. Fits are generally excellent with over 99 percent of the variance explained by the model, and asymptotic t-values of all parameters significant at the 1 percent level or better.

Parameter estimates for a, b and K (available on request) show considerable variation across banks in the time-pattern of adoption of MMDAs. In Table 1, summary statistics for these parameters of the logistic are presented. Comparing the minimum to the maximum, the parameter a varies by over 2 months; the rate of growth parameter b also varies by a factor of 3; and the equilibrium percentage of MMDAs, the parameter K, varies from as low as 7 percent to as high as 35 percent. This wide variation indicates that banks did in fact experience different time-patterns in the adoption of MMDAs.

These results also can be used to compute the time required for MMDAs to reach 90 percent of their equilibrium value by applying the following formula.

TABLE 1

Summary Statistics for the Variation in the Logistic Parameters Across Banks

	Parameter	Mean	Minimum	Maximum	Standard Deviation
a	(origin) (in months)	−2.63	−4.13	−1.34	.74
b	(slope) (percent per month)	1.80	.81	3.08	.62
K	(equilibrium proportion)	.21	.07	.35	.064

$$T^{*90\%} = 1/b \{\log [P/(K-P)] - a\} \quad (2)$$

where: $P/K = .90$

This formula indicates that, for the "average bank" in our sample, MMDA deposits reached 90 percent of their equilibrium level in only 2.7 months, a very rapid rate of adjustment.

To test whether these observed differences in banks' time-pattern of adoption of MMDAs can be explained by differences in banks' strategies, three regressions are performed with parameters of the logistic as the dependent variables. A common set of economic variables, including banks' pricing strategies and other variables intended to capture some of the differences in banks' behavior in attracting MMDA deposits, are the independent variables.

In all three regressions, we allow for parameter differences among the states (in the Twelfth District) by including state dummy variables because banks in different states are likely to be in different deposit and loan markets. We also control for total deposit growth during the year prior to the introduction of MMDAs because it seems likely that banks previously experiencing rapid deposit growth also would experience more rapid MMDA growth. Control variables for the absolute size of the bank, in terms of total deposits in November 1982, the number of branches, and a dummy if the bank has 5 or fewer branches, are also included to capture differences in the nonprice component of payment. The key economic independent variables are the average of the rates offered in the last weeks of December 1982 and January 1983—the average of the initial promotional rates—and the average rate paid during the last weeks of the next 10 months. We expect that banks offering high initial rates would have more rapid initial deposit growth but that the equilibrium percentage of deposits would depend more strongly on the average rate paid over time.

In column 1 of Table 2, we find that banks with more rapid previous total deposit growth also had more rapid MMDA growth, and that the size of the bank in terms of total deposits in November 1982 is also positively related to the growth rate of MMDAs. Banks with more branches experienced less rapid growth of MMDAs perhaps because their greater convenience made them less constrained by interest ceilings and thus less willing to promote MMDAs. As expected, high initial rates on MMDAs did lead to more rapid growth of these accounts, but the average rate paid did not affect growth.

In column 2 of Table 2, the equilibrium percentage of MMDAs is analyzed in terms of the same independent variables. Both Idaho and Utah had significantly smaller percentages of deposits in MMDAs than California. We find that the equilibrium percent of MMDA deposits depends on the mean rate paid over the entire period but not on the initial rates paid in December and January. We also find that banks with a strong wholesale presence initially, measured by the percent of large CDs in total deposits, attracted fewer MMDAs. This might have been expected since banks that focused on wholesale markets probably would be less likely to seek retail deposits.

Finally, in column 3 of Table 2 we analyze the parameter of the logistic that shifts the function horizontally. An increase in this parameter shifts the logistic to the left (and a decrease to the right) indicating an earlier start to whatever pattern of adoption was followed. Banks with more branches did adopt MMDAs earlier. Not too surprisingly, neither the initial nor longer term MMDA rate have a significant effect on when banks adopted the new account.

In sum, this analysis indicates that the adjustment of the market to MMDAs, while rapid, was not instantaneous. This result is consistent with the presence of adjustment costs and differences between short- and long-run interest rate elasticities. Banks that offered higher initial rates attracted MMDAs more rapidly, but the ultimate percentage of their deposits consisting of MMDAs depended on the average rate paid over a longer period of time.

Although this cross-sectional analysis of MMDA deposits does suggest that banks adopted different strategies for attracting deposits and experienced different patterns of acceptance of MMDAs, it does not provide esti-

TABLE 2

Effects of Economic Variables on the Parameters of the Logistic (Standard Errors in Parentheses)

	Speed of Adjustment	Equilibrium Percentage	Origin
	b	K	a
Number of observations	59	59	59
Standard error of estimate	.52	.05	.63
R^2	.46	.62	.46
Mean value of dependent variable	1.80	.21	−2.63
Intercept	3.57 (5.08)	−.83* (.44)	−7.69 (6.10)
State dummy variables: (California is the omitted category)			
Arizona	.21 (.35)	.043 (.03)	−.09 (.42)
Hawaii	−.35 (.40)	−.026 (.035)	.37 (.49)
Idaho	.031 (.44)	−.078** (.04)	−.15 (.53)
Oregon	−.032 (.38)	−.012 (.033)	.34 (.46)
Utah	.30 (.36)	−.12*** (.03)	−1.01** (.43)
Washington	−.43 (.30)	−.014 (.025)	.53 (.35)
Nevada	−1.51** (.70)	−.012 (.061)	2.31*** (.84)
Total November '82 Deposits	$10.03E^{-8}$** ($4.5E^{-8}$)	$7.7E^{-10}$ ($39.4E^{-10}$)	$-1.35E^{-7}$** ($.54E^{-7}$)
Large Certificate of Deposits/Total Deposits	.33 (.52)	−.16*** (.045)	.10 (.62)
November '82 Deposits/November '81 Deposits	1.45* (.77)	.025 (.067)	−.93 (.93)
Number of Branches	$-5.0E^{-3}$** ($1.9E^{-3}$)	$.2E^{-5}$ ($16.3E^{-5}$)	$6.7E^{-3}$*** ($2.3E^{-3}$)
= 1 if 5 or fewer branches	−.35 (.22)	−.008 (.02)	.13 (.27)
A bank's mean MMDA interest rate over Feb. '83–Nov. '83	$-7.9E^{-3}$ ($5.8E^{-3}$)	.0012** (.0005)	.011 (.007)
A bank's mean MMDA interest rate over Dec. '82 and Jan. '83	.004* (.002)	$.74E^{-4}$ ($1.6E^{-4}$)	$-3.2E^{-3}$ ($2.2E^{-3}$)

***Significant at the 1% level
**Significant at the 5% level
*Significant at the 10% level

mates of the short- or long-run cross elasticities with respect to rates on competitive assets (such as the rate on money funds), nor does it provide estimates of the short- or long-run own interest elasticities of MMDA deposits after the initial adjustment occurred. To address these questions, we estimate a stock-adjustment model of the supply of MMDAs to banks by pooling data on our cross-section of banks over time.

A Stock-Adjustment Model of MMDA Deposits

Portfolio theory suggests that the desired stock of a particular asset (MMDA deposits held by households and businesses), A_i^*, will be positively related to its own rate of return and negatively related to the rates of return on substitute assets. Thus,

$$A_i^* = f(r_1, r_2, \ldots, r_i, \ldots, r_n, W), \tag{3}$$

where:

A_i^* = desired holdings of asset i

W = wealth

r_i = expected rate of return on other assets, i.

Because adjustment of actual asset stocks to changes in desired asset stocks is costly, only a fraction of the difference between desired and actual asset stocks will be eliminated each period.[10] Thus, the actual asset stock of MMDAs will behave as follows:

$$\Delta A_{it} \equiv A_{it} - A_{it-1} = \lambda\,(A^*_{it} - A_{it-1}) \tag{4}$$

where λ is the fraction of adjustment per unit of time of the gap between the desired and actual value of the stock. Rewriting equation (4) gives:

$$A_{it} = \lambda A^*_{it} + A_{it-1}\,(1 - \lambda) \tag{5}$$

And substituting equation (3) gives:

$$A_{it} = (1 - \lambda)\, A_{it-1} + \lambda f\,(r_1, \ldots, r_n, W). \tag{6}$$

In this model, $\delta f/\delta r_i$ is the long-run effect of a change in r_i on the (desired) asset stock (which equals the actual asset stock in the long-run), whereas $\lambda \delta f/\delta r_i$ is the short-run, one-period effect.

In estimating this stock-adjustment model, we assume that the public's supply function of deposits is stable compared to individual banks' demand for deposits, so that the observed variation in interest rates is exogenous. For variation in rates over time, this would seem to be a good assumption, but as the preceding analysis suggests, there is also substantial exogenous cross-sectional variation in banks' strategies for attracting MMDAs.

Our empirical version of equation 6, which represents the supply of MMDA deposits to banks, is as follows:

$$\begin{aligned} Ln(Q_{it}) = {} & \alpha + \beta_1 Ln(Q_{it-1}) \\ & + \beta_2 Ln(R_{i,\,MMDA_t}) + \beta_3 Ln(R_{MMF_t}) \\ & + \beta_4 CONTROL_i + e_{it} \end{aligned} \tag{7}$$

where:

Q_{it} = The quantity (stock) of MMDA deposits at bank i in month t.

Q_{it-1} = the lagged quantity of MMDA deposits.

$R_{i,\,MMDA_t}$ = the rate bank i pays on MMDA deposits in month t.

R_{MMF_t} = the average rate on money market funds during month t.

CONTROL = a vector of control variables including: the natural log of total deposits in November 1982, the growth of deposits from November 1981 to November 1982, the percentage of CDs in total deposits in November 1982, the natural log of the number of branches, a dummy for 5 or fewer branches, and dummies for each state in the Twelfth Federal Reserve District with California being the excluded category

e_{it} = a random error term

α, β_s = parameters to be estimated.

This model has four important features. First, as equation 7 indicates, the coefficient of lnQ_{it-1} is an estimate of $1-\lambda$. Also, the coefficients of the interest rate variables are short-run elasticities (one-period elasticities), but the long-run elasticities may be found by dividing the coefficients by λ.

Second, in this model, data from individual banks are pooled over time. This means that the error term is likely to contain a bank-specific permanent component reflecting permanent unmeasured characteristics of the bank. As Balestra and Nerlove (1966) discuss, such a permanent component can cause bias in models like this one with lagged dependent variables because the lagged dependent variable captures the permanent component. To address this problem, we include a variety of control variables including the size of the bank, both in terms of deposits and branches in the month prior to the introduction of MMDAs, the growth of the bank during the year preceding the introduction of MMDAs, the percent of CDs in total deposits in the month prior to the MMDA introduction (to proxy for the bank's retail presence), and a set of state dummies to capture remaining differences in the state market and regulatory environment not captured by the other variables. We hope that by including these control variables most of any permanent component of the error term will be eliminated. In addition, the effects of the control variables are of interest in their own right.

Third, we apply this general model to three different periods: the entire period, the initial adjustment period following the introduction of MMDAs, and the post-adjustment period. These distinctions were made because the logistic analysis suggests that the flow of deposits into MMDAs was much different during the first three months after their introduction than once they had reached an "equilibrium" level. Thus, it is likely that both the own-and cross-interest elasticity of supply of deposits, as well as the speed of adjustment, would differ during the adjustment and post-adjustment periods.

Finally, the functional form of the model is log-log. This functional form has several advantages when bank-level data are used. With this form the coefficients of the interest rate variables can be directly interpreted as elasticities. More importantly, since the banks in our sample vary widely in size (by several orders of magnitude), the constant-elasticity functional form is superior on a priori grounds to the linear form. This is because it is highly unlikely that a 1 percentage point increase in the MMDA rate would have the same absolute effect on MMDA deposits in a $20 billion dollar bank as in a $20 million bank. By using the constant elasticity, log-log functional form, in which all analysis is in percentage terms, this problem is avoided.

In Table 3, coefficient estimates of the model described by equation 7 are presented. In the first column, results are presented for the entire period. The fit is very good and most coefficients are significant at the 1 percent level or better. Generally, there are significant differences between several states and California in the intercept of this model. These differences suggest that MMDAs during the initial adjustment period were more popular in California than in Idaho, Nevada, Oregon, and Utah (holding constant all other variables in the model). This may be because the California banking market is more competitive.

The results for the entire period (Column 1) also suggest that the initial size of the bank (before MMDAs were offered) was an important determinant of MMDAs, with an elasticity of November 1982 total deposits of .30. That is, banks with 1 percent greater total deposits attracted .30 percent more MMDAs, all other things equal. However, as the results in column 3 suggest, this effect of initial deposit size was much smaller (about one-tenth as large) during the post-adjustment period.

In December, banks with larger branch networks were more successful in attracting MMDAs. However, there was not a significant relationship between the number of branches and MMDAs in either the post-adjustment or adjustment periods.

Banks experiencing rapid growth in deposits prior to the introduction of MMDAs also appear to have attracted more MMDAs (although the result is not consistent across all time frames). This result is not surprising since banks situated in rapidly growing markets might also experience more rapid MMDA growth.

TABLE 3

Analysis of the Quantity of MMDA Deposits at Banks Over Time (Standard Errors in Parentheses)

	January '83 December '84	January '83 February '83	March '83 December '84	December '82
Number of observations	1416	118	1288	59
Standard error of estimate	.12	.15	.043	.48
R^2	.996	.995	.9995	.96
Mean of dependent variable	11.83	11.56	11.86	10.48
Anti-log of mean ($1000s)	$137,749	$104,820	$141,492	$35,596
Intercept	−4.81***	396.06***	.08	−22.19**
	(.41)	(58.38)	(.19)	(9.76)
State dummies (California excluded)				
Arizona	.03***	.0038	.006	.58*
	(.01)	(.06)	(.005)	(.29)
Hawaii	−.02	−.15**	.001	−.085
	(.02)	(.06)	(.006)	(.30)
Idaho	−.17***	−.22***	−.006	−.041
	(.01)	(.07)	(.006)	(.41)
Nevada	−.10***	−.37***	−.002	−.72
	(.03)	(.12)	(.01)	(.56)
Oregon	−.05***	−.17***	.002	.62*
	(.01)	(.06)	(.005)	(.34)
Utah	−.24***	−.21***	−.01	−.97***
	(.01)	(.07)	(.006)	(.30)
Washington	−.07***	−.14***	−.0005	.11
	(.01)	(.05)	(.004)	(.26)
Ln (Nov. '82 Deposits)	.30***	.38***	.03***	.56***
	(.01)	(.04)	(.005)	(.17)
Nov. '82 Deposits/Nov. '81 Deposits	.15***	.033	.02*	1.087
	(.03)	(.16)	(.01)	(.71)
Large CDs/Total Deposits, Nov. '82	−.28***	.46***	−.05***	.71
	(.03)	(.13)	(.01)	(.59)
Ln (Number of Branches)	.04***	−.052	.0001	.52***
	(.01)	(.043)	(.004)	(.18)
=1 if 5 or fewer branches	.002	.08	.006	−.023
	(.01)	(.06)	(.005)	(.26)
Ln (Lagged MMDA quantity at Bank i)	.67***	.65***	.97***	
	(.01)	(.04)	(.005)	
Ln (Rate on MMDAs at bank i at time t)	1.56***	.08	.19***	3.18**
	(.08)	(.28)	(.04)	(1.32)
Ln (Rate on MMFs at time t)	−.86***	−.60***	−.21***	
	(.04)	(.09)	(.02)	

The coefficient of the lagged dependent variable does suggest that adjustment is not instantaneous and hence not costless. However, during the initial adjustment period (Column 2), the estimates suggest that an adjustment of about 35 percent of the difference between actual and desired stocks occurred per month. This result indicates that over 70 percent of the adjustment to the new equilibrium occurred within 3 months. This rate of adjustment is not too different, although it is somewhat slower, than that obtained from the logistic analysis. However, during the post-adjustment period (Column 3), adjustment is much more sluggish, with an implied rate of only 3 percent per month. This result also is consistent with the notion that the costs of setting up new accounts are much different than the costs of adjusting deposit balances in existing accounts. In fact, the bank-specific capital model predicts that once new accounts are established, both banks and consumers will behave in such a way as to make relatively few adjustments in the quantity of funds in the accounts. This would lead to high serial correlation and a slow adjustment.

