

Discussion Paper 7915

September 1979

A Short Note on Economies of Scale
in Transactions Demand for Money

by

Ricardo D. Ferrer

Note: UPSE Discussion Papers are preliminary versions circulated privately to elicit critical comment. They are protected by the Copyright Law (PD No. 49) and not for quotation or reprinting without prior approval.

Abstract

This note points to a possible error in Baumol's, and later Brunner's and Meltzer's formulation of the transactions demand for cash. The argument of the paper is rather simple. After showing that the optimum lot size of cash withdrawals, C , is independent of deposits I , one can no longer assume as if I can vary continuously. In other words, differentiation of a cost function with respect to I is illegitimate if I can only vary in multiples of C . On the basis of this argument the transactions demand for cash is reformulated, and the basic result is that the demand function is discontinuous. A method to test the hypothesis is indicated which suggests that Meltzer's cross-section test of the Baumol hypothesis is wrong.

A SHORT NOTE ON ECONOMIES OF SCALE
IN TRANSACTIONS DEMAND FOR MONEY*

by

Ricardo D. Ferrer

The familiar results of Baumol's analysis of the transactions demand for cash¹ has been reexamined recently² and, it appears that some modifications of the earlier results seem called for. However, these reformulations are still inadequate. Consider the more general case treated by Baumol. The individual has a cash influx of T dollars at the beginning of the period which is also equal to intended expenditures for the period. (The actual influx may be greater than T , but the difference "will be invested since, eventually, interest earnings must exceed ("brokerage") cost of investment."³ Hence attention can be focused on T . I dollars of

*Comments on an earlier draft of this paper by Dean José Encarnación and Prof. Edita Tan are gratefully acknowledged. All errors that remain are mine.

¹William J. Baumol, "The Transactions Demand for Cash: An Inventory Theoretic Approach," Quarterly Journal of Economics, LXVI (November, 1952).

²Karl Brunner and Allan H. Meltzer, "Economies of Scale in Cash Balances Reconsidered," Quarterly Journal of Economics, LXXXI (August, 1967); Russell D. Morris, "Transactions Demand for Cash," Quarterly Journal of Economics, LXXXV (August, 1971), and Syed Admad, "Transactions Demand for Money and the Quantity Theory," Quarterly Journal of Economics, XCI (May, 1977).

³Baumol, p. 547.

this amount will be invested and the rest, $T - I = R$ will be withheld, to finance transactions for R/T of the period. After this period, the investment I will be withdrawn in lot sizes C . Let i be the rate of interest, $b_w + k_w C$ be the broker's fee for withdrawals, and $b_d + k_d I$ be a broker's fee for deposits. Rational behavior in this model requires that holding of cash balances to finance the transaction T for the period be minimum.

Now, the total cost of withholding R dollars is

$$a) \quad \frac{T-I}{2} i \frac{T-I}{T} + b_d + k_d I,$$

since disbursements are continuous and average cash balances for the fraction of the period $\frac{T-I}{T} = \frac{R}{T}$ is $\frac{R}{2} = \frac{T-I}{2}$. Similarly, the total cost of obtaining cash for the rest of the period in lot sizes C is

$$b) \quad \frac{C}{2} i \frac{I}{T} + (b_w + k_w C) \frac{I}{C}$$

because the average cash holding for the remainder of the period $\frac{I}{T}$ is $\frac{C}{2}$ and I/C is the number of withdrawals.

Minimizing this cost function (equation a+b) with respect to C yields the familiar result that

$$c) \quad C = \sqrt{2b_w T/i}$$

which does not depend on I . A similar expression for R is obtained by differentiating the cost function with respect to I yielding⁴

⁴ Ibid., p. 548.

$$d) \quad R = C + T \left(\frac{k_w + k_d}{I} \right)$$

It is at this point where we think the analysis goes astray. If investment I will be withdrawn in equal lot sizes C which depends only on T , then I must be equal to integral multiples of C , i.e., $I = NC$ where $N = 1$, or 2 , or 3 , or some whole number. Thus, if we accept the first result (equation c), we must conclude that differentiation of the cost function with respect to I is illegitimate: I can no longer vary continuously. More intuitively, suppose that I is not equal to an integral multiple of C . Then it follows that the size of withdrawals cannot all be equal to $C = \sqrt{2b_w T/i}$ and hence, if C is the lot size that minimizes the cost of holding cash balances, cash withdrawals in amounts less than C will not be optimal. Or, if $I \neq NC \neq N \sqrt{2b_w T/i}$ and I is withdrawn in equal lot sizes C' , and C' is a cost minimizing lot size, then it is not true that the optimum lot size depends only on T because now it also depends on I . Thus, if we accept equation (c), then I must be equal to NC where N is a whole number.⁵ It is for this reason that we reject the validity of Baumol's equation for R (equation d) and of Brunner's and Meltzer's reformulation⁶ of the transactions

