From:
Monetary Economics: Theory + Policy by Bennett T. McCallum
(MacMillan, 1989)

3

The Demand for Money

3.1 Informal Discussion

The purpose of this chapter is to develop an understanding of the determinants of the quantity of money demanded by private agents (i.e., the amount they choose to hold at a point in time). To anyone except an economist, this may sound like a peculiar idea; doesn't everyone always want as much money as possible? But the answer to this hypothetical question is, of course, "no." To someone experienced in the economist's ways of thought, it will readily be recognized that there is an implied choice problem which involves the following consideration: for a given amount of wealth, individuals will normally wish to hold only a fraction of it in the form of money, with the remainder being held in the form of other assets—bonds, stocks, houses, cars, and so on. Almost anyone could be holding more money, more of the medium of exchange, at the present time if he chose to hold less of his wealth in these other forms.

In fact, the surprising thing from one point of view is that individuals hold any money at all. For it pays no interest—at least currency does not—while there are other assets that do pay positive interest. Why, then, hold "barren" money? But from our discussion in Chapter 2, it is clear that this point of view is excessively limited. In particular, it takes account only of pecuniary matters and thereby neglects the most important aspect of money—that it helps to facilitate transactions. But, as we have seen, someone who holds an adequate quantity of money is able to conduct her business—do her "shopping"—with a smaller expenditure of time and energy than one who tries to carry out her transactions without use of the medium of exchange. Consequently, the topic at hand involves an optimization problem—that of balancing the expected transactional benefits of holding an additional unit of money against the cost of doing so, which is the extra interest forgone.

From simply thinking about the tradeoff in this way, we can easily deduce the main characteristics of the demand for money by an individual at a point in time. First, since the purpose of holding money is to facilitate planned transactions, more money will be held the greater is the volume of transactions planned. Second, since it is the real quantities of goods and services that people care about, not their nominal values, the relevant quantity of money demanded will be expressed in real (i.e., price deflated) terms. That is, the behavioral relationship to be studied relates real money balances demanded to real transactions planned. Third, since the drawback to holding money—the cost—is the interest that is sacrificed, the (real) quantity willingly held will be smaller the higher is the rate of interest on alternative assets.

These three properties can be expressed formally in terms of a function relating the quantity of real money demanded by a typical person at time t, M_t/P_t, to her planned spending during period t, y_t, and the prevailing rate of interest on some relevant asset, R_t. Letting L denote the function, we then assume that the person's money demand behavior satisfies

$$\frac{M_t}{P_t} = L(y_t, R_t). \tag{1}$$

Because the left-hand-side variable is written as it is, this relation satisfies the second property mentioned in the preceding paragraph. The first and third properties concern the direction of response of the left-hand-side variable to y_t and R_t, respectively. For them to be satisfied, L must be increasing in y_t and decreasing in R_t. Let us then complete our specification by assuming that the function L possesses partial derivatives, and that the partial derivative with respect to (w.r.t.) y_t, denoted $L_1(y_t, R_t)$, is positive while the partial w.r.t. R_t, denoted $L_2(y_t, R_t)$, is negative. Or, to put it more briefly, we assume that $L_1 > 0$ and $L_2 < 0$.

Now, while we have expressed the conclusions of our reasoning about money demand in a somewhat formal way, the reasoning itself has been highly informal. It would accordingly seem to be desirable to work out some analysis in terms of a more explicit and specific optimizing model—one that is more concrete about the nature of alternative assets, transactional benefits from money holding, and so on. As it happens, there is a quite simple and highly specific model that is presented in most textbooks on macroeconomics or monetary theory, the Baumol–Tobin inventory model.[1] That model is *so* specific, however, that its assumptions seem highly artificial and almost unrelated to reality. For completeness we describe the Baumol–Tobin model in Section 3.6. But for our main discussion of money demand, it seems preferable to focus on a different model, one that is just as explicit as the Baumol–Tobin model but more general (i.e., less specific). The lessons it teaches are the same, but some restrictiveness is avoided by taking this more general approach.

