AN INVESTIGATION OF CASH MANAGEMENT PRACTICES AND THEIR EFFECTS ON THE DEMAND FOR MONEY

Michael Dotsey

I. INTRODUCTION

During the mid-1970s standard regressions explaining the demand for money underwent a- well-documented shift.[1] This shift was largely attributed to the adoption of more sophisticated methods of cash management practices by firms. Specifically, techniques were developed that allowed firms to perform a given level of transactions while holding lower average money balances. Therefore, for given levels of transactions income and interest rates the demand for money was lower than that implied by any historical relationships.

This article investigates the effects that a number of variables related to cash management have on the demand for money. These variables are generally suggested by analyzing specific methods of cash management and developing measures that incorporate the intensity and sophistication of these methods. All of the variables examined help explain the shift in the demand for money. Most notably, the number of electronic funds transfers made over the Federal Reserve's wire system restores stability to the estimated demand for money function.

Since the use of more sophisticated cash management techniques is believed to have its main effect on demand deposit balances, this study concentrates on demand deposits. The empirical work uses annual data over the period 1920-1979. Annual data is used to avoid controversy over the use of lagged dependent variables. The starting date represents the earliest date for which reliable data on all variables could be obtained, while the terminal date was selected to avoid the problems introduced by NOW accounts. This relatively long sample highlights the significant effects that more sophisticated cash management methods have had on the demand for demand deposits.

The paper is organized as follows. Section II provides an overview of some of the more popular methods of economizing on cash balances. Section III discusses a number of proxies that have been employed in attempts to capture the effects of cash management on the demand for demand deposits. Section IV analyzes the relative ability of these proxies to help explain cash management effects, while Section V provides a summary and conclusion.

II. POPULAR METHODS OF CASH MANAGEMENT

This section provides an overview of some popular techniques used by firms in economizing on cash balances. These techniques are an outgrowth of changing technology, as well as a response to the higher opportunity costs of holding transactions or cash balances in the 1970s. The desire to economize on the holdings of demand deposits has always been present. However, changing economic conditions alter the profitability of investing in new methods of managing transactions balances. For example, lower computer costs may make previously unprofitable procedures profitable and advances in computer technology may make new methods in cash management feasible. Firms that are attempting to minimize the costs involved in carrying out transactions, costs that involve the interest foregone on idle balances, will respond to changes in their economic environment by altering the levels and types of cash management services. The degree to which cash management technology is employed will be arrived at in a manner analogous to the choice of any other investment project.

The major changes that have spurred the growth of more sophisticated and more widespread use of cash management techniques in the 1970s have been the improvement of computer technology, the lowering of computer costs, and the rise in market interest rates. These changes have increased the benefit for reducing transactions balances while lowering the

[1] For some examples see Lieberman [16], Kimball [12], Porter, Simpson, and Mauskopf [19], and Simpson and Porter [20].

costs of doing so. As a result, there has been a proliferation of ways in which firms manage their cash balances. Examining these methods will indicate the exact ways that a firm can reduce its demand for money.

The various methods used in managing cash balances can be divided into three basic types. One type speeds up the collection of receivables, another allows firms to consolidate accounts, while a third helps control disbursements. Most of the techniques are linked with improved means of accounting, enabling the firm to more efficiently monitor its cash position. Also, many of the techniques are used together, thereby providing a full range of complementary ways for economizing on transactions balances.

Methods For Speeding Up Receivables

Lock Boxes Essentially a lock box is a centrally located collection post office box selected to minimize the mailing time taken to receive payments from customers. A firm will usually operate a number of lock boxes in various areas of the country. Several times a day, the firm's local bank will open the lock box, sort out the checks and deposit the money in the firm's account. The bank will then send the invoices and a record of the deposit to the firm. Often photocopies of checks will be sent and, in many cases, the information will be processed on magnetic tapes that can directly interface with the firm's accounting system. For a given availability schedule, the firm will have a good idea of the amount of money clearing into each account on any given day. On average, lock boxes can reduce mail float from one to four days.

Preauthorized Checks A preauthorized check is a signatureless check used for accelerating the collection of fixed payments. The customer signs an agreement with the corporation allowing the corporation or the corporation's bank to write a check at specified dates for specified amounts on his account. The corporation, through the use of a computer file, sends the bank the necessary information for performing this function. The bank then informs the firm by means of a computer tape of the deposit and the availability of the funds. This process lowers the uncertainty in income flows as well as reducing mail float.

Preauthorized Debit A preauthorized debit has the same effect as a preauthorized check. In the case of a preauthorized debit, the customer's account is automatically debited on a specified date and funds are electronically wired from the customer's bank to the firm's bank.

Consolidating Cash Balances

Concentration Accounts Concentration accounts allow firms to pool the balances collected by or deposited in local banks. Local banks automatically transfer funds, either by wire or by a depository transfer check to a central concentration bank. This process is advantageous for a number of reasons. It allows the firm to consolidate its cash balances, making it easier and less expensive to switch idle funds into market instruments. It also reduces the amount of total cash balances that need to be held since it allows for some offsetting of local disturbances to transactions balances.

Depository Transfer Checks Depository transfer checks, like preauthorized checks, are a signatureless check. They are issued by the concentration bank against one of the firm's local collection banks based on deposit information sent from the collection bank to the concentration bank, usually over a data processing network. Specifically, the concentration bank receives the deposit data and issues a check the same day for collection. It then sends the information concerning the collected funds and their availability to the firm.

Wire Transfers A wire transfer is a transfer of funds most often sent over either the Federal Reserve's wire system or a bank wire system. In this case, the local bank transfers funds from the corporation's account to the concentration bank. The wire transfer's advantage is that it allows for same day use of funds, while its disadvantage is that it is somewhat more costly than depository transfer checks. Therefore, wire transfers are predominantly used for larger transfers than are depository transfer checks. For example, at an interest rate of 6 percent, it would require a one-day transfer of $36,000 to cover the typical $6.00 wire transfer cost. Naturally, as interest rates rise, the minimum profitable level of the transfer would fall.

There have been a number of innovations in the use of wire transfers. Most notable is the ability of a firm's cash manager to initiate a wire transfer from a computer terminal that either interfaces with the bank's computer controlled wire system or with the data base of a third party that is used by the bank. Often this service is linked with other cash management services, such as programs that forecast a company's cash flows, and produces wire transfers that are less costly and that provide hard copy verification of funds transferred.

Methods For Controlling Disbursements

Controlled Disbursement A firm may also exert more control over its cash balances by being able to better predict disbursements on a day-to-day basis. The firm can achieve this by using a bank that receives only one shipment of checks from the Federal Reserve each morning. The bank informs the firm of the value of checks drawn on its account and the firm then knows, usually before noon, how much of its balances are unnecessary.

Zero Balance Accounts This procedure is a special case of controlled disbursement that allows the firm to maintain zero transactions balances at a number of banks from which it writes checks. When the value of checks presented against the firm's account is tabulated, the appropriate amount of funds are wired from a central account. This allows the firm to greatly economize on the level of balances held at each disbursing bank, and provides centralized data on transactions.

Summary of Cash Management Services

It is clear from the description of the methods used in managing cash balances that many of these procedures will be simultaneously employed. For instance, a firm is likely to have zero balance arrangements with local banks that also provide lock box services. Also, the firm will use both depository transfer checks and wire transfers to facilitate the quick movements of funds. Crucial to the desire to economize on transactions balances is the ability to invest these funds in short-term market instruments at relatively low costs. Otherwise there would be no reason to incur the costs involved in reducing the average level of balances.

III. METHODS OF CAPTURING CASH MANAGEMENT EFFECTS IN MONEY DEMAND EQUATIONS

The preceding discussion described how various cash management techniques are able to reduce the demand for money. Therefore, failure to incorporate cash management effects in a demand deposit regression will result in a misspecified equation. Since the degree to which cash management procedures are used is a choice Variable of the firm and is related to the cost and benefits of investing in these procedures, this misspecification will have serious consequences for any estimated equation.

For example, in inventory models of money demand, either stochastic or nonstochastic, some of the investment in cash management services can be viewed as lowering transactions costs. For instance, a firm having a cash management system that allows it to perform investments in repurchase agreements from a computer terminal has invested in a procedure that greatly reduces transactions costs. In stochastic inventory models, many of the cash management services can be viewed as ways for reducing the variance of cash flows (see Porter and Mauskopf [18]). Therefore, some key elements of the demand for money, namely transactions costs and the variance of cash flows, are not exogenous variables from the standpoint of individual deposit holders, but are variables that can be influenced by the level of cash management sophistication.

This line of reasoning implies that firms are simultaneously choosing the level of investment in cash management services and their average deposit balances. Since a number of the parameters that influence the demand for demand deposits are functions of the level of cash management, the level of cash management should appear in the demand for money equation.[2]

Failure to include a measure of the effects of cash management in the demand for demand deposits will therefore result in a seriously misspecified regression. As a result, coefficient estimates will be biased and predictions from the regression will be inaccurate in periods when cash management practices are changing. Further the regression will appear to be unstable (leading one to believe that the demand for money is unstable), when in fact the instability is totally due to an omission on the part of the econometrician.

