AGGREGATING THE MONETARY AGGREGATES: CONCEPTS AND ISSUES

Alfred Broaddus

In early 1970 the Federal Reserve System modified some of the operating procedures it employs in conducting monetary policy. Specifically, the Federal Open Market Committee (FOMC), which is the System's principal policymaking body, began to place somewhat greater emphasis on what have since come to be known collectively as "the monetary aggregates" as operating variables in formulating and implementing monetary policy.[1] The monetary aggregates are various measures of the nation's stock of money. During the 1950's and 1960's, in contrast, the FOMC had focused primarily on conditions in the money markets, as indexed by member bank reserve positions and certain key short-term interest rates. This shift in procedural emphasis has generated a great deal of interest among and comment from monetary economists, financial market participants, and other observers of System policy. Monetary economists found the change interesting because it suggested that monetarist doctrine, which had achieved considerable prominence in academic circles in the 1960's, had finally attained at least a degree of acceptance in the halls of the nation's principal monetary authority. Market participants, on the other hand, regarded the shift as important from the standpoint of evaluating past and present System policy and making judgments about the likely future course of policy.

The extent to which the Federal Reserve has *in fact* altered its operating strategy since 1970 is the subject of a spirited and sometimes heated debate among economists. Some monetarists claim that although the FOMC now gives lip service to the monetary aggregates in its policy pronouncements, it continues to focus mainly on financial market conditions in practice, thereby relinquishing potentially useful control over the aggregates. Conversely, some nonmonetarists believe the FOMC has paid too much attention to the aggregates to the detriment of the credit markets and, consequently, the general economy. For their part, System officials have made it plain in a number of public statements and articles that as far as the System is concerned, the change that occurred in 1970 represented a shift of emphasis among alternative operating variables rather than any official recognition of a change in economic doctrine. The monetary aggregates, while not emphasized, were by no means ignored prior to 1970.[2] Nor have financial market conditions and interest rates been ignored since 1970.

Whatever the merits of these arguments, it is clear that the monetary aggregates presently play a more important role than earlier, both in the formulation and execution of monetary policy and in public discussions of policy. Perhaps the strongest indication of the increasing prominence of the aggregates is their central position in the Congressional resolution concerning monetary policy passed on March 24, 1975. This resolution calls on the FOMC to maintain longer-run growth in the monetary aggregates at rates consistent with the longer-run potential for growth of the nation's productive capacity. The resolution also requests the Federal Reserve to inform the House and Senate Banking Committees periodically of its targets for growth of the aggregates over the following twelve months. The first such hearings took place on May 1, 1975. At the hearings Chairman Arthur F. Burns of the Federal Reserve announced the System's targets for certain aggregates for the period March 1975-March 1976.[3] The hearings received considerable national attention.

The greater emphasis on the aggregates raises some immediate questions. First, precisely what are the monetary aggregates? As the term implies, they are essentially aggregations or summations of the

[1] More specifically still, the FOMC began to express its operating objectives more frequently in terms of the desired behavior of the monetary aggregates in its instructions to the Manager of the System Open Market Account at the Federal Reserve Bank of New York. Acting as the FOMC's agent, the Manager supervises the System's day-to-day purchases and sales of securities, or open market operations. These operations constitute the FOMC's principal tool for implementing monetary policy. The FOMC normally meets once each month. At the conclusion of each meeting it issues a "Directive" to the Manager containing its operating instructions covering the period until the following meeting.

[2] In 1966 the FOMC began supplementing its instructions in the Directive regarding desired money market conditions with explicit references to the desired behavior of certain monetary aggregates. For an interesting discussion of the Committee's attention to the aggregates during the 1950's, see Elmus R. Wicker, "Open Market Money Supply Strategy," *Quarterly Journal of Economics*, 88 (February 1974), 170-9.

[3] See the Statement by Arthur F. Burns, Chairman, Board of Governors of the Federal Reserve System, before the Committee on Banking, Housing and Urban Affairs, U. S. Senate, May 1, 1975, reprinted in *Federal Reserve Bulletin*, May 1975, pp. 282-8.

public's holdings of various financial assets that appear to function as "money" in household and business portfolios. But this description raises more basic questions. What is money? What are its distinguishing functional characteristics? Exactly which financial assets possess these characteristics?

Unfortunately, economists have not arrived at definitive answers to these questions. As a result, universally agreed definitions of money and the money stock do not exist. In the absence of such definitions, the Federal Reserve has found it necessary to take an eclectic approach in the practical implementation of policy. Accordingly, it has defined several monetary aggregates deemed relevant to policy analysis. Each such aggregate is designated by the letter M and a numerical subscript, higher subscripts indicating more inclusive aggregates.

Table I defines the aggregates M_0 - M_7. Economists have traditionally focused on M_1, M_2 and, to a lesser degree, M_3 as the most useful definitions of the money supply. Among these, M_1 is the definition most frequently referred to in public discussions of money and monetary policy. The specification of the higher numbered aggregates shown in the table is a recent development reflecting the growing belief in some quarters that advanced cash management techniques, the introduction of new financial instruments such as large-denomination negotiable certificates of deposit, and other financial market innovations have broadened the spectrum of assets that serve as money.[4] For this reason, some students of monetary policy believe that explicit consideration of these broader aggregates might increase the effectiveness of monetary policy.[5] Others doubt this contention on the grounds that the Federal Reserve would find it difficult to control these aggregates and that their behavior, in any event, bears a predictable relationship over time to the behavior of the narrower concepts such as M_1.

Whatever the outcome of this relatively technical debate, it seems rather paradoxical that in a policy environment where the money supply is such a central concept, there is no professional consensus as to precisely what the money supply is. This article will not attempt to answer this question. Its purpose, rather, is to indicate to nonprofessional readers—many of whom probably take the existence of an agreed money definition for granted—the difficulties inherent in arriving at an unambiguous answer. The article will also describe recent research aimed at developing new money supply concepts superior to those listed in Table I. It is hoped that this material will assist the nonprofessional in critically evaluating commentary in the financial press and elsewhere on the use of a growing list of monetary aggregates in the conduct of monetary policy.

Table I

EIGHT MONETARY AGGREGATES*

M_0 =	currency
M_1 =	M_0 + demand deposits at commercial banks
M_2 =	M_1 + time deposits at commercial banks other than large negotiable certificates of deposit
M_3 =	M_2 + mutual savings bank deposits, savings and loan association shares, and credit union shares
M_4 =	M_2 + large negotiable certificates of deposit
M_5 =	M_3 + large negotiable certificates of deposit
M_6 =	M_5 + short-term marketable U. S. Government securities and savings bonds
M_7 =	M_6 + short-term commercial paper

* For more precise definitions the reader should consult the footnotes to the table titled "Measures of the Money Stock" in the statistical section of any recent **Federal Reserve Bulletin.**

The article contains four sections. The first section reviews the earlier controversy among economists over the proper definition of money. The second section describes a general and highly flexible procedure for developing so-called weighted monetary aggregates. Such weighted aggregates are refinements of the conventionally-derived aggregates listed in Table I and, in the view of at least some economists, potentially better measures of the money supply. The third section reviews some preliminary empirical efforts to estimate the weights that should be attached to particular categories of financial assets in developing operational weighted monetary aggregates.

I. THE POSTWAR DEBATE OVER THE DEFINITION OF MONEY

When it comes to definitions, money is a little bit like sex appeal: everyone has a fairly clear intuitive idea of what it is, but defining it in precise language is difficult. Economists have been arguing about the best way to define money for centuries.[6] Despite

[4] Money supply statistics are published in the monthly *Federal Reserve Bulletin.* Series for M_1, M_2, and M_3 have been carried in these tables for some time. M_4 and M_5 were added to the tables in April 1975.

[5] For a concise official statement of this attitude, see the Statement by Arthur F. Burns, Chairman, Board of Governors of the Federal Reserve System, before the Committee on Banking, Currency and Housing, House of Representatives, July 24, 1975, reprinted in *Federal Reserve Bulletin,* August 1975, pp. 491-7.

[6] An excellent survey of the historical dialog is contained in Milton Friedman and Anna J. Schwartz, *Monetary Statistics of the United States: Estimates, Sources, Methods,* New York: National Bureau of Economic Research, 1970, pp. 89-198.

their inability to achieve a consensus, the question cannot be abandoned either as a theoretical matter or as a practical matter. Clearly, the concept of money lies at the core of both monetary theory and monetary policy.

The effort to define money has been approached from two directions during the postwar period. One segment of the relevant literature has taken a theoretical approach and has sought to settle the issue on the basis of general principles. Analysts in this camp have commonly begun by specifying their respective views regarding the purpose that money serves from the standpoint of the economic units—households and business firms—that hold money. With these purposes delineated, the analyst has then defined money to include the various categories of deposits and other financial assets that appear to serve the indicated functions. The other approach has been more heavily empirical. Here, the choice among alternative definitions has been made on the basis of such criteria as the stability of the relationship between income and various candidate measures of money as revealed by detailed statistical analysis.

Theoretical Approaches As indicated above, those who have taken a theoretical approach to defining money have often begun by asking what money is used for, or, equivalently, why it is demanded. One obvious response to this question is that money is used to facilitate purchases: that is, money is a means of payment. Money should therefore be defined to include those assets used directly in making purchases and to exclude other assets. On the basis of this criterion, some economists have defined money as the sum of currency in the hands of the public and demand (checking) deposits at commercial banks, or M_1. The appeal of this apparently straightforward logic is so great that M_1 has become the most widely accepted definition of money in the eyes of the general public.[7]

A more thoughtful examination of these points, however, suggests that neither the means of payment criterion nor the M_1 definition is necessarily preferable. From the standpoint of both economic analysis and policy, money is interesting primarily because changes in the stock of money held by the public are likely to affect aggregate spending and hence broader economic conditions respecting such things as the level of output, employment, and prices. There is no reason to believe that the stock of assets relevant to spending decisions is limited to those assets that can be used directly as payments media in the act of exchange itself. For this reason, many economists now regard the essential function of money as extending beyond its service as a means of payment to include its use as a "temporary abode of purchasing power," that is, as a repository bridging the gap between the receipt and disbursement of payments.[8]

This extension of the concept of money's function in the economy might seem at first glance to be a minor refinement. Actually, it constitutes a fundamental break with the narrower view of money as a means of payment. For although only a limited number of assets can be used directly in effecting payments, a wide variety of assets can be used as temporary reservoirs of purchasing power in anticipation of payments. It certainly seems reasonable to suppose that a sizeable portion of household balances in commercial bank time and savings deposits, in mutual savings bank deposits, and in credit union and savings and loan association shares are held in anticipation of specific payments. On these grounds, the view of money as a temporary store of purchasing power suggests that M_2 or M_3, or at least some portion of these aggregates, might properly be regarded as money.[9] Shifting the focus from households to business firms produces further possibilities. It is well known that in the current business environment a major goal of corporate management is to minimize noninterest-bearing cash balances. Using highly sophisticated cash management techniques, large corporations are able to maintain a sizeable fraction of what are effectively transactions balances in various money market instruments such as large-denomination certificates of deposit, short-term commercial paper, and short-term U. S. Government securities. It is on this basis that some analysts would suggest that under present conditions at least a portion of an aggregate as broadly inclusive as M_5

[7] The most comprehensive effort to establish M_1 as the proper definition of money on theoretical grounds is found in the work of Pesek and Saving. See Boris P. Pesek and Thomas R. Saving, *Money, Wealth, and Economic Theory*, New York: The Macmillan Company, 1967, pp. 39-254. For a critique of this analysis see Milton Friedman and Anna J. Schwartz, "The Definition of Money: Net Wealth and Neutrality as Criteria," *Journal of Money, Credit and Banking*, 1 (February 1969), 1-14.

