Behind the Money Market: Clearing and Settling Money Market Instruments

*David L. Mengle**

Whenever a money market instrument is traded, some means must exist for transferring the instrument and for making payment. In other words, there is a necessity for clearing and settling the trade, tasks that are usually referred to as operational, or back-office, functions.

Clearing refers to processing a trade and establishing what the parties to the trade owe each other. Settlement refers to the transfer of value between the parties so the trade is completed (Group of Thirty, 1989, p. 35). The first step in the clearing and settlement process involves conveying the details of the trade from traders to the back office. Second, the details must be compared and matched between the buyer and seller to ensure that both buyer and seller agree on what is to be traded and on what terms. Failure to do so might lead to delivery problems. This article will focus on what happens next: determination of the obligations between the parties and settlement of the trade.

Clearing and settlement systems link the participants in the money market. This article uses examples to describe how clearing and settlement take place for various types of money market instruments.[1] In addition, it discusses risks inherent in clearing and settlement, and the steps being considered to reduce such risks.

WHERE BANKS FIT IN

Banks and the interbank payment system are at the center of the clearing and settlement mechanism for the money market. Banks connect the participants in the money market by acting in three capacities. First, they act as agents for issuers of money market instruments, which means they perform the physical tasks of issuing and redeeming instruments in the market and of maintaining registration records. Second, they act as custodians of instruments, which involves safekeeping them as a service to investors. Like valuables kept in a safe-deposit box, instruments entrusted to a custodian bank do not show up on the bank's balance sheet as either assets or liabilities because they remain the property of their owners.

Finally, and most importantly, some banks specialize in clearing. A clearing bank is responsible for transferring securities from one party to another and for transferring payment for the securities. Dealers maintain two types of accounts at clearing banks: securities accounts and funds accounts. When a clearing bank is instructed to transfer securities from Dealer A's securities account to that of Dealer B, the bank also transfers payment for the securities from Dealer B's funds account to that of Dealer A. If the dealers do not use the same clearing bank, then the transaction involves a transfer of securities and funds between two banks.

Transfers between banks take place at the hub of the money market, the interbank payment system. Even when instruments are cleared outside the banking system, as is the case when a dealer firm clears for itself, payment takes place through banks. The payment system, which links banks to each other, includes both paper checks and electronic funds transfer, although almost all interbank payments now occur electronically over wholesale wire transfer networks.[2]

The main wholesale wire transfer network in the United States is Fedwire, which operates through bank reserve accounts at the twelve Federal Reserve Banks. Fedwire can be used to transfer both funds and book-entry U.S. government securities (to be described presently) between banks and other depository institutions. During 1991, about 260,000 Fedwire funds transfers totaling about $766 billion occurred on an average day. Mean transfer size was about $3 million. In addition, over 44,000 book-entry

* Vice president, Morgan Guaranty Trust Company of New York. The article was written while the author was research officer at the Federal Reserve Bank of Richmond.

[1] For a more detailed description of the operational side of the money market, the reader should consult Marcia Stigum's treatment in *After the Trade.*

[2] Wholesale wire transfer networks link banks with each other. In contrast, retail wire transfer systems, such as automated teller machine networks, link banks with consumers.

securities transfers amounting to about $476 billion occurred daily. The average book-entry transfer was about $10.8 million.

Figure 1 shows how Fedwire is used to complete a federal funds transaction. Assume that Bank of Downtown finds itself with $10 million of excess reserves while Midtown Trust is $10 million short of required reserves. A broker matches the two and arranges for Downtown to sell (lend) $10 million to Midtown, so Downtown's excess reserves will be used to fund Midtown's shortage. Settlement of the transaction will occur through reserve accounts at their Federal Reserve Bank.[3] When Downtown initiates the transfer, its reserve account at the Fed is reduced by $10 million. Within a split second, Midtown's reserve account is increased by the same amount. Once made, the Fedwire payment is final and irrevocable. Notice that on the books of the Fed the transfer simply moves reserves from the account of one bank to that of the other. The next day, Midtown uses Fedwire to repay the funds and essentially reverses the process.

An important feature of Fedwire transfers is that they are settled on a bilateral, trade-for-trade basis, also known as gross settlement. If, instead, transfers were consolidated into net positions between banks or between banks and the network in order to reduce the actual number of interbank transfers that take place, the system would be called a netting system (see box, "Netting and Net Settlement"). Netting can take two forms. Bilateral netting combines gross obligations between banks into net obligations so each pair of banks in a system exchanges only one settlement payment. Multilateral netting combines each bank's bilateral net positions into "net net" obligations between the bank and the other banks in the system. When settlement occurs, each bank is either a net creditor (one that is owed money by the rest of the system) or a net debtor (one that owes money).

The Clearing House Interbank Payments System (CHIPS) is a multilateral netting system. It is owned and operated by the New York Clearing House, a private organization. CHIPS transfers only funds and not securities, and is used largely, although by no means exclusively, in connection with international transactions such as Eurodollars and foreign exchange (Clair, 1991). During 1991 approximately 150,000 transfers totaling about $866 billion took place on an average day on CHIPS. Average transfer size was $6 million. At the end of 1991, 126 depository institutions, many of them branches of foreign banks, participated in CHIPS.

CHIPS is organized in a hierarchical fashion whereby a subset of participating banks (20 out of 126) settle directly with CHIPS while the others must settle on the books of one of the settling banks. Settlement occurs at the end of the day, when settling banks in net debit positions send (over Fedwire) the funds they owe to a special CHIPS net settlement account at the Federal Reserve Bank of New York. CHIPS then wires funds from the account to settling banks in net credit positions. The special account starts out with a zero balance and, when settlement is complete, ends with a zero balance; the CHIPS account is used for nothing else.

The results of a 1987 survey of New York banks highlight the international character of CHIPS payments relative to Fedwire payments (Federal Reserve Bank of New York, 1987-88). According to the survey, 55 percent of the dollar amount of CHIPS payments was related to foreign exchange

[3] If the two banks are in separate Federal Reserve districts, the transaction will involve accounts at two different Federal Reserve Banks.

Figure 1

Fedwire Message between Banks

(in millions of dollars)

Bank of Downtown

Assets	Liabilities
Reserves −10	

Federal Reserve Bank

Assets	Liabilities
	Reserves, Downtown −10 Reserves, Midtown +10

Midtown Trust

Assets	Liabilities
Reserves +10	

Netting and Net Settlement

In order to understand how netting and net settlement work, consider the example of the four banks in Table 1, each of which sends a payment message to each of the other three banks. Bank of Downtown sends transfer messages for $10 million to Midtown Trust, $10 million to Crosstown National Bank, and $10 million to Outatown Bank; Midtown sends $10 million to Downtown, $10 million to Crosstown, and $40 million to Outatown; and so on for a total of 12 separate payments. On a gross settlement system like Fedwire, each of the 12 payment transactions would be settled separately.

Table 1

Payment Messages
(in millions)

Sender	Receiver	Amount
Downtown	Midtown	$10
Downtown	Crosstown	$10
Downtown	Outatown	$10
Midtown	Downtown	$10
Midtown	Crosstown	$10
Midtown	Outatown	$40
Crosstown	Downtown	$40
Crosstown	Midtown	$30
Crosstown	Outatown	$20
Outatown	Downtown	$20
Outatown	Midtown	$20
Outatown	Crosstown	$10

If, instead, each bank's obligations to each of the other banks were combined, that is, netted bilaterally, then the result would be the net positions in the first four columns of Table 2. In such a netting system, each bank (read from the left of the matrix) would be in a net credit or net debit position versus each of the other banks (read from the top of the matrix), and settlement would take place when the banks send net payments to or receive net payments from each of the other banks at the end of the day. Since Downtown sent a payment message for $10 million to Crosstown but received one from Crosstown for $40 million, Downtown will have a net credit of $30 million versus Crosstown (which, correspondingly, has a net debit of $30 million against Downtown). Midtown will send $20 million to Outatown; Crosstown will send $30 million to Downtown, $20 million to Midtown, and $10 million to Outatown; and Outatown will send $10 million to Downtown. Since Downtown's and Midtown's payments to each other cancel out, neither will have to send a payment to the other.

Multilateral netting takes the netting process one step further by combining the bilateral net positions for each bank into a net position versus the network. The network adds up the amounts each owes to and is owed by the other banks (obtained by summing the net positions in a bank's row of the matrix). This results in the net net positions shown in the last column of the matrix: Downtown has a net credit of $40 million coming in, Crosstown has a net debit of $60 million going out, Outatown has a net credit of $20 million, and Midtown's incoming funds are offset by its outgoing funds. Settlement occurs when Crosstown sends the network $60 million and the network wires $40 million to Downtown and $20 million to Outatown.

Moving to bilateral netting and then to multilateral netting can mean substantial reductions in the number of actual exchanges between participants. In Table 1 the gross number of transactions is 12 but the number could be far more. By moving to bilateral netting, the number of exchanges of funds is reduced to a maximum of six or, more generally,

$$\frac{n(n-1)}{2},$$

where n is the number of participating institutions. By moving to multilateral netting, the maximum number of exchanges is reduced to n, which in the example is four. Such reductions in the number of exchanges can mean reductions in operational costs and risk exposures between institutions. For specific examples of how risks can both arise in and be avoided by netting, see Gilbert (1992).

Table 2

Net Bilateral and Net Multilateral Settlement Obligations
(in millions)

	Downtown	Midtown	Crosstown	Outatown	Net Net
Downtown		$ 0	$30	$10	$40
Midtown	$ 0		$20	($20)	$ 0
Crosstown	($30)	($20)		($10)	($60)
Outatown	($10)	$20	$10		$20

Note: Numbers in parentheses denote a net debit; those not in parentheses, a net credit.

transactions; on Fedwire, foreign exchange transactions were negligible. Further, 28 percent of CHIPS dollar value was related to Eurodollar placements; on Fedwire, such transactions were 10 percent of dollar value. Finally, 34 percent of Fedwire dollar value was for federal funds transactions; on CHIPS, the percentage was almost zero.

One last network deserves mention because of its role in international payments. The Society for Worldwide Interbank Financial Telecommunication (SWIFT) is a nonprofit cooperative chartered under Belgian law and owned by 1,885 participating institutions in 73 countries, including the United States. Unlike Fedwire or CHIPS, SWIFT is not a funds transfer system. Instead, SWIFT payment messages instruct banks to transfer funds by means of accounts at correspondent banks.[4] Such a transfer might involve transfers among accounts at the same bank. For example, suppose Bank of Downtown serves as correspondent bank for both Midtown Trust and London Bank and that London Bank wishes to make a payment to Midtown Trust. London makes the payment by sending a SWIFT message instructing Downtown to reduce London's correspondent account and to increase Midtown's by the amount of the payment. Alternatively, a SWIFT message might direct that a payment be made between banks. If London wishes to make a payment to Crosstown National, for example, but Crosstown does not have a correspondent relationship with Downtown, then London's SWIFT message would instruct Downtown to transfer funds (from London's correspondent account) to Crosstown by means of an interbank network like Fedwire or CHIPS.

FORMS OF MONEY MARKET INSTRUMENTS

The form in which a money market instrument is issued and traded largely determines the manner in which it is cleared and settled. Because federal funds are essentially exchanges of bank reserves between accounts at Federal Reserve Banks, they are settled by means of Fedwire transfers. For other money market instruments, how they are cleared and settled depends on whether they are traded in physical (also called "definitive") form or book-entry form. Trades of physical securities may require that paper instruments move between institutions, while trades of book-entry securities only involve changes in computer account entries.

Physical Securities

At present, bankers acceptances, large certificates of deposit (CDs), and some commercial paper issues are issued in physical form; that is, they use paper certificates to represent the obligation of the issuer to the purchaser. Clearing physical securities works as follows. Suppose Hoozon First Securities decides to purchase $10 million of CDs from Watson Second Securities. Suppose also that Hoozon uses Downtown as its clearing bank and Watson uses Midtown. After the securities firms' back offices notify their clearing banks of the trade, Midtown pulls the CDs from the vault and a courier delivers them to Downtown. Downtown then sends over Fedwire $10 million in payment to Midtown. Downtown charges Hoozon for the payment while Midtown credits Watson. The trade between the dealers has been cleared and settled. If Hoozon then sells $5 million of the CDs it bought to Zippi Industries, one of its corporate customers, and if Crosstown National serves as Zippi's custodian bank, it will be necessary for Downtown to deliver the securities to Crosstown for safekeeping and for Crosstown to make a payment to Downtown.

A dealer might elect to clear securities itself. In the above example, self-clearing would mean that securities would be moved directly between the dealers (or between a self-clearing dealer and a clearing bank). Whether a dealer clears for itself or uses a bank depends on whether or not the additional costs of running a clearing operation outweigh the benefits of possibly faster clearing and greater control over the operation. But even if a dealer clears for itself, it will still use a bank for settlement because only banks (or, more accurately, depository institutions) have accounts at the Federal Reserve.[5]

Physical securities by their nature involve handling and delivery costs as well as risks of theft. Consequently, there are incentives for keeping (or "immobilizing") physical securities in depositories instead of requiring that the securities be physically moved each time they are traded. When a security held in a depository is sold, the depository's files are updated to reflect the change of ownership. In other words, a depository effectively converts an exchange of physical securities into an exchange of book-entry securities (McAndrews, 1992). Taking the process

[4] Correspondent banks perform services for other banks in return for fees or minimum deposit balances.

[5] A dealer could avoid using banks for settlement if it physically delivered cash in payment for securities. Transportation costs and theft risks ensure that virtually all payments take place through banks.

one step further, the physical security can be eliminated altogether (or "dematerialized"), and the security can be issued, cleared, settled, and redeemed in book-entry form on the computer files of the depository. As more types of money market instruments become eligible for conversion to book-entry form, cost considerations could quickly turn physical securities into an anachronism.

Book-Entry Securities

Money market instruments have been moving from physical to book-entry form by means of depositories. In particular, the Depository Trust Company (DTC), a New York limited-purpose trust company owned jointly by banks, broker-dealers, and other financial organizations, has been active in making more instruments eligible for conversion to book-entry form. The movement to book entry has been rapid. Municipal securities became eligible for book entry in 1981; by the end of 1991, 77 percent of the value of municipal notes outstanding was issued through DTC in book-entry form and involved no physical securities (DTC, 1991). Commercial paper became eligible for book entry in 1990; by May 1992, 42 percent of the value of the commercial paper market was issued through DTC entirely in book-entry form. And as of this writing, DTC was attempting to make large CDs and bankers acceptances eligible for book entry.

U.S. government securities, including Treasury bills, are now issued only in book-entry form. That is, instead of being represented by paper certificates, obligations of the United States are now recorded as entries on the computer files of the Federal Reserve Banks and commercial banks. The Treasury and Federal Reserve System completed a switch to book-entry securities in 1986 because of concerns about security and the costs of processing and moving huge quantities of paper instruments.

Every Treasury security issue is represented by an entry on a Federal Reserve Bank's computer. The Fed keeps track of which bank holds a particular portion of an issue and, at maturity, transfers funds in repayment to the bank holder. But while the Fed maintains securities accounts in order to keep track of the outstanding issue balance, the accounts do not show up on the Fed's balance sheet. Rather, they reflect the Fed's custody of the Treasury security issue for the various depository institutions. Similarly, when a bank purchases a Treasury security for the account of a customer, the bank is not the actual owner even though the Fed's computer assigns a security balance to that bank.

Now for a transaction. Say that the Bank of Downtown purchases $10 million of Treasury bills from Midtown Trust. When the securities are transferred over Fedwire, two offsetting transactions take place simultaneously: the exchange of securities and the exchange of funds in payment. The movement of Treasury bills takes place by decreasing Midtown's book-entry securities account at the Federal Reserve Bank and by increasing Downtown's by the same amount. Payment occurs as shown in Figure 1 and involves a transfer of funds from Downtown's reserve account to Midtown's. Because funds and securities are transferred at the same time, such a system is called a "delivery versus payment" system.

The preceding example only shows what would happen if the purchasing bank were holding the securities for its own account. Now, suppose that Hoozon First Securities purchases the $10 million of Treasury bills from Watson Second Securities. If Hoozon uses Downtown as its clearing bank and Watson uses Midtown, Downtown increases Hoozon's securities account by $10 million and decreases its funds account by the same amount in payment. At the other end, Midtown decreases Watson's securities account and increases its funds account by $10 million. On Fedwire the securities move from Midtown to Downtown and the payment moves in the opposite direction. Note that actual ownership of the security moves from Watson to Hoozon and does not rest with either bank. The banks and the Federal Reserve are simply the conduit through which ownership of securities is passed.

