

# THEORETICAL ANALYSIS OF THE DEMAND FOR MONEY

*Bennett T. McCallum and Marvin S. Goodfriend\**

In any discussion of the demand for money it is important to be clear about the concept of money that is being utilized; otherwise, misunderstandings can arise because of the various possible meanings that readers could have in mind. Here the term will be taken to refer to an economy's *medium of exchange*; that is, to a tangible asset that is generally accepted in payment for any commodity. Money thus conceived will also serve as a store of value, of course, but may be of minor importance to the economy in that capacity. The monetary asset will usually also serve as the economy's medium of account—i.e., prices will be quoted in terms of money—since additional accounting costs would be incurred if the unit of account were a quantity of some asset other than money. The medium-of-account role is, however, not logically tied to the medium of exchange (Wicksell, 1906; Niehans, 1978).

Throughout much of Western history, most economies have adopted as their principal medium of exchange a commodity that would be valuable even if it were not used as money. Recently, however, fiat money—intrinsically worthless tokens made of paper or some other cheap material—has come to predominate. Under a commodity money arrangement, the exchange value of money will depend upon the demand for the monetary commodity in its non-monetary, as well as its monetary, uses. But in a discussion of money demand, as distinct from a discussion of the price level, any possible non-monetary demand for the medium of exchange—which will be absent anyhow in a fiat money system—can legitimately be ignored.

The quantity of money demanded in any economy—indeed, the set of assets that have monetary status—will be dependent upon prevailing institutions, regulations, and technology. Technical progress in the payments industry will, for instance, tend to alter the quantity of money demanded for given values of determinants such as income. This dependence does not, however, imply that the demand for money is a nebulous or unusable concept, any more than the existence of technical progress and regulatory change in the transportation industry does so for the demand for automobiles. In practice, some lack of clarity pertains to the operational measurement of the money stock, as it does to the stock of automobiles or other commodities. But in an economy with a well-established national currency, the principle is relatively clear: assets are part of the money stock if and only if they constitute *claims* to currency, unrestricted legal claims that can be promptly and cheaply exercised (at par). This principle rationalizes the common practice of including demand deposits in the money stock of the United States, while excluding time deposits and various other assets.

The rapid development during the 1960s and 1970s of computer and telecommunications technologies has led some writers (e.g., Fama, 1980) to contemplate economies—anticipated by Wicksell (1906)—in which virtually all purchases are effected not by the transfer of a tangible medium of exchange, but by means of signals to an accounting network—signals that result in appropriate debits and credits to the wealth accounts of buyers and sellers. If there were literally *no* medium of exchange, the wealth accounts being claims to some specified bundle of commodities, the economy in question would be properly regarded and analyzed as a nonmonetary economy—albeit one that avoids the inefficiencies of crude barter. If, by contrast, the accounting network's credits were claims to quantities of a fiat or commodity medium of exchange, then individuals' credit balances would appropriately be included as part of the money stock (McCallum, 1985).

---

\* Bennett T. McCallum is H. J. Heinz Professor of Economics, Carnegie-Mellon University, and Research Advisor, Federal Reserve Bank of Richmond. Marvin S. Goodfriend is Economist and Vice President, Federal Reserve Bank of Richmond.

This article is reprinted with permission from *The New Palgrave: A Dictionary of Economics*, edited by John Eatwell, Peter Newman, and Murray Milgate. 4 volumes. London: The Macmillan Press; New York: Stockton Press, 1987.

## Basic Principles

An overview of the basic principles of money demand theory can be obtained by considering a hypothetical household that seeks at time  $t$  to maximize

$$(1) \quad u(c_t, l_t) + \beta u(c_{t+1}, l_{t+1}) + \beta^2 u(c_{t+2}, l_{t+2}) + \dots$$

where  $c_t$  and  $l_t$  are the household's consumption and leisure during  $t$  and where  $\beta = 1/(1+\delta)$ , with  $\delta > 0$  the rate of time preference. The within-period utility function  $u(\cdot, \cdot)$  is taken to be well-behaved so that unique positive values will be chosen for  $c_t$  and  $l_t$ . The household has access to a productive technology described by a production function that is homogeneous of degree one in capital and labor inputs. But for simplicity we assume that labor is supplied inelastically, so this function can be written as  $y_t = f(k_{t-1})$ , where  $y_t$  is production during  $t$  and  $k_{t-1}$  is the stock of capital held at the end of period  $t-1$ . The function  $f(\cdot)$  is well-behaved, so a unique positive value of  $k_t$  will be chosen for the upcoming period. Capital is unconsumed output, so its price is the same as that of the consumption good and its rate of return between  $t$  and  $t+1$  is  $f'(k_t)$ .