In this section a number of variables for capturing cash management technology are examined. These candidates are generally related to the actual methods used in cash management and to the underlying costs and benefits associated with investing in techniques that help economize on transactions balances.

A Time Trend (T) The first and simplest way to represent cash management innovations in a money. demand equation is by use of a time trend. This was initially employed by Lieberman [16]. The motivation behind this variable is that the adoption of new technology will be fairly uniform and proceed at an exponential rate. This procedure explicitly, treats the

[2] The result of the optimization process by which firms choose the level of cash management services and average demand deposit balance is a two-equation system that is recursive. For more detail see Dotsey [9].

process of changes in cash management practices as exogenous. It therefore omits from consideration any economic forces, such as changes in costs or returns, that would be expected to alter the rate at which cash management techniques are implemented. However, it serves as a useful benchmark for comparing the effects that more sophisticated methods of incorporating the consequences of cash management have had on the demand for money. One practical problem in using a time trend is choosing the starting date for the trend.

A Ratchet Variable (RATCHET) In general, a firm would adopt new methods of cash management if the expected benefits outweigh the costs. That is, investing in a new cash management system would involve the same considerations as investing in any other project. The motivation behind the use of a ratchet variable constructed from interest rates is to capture some of the economic conditions that would lead to firms' implementing more sophisticated cash management techniques.

Since much of the costs of employing innovations in cash management are start-up costs (e.g., putting in the necessary computer hardware and software), it follows that once a new cash management system is in place it will remain in operation until it is replaced by more advanced technology. For the investment to be profitable, the interest rate savings incurred from lower average money balances must be substantial and expected to last for some time. One would therefore expect that major innovations would occur when long-term interest rates are high relative to their past history, and that these innovations would continue to affect the demand for money once they are initially adopted. Long-term rather than short-term interest rates are the relevant variable, because they indicate that a movement in interest rates is expected to persist. One is also interested in the movement of long-term rates with respect to its past, since upward movements will spur new investment in cash management due to the increased return obtained from economizing on transactions balances.

The preceding discussion suggests that a nondecreasing variable based on long-term interest rates, which increases (or ratchets up) when rates are relatively high, would be helpful in explaining changes in cash management practices and hence changes in the demand for money. The specific formulation investigated in this study is the one derived by Simpson and Porter [20]. Specifically,

$$(1)\quad \text{RATCHET}_t = \text{RATCHET}_{t-1} + \left(r_t - \frac{1}{n}\sum_{i=t-(n-1)}^{t} r_i\right)^+$$

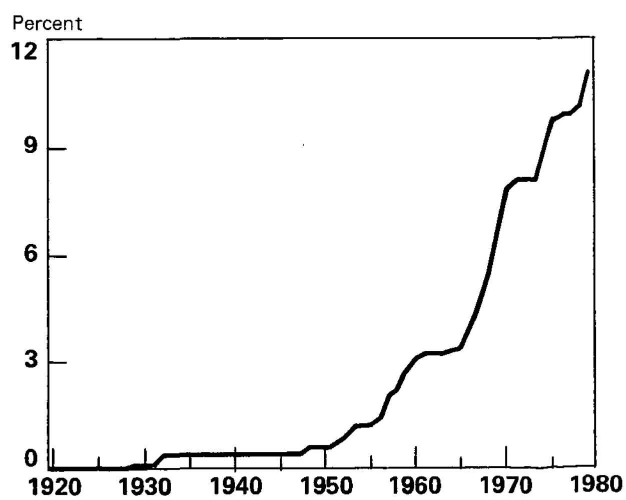
where r_t is the long-term bond rate (Moody Aaa) and the + sign indicates that only positive values of the expression in parentheses are used. The variable is somewhat sensitive to the value of n chosen, so variables using n = 3, 4, 5, 6 were constructed. All gave similar results and only the values for n=4 are reported. A graph of the RATCHET is depicted in Figure 1.

One can see that the formulation given by equation 1 captures the ideas behind the ratchet variable. For example let n=4. Then the current value of RATCHET is equal to last period's value plus an additional term. The additional term reflects the value of today's long-term interest rate relative to its average over the latest four periods. If today's rate is higher than this average, then RATCHET increases indicating an increase in investment in cash management services. If today's rate is lower than the average, then RATCHET remains the same as it was last period. This implies no new investment in cash management technology, and that today's level of technology is the same as last period's level.

Although the ratchet variable possesses some useful features, it does have certain limitations. It only considers the potential benefits of new technology but

Figure 1

THE RATCHET VARIABLE
1920 - 1979



not its cost. Furthermore, the benefits are only potential, since one doesn't know how much economization occurs as a result of new technology. Also, the variable does not consider depreciation.

The Price of Office Computing and Accounting Equipment (P) The discussion in Section II makes it clear that much of the use of cash management techniques involve computers and accounting equipment. Therefore the costs of this equipment will be closely related to the costs of cash management. In constructing a variable that captures these costs, it is important that the variable take into account adjustment in quality. For example, a new computer model may cost slightly more than the one it replaces, but it may be able to perform many more operations in much less time. In terms of what the computer actually does, the newer model is much less expensive than the older model even though its price may be somewhat higher. A true index of the computer's cost will take account of the change in quality. Such an index is referred to as a hedonic price index.

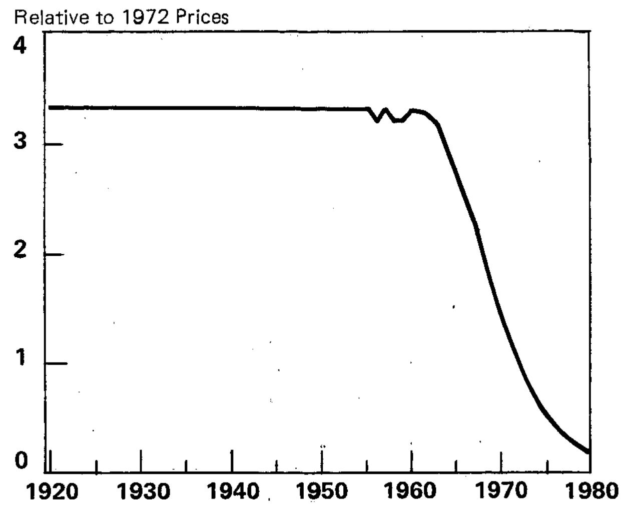
As the cost of technology falls, more firms will adopt the technology thus reducing the demand for demand deposits. Therefore the price of office computing and accounting equipment could help to explain shifts in the demand for money induced by cash management. However, the price variable does have certain limitations. It does not account for technology already in place, nor does it reflect depreciation. Further it does not consider changes in the benefits that occur from the implementation of new cash management services. Therefore, it would be natural to use this variable in conjunction with a ratchet variable.

For the years 1956-1979 data on the hedonic price of office computing and accounting equipment was obtained from McKee [17]. Although his procedures are somewhat rough, they are the best available. For the time period 1920-1955, it is assumed that the real cost of technology remained constant at its 1956 level. A graph of this variable is given in Figure 2.

The Number of Electronic Fund Transfers (EFT) The motivation for this variable is largely attributed to Kimball [12]. The use of many of the cash management techniques discussed in Section II involves the rapid movement of money so that it may be invested in short-term market instruments. In many cases idle transactions balances may only be

Figure 2

REAL PRICE OF OFFICE COMPUTING EQUIPMENT 1920 - 1979



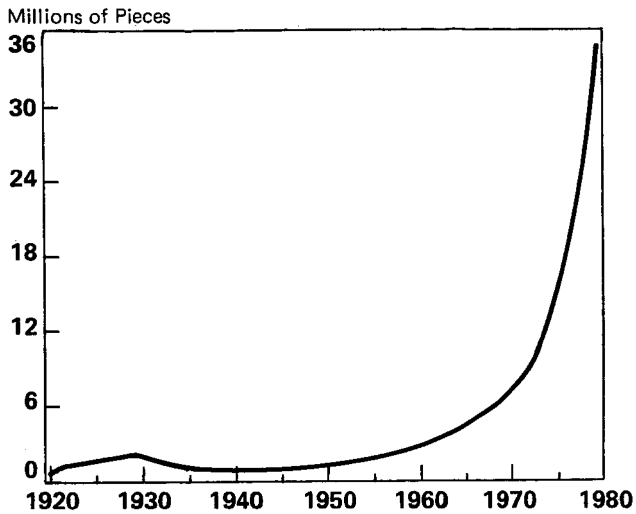
invested overnight: To implement this type of activity often requires the use of immediately available funds. Therefore, much of the transfer of money is done over either the Federal Reserve's wire system or over Bank wire.