EURODOLLARS

Trades involving Eurodollar deposits differ from those of domestic instruments in that they entail corresponding transactions in the United States and overseas and also are likely to involve the CHIPS and SWIFT networks. Eurodollar deposits are dollar deposits held outside the United States in either a foreign bank or an overseas branch of a U.S. bank. Inside the United States, Eurodollars can be held only by international banking facilities of domestic or foreign banks. When Eurodollar deposits move between banks, they normally involve corresponding entries on the balance sheet of some organization located in the United States.

Figure 2 shows an example in which the Bank of Downtown raises $10 million of interbank deposits from London Bank in the Eurodollar market; the transaction takes place through Downtown's

Figure 2

Settlement of a Eurodollar Funds Purchase

(in millions of dollars)

Bank of Downtown, London Branch

Assets	Liabilities
Account at U.S. office +10	Deposit, London Bank +10

London Bank (U.K.)

Assets	Liabilities
Deposit, Midtown Trust −10	
Deposit, Bank of Downtown +10	

Bank of Downtown (U.S.)

Assets	Liabilities
Reserves +10	London branch account +10

Midtown Trust (U.S.)

Assets	Liabilities
Reserves −10	Deposits, London Bank −10

Federal Reserve Bank

Assets	Liabilities
	Reserves, Bank of Downtown +10
	Reserves, Midtown Trust −10

London branch.[6] Because London Bank is not headquartered in the United States, any dollar-denominated transaction in which it engages must ultimately go through a correspondent bank in the United States. London uses Midtown Trust as a correspondent, so the transfer occurs through London's account at Midtown and then through Midtown's and Downtown's reserve accounts at the Federal Reserve Bank.

Once Downtown and London have agreed to the transaction, London sends Midtown a transfer message over the SWIFT network instructing that its balance with Midtown be decreased by the amount of the transfer. In carrying out the transfer of reserves to Downtown, Midtown would normally use the CHIPS network. The transaction is settled at the end of the day when CHIPS goes through net settlement and reserves are transferred from Midtown to Downtown.

There are specialized networks and facilities for clearing and settling other Eurodollar instruments. For example, Euro-commercial paper, Euro-notes, and Eurodollar CDs are commonly cleared and settled in both the Euroclear and CEDEL systems. Euroclear, originally formed to clear Eurobond trades, is owned by a Belgian cooperative and operated under contract by the Brussels branch of Morgan Guaranty Trust Company. Securities are immobilized in a network of depositories and settled in book-entry form; funds transfers in connection with book-entry settlement take place through deposits on Morgan's books.

CEDEL is a Luxembourg corporation, specially chartered as a clearing organization. As with Euroclear, securities settled over CEDEL are immobilized in depositories; unlike Euroclear, funds transfers in connection with book-entry securities settlement take place through deposits with the CEDEL clearing organization itself.

Finally, Eurodollar instruments can be cleared and settled by banks. For example, the First National Bank of Chicago operates the First Chicago Clearing Centre in London in order to provide custodian, agent, and clearing bank services for Eurodollar instruments, primarily dollar-denominated CDs. Funds transfers associated with movements of

[6] The London branch's account with Downtown's headquarters bank in the United States is carried on the liability side of the U.S. bank's books as "due to" its branch and on the asset side of the London branch's books as "due from" its parent bank.

Eurodollar instruments take place on the books of First Chicago's London branch.

RISK AND RISK CONTROLS[7]

Given the daily volume and value of transactions that occur in the money market, the opportunities for loss as the result of default or operational problems are potentially huge. Consequently, over the last decade both market participants and regulators have devoted a great deal of effort to formulating policies for keeping risks within acceptable limits.

Policy discussions often distinguish among several forms of risk (Parkinson et al., 1992). First, credit risk refers to potential losses arising from a clearing and settlement system participant defaulting on some or all of its settlement obligations. Second, liquidity risk arises from the possibility that settlement could be delayed because of temporary unavailability of funds. The distinction between credit risk and liquidity risk lies in the temporary nature of illiquidity as opposed to the permanent nature of default. Third, systemic risk refers to the danger that the failure of one participant to settle its obligations could lead to liquidity problems or settlement failure on the part of others. Finally, operational risk stems from the possible breakdown of computer systems or other elements of the clearing and settling mechanism.

Fedwire provides the most transparent example of credit risk. The Fedwire transaction shown in Figure 1 omits an important point: In order for the transfer to take place, it is not necessary that the sending bank always have sufficient funds in its reserve account to cover the transfer.[8] If at the time of the transfer in Figure 1 the Bank of Downtown has only $5 million on deposit as reserves, Downtown incurs a "daylight overdraft" of $5 million. That is, its reserve account is allowed to go negative during the day so long as the deficit is made up before close of business. Further, the receiving bank will have final payment at the time of the transfer regardless of whether the overdraft is ultimately covered. Until the overdraft is covered, the Federal Reserve Bank assumes the credit risk of Downtown's failing to provide the necessary funds. While credit risk has effectively been socialized by transferring it to the Fed, systemic risk has been eliminated because there is no avenue for losses to spread to other banks in the system.

On CHIPS, credit, liquidity, and systemic risks can all arise. For example, suppose the Bank of Downtown receives a CHIPS transfer message from Crosstown National for a payment to one of its corporate customers. Although CHIPS does not settle until the end of the day, it may be Downtown's practice to allow its customer to withdraw the funds prior to settlement. In allowing such access to funds, Downtown assumes the risk that Crosstown might fail to meet its net settlement obligation at the end of the day. More serious, the failure of Crosstown to settle a particularly large net debit position could conceivably cause a chain reaction of settlement failures among other participants, some of which might depend on the receipt of payments from the failing bank in order to fund their obligations (Humphrey, 1986). Measures to control such risk will be discussed presently.

Finally, operational risks may be illustrated with the following incident that occurred in 1985. The Bank of New York, acting as a clearing bank for book-entry Treasury securities, had an internal computer problem that allowed the bank to accept securities but not to process them for delivery to dealers, brokers, and other market participants. The bank's reserve account was debited for the amount of the securities, but the bank was unable to re-send them and collect payment. The result was a growing daylight overdraft in the Bank of New York's reserve account. As it became increasingly clear that the problem would not be fixed by close of business, the bank borrowed from the discount window. The problem was fixed during the night so the loan was repaid the following day.

As one might guess, the above risk categories overlap considerably. For example, operational problems at a bank could lead to liquidity problems, which in turn might cause systemic problems with other banks. In addition, operational problems could extend to accounting systems and thereby make it difficult for system participants to monitor their credit exposures to other participants. Finally, at the time a participant fails to meet its settlement obligations, the other participants are unlikely to be able to determine whether the problem is the result of default or illiquidity. Still, the distinctions are important to policymakers because each category of risk requires different solutions. For example, operational risks might lead to policies designed to create incentives to develop backup facilities and procedures to keep

[7] For more comprehensive discussions, see Group of Thirty (1989), Juncker, Summers, and Young (1991), and Parkinson et al. (1992).

[8] Exceptions to this general rule include weak institutions whose overdrafts are either prohibited or monitored in real time.

systems running, credit risks might suggest loss sharing arrangements and limits on risk exposure, and liquidity risks might call for emergency lending arrangements.

Risk-control measures cover a wide spectrum. The simplest are membership standards, which seek to head off settlement problems by excluding from a system those participants lacking the financial strength and operational expertise to assure that settlement obligations can be met. Once a participant is admitted, the clearing organization should monitor the participant's financial condition to ensure that it does not pose losses to the other members.

Another form of risk control is quantitative limits on risk exposure. Examples include net debit caps and bilateral net credit limits. Net debit caps are limits on the size of a bank's combined daylight overdraft on Fedwire and net debit position on CHIPS. In other words, they attempt to control the risk a bank poses to the payment system by limiting how much a bank can, on balance, owe others over the wire transfer networks. Bilateral net credit limits specify the maximum net transfer a bank on CHIPS is willing to receive from a particular sending bank; that is, they provide a means for a bank to control its own exposure to other banks. Net debit caps on Fedwire and CHIPS and bilateral net credit limits on CHIPS were part of the original Federal Reserve risk-control policy adopted in 1986.

Risks to a clearing and settlement system can also be limited by requiring system participants to put up collateral to cover their obligations to the system. If a participant defaults, the collateral is used to cover the losses. In effect, such a requirement amounts to a performance bond that a participant forfeits if it defaults on its settlement obligations.

A form of risk-control policy that seeks to create economic incentive to control risks is explicit pricing of daylight overdrafts (Mengle, Humphrey, and Summers, 1987). The rationale for pricing is that it will impose a cost on using intraday credit and thereby provide incentives to reduce risk exposures and to more efficiently allocate intraday credit. In 1992 the Federal Reserve approved a charge on daylight overdrafts that exceed 10 percent of an institution's risk-based capital. By 1996 the charge will be $6.85 per day per $1 million (that is, an annual rate of 25 basis points) of average Fedwire daylight overdrafts arising from funds transfers and book-entry securities transfers that exceed 10 percent of an institution's risk-based capital.

A fifth form of risk-control policy is loss sharing among members of a net settlement system. Under a loss-sharing agreement, banks that are members of a system share the losses caused by another member's failure to settle. A loss-sharing agreement generally requires two characteristics to make it work. The first is settlement finality, that is, assurance that settlement entries will not be reversed in the event of one bank's failure to settle. Second, in order to make the loss-sharing agreement credible, banks are generally required to contribute collateral to a clearing fund, which can be drawn upon in the event of a settlement failure and can also serve as security for an emergency line of credit. By imposing costs on system participants if a failure occurs, a loss-sharing agreement can create incentives for banks to monitor the soundness of other banks in the system. CHIPS adopted settlement finality and a loss-sharing agreement in 1990.

A sixth means of risk control is obligation netting, that is, combining a set of offsetting gross payment of securities obligations into net obligations (see box, "Netting and Net Settlement"). Netting, be it bilateral or multilateral, can reduce operational risks by reducing the volume of transactions that actually pass through a clearing and settlement system. And provided that the underlying legal obligations between participants are netted along with the positions, netting can reduce credit risks between banks by reducing the total amount of funds and securities that actually must be transferred between banks (Gilbert, 1992).

The Government Securities Clearing Corporations (GSCC) was established in 1986 to provide netting of government securities trades for banks and other securities brokers and dealers. It works as follows. Participants submit data on all securities transactions to be settled on a particular day. First, the trades are compared. Then, each participant's transactions of each issue are added up into a net credit or debit security settlement position for each issue and a single funds settlement position. The netting process is the same as the multilateral arrangement shown in the box, except for GSCC the numbers would refer to sales or purchases of a specific issue of government securities instead of CHIPS funds transfers. Settlement occurs over the Fedwire book-entry system: Clearing banks deliver (against payment) net securities positions to GSCC; in turn GSCC sends (against payment) the netted amounts of each issue to receivers.

While netting can reduce operational risks as well as credit risk, it has the potential to increase systemic risk. In response to concerns about systemic risk, the GSCC has adopted three measures to deal with the default of a participant. First, GSCC requires that members contribute to a clearing fund. Second, it maintains a line of credit on which to draw in the event of liquidity problems. Finally, it has in place rules for sharing losses in the event of a default.

A final means of reducing risk, one that is applicable to systems for clearing and settling securities, is moving securities to book-entry, delivery-versus-payment form. Delivery versus payment helps reduce credit risk exposure because making the exchange of funds and securities simultaneous (or nearly so) eliminates (or greatly reduces) the time between delivery of securities and payment of funds during which a participant could fail to meet its obligation. In addition, book entry reduces operational risks by eliminating physical delivery of instruments.

While book entry and delivery versus payment reduce exposure to a defaulting participant, they do not eliminate it entirely. In order to provide additional protection against losses if a participant defaults, the Federal Reserve has issued guidelines for risk controls on privately operated book-entry systems (*Federal Register*, June 21, 1989). A specific example of such controls is in DTC's book-entry commercial paper facility. DTC's safeguards include a clearing fund contributed to by participants, net debit caps and a requirement that a participant maintain collateral on its net debit position (*Federal Register*, October 17, 1990).

To some extent designing a program for risk reduction entails trade-offs between various types of risk. For example, until 1981 CHIPS did not settle until the day after the transfer messages were made. That gave rise to overnight credit risk. When CHIPS moved to same-day settlement, credit risk was reduced (or made shorter in duration), but operational risk most likely increased, at least temporarily, since there was less time to prepare for settlement. In practice, the challenge in developing new clearing system technologies is to reduce credit and systemic risks while avoiding operational risks.

REFERENCES

Clair, Robert T. "The Clearing House Interbank Payments System," *Research in Financial Services*. JAI Press, 1991.

Depository Trust Company. *Annual Report*, 1991.

Federal Reserve Bank of New York. "Large Dollar Payment Flows from New York," Federal Reserve Bank of New York *Quarterly Review*, vol. 12 (Winter 1987-88), pp. 6-13.

Gilbert, W. Alton. "Implications of Netting Arrangements for Bank Risk and Foreign Exchange Transactions," Federal Reserve Bank of St. Louis *Review*, vol. 74 (January/February 1992), pp. 3-16.

Group of Thirty. *Clearance and Settlement Systems in the World's Securities Markets*, March 1989.

Humphrey, David B. "Payments Finality and the Risk of Settlement Failure," in Anthony Saunders and Lawrence J. White, eds., *Technology and the Regulation of Financial Markets*, pp. 97-120. Lexington, Massachusetts: D.C. Heath, 1986.

Juncker, George R., Bruce J. Summers, and Florence M. Young. "A Primer on the Settlement of Payments in the United States," *Federal Reserve Bulletin*, vol. 77 (November 1991), pp. 847-58.

McAndrews, James J. "Where Has All the Paper Gone? Book-Entry Delivery-Against-Payment Systems," Federal Reserve Bank of Philadelphia *Business Review*, November/December 1992, pp. 19-30.

Mengle, David L., David B. Humphrey, and Bruce J. Summers. "Intraday Credit: Risk, Value, and Pricing," Federal Reserve Bank of Richmond *Economic Review*, vol. 73 (January/February 1987), pp. 3-14.

Office of the Federal Register. *Federal Register*, vol. 54, June 21, 1989.

——————. *Federal Register*, vol. 55, October 17, 1990.

Parkinson, Patrick, Adam Gilbert, Emily Gollob, Lauren Hargraves, Richard Mead, Jeff Stehm, and Mary Ann Taylor. "Clearance and Settlement in U.S. Securities Markets," Board of Governors of the Federal Reserve System, Staff Study 163, 1992.

Stigum, Marcia. *After the Trade: Dealer and Clearing Bank Operations in Money Market and Government Securities*. Homewood, Illinois: Dow Jones-Irwin, 1988.

How Useful Is M2 Today?

Robert L. Hetzel

I. INTRODUCTION

The actions of the Federal Reserve System determine the nominal (dollar) expenditure of the public. A key issue for policymakers is what variable best measures the impact of monetary policy actions on nominal expenditure.[1] The press uses changes in the funds rate as an indicator of the thrust of monetary policy. Declines are labeled "easing moves," that is, changes that will augment the rate of growth of nominal expenditure, and conversely with increases. The usefulness of the funds rate as an indicator, however, is contradicted by current experience. The funds rate fell from almost 10 percent in May 1989 to 3 percent in September 1992. Over this same period, however, the trend rate of growth of nominal GDP dropped from 7 percent to around 4 percent.

This paper examines whether the monetary aggregate M2 offers useful information about the impact of monetary policy actions on nominal expenditure.[2] By definition, nominal expenditure equals the amount of dollars in circulation times the average number of times per year those dollars turn over against nominal output. That is, nominal expenditure is the quantity of money times the velocity of circulation of money. M2 is useful as a definition of money if its velocity is a simple, predictable function of a small number of variables. Equivalently, M2 is a useful definition of money if unpredictable changes in M2 velocity are small compared to changes in nominal expenditure.

Section II examines the predictability of M2 velocity. Section III discusses M2 indicator variables. Section IV examines arguments that special factors are currently making M2 velocity less predictable.