Although this setup explicitly recognizes the existence of only one good, it is intended to serve as a simplified representation—one formally justified by the analysis of Lucas (1980)—of an economy in which the household sells its specialized output and makes purchases (at constant relative prices) of a large number of distinct consumption goods. Carrying out these purchases requires *shopping time*,  $s_t$ , which subtracts from leisure:  $l_t = 1 - s_t$ , where units are chosen so that there is one unit of time per period available for shopping and leisure together. (If labor were elastically supplied, then labor time would have to be included in the expression.) In a monetary economy, however, the amount of shopping time required for a given amount of consumption will depend negatively upon the quantity of real money balances held by the household (up to some satiation level). For concreteness, we assume that

$$(2) \quad s_t = \psi(c_t, m_t)$$

where  $\psi(\cdot, \cdot)$  has partial derivatives  $\psi_1 > 0$  and  $\psi_2 \leq 0$ . In (2),  $m_t = M_t/P_t$ , where  $M_t$  is the nominal stock of money held at the end of  $t$  and  $P_t$  is the money price of a consumption bundle. (A variant with  $M_t$  denoting the start-of-period money stock will be mentioned below.) The transaction variable is here specified as  $c_t$  rather than  $c_t + \Delta k_t$  to reflect the idea

that only a few distinct capital goods will be utilized, so that the transaction-cost to expenditure ratio will be much lower than for consumption goods.

Besides capital and money, there is a third asset available to the household. This asset is a nominal bond, i.e., a one-period security that may be purchased at the price  $1/(1+R_t)$  in period  $t$  and redeemed for one unit of money in  $t+1$ . The symbol  $B_t$  will be used to denote the number (possibly negative) of these securities purchased by the household in period  $t$ , while  $b_t = B_t/P_t$ .

In the setting described, the household's budget constraint for period  $t$  may be written as follows:

$$(3) \quad f(k_{t-1}) + v_t \geq c_t + k_t - k_{t-1} + m_t - (1+\pi_t)^{-1}m_{t-1} + (1+R_t)^{-1}b_t - (1+\pi_t)^{-1}b_{t-1}.$$

Here  $v_t$  is the real value of lump-sum transfers (net of taxes) from the government while  $\pi_t$  is the inflation rate,  $\pi_t = (P_t - P_{t-1})/P_{t-1}$ . Given the objective of maximizing (1), first-order conditions necessary for optimality of the household's choices include the following, in which  $\phi_t$  and  $\lambda_t$  are Lagrangian multipliers associated with the constraints (2) and (3), respectively:

$$(4) \quad u_1(c_t, 1-s_t) - \phi_t \psi_1(c_t, m_t) - \lambda_t = 0$$

$$(5) \quad -u_2(c_t, 1-s_t) + \phi_t = 0$$

$$(6) \quad -\phi_t \psi_2(c_t, m_t) - \lambda_t + \beta \lambda_{t+1} (1+\pi_{t+1})^{-1} = 0$$

$$(7) \quad -\lambda_t + \beta \lambda_{t+1} [f'(k_t) + 1] = 0$$

$$(8) \quad -\lambda_t (1+R_t)^{-1} + \beta \lambda_{t+1} (1+\pi_{t+1})^{-1} = 0.$$

These conditions, together with the constraints (2) and (3), determine current and planned values of  $c_t$ ,  $s_t$ ,  $m_t$ ,  $k_t$ ,  $b_t$ ,  $\phi_t$ , and  $\lambda_t$  for given time paths of  $v_t$ ,  $R_t$ , and  $\pi_t$  (which are exogenous to the household) and the predetermined values of  $k_{t-1}$ ,  $m_{t-1}$ , and  $b_{t-1}$ . (There is also a relevant transversality condition, but it can be ignored for the issues at hand.) Also,  $l_t$  values can be obtained from  $l_t = 1 - s_t$  and, with  $P_{t-1}$  given,  $P_t$ ,  $M_t$ , and  $B_t$  values are implied by the  $\pi_t$ ,  $m_t$ , and  $b_t$  sequences.

The household's optimizing choice of  $m_t$  can be described in terms of two distinct concepts of a money demand function. The first of these is a proper demand function, that is, a relationship giving the chosen quantity as a function of variables that are either predetermined or exogenous to the

economic unit in question. In the present context, the money demand function of that type will be of the form

$$(9) \quad m_t = \mu(k_{t-1}, m_{t-1}, b_{t-1}, v_t, v_{t+1}, \dots, R_t, R_{t+1}, \dots, \pi_t, \pi_{t+1}, \dots)$$

where the variables dated  $t+1, t+2, \dots$  must be understood as anticipated values. Now, it will be obvious that this relationship does not closely resemble those normally described in the literature as "money demand functions." There is a second type of relationship implied by the model, however, that does have such a resemblance. To obtain this second expression, one can eliminate  $\beta\lambda_{t+1}(1 + \pi_{t+1})^{-1}$  between equations (6) and (8), then eliminate  $\lambda_t$  and finally  $\phi_t$  from the resultant by using (4) and (5). These steps yield the following:

$$(10) \quad -u_2(c_t, 1 - s_t)\psi_2(c_t, m_t) = [u_1(c_t, 1 - s_t) - u_2(c_t, 1 - s_t)\psi_1(c_t, m_t)] [1 - (1 + R_t)^{-1}]$$