⁵Note that differentiation of the cost function with respect of C is already an approximation. Since C is the lot size, we can think of its variation as the variation of I/N as N (a whole number) decreases from a value, however large, to smaller values, or vice versa. Thus, C can only vary discretely.

⁶Brunner and Meltzer, op. cit., p. 426.

demand for cash which depends on it. However, the point raised by the latter⁷ is important in our own reformulation. The amount of cash demanded for transactions will be a weighted average of R and C .

Since we can now set $R = NC$, where $N = 1, 2, \dots$ a whole number, we must look into the conditions under which each value of N holds.⁸ Suppose $N = 2$, since the case where $N = 1$ seems trivial. If $N = 2$, then a transactor saves an amount, S_2 let us say, given by

$$1) \quad S_2 = b_w + k_w C + k_d C = b_w + C(k_w + k_d)$$

This follows from the fact that the individual has one less withdrawal to make saving him $b_w + k_w C$ and since he invests C less, he will spend Ck_d less on deposit expenses.

However, if $N = 2C$, for the period $2C/T$ his average money balances would be $R/2 = 2C/2 = C$, greater than the average balances he will maintain for the same period equal to $\frac{C}{2}$ if $R = C$. Hence, there is an opportunity cost, L_2 , given by

$$2) \quad L_2 = (C - \frac{C}{2}) i \frac{2C}{T} = \frac{C}{2} i \frac{2C}{T} = C^2 \frac{i}{T}$$

⁷Ibid.

⁸We are proceeding as if T can be broken up integrally by $C = \sqrt{2b_w T/i}$, so that if this were so, then R must at least be equal to C . But see the Appendix on this point.

A rational transactor will withhold $R = 2C$ if savings S_2 is larger than income lost, L_2 , i.e.

$$3) \quad S_2 > L_2,$$

which upon substitution of (1) and (2) gives

$$3a) \quad b_w + C(k_w + k_d) > C^2 \frac{i}{T}$$

But $C = \sqrt{2b_w T/i}$, irrespective of the value of R . Hence,

$$3b) \quad b_w + \sqrt{2b_w T/i} (k_w + k_d) > (\sqrt{2b_w T/i})^2 \frac{i}{T}$$

This expression, upon rearrangement gives us

$$3c) \quad T > \frac{b_w i}{2} \left(\frac{1}{k_w + k_d} \right)^2$$

With the equality sign we have a value of $T = T_1$ when with

$T < T_1$, $R = C$, and with $T = T_2 > T_1$, $R = 2C$.

But under what conditions would the money withheld be equal to $R = 3C$? Here we consider savings S_3 and loss L_3 and similarly compare savings against loss.

$$4) \quad S_3 = 2b_w + 2C(k_w + k_d),$$

Since two withdrawals are avoided by withholding $R = 3C$.

However, with $R = 3C$, the average money holdings for $3C$ of the period is $\frac{3C}{2}$, greater than the $\frac{C}{2}$ money holdings if $R = C$.

Hence, the loss is

$$5) \quad L_3 = \left(\frac{3}{2}C - \frac{1}{2}C \right) i \frac{3C}{T} = Ci \frac{3C}{T} = \frac{3C^2 i}{T}$$

Thus, the condition for $R = 3C$ is that

$$6) \quad S_3 > L_3, \quad ,$$

or

$$6a) \quad 2b_w + 2C(k_w + k_d) > 3 \frac{C^2 i}{T},$$

and the critical value for T is

$$7) \quad T > 2b_w i \left(\frac{1}{k_w + k_d} \right)^2$$

Note that this gives us $T > 4T_1$, where T_1 is the critical value of T when R changes from $R = C$ to $R = 2C$, i.e., between T_1 and $4T_1$, $R = 2C$. With $T > 4T_1$, $R = 3C$.