3.2 A Formal Model

Consider a hypothetical household that seeks at time t to maximize the multiperiod utility function:

$$u(c_t, l_t) + \beta u(c_{t+1}, l_{t+1}) + \beta^2 u(c_{t+2}, l_{t+2}) + \cdots. \tag{2}$$

Here c_t and l_t are the household's consumption of goods and leisure, respectively, during period t. Note that at time t the household is concerned about its consumption of goods and leisure in future periods, as well as in the present. The function $u(c, l)$ is increasing in both c and l; formally, we assume that the partial derivatives u_1 and u_2 are both positive. In addition, we assume that these partials, which reflect marginal utilities, decrease with c and l; in particular, we assume that $u_{11} < 0$ and $u_{22} < 0$.[2] The parameter β in (2) is a discount factor that is positive but smaller than unity. Thus the household has positive time preference[3] and "discounts the future" in the sense that the utility it receives in t from a given c_t, l_t combination is greater than from the same combination if planned for $t+1$: if $c_{t+1} = c_t$ and $l_{t+1} = l_t$, then $\beta < 1$ implies that $u(c_t, l_t) < \beta u(c_{t+1}, l_{t+1})$.

In making its choices, the household is of course restricted by its budget constraint. In fact, it faces a budget constraint currently in t and knows that it will also face such constraints in future periods $t+1$, $t+2, \ldots$. To keep matters as simple as possible, we assume that the household receives in each period real income in the amount y, with

[1] That model was developed by Baumol (1952)) and Tobin (1956). Clear expositions are included in Barro (1984, Chap. 5) and Dornbusch and Fischer (1984, Chap. 8), among others.

[2] Here u_{11} is shorthand notation for $\partial^2 u/\partial c^2$ and u_{22} for $\partial^2 u/\partial l^2$. Our assumptions on u also include the stipulation that it be "well behaved," which implies that corner solutions will not be chosen for either c_t or l_t in any period.

[3] We can use $\beta = 1/(1+\rho)$ to define the time preference parameter ρ, as in Barro (1984). Then $\rho > 0$ implies that $\beta < 1$.

this amount unaffected by the household's choices. The household does, however, have significant choices to make with regard to its borrowing or lending of wealth. Suppose that in t it can borrow or lend at the interest rate R_t, these loan agreements lasting for just one period. Notationally, let B_t be the nominal quantity of loans made ("bonds" purchased) by the household in t (which expire in $t + 1$). This formulation permits borrowing; if B_t is a negative number, the household is borrowing to that extent. Note, finally, that the household begins period t with assets in the amount $B_{t-1} + M_{t-1}$.

Given these assumptions, the household's budget constraint for period t alone can be written as follows:

$$P_t y + M_{t-1} + (1 + R_{t-1})B_{t-1} = P_t c_t + M_t + B_t. \tag{3}$$

Here the left-hand side (abbreviated l.h.s.) totals the resources (in nominal terms) available to the household from current income, money brought into the period, and bonds purchased (loans made) in the past. Similarly, the right-hand side (r.h.s.) totals expenditures on consumption and bonds during t, plus money balances held at the end of the period. Similar constraints will also be faced in each succeeding period. To take account of those constraints, first note that the constraint for $t + 1$ can be written as

$$B_t = \frac{P_{t+1}(c_{t+1} - y) + M_{t+1} - M_t + B_{t+1}}{1 + R_t}. \tag{4}$$

Now the latter can be used to eliminate B_t from (3). That brings in B_{t+1}, but then a similar step can be used to eliminate B_{t+1}, and so on. By successive eliminations of this type, we can finally arrive at the following equation:

$$\begin{aligned}(1 + R_{t-1})B_{t-1} = {} & [P_t(c_t - y) + (M_t - M_{t-1})] \\ & + (1 + R_t)^{-1}[P_{t+1}(c_{t+1} - y) + (M_{t+1} - M_t)] \\ & + (1 + R_t)^{-1}(1 + R_{t+1})^{-1}\,[P_{t+2}(c_{t+2} - y) \\ & + (M_{t+2} - M_{t+1})] + \cdots. \end{aligned} \tag{5}$$

This single equation then describes the household's entire intertemporal budget constraint, incorporating the constraints for each single period, when the planning horizon is infinite.[4]

[4] For a similar development, see Barro (1984, pp. 83–88). It should be said that formulation (5) assumes that the present value of B_T approaches zero as $T \to \infty$.

Next, we bring in the medium-of-exchange role of money by assuming that to acquire its consumption goods[5] the household must expend time (and energy) in shopping. The amount of time (and energy) so spent depends positively on the volume of consumption but, for any given volume, is reduced by additional money holdings. The reason, of course, is that these holdings facilitate transactions in the manner described in Chapter 2. It is the real quantity of money that matters in this regard; with higher prices, greater nominal amounts of money are needed for given real consumption quantities.

