BOJ MONETARY AND ECONOMIC STUDIES · VOL.14 NO.1 · JULY 1996

Production, Financial Sophistication, and the Demand for Money by Households and Firms*

HIROSHI FUJIKI and CASEY B. MULLIGAN

A framework for modeling the demand for money by households and firms is proposed. It allows for both endogenous and exogenous changes in the degree of financial sophistication as well as for multiple monetary assets. The framework is especially useful for interpreting and comparing the many empirical estimates of money demand, as it lists relationships among a variety of empirical and theoretical specifications. We consider a parametric version of the model, and show how the parameters are related to the behavior of various aggregate variables including the aggregate demand for money by firms, the aggregate demand by households, and the aggregate national demand.

Key words: Demand for money; Consumer economics; Household production; Firm behavior

I. Introduction

This paper proposes a theoretical framework for studying the demand for money, but does so with two empirical realities in mind. The first is that a large fraction of the data of interest to monetary economists is aggregate data. By its nature, aggregate data often combines information on both households and firms. Second, the monetary history of the United States and other countries is one of dramatic increase of financial sophistication: NOW accounts, automatic teller machines, and credit cards are few recent examples.

A wide variety of empirical specifications can be found in the money demand literature. Because many of these specifications can be explicitly derived using the theoretical framework of this paper, they can be compared and contrasted in a theoretically rigorous way. This includes a definition of the "demand for money"

*Our research has benefited from the comments of an anonymous referee. Financial support by the Bank of Japan is gratefully acknowledged.

that conforms in an obvious way to the usual price theory meaning of the word "demand," and an explicit consideration of a variety of aggregation issues. We test our theory using Japanese cross prefecture data in Fujiki and Mulligan (1996). Mulligan (1994) uses this framework to propose a parametric model of the demand for money by firms. He estimates the parameters of the model with firm level micro data.

The organization of this paper is as follows. Section II models the usefulness of money by putting money in a production function. Money demand is defined in three ways in section III. First, it is defined as a derived demand - just as economists often model the demand for labor by firms. A second definition is a Marshallian demand - money as a function of "income" and prices. Third, money demand might describe an expansion path, relating money balances to prices and the demand for another input to production. Section III works out and compares some of the properties of these three demand functions. Section IV shows that many of the conventional models of the demand for money are included as special cases of the production model. These include the inventory approaches of Allais (1947), Baumol (1952), Tobin (1956), Miller and Orr (1966), and Barro (1976) as well as other modeling strategies such as those of Whalen (1966), Clower (1967), Lucas and Stokey (1987) and others. Some of the structural parameters - by that we mean parameters of the production functions of agents in the economy - appear in aggregate money demand equations. Section V derives a relationship between the aggregate money stock, aggregate income, and other aggregate variables that is comparable to the demand functions defined in section III. That the results derived from an apparently static model apply in an intertemporal setting is shown in Section VI.

Readers may find the notational appendix to be useful. It collects the definitions of the various symbols used in the mathematical analysis of this paper.

II. The Production Model

A. Final Production Function

Production, broadly defined, is the primary activity of all economic agents. Inputs are rented in markets and used to produce an output. Let this production process be described by the production function f:

$$y_{it} = f(X_{it}, T_{it}, \lambda_f) \tag{1}$$

The quantity y_{it} of output of agent i at date t depends on a J-dimensional vector of inputs X_{it} as well as the quantity of transactions services used at date t, T_{it}. In the case of a firm, y_{it} might be measured as that firm's production or sales. "Household production," on the other hand, may not be observable. λ_f is a parameter of the production function which is assumed - for simplicity only - to be constant over time and identical across agents.

Assumption 1 describes some properties of the production function:

Assumption 1 The production function is continuous, nondecreasing in all arguments and is increasing in *T*.[1]

B. Production of Transactions Services

More will be said about transaction services. In particular, they are produced with money:

$$T_{it} = \phi(m_{it}, X_{it}, A_{it}) \tag{2}$$

m_{it} denotes the quantity of real money balances held by agent i at date t. The agent can change the productivity of a given money stock by renting some or all of the J inputs *X*. Exogenous shifts in the state of financial sophistication also affect the productivity of money and are represented by A_{it}. We assume that $\partial\phi / \partial A > 0$ and $A > 0$. Since *A* is an unobserved technology parameter, these two inequalities are normalizations.

Assumption 2 The production of transactions services is continuous, nondecreasing in each of its arguments and is strictly increasing in *A* and *m*.

C. Cost Minimization

Assumption 3 describes the motivation of agents.

Assumption 3 Agent i's choices of money m_{it} and other inputs X_{it} for period t minimize the rental cost r_{it} of producing output $y_{it} = f(X_{it}, \phi(m_{it}, X_{it}, A_{it}), \lambda_f)$,

[1]The production function need not be differentiable everywhere. If not, the differential statements which follow should be qualified to read "where derivatives exist."

where cost is:

$$r_{it} = q_t' X_{it} + R_t m_{it} \tag{3}$$

where q_t is the J-dimensional column vector of date t rental rates of the J inputs X and R_t is the nominal interest rate at date t. Money is "rented" at rental price equal to the nominal interest rate. This formulation can be justified on the grounds that there exists an alternative asset which pays interest (in units currency units) at rate R_t but does not enter in the production of transactions services. See section IV for one possible derivation of (3).

Assumption 4 All rental rates - including the nominal interest rate - are strictly positive.

The minimum cost achieved is a function of the production level y_{it} and the prices R_t and q_t. This cost function, familiar from standard microeconomic theory (Deaton and Muellbauer, 1980), will be denoted $\Omega(y_{it}, R_t, q_t, A_{it}, \lambda_f)$:

$$\Omega(y_{it}, R_t, q_t, A_{it}, \lambda_f) \equiv \underset{X_{it}, m_{it}}{Min} (q_t' X_{it} + R_t m_{it})$$
$$s.t.\ y_{it} = f(X_{it}, \phi(m_{it}, X_{it}, A_{it}), \lambda_f)$$

Assumptions 1-3 imply that the cost function is homogeneous of degree one in prices q and R, increasing in y, nondecreasing in each of the J rental rates, nondecreasing in the nominal interest rate. Ω is also continuous and concave in (q,R).

The cost function can be used to consider the welfare effects of inflationary monetary policy. For example, suppose that changes in the rate of inflation translate one-to-one into changes in the nominal interest rate. Under this Fisherian hypothesis, we can compute the compensating variation of a μ percentage point change in the rate of inflation:

$$CV = \Omega(y, R+\mu, q) - \Omega(y, R, q)$$

where y is the level of production when the interest rate is R and it is assumed that q is unaffected by a change in the rate of inflation.

Assumptions 5 and 6 are some rather mild restrictions on the two production functions.

Assumption 5 For given rental rates, level of production, and level of financial technology, the elasticity of the production function with respect to transactions services - evaluated at the cost minimizing input depends - approaches zero as λ_f approaches zero.

Assumption 6 The returns to scale of the transactions services production function is bounded above for any positive X and m.

$$\Sigma_j \frac{\partial \phi}{\partial X_j}\frac{X_j}{\phi} + \frac{\partial \phi}{\partial m}\frac{m}{T} \text{ bounded}$$

The sum above is for those X's that appear in the transactions services production function.

III. Definitions and Properties of Money Demand

This section defines the demand for money and compares the definition with some alternatives, with particular attention paid to the choice of "scale variable." This will permit an explicit comparison of some of the scale concepts found in the empirical money demand literature such as income, consumption, sales, and wealth elasticities.

The cost minimizing choices of money and other inputs are functions of output y_{it}, the nominal interest rate R_t, the rental rates of the other inputs q_t, and the level of financial sophistication A_{it}. The Hicksian or derived demand for m_{it} is what we will call the derived demand for money:

$$m_{it} = L(y_{it}, R_t, q_t, A_{it}) = \frac{\partial \Omega(y_{it}, R_t, q_t, A_{it})}{\partial R_t} \tag{4}$$

The second equality follows from Shephard's Lemma.[2]

The properties of the derived money demand function are, of course, related to the properties of the production function f and the possibilities for financial sophistication embodied in the function ϕ. The following section considers special cases of the model; each special case, taken from the money demand literature, places restrictions on the functions f and ϕ. Section V follows by working out some of the relationships between the money demand function and the two production functions.

Lemmas 1 and 2 are useful for deriving some of the properties of the money demand functions. The first shows that the cost function is decreasing in the level of financial technology. The second shows that the relative cost of the production of transactions services goes to zero as the parameter λ_f goes to zero. We put all proofs of the lemmas in a mathematical appendix.

Lemma 1 $d\Omega / dA < 0$

Lemma 2 The relative cost of the production of transactions services$\rightarrow$0 as $\lambda_f \rightarrow 0$.

That "derived money demand slopes down" follows immediately from Shephard's Lemma and the concavity of the cost function:

Proposition 1 The derived demand for money is nonincreasing in the nominal interest rate.

$$\frac{\partial L(y,R,q,A)}{\partial R} < 0$$

One question that we can ask of this derived demand function is "How does the demand for money change as the level of production, *y*, changes?" This is just what we call the **production elasticity of money demand:**

$$\beta_L = \frac{\partial L(y,R,q,A)}{\partial y}\frac{y}{m}$$

---

[2]The first and second derivatives of the cost function with respect to (q,R) exist almost everywhere.