Moreover, if we consider the condition for withholding $R = 4C$ we get

$$8) \quad T > 9 \frac{b_w i}{2} \left(\frac{1}{k_w + k_d} \right)^2 \\ > 9 T_1$$

Notice that the critical values for T , such that $R = (1+N)C$, by induction, is given by

$$9) \quad T > N^2 \frac{b_w i}{2} \left(\frac{1}{k_w + k_d} \right)^2$$

$$\text{i.e., for } R = C, 0 < T < \frac{b_w i}{2} \left(\frac{1}{k_w + k_d} \right)^2 = T_1$$

$$R = 2C, T_1 < T < 4T_1$$

$$R = 3C, 4T_1 < T < 9T_1$$

$$R = 4C, 9T_1 < T < 16T_1$$

N in this case is the integral multiple of C withheld from investment over and above $R = C$.

Now consider the transactions demand for cash. If $R = C$, we get the familiar result, $M = C/2 = \frac{1}{2} \sqrt{2b_w T/i}$, for $0 < T < T_1$. But for $T_1 < T < 4T_1$, $R = 2C$. Hence,

$$\begin{aligned}
 10) \quad M &= \frac{R}{2} \cdot \frac{R}{T} + \frac{C}{2} \left(\frac{T-R}{T} \right) \\
 &= \frac{2C}{2} \cdot \frac{2C}{T} + \frac{C}{2} \left(\frac{T-2C}{T} \right) \\
 &= 2 \frac{C^2}{T} + \frac{C}{2} - \frac{2C^2}{2T} \\
 &= \frac{C^2}{T} + \frac{C}{2} \\
 &= \frac{2b_w}{i} + \sqrt{b_w T/2i} \\
 &= 2\alpha + B T^{1/2} i^{-1/2}
 \end{aligned}$$

where $\alpha = b_w/i$

$$B = \sqrt{b_w/2}$$

Or, if $R = 3C$, for $4T_1 < T < 9T_1$

$$11) \quad M = 6 \frac{b_w}{i} + \sqrt{b_w T/2i}$$

In general, if $R = mC$, $m = 1, 2, \dots$

$$\begin{aligned}
 12) \quad M &= m(m-1) \frac{b_w}{i} + \sqrt{b_w T/2i} \\
 &= (m)(m-1) \alpha + B T^{1/2} i^{-1/2}
 \end{aligned}$$

It is clear that only within the range $0 < T < T_1$ when $M = \sqrt{b_w T/2i}$ will the income elasticity of demand for money for transactions purposes be equal to $1/2$. In general, the income elasticity of demand for money is given by

$$13) \quad E_{MT} = \frac{1}{\frac{4(m-1)\alpha}{\sqrt{2b_w T/i}} + 2} = \frac{1}{A+2}$$

Thus, when $m=1$, i.e., $R=C$, $E_{MT} = 1/2$, and if $m > 1$, $E_{MT} < 1/2$ since $A > 0$, although A approaches zero as T increases. Of course, this expression is valid for the continuous segments of the function, and not at the discontinuities.

Similarly, the interest elasticity of demand for money is given by

$$14) \quad E_{MT} = - \frac{\frac{(m-1) 2b_w}{i} + \frac{1}{2} \sqrt{b_w T/2i}}{(m-1) 2b_w + \sqrt{b_w T/2i}}$$

which, with $m=1$ ($R=C$) equals $-1/2$. Since $\frac{(m-1) 2b_w}{i} > (m-1) 2b_w$ for $m > 1$, then E_{MT} is greater in absolute magnitude than $1/2$, i.e., $E_{MT} < -1/2$.

We can appreciate these results better with the aid of a diagram. With given values of b_w and i , the money demand function will appear as it does in Figure 1.

