Currency Velocity and cash payments in the U.S. Economy: The Currency Enigma

Feige, Edgar L.

University of Wisconsin-Madison

1989

CURRENCY VELOCITY AND CASH PAYMENTS IN THE U.S. ECONOMY: THE CURRENCY ENIGMA

Edgar L. Feige [1]

I. Introduction

In a monetary exchange economy all economic activity can be analyzed from the dual perspectives of payments and transactions. The equality between the aggregate sum of transactions and the payments used to effect them is expressed in Fisher's (1911) famous equation of exchange which states that MV = PT.

A major achievement of the last half century of macroeconomic analysis has been the empirical and theoretical elaboration of the transactions side of the equation of exchange with particular emphasis on the transactions involving expenditures on final goods and services. The empirical decomposition of the value of transactions (PT) took two forms. First, (PT) was disaggregated into major categories representing: final transactions for domestically produced goods and services; intermediate transactions; transactions in real and financial assets and foreign transactions. Estimates of each of these transaction categories have been embodied in the complex systems of national accounts respectively represented by the National Income and Product Accounts; Input-Output Accounts; Flow of Funds Accounts and Balance of Payments Accounts. Second, the nominal value of transactions in various sectors of the economy have been decomposed into price and real components which are universally represented by index numbers reflecting changes in prices and quantities. We now have a plethora of quantity weighed price indices and price weighted quantity indices that enable economists to study both the causes and consequences of changes in nominal and real magnitudes.

A notable oversight of macroeconomics during the last half century has been its failure to empirically elaborate an analogous framework for the analysis of the payments (MV) side of economic activity and to link conceptually consistent estimates of total payments to independent estimates of total transactions.[2] Despite the wealth of data

[2] Feige (1979, 1985) developed provisional time series estimates of MV and its components for both the U.S. and Sweden. Cramer (1980, 1986) has constructed estimates of the total volume of monetary payments associated with the production of final output by employing data from the transactions side of the equation of

pertaining to monetary aggregates, there are as yet no reliable estimates of the total volume of monetary payments.[3] Payments have not been disaggregated by sector and little attention has been given to the appropriate decomposition of nominal payment changes into their constituent components, namely changes in velocity weighted monetary aggregates and money weighted changes in velocities.[4]

The asymmetric development between the elaboration of the payments[5] and transactions components of the equation of exchange is in large part the result of a long standing empirical gap in our knowledge concerning the volume of payments undertaken with currency. The problem of adequately measuring cash payments is due to the absence of measures of the velocity of currency, a lacuna in our knowledge that has remained essentially unresolved since the issue was posed by Jevons in 1895.

> "I have never met with any attempt to determine in any country the average rapidity of circulation, nor have I been able to think of any means whatever of approaching the investigation of the problem, except in the inverse way. If we knew the amount of exchanges effected, and the quantity of currency used, we might get by division the average number of times the currency is turned

---

exchange. Cramer's transactions estimates are not however consistent with estimates of payments by money stock holders because they include imputations for income in kind and federal government expenditures which are not included in conventional monetary payment measures. Cramer's measures also exclude imports which represent legitimate payments to foreigners. Finally, Cramer's measures of transactions employ net rather than the required gross payments of interest and dividends. Corrado and Spindt (1988) adjust Cramer's transactions estimates to bring them into closer conformity with a measure of the monetary payments associated with the production of total output. Corrado and Spindt define T as " a measure of total cash transactions by money stock holders" (p. 3). The "T" defined by Corrado and Spindt should not be confused with the T that appears in the conventional equation of exchange. The former includes monetary payments by money stock holders associated with the production of final output as measured from the transactions side of the equation of exchange accounts. It represents total payments for final goods, intermediate goods and factor payments. The T employed in the equation of exchange is a much broader concept, including all payments by money stock holders associated with transfers of existing real assets and transfers of existing and newly created financial assets and liabilities.

[3] The Federal Reserve does collect debit statistics on demand deposits which reflect check payments by deposit ownership catagories which include households, nonfinancial business and state and local governments. To date, there are however no time series estimates of currency payments.

[4] Provisional attempts to obtain index number estimates of the money stock and velocity have been undertaken by Feige (1985) and Spindt (1985).

[5] The final media of exchange are currency (C) and checkable deposits (D). Aggregate payments effected by the final media of exchange are composed of cash payments (C V_c) and check payments (D V_d), where V_c and V_d respectively represent the "payments velocity" of currency and demand deposits. Check payments by money holder categories are reasonably well documented by statistics on debits to demand deposit accounts but the volume of cash payments requires knowledge of the velocity of currency.

over; but... the data are quite wanting"[6]

The importance of estimating the velocity of currency was clearly recognized by Fisher in the early part of the century.[7] Fisher argued that:

"The importance of such accurate determinations (of the velocity of currency) can scarcely be overestimated.. When we know statistically the velocity of circulation of money we shall be in a position to study inductively the "quantity theory" of money, to discover the significance of that velocity in reference to crises, accumulations of wealth, density of population, rapid transit and communication, as well as many other conditions. In fact a new realm in monetary statistics will have been opened". (p. 618)

Yet, despite Fisher's claims for the theoretical and empirical importance of the payment velocity of currency, and his early efforts to devise a method for calculating currency velocity, little further attention was given to the problem for almost sixty years.[8]

The intervening years did, however, produce a plethora of studies[9] of "income velocity," often in the guise of studies of the demand for money. But, what is commonly known as "income velocity," namely, the ratio of observed income to the money stock, is neither a correct measure of the number of times money is turned over in the process of producing final goods and services,[10] nor is it a useful theoretical construct for predicting the consequences of changes in the money supply on economic activity. What is remarkable, is the tenacity with which economists continue to cling to the "income velocity" concept despite Keynes' own warning that income velocity is "a hybrid conception having no particular significance"[11] whose use, "has led to nothing

---

[6] Jevons, W. S. (1875) quoted in Selden, R. (1956).

[7] Fisher, I. (1909); (1911).

[8] An important exception is a Ph.D dissertation by R. Laurent (1970). Laurent employed an income version Fisher's equation of exchange to estimate currency velocity. More recently, Cramer (1980) and Spindt (1985) employed crude versions of Fisher's procedure to obtain an estimate of currency velocity. The shortcomings of this approach are discussed in the following section.

[9] See Selden (1956) ibid. for an excellent review.

[10] Payments associated with the production of final goods and services must include not only payments for final expenditures on goods and services but also factor payments and intermediate payments associated with production. Money stock measures exclude federal government money holdings yet GNP measures include federal government expenditures on final goods and services.

[11] Keynes, J. M. (1930) p.24

but confusion."[12]

What is needed is a method for directly estimating the payments velocity of the exchange medium without employing the equation of exchange.[13] Rather than estimate velocity indirectly as a ratio of some transaction aggregate to the money supply,[14] we require independent estimates of the turnover of each component of the medium of exchange in order to measure the work money does in the economy to effect payments. The velocity of money would then no longer be a residual measure derived from an income variant of the equation of exchange, but would be directly estimated from observable data pertaining to the turnover of checkable deposits and currency. Such independent estimates of payments velocity would regain a crucial degree of freedom now lost by employing the equation to <u>define</u> velocity. This would free the equation of exchange to perform the functions of an accounting identity and an equilibrium condition on payments and transactions.

Used as a high level accounting identity, the equation of exchange could be used to develop an aggregate accounting system in which independent estimates of the total volume of payments by money holders could be used as a reliability check on all existing macroeconomic accounting systems based on indirect estimates of aggregate transactions.[15] In an age of electronic payments systems, monetary payments could be monitored in real time, permitting high frequency reliable observations of all real and financial economic activity.[16] In principle, it should be possible to develop a real time

---

[12] Keynes, J. M. (1936) p.299

[13] Income velocity as conventionally defined is in fact the product of payments velocity and the ratio of final goods and service transactions to total transactions.

[14] Employing the equation of exchange to <u>define</u> velocity is to sacrifice a crucial degree of freedom. When velocity is so defined, the equation of exchange is deprived of its usefulness as either an accounting identity or as an equilibrium condition.

[15] Total transactions, would comprise the sum of payments for final goods and services, factor payments (GNP + NI) intermediate transactions, (IO accounts) foreign current and capital account transactions, (Balance of Payments Accounts) and <u>gross</u> payments for real and financial assets (A modified Flow of Funds Account) plus other transfer payments. These transactions based accounting systems suffer from a variety of problems. Their accuracy is impaired because of their heavy reliance on survey data that suffer from non response and self selection biases, benchmarks and an increasing amount of imputation. Transaction based data are typically low frequency observations that require repeated revision.

[16] This would require the disaggregation of debits by type of payment. One possible approach would be to identify both payor and payee by SIC codes which can be derived from account identification numbers that are already required by law. Aggregate payments to different SIC classes could be accumulated in electronic

GNP clock from the payments side of the equation of exchange. In short, the equation of exchange could serve as an accounting identity in much the same way as we now employ the income-expenditure identity as a check on the reliability of the two methods for obtaining NIPA estimates of GNP.