Now, the greater the time (and energy) spent in shopping, the smaller the amount left over for leisure. Thus the argument of the preceding paragraph indicates that leisure in period t, l_t, will be negatively related to consumption and positively related to real money holdings. [We are, as suggested by (3), holding constant the amount of labor performed by the household.] Let us then formalize that idea by assuming that the relationship can be expressed in terms of a function, ψ, as follows:

$$l_t = \psi(c_t, m_t). \tag{6}$$

Here $m_t = M_t/P_t$ is real money holdings.[6] The direction of dependence of l_t on c_t and m_t suggests that we would have $\psi_1 < 0$ and $\psi_2 > 0$, in notation that should by now be familiar. The general notion of diminishing marginal effects, moreover, suggests that $\psi_{11} > 0$ and $\psi_{22} < 0$.

We now have the household's situation specified in enough detail to permit the desired analysis. The household's object at t is to choose c_t, M_t, and B_t values, subject to the constraint (5), so as to maximize the value of

$$u\left[c_t, \psi\left(c_t, \frac{M_t}{P_t}\right)\right] + \beta u\left[c_{t+1}, \psi\left(c_{t+1}, \frac{M_{t+1}}{P_{t+1}}\right)\right] + \cdots, \tag{7}$$

[5] Formally, our model recognizes only one type of consumption good, a feature that appears to be inconsistent with our explanation (in Chapter 2) of the use of money, since the latter presumes that households each consume a large number of distinct goods. These notions can be formally reconciled, as explained by Lucas (1980), by assuming that there are many goods but that conditions for the construction of a composite commodity are satisfied. This strategy permits us to think sensibly about a monetary economy with many goods without worrying about relative price changes, or changes in the composition of consumption bundles, which are irrelevant to most of the important issues in monetary theory.

[6] It might be that the real money held at the start of t, rather than at the end, is the relevant magnitude. In that case the theory would be much the same as here presented, but slightly more awkward in appearance. (Actually, some average value of M over the period might be even more appropriate.) The theory would then include as a special case the "cash-in-advance" specification used by Lucas (1980) and in much recent theoretical research.

where we have substituted (6) into (2). To carry out the maximization problem, let us formulate a Lagrangian expression L_t as follows:[7]

$$\begin{aligned} L_t = u\left[c_t, \psi\left(c_t, \frac{M_t}{P_t}\right)\right] + \beta u\left[c_{t+1}, \psi\left(c_{t+1}, \frac{M_{t+1}}{P_{t+1}}\right)\right] + \cdots \\ + \lambda_t\{(1 + R_{t-1})B_{t-1} - [P_t(c_t - y) + (M_t - M_{t-1})] \\ - (1 + R_t)^{-1}[P_{t+1}(c_{t+1} - y) + (M_{t+1} - M_t)] - \cdots\}. \end{aligned} \tag{8}$$

Then by maximizing with respect to λ_t as well as the actual choice variables, we in effect impose the constraint (5) by way of the first-order condition $\partial L_t/\partial\lambda_t = 0$.

The maximization problem at hand requires in principle that we compute the partial derivatives $\partial L_t/\partial c_{t+j}$ and $\partial L_t/\partial M_{t+j}$ for $j = 0, 1, 2, \ldots$ as well as $\partial L_t/\partial\lambda_t$; set all of these equal to zero; and solve the resulting equations for the implied values of c_{t+j} and M_{t+j}. For the purpose of obtaining our money demand function, however, we need only to find $\partial L_t/\partial c_t$ and $\partial L_t/\partial M_t$. Doing so and setting the resulting expressions equal to zero, we have

$$\frac{\partial L_t}{\partial c_t} = u_1[c_t, \psi(c_t, m_t)] + u_2[c_t, \psi(c_t, m_t)]\psi_1(c_t, m_t) - \lambda_t P_t = 0 \tag{9}$$

and

$$\frac{\partial L_t}{\partial M_t} = \frac{u_2[c_t, \psi(c_t, m_t)]\psi_2(c_t, m_t)}{P_t} - \lambda_t + \lambda_t(1 + R_t)^{-1} = 0. \tag{10}$$