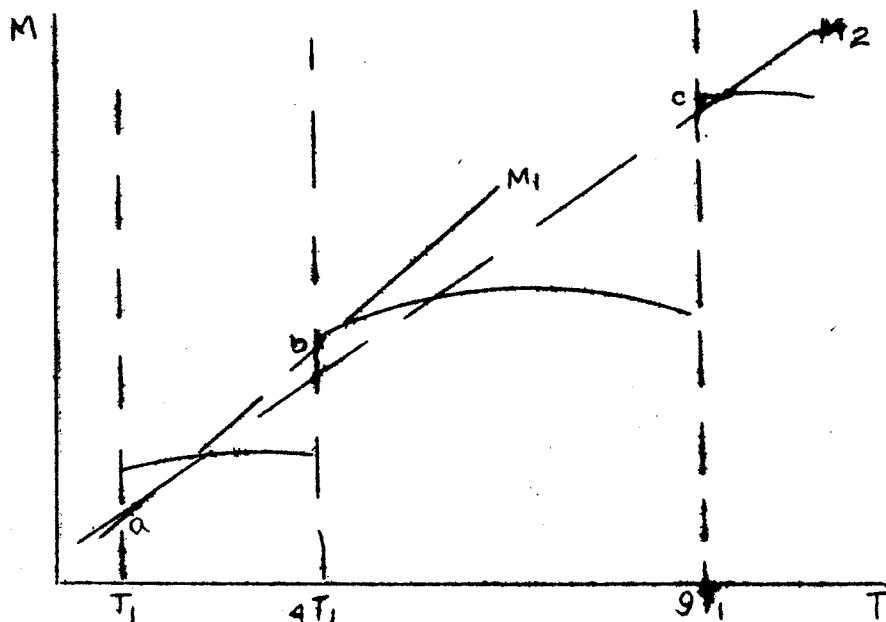


Figure 1

There arise discontinuities in the money demand function at T_1 , $4T_1$, $9T_1$ as R increases from $R = C$ to $R = 2C$; $R = 2C$ to $R = 3C$; etc. Thus, cross-section studies where the values of T range between a value lower than T_1 and a value higher than $4T_1$ will exhibit a demand function such as M_1 in Fig. 1 with a correspondingly higher income elasticity of demand for money.

For example, suppose that a and b are two points in Fig. 1 determining a demand curve for money. Let M_a be the true demand for money at a and M_b that for b .

$$15) \quad M_a = \sqrt{b_w T_1 / 2i} = \frac{b_w}{2(k_w + k_d)},$$

upon substituting the value of $T_1 = \frac{b_w i}{2} \left(\frac{1}{k_w + k_d} \right)^2$.

$$16) \quad M_b = 6 \frac{b_w}{i} + \sqrt{b_w T / 2i} = 6 \frac{b_w}{i} + \frac{b_w}{(k_w + k_d)}$$

upon substituting the value of $T = 4T_1$. Hence,

$$17) \quad \Delta M / M = \frac{M_b - M_a}{M_a} = 12 \left(\frac{k_w + k_d}{i} \right) + 1$$

In this case, however,

$$18) \quad \Delta T / T = \frac{4T_1 - T_1}{T_1} = 3$$

Therefore, the income elasticity of demand for money, E_{MT}

$$19) \quad E_{MT} = \frac{\Delta M / M}{\Delta T / T} = \frac{1}{3} \left[12 \left(\frac{k_w + k_d}{i} \right) + 1 \right]$$

Similarly, the implied elasticity of demand for a line connecting points a and c of Fig. 1 is

$$20) \quad E_{MT} = \frac{1}{4} \left[12 \left(\frac{k_w + k_d}{i} \right) + 1 \right]$$

Thus, with $k_w + k_d = \frac{1}{4} i$, E_{MT} from equation 20 is equal to 1, while equation 19 gives a value for E_{MT} of 1.33. It seems evident that cross-section studies where T ranges from a little below T_1 to a little above $9T_1$ will exhibit elasticities between 1 and 1.33, for the given relative magnitudes of $k_w + k_d$ and i . Now this presentation seems to throw some light on Meltzer's cross-section study.⁹

⁹ Allan H. Meltzer, "The Demand for Money: A Cross-Section Study of Business Firms," Quarterly Journal of Economics, LXXVII (August, 1963), No. 3, pp. 405-22.

In this study the author concludes that there are no economies of scale in the holding of transactions balances, i.e., he thinks that his study supports the conclusion that the income elasticity of demand for money is closer to 1, from the high side, than it is to Baumol's figure of $1/2$. It is clear from our results that the issue is not settled and, more importantly, that the hypothesis regarding economies of scale cannot be tested using Meltzer's procedure, since the Baumol formulation appears to require a modification similar to what we did. Moreover, Meltzer's study did not really rule out the existence of economies of scale.