Furthermore, independent measures of the turnover of each medium of exchange would permit the construction of more meaningful definitions of "the money supply" and "the average velocity of money". These new definitions would be constructed as index numbers involving turnover weighted money aggregates and money weighted turnover rates respectively. Money aggregates would be weighted by the intensity of money use just as money use would be weighted by the alternative media of exchange. These indices would finally produce conceptual symmetry between payments and transactions analysis, suggesting new possibilities for the long sought after integration of monetary and value theory. Just as all transactions are represented as the product of a price index and a real quantity index, payments would be analogously represented as the product of a money index and a velocity index. None of these extensions can be undertaken without independent estimates of currency velocity and hence of cash payments. Finally, estimates of total monetary payments could be used as the appropriate transactions variable in money demand functions that currently employ income as the scale variable.[17]

Interest in estimating the volume of cash payments goes beyond their usefulness in exploiting the implications of the equation of exchange. The past decade has witnessed considerable concern about the existence of a large and growing "underground economy" manifested by under reporting of income on tax returns.[18] Currency is viewed as the preferred means of payment for tax evasion and illicit activities. Cash payments rarely

---

data collection bins appended to the normal clearing mechanism. The moment an individual payment was routed to its appropriate data bin, individual account identification would be obliterated in order to maintain the strict confidentiality of the payment mechanism.

[17] Most money demand functions implicitly assume a time invariant proportional relationship between income and total transactions as the justification for employing income as the scale variable. Since there are as yet no consistent estimates of the total volume of gross transactions (PT), independent estimates of (MV) can be used as a proxy for (PT) in order to permit an explicit examination of the reasonableness of the assumed proportionality between income and total transactions. A recent study (Dotsey, 1988) of the demand function for currency reports that "GNP is likely to be a poor proxy for the value of expenditures" and that, "The poor performance of GNP casts doubts on the empirical procedures used in a number of studies that employ GNP" (p. 23).

[18] According to the Internal Revenue Service (1983) some $580 billion of income is not properly reported on Federal Income Tax returns. See Feige (1989a).

leave an audit trail, giving currency a unique advantage over check payments as a means of avoiding detection. Unreported taxable income reduces tax revenues and thus magnifies the problem of budget deficits. Moreover, since tax information is widely used as a data source in the construction of National Income and Product Account (NIPA) statistics, the growth of unreported income has the insidious effect of corrupting the reliability of the primary data used for most macroeconomic analysis.[19]

Interest in cash payments is also motivated by the failure of widespread predictions concerning the coming of the "cashless society". Per capita currency holdings in the U.S. are now approximately $1000 and one hundred dollar denomination bills make up 45% of the value of the nation's currency supply. Moreover, contrary to most expectations, the stock of U.S. currency has grown dramatically. Yet recent Federal Reserve surveys[20] of currency usage by American households suggest that adult U.S. residents hold only 10-12% of the nation's stock of currency in circulation outside of banks. Allowing for U.S. business holdings of currency, the whereabouts of perhaps as much as 80% of the nation's currency supply is presently unknown. These findings have given rise to what is known the problem of "missing currency".

Finding an adequate solution to the empirical problem of estimating currency velocity is an therefore an important first step in addressing a number of substantive problems in monetary economics and public finance. Independent estimates of currency velocity are required in order to estimate index number measures of monetary aggregates and velocity aggregates and to estimate total cash payments and hence, of total monetary payments. Independent estimates of total payments in turn permit the empirical and conceptual elaboration of the equation of exchange in a manner analogous to that found so fruitful for the income-expenditure equation and allows the re-estimation of money demand functions with an appropriate transactions scale variable. Estimates of total currency payments are also useful as a measure of "underground economic activity" and can be used to analyze the implications of "missing currency" as well as the reliability of present NIPA estimates of income and expenditures.

In light of the aforementioned importance of currency velocity and cash payments for a number of macroeconomic issues, this paper seeks to evaluate earlier methods for

---

[19] IRS estimates of unreported income have necessitated a major, yet largely unnoticed revision of NIPA accounts. For the year 1984, estimates of National Income included a total adjustment of $147 billion to reflect the effect of misreporting on income tax returns. [Survey of Current Business, December, 1985, Table 4].

[20] Federal Reserve Bulletin, (1986) and (1987)

obtaining estimates of currency velocity and cash payments and develops new estimates based on alternative procedures. The paper begins with an institutional description of the environment that determines the provision, maintenance and use of paper currency in the U.S. The specific goal of the paper to is to develop a framework for the analysis of currency transactions that permits estimation of the following particular aspects of the currency circulation system:

1) The average lifetimes of specific currency denominations.

2) The total number of lifetime turnovers that currency notes can sustain.

3) The velocities of specific denominations of currency.

4) Changes in the quality of the outstanding supply of currency.

5) The volume of cash payments effected by each currency denomination.

Section II briefly describes the institutional setting of the currency circulation system. Section III reviews and evaluates earlier attempts to estimate currency velocity. Sections IV and V use a simple payments model to estimate denomination specific average lifetimes of currency, average turnover and average holding periods for currency. Section VI and VII develops comparable estimates based on a less restrictive demographic model of the currency "population" and presents empirical estimates of currency velocity. The final section presents estimates of cash payments made with U.S. currency for the period 1941 - 1985 and the implications of these estimates. The estimated volume of cash payments is so large that it gives rise to what we refer to as the "currency enigma", namely the problems of (1) "missing currency"; (2) missing payments" and (3) "missing income and expenditures".

II. Overview of the Operation of the Currency Circulation System

The supply of currency is ultimately determined by the public's demand for currency. The responsibility for providing the nation's currency supply rests jointly with the U.S. Treasury Department's Bureau of Engraving and Printing [BEP] and the Federal Reserve System [FED]. The BEP designs and prints the currency and therefore establishes the physical specifications that determine currency durability[21] The FED issues currency into circulation and also has the responsibility of maintaining the quality of the circulating stock currency by examining notes that are received from circulation in order to determine their fitness for recirculation. When notes are determined to be

---

[21] Currency includes both coins and paper notes. Since the primary focus of this research is on circulating paper currency, the term currency is here used to denote paper currency of denominations $1; $5; $10; $20; $50 and $100.

unfit for recirculation, they are redeemed and destroyed.[22] Given the physical specifications of currency design and the Federal Reserve's standards for the maintenance of currency quality, the rate of currency redemption will approximately reflect the intensity of currency usage. Data on the rates of currency redemptions can therefore be employed to obtain estimates of currency usage.

III. Earlier Efforts to Estimate Cash Payments
A. Fisher Cash Loop Method:

At the beginning of the century, Irving Fisher (1909) devised a primitive approximation for estimating the velocity of cash payments. Fisher constructed a model of cash flows that described the various types of currency exchanges that take place between the time individuals initially withdraw cash from the banking system and subsequently redeposit it. Sectoring his hypothetical economy into firms, individual depositors and non depositors, Fisher examined all possible cash exchanges between the three sectors, and then derived an expression for the value of cash payments made to acquire goods and services.

Fisher assumed that individuals exclusively acquire cash by withdrawing it from banks.[23] Given information on the value of total withdrawals [ω], and an estimate of the number of payments that occurred between cash withdrawals and subsequent redeposits (the "cash loop" [λ]), the value of cash payments is simply:

1) $C \cdot Vc = \omega \cdot \lambda$

where:

    C represents the stock of currency in circulation;

    Vc represents the average velocity of currency;

    ω represents cash withdrawals from the banking system;

    λ represents the "cash loop".

Fisher's model specification was conditioned by the institutional structure of the

---

[22] The Annual Report of the Federal Reserve System (1987) indicates that management of the currency circulation system absorbs in excess of 25% of the Federal Reserve's net expenses.

[23] A model of currency payments that includes real and financial assets other than "money" would also permit the acquisition of cash from sales of real and financial assets. Such "sales" would of course include cash withdrawals from time and savings deposits.

payments mechanism as it functioned during the early years of the 20th century. The approximate formula he derived for the "total circulation of cash in exchange for goods" equaled the total amount of cash deposited in banks plus total wages paid. Given this estimate of cash payments and information on the stock of currency outstanding, Fisher estimated that the average velocity of currency was approximately 18 turnovers per year, implying an average holding period between cash exchanges of about three weeks, and a "cash loop" [λ] approximately equal to 1.6.

B.) Cramer and Spindt's Applications of the Cash Loop Method.