The own interest elasticity for the entire period (Column 1) suggests that the accounts were sensitive, at least in the long-run, to the rate paid on them. For example, the short-run elasticity was 1.56 and the long-run elasticity 4.73. However, it should be noted that these are firm-level elasticities that are expected to be large in competitive markets.

During the post-adjustment period (Column 3), MMDAs were much less interest-sensitive, with a statistically significant short-run own elasticity of about .19. Also, during this period, the estimated speed of adjustment was dramatically less—only 3 percent per month. This result is not unexpected given the existence of bank-specific capital costs. However, even during this period, the long-run elasticity was approximately 6.33.

To see whether the high initial rates were successful in attracting MMDAs, the model without the lagged dependent variable (which is minus infinity) and the money fund rate (which is constant in any one month) is estimated for December (Column 4). The results suggest a very high initial interest elasticity of over 3. That is, banks with 1 percent higher MMDA rates attracted over 3 percent more MMDA deposits by the end of December.

The results in Table 3 also suggest that money funds provide important competition to MMDAs. We find a statistically significant short-run cross elasticity of MMDAs with respect to rates paid on money funds of about .21 (Column 3), confirming that these two accounts are substitutes. This result is consistent with the sizable initial runoff of money funds into MMDAs.

The money fund rate in this model, however, plays a dual role. Although it is a measure of the return on an alternative substitute asset, it is also a proxy for market rates in general. Thus, we probably have overestimated the cross elasticity of money funds with MMDAs because a higher money fund rate may simply indicate high MMDA rates being paid at other banks. Attempts to measure the effects of both the money fund rate and the average rate of MMDAs proved unsuccessful, probably because of the high correlation and limited independent variation in these two rates.

One of the major uncertainties surrounding the introduction of the MMDA was its interest-rate sensitivity. If MMDAs were very sensitive to interest rates, banks could attract inflows with marginally higher interest rates and MMDAs would be a relatively unstable and costly source of funds whose rate would behave very much like rates on money funds, or other wholesale market return instruments.

If, on the other hand, MMDAs were relatively insensitive to interest rates, then deposits would be less likely to shift from institution to institution without large or permanent rate differences. Thus, institutions potentially would benefit by having a stable source of retail funds whose effective maturity exceeded its stated maturity and whose cost varied much less than wholesale deposits.

Although our results suggest that MMDAs are quite interest-sensitive in the long-run, they also support the notion that MMDAs are not very interest-sensitive in the short-run. Thus, MMDAs appear to behave more like retail deposits with a significant bank-specific capital component than wholesale deposits such as large CDs.

IV. Summary and Conclusions

In this paper, the competition for MMDA deposits—both interbank as well as with the money funds—is analyzed. The analysis focuses on four areas: (1) the sources of the MMDA deposits, (2) the pattern of adoption by banks and the public of these new accounts, (3) the pricing of these accounts and (4) the interest-sensitivity of these accounts in both the short- and long-run and with respect to their own rate as well as the rates on money fund assets. Several findings emerged.

The MMDA deposits came primarily from the money funds, small time deposits, and large CDs. Although MMDAs attracted approximately $90 billion from the money funds, the money funds have continued to prosper in the face of competition from the MMDAs. The MMDA did not, on average, appear to lead banks to increase the overall quantity of their liabilities substantially, but it did enable them to increase substantially their quantity of retail deposits thus reducing their dependence on wholesale deposits (large CDs). To the extent that banks' primary comparative advantage is in providing intermediation services at the retail level, the MMDA has enabled banks to strengthen greatly their competitive position in the retail deposit market. By reducing their reliance on purchased funds, it may actually have improved their ability to borrow in the wholesale markets as well.

This suggests that banks' primary responses to Regulation Q were to substitute wholesale for retail deposits, and nonprice competition for direct price competition in attracting funds to these accounts. Both responses apparently increased banks' deposit costs.

The facts that the money funds lost only a fraction of their deposits to the MMDAs and that their cross elasticity was statistically significant but not too large suggest that the money funds and MMDAs are substitutes, but not as close substitutes as some had anticipated. This is not surprising since money funds had taken numerous actions aimed at reducing potential outflows. With the authorization of the MMDA, many money funds lowered their minimums to well below the statutory MMDA minimum, and increased the services their products provided, for example, by linking accounts to brokerage services and providing easy access to other funds (often called families of funds), and by specializing in short-term investments in tax-exempt securities, riskless securities or high risk/high return securities. We also find in our analysis of aggregate data that because the MMDAs are not substitutes for transaction accounts, there is little reason to expect them to have affected the M1 measure of the money stock.

The adoption of MMDAs was very rapid. Most banks offered such accounts on the day they were authorized and the quantity of funds in these accounts reached over 90 percent of its equilibrium value within 3 months. The rate of adoption by depositors depended on the initial promotional rates offered by banks whereas each bank's equilibrium percentage of deposits in MMDAs depended on the average rate paid over a longer period of time.

Most banks paid very high initial rates on MMDAs, but once the rate at which new accounts formed declined, rates dropped below the level offered by the money funds. This type of pricing behavior is consistent with large bank-specific set-up costs associated with opening new accounts. Theory predicts that such specific capital costs will be shared by banks and their customers and high initial rates are one way of doing this. In addition, the rates paid on MMDAs have been less volatile and generally below wholesale rates after the initial adjustment period—a type of pricing behavior also consistent with the specific capital model.

The speed with which MMDAs were adopted suggests that depositors viewed them as being superior to existing retail accounts, especially small time accounts (from which a significant fraction of the funds came). The fact that banks promoted these accounts so widely and paid such high initial rates suggests that banks had faced substantial costs in their nonprice competition for retail accounts and in their substitution of wholesale for retail accounts to mitigate the economic forces of disintermediation due to Regulation Q.

The MMDAs were fairly interest-sensitive (even in the short-run) during the initial promotional period and this quality made the initial adjustment of actual to desired asset stock levels rapid. However, once MMDAs reached an equilibrium level, further adjustment was much slower because of the existence of significant bank-specific capital costs. These costs meant that, once accounts were opened, deposits would shift only slowly in response to interbank interest differentials. This implies that the effective maturity (or duration) of MMDAs is considerably longer than their stated maturity, or time to repricing.

In sum, MMDAs have been an important innovation in the retail banking market. They have offered retail customers a more valuable package of explicit compensation and implicit services than had existed previously. On the banking side, banks have been able to substitute retail for wholesale deposits and price competition for nonprice competition, thus securing a more stable and lower cost source of deposits.

Data Appendix

Data analyzed in this study were collected by the Federal Reserve Bank of San Francisco's Statistical and Data Services Department for the *Monthly Survey of Selected Deposits and Other Accounts* (FR 2042) and the *Report of Transaction Accounts, Other Deposits, and Vault Cash* (FR 2900) reports. The FR 2042 data are collected from a stratified sample (by size) of sixty-four banks in the eight western states. Total time deposits of these banks account for about eighty-two percent of the total time deposits of all insured banks in the Twelfth District. Both outstanding dollar amounts as of the last Wednesday of the month, and the most common interest rate paid during the week ending on the last Wednesday of the month, are reported for a number of deposit categories, including MMDAs, Super-NOWS and several other time certificate categories. Additional deposit data for these banks were taken from the daily FR 2900 report. In particular, total domestic deposits, and total large denomination time deposits were used as control variables in the study.

Aggregated bank and thrift data for the nation were provided by the Division of Research and Statistics, Board of Governors of the Federal Reserve System. Data were not seasonally adjusted, and most are available from the Board of Governors' H.6 press release entitled *Money Stock Measures*. For our analysis, the large time deposits series used was the gross series, which includes money market fund and thrift holdings of large certificate of deposits.

Additional information on money market fund rates, and bank and thrift interest rates were taken from *Donoghue's Moneyletter* and the *Bank Rate Monitor* respectively.

Bank structure and branch measures are derived from the annual Summary of Deposits Survey taken by the Federal Deposit Insurance Corporation, and published yearly under the title, *Data Book, Operating Banks and Branches.*

APPENDIX
TABLE 1

Features	Account Type		
	Money Market Deposit Account (MMDA)	Super - NOW Account	Money Market Mutual Fund (MMF)
ELIGIBILITY:			
Individuals	Yes	Yes	Yes
Business	Yes	No	Yes
Non-Profit	Yes	Yes	Yes
Government	Yes	Yes	Yes
MINIMUM BALANCE:			
Before Jan. 1, 1985	$2,500	$2,500	$1 and up
Jan. 1, 1985 to Jan. 1, 1986	$1,000	$1,000	$1 and up
INTEREST RATE:	Set by Institution	Set by Institution	Determined by return on fund's portfolio
INSURANCE:	Insured by FDIC or FSLIC	Insured by FDIC or FSLIC	Not Insured*
RESERVE REQUIREMENTS:			
Personal Accounts	None	12%	None
Non-Personal Accounts	3%	12%	None
TRANSACTION FEATURES (Number per month:)			Varies, but most are:
Total Transactions (Including Checks)	Six	Unlimited	Unlimited
Maximum Check Transactions	Three	Unlimited	Unlimited
In Person Transactions	Unlimited	Unlimited	n/a
MINIMUM DEPOSITS:	No statutory minimum	No statutory minimum	Varies, $1 and up
MINIMUM CHECK:	No statutory minimum	No statutory minimum	Varies, $1 and up

*At present only a few money market funds have private insurance coverage.

FOOTNOTES

1. The Depository Institutions Deregulation Committee (DIDC) created the Super-NOW account. This account, which became available on January 5, 1983, was a ceiling-free checking account without limitations on transactions intended to be competitive with the money funds.

2. See Rosen and Katz (1983), Fortune (1975), King (1984), and Garcia and McMahon (1984) for examples of studies of aggregate deposits.

3. The regression model controls for changes in Super-NOWs, total liquid assets in the economy, a time trend, and seasonal factors. The estimation period was from January 1979 to June 1983. The results from this regression are as follows:

A $1.00 Change in MMDAs Has the Following Impact (Standard Errors in Parentheses)

Dependent Variable	*Effect (in dollars)*
Change in Small Denomination Time Deposits	−.52*** (.15)
Change in Large Denomination Time Deposits	−.42*** (.09)
Change in Savings Deposits	−.07 (.15)
Change in Transaction Deposits	+.06 (.07)
Change in Total Deposits Except MMDA and Super-NOW Deposits	−.96*** (.07)
Change in Total Deposits	+.04 (.07)
Change in Money Market Mutual Fund Assets	−.24* (.13)

***Significant at the 1% level
**Significant at the 5% level
*Significant at the 10% level

4. See Furlong (1983).

5. See Keeley (1984) and Keeley and Zimmerman (1984) for a discussion of the effects of ceilings on deposit costs. Also see Benston (1964), Startz (1983) and Rogowski (1984).

6. It should be noted that these marginal and average (interest plus non-interest) costs relate to a given maturity deposit at a given point in time. That is, for a deposit of a given maturity at a particular point in time, the elimination of interest ceilings reduces its marginal and average costs. This concept differs from that of differences in marginal and average costs for long-term deposits at different points in time due to the possibility of attracting new long term deposits at different rates than are being paid on existing long-term deposits that had been acquired earlier.

7. This analysis along with the smaller market for unlimited transaction accounts may explain why Super-NOWs were much less popular than MMDAs. Since Super-NOWs were close substitutes for existing checking and NOW accounts in terms of the services they provided and in terms of reserve requirements, one would expect Super-NOWs to grow rapidly only if the interest ceilings on checking and NOW accounts were binding. However, even with interest ceilings, most banks imposed fees, at least for small depositors, on such accounts. Thus, for such small depositors, the ceilings were not binding. Only for large depositors, for which the ceilings likely were binding, would there be any gain for the banks and depositors in shifting to Super-NOW accounts.

8. The concept of a quasi-fixed factor of production is due to Oi (1962).

9. The logistic function is asymptotic to O and K and symmetric around the inflection point. Its first derivative with respect to time is given by:

$$\frac{dP}{dt}_{t=t^*} = b\,\frac{K\,[K-P(t^*)]}{P(t^*)}$$

That is, the rate of growth of the logistic is inversely proportional to the growth already achieved and directly proportional to the distance from the ceiling.

In other words, $\frac{d \log [P/(K-P)]}{dt} = b$.

10. Griliches (1967) has shown that if the costs of adjustment are a quadratic function of the amount of adjustment, and if the costs of being out of equilibrium are also a quadratic function of the amount one is out of equilibrium, only a fraction of the difference between the desired and actual stock will be eliminated each period.

REFERENCES

Baltensperger, Ernst, "Alternative Approach to the Theory of the Banking Firm," *Journal of Monetary Economics,* Vol. 6, 1980, pp. 1–37.

Becker, Gary S., *Human Capital: A Theoretical and Empirical Analysis with Special Reference to Education,* New York: National Bureau of Economic Research, 1964.

Becker, Gary S., "Investments in Human Capital: A Theoretical Analysis," *Journal of Political Economy,* Vol. 70, Oct. 1962, pp. 4–49.

Benston, George E., "Interest Payments on Demand Deposits and Banks Investment Behavior," *Journal of Political Economy,* October 1964, pp. 431–449.