[8] Friedman and Schwartz, *Monetary Statistics*, pp. 106-7.

[9] Several recent innovations in the financial sector related to the payments services provided by financial institutions to their customers further support this view. For example, so-called NOW (for negotiable order of withdrawal) accounts offered by thrift institutions and banks in New Hampshire and Massachusetts permit depositors to write what are essentially checks on interest-bearing deposits. Also, Federal regulatory authorities recently adopted new regulations allowing banks and thrift institutions to offer preauthorized bill-paying services to savings depositors.

might reasonably be considered money. Nor does the story necessarily end here. For example, bank loan commitments to business firms are sources if not abodes of purchasing power. None of the aggregates listed in Table I captures this additional source.

Although the view of money as a temporary abode of purchasing power has considerably broader implications than the more restrictive means of payment concept, both tend to focus attention on the relationship between money and current transactions. A somewhat different position regarding the basic function of money has been evident in part of the postwar literature. Taking their cue from Keynesian monetary theory, analysts in this group have emphasized the role of money as a store of liquid wealth held to meet unanticipated contingencies necessitating payments as well as expected transactions and to balance illiquid assets such as long-term securities and nonfinancial assets in household and business portfolios.[10] According to this view, "money" is synonymous with "liquidity," although the latter term has never been specified rigorously. Much of the analysis along these lines was published in the late 1950's and 1960's. Writers in this vein argued that the transactions approach to defining money had tended to restrict attention too narrowly to commercial bank deposits, obscuring the significance of the postwar shift of liquid balances from commercial banks to other financial intermediaries such as savings and loan associations and credit unions. Unless money were viewed more broadly as liquidity, and the liabilities of nonbank intermediaries considered part of the money stock, monetary policy would be rendered ineffective.

The more recent extension of the transactions approach described above, which recognizes the possibility that transactions balances may well be held not only in bank and nonbank deposits but also in a variety of money market instruments, has blurred some of the issues that were central to the earlier debate and broadened the scope of the dialog. At this point, many economists would probably acknowledge that *as a purely formal matter* money might be defined more broadly than M_1, or perhaps more broadly than M_2 or M_3. Beyond that, interest in defining money on purely theoretical grounds appears to have waned.

[10] This strain of analysis began with the work of John G. Gurley and Edward S. Shaw in the 1950's. See John G. Gurley and Edward S. Shaw, "Financial Intermediaries and the Saving-Investment Process," *Journal of Finance*, 11 (March 1956), 257-76, and Gurley and Shaw, *Money in a Theory of Finance*, Washington, D. C.: The Brookings Institution, 1960. Similar views were put forward in the Radcliffe Committee report on the British monetary system published in 1959.

Empirical Approaches Since the theoretical approach to defining money has failed to produce any definitive agreement, it is not surprising that economists have attempted to settle the issue empirically. Indeed, Milton Friedman and Anna J. Schwartz, two prominent participants in the discussion, have suggested that the question of the correct definition of money cannot be separated from the question of the practical uses to which such a definition would be put by policymakers or others:

> We conclude that the definition of money is to be sought for not on grounds of principle but on grounds of usefulness in organizing our knowledge of economic relationships. 'Money' is that to which we choose to assign a number by specified operations; it is not something in existence to be discovered like the American continent; it is a tentative scientific construct to be invented, like 'length' or 'temperature' or 'force' in physics.[11]

As suggested above, money is interesting to economists and policymakers primarily insofar as changes in its stock affect basic economic variables such as income, employment, and prices. From this standpoint, the best definition of money might be the definition producing the closest statistical correlation between money so defined and, say, national income. A large number of statistical tests have in fact attempted to determine which money definition yields the closest correlation. Taken as a group, these studies have shown a close relationship between income and several of the narrower money aggregates such as M_1, M_2, M_3, and variants of these measures. But they have been contradictory and inconclusive regarding exactly which concept produces the best fit.[12] In general, the results of these various tests have been quite sensitive to the time period considered and the exact form of the estimating equations used, especially their respective lag structures.

A related but nonetheless distinct empirical approach has focused on the degree of substitutability among various categories of assets considered candi-

[11] Friedman and Schwartz, *Monetary Statistics*, p. 137.

[12] Representative examples are Milton Friedman and David Meiselman, "The Relative Stability of Monetary Velocity and the Investment Multiplier in the United States, 1897-1958," in Commission on Money and Credit, *Stabilization Policies*, Englewood Cliffs, N. J.: Prentice-Hall, 1963, pp. 165-268; George G. Kaufman, "More on an Empirical Definition of Money," *American Economic Review*, 59 (March 1969), 78-87; Frederick C. Schadrack, "An Empirical Approach to the Definition of Money," in *Monetary Aggregates and Monetary Policy*, New York: Federal Reserve Bank of New York, 1974, pp. 28-34; and Jack L. Rutner, "A Time Series Analysis of Income and Several Definitions of Money," *Monthly Review*, Federal Reserve Bank of Kansas City, November 1974, pp. 9-16. The Friedman-Meiselman and Schadrack studies concluded that M_2 is the preferable definition of money. Kaufman's results suggested that a somewhat broader definition along the lines of M_3 is slightly better than either M_1 or M_2 on the basis of certain evaluative criteria. Rutner's work suggested that the correlation between income and alternative money concepts is itself a function of the time frame of the statistical analysis, broader aggregates performing relatively better over longer time horizons.

dates for inclusion in the definition of money. It is generally agreed that demand deposits should be included in any definition of money. A high degree of substitutability between demand deposits and, say, time deposits would suggest that time deposits can satisfy at least partly the purposes for which demand deposits are held and should therefore be considered money. Statistically, the degree of substitutability has commonly been measured by the sensitivity (in technical language the "cross-elasticity") of the demand for agreed money assets, such as demand deposits, to variations in the interest rates paid on candidate categories, such as commercial bank time deposits and the liabilities of nonbank intermediaries. Unfortunately, these substitutability studies, like the money-income correlation studies discussed above, have not produced conclusive results. Some studies have found relatively low cross-elasticities and have concluded that M_1 is the appropriate definition. Others have found higher elasticities, suggesting that M_2 or M_3 might be preferable.[13]

To summarize, neither theoretical nor empirical analysis has produced a concensus among economists as to precisely what collection of financial assets constitutes "money." On reflection, this lack of agreement is not very surprising. For one thing, a given financial asset can serve its holder in more than one fashion. For example, while a savings deposit provides its holder with a store of purchasing power, it also produces income in the form of explicit interest payments. Therefore, savings deposits as a class might be partly money and partly something else. There is no particular reason for insisting that the definition of money either include or exclude the *entire* stock of savings deposits outstanding. More basically, money is fundamentally a social phenomenon, and, like all social phenomena, is subject to continuous change. What appears to be needed is not some final, exclusive catalog of assets labeled money, but a flexible framework aimed at helping analysts and policymakers determine to what extent specific asset classes are functioning as money at particular points in time. The next section describes such a framework.

[13] Two of the most widely discussed of these studies are Edgar L. Feige, *The Demand for Liquid Assets: A Temporal Cross Section Analysis*, Englewood Cliffs, N. J.: Prentice-Hall, 1964, and Tong Hun Lee, "Substitutability of Non-Bank Intermediary Liabilities for Money: The Empirical Evidence," *Journal of Finance*, 21 (September 1966), 441-57. Feige's study indicated that demand deposits and bank time deposits are weak substitutes, suggesting the superiority of a narrow money definition. Lee found significant substitutability between thrift deposits and bank demand and time deposits, indicating that a broader definition such as M_3 might be preferable. See also Franklin R. Edwards, "More on Substitutability between Money and Near-Monies," *Journal of Money, Credit and Banking*, 4 (August 1972), 551-71. A fourth important study dealing with substitutability, by Chetty, will be discussed later in a somewhat different context.

II. A GENERAL AGGREGATION TECHNIQUE

As we have seen, the monetary aggregates presently monitored by policymakers (see Table I) are simple summations of the total stocks of various financial assets. The characteristic feature of this aggregation technique is that the stocks of all assets included in a given aggregate carry equal and unchanging weights, namely unity. This is a convenient procedure, of course, but it raises some rather pressing questions regarding the analytical usefulness of these aggregates when they are expressed quantitatively. Suppose, for example, that an analyst wished to use M_5 as a measure of the money supply. This aggregate includes such diverse assets as currency, savings and loan shares, and large-denomination certificates of deposit. The weighting procedure employed in deriving M_5 would imply that each dollar of each asset class serves as money to the same degree. This implicit assumption would probably be invalid, whatever the analyst's criterion for defining money might be. Therefore, any uncritical use of M_5 as a measure of the money supply would almost certainly be analytically misleading.

This aggregation procedure is obviously a special instance of a more general technique where the weights attached to each asset category are permitted to vary both among categories and over time. For example, if the goal is an improved measure of the money supply, an analyst might want to attach a higher weight to demand deposits and a lower weight to certificates of deposit in compiling M_5. Edward J. Kane has developed a general framework for the weighted aggregation of monetary variables along these lines, and it will be useful to recapitulate briefly the main features of Kane's technique here.[14] It should be noted at the outset that Kane's technique requires that an analyst using it specify precisely his criterion for determining the relative moneyness of asset classes. Kane's own criterion is the extent to which assets are actually used, that is, liquidated, to support expenditures. It is this particular choice among alternative criteria that gives Kane's analysis its substantive content and raises it above the level of a purely mechanical exercise. The following description of the framework employs elementary algebraic notation for generality and simplicity. No high-powered mathematics is involved.