Then  $\psi(c_t, m_t)$  can be used in place of  $s_t$  and the result is a relationship that involves *only*  $m_t, c_t,$  and  $R_t$ . Consequently, (10) can be expressed in the form

$$(11) \quad h(m_t, c_t, R_t) = 0$$

and if the latter is solvable for  $m_t$  one can obtain

$$(12) \quad M_t/P_t = L(c_t, R_t)$$

Thus the model at hand yields a *portfolio-balance* relationship between real money balances demanded, a variable measuring the volume of transactions conducted, and the nominal interest rate (which reflects the cost of holding money rather than bonds). It can be shown, moreover, that for reasonable specifications of the utility and shopping-time functions,  $L(\cdot, \cdot)$  will be increasing in its first argument and decreasing in the second.

There are, of course, two problems in moving from a demand function (of either type) for an individual household to one that pertains to the economy as a whole. The first of these involves the usual problem of aggregating over households that may have different tastes and/or levels of wealth. It is well known that the conditions permitting such aggregation are extremely stringent in the context of any sort of behavioral relation; but for many theoretical purposes it is sensible to pretend that they are satisfied. The second problem concerns the existence of economic units other than households—"firms"

being the most obvious example. To construct a model analogous to that above for a firm, one would presumably posit maximization of the present value of real net receipts rather than (1), and the constraints would be different. In particular, the shopping-time function (2) would need to be replaced with a more general relationship depicting resources used in conducting transactions as a function of their volume and the real quantity of money held. The transaction measure would not be  $c_t$  for firms or, therefore, for the economy as a whole. But the general aspects of the analysis would be similar, so we shall proceed under the presumption that the crucial issues are adequately represented in a setting that recognizes only economic units like the "households" described above.

The distinction between the proper money demand function (9) and the more standard portfolio-balance relation (12) is important in the context of certain issues. As an example, consider the issue of whether wealth or income should appear as a "scale variable" (Meltzer, 1963). From the foregoing, it is clear that wealth is an important determinant of money demand in the sense that  $k_{t-1}, m_{t-1},$  and  $b_{t-1}$  are arguments of the demand function (9). Nevertheless, formulation (12) indicates that there is no separate role for wealth in a portfolio-balance relation if appropriate transaction and opportunity-cost variables are included.

An issue that naturally arises concerns the foregoing discussion's neglect of randomness. How would the analysis be affected if it were recognized that future values of variables cannot possibly be known with certainty? In answer, let us suppose that the household knows current values of all relevant variables including  $P_t, R_t,$  and  $v_t$  when making decisions on  $m_t$  and  $c_t$ , but that its views concerning variables dated  $t+1, t+2, \dots$  are held in the form of nondegenerate probability distributions. Suppose also that there is uncertainty in production, so that the marginal product of capital in  $t+1, f'(k_t),$  is viewed as random. Then the household's problem becomes one of maximizing the expectation of (1), with  $u(\cdot, \cdot)$  a Von Neumann-Morgenstern utility function, given information available in period  $t$ . Consequently, the first-order conditions (4)-(8) must be replaced with ones that involve conditional expectations. For example, equation (7) would be replaced with

$$(7') \quad -\lambda_t + \beta E_t\{\lambda_{t+1}[f'(k_t) + 1]\} = 0$$

where  $E_t(\cdot)$  denotes the expectation of the indicated variable conditional upon known values of  $P_t, R_t, v_t,$  and so on. With this modification, the nature of the

proper demand function becomes much more complex—indeed, for most specifications no closed form solution analogous to (9) will exist. Nevertheless, the portfolio-balance relation (12) will continue to hold exactly as before, for the steps described in its derivation above remain the same except that it is  $E_t[\beta\lambda_{t+1}(1 + \pi_{t+1})^{-1}]$  that is eliminated between equations corresponding to (6) and (8). From this result it follows that, according to our model, the relationship of  $M_t/P_t$  to the transaction and opportunity-cost variables is invariant to changes in the probability distribution of future variables.

Another specification variant that should be mentioned reflects the assumption that it is money held at the start of a period, not its end, that facilitates transactions conducted during the period. If that change in specification were made and the foregoing analysis repeated, it would be found that the household's concern in period  $t$  would be to have the appropriate level of real money balances at the start of period  $t+1$ . The portfolio-balance relation analogous to (12) that would be obtained in the deterministic case would relate  $m_{t+1}$  to  $c_{t+1}$  and  $R_t$ , where  $m_{t+1} = M_{t+1}/P_{t+1}$  with  $M_{t+1}$  reflecting money holdings at the end of period  $t$ . Consequently,  $M_{t+1}/P_t$  would be related to  $R_t$ , planned  $c_{t+1}$ , and  $P_t/P_{t+1}$ . Thus the theory does not work out as cleanly as in the case considered above even in the absence of randomness, and is complicated further by the recognition of the latter. The fundamental nature of the relationships are, however, the same as above.