For instance, a firm may use a number of lock boxes, have a zero balance account with a disbursement bank, and a consolidation account with another bank. On any given day, funds would be wired from the lock box collecting banks to the bank maintaining the consolidation account and from the consolidation account to the zero balance account. Funds may also be wired to another bank for the purpose of executing a repurchase agreement if it can not be done with the consolidating bank. In general there is good reason to believe that the number of electronic funds transfers is largely determined by the degree of cash management practices. Because of this relationship, the number of electronic fund transfers is a logical variable for helping explain the shift in the demand for money (for more detail see Dotsey [9]).

The value for the number of electronic fund transfers used is restricted to funds transfers made over the Federal Reserve's wire transfer system and is depicted in Figure 3. Since there are other wire transfer systems this value is not totally accurate. However, it is believed that the time series properties of the measure is not much different than what would be observed if data on total wire transfers could be obtained.

Figure 3

NUMBER OF ELECTRONIC FUNDS TRANSFERS 1920 - 1979



IV. EMPIRICAL RESULTS

In order to appreciate the severity of the effect of cash management on the demand for demand deposits, a regression explaining demand deposit behavior is run over the period 1920-1965, a period in which cash management innovations are believed to be unimportant. (An examination of Figures l-3 indicates that the various proxies are fairly constant over this time span.) This regression is then rerun over the extended sample period (1920-1979), and the results are compared. This is depicted in Table I.

The regression equation examined is based on an inventory model of the demand for money used in Dotsey [8], [9]. Specifically,

$$(2) \quad LND_t = a_0 + a_1LNC_t + a_2LNRD_t + a_3LNRS_t + a_4LNRCP_t + a_5LNW_t + a_6LNPCR_t + e_t$$

where the letters LN refer to the natural log of a particular variable (i.e., LNX equals the log of X). The letter D represents the level of real demand deposits, C represents the level of real consumption expenditures, RD is the own rate of return on demand deposits calculated using Klein's [14] methodology, RS is a weighted average of the interest rate on passbook savings accounts and money market mutual fund shares, RCP is the commercial paper rate, W is the real wage rate, and PCR is the ratio of credit to consumption where the level of credit in-eludes installment retail credit, noninstallment retail credit, credit outstanding on bank credit cards, credit owed to gasoline companies, and check credit. The letter e refers to the disturbance term.

Consumption expenditures are used to represent transactions income, while RD captures the desirability of holding a demand deposit.[3] RS and RCP are used to capture the return earned on alternative assets held by different classes of economic agents. The real wage rate is a proxy for the value of time and is therefore related to transactions costs, while PCR attempts to net out the percent of transactions income spent via credit.[4]

[3] The use of Klein's rate involves some empirical issues that make interpreting its effect difficult (see Carlson and Frew [6]). However, to the extent that one believes that corporations earn a competitive rate on their deposits, omission of RD leads to specification bias. (For more detail see Dotsey [9], especially footnotes 6 and 7.)

[4] Since the emphasis of this article is to illustrate the effects of cash management practices on the demand for demand deposits, equation 2 is not discussed in detail. For a full discussion see Dotsey [S].

Table I

REGRESSION ANAYLSIS FOR DEMAND DEPOSITS

Independent Variables	1920-1965	1920-1979
CONSTANT	−1.34 (−2.73)**	1.33 (1.71)
LNC	.79 (10.49)**	.31 (2.99)**
LNRD	.075 (2.86)**	.12 (2.51)*
LNRS	−.24 (−13.27)**	−.25 (−8.04)**
LNRCP	−.11 (−3.46)**	−.11 (−1.82)
LNW	.36 (3.73)**	.93 (6.17)**
LNPCR	−.28 (−9.66)**	−.30 (−5.69)**
$\bar{R}^2$	.9959	.9840
D.W.	1.79	.73
S.E.E.	.0306	.0589

The numbers in parentheses are t-statistics.

* indicates significance at the 5 percent level.

** indicates significance at the 1 percent level.

The results of the regression run over the period 1920-1965, yield coefficients that are consistent with an inventory model of money demand. The error term does not exhibit any serial correlation and one can not reject the stability of the regression. The tests for stability used were a standard F-test, a test using the cusum of squares statistic developed by Brown, Durbin and Evans [5], and a test for stability using the varying parameters model of Cooley and Prescott [7].[5] The regression coefficients also converge fairly quickly to their full sample values, when the sample period is continually extended from 1928 to 1965. This combination of evidence strongly implies that the specification in equation 2 is a well-behaved representation of the demand for demand deposits over the period 1920-1965.

When the sample period is extended through 1979 this is no longer the case, and equation 2 is no longer an accurate model of the demand for demand deposits. Most importantly, the error structure of the regression changes. This is evident from the low value of the Durbin-Watson statistic, implying serial correlation in the errors. This means that the standard errors of the regression coefficients are biased making it impossible to state whether the coefficients in column 2 of Table I are significant. After correcting for serial correlation the wage variable becomes insignificant. Also, the presence of serial correlation is often indicative of a missing variable or variables. A good candidate, or candidates, for this missing variable would be variables that take into consideration the effects of cash management.

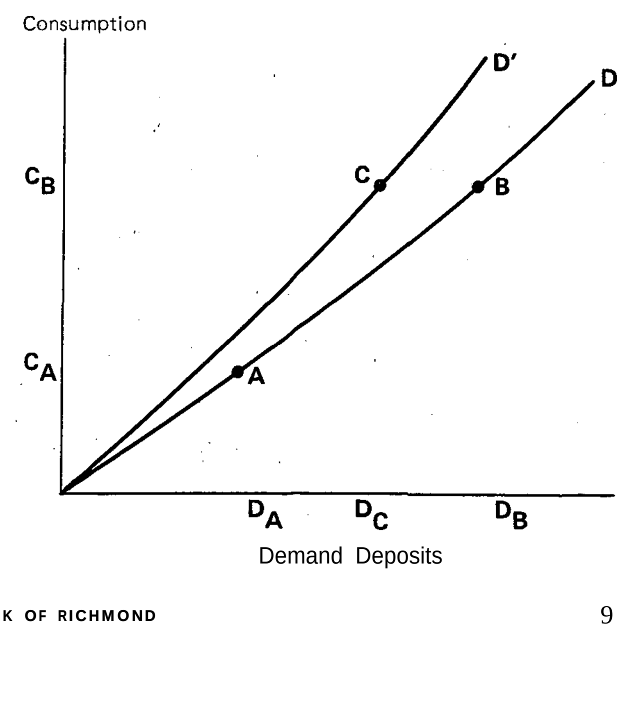
The reduction of the coefficients on consumption and the real wage is consistent with the omission of variables that represent a general decrease in transactions costs, or a lowering of the variance of cash flows associated with a given level of business. Consider the graph in Figure 4. The locus of points labelled D, represents a relationship between real consumption C, and real demand deposit balances with all other variables (i.e., interest rates, PCR, and W) held fixed. The locus D' represents the same relationship depicted for a more sophisticated use of cash management. As shown, less real balances are held for any level of consumption, interest rates, real wages, and the intensity of credit purchases. Now as consumption rises (as it did over the period 1965-1979), instead of moving from point A to point B along D, there is a movement from point A to point C. Thus, excluding the incorporation of cash management implies that demand deposits will appear to be less sensitive to changes in consumption.

A similar argument would apply to the real wage rate. With respect to interest rates, however, the effect of omitting cash management could be ambiguous. This is because interest rates both rose and fell over the period 1966-1979. For example, consider the commercial paper rate. Failure to explicitly include the effects of cash management would imply a greater sensitivity of demand deposits to increases in the commercial paper rates when this rate was rising, while just the opposite would occur when the rate was falling.

Adding Proxies For Cash Management If a movement toward more sophisticated cash management techniques is the sole or primary reason for a shift in, the demand for demand deposits, then incorporating variables that accurately account for this movement should have a pronounced effect on estimated demand for demand deposit equations. Specifically, the errors should be white noise and the coefficients should return to the values found in the regression run over the 1920-1965 period. Further, stability of the regression over the extended 1920-1979 sample should

Figure 4

AN EXAMPLE OF HOW CASH MANAGEMENT AFFECTS THE MEASURED RESPONSE OF DEMAND DEPOSITS TO CONSUMPTION



[5] Descriptions of these test statistics are quite technical and are therefore omitted. The interested reader can find a more detailed discussion in Dotsey [8] or can read the referenced articles. An excellent summary can also be found in Boughton [4].

not be rejected and there should be a significant increase in the predictive power of the equation. The various proxies described in Section III will be examined with respect to all of the properties just listed.

First, examine the regressions in Table II. All of the proxies return the coefficients approximately to the values estimated in column 1, Table I. However, only in the case of EFT, can one reject the presence of first order serial correlation of the residuals[6] Since RATCHET proxies for the potential benefits

[6] The discussion of EFT in Section III indicates that its effects should be examined within the context of a simultaneous system. Since the analysis indicates that this system is recursive, simultaneity bias will occur only if the errors in the two equations are correlated. A two-stage least squares estimation technique gave similar results to those obtained using OLS, implying that simultaneity bias is not a problem. For a more detailed discussion see Dotsey [9].