II. IS M2 VELOCITY PREDICTABLE?

This section examines the predictability of M2 velocity initially by checking whether growth rates of nominal GDP move with growth rates of M2 over long periods of time. It then examines M2 velocity more carefully by estimating an M2 demand regression equation. Table 1 shows annual growth rates of M2 and nominal GDP, with M2 lagged two years.

Table 1

Growth Rates of Nominal GDP and M2 Lagged Two Years

	GDP Growth		M2 Growth	
	annual	average	annual	average
1973	11.8		12.1	
1974	8.1		12.5	
1975	8.7		9.9	
1976	11.5	(10.8)	6.1	(10.9)
1977	11.6		9.3	
1978	13.1		13.0	
1979	11.5		12.7	
1980	8.8		8.5	
1981	11.9		8.3	
1982	3.9		8.0	
1983	8.1		9.4	
1984	10.9	(9.0)	9.3	(7.8)
1985	6.9		12.5	
1986	5.7		8.2	
1987	6.4		8.9	
1988	7.9		8.2	
1989	7.0		6.6	
1990	5.1	(5.2)	5.2	(5.0)
1991	2.9		3.9	
1992			5.3	
1993			3.2	

Notes: The numbers in parentheses show average GDP growth for the years 1973 to 1979, 1980 to 1988, and 1989 to 1991 and average M2 growth for the corresponding periods two years earlier.

[1] As shown by the accounting of the national income and product accounts, aggregate nominal expenditure for final products (aggregate nominal demand) equals aggregate nominal output (aggregate nominal supply) minus changes in business inventories. Aggregate nominal output (within a country's own borders) is the dollar value of gross domestic product (GDP). Again, as an accounting matter, aggregate nominal output equals aggregate nominal income.

[2] It is important to distinguish among instruments, indicators and targets. The usefulness of the funds rate as an instrument is not discussed here: nominal GDP is certainly higher now relative to what it would have been if the Fed had not lowered the funds rate. The use of M2 as a target is also not discussed. The issue is whether the behavior of M2 offers more useful information about aggregate nominal expenditure than the funds rate.

(The lag prevents contemporaneous inverse movements in M2 and its velocity from obscuring the longer-run relationship between growth in M2 and nominal output.) Over the three periods shown, the trend rate of growth of nominal GDP matches fairly closely the trend rate of growth of M2.

If velocity is stable, the rate of inflation will correspond over long periods of time to the excess of the rate of growth of money over output. To illustrate, over the three decades from 1960 through 1990, the excess of the annualized rate of growth of M2 (8.1 percent) over the annualized rate of growth of real GDP (3.0 percent) was 5.1 percent, while annualized inflation (measured by the implicit GDP price deflator) was 4.9 percent.

The inverse of velocity, the fraction of its income the public wants to hold in the form of money, expresses the real value of money. The remainder of the section examines the stability of M2 velocity by examining the stability of M2 demand regressions, which predict the behavior of real M2. Specifically, this section looks at the prediction errors from an updated version of a money demand regression similar to one estimated by Friedman and Schwartz (1982). (Estimation details are in the appendix.) The period of estimation is from 1915 to 1991. Using annual observations, real M2 is regressed on real output and on opportunity cost variables measuring the rate of return on financial market assets and on physical assets. The financial market opportunity cost of holding M2 is proxied for by the difference between the commercial paper rate (R) and a weighted average of the explicit rates of return paid on the components of M2 (RM2). [Hetzel (1989) describes the construction of RM2.]

Following Friedman and Schwartz (1982), the regression employs the percentage change in nominal output as a proxy for the market yield on physical assets. The market yield on physical assets (land, buildings, machinery, consumer durables, etc.) possesses two components: a real rate of return and an anticipated change in dollar value. The percentage change in nominal output also possesses two components: the rate of growth of real output and the rate of inflation. These two components of the percentage change in nominal output proxy for the two components of the market yield on physical assets.[3]

Over the entire estimation period, the fitted M2 demand function exhibits considerable stability. This stability can be observed directly by noting that M2 velocity (nominal output divided by M2) has fluctuated around a value of 1.63 since 1915; M2 (real and nominal) and output (real and nominal) would not gravitate around each other over time unless unpredictable changes in the demand for real M2 cancelled.[4] Particularly in the post-World War II period, prediction errors are relatively small. Over the period 1950 to 1991, the mean absolute errors in predicting the level of real M2 and changes in real M2 are, respectively, 2.2 and 1.0 percent. (Errors are from the regressions in Tables A1 and A2 of the appendix.) The exception to the statement that the public's M2 demand function was stable over the period 1915 to the present is that after the mid-1960s the public's demand for real M2 became less sensitive to variation in market rates.

The regression equation in first differences, shown in Table A2 of the appendix, generates errors in 1989, 1990 and 1991 that cumulate to an overprediction of real M2 of 5.3 percent. This overprediction of real M2 has continued to grow during 1992. As noted above, M2 is a useful definition of money if unpredictable changes in M2 velocity are small

---

[3] This proxy makes two assumptions. First, there is a positive relationship between fluctuations in real output and the economy's equilibrium real rate of interest. When economic activity strengthens, the real rate of interest must rise to achieve macroeconomic equilibrium, and conversely. The real component of changes in nominal output then can proxy for changes in the real rate of return on physical assets. Second, the behavior of inflation is such that the public extrapolates realized inflation in predicting future inflation. The nominal component of changes in nominal output then can proxy for the anticipated change in the dollar value of physical assets.

[4] Judd and Trehan (1992) contend that the long-term stability of M2 velocity is a "statistical artifact" due to the choice of a definition for M2 in 1980 designed to make M2 velocity stable. Their contention is inaccurate. Attempts to circumvent Regulation Q ceilings on interest rates that were kept low relative to market rates led in the 1970s to the appearance of new financial instruments, especially money market mutual funds and NOW accounts. These new instruments necessitated a redefinition of the monetary aggregates. The Board staff did attempt to determine the relative stability of the public's demand for money using alternative definitions of the monetary aggregates, but it was unsuccessful. At the time of the redefinition of the aggregates, the new instruments had been introduced so recently and were still issued in such small amounts that their inclusion in a particular monetary aggregate did not affect econometric analysis of money demand. For example, money market mutual funds were insignificant until 1978 and NOW accounts were not introduced nationwide until 1981. In the end, M2 was constructed a priori to include M1 and savings instruments available in small denominations and basically redeemable at par.

There is, however, a problem in the definition of M2. M2 includes time deposits of less than $100,000. As prices rise, the real value of the cutoff falls. In order to prevent M2 velocity from rising as a consequence of the definition, the $100,000 cutoff should be indexed to the price level.

compared to changes in nominal expenditure.[5] This condition is reasonably well satisfied, despite the recent overprediction of real M2 (underprediction of M2 velocity). Money demand disturbances have not been the primary determinants of the rate of growth of nominal output in recent years. Taken alone, the M2 demand errors for 1989, 1990, and 1991 would have increased the rate of growth of nominal GDP. Instead, beginning in mid-1989, the trend rate of growth of nominal GDP fell from about 7 to 4 percent. The influence of disturbances in the demand for real M2 has been swamped by the reduction in the trend rate of growth of M2 that began in 1987.

III. An M2 Indicator Variable

This section draws on the results of Section II to construct two related measures of the impact of monetary policy actions on nominal expenditure. One, a marginal monetary indicator, measures the effect of contemporaneous policy actions on nominal expenditure. The other, an average monetary indicator, measures the cumulative effect of policy actions, contemporaneous and past, on nominal expenditure. These indicators are suggested by the quantity equation:

$$(1)\quad M \cdot V = Y,$$

where M is M2, V is M2 velocity, and Y is aggregate nominal expenditure (output or income). Equation (1) can be expressed in percentage change form (with continuous compounding) as (2):

$$(2)\quad \Delta m + \Delta v = \Delta y,$$

where Δ indicates a first difference and small letters indicate the natural logarithm of a variable (the change in the logarithm is a percentage change). Setting Δm equal to actual percentage changes in M2 and Δv equal to predicted percentage changes in M2 velocity makes (2) operational as an average monetary indicator variable, Δy^p.

A. An Average Monetary Indicator

The regression results reported in Table A2 of the appendix can give empirical content to Δv^p, predicted percentage changes in M2 velocity. Changes in velocity are a function of changes in the financial market opportunity cost, which is proxied for by the difference between the commercial paper rate (R) and the weighted average of the explicit interest rates paid on the components of M2 (RM2): (R − RM2). Changes in velocity due to changes in this opportunity cost variable are denoted by $\Delta v[\Delta(R - RM2)]$. Equation (2) then becomes

$$(3)\quad \Delta y^p = \Delta m + \Delta v[\Delta(R - RM2)].$$

A proxy for predicted velocity in (3) is constructed as a distributed lag of changes in the financial market opportunity cost variable with the estimated coefficients from the regression in the appendix Table A2 used as weights. The signs of the estimated coefficients reported in Table A2 change because the regression predicts changes in real M2 (the inverse of velocity) and the proxy predicts changes in velocity. Predicted changes in velocity due to changes in the financial market opportunity cost are proxied for by (4) before 1964.

$$(4)\quad \Delta v[\Delta(R - RM2)] = 2.47\ \Delta(R_t - RM2_t) + 2.50\ \Delta(R_{t-1} - RM2_{t-1}) + 1.65\ \Delta(R_{t-2} - RM2_{t-2})$$

Starting in 1964, (4) changes to (5) because of a reduction in the interest sensitivity of M2 demand.[6]

$$(5)\quad \Delta v[\Delta(R - RM2)] = 1.16\ \Delta(R_t - RM2_t) + 1.19\ \Delta(R_{t-1} - RM2_{t-1}) + .34\ \Delta(R_{t-2} - RM2_{t-2})$$

Figure 1 graphs actual annual percentage changes in nominal output and predicted percentage changes in nominal output, Δy^p, given by (3). In (3), Δm is annual percentage changes in M2, and Δv^p is given by (4) before 1964 and (5) thereafter. Predicted changes in nominal expenditure track actual changes in nominal output reasonably well over the period 1918 to 1991. The actual change in nominal output is underpredicted in 1991 by 1.2 percentage points.

[5] Equivalently, M2 offers useful information about the growth of nominal expenditure if unpredictable changes in the public's demand for real M2 are small relative to the sum of changes in M2 and of predictable changes in M2 velocity (real M2 demand).

[6] This reduction in interest sensitivity could reflect an increase in cyclical changes in short-term interest rates. The public began to adjust its money balances less in response to a change in interest rates because it anticipated the change would be reversed in time. Alternatively, the appearance of large negotiable CDs, which are not included in M2, could have drawn interest-sensitive balances out of M2.

The predicted series tracks the sharp fall in the rate of growth of nominal output (GDP) from 1988 to 1991.

B. A Marginal Monetary Indicator

The monetary indicator of nominal expenditure shown in Figure 1 measures the cumulative impact of Fed actions. In particular, the component of this indicator that predicts changes in M2 velocity depends upon the behavior of current and past market rates. This section proposes a monetary indicator that indicates how contemporaneous Fed actions affect the value of this cumulative measure.

The suggested marginal indicator is the difference between the rate of growth of nominal output (GDP) and the short-term rate of interest. As discussed in Section II, Friedman and Schwartz (1982) use the rate of growth of nominal output as a proxy for the market rate of return on physical assets. The short-term rate of interest is the traditional policy instrument of the Fed.[7] An unusually high value for the difference between the rate of growth of nominal output (GDP) and the short-term rate of interest, therefore, indicates that the rate of return on capital is high relative to market rates, and conversely. This proxy variable for a difference in interest rates measures a relative price, not a nominal (dollar) price. The Fed, therefore, cannot control it in a sustained way. It can, however, produce transitory increases by allowing monetary accelerations, and vice versa.

Figure 2 plots annual observations of the marginal indicator (solid line), that is, the difference between the rate of growth of nominal output and the commercial paper rate. It also plots *changes* in the average indicator Δy^P (shaded line), that is, changes in the predicted rate of growth of aggregate nominal expenditure. (The shaded line in Figure 2 shows first differences of the shaded line in Figure 1.) The positive correlation between the series shown in Figure 2 indicates that the Fed can increase temporarily the difference between the market rate of return on capital and the market rate of interest by allowing the rate of growth of aggregate nominal expenditure to increase, and conversely.[8]

Figure 3 displays quarterly observations of the two components of the marginal monetary indicator: the annualized rate of growth of nominal output and the short-term rate of interest. It also shows peaks and troughs of the business cycle. Figure 3 suggests that the Fed has raised the level of short-term rates relative to the rate of growth of nominal output over recovery phases of the business cycle until the thrust of monetary policy became restrictive. With a lag after the decline in the growth of nominal output, it then lowered the level of short-term rates until the thrust of monetary policy became expansionary.

Figure 3 shades in positive differences between the rate of growth of nominal output and the short-term interest rate. Until 1980, during periods of economic recovery, the rate of growth of nominal output exceeded the short-term interest rate. In the 1980s, the economy's underlying real rate of interest rose above its historical average. In the 1980s, therefore, a higher level of short-term rates (relative to nominal GDP growth) was required to maintain a given rate of growth of nominal expenditure.[9]

It is possible that in the 1990s the economy's real rate of interest has fallen back to its longer-run, lower level.[10] One possible explanation for the recent weakness in the growth of nominal expenditure is that a fall in the economy's real rate of interest to

---

[7] The rate of interest in the money market is largely determined by the level of the funds rate, which since the early 1970s the Fed has either targeted directly or indirectly through setting the discount rate and the level of borrowed reserves. Before the 1970s, the Fed used the combination of the discount rate and free reserves (excess reserves minus borrowed reserves) to target the level of money market rates.

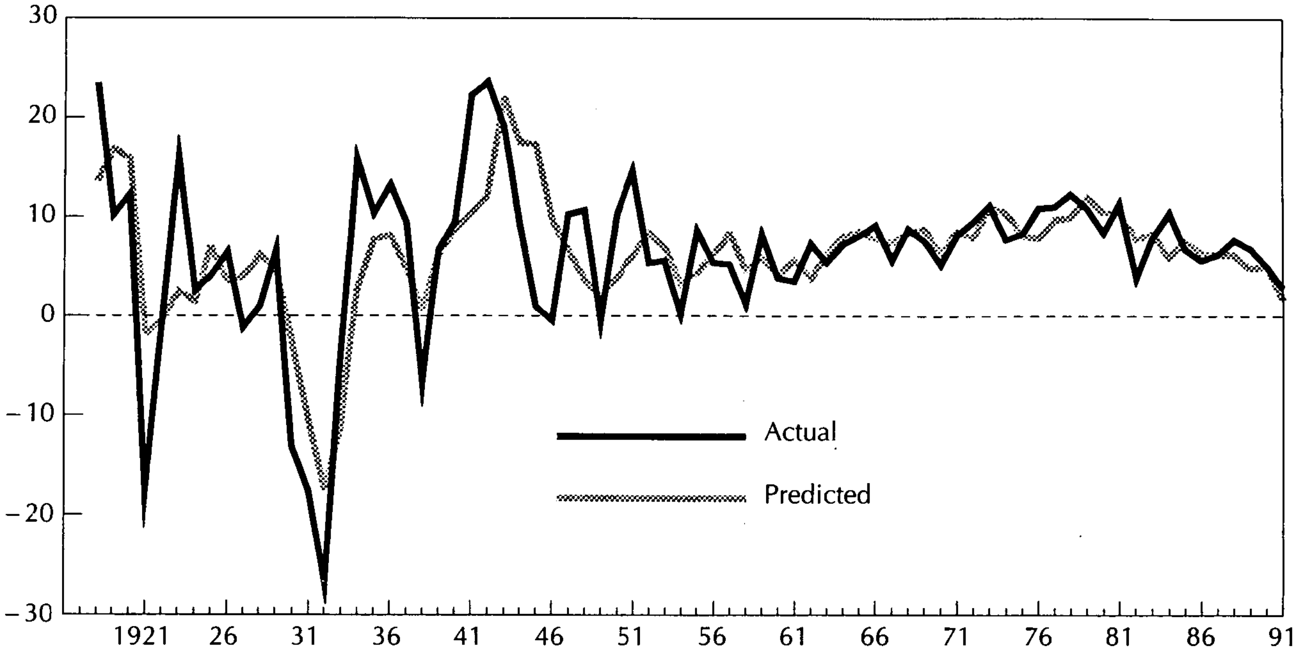
[8] For the period 1950 to 1979, there is a positive correlation like that shown in Figure 2 between the difference in the rate of growth of nominal output and the commercial paper rate and changes in the rate of growth of M1. M1 is a particularly interesting monetary aggregate over the period 1950 to 1979. Because market rates were relatively high and demand deposits could not pay explicit interest, individuals used M1 primarily as a transactions vehicle. For this reason, the interest sensitivity of real M1 demand was low. As a consequence, quantity changes in M1 served as a good proxy for the effect of monetary policy actions on the rate of growth of nominal expenditure. Unlike M2, to use M1 as an indicator for this period, it is not necessary to adjust for velocity changes due to changes in interest rates.