If we consider the special case of firms, where y_{it} is the sales of firm i at date t and X_{it}, is a vector of inputs such as capital and labor, then one might think of the coefficient from a regression of firms' log money balances on their log sales as an estimate of β_L. Such estimates of sales elasticities can be found in Meltzer's (1963) study of firms.[3]

In the case of households, y might correspond to "household production," which is observed neither at the micro or macro levels. To think about alternative scale variables, we define a **Marshallian money demand:**

$$m_{it} = M(r_{it}, R_t, q_t, A_{it}) \equiv L[\Omega^{-1}(r_{it}, R_t, q_t, A_{it}), R_t, q_t, A_{it}]$$

Maintaining the analogy with standard microeconomic theory, we compute the Marshallian money demand function in two steps. First, the cost function Ω is inverted in order to obtain an "indirect production function" y as a function of r, R, q, and A. Second, the indirect production function Ω^{-1} is substituted into the derived money demand function to obtain the Marshallian money demand M.

As will be shown in section IV, our cost variable r_{it} can be identified with income in certain static models of household money demand. In these cases, Propositions 2 and 3 compare the **income elasticity of money demand** with the production elasticity. Proposition 4 compares the elasticities with respect to A of the derived and Marshallian money demand functions.[4]

<u>Proposition 2</u> The production elasticities of money demand, $\beta_M \equiv (\partial M / \partial r) r / M$, and cost elasticities of money demand, $\beta_L \equiv (\partial L / \partial y) y / L$, have the same sign.

<u>Proposition 3</u> If the cost weighted average of the elasticities of the derived demands for the other inputs X with respect to production y is unity and money is a normal good ($\beta_L > 0$), then the cost elasticity of Marshallian money demand is closer to one than is the production elasticity of the derived demand for money. The difference between the production and cost elasticities shrinks as money's share of cost goes to zero.

---

[3]Of course, simple regressions of log money on log sales yield consistent estimates of β_L only if R, q and A are uncorrelated with sales and if money demand is log-linear.

[4]Proofs for propositions 2-4 can be seen in the mathematical appendix .

$$\sum_{j=1}^{J}\left(\frac{\alpha_j}{1-\alpha_m}\right)\left(\frac{\partial X_j}{\partial y}\frac{y}{X_j}\right)=1 \text{ implies}$$

$$\text{(i) } |\beta_M - 1| < |\beta_L - 1|$$

$$\text{(ii) } \lim_{\lambda_f \to 0} |\beta_M - \beta_L| = 0$$

where α_j is X_j' s share of cost and α_m is money's share of cost.

Proposition 4 If money is a normal good, the elasticity of the Marshallian demand for money with respect to the level of financial technology is greater than the derived demand elasticities. The two elasticities are equivalent as money's share of cost approaches zero.

$$\frac{\partial M}{\partial A}\frac{A}{m} > \frac{\partial L}{\partial A}\frac{A}{m}$$

$$\lim_{\lambda_f \to 0}\left|\frac{\partial M}{\partial A}\frac{A}{m} - \frac{\partial L}{\partial A}\frac{A}{m}\right| \to 0$$

Slutsky's equation links the interest elasticity of derived money demand, γ_L, with the interest elasticity of Marshallian money demand:

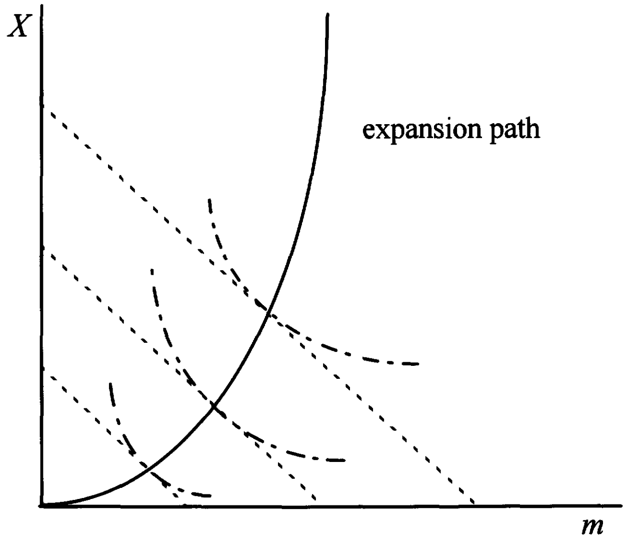
$$\gamma_M = \gamma_L - \beta_M \alpha_m$$

When money is a "normal" good, we see from Slutsky's equation that the interest elasticity of Marshallian money demand is more negative than the corresponding derived elasticity.

In some macroeconomic models, it is important to distinguish income elasticities from **consumption elasticities of money demand**. Section IV will follow Lancaster (1966) and suppose that consumption expenditures can be modeled as inputs into a household production function. Thus, in order to think about the relationship between consumption expenditures and money holdings by households, we need to use our production model to think about the relationship between the derived demand for money and the derived demand for other inputs.

As the level of production increases we can, for given nominal interest rate R, rental rates q, and financial sophistication A, trace out the optimal demands in the

Figure 1
Expansion path for J=1



(X,m) space - an **expansion path**. Figure 1 displays an example for the case J=1.

Now pick a particular input X_j and consider the projection of the expansion path into the (X_j,m) plane. When X_j is a normal good, the expansion path for X_j and m can be expressed as a function g_j:

$$m_{it} = g_j(X_{j,it}, R_t, q_t, A_{it}, \lambda_f) \equiv L(H_j^{-1}(X_{j,it}, R_t, q_t, A_{it}, \lambda_f), R_t, q_t, A_{it}) \tag{5}$$

where $H_j(X,R,q,A,\lambda_f)$ is the derived demand for X_j and $H_j^{-1}(X,R,q,A,\lambda_f)$ is inverse of the derived demand for X_j.

Define the **jth input elasticity of money demand** as the elasticity of this expansion path:

$$\beta_j \equiv \frac{\partial g_j(X_j, R, q, A)}{\partial X_j}\frac{X_j}{m} \quad j = 1,\dots,k$$

Propositions 5-8 show how the various elasticities of an expansion path are related to the corresponding elasticities of the Marshallian and derived money demand functions.[5]

Proposition 5 The jth input elasticity of money demand is proportional to the production elasticity of money demand, where the factor of proportionality is the inverse of the production elasticity of the derived demand for X_j:

$$\beta_j \equiv \frac{\beta_L}{\frac{\partial H_j}{\partial y}\frac{y}{X_j}}$$

Proposition 6 If J = 1 and X is a normal good, the cost elasticity of money demand is closer to one than is the input elasticity of money demand. If J = 1 and β_j is bounded from above, then the cost elasticity approaches the input elasticity as the expenditure on money takes a smaller share of cost.

$$J = 1, \beta_j > 0 \text{ implies } |\beta_j - 1| > |\beta_M - 1|$$
$$J = 1, \beta_j \text{ bdd implies } \lim_{\lambda f \to 0} |\beta_j - \beta_M| = 0$$

Proposition 7 The (point) price elasticities of the derived demands for money and the other inputs are linear combinations of the price elasticities of the expansion paths. The weights depend on the cost shares and the input elasticities:

$$\Sigma = \left[I - \frac{\zeta\alpha'}{\alpha'\zeta} \right] \Pi$$

where Σ is the matrix of price elasticities of the derived demand functions (σ_{ij} is the elasticity of the derived demand for input i with respect to the price of good j), Π is the matrix of price elasticities of the expansion path (π_{ij} is the elasticity of the expansion path for input i with respect to the price of good j), α is a column vector of cost shares (expenditure on an input divided by total cost), and ζ is a column vector

[5]Proofs for propositions 5-8 can be seen in the mathematical appendix.

of scale elasticities of the expansion paths.[6]

Proposition 8 If money and the input X_j are normal goods and X_j does not appear in the production function for transactions services, the elasticity the projection of the expansion path into the (X_j,m) plane with respect to the level of financial technology is greater (less) than the derived money demand elasticity if the elasticity of the derived demand for X_j is less (greater) than 0. If, in addition, this second elasticity approaches zero as $\lambda_f \to 0$, then the two elasticities approach each other.

$$\frac{\partial g_j}{\partial A}\frac{A}{m} \gtrless \frac{\partial L}{\partial A}\frac{A}{m} \quad \text{as} \quad \frac{\partial H_j}{\partial A} \gtrless 0$$

$$\lim_{\lambda_f \to 0}\left|\frac{\partial g_j}{\partial A}\frac{A}{m} - \frac{\partial L}{\partial A}\frac{A}{m}\right| = 0 \text{ when } \lim_{\lambda_f \to 0} \frac{\partial H_j}{\partial A} = 0$$

By designating the price elasticities of any one of the three money demand functions, we can use Proposition 7 and Slutsky's equation to compute the price elasticities of the other two. We find it convenient in sections IV and V to begin with expansion paths, use Proposition 7 to compute the price elasticities of derived money demand, and then use Slutsky's equation to derive the price elasticities of Marshallian money demand.

Various motivations for the demand for money differ on their predictions for the scale elasticity - so it is useful to have a empirical specification that allows for nonunitary scale elasticities. It is well known that scale elasticities of derived or Marshallian demands cannot be constant *and* different than one without the unpleasant prediction that, for some income levels, expenditures on the rental of money will consume the entire budget. Expansion paths, by contrast, can have a constant scale elasticity that differs from one without money rent consuming the entire budget. An empirically oriented study of money demand might therefore begin with expansion paths rather than derived or Marshallian demands so that it might enjoy the empirically convenient constant scale elasticity specification. The results from this section indicate how properties of the two other demand concepts are related to the properties of the expansion path.

[6]For example, consider the three input case. For a given price vector, project the expansion path into the (x_1,x_2) and (x_1,m) and planes. Denote the elasticities of these two projections (with respect to x_1) ζ_2 and ζ_m. Then the column vector ζ is $(1,\xi_2,\xi_m)$.