Table II

REGRESSION ANALYSIS FOR DEMAND DEPOSITS (1920-1979) WHEN VARIOUS PROXIES FOR CASH MANAGEMENT ARE INCLUDED

	I	II	III	IV
CONSTANT	−1.74 (−2.56)*	−.91 (−1.11)	−1.31 (−2.22)**	−1.08 (−2.49)*
LNCON	.93 (9.10)**	.85 (7.05)**	.92 (10.17)**	.78 (12.21)**
LNRSAV	−.22 (−7.31)**	−.19 (−4.24)**	−.23 (−9.10)**	−.21 (−14.84)**
LNRDD	.078 (2.37)*	.088 (2.36)*	.077 (2.52)*	.079 (3.18)**
LNRCP	−.14 (−3.38)**	−.15 (−3.21)**	−.13 (−3.54)**	−.10 (−3.48)**
LNPCR	−.21 (−3.86)**	−.21 (−2.93)**	−.22 (−4.91)**	−.26 (−9.70)**
LNWAGE	.25 (1.86)	.41 (2.64)*	.24 (2.03)*	.40 (4.67)**
T	−.026 (−6.51)			
RATCHET		−.035 (−3.21)**		
LNP			.14 (8.39)**	
EFT				−.013 (−12.71)**
RHO	.56 (5.67)*	.71 (7.15)**	.44 (4.15)**	
$\bar{R}^2$	.9947	.9932	.9928	.9960
D.W.	1.98	1.88	1.99	1.74
S.E.E.	.0325	.03695	.0353	.0293

The numbers in parentheses are t-statistics.

* indicates significance at the 5 percent level.

**indicates significance at the 1 percent level.

RHO is the coefficient for first order autocorrelation.

of cash management, while LNP attempts to capture costs in adopting new technology, it would be natural to use both variables simultaneously. This was attempted, but only LNP retained its significance, perhaps because these variables only reflect general trends and are therefore only picking up an overall tendency toward increasing cash management sophistication. Finally a regression including LNP, RATCHET, and EFT was run with only EFT retaining its significance.

Second, all of the proxies decrease the instability of the regressions in the sense that the cusum of square statistic is lowered. However, only by using EFT could a lack of stability be rejected using the Brown-Durbin-Evans test. Also, one could not reject stability of the regression with EFT under the procedure developed by Cooley and Prescott. However, when the sample was divided in 1949, stability was rejected using a standard F-test. Given that EFT only includes the number of wire transfers over the Federal Reserve wire system, and is therefore an imperfect measure of total wire transfers, the net result of the stability tests is encouraging.

Third, an examination of one step ahead out of sample forecast errors is depicted in Table III. Again, all the proxies generally improve the forecasts, with EFT performing the best. Using EFT resulted in a reduction of the average absolute error of the forecast by 52 percent and a reduction in the root mean square error by 34 percent.

Table III

THE LOGARITHM OF ONE PERIOD AHEAD FORECAST ERRORS

	None	T	RATCHET	LNP	EFT
1966	-.19	.04	-.15	.00	.05
1967	-.24	.04	-.13	.00	.11
1968	-.16	.16	.08	.18	.18
1969	-.15	.14	.16	.17	.23
1970	-.18	.15	.19	.17	.21
1971	-.14	.10	.06	.12	.17
1972	-.25	.05	-.12	-.03	.07
1973	-.17	.07	-.04	.10	.32
1974	-.26	-.02	-.08	.03	.20
1975	-.71	-.04	-.47	-.38	-.28
1976	-.94	-.62	-.72	-.53	-.26
1977	-.63	-.48	-.70	-.36	.02
1978	-.73	-.48	-.80	-.33	.05
1979	-.69	-.30	-.60	-.40	.41
Avg. Error	-.38	-.12	-.24	-.07	.11
RMSE	.29	.28	.35	.24	.19

V. SUMMARY AND CONCLUSION

This article, in a somewhat different empirical setting than that underlying most conventional studies of money demand, presents confirmation of the recent shift in the demand for money. The hypothesis that a shift has taken place in the function is supported by stability tests and the poor predictive performance of the model in the mid- and late 1970s. The empirical evidence combined with documentation on the increased use of sophisticated cash management practices by firms makes changes in cash management techniques a probable explanation for the shift in the historical demand deposit relationship.

An attempt is made to capture this process by the use of variables which are believed to be related to innovations in the management of transactions balances. This process seems to be captured quite well by the variable EFT. The other proxies perform reasonably well in reducing the forecasting errors of the demand deposit relationship, but were not in general able to capture the entire movement in the function. The results are by and large encouraging enough to make future research into transactions technology and its relation to money management a potentially rewarding avenue in helping to explain the current behavior of the demand for money.

References

1. Barger, H. *Outlay and Income in the United States 1921-1938*. New York: National Bureau of Economic Research, 1942.
2. Baumol, W. J. "The Transactions Demand for Cash-An Inventory-Theoretic Approach." *Quarterly Journal of Economics* 66 (1952), 645-56.
3. Beehler, Paul J. *Contemporary Cash Management.* New York: John Wiley and Sons, Inc., 1978.
4. Boughton, J. M. "Recent Instability of the Demand for Money : An International Perspective." *Southern Economic Journal* 48 (1981), 579-97.
5. Brown, R. L.; J. Durbin; and J. M. Evans. "Techniques for Testing the Constancy of Regression Relationships Over Time." *Journal of the Royal Statistical Society,* Series B 37 (1975), 149-63.
6. Carlson, J. A., and J. R. Frew. "Money Demand Responsiveness to the Rate of Return on Money: A Methodological Critique." *Journal of Political Economy* 88 (1980), 598-607.
7. Cooley, T. F., and E. C. Prescott. "Systematic (nonrandom) Variation Models Varying Parameter Regression : A Theory and Some Applications." *Annals of Economic and Social Measurement* 2 (1973), 463-73.
8. Dotsey, Michael. "The Effects of Cash Management Practices on the Demand for Demand Deposits." Working Paper 83-2. Federal Reserve Bank of Richmond, 1983.
9. ———— "The Effects of Cash Management Practices on the Demand for Demand Deposits." 1984. (Processed.)
10. ————; S. Englander; and J. Partlan. "Money Market Mutual Funds and Monetary Control." *Quarterly Review,* Federal Reserve Bank of New York 6 (1982), 9-17.
11. Karni. E. "The Value of Time and the Demand for Money." *Journal of Money, Credit and Banking* 6 (1974), 45-64.
12. Kimball, R. C. "Wire Transfer and the Demand for Money." *New England Economic Review,* Federal Reserve Bank of Boston (March/April 1980), pp. 5-22.
13. Klein, B. "The Payment of Interest on Commercial Bank Deposits and the Price of Money: A Study of the Demand for Money." Unpublished Ph.D. dissertation. University of Chicago, 1970.
14. ————. "Competitive Interest Payments on Bank Deposits and the Long-Run Demand for Money." *American Economic Review* 64 (1971), 931-49.
15. Kuznets, S. *National Income and its Composition.* New York: National Bureau of Economic Research, 1941.
16. Lieberman, D. "The Transactions Demand for Money and Technological Change." *Review of Economics and Statistics* 59 (1977), 307-17.
17. McKee, M. J. "Computer Prices in the National Accounts: Are Our Economic Problems a Computer Error." Council of Economic Advisors, 1981. (Processed.)
18. Porter, R. D., and E. Mauskopf. "Cash Management and the Recent Shift in the Demand for Demand Deposits." Board of Governors of the Federal Reserve System, 1978. (Processed.)
19. ————; T. D. Simpson; and E. Mauskopf. "Financial Innovation and the Monetary Aggregates." *Brookings Papers on Economic Activity* (1: 1979)) 213-29.
20. Simpson, T. D., and R. D. Porter. "Some Issues Involving the Definition of and Interpretation of the Monetary Aggregates" in *Controlling the Monetary Aggregates III.* Federal Reserve Bank of Boston Conference Series no. 23 (1980), pp. 161-234.
21. Tobin. J. "The Interest Elasticity of Transactions Demand for Cash." *Review of Economics and Statistics* 38 (1953), 241-47.

ALGEBRAIC QUANTITY EQUATIONS BEFORE FISHER AND PIGOU

Thomas M. Humphrey

Perhaps the most basic tool of monetary analysis is the quantity equation of exchange MV = PQ, where M is the money stock, V its average turnover velocity, P the price level, and Q the quantity of goods exchanged against money. This equation has at least three alternative interpretations. Stated as the identity MV = PQ (where velocity is defined as V = PQ/M so as to render the equation a tautology), it reminds us that expenditures must equal receipts, that the sum total of monetary payments (MV) must just add up to the aggregate value of goods sold (PQ). Written as M/P = Q/V or M = (Q/V) P, where velocity is now defined independently of the other variables such that the equation is non-tautological, it states that the price level P must adjust to equate the real or price-deflated value of the given nominal money stock M with the given real demand for it, this real demand being the fraction l/V of real transactions Q that the public wishes to hold in the form of real cash balances. In other words, it states that the price level P is determined by the nominal money supply M and real money demand (Q/V), varying directly with the former and inversely with the latter. Alternatively formulated as P = MV/Q, it says that prices are determined by total expenditure (MV) relative to output Q, that is by aggregate demand and supply. Most often the equation is used to expound the celebrated quantity theory of money, which says that, given real money demand, changes in the money supply cause equiproportional changes in prices.