Another point deserving of mention is that if labor is supplied elastically, the portfolio-balance relation analogous to (12) will include the real wage-rate as an additional argument. This has been noted by Karni (1973) and Dutton and Gramm (1973). More generally, the existence of other relevant margins of substitution can bring in other variables. If stocks of commodities held by households affect shopping-time requirements, for example, the inflation rate will appear separately in the counterpart of (12)—see Feige and Parkin (1971).

Finally, it must be recognized that the simplicity of the portfolio-balance relation (12) would be lost if the intertemporal utility function (1) were not time-separable. If, for example, the function  $u(c_t, l_t)$  in (1) were replaced with  $u(c_t, l_t, l_{t-1})$  or  $u(c_t, c_{t-1}, l_t)$ , as has been suggested in the business cycle literature, then the dynamic aspect of the household's choices would be more complex and a relation like (12)—i.e., one that includes only contemporaneous variables—could not be derived.

## Historical Development

The approach to money demand analysis outlined above, which features intertemporal optimization choices by individual economic agents whose transactions are facilitated by their holdings of money, has evolved gradually over time. In this section we briefly review that evolution.

While the earlier literature on the quantity theory of money contained many important insights, its emphasis was on the comparison of market equilibria rather than individual choice, i.e., on “market experiments” rather than “individual experiments,” in the language of Patinkin (1956). Consequently, there was little explicit consideration of money demand behavior in pre-1900 writings in the quantity theory tradition. Indeed, there was little emphasis on money demand *per se* even in the classic contributions of Mill (1848), Wicksell (1906), and Fisher (1911), despite the clear recognition by those analysts that some particular quantity of real money holdings would be desired by the inhabitants of an economy under any specified set of circumstances. Notable exceptions, discussed by Patinkin (1956, pp. 386-417), were provided by Walras and Schlesinger.

In the English language literature, the notion of money demand came forth more strongly in the “cash balance” approach of Cambridge economists, an approach that featured analysis organized around the concepts of money demand and supply. This organizing principle was present in the early (c. 1871) but unpublished writings of Marshall (see Whitaker, 1975, pp. 165-68) and was laid out with great explicitness by Pigou (1917). The Cambridge approach presumed that the quantity of money demanded would depend primarily on the volume of transactions to be undertaken, but emphasized volition on the part of moneyholders and recognized (sporadically) that the ratio of real balances to transaction volume would be affected by foregone “investment income”, i.e., interest earnings. In this regard Cannan (1921), a non-Cambridge economist who was influenced by Marshall, noted that the quantity of money demanded should be negatively related to anticipated inflation—an insight previously expressed by Marshall in his testimony of 1886 for the Royal Commission on the Depression of Trade and Industry (Marshall, 1926). In addition, Cannan developed very clearly the point that the relevant concept is the demand for a *stock* of money.

Although the aforementioned theorists developed several important constituents of a satisfactory money demand theory, none of them unambiguously cast his explanation in terms of marginal analysis. Thus a significant advance was provided by Lavington

(1921, p. 30), in a chapter entitled "The Demand for Money," who attempted a statement of the marginal conditions that must be satisfied for optimality by an individual who consumes, holds money, and holds interest-bearing securities. But despite the merits of his attempt, Lavington confused—as Patinkin (1956, p. 418) points out—the subjective sacrifice of permanently adding a dollar to cash balances with that of adding it for only one period. Thus it was left for Fisher (1930, p. 216) to provide a related but correct statement. The discussions of both Lavington and Fisher are notable for identifying the interest rate as a key determinant of the marginal opportunity cost of holding money.

In a justly famous article, Hicks (1935) argued persuasively that progress in the theory of money would require the treatment of money demand as a problem of individual choice at the margin. Building upon some insightful but unclear suggestions in Keynes's *Treatise on Money* (1930), Hicks investigated an agent's decision concerning the relative amounts of money and securities to be held at a point in time. He emphasized the need to explain why individuals willingly hold money when its return is exceeded by those available from other assets and—following Lavington and Fisher—concluded that money provides a service yield not offered by other assets. Hicks also noted that the positive transaction cost of investing in securities makes it unprofitable to undertake such investments for very short periods. Besides identifying the key aspects of marginal analysis of money demand, Hicks (1935) pointed out that an individual's total wealth will influence his demand for money. All of these points were developed further in Chapters 13 and 19 of Hicks's *Value and Capital* (1939). The analysis in the latter is, some misleading statements about the nature of interest notwithstanding, substantively very close to that outlined in the previous section of this article. Hicks did not, however, provide formal conditions relating to money demand in his mathematical appendix.