Eliminating $\lambda_t P_t$ from these two equations gives

$$\begin{aligned} u_2[c_t, \psi(c_t, m_t)]\psi_2(c_t, m_t) = [1 - (1 + R_t)^{-1}]\,\{u_1[c_t, \psi(c_t, m_t)] \\ + u_2[c_t, \psi(c_t, m_t)]\psi_1(c_t, m_t)\}. \end{aligned} \tag{11}$$

Careful inspection of the latter will reveal that although it is somewhat complicated, this necessary condition for household optimality involves only three variables: c_t, m_t, and R_t. Assuming that it can be uniquely solved for $m_t = M_t/P_t$, we then rewrite (11) as

$$\frac{M_t}{P_t} = L(c_t, R_t). \tag{12}$$

Now this equation is just like (1), the one we set out to rationalize, except that c_t appears instead of y_t. But even that difference is only apparent, for it is clear that in the present model c_t *is* the relevant transaction variable (i.e., the relevant measure of expenditure volume), which is what y_t stands for in (1).

[7] For a very readable explanation of the Lagrangian technique, see Baumol (1977, Chap. 4).

As a theoretical justification for (1), the foregoing development is lacking in one way: we have not shown that c_t enters positively and R_t negatively on the right-hand side of (12). In fact, it is not true that those signs [i.e., $L_1 > 0$ and $L_2 < 0$ in (12)] are strictly implied for all functions satisfying the assumptions that we placed on u and ψ in (2) and (6). But the appropriate signs will, in fact, obtain except in cases with extreme and unrealistic specifications for u and/or ψ.

To illustrate this fact, and provide an example of a relationship like (12), let us consider a case with specific functional forms for $u(c, l)$ and $\psi(c, m)$. In particular, suppose that

$$u(c_t, l_t) = c_t^{1-\alpha} l_t^{\alpha} \tag{13}$$

and

$$\psi(c_t, m_t) = c_t^{-a} m_t^{a} \tag{14}$$

with α and a positive fractions ($0 < \alpha < 1, 0 < a < 1$). Then we have the partials

$$\begin{aligned} u_2 &= \alpha c_t^{1-\alpha} l_t^{\alpha-1} = \alpha c_t^{1-\alpha}(c_t^{-a} m_t^{a})^{\alpha-1}, \\ \psi_2 &= a c_t^{-a} m_t^{a-1}, \\ u_1 &= (1 - \alpha) c_t^{-\alpha} l_t^{\alpha} = (1 - \alpha) c_t^{-\alpha}(c_t^{-a} m_t^{a})^{\alpha}, \\ \psi_1 &= -a c_t^{-(a+1)} m_t^{a}. \end{aligned} \tag{15}$$

Using these, we find that equation (11) becomes

$$\begin{aligned} a\alpha c_t^{1-\alpha} c_t^{a(1-\alpha)} m_t^{a(\alpha-1)} c_t^{-a} m_t^{a-1} \\ = [1 - (1 + R_t)^{-1}]\{(1 - \alpha) c_t^{-\alpha} c_t^{-a\alpha} m_t^{a\alpha} \\ - a\alpha c_t^{1-\alpha} c_t^{a(1-\alpha)} m_t^{a(\alpha-1)} c_t^{-(a+1)} m_t^{a}\}. \end{aligned} \tag{16}$$

Simplifying the latter and solving for m_t, we finally obtain

$$m_t = \frac{a\alpha}{1 - \alpha - a\alpha}\, c_t\left(1 + \frac{1}{R_t}\right). \tag{17}$$

Now the latter will have the appropriate signs—a positive partial derivative w.r.t. c_t and a negative partial w.r.t. R_t—provided that the term $a\alpha/(1 - \alpha - a\alpha)$ is positive; and that will be the case as long as $1 - \alpha > a\alpha$. Looking back at (13) and (14), we see that this condition requires merely that a change in consumption has a *direct* effect on utility that is stronger than its *indirect* effect working through the shopping-time impact on leisure. Such would certainly be the case for any realistic specification of numerical values for α and a. So our formal model does result in a money demand function of the type postulated in (1).

Before continuing, however, we should pause to note that the "money demand function" that we have derived is in fact not a *demand function* in the proper sense of that term. Strictly speaking, that is, a demand function for an individual specifies the quantity chosen of some commodity as a function of variables that are exogenous to this person—taken by him as given beyond his control. Our relation (12) does not accord with that description since the "explanatory" variable c_t is not exogenous but is chosen by the household. It would, therefore, be more appropriate to refer to (12)—or to (1)—by some other name. In fact, the term "portfolio balance relation" is sometimes used for such equations. But the practice of calling relations like (12) money demand functions is extremely common. Indeed, this practice is so widespread that attempting to avoid it would probably cause more confusion than it would eliminate. Accordingly, we shall continue to use that terminology, improper though it may be.