Brainard, William C. and James Tobin, "Econometric Models: Their Problems and Usefulness: Pitfalls in Financial Model Building," *American Economic Review,* Vol. 58, May 1968, pp. 99–122.

Flannery, Mark J., "Retail Bank Deposits as Quasi-Fixed Factors of Production," *American Economic Review,* Vol. 72, No. 3, June 1982.

Flannery, Mark J. and Christopher M. James, "Market Evidence on the Effective Maturity of Bank Assets and Liabilities," *Journal of Money, Credit and Banking,* Vol. 16, No. 4, November 1984, Part 1, pp. 435–445.

Furlong, Frederick T., "New Deposit Instruments," *Federal Reserve Bulletin,* Vol. 69, No. 5, May 1983, pp. 319–326.

Furlong, Frederick T., "New Limits for the New Year," *Weekly Letter,* Federal Reserve Bank of San Francisco, December 28, 1984.

Fortune, Peter, "The Effectiveness of Recent Policies to Maintain Thrift-Deposit Flows," *Journal of Money, Credit and Banking,* August 1975, No. 7(3), pp. 297–315.

Garcia, Gillian and Annie McMahon, "Regulatory Innovation: The New Bank Accounts," *Economic Perspectives,* The Federal Reserve Bank of Chicago, March/April 1984, pp. 12–23.

Griliches, Zvi, "Distributed Lags: A Survey," *Econometrica,* Vol. 35, pp. 16–49, 1967.

Griliches, Zvi, "Hybrid Corn: An Exploration in the Economics of Technological Change," *Econometrica,* Vol. 25, No. 4, October 1957, pp. 501–522.

Keeley, Michael C. and Gary C. Zimmerman, "Deregulation and Bank Profitability," *Weekly Letter,* Federal Reserve Bank of San Francisco, July 13, 1984.

Keeley, Michael C., "Interest-Rate Deregulation," *Weekly Letter,* Federal Reserve Bank of San Francisco, January 13, 1984.

King, Thomas A. "Thrift Institution Deposits: The Influence of MMCs and MMFs," *Journal of Money, Credit and Banking,* Vol. 16, No. 3, August 1984, pp. 328–334.

Klein, Michael A., "A Theory of the Banking Firm," *Journal of Money, Credit and Banking,* Vol. 3, May 1971, pp. 205–218.

Nerlove, Marc and Pietro Balestra, "Pooling Cross-Section and Time Series Data in the Estimation of a Dynamic Model: The Demand for Natural Gas," *Econometrica,* Vol. 34, No. 3, July 1966, pp. 585–612.

Oi, Walter Y., "Labor as a Quasi-Fixed Factor," *Journal of Political Economy,* Vol. 70, December 1962, 538–555.

Rogowski, Robert J., "Pricing the Money Market Deposit and Super-NOW Accounts in 1983," *Journal of Bank Research,* Summer 1984, pp. 72–81.

Rosen, Kenneth T. and Larry Katz, "Money Market Mutual Funds: An Experiment in Ad Hoc Deregulation: A Note," *The Journal of Finance,* Vol. XXXVIII, No. 5, June 1983.

Startz, Richard, "Competition and Interest Rate Ceilings in Commercial Banking," *Quarterly Journal of Economics,* May 1983, pp. 255–265.

Zimmerman, Gary C., "As the Dust Settles," *Weekly Letter,* Federal Reserve Bank of San Francisco, October 1, 1983.

The Information Content of Credit Aggregates

Bharat Trehan*

Knowledge of changes in private credit aggregates is useful in interpreting the money/GNP relationship because it helps to distinguish shifts in asset demands by households from changes in the demand for transactions balances by firms. This is necessary because both these changes have the same impact on money and interest rates but have different implications for future GNP.

There exists a large body of work documenting movements in narrowly defined money (M1) as leading movements in economic activity. The monetarist tradition regards this as evidence of causation, running from money to GNP. However, even casual empiricism suggests that this relationship has not been very stable recently. The first example is the sharp decline in the velocity of M1 over the second half of 1982 and the beginning of 1983 (see Judd, 1983, for a discussion). The second example is the sharp slowdown in the growth rate of money during the second half of 1983 (when money slowed from a 12.4 percent annual rate in the first half to a 7.2 percent rate in the second), which was followed by an unusually high rate of GNP growth in the first half of 1984. (For a more formal analysis of recent shifts in the money-GNP relation see Simpson, 1984.)

These episodes underline the need for obtaining information beyond that contained in the monetary aggregate when predicting future output. Towards that end, this paper examines what information can be obtained from movements in credit aggregates. A simple model is sketched out in which the money-output relationship over the business cycle is motivated in a way that is the opposite of the usual monetarist story. Changes in money growth precede changes in output as firms increase their demand for transactions balances in order to finance plans to increase future output.

Within this framework, we first examine the relationship between money and credit. Changes in both credit and money precede changes in output, and we show that changes in credit provide information in addition to that provided by both monetary aggregates and interest rates. Our findings indicate, in fact, that without knowledge about what has happened to private credit, it is difficult to determine what a change in money growth means for the future course of economic activity. In contrast, the connection between government borrowing and future economic activity is not as clear-cut, and the empirical results indicate that government borrowing does not provide reliable information about future economic activity.

The key point of this paper is that information on credit can help distinguish between disturbances to money demand—money demand "instability", in other words—and disturbances to credit demand, which also affect the stock of money because the demand for credit is in fact a demand for payments media. Positive disturbances to either will lead to increases in the quantity of money and to a rise in interest rates. However, the future course of economic activity depends upon precisely where the disturbance originates. Information on credit aggregates is useful because it provides a means for pinpointing the source of the disturbance. Em-

*Associate Economist, Federal Reserve Bank of San Francisco. David Taylor rendered valuable research assistance.

pirical analysis supports this hypothesis. Section IV presents equations for real and nominal GNP as well as equations for M1 velocity, and shows that changes in several types of private credit are significant in explaining changes in those variables.

I. Households' Demand for Money

Since households operate under a wealth constraint, any change in money holdings not accompanied by a change in wealth must be matched by an opposite change in the holdings of other assets. In the simplified model considered below, the only other asset available to households consists of loans to firms. Thus, an increase in the demand for money must be offset by a decrease in the supply of loans. More generally, the point is that changes in households' asset demand for money will affect credit market conditions. Below, we show how this leads to a role for credit aggregates in predicting future activity.

Before doing that, however, it is useful to examine what factors can cause changes in households' demand for money and to discuss how relevant these factors are likely to be. Intuitively, it appears that expectations about future conditions are important determinants of the households' demand for money. For instance, Friedman and Schwartz state that expectations of instability due to the outbreak of war cause money demand to go up.[1]

Perhaps a more obvious example of a period during which the household demand for money will increase significantly is a recession. When a recession occurs, or is perceived as likely to occur, individuals tend to become more cautious and to retain money balances since they think that there is a greater chance of being unemployed. Furthermore, the more severe or prolonged the recession, the greater the shift in households' expectations of future income. As a consequence, the increase in money demand will be higher as well.

Some evidence consistent with this hypothesis is provided by the behavior of velocity during recessions. (Recall that velocity is defined as the ratio of real GNP to real money balances, so that an increase in money demand due to expectational factors leads to a decline in velocity.) The first example is the behavior of velocity during the period from late 1982 through early 1983. The fact that the 1982 recession was the worst since the Great Depression and that it followed very closely on the heels of the recession in 1980 must have made a substantial psychological impact on households, leading to an increase in money demand.[2] Exactly the same thing happened during the Great Depression: M1 velocity declined practically continually from the first quarter of 1929 to the first quarter of 1933, with the sharpest declines occurring in three of the last four quarters of this period. While both the examples above are rather extreme, they provide some support for the hypothesis that expectational factors are important determinants of money demand.

Previous researchers have, of course, considered the role of movements in various credit aggregates in forecasting economic activity. Perhaps the most well-known is the work done by Benjamin Friedman (see Friedman, 1985, and the references there). In contrast to the approach below, his work focuses on the determinants of asset demands to show why credit aggregates matter. It relies heavily upon the observed stability of the debt-income ratio in the post-war period (see Friedman, 1981). Friedman also showed that movements in domestic nonfinancial debt contained information at least as useful as any of the monetary aggregates about movements in GNP. However, subsequent empirical research has shown that at least some of his results hinge upon econometric technicalities (see Porter and Offenbacher, 1983, and Froewiss and Judd 1979). Moreover, the ratio of nonfinancial debt to income has been rising since 1980.

II. A Simplified Model

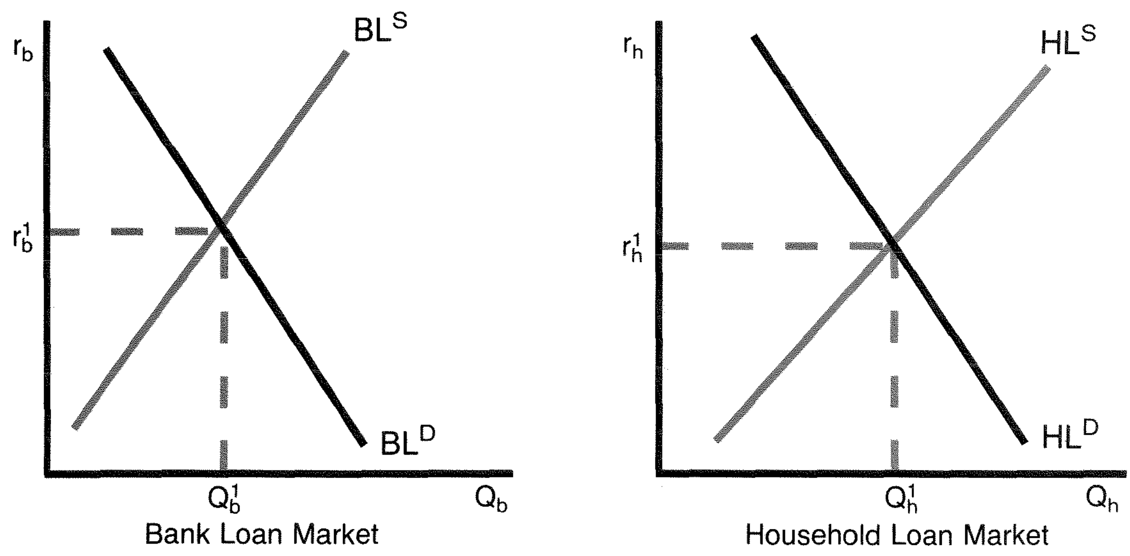
The importance of information about credit aggregates can be shown in a simple framework in which there are only three types of decision-makers: firms, households and banks.

In this framework, we assume that firms desire to increase output levels as a result of positive shocks to productivity (King and Plosser, 1984). Positive shocks to productivity could occur, for instance, when new technology makes it profitable for firms to increase production.

Because production planning and implementation takes time, firms wishing to produce more tomorrow must begin accumulating the needed productive resources today. Not all the funds needed by a firm are likely to be available internally, so it must borrow. It is worth pointing out that this accumulation of money balances for future use is what Keynes called the finance motive for holding money, describing it as the "coping stone" of the liquidity theory of money demand.

Businesses have two sources from which to borrow: banks and households. It is assumed, for simplicity, that households do not borrow from banks, although they add to or reduce their holdings of bank balances by withdrawing or depositing currency. (It is also assumed that the supply of currency is perfectly elastic, that is, the monetary authority supplies the amount of currency demanded.) Thus, firms are the only borrowers. Which source firms draw on for their funds, however, is critical in determining how money and credit behave and, in particular, in determining which will be a better indicator of future economic activity. In the event firms borrow from banks, new transactions deposits are created that add to the stock of money outstanding. In contrast, business borrowing from households simply transfers transactions deposits from households to firms. In the former case, both money and credit are affected; in the latter, only credit.

Figure 1
Credit Market Equilibrium



We turn, now, to a diagrammatic exposition of the analytic framework in this paper. Three markets are of interest in the model: the market for bank loans; the market for household loans, that is, lending by households to firms; and the market for bank deposits. In Figure 1, only the markets for bank loans and household loans are shown. The deposit market is redundant in the sense that developments in the deposit market can be incorporated in what happens in either the household loan or the bank loan market.

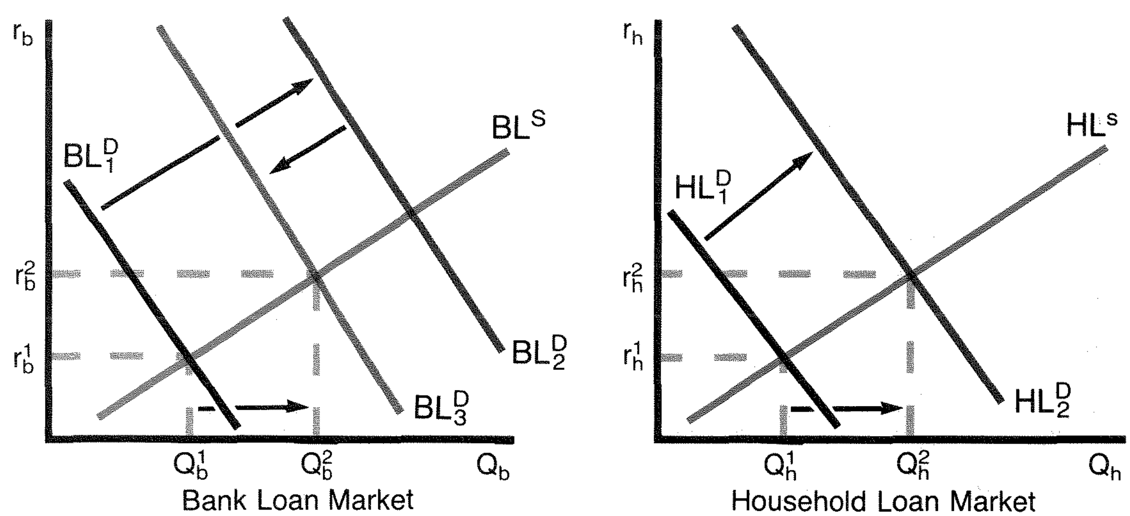
In the market for bank loans, Figure 1a, firms' demand for bank loans is represented by BL^D. The quantity of bank loans demanded increases with a lower bank loan rate, r_b. Also, BL^D is implicitly a function of the rate households charge for loans to businesses, with a higher household loan rate, r_h, increasing firms' demand for banks' loans.

The supply of bank loans varies directly with the rate of interest on loans and inversely with the rate that banks must pay for their deposits, r_d. An increase in r_b induces banks to lend more, creating new deposits for funding. The reserves necessary to support the new deposits come from several sources: from an inflow of currency from the public as banks raise the rate they are willing to pay on deposits; from banks reducing their holdings of excess reserves; and from partial accommodation of the increase in bank credit by the monetary authority.[3] Notice that the bank loan supply curve is based on maximizing behavior by the banks and, in the absence of rigidities or imperfections, this implies equilibrium in the bank reserves market.