Cramer's (1980; 1986) heroic initial efforts to estimate the volume of monetary payments in the U.S employed a variant of Fisher's "cash loop" method to estimate the velocity of currency for the United States. The Fisher method requires both an estimate of the loop length and an estimate of annual cash withdrawals. Cramer (1980) recognized that, "as for the loop length, there is no conceivable method of establishing it empirically, and it must be guessed from casual observation."(p.18) Cramer's "guess" was that the loop length for the U.S. was approximately equal to two. His estimate of the annual cash withdrawals at banks was based on a 1961-1962 survey by the Federal Reserve Bank of Boston (1963) that estimated that during the survey period, cash <u>receipts</u> at banks amounted to approximately 7.5% of total receipts. By further disregarding changes in the stocks of currency and demand deposits, Cramer calculated that cash withdrawals in 1961 (ω) were approximately 7.5% of total withdrawals.[24] Given the ratio of debits to deposits, Cramer concluded that cash withdrawals "<u>must have stood at 2 times average balances.</u> (p.18) " By substituting this estimate into Equation (1), Cramer derived an estimate of Vc for the year 1962 as:

1a) $Vc = 4*D/C \approx 16$

In order to derive a time series of cash payments, Cramer simply assumed that Vc was constant over time, "set by the physical limits to the speed of currency circulation." (p.20). Cramer candidly admitted that, "This is a major assumption, weakly founded, but frankly it is hard to think of something better."(p.20)

Spindt (1985) also required a time series estimate of currency velocity in order to

---

[24] If δ = debits to demand deposits, then Cramer found that: $\omega/(\omega+\delta) = .075$ so that, $\omega = .069\delta$. By dividing both sides of the foregoing equation by demand deposits and noting that $(\delta/D) = 26$, Cramer concluded that the ratio of withdrawals to deposits $(\omega/D) = 1.8 \approx 2$.

construct a monetary aggregate index which he wished to test against conventional money stock aggregates. Spindt, attempted to apply Cramer's procedure which he described as follows:

> "Data on cash withdrawals could presumably be collected from bank's books, but these data are not regularly collected and published along with monetary statistics. Data on the average length of the cash loop are more difficult to obtain. We do, however, have some useful information. Citing bits of survey evidence, <u>Cramer (1980) estimates annual cash withdrawals at banks to be about four times demand deposit balances,</u> and the average length of the cash loop to be two payments. These numbers imply a monthly turnover rate for currency equal to 2/3 times the ratio of demand balances to the currency stock. " (p. 182)

Spindt's resulting equation for the annual velocity of currency is then:

1b) $Vc = 8*D/C$

Since all of the tests employed by Spindt to examine the usefulness of his monetary aggregate index are conditional on the accuracy of his time series estimate of currency velocity, it is important to reexamine his procedure for estimating currency velocity in greater detail. The procedure suffers from several deficiencies. First, Spindt misquotes Cramer and erroneously sets cash withdrawals at <u>four</u> times deposit balances rather than the <u>two</u> times deposit balances derived by Cramer for 1962. This error implies that Spindt's estimated level of currency velocity is overestimated by 100%. Moreover, Spindt's procedure implicitly assumes that the single survey value uncovered by Cramer for 1962 can be assumed constant over time so as to permit estimates of the velocity of currency during the 1970's and 1980's. Finally, Spindt's procedure requires an assumed value of the loop length of two for which there is no independent empirical justification. For the mid 1980's, Spindt's reported equation yields an estimate of currency velocity of approximately twenty turnovers per year. Had Spindt correctly followed Cramer's procedure, his estimate would have been ten annual turnovers. This figure should be compared to the recent Federal Reserve Board's (1986) survey estimate of currency velocity which is approximately fifty turnovers per year.

In light of the obvious shortcomings of the aforementioned efforts to empirically implement Fisher's cash loop method for estimating currency velocity, we now turn to an alternative procedure suggested by Laurent (1970).

C. The Simple Payments Model

Laurent (1970), apparently unaware of Fisher's earlier work, examined the issue of currency payments, and established the rudiments of an alternative method for

estimating currency transfers. Laurent specified a model in which the cumulative number of transfers $T^*_i(t)$ performed by notes of the ith denomination by time (t) is;

$$2) \quad T^*_i(t) = G \left\{ \int_0^t R^n_i(s) \, ds + \int_0^t \{ I^n_i(s) - R^n_i(s) \} \, \gamma_i(s) \, ds \right\}$$

where G represents the total number of turnovers that a note can sustain during its lifetime before being regarded as sufficiently soiled to be withdrawn from circulation; $R^n_i(s)$ and $I^n_i(s)$ are respectively, the number of notes of the ith denomination redeemed and issued, and $\gamma_i(s)$ is the average fraction of total lifetime transfers performed by notes still outstanding at time (s). The first term on the right hand side of equation (2) is the cumulative number of cash transfers undertaken by all notes redeemed up to time (s) and the second term represents the cumulative number of transfers performed by time (s) of all notes still outstanding at time (s). For the simplest case, Laurent assumed that the average note outstanding had performed half of its lifetime transfers so that $\gamma_i = .5$. Under this assumption, equation (2) simplifies to:

$$3) \quad T^*_i(t) = .5 \, G \left\{ \int_0^t R^n_i(s) \, ds + \int_0^t I^n_i(s) \, ds \right\}$$

The change in the cumulative number of transfers from one year to the next represents the total number of transfers $[T_i(t)]$ of the ith denomination undertaken during the year (t). Therefore,

$$4) \quad T_i(t) = .5G \{ R^n_i(t) + I^n_i(t) \}$$

Laurent proceeded to estimate the total value of currency payments, $[C(t) . V_c(t)]$ in terms of the unknown parameter G as;

$$5) \quad C(t) . V_c(t) = \sum T_i(t) . D_i$$
$$= G . \sum .5 \{ [R^n_i(t) + I^n_i(t)] . D_i \}$$
$$= G . U(t)$$

FIGURE 1

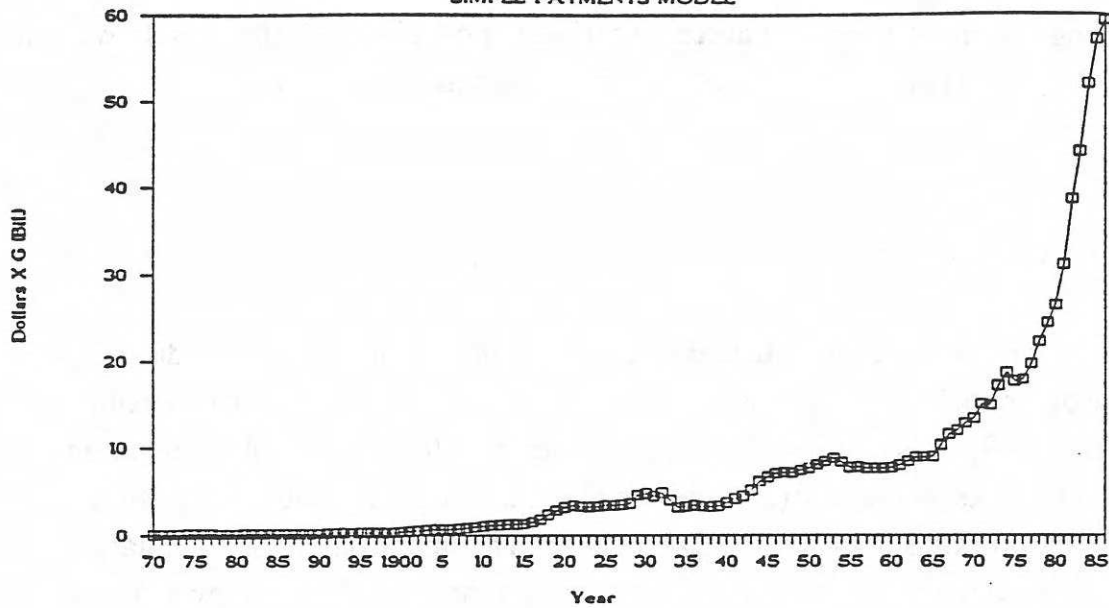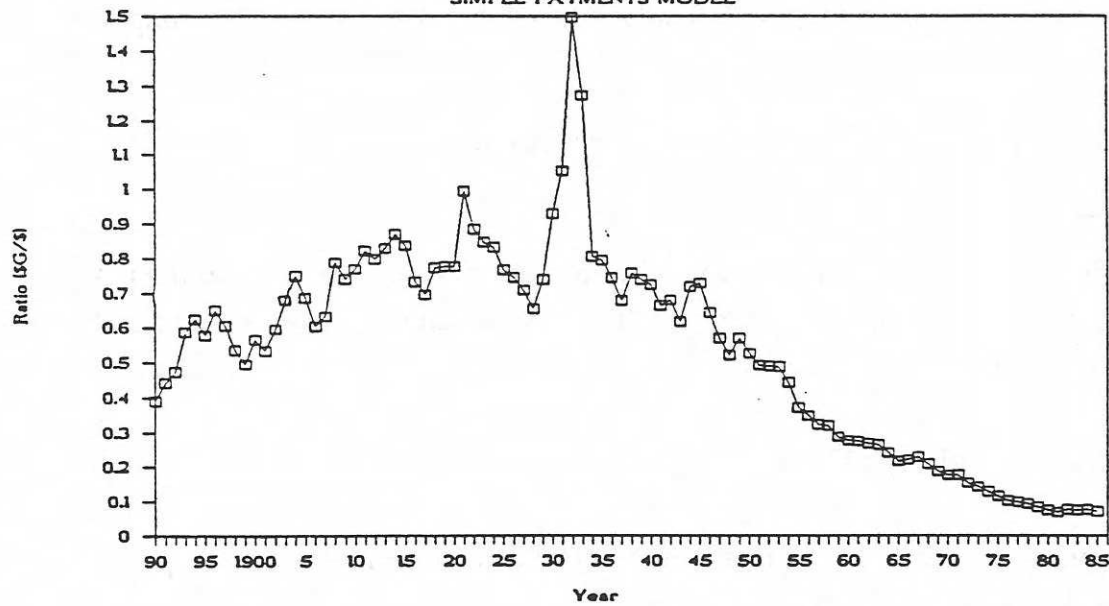# ESTIMATED CURRENCY PAYMENTS
SIMPLE PAYMENTS MODEL

FIGURE 2

RATIO: CASH PAYMENTS TO CHECK PAYMENTS
SIMPLE PAYMENTS MODEL

where D_i denotes the denomination value of the ith denomination and the summation is over all denominations i = 1, 5, 10, 20, 50, 100. Vc(t) denotes the number of times the currency stock turns over during the year (t).