IV. Examples

The previous section defines the demand for money as a producer's derived demand function, or as a Marshallian demand function, or as the relationship between money and another input along a production expansion path. Following Fischer (1974) and Feenstra (1986), we argue that the production function approach includes many of the well-known theoretical approaches to the demand for money.[7] The restrictions that are placed on the production functions f and ϕ in some special cases are shown. The details of the mathematical properties of the production functions used for the following examples can be found in the appendix.

We work out only four examples in detail: money in the utility function, a cash-in-advance model, an inventory model, and a model with two monetary assets. Many more examples, such as Saving (1971) or Whalen (1966) are also possible - we refer readers to Fischer (1974) and Feenstra (1986) for proofs of the equivalence between our approach and many others in the theoretical literature.

A. Money in the Utility Function

Let's suppose that our "economic agent" is an household. A homogeneous commodity c_{it}, together with transaction services T_{it}, produces a composite commodity y_{it}:

$$y_{it} = f(c_{it}, T_{it})$$

The productivity of money is a constant :

$$T_{it} = \bar{\phi} m_{it}$$

As in the analysis of Lancaster (1966), the household's utility function - defined over quantities of the composite commodity - together with the household production function f, defines a reduced form utility function defined over c and m.

[7] Feenstra (1986) shows that transactions-based and other approaches to the demand for money are, for particular reduced-form utility functions, similar to the "money in the utility function" approach. We build on the Feenstra paper in several ways: (1) firms and households are special cases, (2) financial sophistication is explicitly modeled, and (3) multiple monetary assets can be studied.

$$u(c_{it}, m_{it}) \equiv v[f(c_{it}, \bar{\phi} m_{it})]$$

where v(y) is the consumers utility when he consumes an amount y of the composite commodity. v(y) is assumed to be concave.

B. Cash-in-Advance

The cash-in-advance case is rather simple: household i's production for period t is the minimum of desired consumption expenditures and money holdings:

$$y_{it} = Min(c_{it}, m_{it})$$

If desired consumption expenditures are less than or equal to money holdings, then household production is equal to desired consumption expenditures. If money holdings are less, then production is equal to m_{it} because not enough money is on hand to purchase the desired level of consumption.

The three demand functions for the cash-in-advance case are also quite simple:

$$m_{it} = L(y_{it}, R_t, q_t, A_{it}) = y_{it}$$

$$m_{it} = M(r_{it}, R_t, q_t, A_{it}) = r_{it} / (1 + R_t)$$

$$m_{it} = g(c_{it}, R_t, q_t, A_{it}) = c_{it}$$

The expansion path is the 45 degree line in the (c,m) plane. Since the production elasticities of the derived demand for consumption and money are both one, Propositions 3 and 5 imply that the Marshallian and input scale elasticities of the demand for money are also one - a prediction which is confirmed above. Slutksy's equation can be used to derived the Marshallian interest elasticity of $-R/(1+R)$, which can be verified directly by differentiating the Marhsallian demand function displayed above. Proposition 7 predicts that, like the price elasticities of the expansion path g, the price elasticities of the derived demand for money are zero.

C. Allais-Baumol-Tobin

Consider the following production functions:

$$y_{it} = f(c_{it}, T_{it}) = c_{it} \exp(-\frac{1}{T_{it}}); \; T_{it} = 2A_{it}m_{it}$$

Cost in this model is the sum of consumption expenditures and foregone interest: $r_{it} = c_{it} + R_t m_{it}$. Cost minimizing input demands exhibit the familiar square root relationship between money holdings, consumption expenditures, the nominal interest rate, and the transactions technology:

$$m_{it} = \sqrt{\frac{c_{it}}{2R_t A_{it}}}$$

We see that there exists a specification of the production function model that is observationally equivalent to the Allais-Baumol-Tobin model of the demand for money by households.

In the terminology of the previous section, the square root rule is a description of an expansion path, with the input elasticity of money demand (β_j) equal to 1/2. It is a straightforward exercise to compute the other two types of money demand functions. The derived demand for money depends on the level of household production y, the nominal interest rate, and the level of financial sophistication, but (for this model) can only be solved for implicitly:

$$\log R_t = -2\log L(y_{it}, R_t, A_{it}) + \frac{1}{2A_{it}L(y_{it}, R_t, A_{it})} + \log y_{it} - \log 2A_{it}$$

A closed form solution is available for the Marshallian money demand, which depends on cost (cost might be thought of as income in this model), the nominal interest rate, and the transactions technology:

$$m_{it} = \frac{1}{4A_{it}}\sqrt{8A_{it}\frac{r_{it}}{R_t} + 1} - \frac{1}{4A_{it}}$$

Since J=1, $\beta_1 = 1/2 > 0$, and β_1 is bounded above, Proposition 6 predicts that the Marshallian scale elasticity is greater than 1/2 - a fact which can be verified by differentiating the Marshallian demand function displayed above.

D. Two Monetary Assets

Suppose transactions services are produced with two monetary assets - say currency m and demand deposits d:

$$T_{it} = \phi(m_{it}, d_{it}) = [(1-\lambda_\phi)m_{it}^{(\psi_\phi-1)/\psi_\phi} + \lambda_\phi d_{it}^{(\psi_\phi-1)/\psi_\phi}]^{(\psi_\phi-1)/\psi_\phi}$$

Let the rental rate of currency be R_t and the rental rate on demand deposits be $q_{d,t}$. q_3 is measured as the difference between the nominal interest rate and the rate of interest paid on demand deposits (if any). For each of the two inputs, we can define the three types of demand functions: derived demand, Marshallian demand, and expansion path (with scale variable x, an input in the final production function). Section III proves that, as $\lambda_f \to 0$, the price and scale elasticities are identical for each of the three demands for m; those results also go through for the three demands for d. Denoting the scale elasticity as β and the elasticity of substitution of x for T in the final production function as γ, the demands for currency and demand deposits separately are:[8]

$$\log m_{it} = \beta \log y_{it} - \gamma \log R_t + \pi_\phi(\psi_\phi - \gamma)\log\frac{q_{3,t}}{R_t} + (constant)$$

$$\log d_{it} = \beta \log y_{it} - \gamma \log q_{3,t} + (1-\pi_\phi)(\psi_\phi - \gamma)\log\frac{R_t}{q_{3,t}} + (constant)$$

As predicted by Proposition 1, the two demand functions are decreasing in their own price. They are increasing in the price of the other asset relative to its own price, although not in a log-linear way. Log-linear approximations to the relative price terms are displayed above, with the constant π_ϕ arising from that approximation.

Consider now the demand for the sum $(m+d)$.[9] The scale elasticity of the sum is exactly β. A loglinear approximation to the price terms is displayed below:

---

[8] A derivation of β and γ as scale and substitution elasticities for a particular production function can be found in the following section, SectionV.

[9] Using the production function approach, it is straightforward to study the demand for nonlinear monetary aggregates such as those advocated by Barnett, *et al* (1992). A CES aggregate of currency and demand deposits seems particularly relevant for the current example.

$$\log(m_{it}+d_{it}) \approx \beta \log y_{it} - \gamma \log R_t + [\pi_\phi(\psi_\phi - \gamma) - (1-\delta)\psi_\phi]\log \frac{q_{3,t}}{R_t} + (constant)$$

where δ is derived from the log linear approximation and is equal to $d/(m+d)$ at some benchmark value for q_3 / R. We see that demand for this monetary aggregate is similar to the demand for currency, except for its dependence on the relative price. The first order effect of the relative price on the demand for the aggregate depends on the importance of currency relative to demand deposits as well as the ease of substitution of currency for demand deposits.

For a different production function ϕ, we could derive separate scale elasticities for currency and demand deposits. Our homothetic specification above, however, requires that the two scale elasticities be identical.

V. Identification of Parameters from Aggregate Data

Here we begin with a parametric model for production by households and firms. Useful money demand functions are derived for both types of agents. It is then shown how some of the structural parameters (i.e., parameters of the production/utility functions) can be identified from aggregate data.

A. Parametric Model for Households and Firms

Consider the following special cases of the production functions (1) and (2) (all of the Greek parameters are positive constants)[10]

$$y_{it} = f(x_{1,it}, T_{it}, \lambda_f) = [(1-\lambda_f)x_{1,it}^{(\gamma-\beta)/\gamma} + \lambda_f(\frac{\gamma-\beta}{\gamma-1})T_{it}^{(\gamma-1)/\gamma}]^{\gamma/(\gamma-\beta)}$$
$$\lambda_f \in (0,1), \beta > 0, \gamma \in (0, \min(1,\beta)) \tag{1'}$$

$$T_{it} = \phi(m_{it}, x_{3,it}, A_{it}) = A_{it}[(1-\lambda_\phi)m_{it}^{(\psi_\phi-1)/\psi_\phi} + \lambda_\phi x_{3,it}^{(\psi_f-1)/\psi_f}]^{\psi_\phi/(\psi_\phi-1)} \tag{2'}$$

For simplicity (1') parameterizes the function f, considering the case when there

[10] For $\gamma = 1$, log T replaces the power function of T. For $\gamma = \beta$, log replaces the power function whose argument is a term in small square brackets to the $\psi_f / (\psi_f - 1)$ power. For ψ_ϕ or $(\gamma = \beta)$ equal to 1, the corresponding CES aggregator is replaced by a Cobb-Douglas aggregator, with exponents $\lambda_f, (1-\lambda_f)$ or, $\lambda_\phi, (1-\lambda_\phi)$.

is one input other than transactions services in the final production function.[11] As before, *m* denotes real money balances. x_1 and x_3 denote other inputs with only x_3 used in the production of transactions services. The production of transactions services is a CES function, λ_ϕ and λ_f in the interval (0,1). Transactions services are not, however, aggregated with x_1 in a homogeneous way. Notice that the exponent on the first term is $(\gamma - \beta)/\gamma$, whereas the exponent on the second term is $(\gamma - 1)/\gamma$. We will show that scale elasticities will differ from one when these two exponents differ (i.e., when $\beta = 1$).