[9] The difference between GNP growth and the commercial paper rate was 3.3 from 1951 to 1960, 2.0 from 1961 to 1970, 2.7 from 1971 to 1980, but −1.7 from 1981 to 1990.

[10] The merchandise trade deficit provides indirect evidence. It averaged about .5 percent of GDP in the 1970s. It climbed sharply in the 1980s to a level of 3.6 percent of GDP in 1986. It began to fall after 1987 and was about 1.5 percent of GDP in 1991. The trade deficit is the mirror image of capital inflows. The high real rate of return to capital in the United States in the 1980s produced capital inflows that appeared as a trade deficit. The reduction in the trade deficit and the associated reduction in capital inflows suggests that the real rate of interest in the United States is returning to a lower, more normal level.

Figure 1

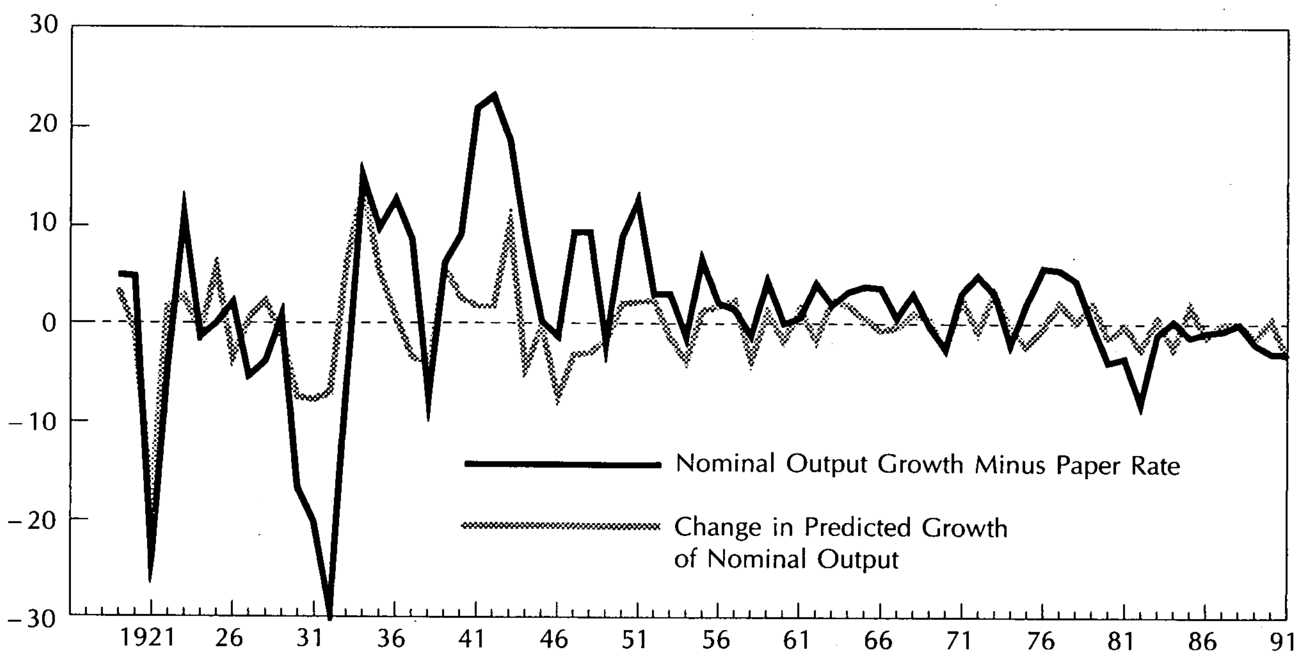
ACTUAL AND PREDICTED NOMINAL OUTPUT GROWTH



Notes: Predictions of nominal output growth are from the M2 indicator variable $\Delta m + \Delta v^p$, where Δm is the percentage growth in M2 and Δv^p is the predicted percentage growth in M2 velocity due to changes in the financial market opportunity cost of holding M2. Actual nominal output growth is the percentage change in GNP before 1959 and GDP thereafter.

Figure 2

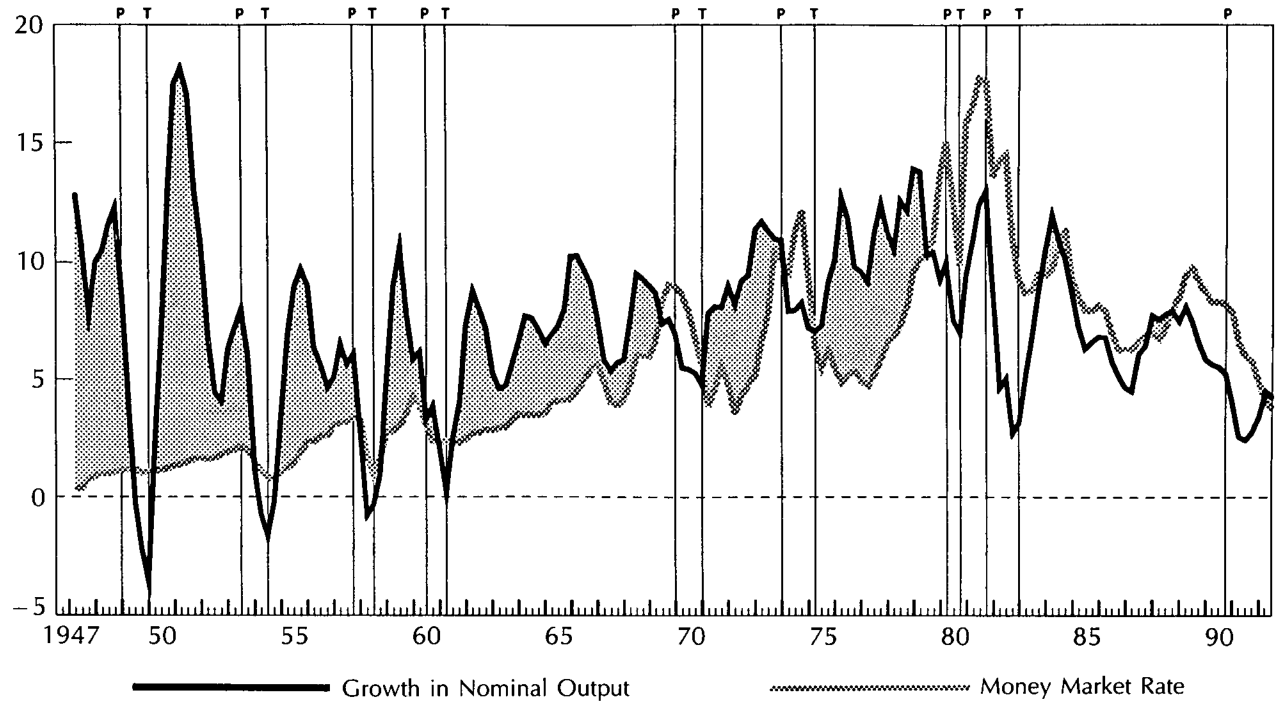
DIFFERENCE BETWEEN RATES OF RETURN ON PHYSICAL AND FINANCIAL ASSETS; CHANGES IN PREDICTED OUTPUT GROWTH



Notes: The solid line is the difference between nominal output growth (GNP before 1959 and GDP thereafter) and the four- to six-month commercial paper rate. The shaded line is the change in predicted growth of aggregate nominal output. That is, it is first differences of the sum of the percentage growth in M2 and the predicted percentage growth in velocity (first differences of the shaded line in Figure 1).

Figure 3

MONEY MARKET RATE AND NOMINAL OUTPUT GROWTH



Notes: Nominal output growth is quarterly observations of four-quarter rates of growth of nominal output (GNP before 1959 and GDP thereafter). Money market rate is the three-month Treasury bill rate for 1947-1963 and the funds rate thereafter. The graph shades in the positive differences in these two series. *Ts* mark business cycle troughs and *Ps* peaks. Heavy tick marks indicate last quarter of year.

a more normal historical level has made it difficult for the Fed to find the level of short-term market rates consistent with this rate. The funds rates in the 3 to 4 percent range that prevailed in 1991 and 1992 seemed low relative to the funds rate peak of almost 10 percent in 1989. Figure 3, however, suggests that these funds rates were low only relative to the unusually high rates of the 1980s. As shown in Figure 3, relative to the rate of growth of nominal GDP in business cycle recoveries before the 1980s, the funds rate has not been low in the current recovery.

C. Inverse Movements in M2 and M2 Velocity

M2 is not widely used as an indicator of the impact of monetary policy actions on the growth of nominal expenditure. The reason may be the low contemporaneous correlation between the rates of growth of M2 and nominal expenditure.[11] The reason for this low contemporaneous correlation is that movements in interest rates initially produce inverse movements in M2 and its velocity.

This inverse relationship is produced by the inertia in the rates paid on many of the deposits in M2 relative to money market rates. Until June 1978, with the issuance of money market certificates by S&Ls, all the deposits in M2 were either subject to Reg Q ceilings or to the outright prohibition of interest payments. Even with the complete phase-out of Reg Q in 1986, banks continue to vary the rates paid on many of the components of M2 (NOWs, MMDAs, and savings deposits) sluggishly. As a consequence, when market rates rise, the cost of holding M2 rises, and depositors move out of M2 into other financial instruments like large CDs. Although M2 growth falls, M2 velocity growth rises because M2 has become more costly to hold. As a consequence, a macroeconomic shock that causes

[11] For example, from first quarter 1965 through second quarter 1992, the correlation between quarterly growth rates of M2 and nominal GDP was .31. This correlation, however, mostly reflected a common trend. When the growth rates are differenced to remove trend, the correlation between M2 and GDP falls to .044. There is almost no contemporaneous relationship between changes in the growth rates of M2 and nominal GDP.

expenditure and market rates to rise is initially associated with a decline in M2 growth, and conversely.[12] Casual observation then suggests that M2 offers little information about the behavior of expenditure.

Figure 4 shows annual observations of rates of growth of M2 and the financial market opportunity cost of holding M2 (the commercial paper rate minus the own rate of return on M2). There is an inverse cyclical relationship between the rate of growth of M2 and the cost of holding M2. Consequently, there is an inverse cyclical relationship between M2 growth and M2 velocity growth. This inverse relationship means that often the contemporaneous behavior of M2 does not give good signals about the contemporaneous rate of growth of nominal output. More generally, cyclical movements in nominal expenditure are largely accounted for by cyclical movements in M2 velocity rather than in M2.

This pattern can be seen in recent years. In 1987, market rates rose absolutely and relative to the rates paid on M2 components like NOWs, savings deposits, and MMDAs; consequently, the rate of growth of M2 fell. This fall, however, was more than offset by a rise in M2 velocity produced by the increased cost of holding real M2. In 1987, therefore, the rate of growth of M2 fell, even though the rate of growth of nominal GDP rose. These inverse movements in M2 and in its velocity, however, are transitory. Sustained changes in the rate of growth of M2 ultimately produce changes in the rate of growth of nominal output. The financial market opportunity cost of holding M2 stopped rising in 1989 and began to fall. In the absence of rising velocity, low M2 growth then began to show through to weakness in the growth of nominal output.

IV. Arguments That M2 Demand Will Be Unstable

In the 1950s, as in the present, many economists argued that the growing importance of nonbank financial intermediation would make money demand unstable. Similar predictions of instability in the demand for money were made in the early 1960s with the appearance of credit cards, in the late 1960s with the emergence of the Eurodollar market, in the mid-1970s with new cash management techniques, and in the 1980s with securitization. The long-term stability of M2 velocity has contradicted these predictions. At present, however, the overprediction of real M2 pointed out in Section II has revived such fears. This section examines five arguments made recently suggesting that M2 demand will be unstable in the future.

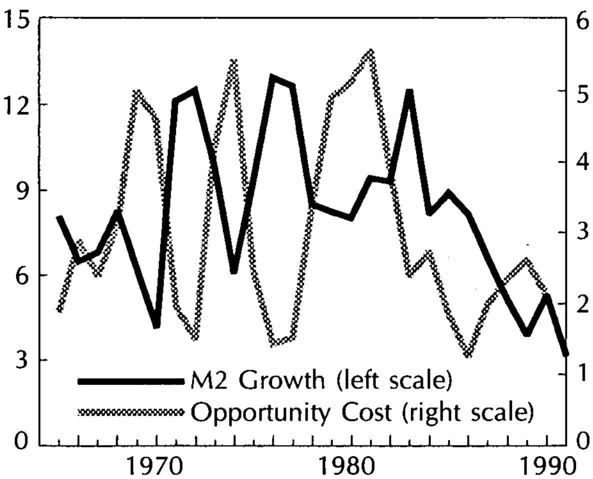
A. Bond Funds

The current weakness in real M2 growth is often attributed to a shift of deposits out of M2 into bond funds prompted by a sharply rising yield curve. It is uncertain, however, whether the magnitude of such transfers is sufficient to explain much of the weakness in real M2. It is true that in 1992 the yield curve has been unusually steep. Weakness in real M2 growth, however, developed before the appearance of a yield spread large by the standards of the

[12] It follows that strength in economic activity is initially associated with a reduction in M2 growth and weakness in economic activity is initially associated with an increase in M2 growth. M2 targeting then would appear to conflict with lean-against-the-wind procedures that call for a rise in the funds rate when economic activity strengthens and a fall when economic activity weakens. This conflict is probably one of the reasons for the relative insignificance of M2 in popular discussions of monetary policy. A substantive target for M2 would provide for a short-term negative elasticity of supply with respect to market rates, but would eliminate long-term base drift in light of the stability of M2 velocity.

Figure 4

M2 GROWTH AND FINANCIAL MARKET OPPORTUNITY COST OF HOLDING M2



Notes: Annual observations of percentage change in M2. The financial market opportunity cost of holding M2 is the difference between the four- to six-month commercial paper rate and a weighted average of the explicit rates of return paid on the components of M2.

1980s.[13] Also, the 5 percentage point change in the yield spread from a −2 in the early 1980s to a +3 in 1985 did not destabilize M2 demand. More generally, over the post-World War II period, the demand for real M2 has not been significantly affected by the shape of the yield curve.

Also, the previous experience with strong growth in bond funds did not weaken real M2 demand. Bond funds increased about $250 billion from 1985 to early 1987. (A strong rally in the bond market made bonds attractive during this period. The 30-year bond rate fell from 11.4 percent in July 1984 to 7.4 percent in September 1986, a decline of 4 percentage points.) In 1985 and 1986, however, M2 grew rapidly at about an 8 percent annualized rate.

If all of the assets of bond funds were included in M2, this augmented monetary aggregate would still have grown only moderately recently. For example, from fourth quarter 1990 through fourth quarter 1991, M2 grew at 2.9 percent while M2 plus bond funds grew at 5.6 percent. It is, however, unlikely that all of the growth in bond funds came at the expense of M2 deposits. It is not plausible that individuals view the deposits in M2 as highly substitutable with bond funds. The value of assets in M2 is not subject to fluctuation as market rates change, while the value of bond funds is. Furthermore, those bond funds that could be defended as substitutes for M2, namely, short-term bond funds, have hardly grown. The amount of money in bond funds with bonds of maturity five years or less, about $20 billion at the end of 1991, is small compared to the amount of M2, $3,438.9 billion in December 1991.[14]

B. Unwinding Debt with M2

Some economists have argued that weakness in real M2 growth is due to the repayment of consumer debt. They argue that individuals experienced an adverse wealth shock in the late 1980s that has made them want to hold less debt. The ratio of consumer debt to household net worth rose from about 15 percent in the 1970s to a peak of 21 percent in 1991. (Consumer debt comprises primarily installment credit and mortgages. Household net worth is the difference between the assets and liabilities of households.) According to the argument, consumers are now reducing their debt by drawing down deposits in M2.