Given data on currency issues and redemptions, one can estimate [U(t)] and obtain estimates of total cash payments in terms of the unknown parameter G. These estimates are displayed in Figure 1 for the period 1870 to 1986. Figure 1 reveals that the estimated volume of currency payments grew dramatically since the mid 1960's. Figure 2 displays the ratio of cash payments in terms of G to total check payments as measured by an estimate of debits to all checkable accounts.[25] Despite the substantial increase in estimated cash payments during the past twenty years, the ratio of cash payments to check payments declined during most of this period but appears to have leveled off during the 1980's.

Since G is unobserved, Laurent attempted to obtain an estimate of G by employing Fisher's equation of exchange,[26]

6) $C(t) \cdot V_c(t) + D_d(t) \cdot V_d(t) = [PT](t),$

namely, that the sum of currency payments $[C(t).V_c(t)]$ and check payments $[D_d(t).V_d(t)]$ must equal total transactions [PT](t). Since the total volume of transactions [PT] is also unobserved, Laurent assumed that:

7) $\dfrac{[PT](t)}{[py](t)} = k_0 \, .$

where:

D_d = Checkable deposits.

V_d = Velocity of checkable deposits.

PT = Value of total transactions.

---

[25] The series on debits is obtained by splicing overlapping series for turnover of demand deposits for non financial centers. By focusing on turnover in non financial centers, the computed series partially eliminates the effect of changes in large financial transactions. The final debits series is obtained by multiplying the derived turnover series by demand deposits and adding debits to other checkable deposits.

[26] In principle, G could be empirically estimated from well designed engineering study that tracked note exchanges in a controlled environment that simulated real world conditions.

py = Value of current income or product [GNP].

k_0 =Constant.

Employing assumption (7) Laurent substituted nominal income as a proxy for total transactions in the equation of exchange (6). Combining equations (5) , (6) and (7) yields:

$$8) \quad py = \theta_1 . [\, U(t)\,] + \theta_2 \,[\, D(t) . Vd(t)\,],$$

where $\theta_1 = G/k_0$ and $\theta_2 = 1/k_0$. Equation (8) can be estimated by least squares regression to yield estimates of $(1/k_0)$ and hence of G.

Instead of estimating equation (8) by least squares methods, Laurent chose the analyze equation (9):

$$9) \quad py = \theta . [(U(t) + D(t) . Vd(t)],$$

Laurent selected that value of θ that maximized the correlation between the two sides of equation (9). He found that a value of $\theta = 129$ maximized the correlation. He then incorrectly interpreted the estimate of θ as being an estimate of G.[27] The only condition under which $\theta = G$ is that $k_0 = 1$.

As an alternative to Laurent's correlation procedure, equation (8) was estimated for the period 1891 - 1984. The resulting estimate of G was 87.1.[28]

Laurent's procedure for estimating G is questionable on both conceptual and empirical grounds. He incorrectly identified his estimate of θ as being an estimate of G, and he imposed the assumption (7) of a constant ratio of total transaction to income. The foregoing estimate of G employing equation (8) eliminates the first error, but even

---

[27] Laurent's correlation analysis examined alternative values of θ ranging from 0-∞. The reported correlation coefficients (p.38) reveal that for $\theta = 25$ the correlation coefficient was .993398 and for $\theta = 400$ it was .989662. The results indicate that Laurent was unable to obtain a robust value of θ.

[28] Equation (8) was first differenced and estimated with a first order ARMA process. The resulting estimated values of θ_1 and θ_2 were 3.03 and .035 respectively. Dotsey (1988:p.27) employs the Laurent method and obtains a value of 85 for G.

this result is still conditioned by the questionable reliability of assumption (7).[29]

IV. Extensions of the Simple Payments Model

Despite the limitations of Laurent's empirical analysis, his innovative conceptual framework for analyzing currency payments can be elaborated and combined with new survey information on cash usage habits of American households obtained by the Federal Reserve Board. The Federal Reserve Board's survey of currency usage in the U.S.[30] provides the first independent point estimate of the average velocity of currency. The survey found that the average turnover of currency for product expenditures in 1984 was approximately 50 turnovers per year and 55 turnovers per year for total currency expenditures. Since the Federal Reserve survey was specifically designed to obtain an estimate of currency velocity, its findings can be used to obtain alternative estimates of the total number of lifetime transactions [G] that notes of various denominations can sustain before being judged unfit for further circulation.

The simple currency payments model assumes that notes are printed by the Bureau of Engraving and Printing and subsequently issued by Federal Reserve Banks. Firms and individuals withdraw currency from banks to effect payments. Note recipients eventually redeposit the currency in financial institutions, which in turn, send notes to Federal Reserve Banks. Federal Reserve Banks, charged with the responsibility for maintaining the fitness standard of the outstanding currency supply, sort notes received from circulation and determine which notes are fit for recirculation and which notes require redemption (R). The redeemed notes are those, that upon inspection, are found to be sufficiently worn to warrant destruction. For given fitness standards and physical characteristics of currency notes, $[G_i(t)]$ represents the total number of lifetime

---

[29] The assumption of proportionality between transactions and income has been implicitly and explicitly utilized in monetary economics for more than half a century, and was rarely questioned prior to its use as a key assumption in the "transactions" method for estimating the size and growth of unrecorded and unreported income. Indeed, the assumption plays an important role in most studies of the demand for money that employ income rather than transactions as the scale variable in the money demand function. The past decade has witnessed an unprecedented degree of financial innovation that may have significantly raised the ratio of transactions to income. Due to the paucity of data on gross financial flows for most nations, it is currently difficult to determine the extent to which the ratio has risen as a result of the understatement of income resulting from growth in the "underground economy" or as the result of a vigorous increase in gross financial transactions. Regardless of how that issue is finally resolved, one is left with the uncomfortable conclusion that (7) is a tenuous assumption for inference in monetary economics. Unfortunately, abandoning this assumption calls into question much of the empirical literature on the demand for money that relies on the assumed proportionality to permit the substitution of income for transactions in the money demand function.

[30] Avery et. al. (1986); Avery et.al. (1987).

turnovers[31] effected by notes of the ith denomination before their destruction. The number of cumulative turnovers $[T^*_i(t)]$ performed by notes of denomination (i) by period (t) is shown in equation (2).

The number of turnovers per year $[T_i(t)]$ is then,

$$10) \quad T_i(t) = \frac{d}{dt} \left\{ G_i(t) \cdot \left\{ \int_0^t R^n_i(s) \, ds + \int_0^t \{ I^n_i(s) - R^n_i(s) \} \cdot \gamma_i(s) \, ds \right\} \right\}$$

Following Laurent in assuming that the average note in circulation has performed one half of its total lifetime turnovers, differentiation with respect to t and division by the number of notes yields the average yearly turnovers per note $Z_i(t)$:

$$11) \quad Z_i(t) = \frac{.5 \cdot G_i(t) \cdot [R^n_i(t) + I^n_i(t)] + \frac{\dot{G}_i(t)}{G_i(t)} \cdot T^*_i(t)}{C^n_i(t)}$$

where $C^n_i(t)$ represents the number of notes of the ith denomination in circulation at time t.

If $G_i(t)$ remains constant over time, the final term of the expression is zero, and we derive the denomination specific average turnover per note $[Z_i(t)]$ as,

$$12) \quad Z_i(t) = G_i(t) / L_i(t)$$

were $[L_i(t)]$ denotes the average lifetime of a note. Under the foregoing assumptions, the average lifetime of a note $[L_i(t)]$ is,

---

[31] $G_i(t)$ can be viewed as the mean of the distribution of lifetime payments.

13) $L_i(t) = C^n{}_i(t) / .5[R^n{}_i(t) + I^n{}_i(t)]$ [32]

Assuming that the physical characteristics of notes are identical for all denominations,[33] ie. that $G_i = G$ for all (i) the average number of turnovers per note $[Z(t)]$ of the entire currency stock is then:

14) $$\begin{aligned} Z(t) &= \Sigma\, x_i(t)\, Z_i(t) \\ &= G \cdot \Sigma\, x_i(t)\, /\, L_i(t) \\ &= G \cdot L_x(t) \end{aligned}$$

where,

$\quad x_i(t) = C^n{}_i(t) / \Sigma\, C^n{}_i(t)$

and $L_x(t)$ is the number share weighted average of the reciprocal of denomination specific lifetimes.