For the case of firms, *y* is observable so the choice of units embodied in the production function (1') has empirical relevance. On the other hand, production is not observed for households so any monotonic transformation of f will not change the empirical implications.

Cost is the sum of rental expenditures for the three inputs and money:

$$r_{it} = q_{1,t}x_{1,it} + R_t m_{it} + q_{3,t}x_{3,it}$$

The functional forms in this example lead to a fairly simple characterization of the expansion path:

$$\begin{aligned}\log m_{it} &= \log g_1(x_{1,it}, R_t, q_{it}, A_{it}) \\ &= \beta \log x_{1,it} - \gamma \log R_t + \pi_\phi(\psi_\phi - \gamma)\log\frac{q_{3,t}}{R_t} + \\ &\quad \gamma \log q_{1,t} - (1-\gamma)\log A_{it} + \varphi(\psi_\phi, \gamma, \beta, \lambda_f, \lambda_\phi)\end{aligned} \qquad (6)$$

φ is an intercept term that is a function of the production parameters only. The projection of the expansion path into the (m, x_1) plane has a constant elasticity of β. Holding constant q_3/R, the elasticity with respect to the nominal interest rate is the constant - γ.[12]

Increases in the level of financial technology, A_{it}, decrease the demand for money when γ is less than one, but increase money demand for $\gamma > 1$. The effect

[11] It is straightforward to allow for more nonmonetary inputs without changing the implications that are derived below. For example, one could replace x_1 with a homogeneous function of several inputs. Instead of representing a single rental rate, q_1 is interpreted as a price index for the rental rates of the several inputs.

[12] The expansion path is not log-linear in q_3/R; the constant π_ϕ in equation (7) is derived from a log linear approximation to this term.

of technology on the demand for money depends on the interest elasticity of money demand. To see this, notice that holding constant q_3 / R, the price of transactions services is the ratio *R/A*. Transactions services are more costly when q_3 / R and *R* increase, but are less expensive when *m* and x_3 are more productive. $\gamma < 1$ means that there are few possibilities for substitution of *T* for x_1, so a change in *R/A* - say because of an increase in *A* - does little to the demand for *T*. The productivity effect of A therefore dictates that the demand for *m* and x_3 fall. For $\gamma > 1$, the substitution towards transactions services outweighs the productivity effect so the demands for *m* and x_3 increase.

x_1 and x_3 need not be the same good for all agents. For example households may use a consumption good and labor while firms use capital and labor. All agents must use money, however. If agents are using different inputs, then (6) should be modified to allow prices q_1 or q_3 to vary across agents.

Section III argues that an expansion path such as (6) is only one way to characterize the demand for money. The derived demand for money and the Marshallian demand for money are alternative characterizations. Propositions 3 and 6 describe some conditions for which the production elasticity of money demand is close to the cost and input elasticities. In particular, when the production elasticities of demand for the other inputs are one (on average), these elasticities approach each other as money's share of cost approaches zero. For the purposes of aggregation, we assume that these three elasticities are in fact equal.[13] Since we can see from (6) that the 1st input elasticity of money demand is β, we can use β to approximate the cost and production elasticities. Proposition 7 and Slutsky's equation tell us that the interest elasticity of the three types of money demand functions are also equal as money's share of cost goes to zero, so we can derive log-linear approximations to all three money demand functions:

[13] The rental cost of money does not appear to be a substantial fraction of GNP in the U.S. The ratio of M1 to GNP is about 0.15 (U.S. Council of Economic Advisors, 1994). Even at a 10% interest rate, the rental cost of money is only 1.5% of GNP. This is an even a smaller percentage if one allows for the fact that the sum of sales of firms and incomes of households would add up to much more than GNP.

The other requirement of Propositions 3, 4, and 5 - that an average of the production elasticities of the nonmonetary inputs be one - also holds as an approximation for this problem. As α_m approaches zero (holding constant relative prices), so must α_3 (because transactions services are produced according to a homothetic production function). The x_1 and x_2 terms will dominate the production function, so they must be used in proportion to production y. Note that $\alpha_m \to 0$ as $y \to \infty$ ($y \to 0$) for $\beta < 1$ ($\beta > 1$).

$$\begin{aligned}\log m_{it} &= \log g_1(x_{1,it}, R_t, q_{it}, A_{it}) \\ &= \beta \log x_{1,it} - \gamma \log R_t + \pi_\phi(\psi_\phi - \gamma)\log\frac{q_{3,it}}{R_t} \\ &\quad +\gamma \log q_{1,it} - (1-\gamma)\log A_{it} + \varphi(\psi_\phi, \gamma, \beta, \lambda_f, \lambda_\phi)\end{aligned} \tag{7}$$

$$\begin{aligned}\log m_{it} &= \log L(y_{it}, R_t, q_{it}, A_{it}) \\ &\approx \beta \log y_{it} - \gamma \log R_t + \pi_\phi(\psi_\phi - \gamma)\log\frac{q_{3,it}}{R_t} \\ &\quad +\gamma \log q_{1,it} - (1-\gamma)\log A_{it} + constant\end{aligned} \tag{8}$$

$$\begin{aligned}\log m_{it} &= \log M(r_{it}, R_t, q_{it}, A_{it}) \\ &\approx \beta \log r_{it} - \gamma \log R_t + \pi_\phi(\psi_\phi - \gamma)\log\frac{q_{3,it}}{R_t} \\ &\quad +(\gamma - \beta)\log q_{1,it} - (1-\gamma)\log A_{it} + constant\end{aligned} \tag{9}$$

where rental rates of the inputs other than money have been subscripted by i to allow for different agents to use different inputs. (8) and (9) are most accurate as approximations to the derived and Marshallian demands for money as the share of money and x_3 in cost approach to zero.[14] Both approximations have the same price elasticities for money and x_3 as does the expansion path (7). The exact (local) price elasticities can be computed using Proposition 7 and Slutsky's equation.[15]

For those agents that are households, *r* is equal to income which will be denoted *I*. For firms, the level of production *y* can be interpreted as sales. It will be assumed that income, rental rates and technology are lognormally distributed across households and that sales, rental rates and technology are lognormally distributed across firms:

[14] For fixed q_3 / R, α_3 must go to zero as α_m does.

[15] Proposition 7 computes the price elasticities of the derived demand for money as a function of the cost shares $(\alpha_1, \alpha_2, \ldots, \alpha_N, \alpha_m)$ and elasticities of an expansion path. Slutsky's equation then computes the Marshallian price elasticities as a function of derived demand elasticities.

for households:

$$\log I_{it} \sim N[\mu_{i,t}(h), \sigma^2_{It}(h)]$$
$$\log q_{j,it} \sim N[\mu_{j,t}(h), \sigma^2_{jt}(h)];\ j = 1,3.$$
$$\log A_{it} \sim N[\mu_{A,t}(h), \sigma^2_{At}(h)]$$

for firms:

$$\log y_{it} \sim N[\mu_{y,t}(f), \sigma^2_{yt}(f)]$$
$$\log q_{j,it} \sim N[\mu_{j,t}(f), \sigma^2_{jt}(f)];\ j = 1,3.$$
$$\log A_{it} \sim N[\mu_{A,t}(f), \sigma^2_{At}(f)]$$

For the sake of generality we allow prices (except the nominal interest rate) to differ across agents. Depending on the application, one might set the variances to zero so that all households face one price and all firms face another or, in addition, set $\mu_{j,t}(f) = \mu_{j,t}(h)$ so that both firms and households face the same price.

B. Firm Aggregates and Household Aggregates

Here we consider aggregation of the derived money demand functions of firms and then the aggregation of the Marshallian demand functions of households. We consider the firm's derived demand function - as opposed to Marshallian demand or an expansion path - for two reasons. First, the derived demand (8) follows the empirical literature by relating money balances to sales of the firm. Second, we argue in subsection V.C. that, because sales is the scale variable, firms' derived demand can be readily combined with households' Marshallian demands to arrive at a national money demand equation that resembles those found in the macro literature. The derived demand for households, on the other hand, is not as useful because household production is unobserved. Fortunately, the three types of money demand functions have some similarities; the similarities can be exploited to derive aggregate relationships that are functions of production parameters such as β and γ.