Figure 5 shows real household net worth (household net worth deflated by the CPI). Although by this measure the increase in the public's wealth slowed in the late 1980s, previous recessions also exhibited such slowdowns. The recent behavior of wealth does not suggest anything unusual about the last recession. Some commentators have referred to a decline in the value of the housing stock. As measured by the index constructed by the National Association of Realtors (median sales price of existing single-family homes), the sales price of existing homes did fall in 1990, after having risen in 1988 and 1989 at a rate of about 5 percent. In 1991, however, home prices rose at about an 8 percent rate.

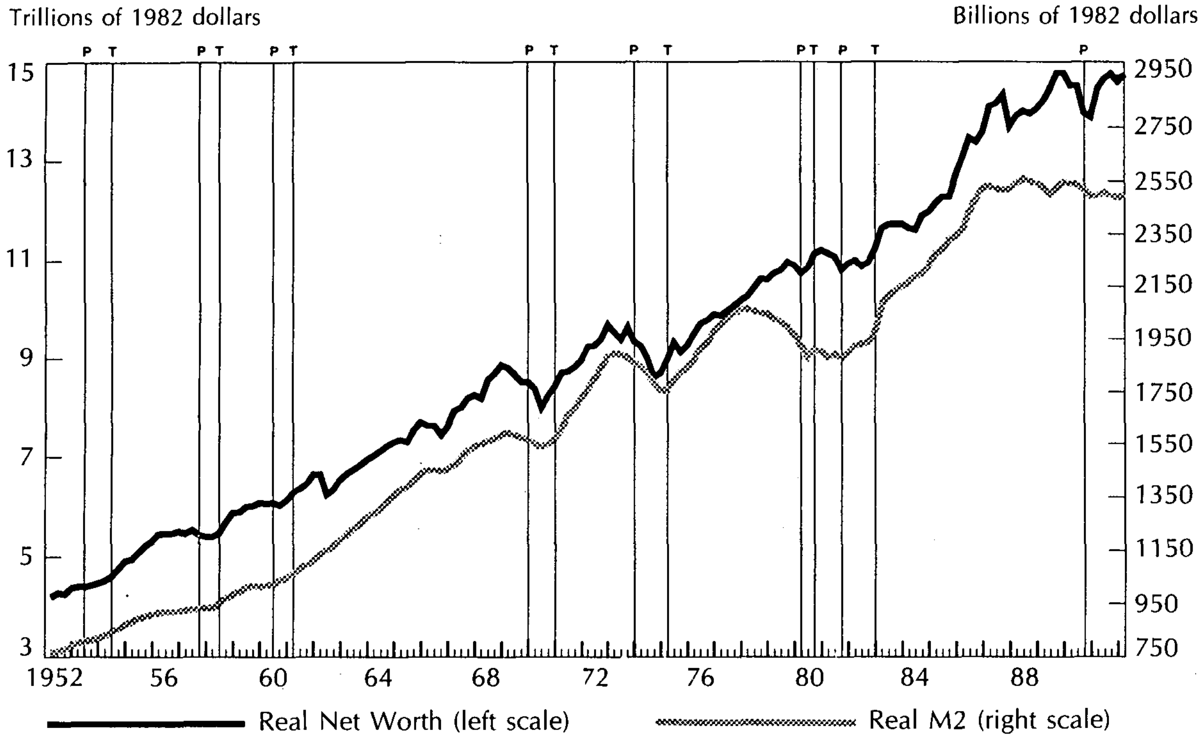
Figure 6 shows the behavior of household debt over recent business cycles (Schreft and Owens, 1991). Household debt (deflated by the CPI) is put into the form of a cycle-relative index for each business cycle by dividing quarterly debt figures by the value of debt six quarters preceding the cycle peak. Figure 6 shows that in the recent cycle consumer debt did rise prior to the cycle peak. At least as of first quarter 1992, however, it has not fallen since the cycle peak as predicted by the debt-unwinding hypothesis. (In the recession that began in fourth quarter 1973, real household debt did fall, but the demand for M2 was not rendered unpredictable.)

The appeal of the debt-unwinding hypothesis may derive in part from a natural tendency to generalize about collective behavior on the basis of individual behavior. An individual who lowers his debt will draw on savings and reduce consumption. It therefore appears plausible to explain both the current weakness in real M2 growth and in real expenditure by an excessive debt level. However, what is true for the individual is not necessarily true for individuals collectively. One person's debt is another person's asset. If debts are high, so are assets. In the aggregate, the level of debt does not affect the level of wealth. Economic theory says that consumers will proportion their holdings of M2 to their total financial wealth, which in the aggregate is not affected by debt creation. The ratio of household net financial

[13] As measured by the difference between the 30-year Treasury bond rate and the six-month commercial paper rate, the yield spread averaged about 2 percentage points from first quarter 1983 to second quarter 1988. After becoming relatively flat in 1989, it began to rise again and reached 2 percentage points again in the middle of 1991. It then rose to about 4 percentage points in third quarter 1992.

[14] The figures on bond funds are from the Investment Company Institute. The figures on short-term bond funds were kindly assembled by Anne Schafer at the Investment Company Institute from individual fund data from Lipper Analytical Securities.

Figure 5
REAL M2 AND REAL HOUSEHOLD NET WORTH



Notes: M2 and household net worth are deflated by the CPI and are in 1982 dollars. Household net worth is from the Federal Reserve Board of Governors, Flow of Funds Accounts, "Balance Sheets for the U.S. Economy." *T*s mark business cycle troughs and *P*s peaks. Heavy tick marks indicate last quarter of year.

wealth to disposable personal income has grown moderately ever since the mid-1970s.[15] It has not exhibited any drops over the last several years that could have caused a reduction in the public's demand for M2.

Similarly, it does not follow that aggregate expenditure will fall when an individual consumes less to reduce his debt. Nothing has changed to cause that individual to work less; he may even work harder. He will save more. In the aggregate, consumption will fall, but saving and investment will increase. The increase in investment will maintain the level of aggregate expenditure.

The behavior of the savings rate contradicts the implication of the debt-unwinding hypothesis that the savings rate should be unusually high. As measured by the National Income and Product Accounts, the savings rate has not risen but has remained around a relatively low level of 5 percent.

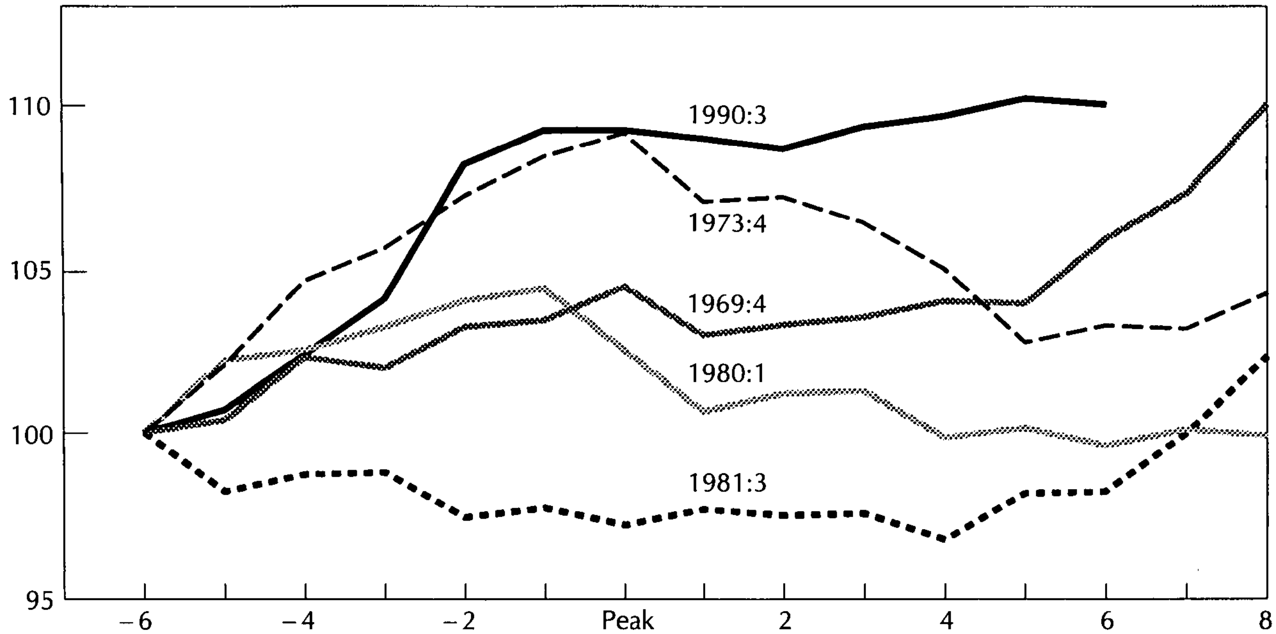
If individuals have experienced an adverse wealth shock, they would want to rebuild their wealth by saving more. Their demand for M2, which is a component of wealth, should increase, not decrease. It has, however, been argued that consumers are using M2 balances to draw down consumer installment debt because the return paid on M2 balances has fallen relative to the cost of installment credit. In particular, the rate paid on a three-month bank CD has fallen from a peak of somewhat more than 10 percent in March 1989 to 3.3 percent in August 1992, while the cost of using a credit card has often remained around 18 percent. This argument, however, assumes that the same individuals hold bank CDs and credit cards. Even when CD rates were at their peak, it is hard to understand why the same individual would borrow at 18 percent while lending at 10 percent.[16]

[15] Household net financial wealth is the difference between "Total Financial Assets" and "Total Liabilities" for households in the table "Financial Assets and Liabilities, Outstandings," from the Federal Reserve Board of Governors Flow of Funds Accounts. Disposable personal income is from the National Income and Product Accounts.

[16] Robert Laurent made this point in personal correspondence.

Figure 6

THE CYCLICAL BEHAVIOR OF REAL HOUSEHOLD DEBT



Notes: Credit market debt owed by households is deflated by the CPI. Debt is put into cycle-relative form by dividing by the value of debt six quarters prior to the cycle peak. Dates by lines indicate the particular business cycle peak. Household debt is from the table "Credit Market Debt Owed by Nonfinancial Sectors, Households," from Federal Reserve Board of Governors, Flow of Funds Accounts, "Financial Assets and Liabilities, Outstandings."

C. The Shrinking Thrift Industry

Some economists have argued that closings of thrifts by the Resolution Trust Corporation (RTC) begun in 1989 have produced slow real M2 growth (Duca, 1992; Kasriel, 1991). Actually, the ratio of thrift deposits to total M2 declined more sharply over the period 1979 through 1982 (about 7.5 percentage points) than over the period 1989 through 1992 (about 5.5 percentage points). The earlier runoff in thrift deposits was not, however, associated with an unpredictable reduction in the public's demand for real M2.

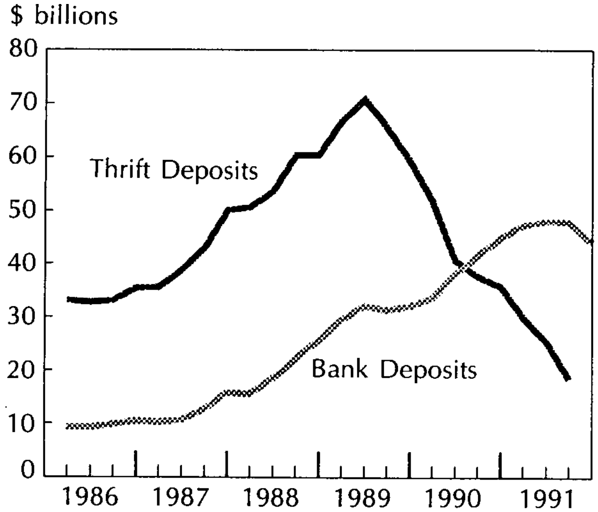
Closing a thrift does not directly affect the money stock. At an aggregate level, closing an insolvent thrift involves replacing a bad asset (a real estate loan in default) on the books of financial intermediaries with a good asset (a Treasury bill). This transaction involves a wealth transfer from taxpayers to thrift depositors. It does not, however, reduce the total assets of financial intermediaries and, therefore, need not affect total deposits.

There may, however, be an indirect effect on the money stock. Because the NOW accounts of a failed thrift are simply transferred to the acquiring institution, these deposits are not lost to M2. When the RTC closes a thrift, however, it may retain some of the thrift's assets. It will fund these assets with government debt, rather than with the high-yielding brokered deposits formerly used by the thrift. The former holders of these brokered deposits may then move into government debt. In this case, the decline in brokered deposits measures the decline in M2.

Figure 7 shows the brokered deposits of thrifts and commercial banks included in M2. Over the period of RTC closures, the decrease in brokered deposits at thrifts minus the increase in these deposits at banks gives a rough estimate of the reduction in M2 that could have arisen from RTC actions. From second quarter 1989, which marked the peak in brokered deposits held by thrifts, to the fourth quarter of 1991, the combined holdings of thrifts and banks fell by $40.3 billion. This figure is small relative to M2. As of fourth quarter 1991, $40.3 billion was only 1.2 percent of M2. Finally, because of a lack of funds, the RTC stopped closing insolvent thrifts after March 1992. The absence of thrift closures, however, did not produce any revival in M2 growth.

Figure 7

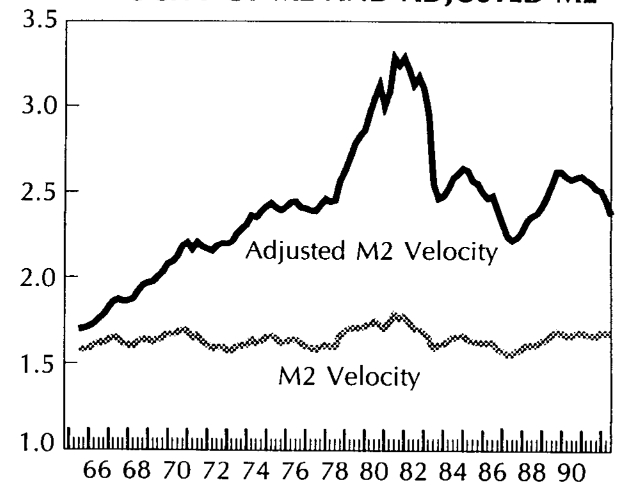
BROKERED DEPOSITS IN M2



Notes: The solid line is brokered deposits included in M2 held by S&Ls and SAIF-insured mutual savings banks. The shaded line is brokered deposits held by commercial banks and FDIC-insured mutual savings banks. Data are from call reports. Heavy tick marks indicate last quarter of year.

Figure 8

VELOCITY OF M2 AND ADJUSTED M2



Notes: Velocity is gross domestic product divided by M2. Adjusted M2 is M2 minus CDs of less than $100,000. Heavy tick marks indicate last quarter of year.

D. The Runoff in Small CDs

Much of the weakness in real M2 growth has been associated with the runoff of small retail CDs (CDs less than $100,000). Some economists have argued that small CDs are "a source of instability in the supply and demand for M2" (Wenninger and Partlan, 1992, p. 34; Citibank, 1992). The concentration of weakness in M2 growth in small CDs, however, does not in itself imply that the public's demand for M2 demand has declined. It is also consistent with a change in M2 from the supply side.

Assume, for example, that the central bank has kept the market rate of interest above the economy's equilibrium rate, so that banks are reducing their assets. As they reduce their assets, they will reduce their deposit liabilities in the least-cost way. Banks buy and sell CDs (large and small) in a spot market. In contrast, their other deposits generally involve a long-term customer relationship. The least-cost way for banks to reduce their liabilities is to let CDs run off by lowering the rate they pay on them.

Figure 8 shows velocity for M2, as well as for a revised M2 defined as M2 less small CDs. Velocity fluctuates less with the current definition of M2 than with a definition excluding small CDs. Money demand regressions using M2 minus small CDs also exhibit a significantly poorer fit than regressions using the current definition of M2.