Correspondingly, the average velocity of the currency stock $V_c(t)$ is defined as;

15) $$\begin{aligned} V_c(t) &= \Sigma\, w_i(t)\, Z_i(t) \\ &= G \cdot \Sigma\, w_i(t)\, /\, L_i(t) \\ &= G \cdot L_w(t) \end{aligned}$$

where,

$\quad w_i(t) = C^n{}_i(t)\, D_i / \Sigma\, C^n{}_i(t)\, D_i$

and $L_w(t)$ is the value share weighted average of the reciprocal of denomination specific lifetimes.

Finally, given an independent estimate of the velocity of currency circulation, an estimate of G is obtained as;

16) $G(t) = V_c(t) / L_w(t)$

---

[32] This is the so called circulation formula for average note lifetime employed by Feige (1979) and Boeschoten and Fase (1984).

[33] In the U.S., notes of all denominations are printed on paper with the same technical specifications.

FIGURE 3
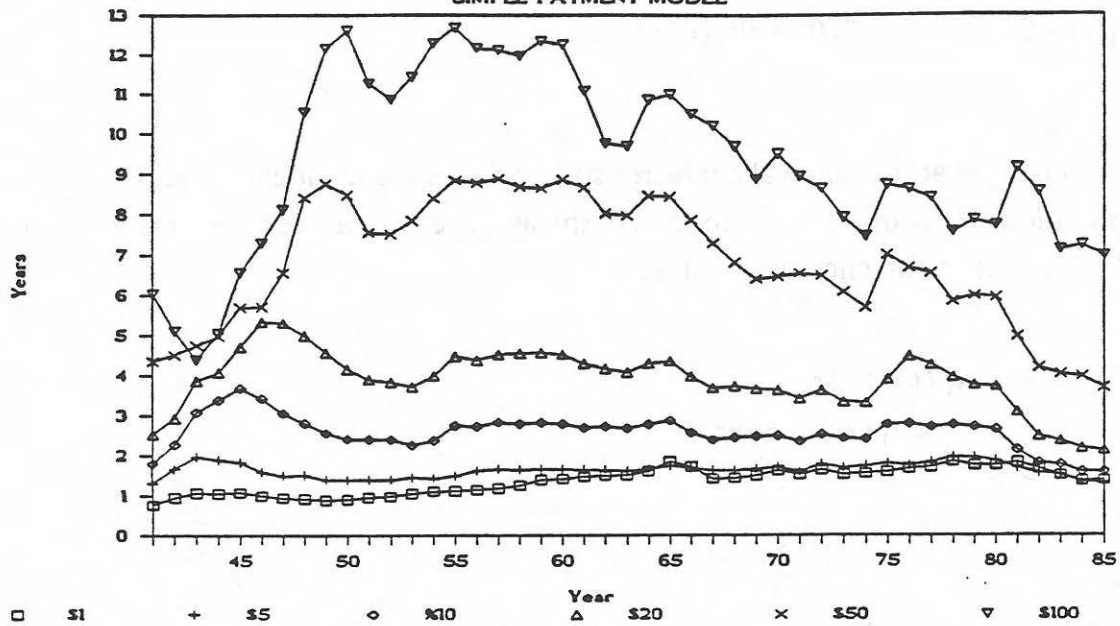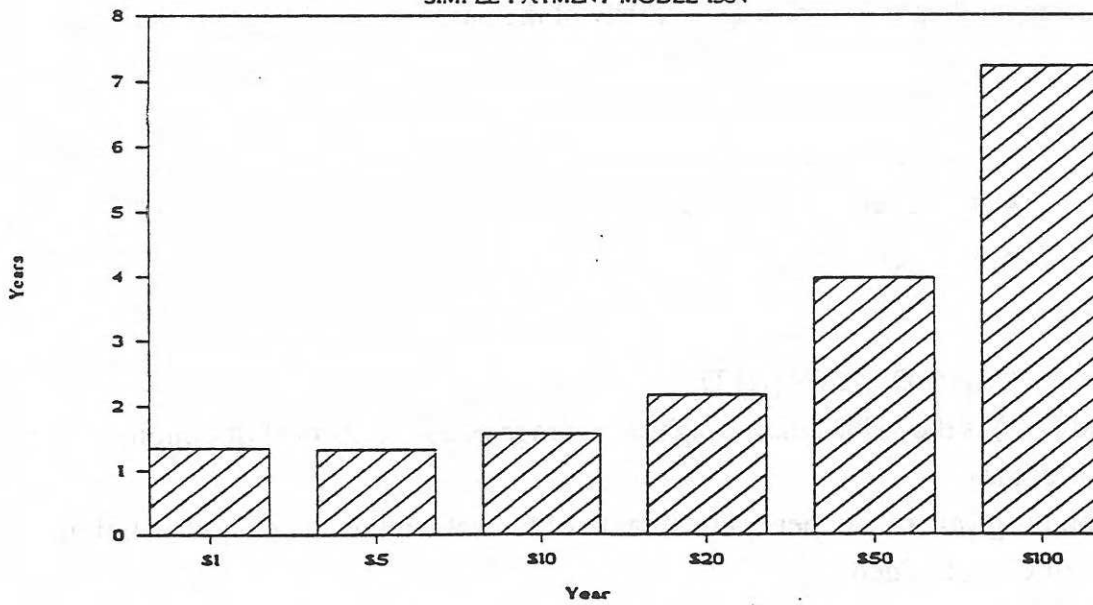
ESTIMATED AVERAGE LIFETIMES
SIMPLE PAYMENT MODEL

FIGURE 4

ESTIMATED AVERAGE LIFETIMES
SIMPLE PAYMENT MODEL 1984

V. Empirical Estimates Based on the Simple Payments Model.

Given time series data on redemptions and issues, the average lifetime of currency is estimated from the circulation formula given in Equation (13). This method of estimating average lifetimes assumes that outstanding bills have completed fifty percent of their lifetime transfers and that the Federal Reserve has successfully maintained its quality standard during the period of observation. Under these assumptions, one can interpret variations in the average lifetimes of notes as solely reflecting changes in the intensity of note usage. However, in the event of departures from the Federal Reserve's quality target standard this interpretation is no longer valid since the estimated values of the average lifetimes of notes depend upon observed rates of currency issue and redemption.

Figure (3) presents the denomination specific estimated average lifetimes of notes based on the simple circulation formula for determining average lifetimes. As expected, the smallest denomination notes have the shortest estimated lifetimes and conversely, the largest denomination notes have the longest average lifetimes. The inter- temporal pattern of the estimated average lifetime series reveal significant short term fluctuations that are inconsistent with the traditional view that payment habits are deeply ingrained and slow to change. One possible explanation of these unexpected fluctuations in the estimated average lifetimes is that currency quality standards themselves changed over time.[34]

Figure (4) displays the estimated values of denomination specific average lifetimes of notes for the year 1984. The smallest denomination notes have an estimated lifetime of approximately one and a half years whereas the $100 denomination notes have an average lifetime of about seven and a half years.

Equation (16) is employed to obtain estimates of the total number of lifetime transactions [G] that notes can sustain. $[L_w(t)]$ is calculated from the foregoing estimates of the average lifetimes of notes and a point estimate of $[V_c]$ for 1984 is obtained from the Federal Reserve's recent survey of currency usage. According to the Federal Reserve survey, the average velocity of currency was 50.4 turnovers per year. Assuming that this estimate of $[Vc(84)]$ applies to the entire circulating stock of currency, then given our estimate of $L_w(84)$, Equation [16] is used to obtain an estimate of [G]. The resulting estimate for the total number of lifetime turnovers that

---

[34] This possibility is further explored in the following section.