Let $N_t(f)$ and $N_t(h)$ denote the number of firms and households in the economy at date t, respectively. $y_t(f)$ and $m_t(f)$ are the average sales and real money balances of firms at date t (ie, the sum of sales and money balances divided by the number of firms). $I_t(h)$ and $m_t(h)$ are date t average household income and real money balances. Using some properties of the lognormal distribution, we arrive at two aggregate money demand functions: one for firms and one for households:

$$
\begin{aligned}
\log m_t(f) = {} & \beta \log y_t(f) - \gamma \log R_t \\
& + \pi_\phi(\psi_\phi - \gamma)\log \frac{q_{3,t}(f)}{R_t} + \gamma \log q_{1,t}(f) - (1-\gamma)\log A_t(f) \\
& + \frac{1}{2}\beta(\beta-1)\sigma_{yt}^2(f) + \frac{1}{2}\pi_\phi(\psi_\phi - \gamma)[\pi_\phi(\psi_\phi - \gamma) - 1]\sigma_{3t}^2(f) \\
& + \frac{1}{2}(1-\gamma)(2-\gamma)\sigma_{At}^2(f) + \frac{1}{2}\gamma(\gamma - 1)\sigma_{1t}^2(f) \\
& + covariances + constant
\end{aligned}
$$

$$
\begin{aligned}
\log m_t(h) = {} & \beta \log I_t(h) - \gamma \log R_t + \\
& \pi_\phi(\psi_\phi - \gamma)\log \frac{q_{3,t}(h)}{R_t} + (\gamma - \beta)\log q_{1,t}(h) - (1-\gamma)\log A_t(h) \\
& + \frac{1}{2}\beta(\beta-1)\sigma_{It}^2(h) + \frac{1}{2}\pi_\phi(\psi_\phi - \gamma)[\pi_\phi(\psi_\phi - \gamma) - 1]\sigma_{3t}^2(h) \\
& + \frac{1}{2}(1-\gamma)(2-\gamma)\sigma_{At}^2(h) + \frac{1}{2}(\gamma - \beta)(\gamma - \beta - 1)\sigma_{1t}^2(h) \\
& + covariances + constant
\end{aligned}
$$

For scale and price elasticities of one and no correlation among the scale and price variables, household and firm aggregate demand functions are identical to their micro counterparts (8) and (9). For scale elasticities different from one, the variance of sales and the variance of household income enter the aggregate equations. Variances and covariances of the price variables also enter the aggregate equations, but disappear if the prices are the same among firms and the same among households.

C. National Aggregates

For scale and price elasticities different from one, the definition of an "economic agent" is important. Section II began with households and firms as agents. Thus, in order to derive macro money demand functions in terms of money and income per capita, we need to keep track of the number of firms and households per capita. Define N_t to be the size of the population at date t. $\eta_t(f) \equiv N_t(f)/N_t$ and $\eta_t(h) \equiv N_t(h)/N_t$ denote the number of firms and households per capita. Let v_t denote aggregate sales as a fraction of aggregate household income:

$$v_t \equiv \frac{N_t(f)}{N_t(h)} \frac{y_t(f)}{I_t(h)}$$

Using the aggregate firm money demands and the aggregate household money demands from the previous subsection - together with a loglinear approximation of $\log[m_t(f) + m_t(h)]$ - we derive an expression (10) for real money balances per capita:

$$\begin{aligned} \log(\frac{M_t}{P_t N_t}) &= \beta \log y_t(h) - \gamma \log R_t \\ &+ \pi_\phi(\psi_\phi - \gamma)[\omega \log \frac{q_{3,t}(f)}{R_t} + (1-\omega) \log \frac{q_{3,t}(h)}{R_t}] \\ &+ \omega\gamma \log q_{1,t}(f) + (1-\omega)(\gamma - \beta) \log q_{1,t}(h) \\ &- (1-\gamma)[\omega \log A_t(f) + (1-\omega) \log A_t(h)] \\ &+ [\omega \log \eta_t(f) + (1-\omega) \log \eta_t(h)] + \beta\omega[\log v_t + \log \frac{\eta_t(h)}{\eta_t(f)}] \\ &+ \frac{1}{2}\beta(\beta - 1)[\omega\sigma_{yt}^2(f) + (1-\omega)\sigma_{It}^2(h)] \\ &+ \textit{other covariances weighted by } \omega, 1-\omega \end{aligned} \tag{10}$$

Beginning with the first four terms (the first line) of equation (10), we see that, like its micro counterparts, the per capita demand for money depends on average household income, the nominal interest rate and the ratio q_3 / R with elasticities β, $-\gamma$ and $\pi_\phi(\psi_\phi - \gamma)$. When the average price q_3 is different for households and firms, however, the geometric mean of the two q_3 / R ratios (one for firms, one for households) enters the aggregate equation. The weight ω, can be approximated by the share of the money stock held by firms (as opposed to households).[16] Terms reflecting averages of the price of x_1 and the level of financial technology enter the aggregate money demand equation separately for firms and households. Per capita money demand also depends on the number of firms and households per capita as well as the ratio of aggregate sales to household income (v_t). These three terms, roughly speaking, represent the degree of vertical integration in the economy. The more stages involved in the production process, the greater the demand for money. This

---

16 ω derives from an approximation to $\log[m_t(f) + m_t(h)]$

vertical integration result follows from the assumption that *a firm* and *a household* are the demanders of money. Economies of scale ($\beta < 1$) cannot be exploited by pooling money holdings across firms or across households while diseconomies of scale ($\beta > 1$) cannot be avoided by subdividing money holdings within the firm or within the household.

Finally, for given average sales and average income, the dispersion of income and sales across agents affects aggregate money demand to the extent there are economies (or diseconomies) of scale in the holding of money. There is some evidence that income distribution terms belong in aggregate money demand equations. For example, Chan and Chen (1992) find that more inequality across geographic regions is associated with a lower aggregate U.S. demand for money. This prediction obtains for our specification if $\beta < 1$ and changes in the variance of income and sales are not correlated with changes in the variance of other arguments of the money demand function (eg., a wage rate). Cover and Hooks (1993), on the other hand, argue that increases in the degree of inequality as measured by household survey data are associated with a higher U.S. demand for money. The Cover and Hooks finding is consistent with $\beta > 1$. If $\beta < 1$, but the wage and income elasticities of money demand summed to greater one, then it is possible that more inequality would be associated with higher money demand.[17]

The aggregate money demand equation (10) indicates that, with enough data, one could obtain consistent estimates of some of the structural parameters of the model such as β, γ, and perhaps ψ_ϕ. The scale elasticity is interesting for economic theory as various models of the demand for money differ on the presence and extent of scale economies. g and ψ_ϕ reflect the own price elasticity of money demand and are therefore indicative of the welfare cost of inflation and relevant for computing the optimal monetary policy (see Lucas (1994) for a discussion and for references). However, estimation of (10) requires that (i) one has a time series on the prices q_1 and q_3 and the level of financial sophistication *or* (ii) all cross-price elasticities are zero *or* (iii) the cross-prices and the level of financial technology are uncorrelated with household income and the nominal interest rate.

Condition (iii) is certainly violated if we estimate equation (10) in levels. Financial technology has grown over time as has household income. Or, if one

[17] See Mulligan (1994) for estimates of the demand for money by firms that indicate that $\beta < 1$ and that β plus a wage elasticity is greater than one.

prefers to think of financial sophistication as endogenous, the rental rate of financial technology (which might be modeled as q_3 in our setup), such as the computer, has fallen over time. One solution to this problem might be to estimate (10) in differences. Perhaps high frequency movements in income are not associated with high frequency movements in financial technology. However, the same might not be true for short-term movements in the nominal interest rate. We can imagine that economy-wide stocks of financial technology (which we might model as the good x_3 are fixed in the short run. A rapid increase in the nominal interest rate will increase the demand for the technology which, because stocks are fixed, must result in an increase (but less than proportional) in the rental rate q_3. In other words, q_3 will be correlated with R at high frequencies.[18]

This has led some studies to use cross-sections of regional aggregates to identify the production parameter β.[19] The idea is that within a country such as the U.S., all agents have fairly equal access to financial technology. Thus it is assumed that the exogenous level of financial technology A, the rental price of financial machines (q_3), and the nominal interest rate are all constant in a cross-section of regions. g and ψ_ϕ can then be estimated in levels using aggregate time series data by imposing that b correspond to its estimate from the cross-sections. We expect consistent estimates as long as q_3 and A are uncorrelated with R in the time series.

VI. Intertemporal Versions of the Model

The analysis in this paper has been static. Here we write down an intertemporal model which requires static cost minimization as described in previous sections. Thus, results for the static model also apply to this particular intertemporal model.

Consider T periods t = 1,2,...,T. The flow of production y for each period is an input into a dynamic production function V:

$$V_0 = V[y_1, y_2, \ldots, y_T]$$

The static production functions (1) and (2) still apply. Agents can carry resources across periods by purchasing (or selling) nominal assets. Assets purchased in period

[18] This problem may also occur with seasonal data because the stock of machines such as computers may not vary across seasons, but the demand for their services might.

[19] Mulligan and Sala-i-Martin (1992) is one example.

t-1 pay interest R_t at the beginning of period t. The agent faces a budget constraint in every period:

$$b_{t+1} = b_t + R_t b_t + P_t(I_t - q_t^{'} X_t) + M_t - M_{t+1}, \quad t = 1,2,\ldots,T. \tag{11}$$

P_t is the price level at date t, I_t is receipts (other than interest receipts) at date t, and M_t is the nominal quantity of money brought into period t. (11) requires that the excess (shortfall) of receipts over the rental of inputs and accumulation of money balances must be used to purchase (sell) the dollar-denominated assets. The sum of initial assets and initial money balances $(b_1 + M_1)$ is given and agents must obey a terminal condition:

$$b_{T+1} + M_{T+1} \geq 0$$

The T period-by-period budget constraints (11), together with the terminal condition, imply a single intertemporal budget constraint (12):

$$\sum_{t=1}^{T} Q_t [q_t^{'} X_t + R_t m_t] = \sum_{t=1}^{T} Q_t I_t + (1 + R_1)(\frac{b_1}{P_1} + m_1) \tag{12}$$

where Q_t 's are real interest rate factors.

<u>Proposition 9</u> An agent that chooses sequences of real money balances $\{m_t\}$ and other inputs $\{X_t\}$ so as to minimize the present value of costs (the LHS of (12)) subject to a given dynamic production level V_0 has period-by-period demands for money and other inputs that are identical to the static derived demands.