Kane begins by defining the money balance held by the j^{th} individual economic unit (perhaps a household or a business firm) as:

[14] Edward J. Kane, "Money as a Weighted Aggregate," *Zeitschrift fur Nationalokonomie*, September 1964, pp. 222-7.

$$(1) \quad m_j = \sum_{i=1}^{N} w_{ij}a_{ij}, \qquad i = 1, \ldots, N; \quad j = 1, \ldots, P;$$

where a_{ij} is the dollar amount of the i^{th} asset (one of N available assets) held by the j^{th} unit (one of P units in the economy) and w_{ij} is the weight. The w_{ij} take on values between zero and unity. Any particular w_{ij} may be interpreted as the proportion of the i^{th} asset regarded by the j^{th} unit as serving a money function. We will adopt Kane's money criterion and regard the w_{ij} as signifying the proportion of the i^{th} asset actually used by the j^{th} unit to support transactions during the time period in question. Any number of alternative interpretations of the weights would be consistent with the framework.

The aggregate money stock, M, can be obtained from (1) by summing over the P economic units in the economy:

$$(2) \quad M = \sum_{j=1}^{P} m_j = \sum_{j=1}^{P} \sum_{j=1}^{N} w_{ij}a_{ij}.$$

This expression can be written equivalently as:

$$(3) \quad M = \sum_{i}^{N} \left[\frac{\sum_{j}^{P} w_{ij}a_{ij}}{A_i} \right] A_i,$$

where A_i is the total dollar amount of the i^{th} asset outstanding in the economy. The weighted aggregate is then:

$$(4) \quad M = \sum_{i}^{N} W_i A_i,$$

where:

$$(5) \quad W_i = \frac{\sum_{j}^{P} w_{ij}a_{ij}}{A_i}.$$

Expressions (4) and (5) appear quite simple on the surface, but they point out with great clarity the fundamental problem facing analysts in monetary aggregation. That problem is to specify the determinants of the individual unit weights (the w_{ij}) and, from these, the determinants of the aggregate weights (the W_i). In the absence of empirical evidence, one can only speculate as to what these determinants might be. Such things as interest rates and the price level and expectations of future changes in interest rates and the price level, however, are likely candidates. Further, since both current and expected interest rates and prices change over time, it seems reasonable to suppose that the weights might change in some systematic and therefore predictable manner over time.

Some simple examples might serve to illustrate the potential analytical usefulness of the weighted aggregate concept. Suppose that some technological innovation or perhaps a regulatory change reduced the cost and inconvenience to households of shifting funds from savings accounts to demand deposits. Under these circumstances, households would have an incentive to hold a greater portion of their transactions balances in savings accounts. Abstracting from any other factors affecting the total volume of savings deposits held by households, the weight attached to savings deposits in calculating the effective money supply would rise.[15]

The preceding example suggests the kinds of factors that might alter the weights over the longer run. A second example will indicate some of the factors that might cause the weights to vary systematically over the business cycle. Suppose that during an expansionary period a general increase in short-term interest rates occurred. Corporations would then have a stronger incentive than during a period of low rates to hold their transactions balances in the form of money market instruments such as Treasury bills or certificates of deposit. Under these circumstances, the monetary weights attached to these instruments would rise.

In view of these examples, it would appear that weighted monetary aggregation of the sort suggested by Kane's framework might be useful in developing improved measures of the money supply. At the same time, it is evident that efforts to apply the technique in practice will confront difficult statistical roadblocks. Nonetheless, the approach has been sufficiently appealing to motivate several preliminary empirical studies. The next section summarizes the results of these studies.

III. EMPIRICAL ANALYSIS USING WEIGHTED AGGREGATION

To date, only a handful of studies have attempted to measure statistically the weights that should be

[15] This example, it should be noted, is more than hypothetical, since the Federal Reserve recently lifted its 39-year-old prohibition of the use of the telephone for transferring funds between savings and demand deposits.

attached to individual asset categories in developing monetary aggregates. The results of these analyses can only be considered preliminary. The studies are interesting, nonetheless, with respect not only to the specific numerical estimates of various weights but also to the methodologies employed.

This section summarizes four such studies published during the 1960's. No attempt will be made to evaluate the studies critically. The purpose of the summary is simply to convey the flavor of their results. For brevity, the following notation is used:

CBDD = demand deposits at commercial banks

CBTD = time and savings deposits at commercial banks

MSD = mutual savings bank deposits

PSD = Postal Savings System deposits

SLS = savings and loan association shares

Several preliminary remarks are in order. First, all of these studies appear to have been stimulated by the debate during the 1960's over whether various categories of consumer-type time deposits should or should not be included along with M_1 in the money supply.[16] They therefore focused on assets included in the M_2 and M_3 aggregates of Table I. No effort has yet been made to measure the weights that might be assigned to other assets included in the more broadly defined aggregates such as M_4 and M_5.

Second, the studies employed different assumptions and statistical procedures, and none adopted Kane's detailed framework and definitional criteria as a starting point. Therefore, differences among the estimated weights for particular assets across the four studies reflect conceptual dissimilarities as well as differences in the data and statistical models used. Still, the underlying concepts are sufficiently alike to permit comparison of the estimates.

The earliest of the four studies was a doctoral dissertation completed by Roy Elliott at the University of Chicago in 1964.[17] The purpose of this study was to investigate whether a money aggregate with nonuniform weights displayed a more stable relationship with income than conventional, uniformly weighted aggregates. To this end, Elliott employed a cross-sectional analysis using per capita deposit and income data by states to estimate the weight for a composite group of assets consisting of CBTD, MSD, and PSD. Three separate cross-sectional estimates were derived for three separate years. The estimated weights were .26 for 1929, .35 for 1937, and .65 for 1954. Each of these estimates was significantly different, statistically, from both zero and unity. A separate time series analysis using aggregate national data for the years 1897-1957 produced an estimated weight of .37. This estimate was also significantly different from zero or unity, and its magnitude was consistent with those obtained from the cross-sectional tests.

Elliott's regression model was of the form:

$$(6)\quad \mathrm{Ln}[\mathrm{CBDD} + w(\mathrm{CBTD} + \mathrm{MSD} + \mathrm{PSD})] = a + b[\mathrm{Ln}(y)],$$

where y is permanent income, and w, a, and b are the parameters to be estimated, with w the desired weight coefficient.[18] Several interpretations of w are consistent with this model. One such interpretation is that w measures the proportion of (CBTD + MSD + PSD) held to support current expenditures. If this interpretation is adopted, Elliott's three cross-sectional results tentatively suggest that the moneyness, in this sense, of time and savings deposits was increasing secularly over the time period considered.

Elliott's results can be compared with the results of two other studies employing roughly similar methodologies by (1) Richard H. Timberlake, Jr. and James Fortson and (2) Gurcharan S. Laumas.[19] Both of these time series studies used the following regression model:

$$(7)\quad \Delta Y = a + b(\Delta M_1) + c(\Delta T),$$

where Δ indicates first differences in the variables, Y is current aggregate income, M_1 is as defined in Table I, and T is time and savings deposits measured in various ways as indicated below. Equation (7) can be rearranged in the following manner:

$$(8)\quad \Delta Y = a + b(\Delta M_1 + c/b\ \Delta T).$$

[16] This debate grew out of Milton Friedman's inclusion of consumer time and savings deposits at commercial banks in his definition of the money supply. Virtually all empirical studies of the money supply, including the four discussed here, include M_1 balances at their full face value. That is, M_1 balances carry a weight of unity.

[17] Roy Elliott, "Savings Deposits as Money" (unpublished Ph.D. dissertation, University of Chicago, 1964).

[18] For statistical convenience, the equation Elliott actually estimated was an approximation of (6) that was linear in the weight coefficient w. All variables in his tests were expressed in real per capita terms.

[19] Richard H. Timberlake and James Fortson, "Time Deposits in the Definition of Money," *American Economic Review*, 57 (March 1967), 190-4: Gurcharan S. Laumas, "The Degree of Moneyness of Savings Deposits," *American Economic Review*, 58 (June 1968), 501-3.

The ratio c/b is then the collective weight for the assets included in T in any given test. Since the Timberlake-Fortson-Laumas tests, like Elliott's test, were based on the correlation of the monetary variable with an income variable, the interpretation suggested above for Elliott's w might also be applied to c/b.[20]

Although the Timberlake-Fortson and Laumas studies, respectively, were based on the same model, they produced very different results. Using *annual* data, Timberlake and Fortson estimated the weight for various subperiods between 1897 and 1960 with T defined, as in Elliott's study, as (CBTD + MSD + PSD). Among the pre-World War II subperiods, the estimated weight was positive only for the years 1933-1938. On these grounds the authors concluded that time and savings deposits did not serve a money function during most of the prewar period. The 1933-1938 result was interpreted as implying that the public associated a greater degree of risk with demand than with time deposits during these years in reaction to the rash of bank failures in the early 1930's. Hence, money balances were held in the form of time deposits during this period. In the postwar era, the weight was estimated at a relatively low .15 for the 1953-1965 subperiod.

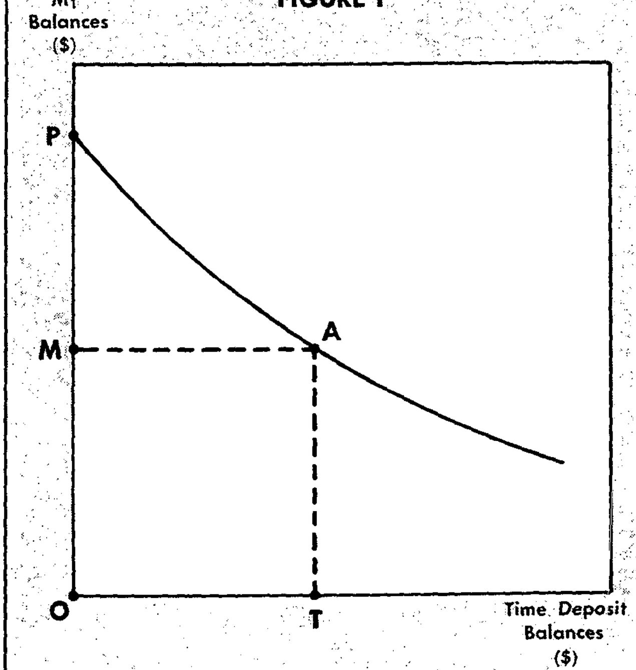
Laumas employed Timberlake and Fortson's technique, but he restricted his study to the postwar era (his tests covered the years 1947-1966), and he used quarterly data and several specifications of T. His results were as follows. Using the Elliott-Timberlake-Fortson specification of T, (CBTD + MSD + PSD), the estimated weight was .48. It is worth noting that this estimate falls about midway between Elliott's cross-sectional estimates for 1937 (.35) and 1954 (.65). Therefore Laumas' results tend to substantiate Elliott's. With T more narrowly defined as CBTD alone, the estimate increased to .58. Defining T more broadly as (CBTD + MSD + PDS + SLS) reduced the estimate to .32. These results imply that the moneyness of CBTD exceeds that of MSD and, in turn, the moneyness of MSD exceeds that of SLS.