FIGURE 5

ESTIMATED TURNOVERS PER NOTE
SIMPLE PAYMENT MODEL 1984

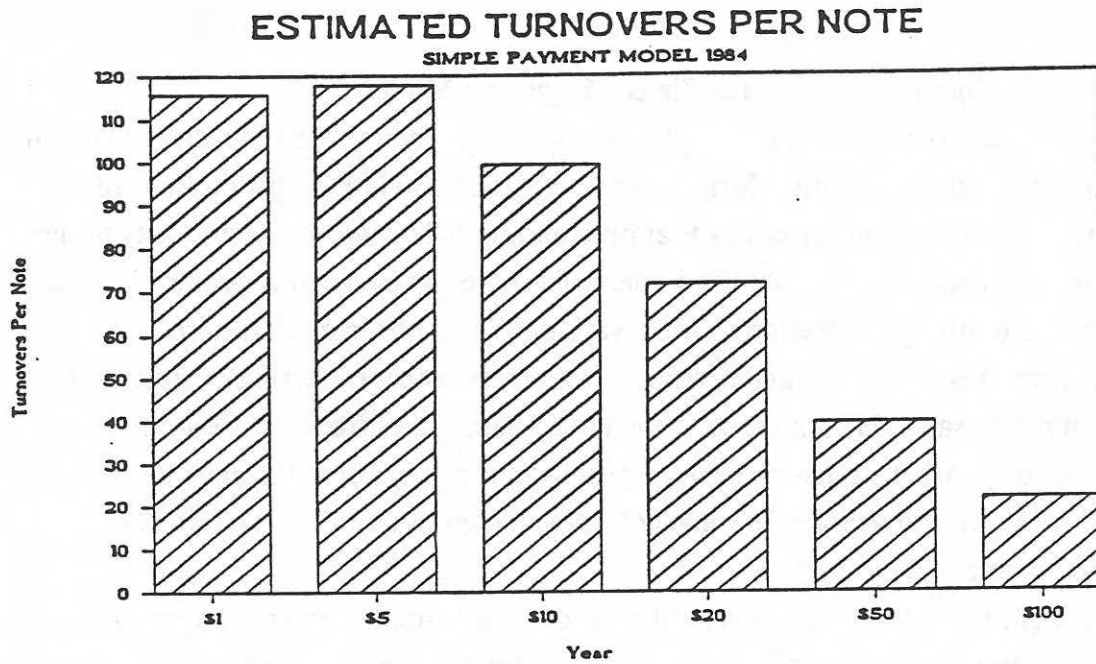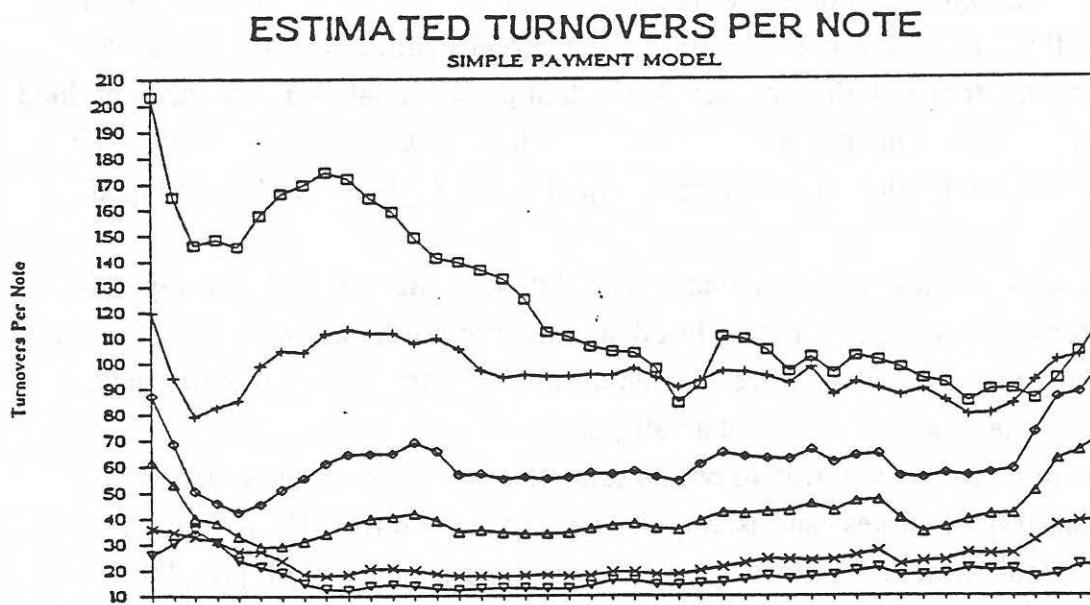FIGURE 6

ESTIMATED TURNOVERS PER NOTE
SIMPLE PAYMENT MODEL

the average currency note supported in 1984 was 156.3 turnovers.[35]

Given this estimate of G, equation (12) is employed to calculate the denomination specific turnovers per note. As displayed in Figure (5), in 1984 small denomination notes turned over approximately 115 times per year for the purchase of goods and services whereas $100 bills turned over approximately 21 times. Assuming no changes in the physical characteristics of notes over time, equation (12) can also be employed to estimate the time series of note turnover for earlier years. The time series estimates presented in Figure (6) suggest that since 1950, turnover of small denomination notes has declined whereas the turnover of larger denomination notes has increased.

Figure (7) presents estimates of the implied holding period for notes of different denominations for the year 1984. The estimates based on the simple payment model suggest that smaller denomination notes are held for approximately three days before being spent whereas $100 bills are held for approximately 17 days before being used to purchase goods and services.

Figure (8) displays the estimated average velocity of currency for the period 1941 - 1985 as derived from equation (15). The average velocity of currency appears to have fallen dramatically during the period 1941 - 1945. Thereafter, the average velocity of currency appears to have fluctuated between thirty five and fifty turnovers per year.

The foregoing estimates of average lifetimes and currency turnovers are conditional on the assumptions implicit in the simple payments model. Some of these assumptions may be violated in the institutional setting of the actual currency circulation system. In particular, the simple payments model assumes that the Federal Reserve rigorously maintains currency quality standards. Any departures from this standard would affect the observed volume of currency issues and redemptions and hence affect the currency velocity estimates. As revealed by Figure (8) the estimated series for currency velocity displays some substantial short run variations which may be attributable to departures from currency quality standards rather than to changes in the intensity of currency usage. For example, a decline in currency fitness standards would reduce the observed volume of redemptions and new issues as older unfit notes are recirculated rather than redeemed. The effect of such a relaxation of quality standards is to reduce the

---

[35] Since the Federal Reserve's estimates of currency velocity refer to the number of times the currency stock turns over for the purchase of final goods and services, the resulting estimate of G should be interpreted as the total number of times a currency note is used to purchase final goods and services in its lifetime.

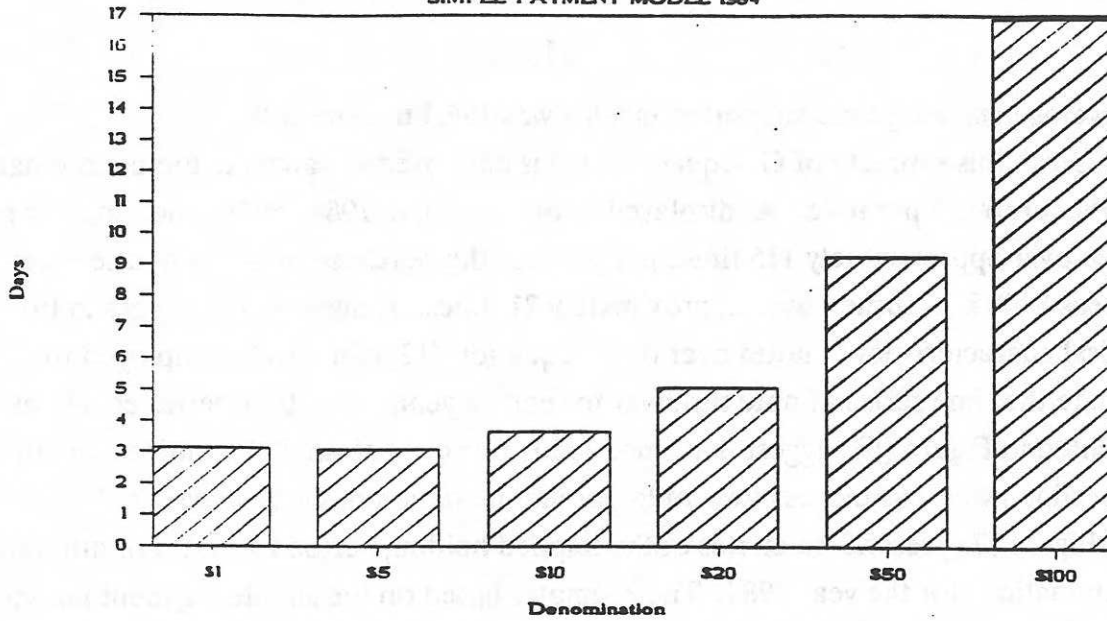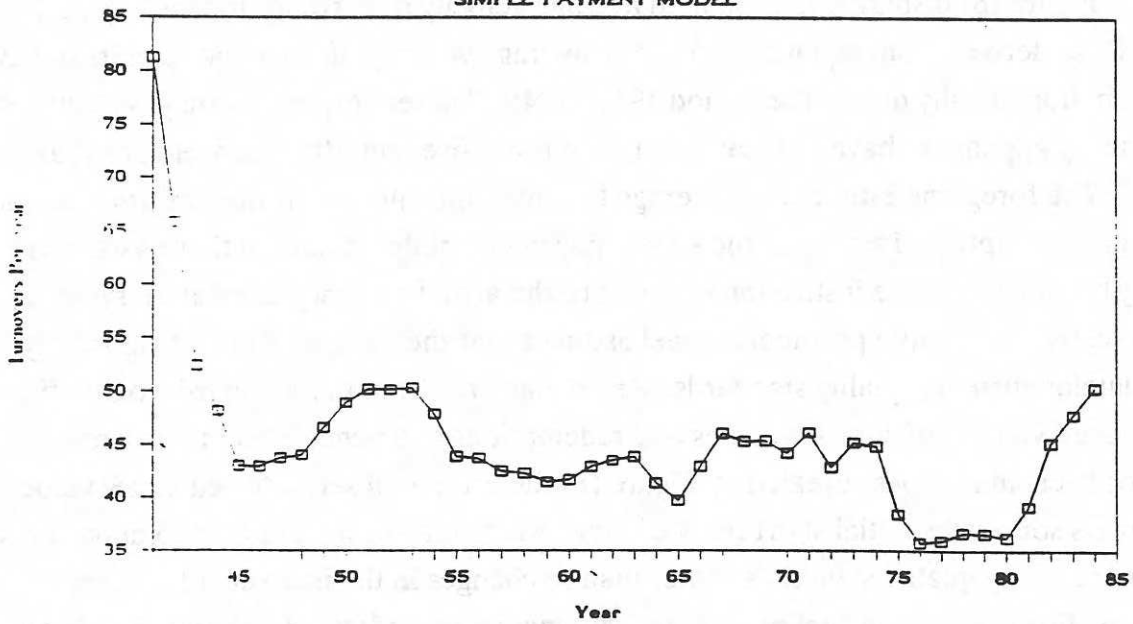FIGURE 7