The equation's applications are of course well-known. Not so well-known, however, is its origin and early history. For the most part, textbooks typically treat it as a product of 20th century monetary thought, usually identifying it with Irving Fisher and A. C. Pigou, its most influential 20th century formulators. Fisher, in his *Purchasing Power of Money* (1911), wrote the equation in its transaction velocity form :

$$(1) \quad MV + M'V' = \Sigma pQ = PT$$

where M is the stock of currency, V its velocity, M' is the volume of checking deposits, V' their velocity, EpQ is the sum of the quantities Q of goods and services sold valued at their market prices p, P is the weighted average of these prices or the general price level, and T is the aggregate of real transactions or the sum of all the Qs. Similarly, Pigou, in his 1917 article "The Value of Money," wrote the equation in its alternative Cambridge cash balance form :

$$(2) \quad \frac{1}{P} = \frac{kR}{M}$$

where l/P, the inverse of the price level, is the value (purchasing power) of the monetary unit ; R denotes real resources; k, the reciprocal of velocity, is the the proportion of those resources that people wish to hold in the form of money; and M is the money stock.[1] Neither Fisher nor Pigou, however, were the first to write such equations. On the contrary, the cash balance equation preceded Pigou by more than thirty years, having been presented by Léon Walras in 1886. Likewise, the transactions velocity equation predated Fisher by more than 100 years, having been fully enunciated in 1804.

In fact, quantity equations are even older than the preceding discussion implies. For rudimentary prototypal versions began to appear as early as the late 17th and 18th centuries, followed by increasingly sophisticated versions in the 19th and early 20th centuries-versions that were often more elaborate and complete than those associated with Fisher and Pigou. These earlier contributions have been largely overlooked. In an effort to correct this oversight and to set the record straight, this article traces the pre-Fisherian, pre-Pigovian development of the quantity

[1] Pigou's equation is virtually the same as the celebrated cash balance expression presented by John Maynard Keynes in his A Tract on Monetary Reform (1923). Keynes's equation is

$$n = pk$$

where n is the nominal money stock (Pigou's M), p is the price level (Pigou's P), and k is that part of real output the command over which people wish to hold in the form of cash balances (Pigou's kR).

equation in the British, German, Italian, French, and American monetary literature. It covers only writers who presented the equation in explicit algebraic form. Unmentioned are the host of analysts (including, among others, Locke, Hume, Smith, Thornton, Ricardo, Mill, and Marshall) who employed the equation in merely arithmetic or verbal form. It shows that earlier economists not only formulated the equation and specified its components; they also interpreted it as an equilibrium condition between the price level (or value of money) and money supply and demand. That is, they viewed it as an algebraic model of equilibrium price level determination.

British Writers

The first rudiments of the quantity equation have their origin in the 17th and 18th century British monetary literature. To John Briscoe [3] in 1694 and Henry Lloyd [14] in 1771 go the credit for presenting the first such equation and also for being the first to interpret it as a model of equilibrium price level determination.[2] Their equation, however, lacked a velocity term, being written in the form

$$(3) \quad M = PQ \text{ or } P = \frac{M}{Q}$$

where P denotes the price level, M the money stock, and Q the quantity of goods exchanged for money. They omitted the velocity term (or implicitly assigned it a magnitude of unity) because they viewed prices as being determined in a single transaction involving the one-time exchange of the entire stock of money for the entire stock of goods. They did not understand that price level determination is a continuous process and that the stock of money turns over several times per period in purchasing goods. Nor did they realize that the volume of goods exchanged against money is a flow and not a stock, and that the stock of money must therefore be multiplied by its average velocity of circulation to make it dimensionally comparable with the flow of goods. Despite this shortcoming, they were able to draw correct conclusions from their equation, namely that prices vary in direct proportion with money and in inverse proportion with output. Lloyd even gave an algebraic proof of this latter conclusion, pointing out that if the quantity of goods increases by a scale factor y while the money stock is held constant, then prices will fall by the inversely proportional scale factor 1/y according to the equation

$$(4) \quad P(\frac{1}{y}) = \frac{M}{yQ}.$$

Lloyd also viewed his equation as embodying an aggregate demand/aggregate supply theory of price determination, a theory in which M serves as the demand variable and Q as the supply variable. That is, he saw M as affecting P through demand just as Q influences P through supply. Although his work had no apparent impact on his fellow countrymen, it did influence his Italian contemporary, the mathematician P. Frisi. The latter, in his review of Lloyd's equation, proposed multiplying the money/goods (M/Q) ratio by the ratio of the number of buyers to the number of sellers (a proxy for real demand and supply) in a crude effort to account for all nonmonetary market forces affecting general prices.[3] He failed to see that Lloyd's commodity (Q) variable already comprehends these forces so that additional variables are superfluous.

After Lloyd, the next British writer to present a quantity equation was Samuel Turner, who, in his *A Letter Addressed to the Right Hon. Robert Peel with Reference to the Expediency of the Resumption of Cash Payments Fixed by Law* (1819), wrote the expression

$$(5) \quad a = bc$$

in which a is the value of commodities exchanged (that is, PQ) over a period of time such as a year, b is the quantity of metallic money in circulation (or M), and c is the circulating power of money or the number of times it changes hands during the year (or V). Turner's formula does not divide the, nominal transactions variable into its price and quantity components.[4] But it does incorporate a velocity term and therefore constitutes an improvement over the primitive equations of Briscoe and Lloyd. Turner also expanded his equation to include a term for paper money, resulting in the augmented expression

$$(6) \quad a = (b+p)c$$

where p is paper money and b is metallic coin. Here is the first quantity equation to contain separate variables denoting different components of the money stock, each multiplied by the same velocity coefficient c.

[2] On Briscoe, see Schumpeter [25, pp. 314-5]. On Lloyd, see Schumpeter [25, p. 315] and Theocharis [26, pp. 30-31].

[3] On Frisi's modification of Lloyd's equation, see Theocharis [26, p. 31] and Marget [18, pp. 154, 270-1, 277].

[4] On Turner's formula, see Theocharis [26, pp. 120-1].

Twenty-one years after Turner, Sir John Lubbock presented in his On Currency (1840) the first quantity equation to incorporate separate velocity coefficients for the different items comprising the media of exchange.[5] Lubbock's equation, which also includes a term for transactions (such as gifts) that do not involve market prices, is

$$(7) \quad \Sigma a\alpha + E = lD + mB + nC$$

where $\Sigma a\alpha$ is the sum of market transactions a at prices α during a given period (Fisher's ΣpQ), E is the sum of transactions or transfers not involving prices (such as gifts, tax payments, the repayment of principal on debts, etc.), D is the amount of checking deposits (Fisher's M'), B is the total amount of bills of exchange (Fisher's M"), C the total amount of cash, or money narrowly defined (Fisher's M), and 1, m, and n are velocity coefficients corresponding to the V', V" and V terms of Fisher's equation. Note that Lubbock distinguishes between the money and near-money (or money-substitute) components of the media of exchange-the near-money component being defined as deposits and bills of exchange. He further decomposes the money or cash component C into its bank note, coin, and cash-reserve constituents according to the equation

$$(8) \quad C = f + g + \frac{D}{k}$$

where f denotes bank notes in circulation, g denotes coin in circulation, and D/k denotes the coin and bullion reserves backing banks' note and deposit liabilities, these reserves being expressed as the ratio of deposits D to the deposit expansion multiplier k.[6] Substituting this last formula into the one immediately preceding it yields the augmented quantity equation

$$(9) \quad \Sigma a\alpha + E = lD + mB + n(f+g+\frac{D}{k}).$$

Here is the first appearance of the deposit expansion multiplier and separate velocity coefficients in a quantity equation.

From his equation, Lubbock concluded as follows: given output and velocities, prices move equiproportionally with changes in the means of payment-but only if the volume of unilateral transfers E is zero or also moves equiproportionally with the means of payment. In these cases the quantity theory holds. It may not hold, however, if E responds disproportionately to monetary changes since such responses require offsetting disproportionate movements in prices. In particular, with E invariant to monetary shocks (a situation Lubbock thought most likely) prices would tend to move in greater proportion than money.