The final study, by V. Karuppan Chetty, took a somewhat different approach.[21] Specifically, Chetty measured the weights for individual time and savings deposit categories (CBTD, MSD, and SLS, respectively) on the basis of prior estimates of their substitutability in demand for M_1 balances.

Chetty's conceptual procedure is illustrated in a simplified manner by Figure 1. This diagram depicts the public's allocation of its liquid balances between M_1-type assets, measured on the vertical axis, and time deposits, measured on the horizontal axis. The sloping line in the figure is what economists refer to as an indifference curve. The curve specifies various combinations of M_1 balances and time deposit balances that are equally satisfactory to the public. It also indicates the rate at which the public is willing to substitute balances in one of the categories for balances in the other.[22]

Let us suppose that the shape and position of the indifference curve are known and that the public is observed to be at point A on the curve. At this point it holds OM dollars of M_1 balances and OT dollars of time deposit balances. The curve indicates that the public would be equally satisfied at point P, where it would hold OP dollars of M_1 balances and no time deposit balances. This implies that the public considers a combination of OM dollars of M_1 balances and OT dollars of time deposit balances to be

FIGURE I



[20] While the concepts are similar, the statistical procedures and data employed in the two sets of studies were vastly different. The present writer believes that Elliott's procedures, as detailed on pages 35-40 of his study, were sounder and that his estimates are therefore more reliable.

[21] V. Karuppan Chetty. "On Measuring the Nearness of Near-Moneys," *American Economic Review*, 59 (June 1969), 270-81.

[22] Indifference curves are explained in most elementary economics textbooks. See for example, Paul A. Samuelson, *Economics*, 8th ed., New York: McGraw-Hill Book Company, 1970, pp. 421-6.

equivalent in moneyness to OP dollars of M_1 balances. That is:

(9) Moneyness of (\$OM + \$OT) = Moneyness of \$OP.

To complete the analysis, a measure of moneyness is needed. Let us arbitrarily assume that one dollar of M_1 balances contains exactly one unit of moneyness. Under this rubric the quantity of moneyness in any combination of M_1 and time deposit balances is measured by the dollar value of the M_1 balance to which the combination is equivalent. In the present example, equation (9) then indicates that the combination of OM dollars of M_1 balances and OT dollars of time deposit balances contains OP dollars of moneyness. That is:

(10) Moneyness of (\$OM + \$OT) = \$OP.

Since an M_1 dollar contains one unit of moneyness, we know that:

(11) Moneyness of \$OM = \$OM.

It then follows that:

(12) \$OM + Moneyness of \$OT = \$OP.

Equation (12) can be rewritten equivalently as:

(13) \$OM + m(\$OT) = \$OP,

where m is the proportion of moneyness in a nominal dollar of time deposits. In other words, m is the weight that would be attached to time deposit balances in monetary aggregation.

It is obvious from equation (13) that the dollar values of the balances OM, OT, and OP are sufficient to determine m. Chetty used actual observations of OM and OT (along with interest rate data) for the years 1945-1966 to determine, in effect, the shape and position of the indifference curve in Figure 1. This procedure fixed the point P and established the value of the hypothetical balance OP from which the weight m was then derived.[23]

By deposit classes, the estimated weights were 1.00 for CBTD, .88 for MSD, and .62 for SLS. For the reader's convenience, Chetty's results are shown in Table II along with the postwar period results of the three other studies discussed above. Chetty's weights were generally higher than those found in the other studies. In particular, Chetty's estimate of the weight for CBTD was unity, implying that these deposits should be included along with M_1 balances at their full dollar value in measuring the aggregate money stock. Apart from this, it is worth noting that Chetty's estimates for the respective asset categories were ordered identically to Laumas' estimates.

As previously stated, the results of these empirical studies are tentative at best. As is common in statistical estimation of this sort, the numerical results are quite sensitive to the methods and data used.[24] Nonetheless the similarities among some of the results shown in Table II are at least mildly encouraging. Moreover, the estimates fall generally within a range that is intuitively plausible. In short, the results of these studies suggest that weighted monetary aggregation might be empirically feasible. In addition, they appear to justify further analysis aimed at developing preliminary estimates of the weights of some of the assets included in the broader aggregates of Table I.

Table II

ESTIMATES OF WEIGHTS FOR VARIOUS ASSETS AND COMBINATIONS OF ASSETS DURING THE POSTWAR PERIOD*

Study	Data	Assets	Weight
Elliott	Cross-sectional (state data), 1954	CBTD + MSD + PSD	.65
Timberlake-Fortson	Annual time series, 1953-1965	CBTD + MSD + PSD	.15
Laumas	Quarterly time series, 1947-1966	CBTD	.58
		CBTD + MSD + PSD	.48
		CBTD + MSD + PSD + SLS	.32
Chetty	Annual time series, 1945-1966	CBTD	1.00
		MSD	.88
		SLS	.62

Notation

CBTD — time and savings deposits at commercial banks
MSD — mutual savings bank deposits
PSD — Postal Savings System deposits
SLS — savings and loan association shares

* All estimates assume a weight of unity for currency and demand deposits.

[23] The foregoing description of Chetty's technique is not precise, but it is a close enough approximation for the purposes of the present survey. In technical language, Chetty measured the weights using a regression model derived by maximizing a CES utility function having M_1, CBTD, MSD, and SLS as arguments. For the detailed derivation, see Chetty, "Near-Moneys," pp. 272-8.

[24] As an example of this sensitivity, Franklin R. Edwards found much less substitutability between M_1 balances and other assets when he applied Chetty's model to cross-sectional metropolitan area data. See Franklin R. Edwards, "More on Substitutability Between Money and Near-Monies," *Journal of Money, Credit and Banking*, 4 (August 1972), 564-6.

IV. SUMMARY AND CONCLUSION

This article has attempted to provide an overview of some of the major conceptual issues associated with the process of monetary aggregation and, as a consequence, with the use of monetary aggregates as they are presently defined in the conduct of monetary policy. The article reviewed the main features of the debate among economists over the proper definition of money. It then described a general framework for weighted aggregation and suggested some of the factors that might influence the weights and their behavior over both the short and long runs. The third section reviewed preliminary efforts to estimate the weights of a limited number of assets statistically.

A major aim of this discussion has been to suggest that despite all of the current public comment about the money supply, there is no firm agreement as to precisely what it is or how it should be measured. As we have seen, this state of affairs reflects the fact that money is simply not as concrete and unambiguous a concept as is commonly believed. Moreover, what serves as money can change over time with longer-run changes in financial technology, financial regulations, and underlying social behavior, and possibly with variations in economic activity and financial conditions over the business cycle.

Do these observations imply that the use of the various monetary aggregates shown in Table I is analytically unsound? Not necessarily. They *do* suggest, however, that the combined behavior of these aggregates as a group may provide a more accurate indication of the effect of monetary policy actions than the behavior of any one of them.

This last comment is not intended to imply that simply monitoring a larger constellation of aggregates is an ideal procedure. Refinements are clearly possible. This is where research along the lines described in Sections II and III is relevant. True, the complexity of weighted aggregation and the measurement difficulties associated with the technique will almost certainly preclude employing any such aggregate as an operational variable in the day-to-day conduct of monetary policy. Nonetheless, this research shows promise of producing new insights that might substantively improve the ability of policymakers to interpret the behavior of the conventional aggregates.

A MONETARIST MODEL OF THE INFLATIONARY PROCESS

Thomas M. Humphrey

Given the inherent complexity of the current inflation problem and the tendency of individuals to differ in their interpretation of events, it is not surprising that a number of competing theories of inflation exist today. This article seeks to explain one of these theories—namely, the monetarist view—with the aid of a simple dynamic macroeconomic model developed by the British economist Professor David Laidler.[1] Laidler's model is enlightening for reasons quite apart from its monetarist orientation. Although exceedingly simple, it nevertheless effectively conveys all the essentials of dynamic process analysis—steady-state solutions, disequilibrium dynamics, stability conditions, etc. It is representative of a whole class of models that deal not with *levels* but rather *rates of change* of economic variables. These models are gradually supplanting the once-popular standard textbook or diagrammatical version of the Hicks-Hansen IS-LM model, whose static equilibrium format is not ideally suited to deal with the phenomenon of continuing inflation or with the dynamics of disequilibrium processes wherein economic variables evolve and interact over time. Therefore, regardless of the particular theory being expounded, Laidler's model can be viewed as an introduction to a distinctive form of macroeconomic analysis that attempts to specify the time paths of the inflation rate and related variables.

A word should be said at the outset about the article's position on rival theories of inflation. Regarding the merits of alternative views, this article takes a deliberately neutral stance. Neither monetarism nor any other theory is advocated as being the most nearly correct. No claims are made for the superiority or indeed even the validity of the monetarist view. The sole aim is to articulate the monetarist interpretation within the framework of a mathematical model whose exposition constitutes a useful exercise in its own right. It should be strongly emphasized, however, that the model constitutes a severe oversimplification of a complex process and thus would probably fit the statistical data poorly. As used in this article, the model is intended solely as an expository device and therefore purposely abstracts from many of the variables and behavior relationships that a well-specified empirical model would contain.

Monetarist Propositions Any mathematical model that purports to convey the essence of monetarism must embody certain key propositions or postulates that characterize the monetarist position. Not all of these propositions, however, can be regarded as exclusively monetarist. Some would be accepted to a greater or lesser degree by nonmonetarists. It is therefore desirable to divide these propositions into two groups, namely, those that are distinctively monetarist and those that are not. A partial listing of the uniquely monetarist propositions would include the following.

1. MONETARY THEORY OF INFLATION. Monetarists hold that inflation is a purely monetary phenomenon that can only be produced by expanding the money supply at a faster rate than the growth of capacity output. Thus at any given time the actual rate of inflation is seen as reflecting current and past rates of monetary expansion. Monetarists reject nonmonetary explanations of inflation—i.e., those that attribute rising prices to such alleged causes as shifts in autonomous private expenditures, government fiscal policies, cost-push influences, food and fuel shortages, etc.—on the grounds that an increased stock of money per unit of output is required in all cases and therefore constitutes the true cause of inflation.[2] In short, the sole necessary and sufficient condition for the generation of inflation is said to be excessive monetary growth.