Nevertheless, one may argue that if the range of T is large such that many segments of the discontinuous demand function for transactions balances are covered, cross-section studies of the demand for transactions balances will result in an estimated demand for money function similar to M^1 in Fig. 2. Obviously, since previous results show that the income elasticity of demand for money of segment 4 is lower than $1/2$ in general, M^1 will tend to approximate the earlier Baumol result that $E_{MT} = 1/2$ and, therefore, Meltzer's results still put a question mark on this current reformulation of the Baumol model. Note, however, that T cannot range far and wide. Under usual competitive assumptions regarding the industry and about production functions, an optimum size firm can be identified, and hence, a limit to T such as T in Fig. 2 exists. In the upswing of a business cycle, even the optimum-size firm will have a volume of transactions greater

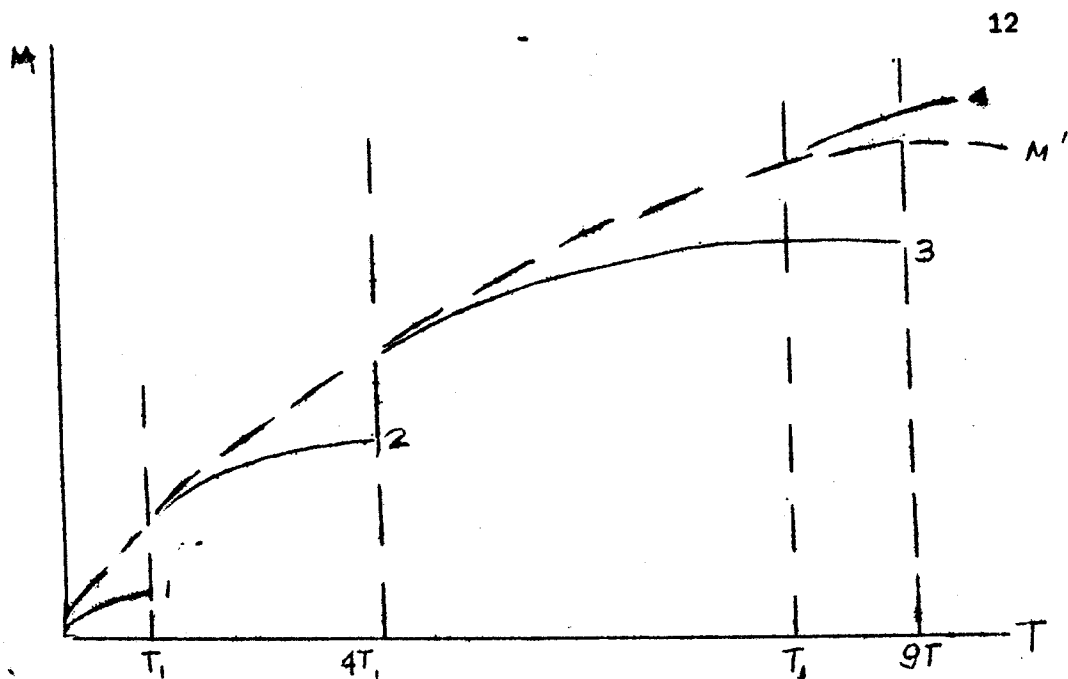


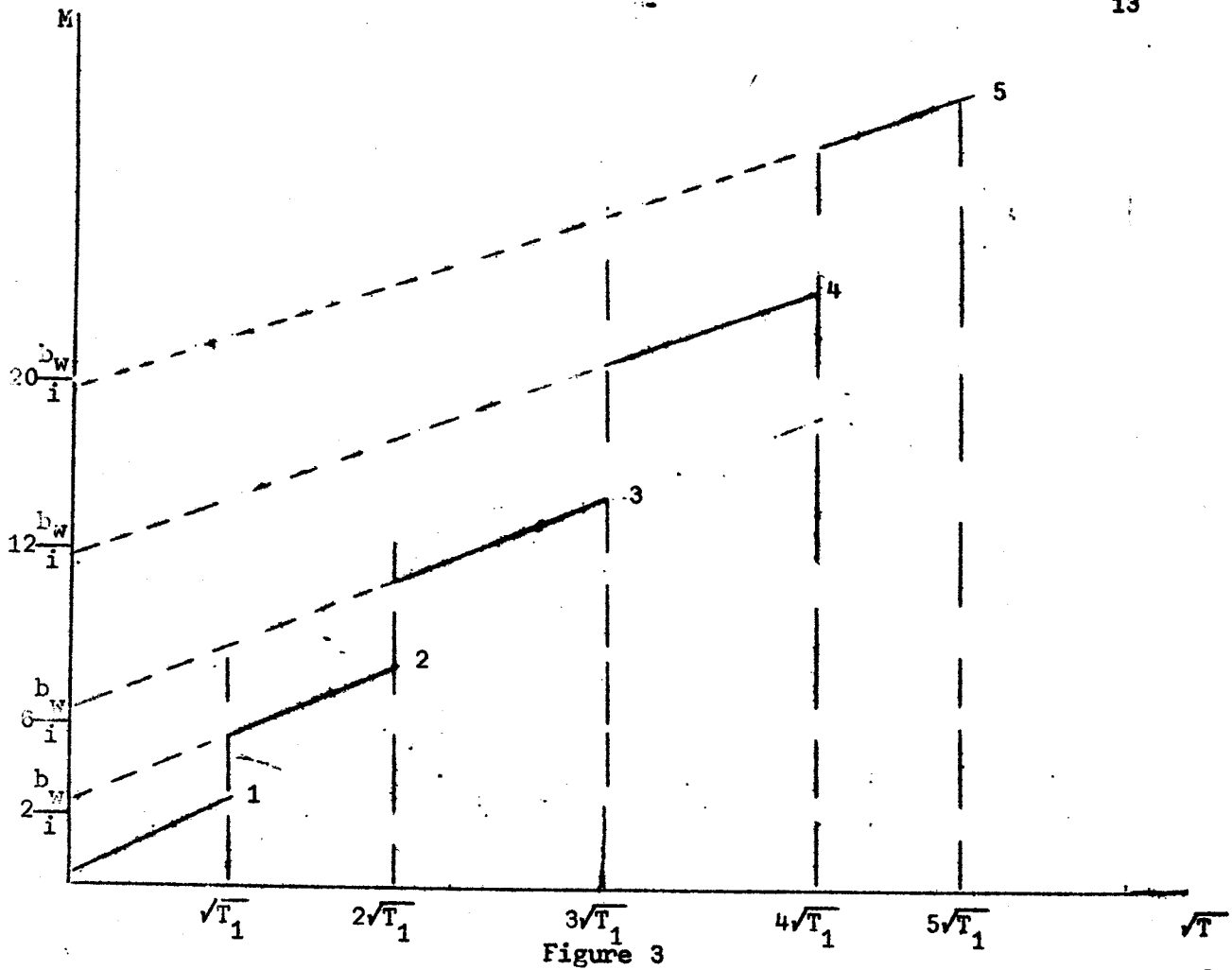
Figure 2

than $\hat{T}$ but if the boom persists long enough, the result will not be an indefinite expansion in T but in the emergence of new firms. Or, given the corporate set-up, the emergence of multi-plant firms with each of these plants, in terms of the model, being equivalent to independent firms. Thus, if only 3 segments of the true demand for money function are practically relevant Meltzer's results are, as indicated, consistent with this current reformulation of the model.

A straightforward test of our reformulation is available.

Note that equation 10 implies that the segments of the money demand function are linear in $\sqrt{T}$, as shown in Fig. 3. The slope of each of the segments is $\sqrt{b_{w/2i}}$. Thus, if one has cross-section data, one can introduce dummy variables to test the significance of these shifts. Meltzer's¹⁰ procedure of testing economies of scale by estimating the

¹⁰ Ibid.



equation $M = a + b\sqrt{T} + cT$ was not appropriate. At any rate, Meltzer discovered that 61 of the 126 cross-sections yielded positive values for b , significantly different from zero.