E. Divergent Growth in M1 and M2

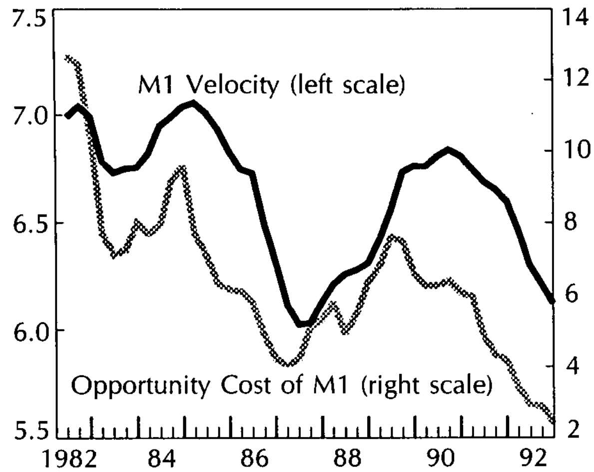
Over the two-year period August 1990 through August 1992, the annualized growth rates of M1 and M2 were, respectively, 9.2 percent and 2.3 percent. Some have argued that this divergence in growth rates indicates instability in the M2 demand function. There is, however, a ready explanation for this divergence. With the nationwide introduction of NOW accounts in 1981, real M1 demand became sensitive to market rates (Hetzel and Mehra, 1989). The recent strength in M1 growth reflects a fall in market rates that has decreased the cost of holding real M1 and increased its demand.

Figure 9 shows M1 velocity and the financial market opportunity cost of holding M1 (the difference between the commercial paper rate and a weighted average of the explicit rates of return paid on the components of M1). The graph starts in 1982 to avoid the distorting effects of the nationwide introduction of NOWs in 1981. As shown, M1 velocity is sensitive to interest rates. Over the 1980s, the fall in the cost of holding M1 has been associated with a fall in M1 velocity (a rise in real M1 demand). During the two periods when the cost of holding M1 rose, 1984 and 1987-1989, M1 velocity ceased falling.

Because banks reduce the rates paid on NOW accounts only with a lag as market rates fall, reductions in market rates make holding NOW accounts more attractive. Also, when market rates fall, corporations

Figure 9

M1 VELOCITY AND THE FINANCIAL MARKET OPPORTUNITY COST OF HOLDING M1



Notes: The financial market opportunity cost of holding M1 is the difference between the rate on four- to six-month commercial paper and a weighted average of the explicit rates of interest paid on the components of M1. Heavy tick marks indicate last quarter of year.

hold a higher level of demand deposits as compensating balances to reimburse banks for various services. Reductions in market rates then increase the demand for M1. When market rates fell beginning in the summer of 1984, M1 growth surged. M1 growth reached 12 percent and 16 percent in 1985 and 1986, respectively. These rates of growth of M1 did not raise the inflation rate because they accommodated an increased demand for M1. Similarly, at present, high M1 growth rates are accommodating an increased demand for M1 produced by the fall in market rates.

Increased M1 growth in turn leads to an increased demand for reserves because of the 10 percent reserve requirement imposed on demand deposits and NOW accounts. At the prevailing funds rate, the Fed accommodates the increased demand for reserves and the rate of growth of bank reserves and the monetary base increases. Higher growth rates of bank reserves and the base, however, do not in themselves indicate that monetary policy actions are expansionary.

V. Concluding Comment

Forecasters have had more than the usual problems in recent years. For example, in its lead-off section entitled "End of Recession Has Arrived on Schedule," the July 10, 1991, *Blue Chip Economic Indicators* (1991) reported consensus forecasts for third and fourth quarter 1991 growth in real GNP of 2.7 percent and 2.9 percent, respectively. The actual growth rates, however, were significantly lower (1.0 and .4 percent, respectively). The forecasters who contributed to these consensus forecasts also ranked as the second most important factor in promoting economic growth "easier monetary policy resulting from more accommodative action by the Federal Reserve," that is, reductions in the funds rate. It now appears that most forecasters were again too optimistic in the spring of 1992 in forecasting growth over the last part of 1992. This article suggests that forecasters would have done better by using the information contained in the behavior of M2.

This article has proposed two related indicators of the impact of monetary policy actions on growth of aggregate nominal expenditure. One, an average indicator, measures the combined impact of the rate of growth of M2 and the rate of growth of M2 velocity produced by contemporaneous and past changes in the cost of holding M2. The other, a marginal indicator, measures the impact of contemporaneous policy actions on this average indicator. The marginal indicator is the difference between the rate of growth of nominal output (a proxy for the rate of return on physical assets) and a short-term interest rate. A large value for this indicator is associated with increases in the rate of growth of aggregate nominal expenditure predicted by the average indicator, and conversely.

Over the last two years, the rates of growth of M2 and nominal GDP have corresponded fairly closely. From second quarter 1990 through second quarter 1992, nominal GDP and M2, respectively, grew at annualized rates of 3.3 percent and 2.7 percent. Given the reduction in the cost of holding M2 due to the fall in interest rates over this period, however, the rate of growth of M2 should have exceeded the rate of growth of nominal GDP. In this sense, the public's demand for real M2 has been unpredictable. Whether M2 conveys useful information about the nominal expenditure of the public, however, depends on the magnitude of unpredictable changes in the demand for real M2 relative to the magnitude of changes in the other determinants of nominal expenditure—changes in nominal M2 and predictable changes in M2 velocity. The regression analysis of Section II indicates that recent unpredictable changes in the public's demand for real M2 have been small relative to these other determinants. In particular, the reduction in the growth rate of nominal expenditure reflects the reduction in the growth rate of M2 rather than an unpredictable increase in M2 velocity.

The relationship between money and nominal output is predictable only over fairly long periods of time. Consequently, inferences about the contemporaneous behavior of money demand are always problematic. For this reason, Section IV examined the plausibility of various reasons advanced for believing that M2 demand is behaving unpredictably at present. Section IV examined the effects on real M2 demand of bond funds, variability in the public's demand for debt, the reduction in the size of the thrift industry, the reduction in bank holdings of small CDs, and divergent growth rates of M1 and M2. None of these phenomena will clearly destabilize real M2 demand. It appears likely that the behavior of M2 will continue to offer useful information about the public's nominal expenditure and output.

APPENDIX

One way to appraise the stability of the public's demand for real M2 is to observe the size of the errors of an M2 demand regression. The regression used here (1) is similar to the one in Friedman and Schwartz (1982). It is also interesting in the present context because its use of percentage changes in nominal output as a regressor measuring the market rate of return on physical assets lends credence to the use of this variable as a component of the marginal indicator variable advanced in Section III.

$$(1)\quad \ln (M2/P)_t = c_0 + c_1 \ln (GDP/P)_t - c_2(R_t - RM2_t) - c_3 \Delta \ln GDP_t + e_t,$$

where M2 is per capita M2; P is the implicit price deflator for GDP (GNP before 1959); GDP is per capita gross domestic product (GNP before 1959); R is the four- to six-month commercial paper rate and RM2 is a weighted average of the own rates of return paid on components of M2. The error term is e. The natural logarithm of a number is indicated by ln and Δ indicates first differences.

An examination of observations of M2 velocity and (R – RM2), the financial market opportunity cost of holding real M2, suggests a reduction in the interest elasticity of real M2 demand after 1963. The large cycles in the cost of holding M2 that began in the mid-1960s induced relatively moderate changes in M2 velocity relative to the earlier period. For this reason, (1) was estimated with a shift dummy on the financial market opportunity cost variable, with the dummy assuming the value one from 1964 through 1991 and zero otherwise.

Tables A1 and A2 exhibit regression equation (1) estimated using annual observations, respectively, in levels and first differences over the period 1915 to 1991. The specification differs from that of Friedman and Schwartz (1982) in two respects. They assume that M2 pays a market rate of return apart from the fraction held in the form of noninterest-bearing base money, H. As a consequence, they use as their opportunity cost variable, R(1 – H/M2). That is, they assume that banks have evaded completely both the prohibition of payment of interest on demand deposits and Regulation Q ceilings. Equation (1) employs instead (R – RM2), which incorporates the assumption that these restrictions were binding. Second, equation (1) omits the dummy variables Friedman and Schwartz use to capture money demand shifts during the Depression and World War II and after World Wars I and II. It adds, however, a shift dummy to capture a reduction in the interest elasticity of real M2 demand beginning in the mid-1960s.

Friedman and Schwartz (1982) use data averaged over phases (contraction or expansion) of the business cycle, while the regressions here are estimated with annual data. Their first observation is for the years 1867 to 1869, while the first observation used here is for the year 1915. The data necessary to estimate the own rate on M2 (RM2), which are used to construct the financial market opportunity cost variable, only become available in 1915. It is necessary to use annual observations because this variable can be constructed quarterly only beginning in the first quarter of 1946.

The parameter values yielded by estimation in level form and in first-differenced form are comparable. Granger and Newbold (1974) point out that regression equations like the one in Table A1 that possess a nonstationary dependent variable and serially correlated errors (as evidenced by a low Durbin-Watson statistic) can yield misleading inferences. After their work, money demand regressions were generally estimated in first-differenced form. First differencing, however, results in a loss of information in the data. For these reasons, recent work has used error-correction models that combine estimation in levels and first differences. [See Engle and Granger (1987),

Table A1

M2 Demand Regression, 1915 to 1991

$$\ln rM2_t = \underset{(.2)}{4.6} + \underset{(46.0)}{.95} \ln rGDP_t - \underset{(12.7)}{7.4} (R_t - RM2_t) - \underset{(7.0)}{.54} \Delta \ln GDP_t + \hat{e}_t$$

Dummy on $(R_t - RM2_t)$ = 5.1 (7.2)

CRSQ = .98 SEE = 5.5 DW = .98 DF = 72

Notes: rM2 is per capita M2 deflated by the implicit GNP deflator before 1959 and by the GDP deflator thereafter; rGDP is real per capita gross national product before 1959 and real per capita gross domestic product thereafter; R is the four- to six-month commercial paper rate expressed as a decimal; RM2 is a weighted average of the own rates of return paid on components of M2; and GDP is nominal gross national product before 1959 and gross domestic product thereafter. ln is the natural logarithm and Δ the first-difference operator. The zero-one multiplicative shift dummy on $(R_t - RM2_t)$ is one from 1964 to 1991 and zero otherwise.

CRSQ is the corrected R-squared; SEE the standard error of estimate; DW the Durbin-Watson statistic; and DF degrees of freedom. Absolute values of t-statistics are in parentheses. Estimation is by OLS. Before 1959, M2 is M4 in Table 1 of Friedman and Schwartz (1970). From 1915 to 1929, GNP is from Balke and Gordon (1989).

Table A2

M2 Demand Regression, First Differences, 1918 to 1991

$$\Delta \ln rM2_t = 1.0\, \Delta \ln rGDP_t - \underset{(6.9)}{6.6}\, \Delta(R_t - RM2_t) - \underset{(11.3)}{.95}\, \Delta^2 \ln GDP_t + \hat{\epsilon}_t$$

Dummy on $(R_t - RM2_t)$ = 3.9 (3.6)

CRSQ = .79 SEE = 2.4 DW = 1.4 DF = 66

Notes: Δ^2 is the second-difference operator. The sum of the estimated coefficients (and absolute value of its t-statistic) is shown. Sum of coefficients on rGDP constrained to sum to one. Estimated coefficients on the contemporaneous and lagged terms (absolute value of t-statistics in parentheses) are as follows:

lag	$\Delta \ln rGDP_t$	$\Delta(R_t - RM2_t)$	$\Delta^2 \ln GDP_t$
0	.83 (13.0)	−2.47 (6.2)	−.46 (8.2)
1	.17 (2.7)	−2.50 (5.9)	−.36 (9.4)
2		−1.65 (3.9)	−.13 (3.6)

The estimated coefficient on the multiplicative shift dummy on $\Delta(R_t - RM2_t)$ was constrained to assume the same value at each lag. Otherwise, see notes to Table A1.

Hendry and Ericsson (1991) and Mehra (1991).] The similarity of the parameter estimates of the regressions shown in Tables A1 and A2, which employ data respectively in levels and first differences, indicates on the one hand that use of nonstationary data is not biasing parameter estimates and on the other hand that differencing is not producing a significant loss of information.

The point estimate of the elasticity of demand for real M2 with respect to real income is .95 using data in levels. The estimate using differenced data was constrained to equal one in order to make the regression analysis conformable to the average indicator, where a 1 percent change in money is associated with a 1 percent change in nominal output. The point estimates of the semi-log slope of demand with respect to the financial market opportunity cost variable are, respectively, -7.4 and -6.6. (This parameter gives the percentage change in real M2 associated with a 1 percentage point change in the cost of holding real M2.) From 1964 on, this parameter is, respectively, estimated at -2.3 and -2.7. Finally, the point estimates of the elasticity of real M2 demand with respect to the market rate of return on physical assets are, respectively, $-.54$ and $-.95$.

REFERENCES

Balke, Nathan S. and Robert J. Gordon. "The Estimation of Prewar Gross National Product: Methodology and New Evidence," *Journal of Political Economy*, vol. 97 (February 1989), pp. 38-92.

Blue Chip Economic Indicators, July 10, 1991.

Citibank. "Small CDs: Chief Suspect in the M2 Mystery," *Economic Week*, June 22, 1992.

Duca, John V. "The Case of the Missing M2," Federal Reserve Bank of Dallas *Economic Review* (Second Quarter 1992), pp. 1-24.

Engle, R. F. and C. W. J. Granger. "Cointegration and Error Correction: Representation, Estimation and Testing," *Econometrica*, vol. 55 (March 1987), pp. 251-76.

Friedman, Milton and Anna J. Schwartz. *Monetary Statistics of the United States*. New York: National Bureau of Economic Research, 1970.

————. *Monetary Trends in the United States and the United Kingdom*. Chicago: The University of Chicago Press, 1982.

Granger, C. W. J. and Paul Newbold. "Spurious Regressions in Econometrics," *Journal of Econometrics*, vol. 2 (July 1974), pp. 111-20.

Hendry, David F. and Neil R. Ericsson. "An Econometric Analysis of U.K. Money Demand in *Monetary Trends in the United States and the United Kingdom* by Milton Friedman and Anna J. Schwartz," *The American Economic Review*, vol. 81 (March 1991), pp. 8-38.

Hetzel, Robert L. "M2 and Monetary Policy," Federal Reserve Bank of Richmond *Economic Review*, vol. 75 (September/October 1989), pp. 14-29.

———— and Yash P. Mehra. "The Behavior of Money Demand in the 1980s," *Journal of Money, Credit and Banking*, vol. 21 (November 1989), pp. 455-63.

Judd, John P. and Bharat Trehan. "Money, Credit and M2," *FRBSF Weekly Letter*, September 4, 1992.

Kasriel, Paul. "The Fed Can't See the Forest for the Trees," Northern Trust *Money Market Commentary*, October 24, 1991.

Mehra, Yash P. "An Error-Correction Model of U.S. M2 Demand," Federal Reserve Bank of Richmond *Economic Review*, vol. 77 (May/June 1991), pp. 3-12.

Schreft, Stacey and Ray Owens. "The Dangerous Debt-Level Debate," Federal Reserve Bank of Richmond Memo, October 29, 1991.

Wenninger, John and John Partlan. "Small Time Deposits and the Recent Weakness in M2," Federal Reserve Bank of New York *Quarterly Review*, vol. 17 (Spring 1992), pp. 21-35.

Has M2 Demand Become Unstable?

Yash P. Mehra

I. INTRODUCTION

An important issue underlying the current discussion of monetary policy is the interpretation of the recent weakness in the monetary aggregate, M2. Since about 1990, standard money demand regressions have overpredicted M2 growth. The dilemma for policymakers is to determine whether this shortfall in M2 growth has resulted from a shift in money demand or whether it indicates that the Federal Reserve has been supplying an inadequate amount of money to the economy.

A number of analysts contend that the size of the recent shortfall in M2 growth is large and unpredictable. They therefore conclude that the public's M2 demand function has shifted leftward.[1] Those who hold this view believe that M2 is no longer useful as an indicator variable for the thrust of monetary policy.