2. LONG-RUN STABILITY (NEAR-CONSTANCY) OF VELOCITY. The proposition of a near-constant circulation velocity or rate of turnover

[1] Laidler presents his model in two papers: "The 1974 Report of the President's Council of Economic Advisers: The Control of Inflation and the Future of the International Monetary System," *American Economic Review*, 64 (September 1974), pp. 535-43, and "The Influence of Money on Real Income and Inflation: A Simple Model with some Empirical Tests for the United States 1953-72," *Manchester School of Economic and Social Studies*, 41 (December 1973), pp. 367-95. The version of the model contained in the present paper differs from Laidler's in at least five respects. First, it is simpler and employs a different notation. Second, its numerous close linkages with monetarism are identified. Third, it is employed solely to explain the monetarist view of inflation. Fourth, an explicit derivation of the equations is provided. Finally, the model and its components are expounded in considerably greater detail than in Laidler's rather terse treatment.

[2] Monetarists readily admit that nonmonetary influences—e.g., union wage pressure, monopoly (administered) pricing policies, OPEC cartels, oil embargoes, crop failures, commodity shortages, and the like—can directly affect particular prices. But they argue that without excessive monetary growth such nonmonetary-induced rises in the prices of some commodities eventually would be offset by declines in the prices of others, leaving the average price level unchanged.

of money follows logically from the monetarist view that inflation stems solely or largely from excessive monetary growth. For if velocity were not a constant it would exhibit a non-zero rate of change that would supplement monetary growth as a separate and independent determinant of inflation. It follows, therefore, that monetarists must assume that velocity is at least a quasi-constant if they are to assert that inflation stems solely or primarily from changes in the stock of money per unit of output.

3. EXOGENEITY OF THE NOMINAL STOCK OF MONEY. Monetarists treat the quantity of money and its rate of growth as variables whose magnitudes are fixed outside the system.[3] This view contrasts sharply with the nonmonetarist treatment of money as an endogenous variable determined within the system by the level of economic activity and by the public's preferences for money and for liquid-asset money substitutes. The exogeneity postulate implies that monetary growth enters the system as a datum to determine the growth rates of spending, prices, and nominal income. The postulate is therefore consistent with the monetarist view of monetary growth as the independent causal factor governing the rate of inflation.

4. ABSENCE OF REVERSE CAUSALITY RUNNING FROM INCOME TO MONEY. Implied by the exogeneity condition, this proposition rejects the notion of passive income-determined monetary growth and asserts the monetarist view of the unidirectional channel of influence or flow of causation running from money to spending to income to prices. Monetary growth is seen as entering this sequence not as a dependent or accommodative variable responding passively to prior income growth but rather as the active independent variable that precedes and causes inflation. It is true that monetarists, in their asides and qualifications, acknowledge that income may influence money indirectly through the policymakers' reactions to changes in the economy. But for the most part they have not incorporated such policy response functions into their formal models, and they continue to treat monetary policy as largely exogenous.

The preceding constitutes the group of uniquely monetarist tenets. As for the remaining key propositions, i.e., those that monetarists share with at least some nonmonetarists, they can be listed briefly. They include the following: (5) the non-neutrality of money in the short run (i.e., the tendency for changes in monetary growth to have substantial effects on real output and employment in the short run); (6) the long-run neutrality of money (i.e., the tendency for changes in monetary growth to have no lasting impact on real output and employment but only on the rate of inflation); (7) the view of erratic and volatile monetary growth as the prime cause of business cycles; (8) the inherent stability of the economy (i.e., the view of the system as a self-regulating mechanism, perturbations of which tend to generate only damped cycles about full-employment equilibrium); (9) the existence of long lags in the response of inflation to changes in the rate of monetary growth; and finally (10) the importance of inflationary expectations in determining market wage- and price-setting behavior. As shown below, Laidler's model is capable of accommodating all these propositions.

The Model and Its Components The model itself is composed of three equations, the first being the *monetary growth equation.* A dynamic version of the static Cambridge cash-balance formula, this equation relates the rate of growth of real (price-deflated) cash balances to the growth rate of real output. The second relation in the model is a *price-adjustment equation* that explains the determination of the current rate of inflation. The third component is an *expectations-formation equation* that embodies a particular hypothesis about how people formulate their expectations of the future rate of inflation. Using these three equations one can solve for the three endogenous variables of the model, namely, (1) the current rate of inflation, (2) the expected rate of inflation, and (3) an excess demand variable represented by the gap between actual and capacity real income. In addition to these endogenous variables there is one exogenous variable, the growth rate of the nominal money stock, and one exogenous constant, the growth rate of full-capacity real income. This treatment of the monetary variable reflects the monetarist view of the money stock and its growth rate as largely exogenous magnitudes determined by an autonomous central bank via its control over a base of so-called high-powered money, consisting of currency and bank reserves. It also effectively rules out any reverse-causation feedbacks running from income to money. The assumption of a fixed capacity growth rate also squares with monetarist doctrine, which holds that the long-run path of potential output is independently determined by fundamental real economic conditions including technological progress and labor force growth.

Three other features of the model should be mentioned at the outset. First, all relations are linear and are expressed in logarithmic form. There is a specific reason for this formulation. Modern monetarist analysis is usually stated in terms of percentage rates of change of the relevant variables. And since the percentage change of any variable over a given interval of time can be represented mathematically by the first time difference of its logarithm, it follows that a log-linear formulation facilitates the analysis.

A second feature of the model is the introduction of time delays in the form of lagged relationships among the variables. These lags reflect the mone-

[3] The exogeneity condition applies only to the *nominal* and not to the *real* (price-deflated) stock of money. Unlike the nominal stock, the real stock is treated as an endogenous variable determined by the public's demand for real balances. The public, via the impact of its spending on the price level, can make the real value (purchasing power) of any given nominal stock of money conform to whatever magnitude it desires.

tarist view of the many delays or frictions inherent in the inflationary process. Their inclusion also permits the analyst to describe the time paths taken by output and prices following a monetary disturbance.

The third feature of the model is its extreme simplicity, as manifested by the minimal number of variables it contains. In particular, the model possesses neither an interest-rate variable nor a variable to represent a discrepancy between actual and desired real cash balances. As a result, the model ignores two potentially important elements in the inflationary process, namely, (1) changes in the rate of interest and (2) the transitory rise in real cash balances (or the temporary fall in the velocity of money) that occurs at the beginning of inflationary periods immediately following a rise in the growth rate of money. These elements could of course be explained in a more complex model, but such a model would lose in simplicity, manageability, and ease of comprehension what it gains in completeness. Moreover, Laidler's model, despite its simplicity, is capable of explaining a large part of the inflationary process, namely, how variations in the growth rate of the money stock are divided between changes in real output and prices both in the short and the long run.

As for notation, the model employs the following symbols. Let m be the money stock, y actual real income, y_c standard or normal capacity real income, x the excess demand variable represented by the difference between actual and capacity income, i.e., $x = y - y_c$, and p the price level—with all variables expressed as logarithms. Actual real income, y, can exceed capacity, y_c, because the latter is defined not as the absolute physical limit or maximum ceiling level of output but rather as the output associated with the economy's normal or standard level of operation. This concept of capacity or potential output corresponds roughly to the monetarist notion of the *natural rate of unemployment,* i.e., the unemployment rate that, given the inevitable frictions, rigidities, and market imperfections existing in the economy, is just consistent with equilibrium between demand and supply in the labor market. The superscript e denotes the expected value of a variable, and the subscripts -1 and -2 denote time lags of one and two periods, each defined as being a year in length.[4] The symbols Δ and Δ^2 appearing before a variable denote first and second time differences, respectively, so that the model is effectively expressed in terms of proportional rates of change and rates of acceleration or deceleration of those rates of change. Finally, a bar over a variable indicates that it is exogenous, i.e., determined outside the system.

The Monetary Growth Equation The first equation of the model is the monetary growth equation:

$$(1) \quad \overline{\Delta m} - \Delta p = \Delta y = \Delta x + \overline{\Delta y_c}.$$

This equation states that the rate of growth of the real money stock—i.e., the percentage rate of nominal money growth, Δm, less the percentage rate of price inflation, Δp—determines the percentage change in real expenditure and hence real income, Δy, that occurs during the given period. More precisely, a rate of growth of the real money stock, $\Delta m - \Delta p$, in excess of the growth rate of capacity output, $\overline{\Delta y_c}$, causes the growth rate of actual output, Δy, to deviate from the capacity growth rate, where the deviation is represented by the variable Δx, i.e., $\Delta x \equiv \Delta y - \overline{\Delta y_c}$.

Equation (1) implies a constant unitary income elasticity of demand for real (price-deflated) money balances. This condition follows from the notion—associated with the old Cambridge cash-balance version of monetarism—that people desire to maintain a stable (constant) proportional relationship between their real cash balances and real income. If the ratio of real balances to real income is to remain fixed, then both elements of the ratio must grow at the same percentage rate, as in equation (1). The monetary growth equation also expresses the strong monetarist view of a stable equiproportional relationship between changes in nominal money and nominal income and likewise between changes in nominal money per unit of output and the price level. The equation predicts that a given percentage change in nominal money will be matched by an identical percentage change of nominal income. The same holds for percentage changes of nominal money per unit of output and of prices. Note, however, that the equation, by itself, is incapable of expressing a stable predictable short-run relationship between nominal monetary growth and the inflation rate. This is because, in the short run, monetary growth may stimulate output as well as prices. And one cannot determine from equation (1) alone the proportions in which the stimulus will be divided between price changes and output changes. One has to supplement equation (1) with the price-adjustment equation to explain this division.

Equation (1) may also be interpreted as embodying a crude monetarist view of the *direct expenditure mechanism* whereby monetary impulses are transmitted directly to income via a prior effect on the

[4] The time period in Laidler's model is not specified but is defined here as one year to conform to the monetarist interpretation that this article is developing.

demand (spending) for goods. The direct mechanism should be contrasted with the *indirect interest-rate mechanism*—often stressed by nonmonetarists—in which monetary changes influence income indirectly via a prior effect on the rate of interest.[5] As shown in Appendix A, the money growth equation is derived from the celebrated Cambridge cash-balance equation and assumes that the velocity of money (or the Cambridge K) is constant and that the money market clears with sufficient rapidity to maintain equality between money demand and supply. The constant-velocity assumption is what insures that given rates of monetary growth, real and nominal, will be matched in equation (1) by corresponding identical rates of income growth, real and nominal.