The derived demand functions for the intertemporal cost minimization problem are functions of V_0, the sequence of nominal interest rates $\{R_t\}$, the sequence of rental rates $\{q_t\}$, and the sequence of financial productivities $\{A_t\}$. We can consider, for example, the demand for money in the initial period:

$$m_{i1} = \tilde{L}[V_0, R_1, q_1, A_{i1}, \ldots, R_T, q_T, A_{iT}] \tag{13}$$

Proposition 9, however, shows that period-by-period money demands are related to other contemporaneous variables (y, R and q) in the same way as in the static problem. A conceptual difference with the intertemporal formulation is that current production is "endogenous" so that conventional money demand relationships such as (4) are not, strictly speaking, "derived demands."

In the intertemporal model, we can distinguish a **wealth elasticity of money demand** from our previously defined production, cost input, and income elasticities. First, consider the intertemporal version of the cost elasticity of demand. As in the static case, this can be defined by composing the relationship between (dynamic) cost and (dynamic) production V_0 with the derived demand function (13). If we define wealth to be the RHS of the intertemporal budget constraint (12), the wealth must equal dynamic cost.[20] Denote the wealth elasticity of (initial) money demand β_W,

$$\beta_W = \frac{\partial \tilde{L}[V_0, R_1, q_1, A_{i1}, \ldots, R_T, q_T, A_{iT}]}{\partial V_0} \frac{dV_0}{dW_1} \frac{W_1}{m_1} \tag{14}$$

(14) is the elasticity of first period money holdings with respect to initial wealth. If the dynamic production function were recursive, then we could imagine the agent making decisions about current money balances and other inputs period-by-period in a time consistent fashion.[21] Then, in an infinite horizon setting with constant prices, (14) describes the wealth elasticity of money demand for any period (eg., the elasticity of m_4 with respect to W_4).

VII. Conclusion

Money is assumed to enter a firm or a household's production function. Money in the production function is not the best way to think about all issues in monetary economics. Complete knowledge about the form of the production function is not enough to describe many of the details of an agent's monetary dealings, but neither is

---

[20] To be a little less abstract, consider the household case with period-by-period household production given by $f(c,l,m)$ where c is expenditures on market goods, l is leisure and m is real money balances. Then the RHS of (15) is the present value of full income (wages time the time endowment) plus initial assets. The LHS, "dynamic cost," is the present value of expenditures on market goods, leisure, and the "rental" of money. Wealth might be called "full wealth" in this case since it includes the present value of leisure.

[21] Examples of recursive dynamic production functions are the exponentially (E S) discounted utility function $V_0 = u(y_1) + \beta u(y_2) + \beta^2 u(y_3) + \beta^3 u(y_4)\ldots$ or nonlinear recursive aggregators such as those used by Koopmans (1960).

it enough to describe how a firm treats its workers, how a household cooks a meal, or how a company installs a new copier. In Stanley Fischer's words, "to know that a physical production function is Cobb-Douglas is not to know how to run a factory" (1974, p. 525).

Nevertheless, the formulation is shown to be useful for certain theoretical and empirical problems. On the theoretical side, treating money as an input to production allows one to rigorously define the "demand for money" and to discuss its properties using the standard tools of microeconomics. For a variety of microeconomic - or should we say picoeconomic - motivations for the usefulness of money, money demand is thought of as a derived demand. Concepts that appear in the empirical literature - such as the income elasticity, sales elasticity and consumption elasticities - are related in an explicit way. Because one goal of empirical work on money demand is to distinguish some of the picoeconomic stories, it is convenient to have an theoretical framework that embodies each story as a special case.

Mathematical Appendix

This appendix proves the Lemmas and Propositions from the text for those cases of the cost minimization problem when (i) the output constraint $y = f(X, T, \lambda)$ is binding and (ii) optimal input demands are interior. Attention is restricted to situations where (i) is satisfied because it is our belief that these are the empirically relevant ones. For the sake of brevity, we omit proofs for situations where solutions fail to be interior.

Lemma 1 $\quad d\Omega / dA < 0$

Proof Fix y and the rental rates and define X^* and m^* to be the cost minimizing input demands when the level of financial technology is A. T^* is defined as the associated quantity of transactions services:

$$T^* \equiv \phi(m^*, X^*. A)$$

Because $\partial\phi / \partial m > 0$, the implicit function theorem guarantees that we can define a function h(X,A) and that $\partial h / \partial A < 0$:

$$T^* \equiv \phi[h(X^*, A), X^*, A]$$

If we choose a higher level of financial technology $A + \partial A$, we know that the quantity of money m' required to achieve the level of transactions services T^* when $X = X^*$ is less than m^*. Since $R > 0$, the cost function therefore cannot increase when the level of financial technology increases from A to $A + \partial A$.

Lemma 2 The relative cost of the production of transactions services→0 as $\lambda_f \to 0$.

Proof For this proof and the proof of Lemmas 3 and 4, we distinguish those elements of the vector X that appear in the final production function from those that appear in the production of transactions services. The former are denoted by the vector X_f and the latter by X_ϕ. The corresponding rental rates are denoted q_f and q_ϕ.

We consider the case when the first order conditions of the cost minimization problem with respect to X_ϕ and m hold with equality and use μ to denote the inverse of the LaGrange multiplier. The first order conditions are:

$$\frac{\partial f}{\partial T} D_{X_\phi} \phi(m, X_\phi, A) = \mu q_\phi$$

$$\frac{\partial f}{\partial T}\frac{\partial \phi}{\partial m} = \mu R$$

Summing the first order conditions term by term, dividing by r, and taking the limit as $\lambda_f \to 0$:

$$\lim_{\lambda f \to 0}[\frac{\partial f}{\partial T}\frac{T}{y}][\frac{y}{r}][\Sigma_j \frac{\partial \phi}{\partial X_j}\frac{X_j}{T} + \frac{\partial \phi}{\partial m}\frac{m}{T}] = \lim_{\lambda f \to 0} \mu[\Sigma_j \alpha_j + \alpha_m]$$

$$0 = \lim_{\lambda f \to 0} \mu[\alpha_\phi + \alpha_m]$$

$$0 = \lim_{\lambda f \to 0} [\alpha_\phi + \alpha_m]$$

where the sums are over the X's that enter the transactions services production function ϕ. α_m is money's share of cost while α_ϕ is the share of those X's that appear in ϕ. On the LHS of the first line, we have three terms. The first goes to zero by Assumption 5. The second must be bounded for $y>0$ while Assumption 6 requires that the third be bounded. Since μ is bounded below as $\lambda_f \to 0$, the cost share of the inputs required to produce transactions services must approach zero as $\lambda_f \to 0$. Since cost shares must be nonnegative, it follows that the cost share of any particular input to the production of transactions services approaches zero as $\lambda_f \to 0$.

Lemma 3 For any fixed A and any fixed prices, $\lim\limits_{\lambda_f \to 0} \frac{\partial\Omega}{\partial A}\frac{A}{r} = 0$

Proof By definition of the cost function and the cost shares, we can write the cost function as:

$$[1-\alpha_\phi(y,q_f,q_\phi,R,A,\lambda)-\alpha_m(y,q_f,q_\phi,R,A,\lambda)]\cdot\Omega(y,q_f,q_\phi,R,A,\lambda)$$
$$=q_f{}' H_{X_f}(y,q_f,q_\phi,R,A,\lambda)$$

where α_ϕ is the sum of the cost shares of those inputs which enter the production function ϕ and H_{Xf} is a vector of the derived demands of those inputs which appear in the final production function. From the above expression, we compute the elasticity of the cost function with respect to A:

$$\frac{\partial\Omega}{\partial A}\frac{A}{r} = \frac{q_f{}'\frac{\partial H_{X_f}}{\partial A}\frac{A}{r} + (\alpha_\phi + \alpha_m)\frac{\partial(\alpha_\phi+\alpha_m)}{\partial A}\frac{A}{\alpha_\phi+\alpha_m}}{1-\alpha_\phi-\alpha_m}$$

According to Assumption 7, the elasticity of $\alpha_\phi + \alpha_m$ with respect to A is bounded above. Taking the limit as $\lambda_f \to 0$,

$$\lim_{\lambda_f \to 0}\frac{\partial\Omega}{\partial A}\frac{A}{r} = \lim_{\lambda_f \to 0}\Sigma_{X_f}\alpha_j\frac{\partial H_j}{\partial A}\frac{A}{X_j}$$

where the sum on the RHS is taken over those inputs which appear in the final production function. To evaluate the sum on the RHS, we totally differentiate the production function and substitute the first order conditions for cost minimization:

$$0=\mu\Sigma_{X_f}\alpha_j\frac{\partial H_j}{\partial A}\frac{A}{X_j}+[\frac{\partial f}{\partial T}\frac{T}{y}][\frac{d\phi}{dA}\frac{A}{T}]$$

where μ is the LaGrange multiplier for the cost minimization problem and $d\phi/dA$ is the total derivative of ϕ with respect to A, including the indirect effects of A on ϕ via the derived demands for X_ϕ and m. As $\lambda_f\rightarrow 0$, the elasticity of f w.r.t T approaches zero by Assumption 5, the elasticity of ϕ is bounded above by Assumption 8, and μ>0. Therefore the sum approaches zero which, according to (A-1), proves that the elasticity of the cost function with respect to A approaches zero.

Proposition 1 The derived demand for money is nonincreasing in the nominal interest rate.

$$\frac{\partial L(y,R,q,A)}{\partial R}<0$$

Proof See text.

Proposition 2 The production elasticities of money demand, $\beta_M\equiv(\partial M/\partial r)r/M$, and cost elasticities of money demand, $\beta_L\equiv(\partial L/\partial y)y/L$, have the same sign.