Figure 1b shows the market for household lending to firms. Business demand for loans from households, HL^D, is negatively related to the rate charged on these loans, r_h, and (implicitly) positively related to the rate firms must pay for bank loans, r_b. The supply of household loans, HL^S, responds positively to r_h. It is negatively related to the rate banks pay on deposits, r_d, with households offering a smaller supply of loans to businesses when banks pay a higher return on deposits.

We are now in a position to examine why changes in the quantity of credit provide useful

Figure 2
An Increase in the Demand for Credit



information about the future course of the economy. To see this, two different situations are contrasted below.

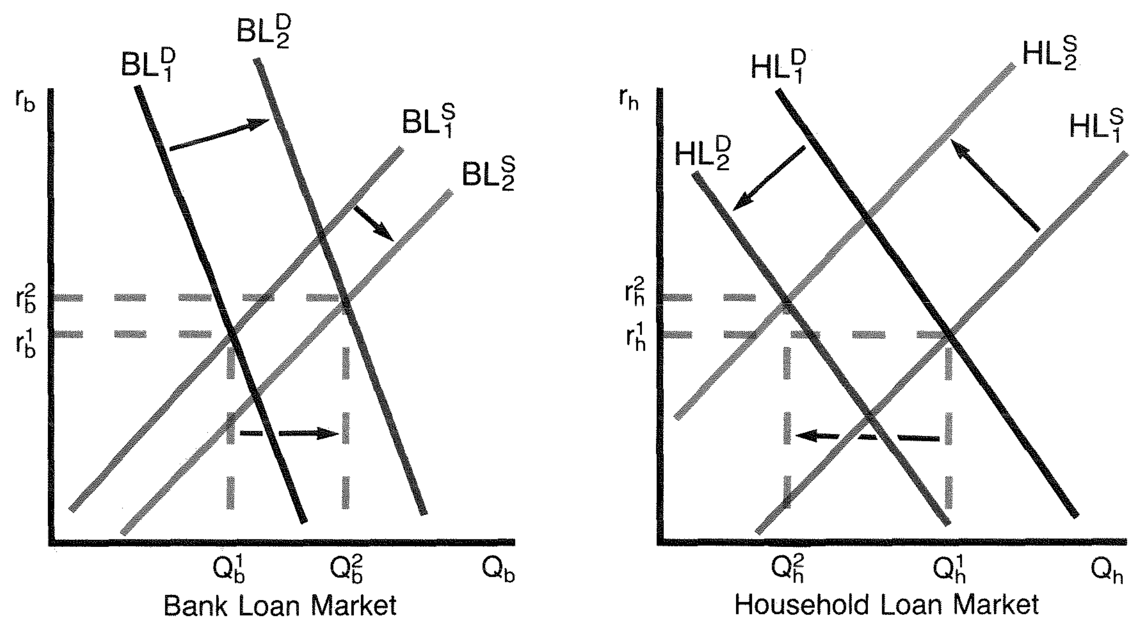
Consider, first, what happens when firms decide to supply greater output in the next time period. If the increased demand for credit is manifested first in the market for bank loans, the demand curve BL^D shifts from BL^D_1 to BL^D_2. The resulting increase in r_b forces some firms into the nonbank market, that is, HL^D shifts out. Consequently, r_h increases. Arbitrage between the two loan markets will continue until interest rates are brought back into equality.[4] In equilibrium, the quantity of both bank and nonbank loans has increased and so has the rate of interest. Since loans are positively related to deposits, both deposits and money supply are higher as well. In the context of the present model, both the interest rate and the money supply provide evidence of increased demand for credit and predict increased output in the next period.

Now, consider an alternative scenario. Assume that household demand for money increases, so that the supply of credit by households declines. In Figure 3 below, two things happen. First, the HL^S curve shifts from HL^S_1 to HL^S_2.

Second, since households do not lend this money to firms, but hold it as deposits, banks can use the money to make new loans. Thus, the bank loan supply curve shifts from BL^S_1 to BL^S_2. Note that the increase in loan supply by banks will be smaller (in absolute terms) than the decrease in loan supply by households for two reasons. First, households will not increase deposit holdings by the exact amount of the increase in money demand (or decrease in loan supply) because part of their money holdings will be held as currency. Second, banks will not be able to lend out the entire amount of the increase in deposits because of the existence of reserve requirements.

Figure 3

An Increase in the Household's Demand for Money



Now, as a result of these shifts, r_h exceeds r_b. Therefore, firms will begin to move from the household loan market to the bank loan market. As a result HL^D moves in and BL^D moves out. Thus, the two rates move towards each other and equilibrium is restored when the two are equal. At the new equilibrium, interest rates will be higher than before. This follows because as a result of a decrease in the willingness to lend, total credit supply has declined with no shift in the demand for credit. The new equilibrium interest rates are r_b^2 and r_h^2. Since bank loans have increased, the quantity of money will be higher as well.

Thus, as far as the impact on money and interest is concerned, this scenario is no different from the first one. But the implications for future GNP are entirely different. In the first case, higher interest rates and money were indicators of a future rise in GNP. In the second, there is no such implication. Indeed, to the extent that higher interest rates discourage economic activity, future GNP will be *lower* in this case.

The only way to discriminate between the two cases is to look at the credit aggregates. Non-bank credit (household lending) increases in the first, but declines in the second. This difference in performance provides a way of discriminating accurately between the two cases. Movements in total private credit—bank plus non-bank lending—similarly allow one to discriminate between the two.[5]

It is tempting to conclude that changes in government borrowing will play the same sort of role as private borrowing did in the model above. However, for this to be true, changes in government borrowing must be causally related to changes in future economic activity. An important reason that this may not be true has to do with the procyclical nature of the budget deficit. During recessions, for example, tax revenues decrease while outlays increase because of higher cyclically sensitive government expenditures such as unemployment benefits. Thus, government borrowing goes up during periods when income is low. However, this borrowing is intended to cushion household income from cyclical vagaries and does not directly influence future output. This source of borrowing will, therefore, offset any positive correlation between future output and federal borrowing due to the other federal expenditures. Thus, it is likely that changes in federal borrowing will not provide useful information about changes in future output.

III. Empirical Tests

The discussion above is based primarily on two hypotheses. First, firms borrow money to increase production over the course of the business cycle. Thus, changes in credit lead changes in economic activity. Further, since firms satisfy part of their needs by borrowing from banks, and since loans by banks and demand deposits are positively correlated, changes in money lead changes in economic activity as well. Second, the household demand for money function is subject to shifts due to changes in expectations. These shifts in the money demand function will be reflected in the demands for other financial assets, such as the financial liabilities of firms.

Thus, movements in private credit should provide significant information about future activity and also be important in explaining money growth. This section presents empirical tests for both these propositions.

First, several alternative "forecasting" equations for real output have been estimated with the intent of testing whether changes in private credit predict changes in output. These equations are similar to those presented in earlier work, for instance, Friedman (1983). The basic equation includes real federal high employment expenditures and real money balances as explanatory variables. To this, the real rate of interest and alternative credit measures have been successively added and tests made to determine the significance of the additional variables.

Second, essentially the same equation has been estimated for nominal GNP. The attraction in estimating such an equation is that the results are directly comparable to previous research. For instance, the well-known "St. Louis" equations (from the Federal Reserve Bank of St. Louis) also regress GNP on money and High Employment Federal Expenditures. (Note, however, that the St. Louis equations view money as being exogenous, that is, all changes in money are viewed as being policy-induced.)

Finally, an equation for velocity is estimated to test the money demand implications. The velocity equation is derived from a money demand specification in which the demand for money is expressed as a function of income and interest rates and in which nominal money balances adjust to desired balances with a lag. Different credit variables are then included in this basic equation to determine whether they help predict movements in velocity.

Quarterly data was used over the period 1959:Q1 to 1984:Q2. The beginning date is dictated by the availability of the M1 series. The interest rate used is the three-month Treasury bill rate. All credit variables are expressed as flows. Of the different measures of credit employed below, the widest aggregate is Domestic Nonfinancial Debt (which is the variable used by Benjamin Friedman). This is decomposed into Federal Debt and Private Debt (the latter is not strictly accurate since it includes state and local government borrowing).

Finally, two other measures are also considered—total loans by commercial banks and total loans by financial institutions other than banks. This is in line with the discussion above, which distinguished between bank and non-bank sources of credit. Obviously, these variables are not ideal for the purpose at hand. For instance, because banks may, in the short run, vary managed liabilities such as Certificates of Deposit when the volume of loans changes, the loans-to-money link may not be as tight. Similarly, loans by financial institutions serve only as a proxy for loans by households.

Consider now, the real GNP equations in Table 1. The dependent variable is the rate of growth of real GNP. For all independent variables, except the rate of interest, the GNP deflator has been used to transform nominal values to real values. The real rate of interest has been obtained by subtracting the expected rate of inflation (in terms of the GNP deflator again) from the nominal rate of interest. The expected rate of inflation is itself obtained by estimating a univariate time series equation for inflation. (An alternate method for obtaining the real rate, where actual inflation was used instead of expected inflation, produced results that were essentially the same as those reported below.) All independent variables are included in growth rate terms, except for the interest rate, which is included as a difference.

Current and two lagged values have been included for all explanatory variables (except the lagged dependent variable and the time trend). For four of these variables (money, private credit, the Treasury bill rate and loans by financial institutions), this length was selected by imposing the condition that the F statistic (for the null hypothesis that current and lagged values of the variable being tested are all zero) have a marginal significance level of at most 0.05 and that the standard error of the equation not increase when additional lags were added. The same lag length was chosen for the other credit aggregates to ensure comparability. It should be pointed out that for these latter variables, the results will not change if the lag length is altered. Finally, only one lag for the dependent variable was included since the second lag is insignificant across all specifications.

Table 1 presents summary statistics for these equations. For each independent variable (including lags), I report the marginal significance levels for the F test. The marginal significance level (M.S.L.) can be can be interpreted as the probability that the variable under consideration has no impact on GNP. Conventionally, the variable is regarded as significant if this probability is less than 0.05. Thus, in the first equation the M.S.L. for the F test on money is 0.0001, which implies that the probability that changes in money have no impact on GNP is extremely small. Equation 1 also includes real federal high employment expenditures, which do not seem to affect output significantly. Notice that Durbin's h statistic shows significant evidence of serial correlation. The second equation adds the real rate of interest to Equation 1. The explanatory power of the equation increases, while the serial correlation declines.

In Equations 3 through 7, different credit variables are added to the set of explanatory variables in Equation 2. Equation 3 includes the rate of growth of private credit. Notice that the credit variable is highly significant and that the h statistic (testing for the presence of serial correlation) is close to zero. Also notice that the M.S.L. on lagged real GNP jumps to 0.8. In Equation 4, the rate of growth of federal borrowing is included. This variable is clearly insignificant and $\bar{R}^2$ is actually lower than in Equation 2. In the next equation, total domestic nonfinancial debt is also insignificant, although it is more "significant" than federal debt. From the last two equations, it can be seen that loans by financial institutions have somewhat greater explanatory power for real GNP than do loans by commercial banks.

Some simple tests to examine the stability of the coefficients on the credit variables were also carried out. The sample was split at two different places to see whether there was any evidence of a structural break. The first split was at 1971:Q2, where Sims (1980) found evidence of a structural break in a system that included output, money growth, and prices. Second, the data was also split at 1979:Q3 to examine whether the change in operating procedures by the Federal Reserve at that time had any impact. Tests were then carried out to examine whether the coefficients of the credit variables (at a specific lag as well as for all lags taken together) had changed.

TABLE 1

Real GNP Equations—Summary Statistics
Sample 1959(Q4)–1984(Q2)

	Equations						
	1	2	3	4	5	6	7
Credit Variable Included	—	—	PVT	FED	TOT	TLFI	TLCB
Explanatory Variables	Marginal Significance Levels of Explanatory Variables						
M1	.0001	.0001	.001	.0001	.0001	.0001	.0001
HEXP	.13	.12	.03	.13	.10	.02	.051
TBR		.007	.04	.01	.007	.12	.02
CDT			.0002	.76	.18	.003	.04
RGNPL1	.055	.04	.8	.03	.056	.058	.17
TIME	.27	.18	.03	.24	.13	.09	.09
$\bar{R}^2$	.370	.433	.532	.415	.440	.498	.464
MSE(X10^2)	.661	.596	.487	.608	.582	.522	.558
Durbin's h	1.69	1.22	0.34	2.09	.07	.84	.98

Notes: (1) All variables are in real terms and in rates of growth, except that the rate of interest term is included as a difference. HEXP is federal high employment expenditure. CDT represents the credit variable. To see which credit variable is included in a particular equation, look at the top of the relevant column. Thus, in equation 3, PVT (that is, private credit) is included. FED is federal borrowing and TOT is total domestic nonfinancial borrowing. FITL is total loans by financial institutions and CBTL is total loans by commercial banks. The current value and two lags have been included for each of these variables. RGNPL1 is the lagged dependent variable.

(2) The Marginal Significance Level (M.S.L.) for a particular variable can be interpreted as the probability that the variable has no impact on the dependent variable (real output in this case). Conventionally, a variable is considered important in a particular equation if it has a M.S.L. of .05 or less.

For the split at 1971:Q2, it was possible to reject the hypothesis of no shift at the 5 percent level only for the contemporaneous value of domestic nonfinancial debt (TOT) and for the first lag of loans by commercial banks. For the split at 1979:Q3, the federal debt variable shows evidence of a shift at all lags individually and together, a result that is not very surprising given the large Treasury borrowings of recent years.

Summary statistics for the nominal GNP equations are presented in Table 2. The variables are defined in the same way as in Table 1, with the exception that all variables are now expressed in nominal terms. (Neither lagged values of GNP nor a time trend were significant here.)

Once again, the rate of interest is significant in predicting changes in nominal GNP. Adding private credit improves the fit of the equation even futher. Notice that adding federal borrowing reduces $\bar{R}^2$ again and that the total debt variable has a M.S.L. of .25. Both the loan variables, total loans by financial institutions (TLFI) and total loans by commercial banks (TLCB), are significant at 1 percent.

The significant credit variables were also tested to see if the coefficients had shifted over time. For the break at 1971:Q2, it was not possible to reject stability for any of the coefficients. For the break at 1979:Q3, the private debt and commercial bank loan variables show no evidence of a shift, while TLFI does.