This paper presents the results of empirical tests of the stability of M2 demand over the period 1990Q1 to 1992Q2. Standard M2 demand regressions typically include a scale variable measured by real GDP and an opportunity cost defined as a short-term nominal rate minus the rate of return on M2 itself (the so-called own rate). The regressions presented here do indeed generate prediction errors in 1990, 1991, and 1992 that cumulate to an overprediction of M2 of about \$144 to \$149 billion (4.2 to 4.3 percent) by the second quarter of 1992. The Dufour test, which is a version of the Chow test, indicates that the prediction errors of this magnitude are not statistically significant. These test results are consistent with the hypothesis that the standard M2 demand regression is stable over the period 1990Q1 to 1992Q2.

Although the prediction errors are not large by the Dufour test, they have been consistently negative. This may indicate that some alternative factors not accounted for in standard M2 demand regressions have been depressing M2 growth in recent years. The appendix to this paper examines the role of a yield curve variable, namely, the long-term nominal interest rate minus the own rate on M2. This variable captures substitutions by households out of M2 into long-term financial assets. The empirical work shows that the yield curve variable is significant in a money demand regression that includes post-1989 data, but not pre-1989 data. Such a money demand regression can account for most of the "unexplained" weakness of M2 during the current period. This result is consistent with the hypothesis that M2 demand in recent years has been affected by portfolio substitutions. The hypothesis needs to be confirmed with more out-of-sample data and must therefore be considered tentative. In any event, the size of the current shortfall in M2 that can be attributed to these portfolio substitutions is not so large as to render irrelevant the short-run behavior of M2.

The plan of this article is as follows. Section II presents the error-correction model of M2 demand used here. Section III presents the empirical results. Concluding observations are given in Section IV. The appendix examines whether adding a yield curve variable to a standard money demand regression can account for the recent shortfall in M2 growth.

II. THE MODEL AND THE METHODOLOGY

An M2 Demand Model

The error-correction money demand model used here is reproduced below (Mehra, 1991 and 1992).

$$\ln(rM2)_t = a_0 + a_1 \ln(rY)_t + a_2 (R - RM2)_t + U_t \tag{1}$$

$$\Delta\ln(rM2)_t = b_0 + \sum_{s=1}^{n1} b_{1s} \Delta\ln(rM2)_{t-s} + \sum_{s=0}^{n2} b_{2s} \Delta\ln(rY)_{t-s} + \sum_{s=0}^{n3} b_{3s} \Delta(R - RM2)_{t-s} + \lambda U_{t-1} + \epsilon_t, \tag{2}$$

[1] See, for example, Carlson and Parrott (1991) and Duca (1992), who use the M2 demand model given in Small and Porter (1989) to demonstrate that M2 demand is seriously overpredicted in recent years. The M2 demand regression given in Small and Porter (1989) is based on an error-correction model of nominal M2 demand. The model includes a linear time trend and is estimated under the assumption that nominal M2 and GNP are cointegrated.

where rM2 is real M2 balances; rY real income; R a short-term nominal interest rate; RM2 the own rate on M2; U and ϵ the random disturbance terms. Δ is the first-difference operator and ln the natural logarithm. Equation 1 is the long-run equilibrium M2 demand function and is standard in the sense that the public's demand for real M2 balances depends upon a scale variable measured by real GNP and an opportunity cost variable measured as the differential between a short-term nominal rate of interest and the own rate of return on M2. The parameter a_1 measures the long-run income elasticity and a_2 the long-run opportunity cost parameter. Equation 2 is the short-run money demand equation, which is in a dynamic error-correction form. The parameter b_{is} ($i = 2, 3$) measures short-run responses of real M2 balances to changes in income and opportunity cost variables. The parameter λ is the error-correction coefficient. It is assumed that if the variables in (1) are nonstationary, they are cointegrated (Engle and Granger, 1987). Under this assumption, the parameter λ that appears on U_{t-1} in (2) is likely to be non-zero.

Estimating the Money Demand Model: Imposing the Convergence Condition

The long- and short-run money demand equations given above can be estimated jointly. This is shown in (3), which is obtained by solving for U_{t-1} in (1) and substituting in (2).

$$\begin{aligned}\Delta\ln(rM2)_t = {} & d_0 + \sum_{s=1}^{n1} b_{1s}\,\Delta\ln(rM2)_{t-s} \\ & + \sum_{s=0}^{n2} b_{2s}\,\Delta\ln(rY)_{t-s} \\ & + \sum_{s=0}^{n3} b_{3s}\,\Delta(R-RM2)_{t-s} \\ & + d_1 \ln(rM2)_{t-1} + d_2 \ln(rY)_{t-1} \\ & + d_3 (R-RM2)_{t-1} + \epsilon_t, \qquad (3)\end{aligned}$$

where $d_0 = (b_0 - a_0\lambda)$
$d_1 = \lambda$
$d_2 = -\lambda a_1$
$d_3 = -\lambda a_2$.

As can be seen, the long- and short-run parameters of the money demand model now appear in (3). The key parameters of (1) and (2) that pertain to income and opportunity cost variables can be recovered from (3).

The long-run income elasticity can be recovered from the long-run part of the money demand equation (3), i.e., a_1 is d_2 divided by d_1. The short-run part of (3) yields another estimate of the long-term income elasticity, i.e., $(\sum_{s=0}^{n2} b_{2s})/(1 - \sum_{s=1}^{n1} b_{1s})$.

If the same scale variable appears in the long- and short-run parts of the model, then a convergence condition can be imposed in equation (3) to ensure that one gets the same point-estimate of the long-run scale elasticity. To explain further, assume that the long-run income elasticity is unity, i.e., $a_1 = 1$ in (1). This assumption implies the following restriction on the long-run part of equation (3).

$$d_1 + d_2 = 0 \qquad (4)$$

Equation (4) says that coefficients that appear on $\ln(rY)_{t-1}$ and $\ln(rM2)_{t-1}$ in (3) sum to zero. The convergence condition implies another restriction (5) on the short-run part of equation (3).

$$\left(\sum_{s=0}^{n2} b_{2s}\right)/\left(1 - \sum_{s=1}^{n1} b_{1s}\right) = 1.0 \qquad (5)$$

Equivalently, (5) can be expressed as

$$\sum_{s=0}^{n2} b_{2s} + \sum_{s=1}^{n1} b_{1s} = 1.0.$$

That is, coefficients that appear on $\Delta\ln(rM)_{t-s}$ and $\Delta\ln(rY)_{t-s}$ in (3) sum to unity. This study examines whether the test results of stability are sensitive to the convergence condition imposed.

Data and Definition of Variables

The empirical work reported here uses quarterly data over the period 1953Q1 to 1992Q2. The variable rM2 is measured as nominal M2 deflated by the implicit GDP price deflator; rY by real GDP; R by the four- to six-month commercial paper rate; and RM2 by the weighted average of the explicit rates paid on the components of M2.

Real income appears as a scale variable in both the long- and short-run parts of the money demand regression (3). In contrast, the empirical work reported by Small and Porter (1989) uses consumer spending as the short-run scale variable and GNP as the long-run scale variable. They reason that some components of GNP, such as business fixed investment and changes in inventories, do not generate as much increase in money balances in the short run as does consumer expenditure. Equation (3) is alternatively estimated using real consumer spending as the short-run scale variable and real GNP as the long-run scale variable. Real consumer expenditure is hereafter denoted as rC.[2]

III. Empirical Results

Estimated Standard M2 Demand Regressions

Table 1 presents results of estimating the standard money demand regression (3) over the period 1953Q1 to 1989Q4. Regression A in Table 1 gives unrestricted estimates of the money demand regression, whereas regression B gives estimates that satisfy the convergence condition. That is, the regression satisfies the restrictions (4) and (5). The regressions reported in Table 1 use real GDP as the short- and the long-run scale variables, whereas the regressions reported in Table 2 use real consumer expenditure as the short-run scale variable and real GDP as the long-run scale variable.

The unrestricted estimates of the money demand regressions reported in Tables 1 and 2 indicate that the long-run GDP elasticity calculated from the long-run part of the model is unity (see regressions A and C in Tables 1 and 2). This result indicates that it

[2] All the data for the post-1959 period is from the Citibank data base with the exception of RM2. M2 for the pre-1959 period and RM2 are constructed as described in Hetzel (1989). Real GDP for the pre-1959 period are constructed by applying growth rates of real GNP to the real GDP series. Real consumption expenditure for the pre-1959 period are analogously constructed.

Table 1

Error-Correction Standard M2 Demand Regressions; 1953Q1 to 1989Q4
Real GDP in the Short- and Long-Run Parts of the Model

Regression A. Estimates without the Convergence Condition

$$\begin{aligned}\Delta\ln(rM2)_t = & \underset{(1.2)}{-.02} + \underset{(4.4)}{.31}\,\Delta\ln(rM2)_{t-1} + \underset{(1.9)}{.14}\,\Delta\ln(rM2)_{t-2} + \underset{(1.2)}{.07}\,\Delta\ln(rY)_t + \underset{(.9)}{.05}\,\Delta\ln(rY)_{t-1} - \underset{(4.6)}{.003}\,\Delta(R-RM2)_t \\ & - \underset{(5.1)}{.004}\,\Delta(R-RM2)_{t-1} - \underset{(2.1)}{.05}\ln(rM2)_{t-1} + \underset{(2.1)}{.05}\ln(rY)_{t-1} - \underset{(3.3)}{.002}\,(R-RM2)_{t-1} - \underset{(2.1)}{.012}\,CC1 \\ & - \underset{(0.0)}{.001}\,CC2 + \underset{(3.0)}{.020}\,D83Q1\end{aligned}$$

CRSQ = .64 SER = .00551 DW = 2.1 Q(36) = 25.4

$N_{rY} = 1.0$ $N_{(R-RM2)} = -.10$ [evaluated at the sample mean]

Regression B. Estimates with the Convergence Condition

$$\begin{aligned}\Delta\ln(rM2)_t = & \underset{(3.7)}{-.04} + \underset{(6.3)}{.43}\,\Delta\ln(rM2)_{t-1} + \underset{(3.5)}{.25}\,\Delta\ln(rM2)_{t-1} + \underset{(3.0)}{.17}\,\Delta\ln(rY)_t + \underset{(2.7)}{.15}\,\Delta\ln(rY)_{t-1} - \underset{(4.6)}{.003}\,\Delta(R-RM2)_t \\ & - \underset{(6.4)}{.005}\,\Delta(R-RM2)_{t-1} - \underset{(3.6)}{.08}\ln(rM2)_{t-1} + \underset{(3.6)}{.08}\ln(rY)_{t-1} - \underset{(1.56)}{.001}\,(R-RM2)_{t-1} - \underset{(2.2)}{.01}\,CC1 \\ & + \underset{(.2)}{.001}\,CC2 + \underset{(3.1)}{.02}\,D83Q1\end{aligned}$$

CRSQ = .58 SER = .00578 DW = 2.2 Q(36) = 31.7

$N_{rY} = 1.0$ $N_{(R-RM2)} = -.03$ [evaluated at the sample mean]

Notes: rM2 is real M2 balances; rY real GDP; R the four- to six-month commercial paper rate; RM2 the own rate on M2; ln the natural logarithm; and Δ the first-difference operator. CC1, CC2, and D83Q1 are, respectively, one in 1980Q2, 1980Q3 and 1983Q1 and zero otherwise. CRSQ is the corrected R-squared; SER the standard error of regression; DW the Durbin-Watson Statistic; Q(36) the Ljung-Box Q-statistic based on 36 autocorrelations of the residuals. The long-term income elasticity, N_{rY}, is given by the estimated coefficient on $\ln(rY)_{t-1}$ divided by the estimated coefficient on $\ln(rM2)_{t-1}$.

Table 2

Error-Correction Standard M2 Demand Regressions; 1953Q1 to 1989Q4 Real GDP as the Long-Run Scale Variable and Real Consumer Expenditure as the Short-Run Scale Variable

Regression C. Estimates without the Convergence Condition

$$\Delta\ln(rM2)_t = \underset{(1.8)}{-.04} + \underset{(4.3)}{.30}\,\Delta\ln(rM2)_{t-1} + \underset{(1.9)}{.14}\,\Delta\ln(rM2)_{t-2} + \underset{(2.1)}{.17}\,\Delta\ln(rC)_t + \underset{(2.5)}{.20}\,\Delta\ln(rC)_{t-1} - \underset{(4.9)}{.003}\,\Delta(R-RM2)_t$$

$$- \underset{(5.2)}{.004}\,\Delta(R-RM2)_{t-1} - \underset{(2.6)}{.06}\,\ln(rM2)_{t-1} + \underset{(2.6)}{.06}\,\ln(rY)_{t-1} - \underset{(3.0)}{.002}\,(R-RM2)_{t-1} - \underset{(1.7)}{.01}\,CC1$$

$$+ \underset{(.2)}{.001}\,CC2 + \underset{(3.4)}{.02}\,D83Q1$$

CRSQ = .66 SER = .00534 DW = 2.1 Q(36) = 23.6

$N_{rY} = 1.0$ $N_{(R-RM2)} = -.08$ [evaluated at the sample mean]

Regression D. Estimates with the Convergence Condition

$$\Delta\ln(rM2)_t = \underset{(3.0)}{-.03} + \underset{(4.9)}{.33}\,\Delta\ln(rM2)_{t-1} + \underset{(2.6)}{.17}\,\Delta\ln(rM2)_{t-2} + \underset{(3.5)}{.23}\,\Delta\ln(rC)_t + \underset{(3.8)}{.26}\,\Delta\ln(rC)_{t-1} - \underset{(4.8)}{.003}\,\Delta(R-RM2)_t$$

$$- \underset{(5.8)}{.004}\,\Delta(R-RM2)_{t-1} - \underset{(3.1)}{.06}\,\ln(rM2)_{t-1} + \underset{(3.1)}{.06}\,\ln(rY)_{t-1} - \underset{(2.5)}{.001}\,(R-RM2)_{t-1} + \underset{(1.5)}{.008}\,CC1$$

$$+ \underset{(.5)}{.003}\,CC2 + \underset{(3.4)}{.02}\,D83Q1$$

CRSQ = .66 SER = .00536 DW = 2.1 Q(36) = 23.6

$N_{rY} = N_{rC} = 1.0$ $N_{(R-RM2)} = -.02$ [evaluated at the sample mean]

Notes: See notes in Table 1. rC is real consumption expenditure.

is appropriate to impose the convergence condition if real GDP is also the short-run scale variable (see regression B in Table 1). The empirical results reported in Mankiw and Summers (1986) indicate that the long-run real consumption expenditure elasticity is not different from unity. Hence, the convergence condition is imposed even when real consumer expenditure is the short-run scale variable (see regression D in Table 2).

The estimated money demand regressions B and D look reasonable. The coefficients that appear on the scale and opportunity cost variables have theoretically correct signs and are statistically significant. The use of real consumption expenditure in the short-run part of the model does reduce somewhat the standard error of the regression, suggesting real consumption expenditure may be a better short-run scale variable than real GDP.

Evaluating Standard Money Demand Regressions

Is the actual behavior of real M2 balances over 1990Q1 to 1992Q2 consistent with stable M2 demand behavior? This question is investigated by using the Dufour test (Dufour, 1980), which is a variant of the Chow test. It uses an F-statistic to test the joint significance of dummy variables introduced for each observation of the interval for which structural stability is examined. A small F-statistic indicates structural stability.