Proof Apply the chain rule to the definition of M:

$$\frac{\partial M}{\partial r}=\frac{\partial L}{\partial y}\frac{\partial\Omega^{-1}}{\partial r}=\frac{\partial L/\partial y}{\partial\Omega/\partial y}$$

The second equality follows from the implicit function theorem. The fact that Ω is increasing in y proves the Proposition.

Proposition 3 If the cost weighted average of the elasticities of the derived demands for the other inputs X with respect to production y is unity and money is a normal good ($\beta_L>0$), then the cost elasticity of Marshallian money demand is closer to one than is the production elasticity of the derived demand for money. The difference between the production and cost elasticities shrinks as money's share of cost goes to zero.

$$\sum_{j=1}^{J}\left(\frac{\alpha_j}{1-\alpha_m}\right)\left(\frac{\partial X_j}{\partial y}\frac{y}{X_j}\right)=1 \ \textit{implies}$$

$$\textit{(i)}\ |\beta_M-1|<|\beta_L-1|$$

$$\textit{(ii)}\ \lim_{\lambda_f\to 0}|\beta_M-\beta_L|=0$$

Proof From Proposition 2, $\beta_M=\beta_L/\varepsilon_{\Omega_y}$, where ε_{Ω_y} is the elasticity of the cost function with respect to y. Applying the chain rule to the cost function, we compute the Marshallian scale elasticity as a function of the scale elasticities of each of the J+1 derived demand elasticities:

$$\beta_M=\frac{\beta_L}{(1-\alpha_m)\sum_{j=1}^{J}[\frac{\alpha_j}{1-\alpha_m}][\frac{\partial X_j}{\partial y}\frac{y}{X_j}]+\alpha_m\beta_L}=\frac{\beta_L}{1-\alpha_m+\alpha_m\beta_L}\to\beta_L$$

The second equality follows from the assumption that the derived demand scale elasticities, except for β_L, average 1. Item (i) of the Proposition follows from the third term and the fact that money's share of cost is nonnegative. Item (ii) follows from the third term and the fact that $\alpha_m\to 0$ as $\lambda_f\to 0$.

Proposition 4 If money is a normal good, the elasticity of the Marshallian demand for money with respect to the level of financial technology is greater than the derived demand elasticities. The two elasticities are equivalent as money's share of cost approaches zero.

$$\frac{\partial M}{\partial A}\frac{A}{m}>\frac{\partial L}{\partial A}\frac{A}{m}$$

$$\lim_{\lambda_f\to 0}\left|\frac{\partial M}{\partial A}\frac{A}{m}-\frac{\partial L}{\partial A}\frac{A}{m}\right|\to 0$$

Proof Applying the chain rule to the definition of the Marshallian money demand function and using the implicit function theorem to evaluate the derivatives of Ω^{-1}:

$$\frac{\partial M}{\partial A}\frac{A}{m}=\frac{\partial L}{\partial A}\frac{A}{m}-\frac{\partial\Omega/\partial A}{\partial\Omega/\partial y}\frac{A/r}{y/r}\frac{\partial L}{\partial y}\frac{y}{m}$$

$\partial L / \partial y > 0$ because money is assumed to be normal. That $\partial \Omega / \partial y > 0$ is a property of cost functions. Lemma 1 demonstrates that $\partial \Omega / \partial A < 0$, from which the first item of the proposition follows. Similarly the difference between the Marshallian and derived demand elasticities must approach zero because $\partial \Omega / \partial A < 0$ approaches zero with λ_f.

Proposition 5 The jth input elasticity of money demand is proportional to the production elasticity of money demand, where the factor of proportionality is the inverse of the production elasticity of the derived demand for X_j:

$$\beta_j \equiv \frac{\beta_L}{\dfrac{\partial H_j}{\partial y}\dfrac{y}{X_j}}$$

Proof Apply the chain rule to the definition of the function $g_j(X_j, q, R, A)$, using the implicit function theorem to evaluate $\partial H^{-1} / \partial X_j$.

Proposition 6 If J = 1 and X is a normal good, the cost elasticity of money demand is closer to one than is the input elasticity of money demand. If J = 1 and β_j is bounded from above, then the cost elasticity approaches the input elasticity as the expenditure on money takes a smaller share of cost.

$$J = 1, \beta_j > 0 \;\; implies \;\; |\beta_j - 1| > |\beta_M - 1|$$
$$J = 1, \beta_j \; bdd \;\; implies \;\; \lim_{\lambda f \to 0} |\beta_j - \beta_M| = 0$$

Proof From the first equality displayed in the proof of Proposition 3,

$$\beta_M = \frac{\beta_L}{(1-\alpha_m)(\dfrac{\partial X}{\partial y}\dfrac{y}{X}) + \alpha_m \beta_L}$$

From proposition 5, the RHS can be expressed as a function of

$$\beta_M = \frac{\beta_j}{(1-\alpha_m) + \alpha_m\ \beta_j}$$

The first item of the proposition follows from the positivity of β_j. The second item follows because $\alpha_m\beta_j \to 0$ as $\lambda_f \to 0$

Proposition 7 The (point) price elasticities of the derived demands for money and the other inputs are linear combinations of the price elasticities of the expansion paths. The weights depend on the cost shares and the input elasticities:

$$\Sigma = \left[I - \frac{\zeta\alpha'}{\alpha'\zeta} \right] \Pi$$

where Σ is the matrix of price elasticities of the derived demand functions (σ_{ij} is the elasticity of the derived demand for input i with respect to the price of good j), Π is the matrix of price elasticities of the expansion path π_{ij} is the elasticity of the expansion path for input i with respect to the price of good j), α is a column vector of cost shares (expenditure on an input divided by total cost), and ζ is a column vector of scale elasticities of the expansion paths.

Proof A standard result from microeconomic theory is that price elasticities of derived demands satisfy three properties: (i) symmetry, (ii) adding up, and (iii) homogeneity:

$$\begin{aligned}
&(i)\ \alpha_i\sigma_{ij} = \alpha_j\sigma_{ji},\ \ all\ i,j \in \{1,2,\dots,J\} \\
&(ii) \sum_{j=1} \alpha_j\sigma_{ji} + \alpha_m\sigma_{mi} = 0,\ \ all\ i \in 1,2,\dots,J,m \\
&(iii) \sum_{j=1} \sigma_{ij} + \sigma_{im} = 0,\ \ all\ i \in 1,2,\dots,J,m
\end{aligned}$$

where σ_{ij} is the elasticity of the derived demand for good X_i with respect to the rental rate q_j. $\sigma_{mj} \equiv \partial L / \partial q_j$ is the elasticity of the derived demand for money with respect to q_j. σ_{im} is the elasticity of the derived demand for good X_j with respect to the rental rate of money, R. From the definition of the expansion path, we compute jth price elasticity of the expansion path for good i (where X_1 is the scale variable) as a function of σ_{ij}, σ_{1j} and the elasticity of the projection of the expansion path into the (x_1,x_j) plane:

$$\pi_{ij} = \sigma_{ij} - \zeta_i\sigma_{1j} / \zeta_1 \tag{A-1}$$

This follows from the implicit function theorem and the chain rule of calculus. π_{ij} denotes the jth price elasticity of the expansion path for good i (where X_1 is the scale variable) and ζ_i the elasticity of the projection of the expansion path into the (x_1,x_j) plane. Substituting (A-1) into the adding up condition (ii), we find:

$$\sigma_{1k} = \zeta_1 \frac{\sum_{j=1}^{J}\alpha_j\pi_{jk} + \alpha_m\pi_{mk}}{\sum_{j=1}^{J}\alpha_j\zeta_j + \alpha_m\zeta_m}$$

Substituting this expression back into (A-1), we find:

$$\sigma_{ik} = \pi_{ik}\zeta_k \frac{\sum_{j=1}^{J}\alpha_j\pi_{jk} + \alpha_m\pi_{mk}}{\sum_{j=1}^{J}\alpha_j\zeta_j + \alpha_m\zeta_m}$$

The matrix version of the above expression is:

$$\Sigma = \left[I - \frac{\zeta\alpha'}{\alpha'\zeta} \right]\Pi$$

where Σ and Π are formed in the obvious way from their elements σ and π. ζ is $(1,\zeta_2,\zeta_3,\ldots,\zeta_J,\zeta_M)'$.

<u>Proposition 8</u> If money and the input X_j are normal goods and X_j does not appear in the production function for transactions services, the elasticity the projection of the expansion path into the (X_j,m) plane with respect to the level of financial technology is greater (less) than the derived money demand elasticity if the elasticity of the derived demand for X_j is less (greater) than 0. If, in addition, this second elasticity approaches zero as $\lambda_f \rightarrow 0$, then the two elasticities approach each other.

$$\frac{\partial g_j}{\partial A}\frac{A}{m} \gtrless \frac{\partial L}{\partial A}\frac{A}{m} \quad as \quad \frac{\partial H_j}{\partial A} \gtrless 0$$

$$\lim_{\lambda_f \to 0}\left|\frac{\partial g_j}{\partial A}\frac{A}{m} - \frac{\partial L}{\partial A}\frac{A}{m}\right| = 0 \ \ when \lim_{\lambda_f \to 0} \frac{\partial H_j}{\partial A} = 0$$

Proof Applying the chain rule to the definition of the money demand function g_j and using the implicit function theorem to evaluate the derivatives of H_j^{-1}

$$\frac{\partial g_j}{\partial A}\frac{A}{m} = \frac{\partial L}{\partial A}\frac{A}{m} - \frac{\partial H_j / \partial A}{\partial H_j / \partial y}\frac{\partial L}{\partial y}\frac{A}{m}$$

Both items of the proposition follows from the normality assumption $\partial L / \partial y$ and $\partial H_j / \partial y > 0$.