Consider, now, Table 3 which presents the results for the velocity equation. As discussed above, the estimated equation is derived from a money demand equation, with the additional constraint that the coefficient on real income in the estimated money demand equation equal 1.[6] The first equation includes only the rate of interest and a lagged money term. Successive equations then add different credit variables. Table 3 also shows that while private credit is significant in predicting velocity, neither federal nor total credit are. Total loans by financial institutions are significant here (as in the GNP equations) but loans by commercial banks are not.

Credit variables in the velocity equation were also tested to see if they showed any signs of a shift. However, for neither the break at 1971:Q2 nor the break at 1979:Q3 can the hypothesis of no shift be rejected.

TABLE 2
Nominal GNP Equations—Summary Statistics
Sample 1959(Q4)–1984(Q2)

	Equations						
	1	2	3	4	5	6	7
Credit Variable Included			NPVT	NFED	NTOT	NTLFI	NTLCB
Explanatory Variables	Marginal Significance Levels of Explanatory Variables						
M1	.0001	.0001	.0002	.0001	.0001	.0001	.0001
NHEXP	.10	.04	.0052	.064	.056	.005	.013
NTBR		.0001	.0008	.0001	.0002	.002	.0001
NCDT			.0005	.80	.25	.010	.013
$\bar{R}^2$	.318	.451	.539	.439	.458	.502	.499
MSE($X10^2$)	.753	.607	.509	.620	.599	.550	.554
D.W.	1.54	1.63	1.86	1.59	1.74	1.72	1.81

Notes: See notes to Table 1 for explanations. The variables are the same as in Table 1 except that they are all measured in nominal terms.

TABLE 3
M1 Velocity
Sample 1959(Q4)–1984(Q2)

	Equations					
	1	2	3	4	5	6
Credit Variable Included		PVT	FED	TOT	TLFI	TLCB
Explanatory Variables						
Constant	.623	.578	.609	.595	.596	.563
NTBR	.267	.206	.267	.255	.216	.249
NTBR1L	.481	.438	.480	.475	.430	.478
NTBR2L	.064	.076	.068	.053	.006	.079
CDT		.015	−.0002	.005	.0045	.0017
CDT1L		.016	−.0001	.004	.0015	.0016
CDT2L		−.003	.00	.005	.0067	−.0002
M1LP	−.126	−.253	−.105	−.165	−.216	−.122
			Marginal Significance Levels			
NTBR	.0001	.0001	.0001	.0001	.0001	.0001
CDT		.01	.65	.72	.02	.10
M1LP	.22	.02	.31	.14	.04	.23
$\bar{R}^2$	.363	.413	.353	.351	.405	.385
MSE(X10^2)	.775	.714	.786	.789	.724	.749
D.W.	1.80	1.95	1.76	1.86	1.92	1.94

Notes: The dependent variable is the growth of velocity. See Tables 1 and 2 for an explanation of variables. M1LP is the lagged money term. NTBR1L is NTBR lagged one quarter, NTBR2L is NTBR lagged 2 quarters.

IV. Conclusions

This paper has presented some theoretical arguments and empirical evidence to show that changes in private credit will provide useful information about changes in future output and M1 velocity. Previous analyses of the money/GNP relationship have often tended to focus on the asset demand for money and, consequently, emphasized the substitutability between money and credit. In contrast, explicit attention was paid earlier in this paper to the need for money to carry out transactions. In this framework, it is easy to see how money and credit can vary in the same direction—because the demand for credit can be viewed as a demand for payments media.

For the policymaker, this means that information about changes in money and interest rates alone is not sufficient for predicting what will happen to output. To determine the implications of a change in money, it is important also to know how the credit aggregates are behaving. The evidence presented above is also specific about what credit aggregates are useful. It indicates that changes in federal government borrowing are not significantly related to GNP, while several measures of private credit are. The most recent example of the phenomenon captured in these tests is what happened in late 1982 when declining output was accompanied by rising money but falling private borrowing.

Although the analysis above has focused on recessions as periods when changes in (private) credit aggregates are likely to provide significant information, the underlying logic can be applied more widely. The argument of this paper has been that changes in credit provide a direct means to determine whether the money demand function has shifted (regardless of the source of the change), and that this knowledge is necessary to interpret the money-output relationship properly.

FOOTNOTES

1. Friedman and Schwartz (1982, p. 39), when specifying the arguments of the money demand function, state "another variable that is likely to be important empirically is the degree of economic stability expected to prevail in the future. Wealthholders are likely to attach considerably more value to liquidity when they expect economic conditions to be unstable than when they expect them to be highly stable . . . For example, the outbreak of war clearly produces expectations of instability, which is one reason war is often accompanied by a notable increase in real balances. . ."

2. Others have also suggested the possibility of a shift in the money demand function during the 1982 recession. For example, Axilrod (1984), when discussing the decline in velocity in 1982, states "During part of the period, economic uncertainties may have heightened precautionary demands for cash." Later, he says that he expects velocity to increase in the near future, which "would be consistent with the view that some of the previous decline was a reflection of precautionary demand for cash balances, balances that can be expected to be unwound as confidence in the economy is restored." Similarly, Simpson (1984, p. 259) says "In late 1981 and early 1982, the demand for NOW accounts, passbook savings, and other very liquid assets in household portfolios strengthened while transactions demands weakened and rates dropped only moderately, perhaps reflecting a desire to be better able to cushion an earnings disruption, which at that time seemed more likely."

3. The shape of the bank loan supply curve depends upon the monetary authority's behavior. To see this, consider the two extremes of behavior by the monetary authority. Assume first that the monetary authority accommodates all increases in credit demand, which would happen, for instance, if it were trying to peg the interest rate. In such a situation, the supply curve of bank loans would be horizontal, because the monetary authority stands ready to supply all the reserves for deposit (and loan) expansion.

The other extreme is where the monetary authority does not accommodate any cyclical increase in credit demand. Such a situation may occur, for instance, if the authority is following a fixed money growth rule. In this case, banks can increase loans only by inducing the public to hold more deposits. The supply curve for loans would then be much steeper and, given a limit to the amount of deposits that individuals wish to hold, would ultimately become vertical.

The assumption in the text is that the authority's behavior lies somewhere in between these two extremes. It can also be shown that changes in the credit aggregate convey significant information even if the monetary authority follows one of the above policies.

4. The analysis suppresses the shift in bank loan supply due to a change in r_h and a similar impact of r_b on household loan supply. These effects arise through the deposit market. For example, if the rate on bank loans goes up, banks are likely to begin offering higher rates on demand deposits. As a result, households will decrease loan supply and increase deposit holdings.

5. It is interesting to examine whether the model sketched above is robust to some generalizations. Consider, first, the assumption that demand deposits are the only liabilities of banks. In a more general setting, one would also have to consider other liabilities such as certificates of deposit (CDs). Does the existence of CDs destroy the positive link between loans and demand deposits? Intuitively, the answer appears to be no. If it is true that banks face rising marginal costs to increasing either demand deposits or CDs, banks will increase both types of liabilities together. In equilibrium, the bank must face the same marginal cost for both liabilities, otherwise it is always possible to decrease costs by substituting the cheaper liability for the more expensive. Thus, it is unlikely that the amount of bank loans will increase significantly without an increase in demand deposits.

Consider next, the implications of allowing households to hold a third asset in addition to loans and money, say equity. In this case, increased demand for liquidity will not be matched exactly by a decrease in the supply of loans to firms. Instead, households will reduce equity holdings as well. Once again, the household is unlikely to obtain the necessary balances by selling equity holdings only. Since the shift in liquidity preference does not alter the relative price of loans to equity, holdings of both will be reduced. Thus, the qualitative result is unchanged—firms must still turn to the banking sector for loans.

6. The demand for nominal money can be written (in log form) as

$$M_t^* = \alpha y_t - \beta R_t + P_t$$

where M_t^* denotes desired nominal money balances,
y_t denotes real income,
R_t denotes the nominal rate of interest, and
P_t denotes the price level

Then, under the assumption that actual money balances do not adjust at once to desired, we have

$$M_t - M_{t-1} = \lambda(M_t^* - M_{t-1}).$$

Substituting this in the equation above gives

$$M_t = \lambda(\alpha y_t - \beta R_t) + (1-\lambda)\, M_{t-1} + \lambda P_t.$$

Next, subtract P_t from both sides to obtain an expression for real balances.

$$M_t - P_t = \lambda(\alpha y_t - \beta R_t) + (1-\lambda)\,(M_{t-1} - P_t).$$

Imposing the condition that the coefficient on real income is 1 and transposing gives an expression for velocity that is the estimated Equation 1 of Table 3.

REFERENCES

Axilrod, S.H. "Issues in Monetary Targeting," in *Proceedings of the Conference on Monetary Targeting and Velocity*, Federal Reserve Bank of San Francisco, 1984.

Friedman, B. "Portfolio Choice and the Debt to Income Relationship," NBER Working Paper No. 1545, January 1985.

———. "The Roles of Money and Credit in Macroeconomic Analysis," in "Macroeconomics, Prices and Quantities", James Tobin, ed. Washington, D.C.; Brookings Institution.

———. "The Relative Stability of Money and Credit 'Velocities' in the United States: Evidence and Some Speculations," NBER Working Paper No. 645, March 1981.

Friedman, M. and A.J. Schwartz. "Monetary Trends in the United States and the United Kingdom," Chicago: National Bureau of Economic Research, 1982.

Froewiss, K.C. and J.P. Judd. "Optimal Control and Money Targets: Should the Fed Look at 'Everything'?," *Economic Review*, Federal Reserve Bank of San Francisco, Fall 1979.

Judd, J.P. "The Recent Decline in Velocity: Instability in Money Demand or Inflation," *Economic Review*, Federal Reserve Bank of San Francisco, Spring 1983.

King, R. and C. Plosser. "Money, Credit and Prices in a Real Business Cycle," *American Economic Review*, Vol. 74, June 1984.

Porter, R.D. and Offenbacher, E.K. "Empirical Comparisons of Credit and Monetary Aggregates Using Vector Autoregressive Methods," *Economic Review*, Federal Reserve Bank of Richmond, November/December 1983.

Simpson, T.D. "Changes in the Financial System: Implications for Monetary Policy," *Brookings Papers on Economic Activity*, 1984:1.

Sims, C.A. "Macroeconomics and Reality," *Econometrica*, Vol. 48, Jan. 1980.

Whither the Unemployment Rate?

Brian Motley*

This article develops a model of short-run changes in the unemployment rate and uses it to make forecasts of the rate in 1985. The model is based on Okun's Law which relates changes in unemployment to the growth rate of aggregate demand. It differs from earlier models, including Okun's own work, because it estimates explicitly the growth rate of demand that is required to offset increases in labor force participation and labor productivity rather than assuming that growth rate to be constant. The unemployment rate changes in response to the differential between the actual growth of GNP and this "required" growth rate.

Between December 1982 and June 1984, the unemployment rate in the U.S. declined from 10.7 percent to 7.2 percent of the civilian labor force. Over this same period, real GNP grew at a rapid 6.8 percent annual rate. Since last June, however, real GNP growth has slowed and no further progress has been made in lowering the unemployment rate. Moreover, most economic forecasters do not expect real growth to pick up in 1985, with most estimates for the year in the 3–4 percent range.

An important issue facing economic policymakers is whether real growth in this range would be sufficient to bring about significant further reductions in the unemployment rate. Many economists argue that it probably would not be, but that any attempt to pursue more rapid real growth would risk jeopardizing the hard-won gains in bringing down inflation in recent years. Others agree that faster real growth is required to reduce unemployment to any significant extent, but argue that the risk of faster inflation is worth running, in view of an unemployment rate that remains high by historical standards. In the twenty-five years before 1975, unemployment exceeded six percent of the civilian labor force in only two years, 1958 and 1961, but in the last ten years it has been *below* six percent only once.

One piece of information that is required to make a judgment on this issue is an estimate of the response of the unemployment rate to changes in the growth rate of real GNP. To this end, this paper develops a model that provides short-term predictions of the unemployment rate given expectations of the growth rate of real GNP. This model extends the work reported in a recent *Economic Review* article[1] that developed long-term projections of the unemployment rate. Like that earlier long-run model, the analysis in this paper is based on the observed relation between changes in the unemployment rate and the rate of growth of real GNP, also known as *Okun's Law.*

To bring down the unemployment rate, the real demand for the economy's output of goods and services must increase. Indeed, a certain minimum rate of economic growth is required simply to prevent the unemployment rate from rising. For example, increases in the total population and in the proportion of the population that wants to work mean that to prevent an increase in unemployment, the demand for output must grow enough to create jobs for these new entrants to the labor force. Similarly, the productivity of labor (that is, output per employed worker) generally rises through time, so

*Senior Economist, Federal Reserve Bank of San Francisco. Research assistance was provided by Kenneth Khang.

that unless the demand for goods and services increases at least as rapidly as output per worker, the demand for labor will decline and unemployment will mount.

In this article, the rate of growth in the demand for real GNP that is needed to offset changes in the labor force and productivity exactly—and thus to hold the unemployment rate constant—will be termed the *required GNP growth rate.*[2] To predict the impact on unemployment of a particular rate of growth of real GNP, an estimate of this required growth rate is needed. This article develops a set of equations that explain changes in labor productivity and in the size of the labor force, and uses these equations to derive estimates of the required GNP growth rate.

Over the business cycle, the actual growth rate of real GNP diverges from the required growth rate and, as a result, unemployment rises and falls. In the recovery phase of the cycle, for example, output increases more rapidly than the required rate and the unemployment rate consequently declines. During the recession phase, the reverse occurs. Okun's Law (See Box 1) summarizes the relationship between changes in the unemployment rate and cyclical variations in the rate of GNP growth relative to the required rate. It provides a "rule of thumb" for estimating how much the unemployment rate will change in response to a given change in real GNP. For example, Okun's own estimate of this rule of thumb was that a *three* percentage point increase in the growth rate of real GNP above the required rate would be associated with a *one* percentage point decline in the unemployment rate.

However, most previous estimates of this relationship, including Okun's own estimates, have assumed that the required rate of GNP growth remained constant over the sample period. If this assumption were not correct, the estimates of the relation between GNP growth and changes in the unemployment rate might be biased. The Okun's Law equation developed in this paper avoids this assumption by using the estimates of the required rate derived from the analysis of the determinants of labor force participation and labor productivity.

The Model

The accounting relation between real GNP and total employment may be represented in the following identity:

$$Y/Pop \equiv (Y/E) \times (E/L) \times (L/Pop) \qquad (1)$$

where

Y = Real GNP
E = Civilian employment
L = Civilian labor force
Pop = Adult population[3]

This identity shows that real output per capita, Y/Pop, may be decomposed into the product of (i) output per employed worker, Y/E,[4] (ii) employment as a proportion of the labor force, E/L, and (iii) the labor force as a proportion of the population, L/Pop. Using lower case letters to represent the ratios in Equation 1, this identity also may be written in terms of growth rates:

$$\text{dln } y \equiv \text{dln } q + \text{dln } e + \text{dln } p \qquad (2)$$

where

y = real GNP per capita, Y/Pop
q = labor productivity, or real GNP per employed worker, Y/E
e = the employment ratio, or the proportion of the labor force that is employed, E/L
p = the participation rate, or the proportion of the population that is in the labor force, L/Pop, and

dln represents the change in the logarithm, and thus the growth rate, of each variable.