Appendix

Let $2b_w/i = T^*$ and $T = nT^*$ where $n > 0$. Then

$$1) \quad C = \sqrt{2b_w T/i} = \sqrt{nT^*2} = T^* \sqrt{n}$$

Therefore, the number of withdrawals N is given by

$$2) \quad N = T/C = nT^*/\sqrt{n} T^* = \sqrt{n}$$

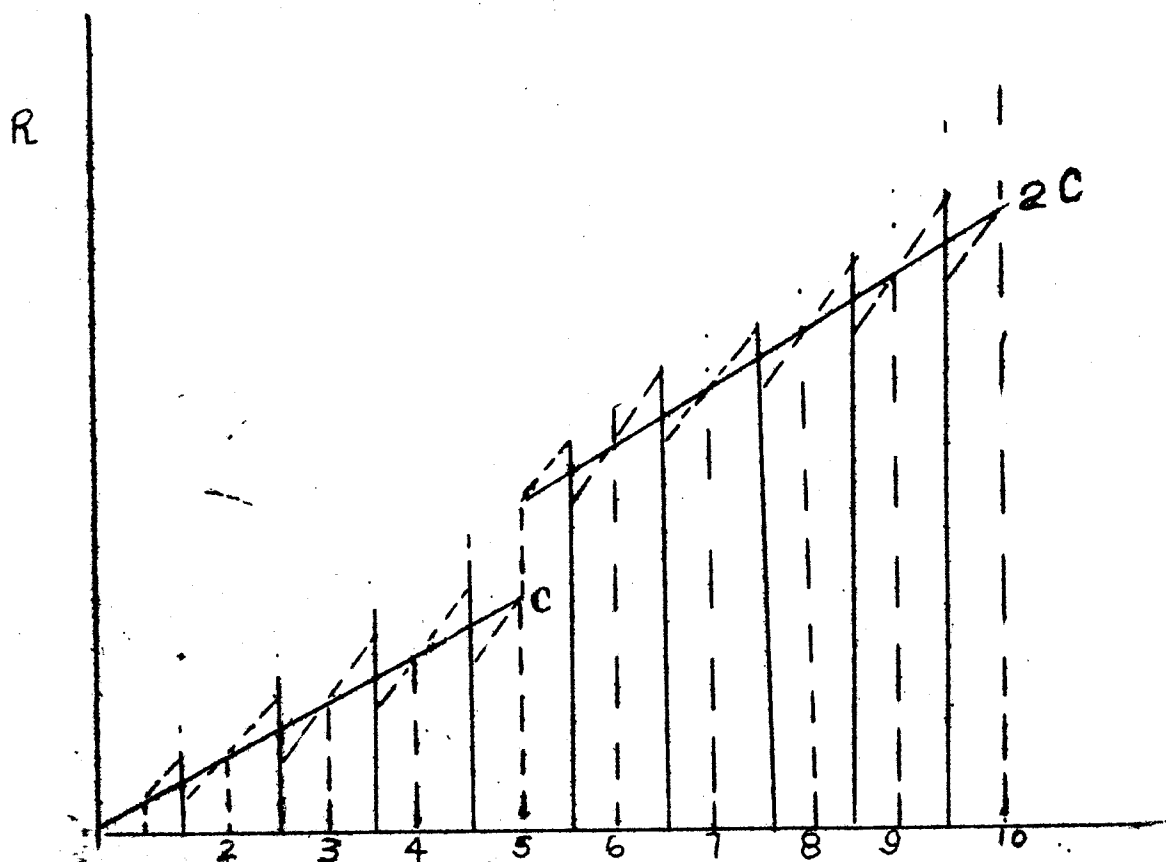
Note that N is an integral number if and only if n is a perfect square, while the only requirement of the model is that n be a positive real number. This is a paradoxical result since Baumol's model is premised on N being an integral number. This can be resolved by the fact that the algorithm of the calculus requires that the variables in the function be continuous which, in the present case, is not so (cf. fn. 5, p. 3). Thus, the discussion in the text proceeds as if N is a whole number. But what is the significance of the fact that this assumption is not true?