The period between 1935 and 1939 witnessed, of course, the publication of Keynes's *General Theory* (1936). That work emphasized the importance for macroeconomic analysis of the interest-sensitivity of money demand—"liquidity preference," in Keynes's terminology—and was in that respect, as in many others, enormously influential. Its treatment of money demand *per se* was not highly original, however, in terms of fundamentals. (This statement ignores some peculiarities resulting from a presumably inadvertent attribution of money illusion; on this topic, again see Patinkin, 1956, pp. 173-74).

The importance of several items mentioned above—payments practices, foregone interest, and transaction costs—was explicitly depicted in the formal optimization models developed several years later by Baumol (1952) and Tobin (1956). These models, which were suggested by mathematical inventory theory, assume the presence of two assets (money and an interest-bearing security), a fixed cost of making transfers between money and the security, and a lack of synchronization between (exogenously given) receipt and expenditure streams. In addition, they assume that all payments are made with money. Economic units are depicted as choosing the optimal frequency for money-security transfers so as to maximize interest earnings net of transaction costs.

In Baumol's treatment, which ignores integer constraints on the number of transactions per period, the income and interest-rate elasticities of real money demand are found to be  $\frac{1}{2}$  and  $-\frac{1}{2}$ , respectively. Thus the model implies "economies of scale" in making transactions. Tobin's (1956) analysis takes account of integer constraints, by contrast, and thus implies that individuals respond in a discontinuous fashion to alternative values of the interest rate. In his model it appears entirely possible for individual economic units to choose corner solutions in which none of the interest-bearing security is held. A number of extensions of the Baumol-Tobin approach have been made by various authors; for an insightful survey the reader is referred to Barro and Fischer (1976).

Miller and Orr (1966) pioneered the inventory approach to money demand theory in a stochastic context. Specifically, in their analysis a firm's net cash inflow is generated as a random walk, and the firm chooses a policy to minimize the sum of transaction and foregone interest costs. The optimal decision rule is of the (S,s) type: when money balances reach zero or a ceiling, S, the firm makes transactions to return the balance to the level s. In this setting there are again predicted economies of scale, while the interest-rate elasticity is  $-\frac{1}{3}$ . For extensions the reader is again referred to Barro and Fischer (1976).

The various inventory models of money demand possess the desirable feature of providing an explicit depiction of the *source* of money's service yield to an individual holder. It has been noted, e.g., by Friedman and Schwartz (1970), that the type of transaction demand described by these models is unable to account for more than a fraction of the transaction balances held in actual economies. Furthermore, their treatment of expenditure and receipt streams

as exogenous is unfortunate and they do not generalize easily to fully dynamic settings. These points imply, however, only that the inventory models should not be interpreted too literally. In terms of fundamentals they are closely related to the basic model outlined in the previous section.

A quite different approach was put forth by Tobin (1958), in a paper that views the demand for money as arising from a portfolio allocation decision made under conditions of uncertainty. In the more influential of the paper's models, the individual wealth-holder must allocate his portfolio between a riskless asset, identified as money, and an asset with an uncertain return whose expected value exceeds that of money. Tobin shows how the optimal portfolio mix depends, under the assumption of expected utility maximization, on the individual's degree of risk aversion, his wealth, and the mean-variance characteristics of the risky asset's return distribution. The analysis implies a negative interest sensitivity of money demand, thereby satisfying Tobin's desire to provide an additional rationalization of Keynes's (1936) liquidity preference hypothesis. The approach has, however, two shortcomings. First, in actuality money does not have a yield that is riskless in real terms, which is the relevant concept for rational individuals. Second, and more seriously, in many actual economies there exist assets "that have precisely the same risk characteristics as money and yield higher returns" (Barro and Fischer, 1976, p. 139). Under such conditions, the model implies that no money will be held.

Another influential item from this period was provided by Friedman's well-known "restatement" of the quantity theory (1956). In that paper, as in Tobin's, the principle role of money is as a form of wealth. Friedman's analysis emphasized margins of substitution between money and assets other than bonds—e.g., durable consumption goods and equities. The main contribution of the paper was to help rekindle interest in monetary analysis from a macroeconomic perspective, however, rather than to advance the formal theory of money demand.

A model that may be viewed as a formalization of Hicks's (1935, 1939) approach was outlined by Sidrauski (1967). The main purpose of Sidrauski's paper was to study the interaction of inflation and capital accumulation in a dynamic context, but his analysis gives rise to optimality conditions much like those of equations (4)-(8) of the present article and thus implies money demand functions like (9) and (12). The main difference between Sidrauski's model and ours is merely due to our use of the "shopping-time" specification, which was suggested by Saving (1971). That feature makes real balances an argu-

ment of each individual's utility function only indirectly, rather than directly, and indicates the type of phenomenon that advocates of the direct approach presumably have in mind. Thus Sidrauski's implied money demand model is the basis for the one presented above, while a stochastic version of the latter, being fundamentally similar to inventory or direct utility-yield specifications, is broadly representative of current mainstream views.