Let us conclude this section by emphasizing the economic, as opposed to mathematical, content of our derivation of the money demand relation (12). That relation, it should be recalled, is based entirely upon the two first-order optimality conditions (9) and (10), which we rewrite for convenience as follows:

$$u_1(c_t, l_t) + u_2(c_t, l_t)\psi_1(c_t, m_t) = \lambda_t P_t \qquad (9')$$

$$\frac{u_2(c_t, l_t)\psi_2(c_t, m_t)}{P_t} = \lambda_t\left(1 - \frac{1}{1 + R_t}\right). \qquad (10')$$

Consequently, the money demand function (12) is simply the relationship between m_t, c_t, and R_t that holds when both (9′) and (10′) are satisfied. To appreciate the economic content of these, we begin by observing that the first term on the left-hand side of (9′) is the utility provided by an extra unit of consumption, while the second term is the utility provided by an extra unit of leisure multiplied by the loss of leisure necessitated by an extra unit of consumption (a negative number). So the right-hand side of (9′), $\lambda_t P_t$, is equal to the *net* marginal utility of consumption, which is the utility obtained directly from an incremental unit of consumption minus the cost (in utility units) of the leisure sacrificed as a result of the extra consumption. We can then divide $\lambda_t P_t$ by P_t, which is the number of units of money that correspond to one unit of the consumption good, and find that λ_t is the net utility provided by an incremental unit of money holdings.

Turning now to (10′), we see that the left-hand-side numerator is the utility of an extra unit of leisure times the addition to leisure provided by an incremental unit of real money holdings. Division by P_t then makes this pertain to an incremental unit of nominal money holdings. On the right-hand side we have λ_t, the net marginal utility of a unit of money, times $[1 - 1/(1 + R_t)]$. But the latter is approximately equal to R_t, so equation (10′) can be interpreted as follows:

$$\begin{pmatrix}\text{utility from extra leisure provided}\\ \text{by an incremental unit of money held}\end{pmatrix} = \begin{pmatrix}\text{utility from an}\\ \text{extra unit of money}\end{pmatrix}\begin{pmatrix}\text{interest earnings forgone}\\ \text{per unit of money}\end{pmatrix}.$$

In other words, the gain in leisure from holding an additional unit of money must, for optimality to obtain, equal the interest lost, both sides being evaluated in utility units. This is, of course, the type of equality "at the margin" that is characteristic of optimal choices in general, in the manner made familiar by basic microeconomic analysis.

3.6 The Baumol–Tobin Model

It was mentioned above that there exists a theoretical model of money demand behavior that appears frequently in textbooks and other treatments of monetary economics. This model, which was developed by Baumol (1952) and Tobin (1956), is rather different in detail from the "shopping-time" model of Section 3.2 but is quite similar in spirit. The reader's understanding of money demand analysis should be enhanced, therefore, by a familiarity with the Baumol–Tobin model. In this section we begin with a description of the latter and then provide some comments on its relation to the empirical evidence and to the theoretical approach of Section 3.2.

In the Baumol–Tobin analysis it is assumed that the typical household's expenditure during a given period is c in real terms or cP in nominal terms. Also, its purchases are spread evenly in time throughout the period, and these purchases must be paid for entirely with money. Now suppose that the household's income for the period is

received,[21] at the start of the period, in the form of an interest-earning deposit—a "savings account"—that pays interest at the rate R per period but which cannot be used for making purchases. In addition, assume that each "withdrawal" of money from this account or payment into it—that is, each transfer between the earning asset and money—costs the household the lump-sum amount δ in real terms (or δP in nominal terms). This cost is perhaps best thought of as involving an expenditure of time and/or energy, rather than an explicit charge, but that interpretation is not essential.

In the setting described, the household obtains cash for the period's purchases by making one withdrawal in the amount cP, or two withdrawals in the amount $cP/2$, or three in the amount $cP/3$, and so on. Such withdrawals are evenly spaced, so with a steady outflow of money for purchases the household's cash holdings will be as shown in the appropriate panel of Figure 3-1. The different cash-management strategies imply, as a glance at Figure 3-1 will indicate, different *average* levels of money holdings over the period. In particular, if the household makes n withdrawals per period, its average level of money balances during the period will be

$$M = \frac{cP/n}{2}. \tag{29}$$

Thus the larger the number of transactions (withdrawals) per period, the smaller will be the level of money balances held on average over the period.