Since our principal interest is in the growth of employment, and hence of unemployment, it is useful to rearrange this equation and write it as:

$$\text{dln } e \equiv \text{dln } y - (\text{dln } q + \text{dln } p) \qquad (3)$$

If the growth rates of labor productivity (dln q) and labor force participation (dln p) were to depend only on technological, demographic and other non-economic factors that remained constant over time, the forecasting of the employment ratio would be relatively straightforward. Suppose, for example, that la-

Box 1
Okun's Law

When first introduced in 1962, Okun's Law was designed to estimate *potential output.** Okun's objective was to compute the level of real GNP that would be produced if the economy were operating at full employment, and thus to estimate the "costs"—in terms of lost output—incurred when it was at less than full employment. Okun hypothesized that the "gap" between actual and potential output was proportional to the difference between the actual unemployment rate and the rate that would be observed at "full employment." Thus,

$$u - u^F = b\left[\frac{P - A}{A}\right] \qquad \text{(B1)}$$

where u and u^F represent the actual and "full employment" values of the unemployment rate and P and A represent potential and actual values of real GNP.** Okun assumed that full employment corresponded to a measured unemployment rate of four percent and thus that potential output corresponded to the level of GNP that would be reached at a four percent unemployment rate.

Using a variety of techniques, Okun estimated that the value of b in this expression was close to one-third, giving rise to his rule of thumb that "each percentage point in the unemployment rate above 4 percent has been associated with about a 3 percent decrement in real GNP."

In discussions of short-run policy, this rule frequently is expressed in terms of the growth rate of real GNP rather than of the gap between actual and potential output. In this form, the rule states that a 3 percentage point reduction in the annual *growth rate* of real GNP will be associated with a one percentage point increase (per year) in the unemployment rate. It is frequently more informative to the policymaker to know the likely impact on the unemployment rate of a given *change* in real GNP growth than to know the costs in terms of lost potential output of a given unemployment rate. Moreover, there is less unanimity among economists today on what measured rate of unemployment corresponds to "full employment", and hence what "potential GNP" is, than there was when Okun wrote his article. This makes it more difficult to estimate the Law in the form of Equation B1.

Although, in this paper, the Law has been stated and estimated in its "growth rate" form, it is important to recognize that the "growth rate" and "gap" versions of the Law are simply alternative ways of stating the same basic relationship.*** Clearly, if the potential (or required) growth rate of the economy does not change, a three percentage point increase in the annual growth rate of actual GNP will reduce the "gap" between potential and actual GNP by three percentage points. In either formulation, this will lower the unemployment rate by one percentage point.

---

*Arthur M. Okun, "Potential GNP: its Measurement and Significance," *Proceedings of the Business and Economic Statistics Section of the American Statistical Association,* 1962.

**Since u is approximately equal to $-\ln e$ and $(P - A)/A$ is approximately equal to $-\ln(A/P)$, this hypothesis also may be written as

$$\ln e - \ln e^F = b(\ln A - \ln P).$$

By assuming that potential GNP grows at an annual rate r, this expression may be written as

$$\ln e = (\ln e^F - b \ln P_0) + b \ln A - b\, rT,$$

where T represents time and P_0 is the level of potential GNP when T is zero.

In his empirical estimates, Okun used this version of this relation as well as the one represented by Equation B1 with similar results.

***Note that the formulation used in the body of this paper may be obtained from the expression in the preceding footnote by differentiating it with respect to time.

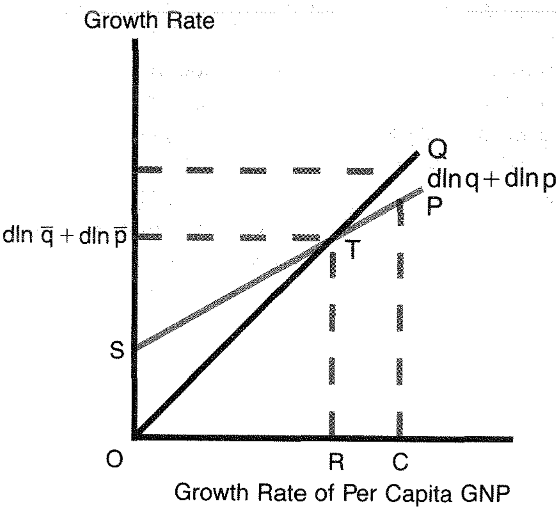
bor productivity were known to rise at a constant *two* percent a year and the available labor force at *one* percent. In this case, the required GNP growth rate would be *three* percent since if real aggregate demand were to increase at that rate, the growth in the demand for labor would exactly match the growth in the supply, and the proportions of the work force that were employed and unemployed would remain constant. In terms of Equation 3, if real GNP were to grow at three percent a year, dln e would be zero because dln y would be exactly equal to the sum of dln q and dln p.

If real GNP were to increase by more than three percent, the proportion of the labor force employed would rise. In particular, Equation 3 shows that, in the special case in which the growth rates of participation (dln p) and of productivity (dln q) are constant, an increase in the annual GNP growth rate of, for example, *one* percentage point (from three percent to four percent), would cause the employment ratio to grow at an annual rate of *one* percent and hence would cause the unemployment rate to decline by *one* percentage point per year.[5]

Thus, if the growth rates of productivity and the labor force were constant, the required GNP growth rate also would be constant and each *one* percentage point increase in the actual GNP growth rate above the required rate would produce a *one* percentage point decline in the unemployment rate.

In fact, the growth rates of productivity and the labor force are *not* constant. The demographic, technological and other non-economic factors that affect labor force participation and productivity growth vary over time, and these variations lead to changes in the required GNP growth rate. In addition, participation and productivity also respond to changes in the growth rate of real GNP over the business cycle. More rapid GNP growth during a business cycle expansion, for example, tends to be associated with faster growth both in output per worker (partly because hours of work increase) and in labor force participation. This means that a given increase in real GNP growth leads to a smaller increase in the employment ratio than

Figure 1
Determining the Required Growth Rate



if the growth rates of productivity and participation were unchanged. In terms of Equation 3, since the rise in dln y associated with a cyclical upswing typically is accompanied by increases in both dln q and dln p, the increase in dln e is correspondingly smaller.

Figures 1 and 2 illustrate these arguments graphically. In these figures, the horizontal axis represents the growth rate of per capita GNP. The 45-degree ray OTQ identifies points at which the growth rates measured on the vertical and horizontal axes are equal. Figure 1 illustrates the determination of the required GNP growth rate and shows how changes in the growth rate of GNP over the business cycle lead to increases and decreases in the employment rate, while Figure 2 illustrates the effect of demographic, technological and other non-cyclical factors on the required GNP growth rate.

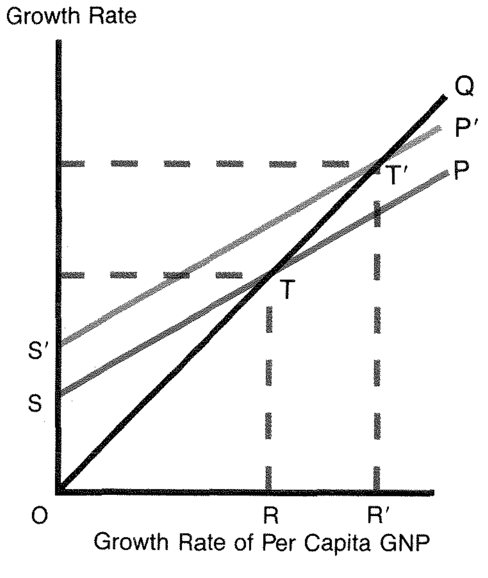
In Figure 1, the curve STP, labeled dln q + dln p, represents the combined growth rate of productivity and participation. This curve slopes upward to illustrate the tendency for the growth of both productivity and participation to increase and decrease as the growth rate of real GNP rises and falls over the business cycle.

For simplicity, STP is represented as a straight line. The accounting identity in Equation 3 implies that the vertical distance of the curve STP above or below the ray OTQ represents the rate of change of the employment ratio. Hence, the intersection of the curve STP with the ray OTQ at the point T identifies the growth rate of per capita real GNP at which the employment ratio remains unchanged. At this intersection, the combined growth rate of productivity and participation is exactly equal to the growth rate of per capita GNP. This growth rate of productivity and participation is labeled $\text{dln}\,\bar{q} + \text{dln}\,\bar{p}$.

The GNP growth rate which holds the employment ratio constant is the "required growth-rate." In Figure 1 this growth rate is OR. If real aggregate demand per capita grows at the required rate, the combined growth rate of productivity and participation, RT, is exactly equal to the growth rate of per capita GNP, OR. Hence, the demand for labor rises at the same rate as the supply, and the employment ratio remains unchanged.

Figure 2

An Increase in the Required Growth Rate



In a cyclical upswing, when the actual growth rate of real GNP rises above the required growth rate, OR, the combined growth rate of productivity and participation also increases *but by a lesser amount*. Hence, the employment ratio increases. For example, if GNP per capita grows at rate OC, productivity and participation together grow at rate CP. Although this growth rate is above the required GNP growth rate, OR, it is less than the actual GNP growth rate, CQ. Hence, the employment ratio increases at rate PQ. Conversely, when real GNP growth is less than OR, the growth rate of productivity and participation is greater than that of GNP so that the employment ratio declines.

When the growth rates of participation or productivity increase or decrease for reasons that are *not* related to the business cycle, the required GNP growth rate will change. An increase in the trend rate of growth of labor force participation, for example, adds to the supply of labor, which means that real aggregate demand must increase more rapidly if there is to be no increase in unemployment. Similarly, faster trend growth in labor productivity reduces the demand for labor; if unemployment is to remain unchanged, this must be offset by faster output growth. Thus, in both of these instances, the required GNP growth rate rises. In Figure 2, such changes are represented by an upward shift of the curve STP to S′T′P′. As a result, the intersection point with OTQ is shifted from T to T′ and the required growth rate increases from OR to OR′. The empirical section below attempts to quantify such shifts and to derive estimates of how the required growth rate has changed over time.

The preceding argument also may be stated in algebraic terms. The hypothesis that a cyclical increase (decrease) in the growth rate of per capita GNP leads to a lesser increase (decrease) in the combined growth rate of productivity and participation may be written as

$$\text{dln}\,q + \text{dln}\,p = \alpha + \beta\,\text{dln}\,y \quad \text{where } 0 < \beta < 1 \qquad (4)$$

This equation represents the curve STP in Figure 1. The intercept term, α, represents the effect of technological, demographic or other noncyclical factors that affect the growth rates of productivity and participation. The slope coefficient, β, represents the response of productivity and participation to variations in the growth rate of per capita GNP over the business cycle. Substituting this equation into Equation 3 and re-arranging terms yields

$$\text{dln } e = -\alpha + (1 - \beta) \text{ dln } y \quad (5)$$

As illustrated in Figure 1, when per capita GNP is growing at the required rate, dln y^R, the employment ratio is constant and the growth rate of per capita GNP is equal to the combined growth rate of productivity and participation. Thus,

$$\text{dln } y^R = \text{dln } \bar{q} + \text{dln } \bar{p} = \alpha + \beta \text{ dln } y^R \quad (6)$$

where dln $\bar{q}$ and dln $\bar{p}$ represent the growth rates of productivity and participation when per capita GNP is growing at the required rate. Equation 6 represents the growth rate of per capita GNP at the intersection point T in Figure 1. As was illustrated in Figure 2, a change in the value of the intercept term, α, which represents the effect of non-cyclical variables on the growth of productivity and participation, alters the required growth rate.

When Equation 6 is solved for α and the resulting expression is substituted into Equation 5 it yields:

$$\text{dln } e = (1-\beta)(\text{dln } y - \text{dln } y^R) \quad (7)$$

This equation, which is a form of Okun's Law, shows that the growth rate of the employment ratio depends on the *differential* between the actual and required growth rates of real GNP. However, most estimates of this equation, including Okun's own, have assumed that the required growth rate was constant and hence that changes in the employment ratio depend only on the *actual* GNP growth rate.

More recent research[6] suggests that this assumption that the required growth rate does not change over time may not be an accurate one and hence that estimates of Equation 7 made under that assumption may be biased. This suggests that an alternative and preferable procedure is to construct a statistical series for the required growth rate, dln y^R, and to use it to estimate Equation 7. This is the procedure employed in the following empirical section.

Empirical Results

To estimate Equation 7, a statistical series for the required GNP growth rate must be constructed. The previous section showed that this growth rate varies in response to changes in the demographic, technological and other non-cyclical factors that influence productivity and labor force participation. This argument suggests that a statistical series for the required growth rate, dln y^R, may be constructed in a series of steps. First, separate equations are estimated to explain labor productivity and labor force participation in terms of both cyclical and non-cyclical variables. Second, these equations are simulated over the sample period holding the cyclical variables constant. The growth rates of these simulated values are interpreted as estimates of dln $\bar{q}$ and dln $\bar{p}$—the growth rates of productivity and participation that would arise if there were no cyclical variations in the economy and hence a situation in which the unemployment rate remained constant. On this interpretation, the sum of these simulated growth rates represents dln y^R, the required GNP growth rate.

Separate equations were estimated for the female and male participation rates and for labor productivity. Earlier research[7] suggested that both participation and productivity may be adequately modeled using a cyclical variable, a few demographic variables and a series of trend variables. In the present context, it was natural to follow this previous research and choose the employment ratio as the cyclical variable since the required growth rate is defined as the rate that holds that ratio constant.[8] Full details of the estimated equations are shown in the table in Box 2. The estimation period was from the first quarter of 1953 to the last quarter of 1982.

Each estimated equation was simulated dynamically over the sample period, holding the employment ratio constant at its 1953(Q1) level[9]. This procedure computes how productivity and participation would have changed

Box 2
Estimating the Required GNP Growth Rate

The required GNP growth rate is the rate at which real aggregate demand must grow in order that the employment ratio remains constant. It is the growth rate which just suffices to provide jobs for new entrants to the labor force and to offset increases in output per worker. The principal difficulty involved in estimating this growth rate is that the rates at which labor force participation and labor productivity increase are themselves influenced by changes in the rate of output growth over the business cycle. This was illustrated in Figure 1.