### Ongoing Controversies

Having outlined the current mainstream approach to money demand analysis and its evolution, we now turn to matters that continue to be controversial. The first of these concerns the role of uncertainty. In that regard, one point has already been developed, i.e., that rate-of-return uncertainty on other assets cannot be used to explain why individuals hold money in economies—such as that of the United States—in which there exist very short-term assets that yield positive interest and are essentially riskless in nominal terms. But this does not imply that uncertainty is unimportant for money demand in a more general sense, for there are various ways in which it can affect the analysis. In the basic model outlined above, uncertainty appears explicitly only by way of the assumption that households view asset returns as random. In that case, if money demand and consumption decisions for a period are made simultaneously then the portfolio-balance relation (12) will be—as shown above—invariant to changes in the return distributions. But the same is not true for the proper demand function (9). And the arguments  $c_t$  and  $R_t$  of (12) will themselves be affected by the extent of uncertainty, for it will affect households' saving—as well as portfolio—decisions. The former, of course, impact not only on  $c_t$  but also on the economy's capital stock and thus, via the equilibrium real return on capital, on  $R_t$ . In addition, because  $R_t$  is set in nominal terms, its level will include a risk premium for inflation uncertainty (Fama and Farber, 1979).

Furthermore, the invariance of (12) to uncertainty breaks down if money must be held at the start of a period to yield its transaction services during that period. In this case, the money demand decision temporally precedes the related consumption decision so the marginal service yield of money is random, with moments that depend on the covariance matrix of forecast errors for consumption and the price level. Thus the extent of uncertainty, as reflected in this covariance matrix, influences the quantity of real balances demanded in relation to  $R_t$  and plans for  $c_{t+1}$ .

There is, moreover, another type of uncertainty that is even more fundamental than rate-of-return randomness. In particular, the existence of uncertainty regarding exchange opportunities available at an extremely fine level of temporal and spatial disaggregation—uncertainties regarding the “double coincidence of wants” in meetings with potential exchange partners—provides the basic *raison d'être* for a medium of exchange. In addition, the ready verifiability of money enhances the efficiency of the exchange process by permitting individuals to economize on the production of information when there is uncertainty about the reputation of potential trading partners. Thus uncertainty is crucial in explaining why it is that money holdings help to facilitate transactions—to save “shopping time” in our formalization. In this way randomness is critically involved, even when it does not appear explicitly in the analysis. (Alternative treatments of uncertainty in the exchange process have been provided by Patinkin, 1956; Brunner and Meltzer, 1971; and King and Plosser, 1986).

An important concern of macroeconomists in recent years has been to specify models in terms of genuinely structural relationships, i.e., ones that are invariant to policy changes. This desire has led to increased emphasis on explicit analysis of individuals' dynamic optimization problems, with these expressed in terms of basic taste and technology parameters. Analysis of that type is especially problematical in the area of money demand, however, because of the difficulty of specifying rigorously the precise way—at a “deeper” level than (2), for example—in which money facilitates the exchange process. One prominent attempt to surmount this difficulty has featured the application of a class of overlapping-generations models—i.e., dynamic equilibrium models that emphasize the differing perspectives on saving of young and old individuals—to a variety of problems in monetary economics. The particular class of overlapping-generations models in question is one in which, while there is an analytical entity termed “fiat money,” the specification deliberately excludes any shopping time or related feature that would represent the transaction-facilitating aspect of money. Thus this approach, promoted most prominently in the work of Wallace (1980), tries to surmount the difficulty of modeling the medium-of-exchange function of money by simply ignoring it, emphasizing instead the asset's function as a store of value.

Models developed under this overlapping-generations approach typically possess highly distinctive implications, of which three particularly striking examples will be mentioned. First, if the

monetary authority causes the stock of money to grow at a rate in excess of the economy's rate of output growth, no money will be demanded and the price level will be infinite. Second, steady-state equilibria in which money is valued will be Pareto optimal if and only if the growth rate of the money stock is nonpositive. Third, open-market changes in the money stock will have no effect on the price level. It has been shown, however, that these implications result from the models' neglect of the medium-of-exchange function of money. Specifically, McCallum (1983) demonstrates that all three implications vanish if this neglect is remedied by recognition of shopping-time considerations as above. That conclusion suggests that the class of overlapping-generations models under discussion provides a seriously misleading framework for the analysis of monetary issues. This weakness, it should be added, results not from the generational structure of these models, but from the overly restrictive application of the principle that assets are valued solely on the basis of the returns that they yield; in particular, the models fail to reflect the nonpecuniary return provided by holdings of the medium of exchange. On these points also see Tobin (1980).