We now ask: What is the *optimum* number of transactions for the household? This can be determined by noting that the relevant tradeoff is between the cost of making transactions and the opportunity cost of forgone interest that occurs when wealth is held as cash rather than in the interest-earning deposit. In particular, for n withdrawals the transaction cost over the period will be $n\delta P$ in nominal terms, while the interest forgone will be RM. Using (29), the transaction cost can alternatively be expressed as $n\delta P = (cP/2M)\,\delta P = c\,\delta P^2/2M$. Consequently, the magnitude of the two relevant costs together is

$$RM + \frac{c\,\delta P^2}{2M}. \tag{30}$$

To find the household's minimum-cost cash management strategy, we

[21] The model would work the same, but the explanation would be slightly more awkward, if the income were received in the form of money. Then the household for which $n = 2$, for example, would make one deposit into the earning asset (keeping $Pc/2$ in the form of cash) and one withdrawal—two transfers—and would have an average money balance of $Pc/4$, just as it would with the assumption of the text.

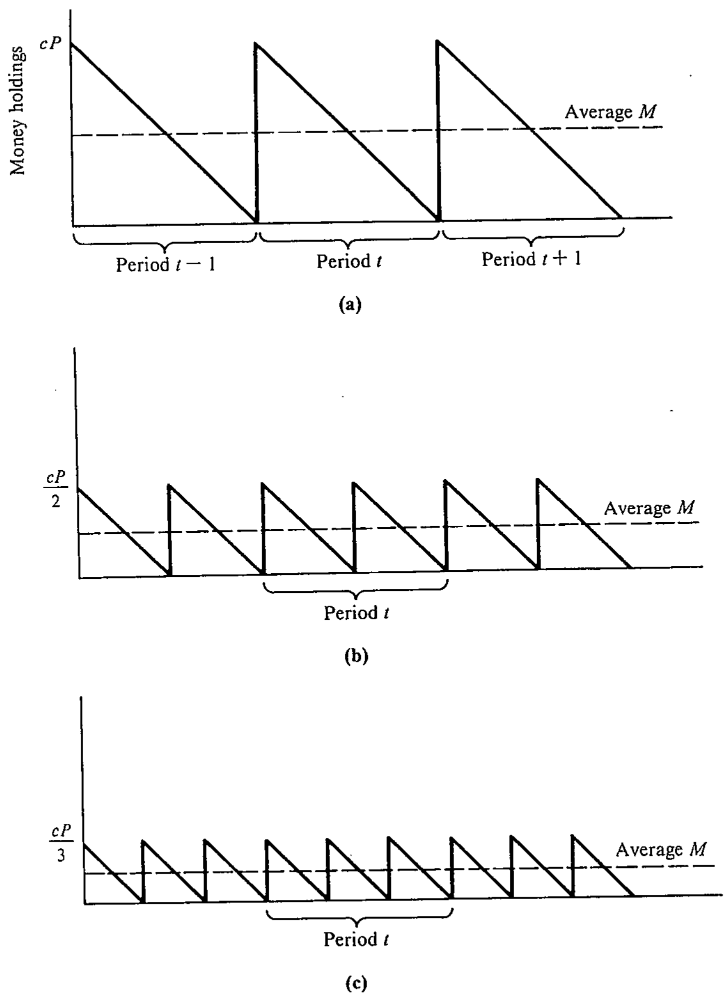

Figure 3-1

can then differentiate this expression with respect to M and set the result equal to zero.[22] Doing so, we obtain

$$R - \frac{c\ \delta P^2}{2M^2} = 0, \tag{31}$$

[22] Since the second derivative is $\delta P^2 c/M^3$, which is necessarily positive, a minimum will be obtained.

and solving the latter for M/P yields

$$\frac{M}{P} = \sqrt{\frac{c\ \delta}{2R}}. \tag{32}$$

Introducing time-period subscripts and using exponential notation, we then have

$$\frac{M_t}{P_t} = \left(\frac{\delta}{2}\right)^{0.5} c_t^{0.5} R_t^{-0.5} \tag{33}$$

as the money demand formula implied by the Baumol–Tobin model. Taking logarithms, this formula can equivalently be expressed as

$$\log \frac{M_t}{P_t} = 0.5 \log \frac{\delta}{2} + 0.5 \log c_t - 0.5 \log R_t. \tag{34}$$

Inspection shows that the latter is a special case of equation (20), which verifies the claim of footnote 12.