German Writers

In the 48-year interval separating the contributions of Lloyd and Turner, German economists made significant advances in the formulation and analysis of algebraic quantity equations. Claus Kröncke, in his *Das Steuerwesen* nach *seiner Natur und* seinen *Wirkungen untersucht* (1804), was the first writer to introduce a velocity term into the equation, doing so fifteen years before Turner.[7] Kröncke's equation is

$$(10) \quad r = \frac{\phi}{m}$$

where r is the money stock needed in a country, ø is the nominal value of all goods sold during a certain period of time, and m is the number of times on the average that money turns over in purchasing goods during the period. This is the same as the conventional quantity equation M = PQ/V, where r = M, ø = PQ, and m = V. Although Kröncke did not divide his nominal transactions variable into its price and quantity components, he did state that if output and velocity are given, prices must vary directly with the money stock. That is, he used his equation to help illustrate the quantity theory of money. He also recognized that monetary contraction could occur without depressing nominal activity only if there were offsetting rises in velocity. Because he wished to maintain the level of activity while simultaneously minimizing the quantity of gold in circulation (so that the excess could be exported for consumption goods), he advocated policies to increase velocity.

In 1811, Kröncke's compatriot Joseph Lang made two key contributions to quantity-equation analysis.[8] He was the first to include separate terms for the four crucial variables M, V, P, and Q, thereby improving upon Kröncke's three-variable formulation. He was also the first mathematical economist to employ finite difference notation in deriving the quantity theory prediction that prices vary equiproportionally with money. He writes the equation in his *Grundlinien der politischen Arithmetik* (1811) as

[5] On Lubbock, see Marget [17, pp. 11-12].

[6] See Marget [17, pp. 152-3].

[7] On Kröncke, see Theocharis [26, pp. 102-3].

[8] On Lang, see Theocharis [26, pp. 109-10].

$$(11) \quad yZ = Px$$

where according to his symbols, y is velocity, Z is money, P is real output, and x is the price level. His equation can be translated into the conventional formula MV = PQ. Having written the equation, he then solves for the price level or

$$(12) \quad P = \frac{MV}{Q}$$

from which he concludes that prices P vary in direct proportion to M and V and in inverse proportion to Q.

Then, for the first time in the history of mathematical economics, he employs finite difference, or delta (A), notation to demonstrate rigorously that, with V and Q given, prices vary in exact proportion to money.[9] Starting with his equation

$$(13) \quad MV = PQ$$

he supposes money to increase by a small amount AM, where the delta symbol denotes an incremental change in the attached variable. Assuming V and Q fixed, he notes that only prices can respond. Denoting this price response by ΔP and inserting it and the corresponding monetary increment ΔM into his quantity equation yields

$$(14) \quad (M + \Delta M)V = (P + \Delta P)Q.$$

Expanding this equation, subtracting the preceding equation MV = PQ from the result, and then solving for the increment in prices gives him the expression

$$(15) \quad \Delta P = (\frac{V}{Q})\, \Delta M$$

which states that the incremental variation in prices is exactly proportional to that of money, with the ratio V/Q (the inverse of the demand for real balances) being the factor of proportion. Here is the first rigorous algebraic statement of the quantity theory of money.

Other 19th century German writers who employed quantity equations include K. Rau and W. Roscher.[10] Little need be said about them, however, as they added virtually nothing to the earlier formulations of Kröncke and Lang. Rau, in his *Grundsätze der Volkswirtschaftslehre* (1841), stated the formula MV = PQ, prompting Friedrich Lutz in 1936 to suggest that it thereafter be called the "Rau-Fisher equation."[11] Similarly, Roscher, in his *Grundlagen der Nationalökonomie* (1854), presented an equation similar to Kröncke's, namely

$$(16) \quad u = ms$$

where u is the monetary sum of transactions (or PQ), m is the quantity of money (or M), and s is the velocity of circulation (or V). Roscher's three-variable formula, of course, was already obsolete at the time he published it, having been superseded by Lang's four-variable formulation forty-three years before. Nevertheless, the quantity equation's appearance in the popular textbooks of Rau and Roscher indicates that it had gained thorough acceptance in Germany by the middle of the 19th century.

Italian Writers

At least three pre-twentieth century Italian writers presented versions of the quantity equation. They include P. Frisi in 1772, L. Cagnazzi in 1813, and M. Pantaleoni in 1889. Of these, Frisi has already been discussed above and for that reason will be treated only briefly here. As previously mentioned, his equation, as presented in his review of Henry Lloyd's *An Essay on the Theory of Money* (1771), is

$$(17) \quad P = \frac{MC}{QV}$$

where P is price, M is money, Q is quantity of goods, C is number of buyers (a crude proxy for real demand), and V is number of sellers (a proxy for real supply). In essence, Frisi's equation constitutes a naive attempt to decompose the price level into its nominal (monetary) and real determinants. In this connection, he argues that the ratio of money to output M/Q captures the monetary factors affecting prices while the ratio of buyers to sellers C/V captures the real factors.[12] What he overlooks is that the real factors underlying prices are already accounted for by the output variable Q so that the other variables C and V are unnecessary. This, plus the omission of a velocity term, renders his equation defective.

Also defective is Cagnazzi's equation, but for a different reason : it omits the price variable. No price term appears in his formula, which he presents in his *Elementi di Economia Politica* (1813), namely

[9] See Theocharis [26, p. 109].

[10] On Rau and Roscher, see Marget [17, pp. 10-11].

[11] Marget [17, p. 10].

[12] On Frisi's equation, see the references cited in footnote 3.

$$(18) \quad Mc = DC$$

where M is the money stock, c its velocity of circulation, D the quantity of goods, and C their velocity of circulation.[13] Cagnazzi claims that his equation describes market equilibrium between the flow of money and the flow of goods. Without a price term, however, his equation makes little sense since it equates dimensionally dissimilar magnitudes. It equates one flow having the dimensions dollars per unit of time with another flow having the dimensions real quantity per unit time. To render the latter flow dimensionally comparable to the former, he should multiply goods by their dollar prices.

Cagnazzi's equation was the first to include a velocity coefficient on the goods variable. Conventional quantity equations of course dispense with that coefficient (or implicitly assign it a magnitude of unity). They do so on the grounds that since the PQ side of the equation summarizes a continuing process, i.e., an ongoing flow of physical goods and services sold, each item transferred should be treated as if it were sold but once before disappearing from economic circulation. That is, each good should be treated as if it had a turnover velocity of one. On this logic, items transferred more than once are to be counted as additional goods each time they are sold. For example, if a single item such as a house were sold four times during the period for which PQ is measured, it would be counted in the Q variable as four houses. In this way, the goods variable itself registers commodity turnover; no velocity coefficient is needed. Cagnazzi, however, proposed that such transfers be registered by a velocity coefficient. Here is the first appearance in the equation of a term denoting the velocity of circulation of goods, a concept later embodied in the quantity equations of the Frenchmen Levasseur and Walras and of the Americans Bowen and Kemmerer.

Maffeo Pantaleoni, in his *Pure Economics* (1898), also endorsed the goods-velocity concept. The volume of business transactions, he said, resolves itself into two elements: the quantity of goods offered for sale and the number of times each good is bought and sold for money. Having acknowledged the goods turnover concept, however, he failed to assign it a specific symbol in his quantity equation

$$(19) \quad v = \frac{m}{qr}$$

where v is the value of the monetary unit (or inverse of the price level l/P), m is the volume of business transactions (or Q), q is the quantity of money (or M), and r is its rapidity of circulation (or V). He did, however, present his equation as a money demand/money supply theory of price level determination. He defined the numerator m of the right hand side of his equation as real money demand and the denominator qr as nominal money supply. Today we would define m/r as real money demand and q as nominal money supply. Their quotient-the ratio of money demand to money supply-determines the value of money and hence the price level. He also stated the quantity theory of money according to which, for given values of the transactions and velocity variables, the value of money varies equiproportionally with its quantity.

French Writers

Quantity equations made their debut in the English, German, and Italian literature no later than the early 1800s. Not until the middle of the century, however, were they first seen in French monetary texts. E. Levasseur in his *La Question de l'Or* (1858) was the first French writer to present a quantity equation.[14] Like Pantaleoni, he argued that the value of money is determined by the ratio of real money demand to nominal money supply, the former defined by him as the quantity of goods for sale times their rate of turnover and the latter defined as the money stock times its circulation velocity. To illustrate this proposition he writes the equation

$$(20) \quad \text{value of money} = \frac{1}{P} = \frac{TC}{(M-R)C' + C_r}$$

where P is the price level, T is the total sum of goods and services for sale, C their circulation velocity, (M-R) is the portion of the total quantity of precious metals M that circulates as money-the remainder R being reserved for nonmonetary uses-, C' is the circulation velocity of metallic money, and C_r denotes credit instruments serving as nonmetallic means of payment multiplied by their velocities. Except for the inclusion of the velocity of circulation of goods C, Levasseur's equation is virtually the same as Irving Fisher's equation. This can be seen by omitting the C variable and replacing Levasseur's terms T, (M-R), C', and C_r by their Fisherian counterparts T, M, V, and M'V' to obtain Fisher's formula

[13] On Cagnazzi, see Marget [17, p. 11] and Theocharis [26, pp. 39-40].