The first stage in the estimation process is to estimate equations to explain output per employed worker and the female and male participation rates in terms of both cyclical and noncyclical variables. The variables used in these equations were chosen on the basis of an examination of earlier research (See Footnote 7). In this research, the effect of the business cycle generally is represented by some indicator of labor market tightness. In this paper, both the level and the quarterly change in the employment ratio were included as regressors to represent cyclical influences.

Labor productivity and labor force participation exhibit significant time trends. It is widely argued that the trend of participation changed during the sixties and that of productivity slowed in 1969 and perhaps again in 1974. Hence, an initial version of the equations included a series of dummy variables representing linear trends that shifted at various dates. A second version sought to capture these varying trends by a fourth-order polynomial in time. This second version provided a closer fit to the sample data and also yielded coefficients on the demographic variables that were statistically significant and of the correct sign. This is the version reported in the table below.

The equation explaining output per worker includes a special dummy variable to capture a bias in the measurement of labor productivity resulting from the imposition and subsequent removal of price controls in 1971–1974.* Since such controls can never be perfectly effective and probably became less effective as time passed, it is likely that "announced" prices increasingly understated "true" prices during the control period. Since real GNP is measured by deflating nominal GNP by a price index that represented announced prices, this would cause the growth of output per worker to be over-estimated during the control period and correspondingly under-estimated while the controls were being dismantled. The dummy variable, DPC, which seeks to capture this effect takes the value 1 from 1971(Q3) to 1973(Q1), −1 from 1973(Q2) to 1974(Q4), and 0 at all other dates. The above argument implies that the coefficient on this variable is expected to be positive.

Overall male participation in the labor force has been declining in the last thirty years. However, to a significant extent, this decline reflects the falling participation of older males which probably represents an exogenous demographic trend. The decrease in the participation of prime-age males has been much smaller. To capture this demographic change, the equation explaining the male participation rate includes the participation rate of males over age 55, PARTM55, as an independent regressor.

For similar demographic reasons, the equation explaining female participation includes a variable, INFANT, representing the number of young children relative to the female population. This variable began to decline at about the same time as the upward trend of female participation accelerated in the 1960s. Although a more complete analysis of the determinants of female labor force participation would treat decisions on child-rearing and participation as jointly determined, the number of children is effectively pre-determined in a single quarter. Hence, this variable also is regarded as an exogenous demographic factor.

Hence, the equations estimated were as follows:

Output Per Worker

$$\ln q_T = g_0 + g_1T + g_2T^2 + g_3T^3 + g_4T^4 + g_5DPC_T + g_6\ln e_{T-1} + g_7d\ln e_T + g_8\ln q_{T-1}$$

Female Participation

$$\ln p_T^F = f_0 + f_1T + f_2T^2 + f_3T^3 + f_4T^4 + f_5INFANT_T + f_6\ln e_{T-1} + f_7d\ln e_T + f_8\ln p_{T-1}^f$$

Male Participation

$$\ln p_T^M = m_0 + m_1T + m_2T^2 + m_3T^3 + m_4T^4 + m_5PARTM55_T + m_6\ln e_{T-1} + m_7d\ln e_T + m_8\ln p_{T-1}^m$$

The table below reports the results of estimating these equations.

In constructing the series for the required GNP growth rate, each estimated equation was simulated over the sample period holding $\ln e_{T-1}$ equal to its value in the second quarter of 1953 and $d\ln e_T$ equal to zero. The simulation was dynamic in the sense that in each quarter the value of the lagged dependent variable was set equal to its *simulated* value in the preceding quarter. The simulated male and female participation rates were combined to form the total participation rate. The sum of the growth rates of these simulated values of overall participation and of productivity represents the estimated growth rate of per capita GNP that would hold the employment rate constant at its 1953(Q1) level. This growth rate, $d\ln \hat{y}_T^R$, was used in the estimates of Equation 7.

*For a detailed discussion of the impact of price controls on measured productivity growth, see Charles S. Morris, "The Productivity 'Slowdown': A Sectoral Analysis," *Economic Review,* Federal Reserve Bank of Kansas City, April 1984.

Female and Male Participation and Labor Productivity

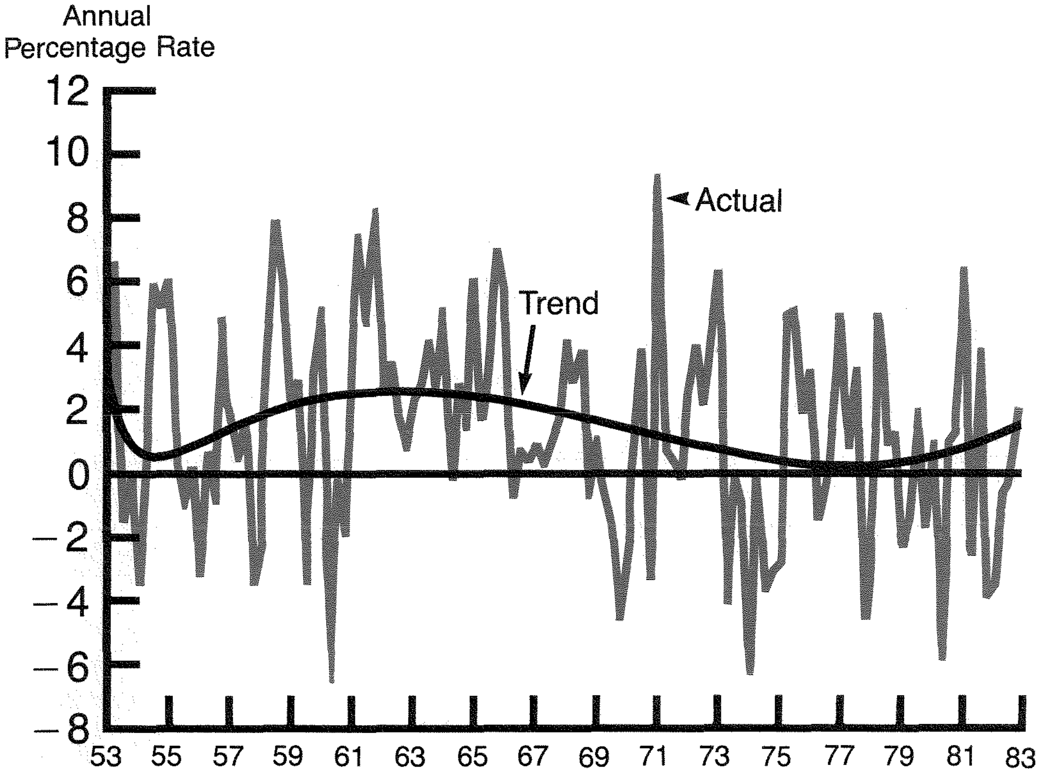
Independent Variables	Female Labor Force Participation Rate	Male Labor Force Participation Rate	Real GNP Per Employed Worker
Constant	−0.636	−0.428	2.573
	(6.26)	(2.434)	(4.348)
$\ln e_{T-1}$	0.236	0.073	−0.019
	(3.955)	(3.906)	(0.243)
$d\ln e_T$	−0.130	−0.060	1.058
	(1.01)	(1.105)	(6.445)
Price Control Dummy (DPC)			0.00328
			(1.880)
INFANT	−0.187		
	(3.695)		
LPARTM55		0.0558	
		(2.134)	
TIME	4.729E−3	3.620E−4	−6.984E−4
	(6.250)	(1.543)	(1.766)
$TIME^2$	−1.107E−4	−3.102E−5	8.048E−5
	(5.834)	(3.676)	(4.508)
$TIME^3$	9.318E−7	4.104E−7	−1.033E−6
	(4.952)	(3.898)	(4.650)
$TIME^4$	−2.311E−9	−1.606E−9	3.944E−9
	(2.829)	(3.921)	(4.582)
Lagged Dependent Variable	0.612	0.539	0.722
	(10.427)	(6.922)	(11.304)
SEE	0.00615	0.00822	0.00627
DW	2.06	1.73	1.91

over the sample period if the employment ratio had remained constant. The simulated values of the male and female participation rates were combined into an overall participation rate. Finally, the growth rates of simulated total participation and labor productivity were summed to yield a series of the required growth rate of per capita real GNP that would hold the employment ratio constant at its 1953(Q1) level.

Charts 1–3 show the actual and simulated values of productivity and of male and female participation. Although most of the variation in all three variables represents the business cycle, it is clear that even when the effects of the cycle are removed, a significant amount of variation remains. Chart 4 shows the actual and required growth rates of per capita real GNP. Toward the end of the sample period, the required per capita GNP growth rate was approximately two percent, but it was significantly lower through most of the 1970s. This constructed series of the required GNP growth rate was used to estimate an empirical version of Equation 7.

Most previous estimates of Okun's Law have found that the employment ratio responds to changes in the GNP growth rate with a lag. The theoretical model represented in Equation 7 implies that this lagged response should refer to the *differential* between the actual and required growth rates, suggesting that the empirical form of Equation 7 should include current and lagged values of both the actual and the required growth rates of per capita GNP. In

Chart 1
Growth Rate of Productivity
Actual vs. Trend



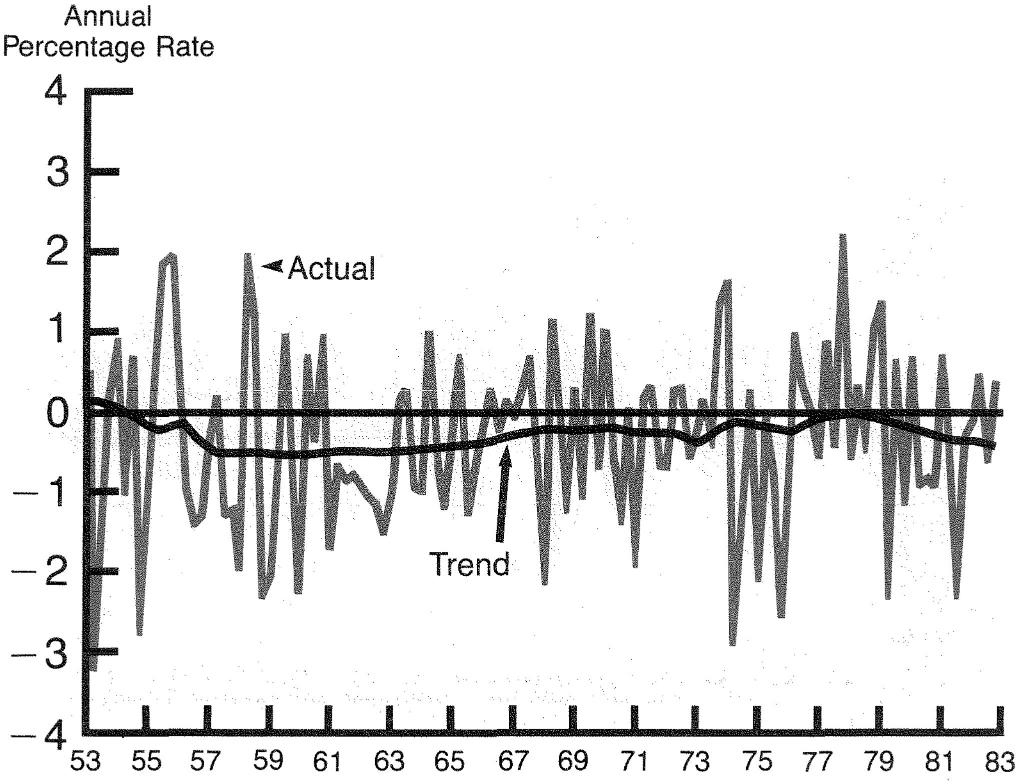
practice, because the required growth rate is a smooth series, its current value is a good proxy for its lagged values.[10] Hence, the estimated equation was:

$$\text{dln } e_t = a + b_0 \text{ dln } y_t + b_1 \text{ dln } y_{t-1} + b_2 \text{ dln } y_{t-2} - c \text{ dln } \hat{y}^R_t \quad (8)$$

where $\text{dln } y^R_t$ represents the constructed series of the required GNP growth rate. The proposition embodied in Equation 7 that growth in the employment ratio is proportional to the differential between the actual and required GNP growth rates implies that the intercept term in Equation 8 should be zero and that the sum of the coefficients on the current and lagged growth rates of per capita GNP should be equal to the coefficient on the required growth rate, that is, $b_0 + b_1 + b_2 = c$.

The results of estimating this equation, and testing these hypotheses, are set out in Table 1. In that Table, Equation A shows the estimated coefficients of Equation 8 with no restrictions. All coefficients carry the signs predicted by the theory. Although the sum of the estimated coefficients on the current and lagged values of the GNP growth rate is not exactly equal to the coefficient on the required rate, the hypothesis that they are equal cannot be rejected at conventional significance levels. In addition, Equation A confirms the prediction that the intercept term should be zero; the estimated intercept is small and not statistically significant. When the intercept is eliminated in Equation B, the sum of the coefficients on $\text{dln } y_t$ is much closer to that on $\text{dln } \hat{y}^R_t$. Finally, constraining these values to be equal, as in

Chart 2
Growth Rate of Male Participation Actual vs. Trend



Equation C, has no noticeable effect on the estimated coefficients.

For comparison, Equation D reports the result of restoring the intercept but excluding the variable dln $\hat{y}^R_t$. This equation corresponds to a specification in which the required growth rate of real GNP is constant and equal to $-a/(b_0 + b_1 + b_2)$. As pointed out above, earlier research suggested that this specification is not supported by post-war U.S. data[11]. This finding is confirmed by its standard error, which is slightly larger than those for the earlier equations.

Equation C is the empirical counterpart of Equation 7 and incorporates the coefficient restrictions suggested by the theory. It implies that, in order to increase the annual growth rate of the employment ratio by one percent (that is, to lower the unemployment rate by one percentage point per year), actual per capita GNP must increase at a rate two percentage points above the required rate. This compares with Okun's estimate of three percentage points. Several other studies made since Okun's initial work, which used data from the 1950s, also have suggested that the unemployment rate has become more responsive to changes in the GNP growth rate.[12]

Chart 5 shows the quarter-to-quarter changes in the unemployment rate and compares them to those derived from the fitted values of Equation C in Table 1[13]. Given the substantial volatility of the unemployment rate, the fit of the equation appears to be quite good.