Proposition 9 An agent that chooses sequences of real money balances $\{m_t\}$ and other inputs $\{X_t\}$ so as to minimize the present value of costs (the LHS of (15)) subject to a given dynamic production level V_0 has period-by-period demands for money and other inputs that are identical to the static derived demands.

Proof Because both the objective function and the constraints are time separable, first-order conditions can be grouped by period and are identical to the static first-order conditions (where the first-order conditions equate static marginal rates of transformation with intraperiod relative prices).

Notional Appendix

(symbols appear in roughly the order that they first appear in the text)

Symbol	Definition
y_{it}	level of production by agent i at date t
m_{it}	real money balances held by agent i at date t
X_{it}	other inputs used in production by agent i at date t
J	dimension of the vector X
T_{it}	transactions services

λ_f	parameter dictating the importance of transactions services in the final production function
A_{it}	level of financial sophistication of agent i at date t
$f(X,T,\lambda_f)$	production function for final goods
$\phi(m,X,A)$	production of transactions services
q_t	J*1 vector of rental rates of the inputs X
R_t	nominal interest rate
r_{it}	cost of production
$\Omega(y,R,q,A)$	cost of producing y when prices are R, q and financial technology is A
$L(y,R,q,A)$	derived demand for money
$\beta_L \equiv (\partial L/\partial y)/y/L$	production elasticity of derived money demand
$M(r,R,q,A)$	Marshallian demand for money
$\beta_M \equiv (\partial M/\partial r)/r/M$	cost elasticity of Marshallian money demand
$\alpha_j \equiv (q_j X_j)/r$	input j's share of cost
$\alpha_m \equiv Rm/r$	money's share of cost
$\varepsilon_{LR} = (\partial L/\partial R)R/L$	interest elasticity of derived money demand
$\varepsilon_{MR} = (\partial M/\partial R)R/M$	interest elasticity of Marshallian money demand
$H_i(y,R,q,A)$	the derived demand for x_i
$H_i^{-1}(y,R,q,A)$	inverse of the derived demand for the derived x_i
$\zeta_i = (\partial H_i/\partial y)y/X_i$	production elasticity of the derived demand for X_i
$g_i(X_i,R,q,A)$	projection of the expansion path into the (m, X_i) plane
$\beta_i \equiv (\partial g_i/\partial X_i)X_i/g_i$	elasticity of the (m, X_i) projection of the expansion path
c_{it}	consumption of household i at date t (an input into household production)
λ_ϕ	parameter dictating the importance of time in the production function for transactions services elasticity of substitution of other inputs for transactions
γ	elasticity of substitution of other inputs for transactions
β	scale elasticity of the demand for money
d_{it}	demand deposits of household i at date t
δ	ratio of demand deposits to the sum of currency and demand deposits

ψ_ϕ	elasticity of substitution of money for time (or, in Section VI.D, currency for demand deposits)
π_ϕ	labor's share of the cost of production of transactions services
$N_t(f)$	number of firms in the economy at date t
$N_t(h)$	number of households in the economy at date t
N_t	number of people in the economy at date t
$\eta_t(f) \equiv N_t(f) / N_t$	number of firms per person
$\eta_t(h) \equiv N_t(h) / N_t$	number of households per person
P_t	price level at date t
I_{it}	income of household i at date t
$I_t(h)$	average household income at date t
$y_t(f)$	average sales of firms at date t
$m_t(h)$	average household real balances at date t
$m_t(f)$	average real balances of firms at date t
$v_t \equiv N_t(f)y_t(f)/N_t(h)I_t(h)$	aggregate sales as a fraction of aggregate income
M_t	nominal money stock at date t
ω	share of the aggregate money stock held by firms (in some benchmark year)
V_0	intertemporal production
$V(y_1,\ldots,y_T)$	intertemporal production function
b_t	quantity of dollar-denominated assets held at date t
Q_t	real discount factor
β_w	wealth elasticity of money demand

Hiroshi Fujiki: Research Division I, Institute for Monetary and Economic Studies, Bank of Japan

Casey B. Mulligan : Department of Economics, University of Chicago

References

Allais, Maurice, *Economie et Interet,* Paris: Imprimerie Nationale, 1947.

Barro, Robert J., "Integral Constraints and Aggregation in an Inventory Model of Money Demand," *Journal of Finance,* 31, 1976, pp. 77-87.

Baumol, William J., "The Transactions Demand for Cash: An Inventory Theoretic Approach," *Quarterly Journal of Economics*, 66 (4), 1952, pp. 545-556.

Chan, Anthony and Carl R. Chen, "Money Demand: Aggregation Bias and the Dispersion of Income," *Quarterly Review of Economics and Finance*, 32 (2), Summer 1992, pp. 3-29.

Clower, Robert, "A Reconsideration of the Microfoundations of Monetary Theory," *Western Economic Journal*, 6, 1967, pp.1-9.

Cover, James P. and Donald L. Hooks, "Money Demand and Income Distribution: Evidence from Annual Data," *Review of Economics and Statistics*, 75 (3), August 1993, pp. 520-523.

Deaton, Angus and John Muellbauer, *Economics and Consumer Behavior,* Cambridge: Cambridge University Press, 1980.

Feenstra, Robert C., "Functional Equivalence between Liquidity Costs and the Utility of Money," *Journal of Monetary Economics*, 17, 1986, pp. 271-291.

Fischer, Stanley, "Money and the Production Function," *Economic Inquiry*, 12, 1974, pp. 517-533.

Friedman, Milton and Anna J. Schwartz, *Monetary Trends in the United States and the United Kingdom: Their Relation to Income, Prices, and Interest Rates, 1867-1975,* Chicago: University of Chicago Press, 1982.

Fujiki, Hiroshi and Casey B. Mulligan, "A Structural Analysis of Money Demand: Cross-Sectional Evidence from Japan," Bank of Japan IMES Discussion Paper 96-E-7, January 1996.

Fuss, Melvyn and Daniel McFadden, *Production Economics: A Dual Approach to Theory and Applications*, New York: North-Holland, 1978.

Karni, Edi, "The Transactions Demand for Cash: Incorporation of the Value of Time into the Inventory Approach," *Journal of Political Economy*, 81, 1973, pp. 1216-1225.

Koopmans, Tjalling.C., "Stationary Ordinal Utility and Impatience," *Econometrica*, 28, 1960, pp. 287-309.

Lancaster, Kelvin. J., "A New Approach to Consumer Theory," *Journal of Political Economy*, 74, 1966.

Lucas, Robert E., Jr., "On the Welfare Cost of Inflation," Working paper, University of Chicago, February 1994.

———, "Money Demand in the United States: A Quantitative Review," *Carnegie-Rochester Conference Series on Public Policy*, 29, 1988, pp.137-168.

Lucas, Robert E., Jr. and Nancy L. Stokey, "Money and Interest in a Cash-in-Advance Economy," *Econometrica*, 55 (3), May 1987, pp. 491-513.

Mairesse, Jacques, "Time-Series and Cross-Sectional Estimates on Panel Data: Why are They Different and Why Should They be Equal?" in J. Hartog, G. Ridder and J. Theeuwes, eds., *Panel Data and Labor Market Studies,* Elsevier Science Publishers B.V., 1990.

McKenzie, Collin R., "Money Demand in an Open Economy," *Journal of the Japanese and International Economies*, 6, 1992, pp. 176-198.

Meltzer, Allan H., "The Demand for Money: Evidence from Time Series," *Journal of Political Economy*, 71 (3), 1963, pp. 219-46.

——————,"The Demand for Money: A Cross-Section Study of Business Firms," *Quarterly Journal of Economics* , 77 (3), August 1963, pp. 405-422.

Miller, Merton M. and Daniel Orr, "A Model of the Demand for Money by Firms," *Quarterly Journal of Economics*, August 1966, 80, pp. 413-435.

Mulligan, Casey B., "Scale Economies, the Value of Time, and the Demand for Money: Longitudinal Evidence from Firms," manuscript, University of Chicago, July 1994.

Mulligan, Casey B. and Xavier Sala-i-Martin, "U.S. Money Demand: Surprising Cross-Sectional Estimates," *Brookings Papers on Economic Activity,* 2, 1992, pp. 285-343.

Poole, William, "Monetary Policy Lessons for Recent Inflation and Disinflation," *Journal of Economic Perspectives*, 2, 1988, pp. 73-100.

Rasche, Robert H., "Equilibrium Income and Interest Elasticities of the Demand for M1 in Japan," *Monetary and Economic Studies,* 8 (2), Institute for Monetary and Economic Studies, Bank of Japan, 1990, pp. 31-55.

Saving, Thomas R., "Transactions Costs and the Demand for Money," *American Economic Review*, 61 (3), June 1971, pp. 407-420.

Suzuki, Yoshio, ed., *The Japanese Financial System*, Oxford: Oxford University Press, 1991.

Tobin, James, "The Interest-Elasticity of Transactions Demand for Cash," *Review of Economics and Statistics*, 38 (3), 1956, pp. 241-47.

——————, "The Interest Elasticity of the Transactions Demand for Money," *Review of Economics and Statistics*, 38, August 1956.

Yoshida, Tomoo and Robert H. Rasche, "The M2 Demand in Japan: Shifted and Unstable?" *Monetary and Economic Studies,* 8 (2), Institute for Monetary and Economic Studies, Bank of Japan, 1990, pp. 9-30.

Whalen, Edward L., "A Rationalization of the Precautionary Demand for Cash," *Quarterly Journal of Economics*, 80, 1966, pp. 314-324.