At this point it is necessary to mention that there is actually a flaw—a logical error—in the foregoing derivation. The problem is that the number of transactions per period by a household, n, should be constrained to be an integer value—one cannot make (say) 1.37 transactions in a single period. But by using (29) to eliminate n in our derivation of the total cost expression (30), and then taking the derivative with respect to M, we have implicitly treated n as if it were a continuous variable. Thus the final formula (32) cannot be exactly correct.

The main practical implication of this problem is, as it happens, one that makes the Baumol–Tobin model more nearly consistent with the empirical evidence than formula (34) would suggest. The point is that recognition of the integer restriction on n leads to the conclusion that many households will choose to make only one withdrawal, as in the first panel of Figure 3-1.[23] In this case, the household's average money balance during the period will be $M = cP/2$, which features no response at the margin to changes in R.[24] The implied expressions analogous to (33) and (34) are then $M_t/P_t = 0.5c_t$ and $\log\ (M_t/P_t) = \log 0.5 + \log c_t$. Thus the coefficient corresponding to γ_1 in expression

[23] Under the interpretation of footnote 21, these households will make either one or zero transfers, depending on whether or not they deposit part of their income in the interest-earning account. Either way, their average money holdings will agree with the formula given in the next sentence.

[24] There is also no response at the margin to changes in R for some households for which n is greater than 1. But a few such households will make substantial responses as they move to different values of n. For more detail concerning these issues, see Barro and Fischer (1976) and references provided therein.

(20) is 1.0, rather than 0.5, and the coefficient corresponding to γ_2 is zero rather than -0.5. So with some households falling into this category of making the fewest possible number of transfers between money and the interest-earning deposit, the measured economy-wide value for γ_1, which will be an average of the values for all households, will be greater than the value 0.5 given in (34). Similarly, the measured aggregate value for γ_2 will be closer to zero than -0.5. In fact, this is what has been found in the econometric studies mentioned above.[25]

3.7 Conclusions

Before leaving the subject of money demand, it is important to emphasize that the two theoretical models here described—the shopping-time model of Section 3.2 and the Baumol–Tobin model of Section 3.6—are quite similar in terms of their fundamental precepts.[26] Specifically, both of these models presume that money is held (even though higher-yielding assets are available) because it helps to facilitate transactions: in Section 3.2 increased money holdings serve to reduce the time and energy that must be devoted to shopping, while in Section 3.6 money is an absolute necessity for making purchases. And in each case there are ways of getting by with smaller average money holdings when interest opportunity costs are high. Consequently, both models lead to demand functions in which real money balances are positively related to the volume of transactions and negatively related to the interest opportunity-cost variable.

That is not to say that there are no differences between the models; obviously, there are. But these involve details of timing and the precise way in which transaction costs impinge on individuals. Choosing between the models on the basis of these details would seem to be unwise, for the precise ways in which money facilitates exchange are too complex and multifaceted to be represented accurately in any model that

[25] A remaining flaw is that the total amount of money held by households in the United States is larger than the model can account for even with the minimum value of n. But it is widely believed that the surprisingly large magnitude of the U.S. money stock is to be explained by large cash holdings on the part of participants in criminal activities, who wish to keep their transactions unrecorded. Such behavior cannot be explained by any theory that fails to take account of the special nature of criminal activities.

[26] There is, by contrast, a very different approach to money demand analysis known as the "overlapping generations model of money." This approach ignores the fact that money helps to facilitate transactions, emphasizing instead its role as a store of value. Consequently, the overlapping generations model fails to rationalize the observed tendency for agents to hold money even when there are other assets that are free of nominal risk and pay positive interest rates. A relatively nontechnical exposition of this approach is provided by Wallace (1977).

is simple enough to be manageable. Any manageable model, then, should be thought of as a potentially useful parable rather than as a literal description of the exchange process. Choice among different models must consequently be made on the basis of fundamental precepts, consistency with empirical evidence, and analytical convenience. In the case of the shopping-time and Baumol–Tobin models, there is no basis for choice provided by the first two of these criteria. The decision to use one in preference to the other should, therefore, be made on the grounds of relative analytical convenience—which will depend on the issue being studied.