Chart 3
Growth Rate of Female Participation Actual vs. Trend

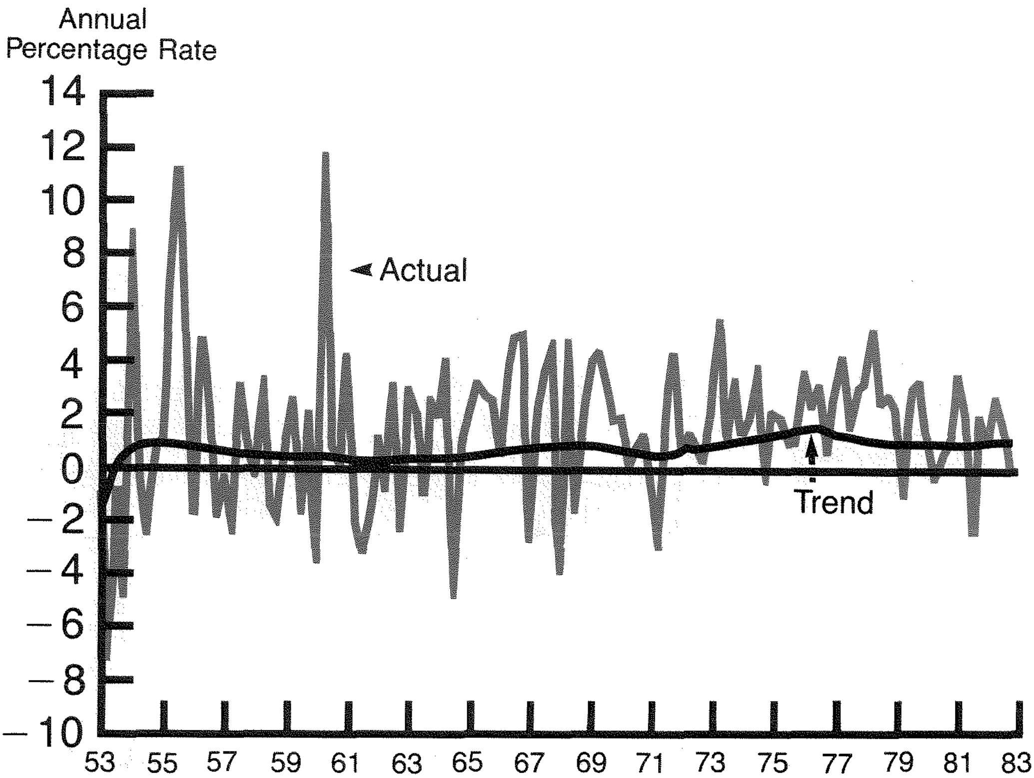


TABLE 1
Estimates of Okun's Law

	Equations A	B	C	D
Constant	−0.0010			−0.0023
	(1.49)			(10.25)
$d\ln y_t$	0.253	0.253	0.254	0.250
	(11.94)	(11.90)	(12.01)	(11.69)
$d\ln y_{t-1}$	0.179	0.181	0.181	0.176
	(7.98)	(8.01)	(8.04)	(7.76)
$d\ln y_{t-2}$	0.066	0.067	0.067	0.062
	(3.13)	(3.16)	(3.19)	(2.91)
Sum	0.498	0.501	0.503	0.487
	(18.71)	(18.78)	(18.91)	(18.44)
$d\ln\hat{y}_t^R$	−0.306	−0.533	−0.503	
	(1.92)	(10.38)	(18.91)	
SEE	0.00225	0.00226	0.00225	0.00227
DW	1.65	1.61	1.62	1.60

TABLE 2
Unemployment Forecasts 1983–84

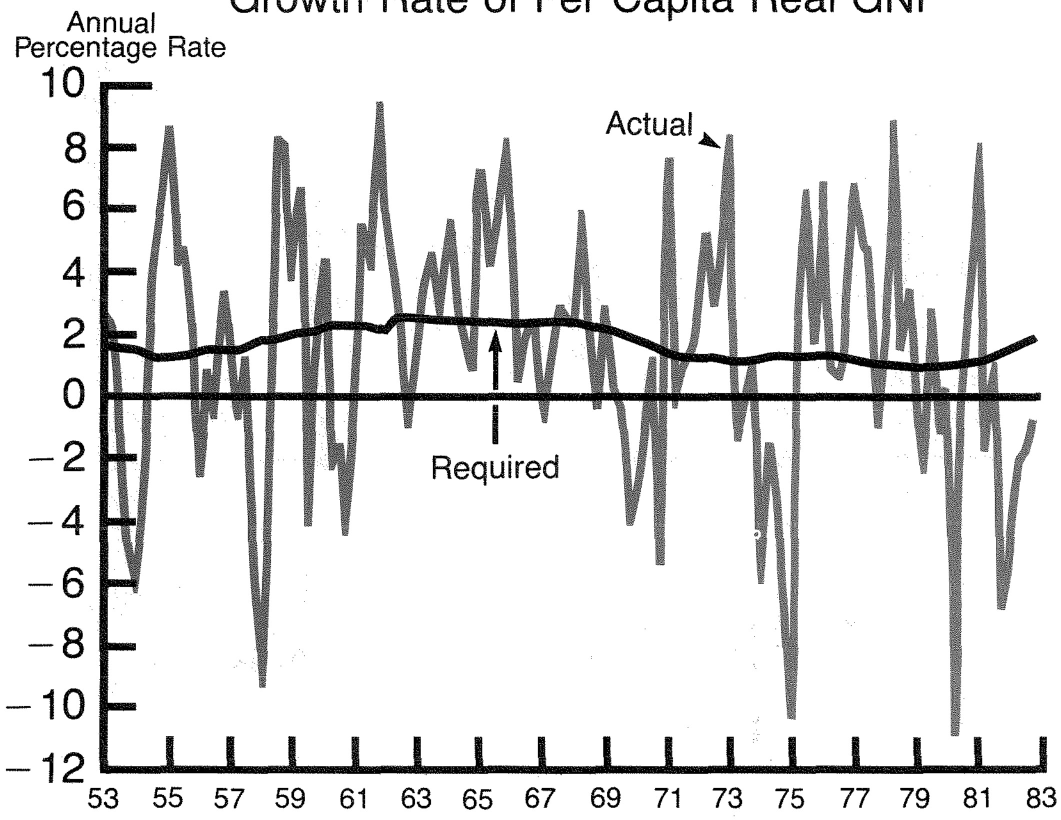
	Change in Unemployment Rate (Percentage Points) Predicted	Actual	Error
1983 Q1	+0.16	−0.23	+0.39
Q2	−0.30	−0.24	−0.06
Q3	−0.43	−0.80	−0.37
Q4	−0.37	−0.86	−0.49
1984 Q1	−0.48	−0.60	−0.12
Q2	−0.49	−0.37	+0.12
Q3	−0.13	−0.09	−0.04
Q4	+0.01	−0.22	−0.23
1982 Q4/1983 Q4	−0.95	−2.13	−1.18
1983 Q4/1984 Q4	−1.09	−1.27	−0.18

Predictions and Policy Implications

To test its predictive power, the model was used to forecast the unemployment rate over the period from 1984(Q1) to 1984(Q4). The forecast was made in two stages. In the first stage, the equations estimated in Box 2 were simulated over the forecast period holding the employment ratio constant at its 1953(Q1) level, and the resulting projections of labor productivity and participation were combined to yield quarterly estimates of the required per capita GNP growth rate. Over the eight-quarter forecast period, this required growth rate was estimated to increase modestly and to average slightly above two percent. In the second stage of the forecasting procedure, these projections of the required rate were entered into Equation C in Table 1 and that equation was simulated to produce forecasts of the employment ratio. Finally, these estimates were transformed into forecasts of the unemployment rate. These forecasts are shown in Table 2.

Over the eight-quarter period, actual per capita GNP growth averaged 5.2 percent. Simulation of the model predicted a decline in the unemployment rate of 2 percentage points. In fact, the unemployment rate declined by more than this: by 3.4 percentage points. The underprediction of the improvement in the employment ratio implies corresponding overpredicitons of the other components of real output growth. Examination of unrestricted simulations of the productivity and participation rate equations (that is, allowing the employment rate to vary rather than holding it constant) indicates that both female participation and labor productivity increased less rapidly over this period than historical relations would have pre-

Chart 4
Growth Rate of Per Capita Real GNP

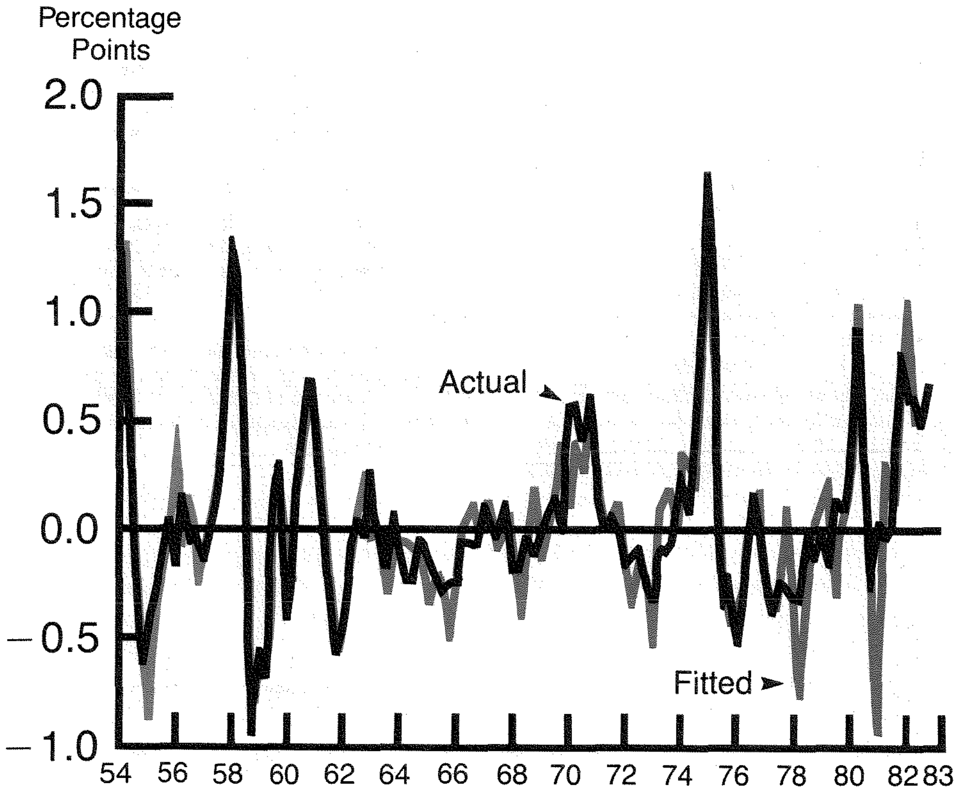


dicted. Thus, although the strong economic expansion did produce more rapid growth in productivity, participation and employment, the gain in productivity and participation was smaller than usual and hence, given the GNP growth that actually occurred, the gain in employment and the corresponding decline in the unemployment rate were larger. This outcome is somewhat ironic in view of earlier expectations that changes in tax policy would lead to *faster* productivity and labor supply growth.

However, Table 2 indicates that most of the prediction error occurred in 1983 when the unemployment rate fell much more rapidly than the model would have predicted. The error in 1984 was significantly smaller: from 1983(Q4) to 1984(Q4) the model predicted a decline in the unemployment rate of 1.1 percentage points compared to the actual decline of 1.3 percentage points. In view of this result, the model has been used to make projections of the unemployment rate over the four quarters of 1985.

To do so, the model was re-estimated through the fourth quarter of 1984. The estimated coefficients of the Okun's Law equation were essentially unchanged although the computed values of the required per capita GNP growth rate over 1983–84 were slightly lower than those forecasted on the basis of pre-1983 data. In making the forecasts for 1985, the required per capita GNP growth rate was assumed to remain constant at its 1984(Q4) level, namely two percent per annum. Real GNP was assumed to grow by four percent over the four quarters of 1985. Given the Census Bureau estimate that the adult population will rise 1.1 percent, this real growth assumption implies that per capita GNP will increase by 2.9 per-

Chart 5
Change in the Unemployment Rate Actual vs. Fitted Values



cent. On the basis of these assumptions, simulation of the model indicated that the unemployment rate would decline modestly from its level of 7.2 percent in 1984(Q4) to 6.8 percent in the fourth quarter of this year.

Most economic forecasters outside the Administration expect real GNP growth in 1985 to be less than the four percent rate assumed in making this forecast. Most forecasts cluster around 3½ percent growth. Thus, one possible conclusion from these estimates would be that it is unlikely that much further progress will be made toward lowering the nation's unemployment rate this year. An alternative conclusion would be that a somewhat more rapid rate of real growth would not bring the economy significantly closer to a level of the unemployment rate at which the inflation rate would be likely to rise. This appears to be the Administration's position as it has suggested as a target the four percent growth rate assumed above. The estimates developed in this paper suggest that even a four percent growth rate would produce only a relatively modest decline in the unemployment rate and hence would not add significantly to the risks of inflation.

FOOTNOTES

1. Brian Motley, "How Soon will the U.S. Reach Full Employment? An Assessment Based on Okun's Law" *Economic Review,* Federal Reserve Bank of San Francisco, Number 3, Summer 1984.

2. For the purposes of this paper, the phrase "required growth rate" is more descriptive than "potential growth rate." The latter describes the rate at which the *supply* of output *could grow* while the former is that at which *demand needs to grow* to hold the unemployment rate constant.

3. Throughout this paper the phrase "adult population" refers to the civilian non-institutional population.

4. Output per worker may, in turn, be decomposed into output per hour and hours per worker. For simplicity, this decomposition is not made in this paper.

5. Representing the unemployment rate by u and the employment ratio by e, it can be shown that dln e is approximately equal to $-du$. This means that if the employment ratio increases at an annual rate of one percent, the unemployment rate declines by one percentage point per year.

6. Motley, *op. cit.* and Douglas G. Woodham, "The Changing Relationship between Unemployment and Real GNP in the United States," Research Paper No. 8407, Federal Reserve Bank of New York, September 1984.

7. Rose McElhattan, "Is the Economy Overheating?" unpublished paper, Federal Reserve Bank of San Francisco, March 1984; George L. Perry, "Potential Output and Productivity," *Brookings Papers on Economic Activity,* 1, 1977; Charles S. Morris, "The Productivity 'Slowdown': A Sectoral Analysis," *Economic Review,* Federal Reserve Bank of Kansas City, April 1984.

8. In each of the studies cited in the preceding footnote, the influence of the business cycle on labor force participation and productivity is represented by changes in employment or unemployment.

9. The equations also include the *change* in the employment rate. In the simulations this term was set to zero after 1953(Q1).

10. When both the current and lagged values of the required growth rate are included in the estimated equation, none of their coefficients is individually significant, but their sum is very close to the coefficient on the current value when it alone is included.

11. See Motley, *op. cit.,* and Woodham, *op. cit.*

12. Woodham's results, for example, imply that over the 1974–1983 period it would have required a 2.3 percentage point increase in the GNP growth rate to lower the unemployment rate by one point. See Woodham, *op. cit.,* Table 4.

13. The dependent variable in the estimated equation is the quarterly change in the logarithm of the employment ratio. For purposes of Chart 5, the fitted values have been transformed into quarterly percentage point changes in the unemployment rate.