[14] On Levasseur, see Wu [31, pp. 191-3].

$$(21)\quad \frac{1}{P} = \frac{T}{MV + M'V'}$$

which implies that the price level adjusts to equilibrate money demand and supply.

Sixteen years after Levasseur, Leon Walras, in the first edition of his *Eléments d' économie politique pure* (1874), also presented a Fisherian equation.[15] In addition, he formulated the quantity equation in its alternative cash balance form, becoming the first person to do so. Also, he augmented the latter equation with a base/multiplier component to account for the relationship between high-powered (metallic) money and the rest of the money stock. His contributions are outlined below.

Regarding the Fisherian equation, he derives it in two steps. First, he assumes that the means of payment consists solely of metallic money so that the equation is :

$$(22)\quad \alpha'' Q''_a = \alpha' Q'_a + \beta Q_b p_b + \gamma Q_c p_c + \delta Q_d p_d \dots.$$

where Q''_a is the stock of (metallic) money; Q'_a the quantity of the monetary metal in its nonmonetary commodity uses; $Q_b, Q_c, Q_d \dots.$ the quantities of the other goods B, C, D . . . . exchanged against money; α'' is the velocity of money; α' is the velocity of monetary metal in its commodity uses; $\beta, \gamma, \delta \dots.$ are the velocities of circulation of the goods B, C, D . . . .; and $p_b, p_c, p_d \dots.$ are the prices of goods B, C, D . . . . stated in terms of money (p_a, the price of the monetary metal in terms of itself, by definition being unity).[16] In Fisher's notation:

$\alpha'' = V$,

$Q''_a = M$,

$Q'_a + Q_b p_b + Q_c p_c + Q_d p_d + \dots. = \Sigma pQ$ and

$\alpha', \beta, \gamma, \delta \dots. = v$ (the velocity of circulation of goods).

Expressed this way, Walras' equation is :

$$(23)\quad MV = \Sigma pQv.$$

As a second step, Walras adds to the left-hand side of his equation the term F (or M'V' in Fisher's notation) to represent the value of exchanges effected by means of fiduciary (nonmetallic) money. The result is the augmented expression

$$(24)\quad MV + M'V' = \Sigma pQv$$

which, except for the v or goods-velocity term, is the same as Fisher's formula.

Walras' next contribution is his cash balance equation. This states that the nominal stock of money M must just equal the demand for it, this demand being the aggregate nominal value of goods kPQ the command over which people desire to hold in the form of cash. In his *Théorie de la Monnaie* (1886) he writes the cash balance equation as

$$(25)\quad Q''_a = \alpha + \beta p_b + \gamma p_c + \delta p_d$$

where $\alpha, \beta, \gamma, \delta \dots.$ denote the respective quantities of the goods A, B, C, D . . . . the money value of which people require to hold in the form of cash; Q''_a is the quantity of money needed to satisfy these requirements; and $p_b, p_c, p_d \dots.$ are the money prices of the goods B, C, D.[17] This expression is essentially the same as Keynes' famous cash balance equation n = kp presented almost 37 years later in his *Tract on Monetary Reform* (1923), where n is money, k is the collection of goods the command over which people desire to hold in money form, and p is the price of those goods. Indeed, Walras elsewhere presents his equation in Keynesian form, writing it as

$$(26)\quad Q''_a P_a = H$$

where H is the demand for real balances (Keynes' k), Q''_a is the quantity of metallic money (Keynes' n), and P_a is the value of money (the inverse of Keynes' p). From this equation Walras reached the strict quantity theory conclusion: given the demand for real balances H, the value of money P_a varies in inverse proportion to its quantity Q''_a.

Finally, in the 2nd ed. of his *Eléments* and in his 1898 *Etudes* [30], Walras adds to his cash balance equation the term F resulting in the augmented expression

$$(27)\quad (Q''_a + F) P_a = H$$

where F is defined as the stock of fiduciary (nonmetallic) money in circulation. Denoting such fiduciary money F as a fixed multiple f of the stock of metallic money Q (i.e., F = fQ) and substituting this expression into the one preceding it yields

$$(28)\quad Q''_a (1+f) P_a = H.$$

Here are four key ingredients of modern monetarist analysis, namely the stock of high-powered or base money Q''_a, a money multiplier (1 + f), the demand

[15] What follows draws heavily from Marget's [16] classic study of Walras' work on quantity equations.

[16] Marget [16, p. 577].

[17] Marget [16, p. 580].

for real balances H, and the value of money P_a or its inverse, the general price level.[18] All this in an equation presented in 1898, fully 13 and 19 years, respectively, before the appearance of Fisher's and Pigou's equations.

Additional evidence that French monetary theorists had fully developed algebraic quantity equations before Fisher and Pigou comes from A. de Foville. His book *La Monnaie* (1907) contains the expression

$$(29) \quad \frac{P}{p} = \frac{M}{m} \frac{c}{C} \frac{V}{v}$$

where P denotes prices, M the money stock, V its velocity, C the quantity of commodities exchanged against money, and the upper-case and lower-case letters refer to the magnitudes of these variables on any two different dates. Written in ratio form, Foville's expression explains the relative change in the price level between any two dates as the product of the underlying relative changes in its money, velocity, and output determinants.

American Writers

After a late start, the quantity equation developed rapidly in the United States, progressing from an initial incomplete version in the mid-1850s to an elaborate disaggregated version in the early 1900s. The major steps in this progression can be outlined briefly. In 1856 Francis Bowen presented in his *The Principles of Political Economy* the equation

$$(30) \quad gs = mr$$

where g is the quantity of goods sold, s is the number of times the goods are sold, m is the quantity of money in circulation, and r is its rapidity of circulation. Bowen's equation, expressing as it does an equivalence between a flow of goods and a flow of money, is the same as that presented earlier by Cagnazzi and suffers from the same defect, namely the omission of a price-level variable necessary to render the two sides dimensionally comparable.

Simon Newcomb corrected this defect in his "equation of societary circulation" which he presented in his *Principles of Political Economy* (1885). Newcomb's equation is

$$(31) \quad VR = KP$$

where V is the volume of the currency, R its average rapidity of circulation, K the number of real transactions, and P the price level. From his equation he concluded that prices vary equiproportionally with changes in the money stock since the latter can have no lasting effect on the steady-state levels of the real variables R and K. Equilibrium values of these real variables, he said, are immune to monetary change such that the latter registers its full impact on prices only. To explain how money affects prices, he constructs an aggregate demand function from the components of his quantity equation. Like modern monetarists who define real aggregate demand as money stock times velocity divided by prices (MV/P) he writes the demand function as:

$$(32) \quad D = \frac{NVR}{P}$$ [19]

where D is the quantity of goods demanded, N is a fixed constant, and V, R, and P are the volume-of-currency, rapidity-of-circulation, and price-level variables as defined above. This equation says that, whereas demand varies directly with money and inversely with prices, it is unaffected by equiproportional changes in both variables. Thus, according to Newcomb, a monetary expansion initially puts upward pressure on real demand. But the resulting rise in demand subsequently bids up prices, which eventually rise equiproportionally with money, thus restoring real demand to its original level. In steady-state equilibrium, prices vary proportionally with money, and the latter is neutral in its effect on real variables-just as the quantity theory predicts.

While endorsing the quantity theory, however, he was quick to point out that it holds only if prices are flexible. He put his quantity equation VR = KP to work in demonstrating that price inflexibility would render monetary changes nonneutral in their effect on real activity. For, with prices P slow to adjust to monetary shocks, the real transactions K term of the equation VR = KP would have to bear some of the burden of adjustment to currency contraction. Also, he noted that, with velocity given, autonomous rises in prices P engineered by monopolistic sellers would result in compensating falls in real activity K if the money stock V were held constant. Despite this, he warned that a policy of validating or

[18] Marget [16, p. 585].

[19] More precisely, he writes

$$D = \frac{NF}{P}$$

where F ("the flow of the currency") is defined as

$$F = VR.$$

Substituting this latter expression into the former yields equation 32 of the text.

underwriting such price increases with money growth in an effort to maintain full employment would only serve to perpetuate inflation. The full employment guarantee, he claimed, would encourage sellers to raise prices repeatedly. Each time accommodating money growth would follow. In this way prices and money would chase each other upward ad infinitum in a cumulative inflationary spiral. Newcomb's work strongly influenced Irving Fisher, who derived his famous equation of exchange from Newcomb's formulation and who dedicated his *The Purchasing Power, of Money* (1911) to Newcomb.