In the text we derived a critical value for T equal to T_1 where, with $T < T_1$, $R = C$, and with $T_1 < T < 4T_1$, $R = 2C$, and R is the amount of cash withheld from investment or deposit at time zero. Now T_1 can be expressed as

$$3) \quad T_1 = \frac{b_w}{2} \left(\frac{i}{k_w + k_d} \right)^2 = \frac{T^*}{4} \left(\frac{i}{k_w + k_d} \right)^2$$

Assume for purposes of illustration, that $k_w + k_d = .01$, and

$i = .10$, under which assumption $T_1 = 25T^*$. Under the "as if" assumption of the text, R is depicted in Fig. 1 by the solid lines C for $T < T_1 = 25T^*$ and $2C$ for $25T^* < T < 4T_1 = 100T^*$. But we



know now that only if $T = T^*, 4T^*, 9T^*$, etc., will N be an integral number or, if T is between $16T^*$ and $25T^*$, say, $T = 4C + R'$, where $0 < R' < C$, increasing from zero at $T = 16T^*$ and approaching C as T approaches $25T^*$. Now what is a transactor supposed to do? One approach is for the transactor to approximate C with C' such that $N'C' = T$ where N' is a whole number closest to $N = \sqrt{n}$. For instance between $T = T^*$ and $T = 4T^*$, C' can be constrained to equal T , i.e., $N' = 1$ till midway between T^* and

4T* when N' is made equal to 2. Given this behaviour, with $R = C'$ for values of T between T^* and $25T^*$, R will be depicted by the jagged line about C in Fig. 1. A similar line is drawn about $2C$ when R is set equal to $2C'$.

Note, incidentally, that for $T < T^*$ the value of C from equation (1) is $\sqrt{n} T^*$ for $0 < n < 1$. This, obviously is not possible. Thus, C is simply constrained to be equal to T when $T \leq T^*$.

A deeper question to ask at this point is why, if approximation to the Baumol model is to be made in the real world anyway, cannot that approximation be made theoretically, so that our criticism of Baumol's equation that $R = T(k_w + k_d)/i + C$ following the differentiation of the cost function loses its strength. Our reply to this is that at the theoretical level, either one approach is wrong or not and we think that the equation for R of Baumol is incorrect. Nevertheless, Tobin's approach¹ to the transactions demand for cash is free of the deficiency of Baumol's model. In Tobin's model $Y = T$, $b = k_w = k_d$ and $r = i$, notationally. For Tobin, the objective is to maximize net revenue by exchanging cash for bonds at $t = 0$, where the maximum time available is set equal to 1. Under this condition, bonds will break even only if $r > 2b$. Thus, of the initial

¹James Tobin, "The Interest Elasticity of the Transactions Demand for Cash," The Review of Economics and Statistics, XXXVIII (August, 1956), pp. 241-47.

cash available at $t = 0$ equal to T , at least $(2b/r)T$ must be withheld from bonds and the problem is to optimize the scheduling of n transactions (from cash into bonds and back to cash) for the effective transactions balance equal to $T(1 - 2b/r)$. Note that this amount, $T(1 - 2b/r)$ is identical to Baumol's and yet this is derived from common statistical economic reasoning (not with the aid of differentiation as in Baumol's).

The main result from Tobin's model is that the demand for cash, M_T is given by

$$4) \quad M_t = kT + C(T)$$

where k , a constant, depends on b and r , and $C'(T) > 0$, $C''(T) < 0$, and $C(T)$ breaks up the effective balances integrally without any remainder. There is no specified form for $C(T)$ such as $C(T) = \sqrt{2b_w T/r}$. Equation (4) does not show any discontinuity as in our reformulation of Baumol's. Moreover, the income-elasticity of demand implied by (4), E_{MT} , is

$$5) \quad E_{MT} = \frac{k + C'(T)}{k + \frac{C(T)}{T}}$$

Given the specification of $C(T)$ above, E_{MT} is always less than unity since $C'(T) < C(T)/T$.

Thus, this specification (equation 4) of the transactions demand for cash cannot provide justification for Brunner's and

Meltzer's model predicts that the elasticity of demand for transactions balances increases as T increases. Neither Meltzer's model nor his explanation of Meltzer's³ cross-section results can account for the inelasticity of demand for transactions balances as T increases (from the high side). However, as we have seen, our reformulation of Baumol's model can account for this empirical result. Now, what is it that we are testing when testing theoretical models? It seems to me that what we are testing mainly is whether the behaviour of economic agents postulated in the model is true or not. It may be granted that Tobin's model is more elegant than Baumol's (and Baumol's model modified, as in the text), but if Meltzer's empirical results indicate anything at all, it is that the modified model seems to have been confirmed.

²Brunner and Meltzer, loc. cit.

³Meltzer, loc. cit. Note however that this provides the basis for Meltzer's equation $M = a + b\sqrt{T} + cT$.