The Price-Adjustment Equation The second equation of the model explains how the current rate of inflation is determined, i.e., the rate at which businessmen mark up their product prices. The price-adjustment equation is written in the following way:

$$(2) \quad \Delta p = ax_{-1} + \Delta p^e_{-1}$$

where Δp is the current rate of inflation, x_{-1} is excess demand lagged one period, and Δp^e_{-1} is the rate of inflation expected to prevail in the present period as of the preceding period. The price-adjustment equation expresses a short-run relationship between the rate of inflation, Δp, and excess demand, x, the latter measured by the gap between actual and potential (i.e., normal capacity) output. The existence of a gap implies that businessmen are straining productive capacity in an effort to meet demand. Spare plant and equipment are being drawn into use and increasing resort is being had to overtime and marginal labor. In brief, resources become increasingly scarce relative to demand as production approaches and then surpasses standard capacity output. The size of the gap measures the pressure of resource scarcity on prices. The larger the gap, the greater the pressure. As the gap expands, wages are bid up, labor-hour productivity falls, unit costs rise, bottlenecks develop, and the backlog of unfilled orders mounts. All these forces combine to cause prices to rise at an increasingly rapid rate. Thus inflation accelerates as the gap expands.

From the preceding discussion, it is evident that the price-adjustment equation is similar to so-called Phillips-curve equations that state a trade-off relationship between the rate of wage increase and the unemployment rate. In the price equation, however, excess demand replaces the unemployment rate as the indicator of the level of economic activity, and the rate of price inflation replaces the rate of wage inflation as the dependent variable. It is, of course, assumed that rates of wage increase in excess of productivity growth eventually tend to be incorporated in rates of price inflation as businessmen raise their prices to cover increases in unit labor costs.

According to the price-formation equation the rate at which businessmen mark up their prices depends upon two influences, namely, the level of excess demand, x, and the expected rate of inflation, Δp^e. The equation states that if aggregate supply and demand are equal so that there exists no excess demand ($x =$ zero), then actual price inflation will just equal expected inflation. If, however, product demand exceeds supply at the economy's natural or normal capacity level of operation, businessmen eventually will react to the excess demand by raising prices at a faster rate than the expected rate of inflation. This price response, however, is not instantaneous. For a while, quantities rather than prices tend to absorb the impact of excess demand as businessmen temporarily expand output and perhaps allow their inventories to be depleted. These quantity changes signal the desirability of raising the rate at which prices are marked up. Later, therefore, businessmen respond to the excess demand by raising prices. The one-period lag—again defined as a year—on the excess demand variable is meant to account for the time it takes for a shift in demand to affect prices. The coefficient a, attached to the excess demand variable, measures the magnitude of the impact that any given volume of excess demand has on the rate of inflation. The higher the numerical value of a, the greater the impact. This coefficient, of course, must be a positive number, i.e., $a > 0$.

Expectations-Formation Equation The third equation of the model is the expectations-formation equation. It is written as follows:

$$(3) \quad \Delta p^e = b\Delta p + (1-b)\Delta p^e_{-1}$$

or, alternatively, as:

$$(3a) \quad \Delta p^e - \Delta p^e_{-1} = b(\Delta p - \Delta p^e_{-1}).$$

This implies that the *change* in the expected rate of inflation, $\Delta p^e - \Delta p^e_{-1}$, is proportional to the amount by which this period's actual inflation, Δp, deviated from expected inflation as forecast one year ago, Δp^e_{-1}, with the factor of proportionality, b, having a value between zero and unity.

[5] Modern monetarists acknowledge that interest rate effects are always present. They view the direct mechanism merely as an empirical proxy for the indirect mechanism in which many specific interest rate effects cannot be captured statistically either because they are implicit and hence unobservable or because they are too weak and too brief to be measured.

Embodied in the equation is a particular theory—the so-called adaptive-expectations or error-learning hypothesis—of how inflationary expectations are formed. According to the error-learning hypothesis, people formulate expectations about the inflation rate, observe the discrepancy between the actual and anticipated rates, and then revise the anticipated rate by some fraction of the error between the actual and anticipated rates. It can also be shown that the adaptive-expectations hypothesis is equivalent to the theory that people formulate price-expectations by looking at a geometrically-weighted average of current and past rates of inflation with the weights diminishing exponentially as time recedes. This weighting scheme implies that people assign higher weights to more recent phenomena when forming expectations.

How realistic is the error-learning hypothesis? Some economists claim that it is not an accurate description of how anticipations are formed. These analysts argue that expectations are as likely to be generated from direct forecasts of the future as from mere projections of the past. Moreover, they assert that people probably base anticipations at least as much on current information about a variety of developments as on old data pertaining solely to past price changes. There is undoubtedly much truth in these observations. Nevertheless, the error-learning formulation will be retained in this article subject to the caveat that purely extrapolative price forecasts may be modified by additional information.

The Complete System Taken together, the *money growth, price-adjustment,* and *expectations-formation* equations form a simple three-equation system that embodies a monetarist view of the inflationary process. To recapitulate, the complete system is written as follows:

(1) $\overline{\Delta m} - \Delta p = \Delta x + \overline{\Delta y_c} = \Delta y$

(2) $\Delta p = a x_{-1} + \Delta p^e_{-1} \qquad a > 0$

(3) $\Delta p^e = b \Delta p + (1-b) \Delta p^e_{-1}. \qquad 0 < b < 1$

The variables in this system of equations interact to determine the rates of expected and actual inflation and the short-run growth rate of real income. The logic of the system implies that variations in the money growth rate initially affect excess demand, thereby inducing real income to deviate from its full-employment path. Lagged excess demand interacts with lagged price-expectations in equation (2) to determine the current rate of inflation. The current rate of inflation enters equation (3) to influence the expected rate, which in turn feeds back into equation (2) to become a determinant of next period's inflation rate. Finally, in equation (1) the current rate of inflation interacts with the given rate of monetary growth to determine the growth rate of real income. In this manner the system and its constituent elements determine the division of monetary growth, Δm, between price and output growth, Δp and Δy.

Less formally, the model implies the following causal chain.

1. Inflation is determined by excess demand and inflationary expectations.

2. Inflationary expectations are generated by previous inflationary experience and hence by previous excess demand.

3. Excess demand is created by excessive monetary growth.

4. Therefore, excessive monetary growth—past and present—is the root cause of inflation.

The Long Run and the Short It is useful at this point to distinguish between the long-run and the short-run properties of the system of equations. This dichotomy, of course, corresponds to the two main stages or phases of the inflationary process, i.e., the temporary or transition phase in which changes in monetary growth affect real output and employment and the final or permanent stage in which the sole impact is on the rate of inflation. It also corresponds to the monetarist distinction between the long-run neutrality and the short-run non-neutrality of money. In the context of the model, the long run refers to the equilibrium or steady-state solution of the system after it has completely adjusted to a monetary disturbance. By contrast, the short run refers to the disequilibrium transitional adjustment period between successive long-run equilibria. Regarding the long run, the relevant question is whether a monetary shock has any lasting impact on real variables, i.e., is there a permanent trade-off between inflation and output. As for the short run, one should focus on the type of monetary shocks that initially disturb the system and upon the subsequent reaction of the system to those shocks. Does a monetary disturbance affect output as well as prices in the short run? What types of time paths do the variables describe in disequilibrium? How do the variables interact to produce these paths? Finally and most important, do these paths tend to converge on the long-run equilibrium, i.e., is the system stable?

Long-Run Steady-State Solution of the System According to monetarist doctrine, long-run monetary equilibrium is characterized by the following conditions:

1. Equality between actual and expected rates of price change, reflecting the long-run tendency of people to correctly anticipate inflation and fully adjust to it;

2. The absence of any trade-off between inflation and output, reflecting the tendency of monetary shocks to have no lasting impact on real variables but only on prices;

3. A constant steady-state (non-accelerating, non-decelerating) rate of inflation equal to the difference between the growth rate of the money stock and the growth rate of capacity output;

4. Attainment of full-capacity real income reflecting the long-run tendency of actual output to adhere to its full-employment growth path.

Does the model yield these conditions? Only a look at its steady-state properties will tell, i.e., the model must be analyzed at its long-run equilibrium position. The concept of equilibrium, of course, implies equality between aggregate demand and supply, i.e., a state of zero excess demand. Setting the excess demand variable, x, equal to zero in the price-adjustment equation yields $\Delta p = \Delta p^e_{-1}$. Thus, actual and expected inflation are equal, as required. Moreover, the zero numerical value of the excess demand variable (an index of real economic activity) in the price equation signifies the absence of long-run inflation-output trade-offs, as required. Money growth has a neutral long-run impact on real variables, at least in the model.

The next step is to set the first difference of excess demand, Δx, at zero in the money growth equation. Doing so enables one to solve for the steady-state rate of inflation, which is $\Delta p = \overline{\Delta m} - \overline{\Delta y_c}$. In brief, the model *does* yield the monetarist conclusion that the equilibrium rate of inflation is the difference between the respective growth rates of the money stock and full-capacity income. The final step is to recognize that when excess demand goes to zero, actual output growth converges on its full-capacity path, consistent with the fourth condition of monetary equilibrium. Therefore, the model contains all the equilibrium conditions required by monetarist doctrine.

Disequilibrium Dynamics of the System in the Short Run So much for equilibrium analysis, which is a relatively simple and straightforward exercise. The next stage is disequilibrium dynamic analysis. Unfortunately, the analytics of the short-run disequilibrium behavior of the system are somewhat more complex and involved. For one thing, the excess demand variable does not drop out of the short-run analysis as it does in the long-run equilibrium case; nor is the current rate of inflation stationary and identical to the expected rate.

The short-run analysis involves at least two steps. First, because interest centers on the time-paths of (1) inflation and (2) the excess demand gap between actual and capacity income, one must derive expressions for the dynamic behavior of these two variables. This derivation is accomplished in Appendix B. Second, the resulting expressions must be analyzed to determine whether the system is dynamically stable, i.e., whether the variables will eventually converge on their long-run equilibrium values.

Disequilibrium Dynamical Equations As shown in Appendix B, the expressions for the respective short-run time paths of the inflation rate and excess demand are:

$$(4)\quad \Delta p = ax_{-1} - a(1-b)x_{-2} + \Delta p_{-1}$$

and

$$(5)\quad x = \overline{\Delta^2 m} - \overline{\Delta^2 y_c} + (2-a)x_{-1} - [1-a(1-b)]x_{-2}.$$

Two monetarist features are immediately apparent even from the most casual inspection of these equations. First is the appearance of the second time difference, Δ^2, of the money stock variable in the excess demand equation. This second difference, of course, measures the rate of change (i.e., acceleration or deceleration) of the money stock growth rate. Its role in the equation as an active independent variable and determinant of the excess-demand gap is consistent with several monetarist propositions. It squares with the monetarist view of variation in the growth rate of money as the prime initiating cause of business cycles. It corresponds with the monetarist argument that sharp changes in money growth can disturb real income in the short run. In general, it is consistent with the monetarist focus on changes in the *growth rate* rather than the *level* of money as a key indicator of recent policy shifts and future price movements.