Following Newcomb, the quantity equation appeared with increasing frequency in the U. S. monetary literature. Arthur Hadley helped to popularize it by incorporating it into his well-known textbook *Economics* (1896) in the form

$$(33)\quad RM = PT$$

with M being money, R its rapidity of circulation, P prices, and T the volume of real transactions. Similarly, Edwin W. Kemmerer employed it in his *Money and Credit Instruments in Their Relation to General Prices* (1907), stating it alternatively as a price equation and a money supply/demand equation. The price equation version he writes as

$$(34)\quad P = \frac{MR + CR_c}{NE + N_cE_c}$$

where P is the price level; M the money stock (coin, currency, bank notes), R its average rate of turnover, C the dollar amount of checks in circulation, R_c the average rate of check turnover, N and N_c are the number of commodities exchanged by means of money and checks, respectively, and E and E_c are the average number of exchanges of these goods (that is, their velocities of circulation). This equation expresses the equilibrium price level as the quotient of its monetary and real determinants, which he identifies with money supply and demand. He obtains his alternative money supply/demand expression by rewriting the equation as

$$(35)\quad MR + CR_c = P(NE + N_cE_c).$$

According to him, the left-hand side measures money supply, the right-hand side measures money demand, and the price level P adjusts to equilibrate the two. Except for the inclusion of velocity in his concept of the money supply, his analysis is the same as modern monetarists'.

Five years before Kemmerer, John P. Norton, in his *Statistical Studies in the New York Money Market* (1902), presented perhaps the most elaborate version of the quantity equation to be found in the literature. He includes separate terms for each type of coin and currency in circulation. He distinguishes between the velocity of demand deposits and the velocity of coin and currency. He expresses the total of demand deposits, in terms of its three underlying components, namely bank reserves, the deposit-expansion multiplier, and the proportion of maximum allowable deposits that banks actually create. Lastly, he shows the effect of loan extension and repayment on the equation. He does all this in the following way.

First, he starts with the equation's money or MV side, writing it as

$$(36)\quad E = (MV + DU)T$$

where E is total monetary expenditure, M is money narrowly defined (coin and notes), V is its velocity, D is the volume of demand deposits, U is the turnover velocity of deposits, and T is the number of units of time for which these variables are measured. He then disaggregates the money variable M into its constituent components, namely gold coin G, silver coin S, silver certificates C, United States notes N, National bank notes B, and all other forms of currency L. That is, he defines money M as

$$(37)\quad M = G + S + C + N + B + L.$$

Third, having expressed money in terms of its coin and currency components, he next expresses demand deposits in terms of the reserves backing them. More precisely, he defines such deposits D as the product of the reserve base R, the deposit-expansion multiplier Z (which determines the maximum of deposits per dollar of reserves), and the proportion K of maximum allowable deposits that banks actually create. In short, D = ZKR. He multiplies these deposits by their turnover velocity U and aggregates over the four classes of banks in existence in 1902 to obtain the expression

$$(38)\quad DU = \sum_{i=1}^{4} Z_iK_iR_iU_i$$

where the subscript i indexes the type of bank (country, reserve city, central reserve, and state). He then substitutes this equation and the one immediately preceding it into equation (36) to get his final expression for the MV side:

$$(39)\quad E = [(G+S+C+N+B+L)V + \sum_{i=1}^{4} Z_iK_iR_iU_i]T$$

where E is total expenditure.[20]

He next attempts to show how bank loan extension and repayment affects the equation. He argues that loan repayment temporarily absorbs expenditures that otherwise would be directed toward goods just as loan extension expands them. To show the effect of loan retirement, he adds to the goods or PQ side of the equation the term ΣM_s denoting banks' receipt of "spot" or current dollars M_s as borrowers repay loans. Similarly, to show how loan creation increases expenditure he adds to the opposite side of the equation the term $\Sigma(1-d)M_f$, denoting banks' acquisition of loan assets or "claims to future dollars" M_f, each such dollar valued at its discounted price (1-d), where d is the discount rate on loans. For convenience, he then transposes this term to the goods side of the equation such that the latter reads

$$(40)\quad \Sigma pQ + \Sigma M_s - \Sigma(1-d)M_f$$

where the first term denotes expenditure on goods and the last two terms denote net debt repayment. Finally, he equates both sides of the equation to obtain the entire expression

$$(41)\quad \Sigma pQ + \Sigma M_s - \Sigma(1-d)M_f = [(G+S+C+N+B+L)V + \sum_{i=1}^{4} Z_iK_iR_iU_i]T.$$

[20] Note that Norton defines the velocity V term of equation (39) as the average of the individual velocities of each currency component weighted by each component's share in the entire stock. He writes

$$V = \frac{GV_G+SV_S+CV_C+NV_N+BV_B+LV_L}{G+S+C+N+B+L}.$$

Here is another example of his elaborate derivation of the equation's components.

This equation, together with those of Newcomb, Hadley, and Kemmerer, prepared the way for the appearance of Fisher's equation in 1911.

Conclusion

Irving Fisher and A. C. Pigou presented their famous quantity equations in the second decade of the 20th century. By that time, however, their contributions had already been largely or fully anticipated by at least 19 writers located in five countries over a time span of at least 140 years. Except for some primitive initial versions, these writers formulated equations that in all essential respects were virtually the same as their Fisherian and Pigovian counterparts, and in at least two cases were even more detailed and sophisticated than the latter.

Not only did these earlier equations include the same variables and possess the same properties as their celebrated modern counterparts, they also embodied the same analysis. Their authors presented them either as price equations expressing P as a mathematical function of the variables M, V, and Q, or as money-supply-and-demand equations expressing an equilibrium condition between the money stock and the underlying determinants of the demand to hold it. In any case, earlier writers perceived their quantity equations as functional relationships and not as mere identities, just as Fisher and Pigou likewise were to do. Recognition of this fact renders invalid the typical textbook identification of Fisher and Pigou as "the original sources of the equation of exchange and the cash balance equation" [1, p. 98]. Far from being the source of such equations, those writers were the recipients or inheritors of them. In short, whereas quantity equations may have *culminated* in the writings of Fisher and Pigou, they did not begin there. As documented in this article, their source is to be found elsewhere.

References

1. Barger, H. *Money, Banking, and Public Policy.* Chicago, 1962.
2. Bowen, F. *The Principles of Political Economy.* London, 1856.
3. Briscoe, J. *Discourse on the Late Funds . . .* London, 1694.
4. Cagnazzi, L. de. *Elementi di Economia Politica.* Naples, 1813.
5. Fisher, I. *The Purchasing Power of Money.* New York, 1911.
6. Foville, A. de. *La Monnaie.* Paris, 1907.
7. Frisi, P. "Estratto del saggio sulla teoria della moneta de1 General Lloyd." Reproduced in M. Fasiani,. "F. Fuoco : Applicazione dell' Algebra all' Economia Politica." Genoa, 1937.
8. Hadley, A. T. Economics. New York, 1896.
9. Kemmerer, E. W. *Money and Credit Instruments in Their Relation to General Prices.* New York, 1907.
10. Keynes, J. M. *A Tract on Monetary Reform.* London, 1923.
11. Kröncke, *C. Das Steuerwesen nach seiner Natur und seinen Wirkungen untersucht.* Darmstadt and Giessen, 1804.
12. Lang, J. *Grundlinien der politischen Arithmetik.* Kharkov, 1811.
13. Levasseur, E. *La Question de l'Or: les mines de Californie et d' Australie.* Paris, 1858.
14. Lloyd, H. *An Essay on the Theory of Money.* London, 1771.
15. Lubbock, J. On Currency. London, 1840.
16. Marget, A. W. "Léon Walras and the 'Cash-Balance Approach' to the Problem of the Value of Money." *Journal of Political Economy* 39 (1931), 569-600.
17. ——— *The Theory of Prices.* Vol. 1. New York, 1938:
18. ——— *The Theory of Prices.* Vol. 2. New York, 1942:
19. Newcomb, S. *Principles of Political Economy.* New York, 1885.
20. Norton, J. P. *Statistical Studies in the New York Money Market.* New York, 1902.
21. Pantaleoni, M. *Pure Economics.* trans., T. B. Bruce, London, 1898. First Italian edition, 1889.
22. Pigou, A. C. "The Value of Money." Quarterly *Journal of Economics* 32 (1917-18). Reprinted in *Readings in Monetary Theory*, ed. F. A. Lutz and L. W. Mints. Philadelphia, 1951, pp. 162-183.
23. Rau, K. H. *Grundsätze der Volkswirtschaftslehre.* 4th ed. Leipzig and Heidelberg, 1842.
24. Roscher, W. *Die Grundlagen der Nationalökonomie.* Stuttgart, 1854.
25. Schumpeter, J. A. *History of Economic Analysis.* New York, 1954.
26. Theocharis, R. D. *Early Developments in Mathematical Economics.* 2nd ed. Philadelphia, 1983.
27. Turner, *S. A Letter Addressed to the Right Hon. Robert Peel with Reference to the Expediency of the Resumption of Cash Payments at the Period Fixed by Law.* London, 1819.
28. Walras, L. *Eléments d'* économie *politique pure.* 1st ed. Lausanne, 1874.
29. ——— *Théorie de la monnaie.* Lausanne, 1886.
30. ——— *Etudes d'économie politique appliquée.* Paris, 1898.
31. Wu, *C-Y. An Outline of International Price Theories.* London, 1939.