Recognizing this problem but desiring to avoid specifications like (2), some researchers have been attracted to the use of models incorporating a *cash-in-advance* constraint (e.g., Lucas, 1980; Svensson, 1985). In these models, it is assumed that an individual's purchases in any period cannot exceed the quantity of money brought into that period. Clearly, imposition of this type of constraint gives a medium-of-exchange role to the model's monetary asset and thereby avoids the problems of the Wallace-style overlapping-generations models. Whether it does so in a satisfactory manner is, however, more doubtful. In particular, the cash-in-advance formulation implies that start-of-period money holdings place a *strict* upper limit on purchases during the period. This is a considerably more stringent notion than that implied by (2), which is that such purchases are possible but increasingly expensive in terms of time and/or other resources. Thus the demand for money will tend to be less sensitive to interest-rate changes with the cash-in-advance specification than with one that ties consumption and money holdings together less rigidly. More generally, the cash-in-advance specification can be viewed as an extreme special case of the shopping-time function described in (2), in much the same way as a fixed-coefficient production function is a special case of a more general neo-classical technology. For some issues, use of the special case specification will be convenient and not



misleading, but care must be exerted to avoid inappropriate applications. It seems entirely unwarranted, moreover, to opt for the cash-in-advance specification in the hope that it will be more nearly structural and less open to the Lucas critique (1976) than relations such as (2). Both of these specificational devices—and probably any that will be analytically tractable in a macroeconomic context—should be viewed not as literal depictions of technological or social constraints, but as potentially useful metaphors that permit the analyst to recognize in a rough way the benefits of monetary exchange. (On the general topic, see Fischer, 1974).

A final controversy that deserves brief mention pertains to an aspect of money demand theory that has not been formally discussed above, but which is of considerable importance in practical applications. Typically, econometric estimates of money demand functions combine “long-run” specifications such as (12) with a *partial adjustment* process that relates actual money holdings to the implied “long-run” values. Operationally, this approach often results in a regression equation that includes a lagged value of the money stock as an explanatory variable. (Distributed-lag formulations are analytically similar.) Adoption of the partial adjustment mechanism is justified by appeal to portfolio-adjustment costs. Specifically, some authors argue that money balances serve as a “buffer stock” that temporarily accommodates unexpected variations in income, while others attribute sluggish adjustments to search costs.

From the theoretical perspective, however, the

foregoing interpretation for the role of lagged money balances (or distributed lags) appears weak. It is difficult to believe that tangible adjustment costs are significant, and in their absence there is no role for lagged money balances, in formulations such as (12), when appropriate transaction and opportunity-cost variables are included. Furthermore, typical estimates suggest adjustment speeds that are too slow to be plausible.

These points have been stressed by Goodfriend (1985), who offers an alternative explanation for the relevant empirical findings. A model in which there is full contemporaneous adjustment of money holdings to transaction and opportunity-cost variables is shown to imply a positive coefficient on lagged money when these determinants are positively autocorrelated and contaminated with measurement error. Under this interpretation, the lagged variable is devoid of behavioral significance; it enters the regression only because it helps to explain the dependent variable in a mongrel equation that mixes together relations pertaining to money demand and other aspects of behavior. (This particular conclusion is shared with the buffer stock approach described by Laidler (1984), which interprets the conventional regression as a confounding of money demand with sluggish price-adjustment behavior.) Furthermore, the measurement error hypothesis can account for positive autocorrelation of residuals in the conventional regression and, if measurement errors are serially correlated, the *magnitude* of the lagged money coefficient typically found in practice.

## References

- Barro, R. J., and S. Fischer. “Recent Developments in Monetary Theory.” *Journal of Monetary Economics* 2 (April 1976): 133-67.
- Baumol, W. J. “The Transactions Demand for Cash: An Inventory Theoretic Approach.” *Quarterly Journal of Economics* 66 (November 1952): 545-56.
- Brunner, K., and A. H. Meltzer. “The Uses of Money: Money in the Theory of an Exchange Economy.” *American Economic Review* 61 (December 1971): 784-805.
- Cannan, E. “The Application of the Theoretical Apparatus of Supply and Demand to Units of Currency.” *Economic Journal* 31 (December 1921): 453-61.
- Dutton, D. S., and W. P. Gramm. “Transactions Costs, the Wage Rate, and the Demand for Money.” *American Economic Review* 63 (September 1973): 652-65.
- Fama, E. F. “Banking in the Theory of Finance.” *Journal of Monetary Economics* 6 (January 1980): 39-57.
- Fama, E. F., and A. Farber. “Money, Bonds, and Foreign Exchange.” *American Economic Review* 69 (September 1979): 639-49.
- Feige, E., and M. Parkin. “The Optimal Quantity of Money, Bonds, Commodity Inventories, and Capital.” *American Economic Review* 61 (June 1971): 335-49.
- Fischer, S. “Money and the Production Function.” *Economic Inquiry* 12 (December 1974): 517-33.
- Fisher, I. *The Purchasing Power of Money*. New York: Macmillan, 1911.
- . *The Theory of Interest*. New York: Macmillan, 1930.
- Friedman, M. “The Quantity Theory of Money: A Restatement.” In *Studies in the Quantity Theory of Money*, edited by M. Friedman. Chicago: University of Chicago Press, 1956.
- Friedman, M., and A. J. Schwartz. *Monetary Statistics of the United States*. New York: National Bureau of Economic Research, 1970.