The second conspicuous monetarist feature is the appearance of lagged values of excess demand in the price-change equation. The equation states that demand leads inflation by as much as two periods, each defined as a year—another manifestation of the monetarist view of the tendency for shifts in demand to influence quantities first, prices only later. This lead-lag relationship corresponds to the monetarist notion of long and complex lags in the monetary transmission mechanism.

The lag structure of the model carries some important policy implications. Given the long lag in the response of prices to changes in demand—not to mention additional delays in the influence of money on demand—inflation will be slow to respond to contractionary policy. This is especially true if inflationary expectations have become firmly embedded in behavior patterns. It is a generally-accepted principle that an inflation rate that comes to be anticipated will resist a period of deficient demand much longer than a rate that is not anticipated. To reduce the actual rate of inflation one must reduce the expected rate, since the latter is a determinant of the former. This requires a recession during which the actual rate falls below the expected rate, inducing a gradual downward revision of the latter. According to the adaptive expectations hypothesis, however, expectations are based on a weighted average of current and past rates of inflation. And it may take a long time before the decelerating current rate begins to outweigh the lagged influence on expectations of accelerating past rates. During this time there exists the danger that the authorities, observing the failure of their actions to achieve quick results, may be tempted to abandon monetary restraint as ineffective. Monetarists, however, would counsel perseverance, believing that contractionary policy, if adhered to long enough, would eventually bring down the rate of inflation. Monetarists would argue, moreover, that there is no other option—continued monetary restraint is the *only* way to reduce inflation permanently.

Stability Analysis of the System The last step in the analysis of the model is to examine the dynamic *stability* of the system. Here the term stability means the tendency of the system when in disequilibrium to converge on its long-run steady-state equilibrium. The concept of stability is central to the rules versus discretion debate between monetarists and nonmonetarists. Some of the latter group claim that the economic system may be inherently unstable such that once disturbed it tends either to oscillate ceaselessly about equilibrium in cycles of regular or increasing amplitude, or alternatively, to move steadily away from equilibrium via a divergent monotonic path. Other nonmonetarists believe that, while the system is stable, the adjustment process takes too long to be left to itself. These views lead to the advocacy of discretionary stabilization policy to counter or smooth the cycle. By contrast, the monetarist group views the economy as an inherently stable self-regulating mechanism capable of restoring equilibrium without the intervention of discretionary policy. In fact, monetarists contend that due to the existence of long, variable, and unpredictable lags in the monetary transmission mechanism, discretionary stabilization policy has a capricious and often *destabilizing* impact on the economy, amplifying rather than dampening cyclical swings. This argument forms the basis of the monetarist advocacy of a rigid policy rule fixing the growth rate of the money stock.

What about the stability of the model? Will output converge on its capacity growth path and will the excess demand gap vanish as the monetarists predict? To answer these questions one must analyze the excess demand equation

$$(5) \quad x = \overline{\Delta^2 m} - \overline{\Delta^2 y_c} + (2-a)x_{-1} - [1-a(1-b)]x_{-2}.$$

It is assumed that the initial monetary disturbance has ended and that, consequently, money is now growing smoothly at a constant rate. In other words, the *rate of change* of the money growth rate—$\Delta^2 m$—is zero. Moreover, it is also assumed that the growth rate of capacity output is a constant, i.e., that the rate of change of the capacity growth rate—$\Delta^2 y_c$—is also zero. Setting these first two terms on the right-hand side of the equation at zero leaves the second-order difference equation:

$$(6) \quad x = (2-a)x_{-1} - [1-a(1-b)]x_{-2}.$$

Specialists in dynamic models have worked out a set of stability conditions for this type of equation. These conditions are listed in Appendix C. By referring to the stability criteria, it can be shown that, given plausible values of the coefficients a and b, the system will be stable. Depending upon the specific magnitudes of the coefficients, the system may approach long-run equilibrium either monotonically or cyclically, but it will always converge upon it.[6] Hence the model conforms to the monetarist specification of an inherently stable system.

Monetarist View of the Inflationary Process The foregoing section completes the analysis of the steady-state and disequilibrium dynamical properties of the model. These properties were shown to be consistent with the basic postulates of monetarist doctrine. It remains to compare Laidler's formal model with a leading monetarist's verbal description of the inflationary process to see if the two agree with regard to treatment of timing, direction of causation, and pattern of interaction of key variables.

---

[6] Only the excess demand equation is examined here. Exactly the same type of analysis can be performed on the difference equation expressing the behavior of the inflation rate following a step increase in the monetary growth rate. Such an analysis reveals that the rate of inflation eventually stabilizes at a level equal to the difference between the new monetary growth rate and the growth rate of capacity output.

Professor Milton Friedman, perhaps America's foremost monetarist, summarizes the inflationary process as a stylized sequence of events.

> Start from a hypothetical, reasonably balanced situation when monetary growth has been proceeding for some time at a constant rate so that the public in general has adjusted to that rate. GNP in nominal terms will then be growing at about the same percentage rate as M_2, prices at about 3.0 to 4.0 percentage points less. Let the growth rate of M_2 accelerate. For something like six months, the main effect will be that actual balances will exceed desired balances, which may temporarily depress short-term interest rates but will have little other effect. After about six to nine months, the rate of growth of nominal GNP will accelerate, as holders of the excess cash seek to dispose of it. The increased spending . . . will 'excite industry,' as producers facing unexpectedly high nominal demands treat the increase as special to them and so seek to expand output. For a time they can do so, because their suppliers too, including laborers, take the increase in demand as special and temporary and do not alter their anticipations. This, if you will, is the temporary Keynesian phase, where output responds more quickly than prices. In its course, prices do respond, rising more rapidly than before, and interest rates stop falling and start to rise. But it takes about eighteen months after output starts to quicken—or two years after money accelerates—for the main effect to have shifted from output to prices. During this period, anticipations are changing, reflected most sensitively perhaps in interest rates, but even after prices have started to absorb the bulk of the acceleration in money, anticipations have not fully caught up. In the next year or so they will, which will force a decline in the rate of growth of output back to or below the 'natural level,' producing the stagflation stage.[7]

Friedman's description clearly implies a chain of causation running from money to spending to output to prices to inflationary expectations, with deviations between actual and expected rates of inflation feeding back into the process to determine the division of the increase in spending between price and output growth. Moreover, there are substantial time lags operating in each link of the chain or stage of the inflationary process. Together, these feedbacks and time lags produce growth cycles, i.e., oscillations of output growth about the equilibrium or full-capacity growth rate.

How does the formal model compare with Friedman's description? Two differences are immediately apparent. The first relates to the initial (money-spending) link. Friedman asserts the existence of a six-to-nine month lag in the response of spending to monetary stimuli. During this interval, the total impact of the monetary shock is absorbed by a passive rise in undesired cash balances; none of the shock is transmitted to spending. By contrast, the model implies an instantaneous first-round response of spending to changes in money growth. The difference stems from the model's simplifying assumption that actual and desired real cash-balances are always identical, implying the absence of an adjustment lag for real balances. As a second departure from Friedman's version, the model—again for purposes of simplicity—contains no interest rate variables and therefore cannot describe the impact of inflation on interest rates. In brief, Friedman's description implies the existence of one additional time-lag and one additional variable absent from the model.[8]

As for (1) direction of causation and (2) pattern of interaction of variables, however, the model is quite similar to Friedman's description. Causation runs from money to output to prices to expected inflation and back again to real income. Specifically, in the model the sequence is as follows.

> (1) Accelerated money growth generates excess demand, thus causing real output to rise above its full-capacity growth path. See equation (5).
>
> (2) After a lag, excess demand begins to influence the current rate of inflation, causing it to rise above the expected rate. See equation (2).
>
> (3) The rise in the actual inflation rate in turn influences the expected rate, which will feed back into next period's actual rate. See equations (3) and (2).
>
> (4) The rate of inflation interacts with the given rate of money growth to determine the growth rate of output. See equation (1). Moreover, since the rate of inflation itself is determined by the level of excess demand and by expected inflation, these two variables may be regarded as determining the division of monetary growth between output and price level growth.
>
> (5) Finally, current output growth as determined in equation (1) feeds back into equation (2) to become

[7] "Rediscovery of Money—Discussion," *American Economic Review*, 65 (May 1975), 178.

[8] It should be noted that Friedman's explanation of the expectations-formation mechanism is consistent with the so-called *rational expectations hypothesis* and thus may differ from the adaptive expectations model employed by Laidler. According to the rational expectations hypothesis, the inflationary expectations that individuals formulate represent the most-accurate (unbiased) forecasts given the available market information on the stochastic process generating the inflation. By contrast, the adaptive expectations hypothesis may imply nonrational forecasting behavior. That is, it can be shown that under certain conditions, the adaptive expectations mechanism will produce forecasts that are systematically wrong. For example, suppose the monetary authority follows a policy rule of continually accelerating the rate of inflation. In this case the backward-looking adaptive expectations model will yield a predicted rate of inflation that lags consistently behind the actual rate, i.e., inflation will be systematically underestimated. Adherence to the adaptive expectations model despite persistent forecasting errors implies nonrational behavior. Rational individuals would revise their forecasting model to produce unbiased predictions. Once rational individuals learn of the policy rule, they will adopt it as their optimal forecasting model. Under other very restrictive conditions, however, the adaptive expectations model will yield rational (i.e., unbiased) predictions. This would be the case if the time path of inflation is generated by random shocks of a permanent and transitory nature. The notion of the inflation-generating process as a random-walk with noise superimposed would seem to correspond closely to the monetarist view of the capricious and unpredictable impact of discretionary monetary policy. If so, then the adaptive expectations mechanism would be consistent with rational behavior, at least within the context of monetarist models.

a determinant of next period's inflation rate, etc. As mentioned in the preceding section, this iterative process is capable of producing oscillations much like those mentioned by Professor Friedman.

To summarize, both Friedman and Laidler agree that, owing to the operation of lags in the monetary transmission mechanism, the effect of money growth on the rate of inflation is spread over substantial periods of time. During the interim, quantities as well as prices are affected, i.e., variations in money growth can produce business fluctuations. But changes in money growth have no lasting impact on output. Ultimately, the entire effect is on the rate of inflation.