ESTIMATED HOLDING PERIOD
SIMPLE PAYMENT MODEL 1984

FIGURE 8

ESTIMATED AVERAGE VELOCITY
SIMPLE PAYMENT MODEL

estimated average velocity of currency even if currency usage remains the same.[36] Conversely, an effort to upgrade the quality of circulating currency requires the Federal Reserve to redeem and issue a higher than normal number of notes. Such temporary increases in issues and redemptions shortens the effective lifetime of notes and hence increases the estimated velocity of circulation.[37] In order to take explicit account of possible changes in fitness standards we turn to an alternative model of the currency circulation system which relaxes some of the restrictive assumptions of the simple payments model.

VI. A Demographic Framework For The Analysis of Currency.

Demographic concepts typically applied to the study of human populations, can be usefully adapted to the study of currency populations in order to develop an alternative procedure for estimating the average lifetime of notes of specific denominations.

The total population of outstanding currency notes consists of the aggregate of denomination specific sub-populations. Notes are born when they are issued and die when they are redeemed from circulation. The change in the number (population) of notes of each denomination of currency from one year to the next is simply the difference between annual issues and redemptions. Since the government is obligated to supply the amount of currency demanded by the public, the number of gross currency issues (births) in any year is determined by the growth in the demand for notes from one year to the next plus an amount required to replace notes that have been redeemed (died) during the year.

The average lifetime of notes in circulation is determined by the intensity with which notes are used and by the institutional arrangements of currency management. The Federal Reserve is assumed to be the omnipotent force that determines when a note must die. Its criterion for death is that a note will die once it has become sufficiently soiled to be regarded as unfit for further circulation. Assuming that notes are soiled in direct proportion to the number of turnovers performed, the FED's monitoring of notes

---

[36] A search of the archives of the Federal Reserve Bank of New York produced a number of letters and circulares suggesting that currency quality standards were significantly relaxed during World War II in an effort to conserve material and labor in the printing of currency.

[37] Interviews with officials of the Federal Reserve System suggest that during the period 1975 - 1979 currency fitness standards were significantly relaxed as some district banks attempted to minimize expenses related to currency issue. In response to widespread complaints concerning the apparent deterioration of currency quality, a committee was established to assess quality standards and to insure a system wide return to earlier fitness standards. These higher standards were implemented during the 1980's.

received from circulation ensures that on average, notes that have performed the requisite total of G lifetime turnovers will be permanently retired from circulation. The Federal Reserve is not however omniscient, since it can not observe the exact number of turnovers that each note has performed. Moreover, the FED is not always consistent in uniformly applying its criterion for a note's death sentence.

The assumptions of the demographic model of currency differ from the simple payments model of currency in several important respects. The simple payment model assumes that the Federal Reserve strictly adheres to the currency fitness standard in every time period whereas the demographic model allows the Federal Reserve to depart from its fitness standard. The assumption of a strict adherence to the quality standard is replaced by the weaker and more realistic assumption that the expected value of departures from the quality standard are zero. Let [Ga] represent the actual total number of turnovers that a note sustains during its lifetime, and let [G] represent the total number of lifetime turnovers regarded by the Federal Reserve as making the currency sufficiently soiled to be retired from circulation. The simple payment model assumes that,

17) $G_a = G$

that is, all notes are retired when they have sustained the number of lifetime transfers established by the Federal Reserve's quality standard.

Departures from the Federal Reserve quality standard are denoted by (μ_t) where;

18) $\mu_t = (G - G_a)$ and $E(\mu_t) = 0$.

When the Federal Reserve strictly adheres to its fitness standards it will retire those notes that have sustained a total of $[G = G_a]$ actual turnovers. A relaxation of fitness standards implies that $[G_a > G]$ and conversely, tightening of fitness standards implies that $[G_a < G]$. The first modification of the simple payments model is to permit departures from the Federal Reserve's quality standard.

The second modification of the simple payments model involves the method employed to obtain estimates of the average lifetime of a currency note. In demographic theory, the expected lifetime of an individual is obtained by constructing a life table which depends upon age specific rates of mortality. Age specific mortality rates are available for many human populations since data are routinely collected on the number of annual

deaths by age of the deceased. Similar age specific mortality rates for currency could be estimated by sampling the population of notes that are selected for redemption and determining from the serial numbers on the notes, the date of the note's birth. The difference between the date of birth and the date of death determines the note's age. Given knowledge of the age distribution of redeemed notes, one could calculate the average lifetime of notes and the age specific mortality rates. In the absence of such age specific mortality rates derived from samples of notes selected for redemption, it is necessary to obtain indirect estimates of age specific mortality rates from the temporal history of issues and redemptions.

In the simple payments model, the average lifetime of currency is derived by the circulation formula [Equation 13]. In the demographic model, the average lifetime of currency is determined from estimates of age specific mortality rates. The number of notes redeemed in any year depends upon the past history of note issue. If $R^n_i(t)$ represents the number of notes of a particular denomination redeemed during year (t) and $I^n_i(t)$ represents the number of notes of that denomination issued during period t, then it is possible to express this year's redemptions as a polynomial distributed lag [PDL] function of past issues such that;

19) $R^n_i(t) = \beta[w_0 \, I^n_i(t) + w_1 \, I^n_i(t-1) + w_2 \, I^n_i(t-2) + ...] + \mu(t)$

where we assume that,

20) $w_j = c_0 + c_1 j + c_2 j^2 + c_3 j^3 + ..$

and that the c_k's are non negative and sum to unity. For example, if Equation (19) takes the form of a second degree polynomial with a five period lag, then,

21) $R^n_i(t) = \beta[w_0 \, I^n_i(t) + w_1 \, I^n_i(t-1) + ... + w_4 \, I^n_i(t-4)] + \mu(t)$

and,

22) $w_j = c_0 + c_1 j + c_2 j^2$

The coefficients (βw_j) of Equation (19) represent the fraction of notes previously issued that are redeemed during the year (t). Notes that are redeemed in the same year

in which they are issued are assumed to have lived for half a year. Similarly, notes that are redeemed in the year after they have been issued are assumed to have lived an average of 1.5 years. Each βw_j thus represents the probability of death for notes of various ages. The average lifetime (L_i) of a particular denomination can then be calculated as:

23) $L_i = \beta w_0(.5) + \beta w_1(1.5) + \beta w_2(2.5) + \ldots\ldots$

Notes that are used intensively, ie. those that turn over frequently, will have a short average lifetime, whereas infrequently used notes will have longer average lifetimes. Following Fisher (1911) we assume that payment habits are deeply ingrained and that notes of a particular denomination exhibit a stable inter-temporal pattern of turnovers. Small denomination notes are expected to turn over frequently whereas larger denomination notes will turnover more slowly. Changes in payment habits are manifested in changes in the overall velocity of currency circulation. When payment patterns change, individuals are assumed to alter the denomination structure of their note portfolio rather than change the number of turnovers that notes of a particular denomination perform. Under these assumptions, we would expect the polynomial distributed lag function (equation 19) for small denomination bills to have a relatively small number of lags, whereas large denomination bills may require long lags before all notes of a particular issue cohort are redeemed. If the number of turnovers that notes of a particular denomination sustains each year is invariant over time, then the age specific mortality rates for that denomination can be obtained from estimates of equation (19).

If currency issues and redemptions are measured without error, the estimated residual error term $\mu(t)$ in equation (19) can be interpreted as an index of estimated departures from the Federal Reserve's quality standards. Because such departures have an expected value of zero, equation (19) can be estimated by ordinary least squares methods. An observed residual $\mu(t) > 0$ suggests that the FED has temporarily increased its quality standard. In this case, the Federal Reserve redeems a larger number of notes during the year (t) than would be dictated by its normal criterion for destroying soiled notes. Similarly, a relaxation of quality standards [$\mu(t) < 0$] implies that the FED will redeem a smaller number of notes than would be dictated purely by note use.

The empirical problem is to estimate polynomial distributed lag functions for each note denomination in order to obtain estimates of the probability that a note will be

redeemed after living a certain number of years. The residuals from the polynomial distributed lag model are interpreted as estimates of the departures from the FED's long range quality standards.

Given the foregoing framework, the mortality of notes destroyed in any year depends upon two factors:

a) The condition of notes in circulation as determined by note usage, and

b) The strictness with which the FED adheres to its quality sorting standards.