- Goodfriend, M. "Reinterpreting Money Demand Regressions." *Carnegie-Rochester Conference Series on Public Policy* 22 (Spring 1985): 207-42.
- Hicks, J. R. "A Suggestion for Simplifying the Theory of Money." *Economica* 2 (February 1935): 1-19.
- . *Value and Capital*. London: Oxford University Press, 1939.
- Karni, E. "The Transactions Demand for Cash: Incorporation of the Value of Time into the Inventory Approach." *Journal of Political Economy* 81 (September/October 1973): 1216-25.
- Keynes, J. M. *A Treatise on Money*, 2 vols. London: Macmillan, 1930.
- . *The General Theory of Employment, Interest, and Money*. London: Macmillan, 1936.
- King, R. G., and C. I. Plosser. "Money as the Mechanism of Exchange." *Journal of Monetary Economics* 17 (January 1986): 93-115.
- Laidler, D. "The 'Buffer Stock' Notion in Monetary Economics." *Economic Journal* 94 (Supplement 1984): 17-34.
- Lavington, F. *The English Capital Market*. London: Methuen, 1921.
- Lucas, R. E., Jr. "Econometric Policy Evaluation: A Critique." *Carnegie-Rochester Conference Series on Public Policy* 1 (Autumn 1976): 19-46.
- . "Equilibrium in a Pure Currency Economy." In *Models of Monetary Economics*, edited by J.H. Kareken and N. Wallace. Minneapolis: Federal Reserve Bank of Minneapolis, 1980.
- Marshall, A. *Official Papers by Alfred Marshall*, edited by J.M. Keynes. London: Macmillan, 1926.
- McCallum, B. T. "The Role of Overlapping-Generations Models in Monetary Economics." *Carnegie-Rochester Conference Series on Public Policy* 18 (Spring 1983): 9-44.
- . "Bank Deregulation, Accounting Systems of Exchange, and the Unit of Account: A Critical Review." *Carnegie-Rochester Conference Series on Public Policy* 23 (Autumn 1985): 13-45.
- Meltzer, A. H. "The Demand for Money: The Evidence From the Time Series." *Journal of Political Economy* 71 (June 1963): 219-46.
- Mill, J. S. *Principles of Political Economy*, 2 vols. London: John W. Parker, 1848.
- Miller, M. H., and D. Orr. "A Model of the Demand for Money by Firms." *Quarterly Journal of Economics* 80 (August 1966): 413-35.
- Niehans, J. *The Theory of Money*. Baltimore: Johns Hopkins University Press, 1978.
- Patinkin, D. *Money, Interest, and Prices*. New York: Harper and Row, 1956.
- Pigou, A. C. "The Value of Money." *Quarterly Journal of Economics* 32 (November 1917): 38-65.
- Saving, T. R. "Transactions Costs and the Demand for Money." *American Economic Review* 61 (June 1971): 407-20.
- Sidrauski, M. "Rational Choice and Patterns of Growth in a Monetary Economy." *American Economic Review* 57 (May 1967): 534-44.
- Svensson, L. E. O. "Money and Asset Prices in a Cash-in-Advance Economy." *Journal of Political Economy* 93 (October 1985): 919-44.
- Tobin, J. "The Interest-Elasticity of Transaction Demand for Cash." *Review of Economics and Statistics* 38 (August 1956): 241-47.
- . "Liquidity Preference as Behavior Toward Risk." *Review of Economic Studies* 25 (February 1958): 65-86.
- . "Discussion." In *Models of Monetary Economics*, edited by J.H. Kareken and N. Wallace. Minneapolis: Federal Reserve Bank of Minneapolis, 1980.
- Wallace, N. "The Overlapping Generations Model of Fiat Money." In *Models of Monetary Economics*, edited by J.H. Kareken and N. Wallace. Minneapolis: Federal Reserve Bank of Minneapolis, 1980.
- Whitaker, J. K., ed. *The Early Economic Writings of Alfred Marshall, 1867-1890*, 2 vols. New York: Free Press, 1975.
- Wicksell, K. *Lectures on Political Economy*, Vol. 2 (1906). Translated by E. Classen. London: Routledge & Kegan Paul, 1935.