The results of the Dufour test for the period 1990Q1 to 1992Q2 appear in Table 3. To carry out the test, the regressions in Table 1 and 2 were reestimated over the period 1953Q1 to 1992Q2 with separate shift dummies introduced for each quarter from 1990Q1 to 1992Q2. As can be seen, the

Table 3

Evidence on Stability in Standard M2 Demand Regressions over 1990Q1 to 1992Q2
Coefficients (t-values) on Dufour Dummies

Year/Quarter	Regression A	Regression B	Regression C	Regression D
1990Q1	-.006 (1.1)	-.005 (.9)	-.005 (.9)	-.003 (.7)
1990Q2	-.009 (1.6)	-.008 (1.3)	-.008 (1.4)	-.006 (1.3)
1990Q3	-.008 (1.5)	-.005 (.9)	-.007 (1.3)	-.005 (1.0)
1990Q4	-.009 (1.3)	-.003 (.4)	-.006 (1.0)	-.003 (.6)
1991Q1	-.009 (1.5)	-.003 (.5)	-.005 (.1)	-.001 (.3)
1991Q2	-.005 (.9)	-.005 (.1)	-.003 (.5)	-.003 (.1)
1991Q3	-.012 (2.1)	-.008 (1.4)	-.011 (2.0)	-.009 (1.7)
1991Q4	-.007 (1.1)	-.002 (.4)	-.005 (.9)	-.002 (.4)
1992Q1	-.008 (1.3)	-.003 (.5)	-.007 (1.2)	-.004 (.9)
1992Q2	-.018 (3.0)	-.014 (2.4)	-.017 (3.0)	-.015 (3.0)
FD(10,137)		1.06		1.50
FD(10,135)	1.66		1.45	

Notes: The regression equations A, B, C, and D above correspond, respectively, to regressions reported in Tables 1 and 2. These regressions are reestimated including Dufour dummy variables over the period 1953Q1 to 1992Q2. Dufour dummies are zero-one dummy variables defined for each observation over 1990Q1 to 1992Q2. FD is the F-statistic that tests the null hypothesis that all Dufour dummies are not significant as a group. The degrees of freedom for the F-statistics are in parentheses.

individual coefficients that appear on the shift dummies are generally not statistically significant with the exception of the one for the second quarter of 1992. FD is the F-statistic that tests the null hypothesis that these shift dummies are not significant as a group. These F-statistics are small (the 5 percent critical value is 1.9) and thus indicate that the standard M2 demand regression is stable. The stability result is not sensitive to the short-run scale variable used or to whether the convergence condition is imposed or not. (The conventional Chow test with the shift point located at or before 1990Q1 also indicates that the M2 demand regression is stable.)[3]

The coefficients that appear on the Dufour dummies measure (static) errors that occur in predicting real M2 balances over the period 1990Q1 to 1992Q2. As can be seen, these prediction errors, though small, are consistently negative, suggesting that the standard money demand regression used here consistently overpredicts real M2 balances over this period. In order to provide a different insight into the magnitude of the prediction error, Table 4 presents static simulations of M2 growth conditional on actual values of scale and opportunity cost variables. The predicted values are generated using the regressions reported in Tables 1 and 2. (The regressions are estimated over 1953Q1 to 1989Q4 and then simulated over 1981Q1 to 1992Q2.) Actual M2 growth and prediction errors (with summary statistics) are also reported.

The results reported in Table 4 suggest two observations. The first is that the imposition of the convergence condition raises substantially the accuracy of M2 forecasts from the standard M2 demand regression. The root mean squared error (RMSE) declines by about 30 percent when the long-run real GDP elasticity is constrained to be unity. (Compare the RMSEs of regressions A with B and C with D in Table 4.) Over the recent period 1990Q1 to 1992Q2, regressions A and C, which ignore the convergence condition, generate prediction errors in 1990, 1991, and 1992 that cumulate to an overprediction of the level of M2 of about \$324 to \$257 billion, or 9.3 to 7.4 percent, by the second quarter of 1992. These results suggest that the public's M2 demand function experienced a large leftward shift. However, regressions B and D, which impose the convergence condition, indicate a much smaller leftward shift. Prediction errors from the latter regressions cumulate to an overprediction of M2 of only \$144 to \$149 billion, or 4.2 to 4.3 percent.

The second observation is that standard M2 demand regressions systematically overpredict real M2

[3] Bleaney (1990) notes that when the shift point is close to the end of the data set, the appropriately located Chow test is more powerful than some other general tests for structural change.

Table 4

Actual and Predicted M2 Growth; Standard M2 Demand Regressions

Year	AG	Regression A PG	Regression A E	Regression B PG	Regression B E	Regression C PG	Regression C E	Regression D PG	Regression D E
1981	8.9	8.5	.3	9.7	−.8	8.5	.4	8.6	.2
1982	8.7	7.8	.7	8.4	.3	8.3	.4	8.4	.3
1983	11.5	12.3	−.7	13.5	−1.9	12.4	−.9	12.7	−1.2
1984	7.7	7.3	.3	7.6	.1	7.5	.2	7.4	.3
1985	8.3	8.8	−.5	8.9	−.6	9.1	−.8	9.0	−.7
1986	8.8	7.8	1.0	7.3	1.5	7.9	.9	7.5	1.3
1987	4.2	5.4	−1.2	4.3	−.1	4.7	−.5	4.0	.2
1988	5.1	6.2	−1.1	5.3	−.2	6.2	−1.1	5.7	−.6
1989	4.7	6.1	−1.5	5.1	−.4	5.8	−1.1	5.1	−.4
1990	3.9	7.1	−3.2	6.0	−2.1	6.6	−2.6	5.7	−1.8
1991	2.8	6.1	−3.3	4.2	−1.4	5.2	−2.4	3.9	−1.2
Mean Error			−.8		−.5		−.7		−.3
RMSE			1.61		1.12		1.29		.89
Cumulative Error by 1992Q2									
Level (billions)			−323.5		−144.3		−257.3		−148.9
Percentage			9.3		4.2		7.4		4.3

Notes: AG is actual M2 growth; PG predicted M2 growth; and E the prediction error. The predicted values are generated using the money demand regressions reported in Tables 1 and 2. The money demand regressions are estimated over 1953Q1 to 1989Q4 and simulations begin in 1981. RMSE is the root mean squared error.

demand in recent years. This indicates that some additional factors not accounted for in standard M2 demand regressions may be depressing M2 growth in recent years. The appendix to this paper examines the role of a yield curve variable.

IV. Concluding Observations

Since about 1990, standard money demand regressions have overpredicted M2 growth. The empirical results presented here indicate that the size of these prediction errors is consistent with the presence of a stable M2 demand function over the period 1990Q1 to 1992Q2.

The error-correction money demand regressions estimated without the convergence condition do not predict well the current slowdown in M2 growth. The reason is that in such regressions the coefficients on the short-run scale variables are small in magnitude and at times even statistically insignificant. Such estimated short-run coefficients do not cumulate to satisfy the long-run constraint that the long-term scale elasticity is unity. As a result, such regressions may indicate that the short-run changes in real M2 balances are not closely related to short-run changes in the scale variable.

However, not all of the recent slowdown in M2 is predicted by standard M2 demand regressions. The expanded M2 demand regressions reported in the appendix indicate that the recent unexplained weakness in M2 may be due to portfolio substitutions triggered by the steepening of the yield curve. Nevertheless, the size of the current shortfall in M2 that is due to these portfolio substitutions is not so large as to render irrelevant the short-run behavior of M2. M2 has been weak primarily because economic activity has been weak.

APPENDIX

This appendix examines whether a yield curve variable added to M2 demand regressions can account for the recent shortfall in M2 growth.

One of the explanations that has been offered for the recent shortfall in M2 growth is that households have substituted out of M2 into long-term financial assets such as bond and equity funds.[4] These portfolio substitutions were triggered in part by declines in short-term interest rates in general and deposit rates on components of M2 in particular. The steepening of the yield curve encouraged investors to substitute into non-M2 assets.

The slope of the yield curve variable is measured by the long-term bond rate minus the own rate on M2. This variable is used to test whether substitutions by households out of M2 into long-term financial assets can account for the recent money demand prediction error.[5] The yield curve variable is usually not significant in M2 demand regressions if the estimation period excludes the post-1989 data. This result means that long-term interest rates did not influence M2 demand prior to 1989. Hence, these regressions cannot account for the weakness in M2 over the post-1989 sample period. (These results are not reported.)

The yield curve variable enters significantly in money demand regressions if the estimation period includes the post-1989 data. Table 5 reports regression results when the most recent data are used to estimate the influence of the yield curve variable on money demand. In particular, the yield curve measure is entered in the money demand regression as the product of the long-term cost measure and a zero-one dummy that is unity in 1989Q1 to 1992Q2 and zero otherwise. The regressions are estimated over 1954Q2 to 1992Q2.[6] The regression F in Table 5 uses real GDP as the scale variable, whereas the regression G uses real consumer spending as the short-run scale variable and real GDP as the long-run scale variable. Both regressions are estimated under the constraint that the long-run scale elasticity is unity.[7] As can be seen, the yield curve measure enters with the theoretically correct sign and is statistically significant in both regressions. (The yield curve variable is significant even when it is entered in money demand regressions without the interactive dummy.)

Table 6 evaluates whether the regressions reported in Table 5 can eliminate the prediction error over the period 1990Q1 to 1992Q2. In particular, the regressions reported in Table 5 were simulated over 1981Q1 to 1992Q2. The resulting within-sample forecasts of M2 growth are reported in Table 6. As can be seen, the expanded M2 demand regression explains most of the current shortfall in M2. The cumulative overprediction of M2 is now about $8 to $11 billion by the second quarter of 1992. (The cumulative overprediction of M2 is $84 to $86 billion or about 2.5 percent when the yield curve variable is added to money demand regressions without the interactive dummy.)

In sum, the yield curve variable captures substitutions by households between M2 and other long-term financial assets. The empirical work shows that this variable is significant in money demand regressions estimated including the post-1989 data. This result implies that M2 demand in recent years has been affected by portfolio substitutions. However, one needs more observations before one can reliably conclude whether this variable is capturing the random variation in money demand or whether it is capturing the recent systematic influence of the long-term rate on money demand.[8]

[4] Hetzel (1992) provides a thorough review of these alternative explanations. He argues that no single explanation appears to account for the "missing M2" during the recent period.

[5] Others have followed a different approach. For example, Duca (1992) redefines M2 to include funds held in bond and equity mutual funds and then examines whether money demand regressions estimated using mutual funds adjusted M2 series can account for the "missing M2" in recent years. He concludes that the growth of these mutual funds accounts for only a small part of the "missing M2." Hetzel (1992) arrives at a similar conclusion.

[6] The sample period begins in 1954Q2 because the data on the ten-year bond rate used here begins in 1953Q4.

[7] The unconstrained estimate of the long-run part of the money demand regression indicates that the long-run GDP elasticity is unity.

[8] The portfolio substitutions emphasized here are not the only explanation offered for the current weakness in M2. Some have argued that households experienced an adverse shock to their wealth that caused them to desire a smaller amount of debt. They are now reducing their debt by drawing down deposits in M2. Others have suggested that a number of regulatory and economic pressures have reduced the size of the depository system, thereby rechanneling credit flows away from depository institutions and lessening their need to issue monetary liabilities included in M2.

The standard M2 demand regression was alternatively estimated including a lagged value of the level and/or the change in real household net worth. These variables entered with the wrong sign and in general were not significant in the regressions. Similarly, changes in the size of the depository sector were captured by changes in the ratio of deposits in thrift institutions to M2. This variable when included in M2 demand regressions was also not significant.

Table 5

Expanded M2 Demand Regressions; 1954Q2 to 1992Q2

Regression F. Real GDP in the Short- and Long-Run Parts of the Model

$$\underset{}{\Delta\ln(rM2)_t} = \underset{(3.5)}{-.04} + \underset{(6.7)}{.45}\,\Delta\ln(rM2)_{t-1} + \underset{(3.5)}{.25}\,\Delta\ln(rM2)_{t-2} + \underset{(2.7)}{.15}\,\Delta\ln(rY)_t + \underset{(2.7)}{.14}\,\Delta\ln(rY)_{t-1} - \underset{(4.7)}{.003}\,\Delta(R-RM2)_t$$

$$- \underset{(6.5)}{.005}\,\Delta(R-RM2)_{t-1} - \underset{(3.4)}{.07}\,\ln(rM2)_{t-1} + \underset{(3.4)}{.07}\,\ln(rY)_{t-1} - \underset{(1.4)}{.0001}\,(R-RM2)_{t-1} - \underset{(2.0)}{.012}\,CC1$$

$$- \underset{(.1)}{.001}\,CC2 + \underset{(3.2)}{.02}\,D83Q1 - \underset{(1.8)}{.001}\,(R10-RM2)_{t-1} * D_{t-1} - \underset{(2.3)}{.009}\,\Delta(R10-RM2)_{t-1} * D_{t-1}$$

CRSQ = .64 SER = .00555 DW = 2.1 Q(36) = 33.3

Regression G. Real Consumption Expenditure as the Short-Run Scale Variable and Real GDP as the Long-Run Scale Variable

$$\Delta\ln(rM2)_t = \underset{(2.7)}{-.02} + \underset{(5.4)}{.35}\,\Delta\ln(rM2)_{t-1} + \underset{(2.6)}{.17}\,\Delta\ln(rM2)_{t-2} + \underset{(3.7)}{.23}\,\Delta\ln(rC)_t + \underset{(3.7)}{.24}\,\Delta\ln(rC)_{t-1} - \underset{(5.0)}{.001}\,\Delta(R-RM2)_t$$

$$- \underset{(6.1)}{.005}\,\Delta(R-RM2)_{t-1} - \underset{(2.8)}{.06}\,\ln(rM2)_{t-1} + \underset{(2.8)}{.06}\,\ln(rY)_{t-1} - \underset{(2.3)}{.001}\,(R-RM2)_{t-1} - \underset{(1.6)}{.009}\,CC1$$

$$+ \underset{(.5)}{.002}\,CC2 + \underset{(3.5)}{.02}\,D83Q1 - \underset{(1.8)}{.001}\,(R10-RM2)_{t-1} * D_{t-1} - \underset{(2.4)}{.009}\,\Delta(R10-RM2)_{t-1} * D_{t-1}$$

CRSQ = .69 SER = .00511 DW = 2.1 Q(36) = 24.7

Notes: R10 is the ten-year bond rate; D a zero-one dummy that is one over 1989Q1 to 1992Q4 and zero otherwise. All other variables are as defined before.

REFERENCES

Bleaney, Michael. "Some Comparisons of the Relative Power of Simple Tests for Structural Change in Regression Models," *Journal of Forecasting*, vol. 9 (1990), pp. 437-44.

Carlson, John B. and Sharon E. Parrott. "The Demand for M2, Opportunity Cost, and Financial Change," Federal Reserve Bank of Cleveland *Economic Review*, vol. 27 (Quarter 2, 1991), pp. 2-11.

Duca, John V. "The Case of the Missing M2," Federal Reserve Bank of Dallas *Economic Review* (Quarter 2, 1992), pp. 1-24.

Dufour, Jean-Marie. "Dummy Variables and Predictive Tests for Structural Change," *Economic Letters*, vol. 6 (1980), pp. 241-47.

Engle, R. F. and C. W. J. Granger. "Cointegration and Error-Correction: Representation, Estimation and Testing," *Econometrica*, vol. 55 (March 1987), pp. 251-76.

Hetzel, Robert L. "M2 and Monetary Policy," Federal Reserve Bank of Richmond *Economic Review*, vol. 75 (September/October 1989), pp. 14-29.

__________. "How Useful is M2 Today?" Federal Reserve Bank of Richmond Memo, September 25, 1992.

Mankiw, N. Gregory and Lawrence H. Summers. "Money Demand and the Effects of Fiscal Policies," *Journal of Money, Credit and Banking* (November 1986), pp. 415-29.

Mehra, Yash P. "An Error-Correction Model of U.S. M2 Demand," Federal Reserve Bank of Richmond *Economic Review*, vol. 77 (May/June 1991), pp. 3-12.

__________. "The Stability of the M2 Demand Function: Evidence from an Error-Correction Model," *Journal of Money, Credit and Banking*, forthcoming 1993.

Small, David H. and R. D. Porter. "Understanding the Behavior of M2 and V2." *Federal Reserve Bulletin* (April 1989), pp. 244-54.

Table 6

Actual and Predicted M2 Growth; 1981 to 1991

Year	AG	Regression F PG	Regression F E	Regression G PG	Regression G E
1981	8.9	9.7	−.7	8.6	.3
1982	8.7	8.5	.2	8.4	.3
1983	11.5	13.6	−2.0	12.7	−1.2
1984	7.7	7.7	.0	7.3	.3
1985	8.3	8.9	−.6	9.0	−.7
1986	8.8	7.3	1.5	7.5	1.3
1987	4.2	4.3	−.1	4.0	.1
1988	5.1	5.3	−.2	5.7	−.6
1989	4.7	5.5	−.8	5.5	−.8
1990	3.9	4.0	−.1	3.7	.2
1991	2.8	2.2	.6	2.1	.7
Mean Error			−.20		−.02
RMSE			.86		.66
Cumulative Error by 1992Q2					
Level (billions)			−8.0		−11.5
Percentage			.2		.3

Notes: The predicted values are generated using regressions F and G reported in Table 5. These regressions are estimated over 1954Q2 to 1992Q2 and simulated over 1981Q1 to 1992Q2.