Policy Implications of the Model Since much of the monetarist discussion of inflation tends to be strongly policy-oriented, it is appropriate to close the article with a brief mention of some of the policy implications of Laidler's model. From the point of view of the policymaker, two features of the model are of particular interest. The first feature is the time it takes for changes in the rate of money growth to work through to the rate of inflation. The second feature is the marked short-run impact of changes in money growth on real output. These features combine to produce in the model dissimilar patterns of response of output and prices to the monetary change. These response patterns have important implications for monetary stabilization policy.

First, owing to the slow response of inflation to a monetary change, it necessarily takes a long time for anti-inflationary monetary policy to work. Quick monetary remedies for inflation do not exist. Moreover, since the first effect of a change in the growth rate of money is on output and employment rather than on prices, monetary restraint would almost surely entail a recession or at least a marked retardation in the expansion of the economy. In sum, a temporary but protracted period of high unemployment and sluggish growth would have to be tolerated if monetary policy were to be successful in permanently lowering the rate of inflation.

Second, due to the difference in timing of the responses of output and prices to a monetary change, anti-inflationary monetary policy may appear impotent or, even worse, counter-productive and perverse. Because inflationary movements tend to subside so slowly, prices may continue to rise long after output and employment have turned down. Thus inflation can persist even in slack markets—a condition variously known as inflationary recession, stagflation, or slumpflation. During such periods, monetary restraint may be wrongly blamed for causing both the slump and the accompanying inflation, and the temptation may be strong to abandon prematurely the policy of monetary restraint as ineffective at best and harmful at worst.

Third, the same asymmetrical pattern of response —output first, prices only much later—may create the dangerous illusion that expansive policy in the upswing can achieve permanent gains in output and employment at the cost of very little additional inflation. This view may have unfortunate consequences. For monetarist reasoning teaches that stimulative policy can peg output and employment above their natural or equilibrium levels only by continuously accelerating the rate of inflation. In any case, time lags may well compound the problem of curbing inflation by leading to the undue prolongation of expansive policy, thus increasing the momentum behind inflation when it finally occurs. In sum, given the commitment to full employment, the tendency for output to respond quickly and prices sluggishly to both monetary ease and tightness is sufficient to bias monetary policy toward inflation over the entire policy cycle.

A fourth policy implication is that direct controls cannot permanently reduce inflation within an environment of expansionary monetary and fiscal policy. As previously mentioned, the elimination of inflation requires the eradication of inflationary expectations, since the latter are a determinant of the former. According to the model, however, the only way to dampen expectations is to create slack (excess supply) in the economy, thus causing the actual rate of inflation to fall below the expected rate, which in turn leads to a downward revision of the latter. Here direct controls are sometimes advocated as a means of speeding the fall of expectations and thus reducing the duration and severity of the recession necessary for the dampening of inflation. The idea is that controls would influence inflationary anticipations independently of the adaptive expectations mechanism described in equation (3). To be successful, however, the controls program must be supported by restrictive monetary-fiscal policy that eliminates excess demand. For, as shown in equations (2) and (3), unless excess demand is eliminated, actual inflation will lie above expected inflation leading to an upward revision of the latter. Of course controls might conceivably lower expectations by reducing the current rate of inflation itself, but only if people are convinced that the lowered rate will likely continue after the controls are lifted. It is useless to endeavor to dampen expectations via controls while simultaneously pursuing demand-expansion policies that lead inevitably to their disappointment and subsequent resurgence. In short, the elimination of excess demand is the *sine qua non* for the success of a controls program. And the excess demand problem may be compounded by the inevitable shortages created by controls.

Summary This article has expounded the principal postulates of monetarist doctrine within the context of Professor David Laidler's three-equation macroeconomic model. This model can account for the phenomena of stagflation (i.e., the persistence of inflation long after aggregate demand has slackened), for the entrenchment of inflationary expectations, for the intractability or resistance of inflation to anti-inflationary monetary policy, and, finally, for the output and employment effects of such a policy. Since the model embodies virtually all of the monetarist predictions relating to the long-run neutrality and short-run non-neutrality of monetary disturbances, it can be interpreted as capturing the essence of the monetarist view of the inflationary process. Moreover, the very simplicity of the model renders it a pedagogically useful introduction to the economics of long-run steady-state equilibrium and of short-run dynamic disequilibrium processes in which economic variables interact and evolve over time. It also provides a framework for stating clearly the public-policy issues involved in the monetarist-nonmonetarist controversy.

APPENDIX A

Derivation of the Monetary Growth Equation from the Cambridge Cash-Balance Equation

Let M be the money stock, P the price level, Y the level of real national income, and K the desired ratio of real cash balances to real income. This Cambridge K, the reciprocal of the income velocity of money, is treated as a fixed constant. These elements comprise the Cambridge cash-balance equation, $M/P = KY$. This equation is interpreted as the equilibrium solution of a three-equation demand-supply system. Specifically, the Cambridge formulation implies: (1) a relation expressing the demand for real balances as a function of income, $M_d/P = KY$; (2) an exogenously-determined nominal money supply, $M_s = M$; and (3) an equilibrium (market-clearing) condition stating that nominal money supply must equal nominal money demand, $M_s = M_d$, resulting in the Cambridge cash-balance formula, $M/P = KY$.

To transform the Cambridge formula into the money growth equation of the text, simply take the logarithm of both sides of the formula. Remembering (1) that the logarithm of a ratio is equivalent to the logarithm of its numerator minus the logarithm of its denominator, and (2) that the logarithm of the product of two terms is equal to the sum of their respective logarithms, one obtains $\log M - \log P = \log K + \log Y$. Expressing the logarithms of the variables as lower-case letters allows the preceding relation to be expressed more simply as $m - p = k + y$. Taking the first difference of this equation yields $\Delta m - \Delta p = \Delta y$, the money growth equation of the text. The first difference of k is of course zero and thus drops out of the money growth equation, i.e., Δk = zero, since k is a constant by definition.

APPENDIX B

Derivation of the Expressions for the Disequilibrium Time Paths of the Inflation Rate (Δp) and Excess Demand (x)

(I) **Derivation of the expression for Δp.**

The model in the text is

(1) $\Delta m = \Delta x + \Delta y_c + \Delta p$

(2) $\Delta p = ax_{-1} + \Delta p^e_{-1}$

(3) $\Delta p^e = b\Delta p + (1-b)\Delta p^e_{-1}$.

First, lag equation (3) one time period to get

(4) $\Delta p^e_{-1} = b\Delta p_{-1} + (1-b)\Delta p^e_{-2}$.

Next, substitute (4) into (2) to get

(5) $\Delta p = ax_{-1} + b\Delta p_{-1} + (1-b)\Delta p^e_{-2}$.

Then, rewrite (2) as

(6) $\Delta p^e_{-1} = \Delta p - ax_{-1}$.

Next, lag (6) one time period to obtain

(7) $\Delta p^e_{-2} = \Delta p_{-1} - ax_{-2}$.

Next, substitute (7) into (5) to get

(8) $\Delta p = ax_{-1} + b\Delta p_{-1} + (1-b)(\Delta p_{-1} - ax_{-2})$.

Finally, expand (8) and simplify to obtain

(9) $\Delta p = ax_{-1} - a(1-b)x_{-2} + \Delta p_{-1}$.

Equation (9) is the expression for the disequilibrium time path of the inflation rate that appears in the text. Recognizing that $\Delta p - \Delta p_{-1} = \Delta^2 p$, one can also express (9) as

(10) $\Delta^2 p = ax_{-1} - a(1-b)x_{-2}$.

(II) **Derivation of the expression for x.**

First, start with equation (1) again, i.e.,

(1) $\Delta m = \Delta x + \Delta y_c + \Delta p$.

Then, take the first difference of (1) to get

(11) $\Delta^2 m = \Delta^2 x + \Delta^2 y_c + \Delta^2 p$.

Next, expand $\Delta^2 x$ to obtain

$$\Delta^2 x = \Delta x - \Delta x_{-1} = (x - x_{-1}) - (x_{-1} - x_{-2}) = x - 2x_{-1} + x_{-2}$$

or

(12) $\Delta^2 x = x - 2x_{-1} + x_{-2}$.

Now, substitute (12) and (10) into (11) to get

(13) $\Delta^2 m = [x - 2x_{-1} + x_{-2}] + \Delta^2 y_c + [ax_{-1} - a(1-b)x_{-2}]$.

Finally, solve (13) for x and simplify to obtain

(14) $x = \Delta^2 m - \Delta^2 y_c + (2-a)x_{-1} - [1 - a(1-b)]x_{-2}$.

Equation (14) is the expression for the disequilibrium time path of the excess demand variable, as stated in the text.

APPENDIX C

Stability Conditions for Second-Order Homogeneous Difference Equations

The general homogeneous second-order difference equation $x + a_1 x_{-1} + a_2 x_{-2} = 0$ has two solutions or roots (r) which can be found by solving the quadratic *characteristic equation* $r^2 + a_1 r + a_2 = 0$ corresponding to the difference equation. Depending on the numerical values of the roots, the time path of x will move toward, away from, or around equilibrium. It is not, however, necessary to solve for the roots of the equation to determine if the system is dynamically stable, i.e., tends to converge on equilibrium either via damped-oscillatory or monotonic paths. One needs only to refer to the *stability conditions* pertaining to the difference equation. For stability, all of the following conditions must be met[1]:

$$1 + a_1 + a_2 > 0$$

$$1 - a_2 > 0$$

$$1 - a_1 + a_2 > 0.$$

In the excess demand equation of the text, the term $-(2-a)$ corresponds to the coefficient a_1 of the stability conditions and the term $[1-a(1-b)]$ corresponds to coefficient a_2. Substitution of these terms for a_1 and a_2 in the stability conditions quickly reveals that the first and second conditions are automatically satisfied as long as $a > 0$ and $0 < b < 1$, the range of values specified in the model of the text. The third stability condition will be satisfied if $a > [4/(2-b)]$.

To determine whether the stable path is oscillatory or monotonic, one must analyze the characteristic roots of the system. The roots of the characteristic equation $r^2 + a_1 r + a_2 = 0$ are

$$r_{1,2} = \frac{-a_1 \pm \sqrt{a^2_1 - 4a_2}}{2}$$

where $a_1 = -(2-a)$ and $a_2 = [1 - a(1-b)]$. The system will exhibit oscillatory behavior if the roots are *complex*, i.e., if $4a_2 > a^2_1$, or in terms of the model, if $[4 - 4a(1-b)] > [-(2-a)]^2$. The latter inequality reduces to $a^2 < 4ab$, hence oscillatory behavior is obviously possible for $a > 0$ and $0 < b < 1$.

[1] Paul A. Samuelson, *Foundations of Economic Analysis*, Cambridge: Harvard University Press, 1947, p. 436.