The condition of notes of any particular denomination in circulation is determined by the unobserved number of turnovers $[Z_i]$ that such notes have sustained. If payments habits relating to a particular denomination of notes are invariant over time, changes in the aggregate average velocity of currency will be reflected in a shifting pattern of the denomination structure of circulating notes. For example, if the public chooses to reduce the intensity of overall currency usage, that is, to reduce the number of times it turns over the currency stock in a given year, it will shift its portfolio of notes toward higher denomination notes. Less frequent trips to the store suggest that the average size of each cash transaction will increase, and hence it is more efficient to pay for these larger average purchases with larger denomination notes.

The average lifetime $[L_i]$ of notes of denomination (i) can be estimated from equation (23). The number of turnovers per note $[Z_i]$ sustained by the ith denomination is then,

24) $Z_i = G / L_i$

Polynomial distributed lag [PDL] functions represented by equation (19) were estimated for each denomination for the period 1940 - 1974. The particular form of the polynomial function estimated for each denomination is displayed in Table 1.

TABLE 1

Denomination	Length of Lag	Order of Polynomial
$1	5	2
$5	5	2
$10	10	2
$20	10	2
$50	15	3
$100	25	3

FIGURE 9
AGE SPECIFIC DEATH RATE
PDL MODEL - SMALL DENOMINATIONS

FIGURE 10
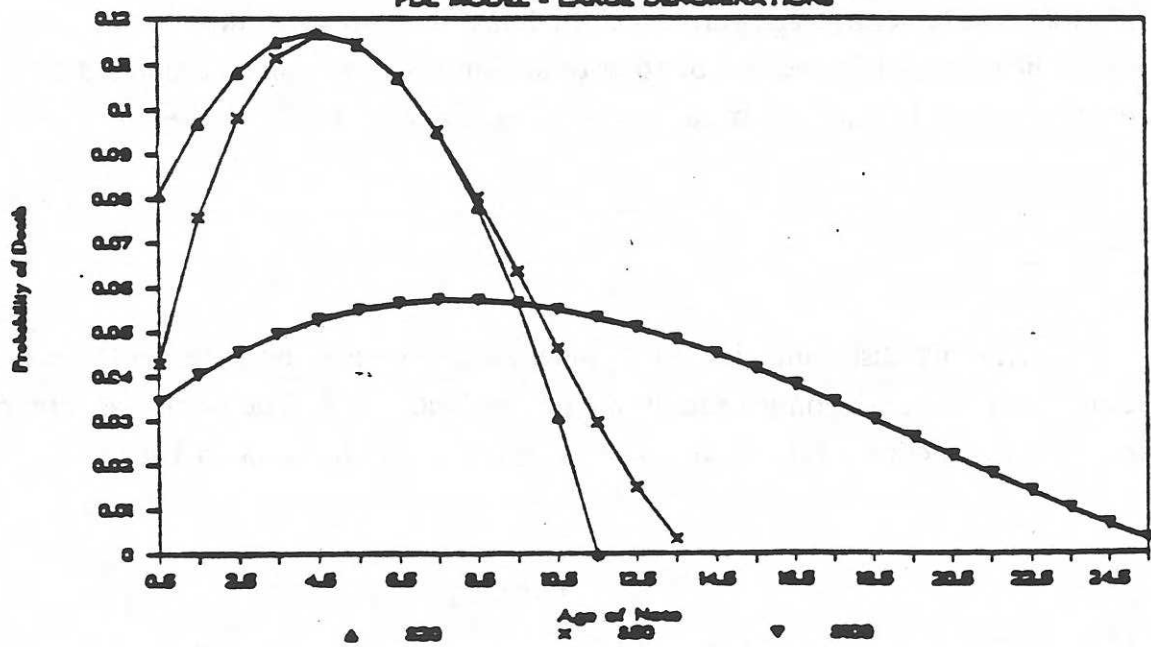## AGE SPECIFIC DEATH RATE
PDL MODEL - LARGE DENOMINATIONS
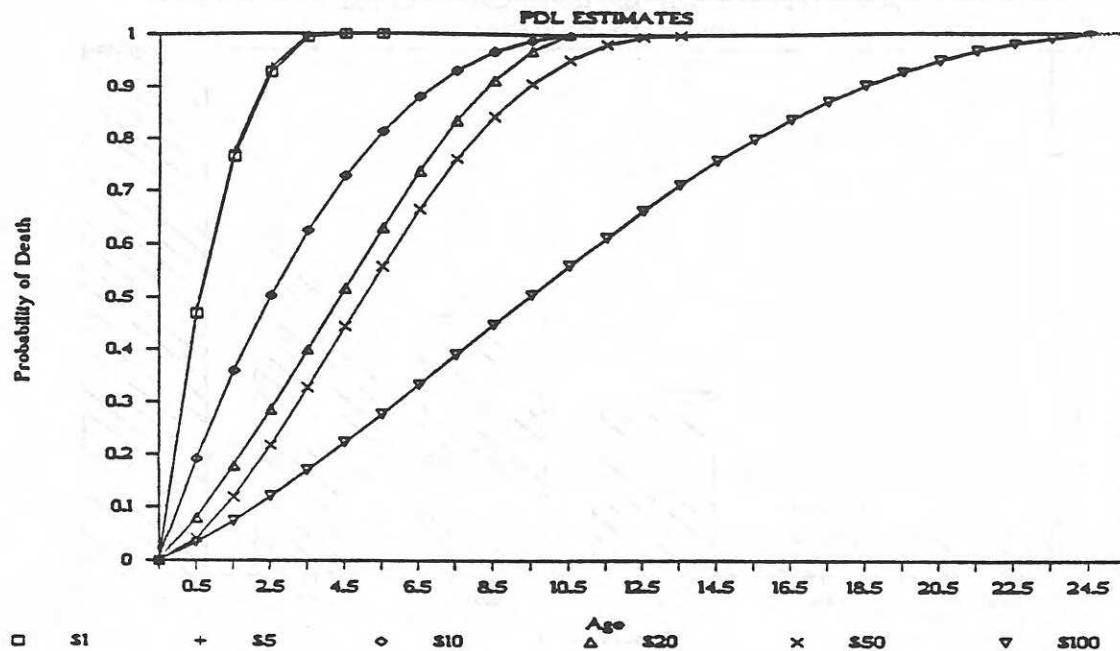
FIGURE 11

CUMULATIVE AGE SPECIFIC MORTALITY RATES
PDL ESTIMATES

FIGURE 12

PROBABILITY OF SURVIVING TO AGE X
PDL ESTIMATES

FIGURE 13

AVERAGE LIFETIME OF NOTES
PDL MODEL ESTIMATES

FIGURE 14

ESTIMATED AVERAGE LIFETIME OF CURRENCY
SIMPLE PAYMENT AND PDL MODEL ESTIMATES

□ Simple Payment   + PDL Model
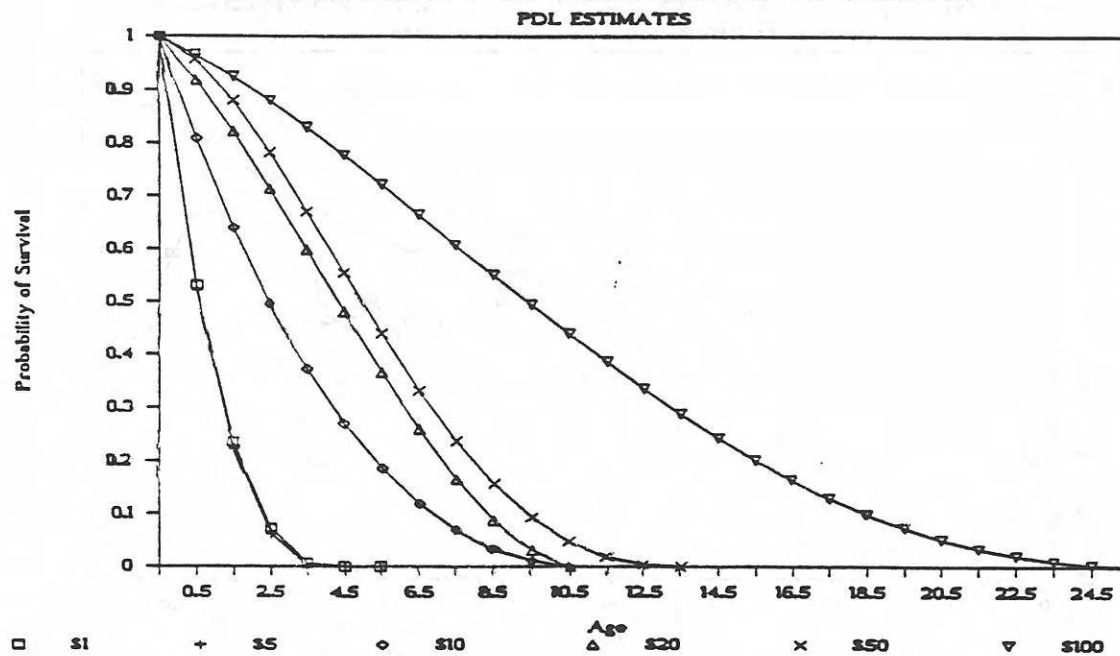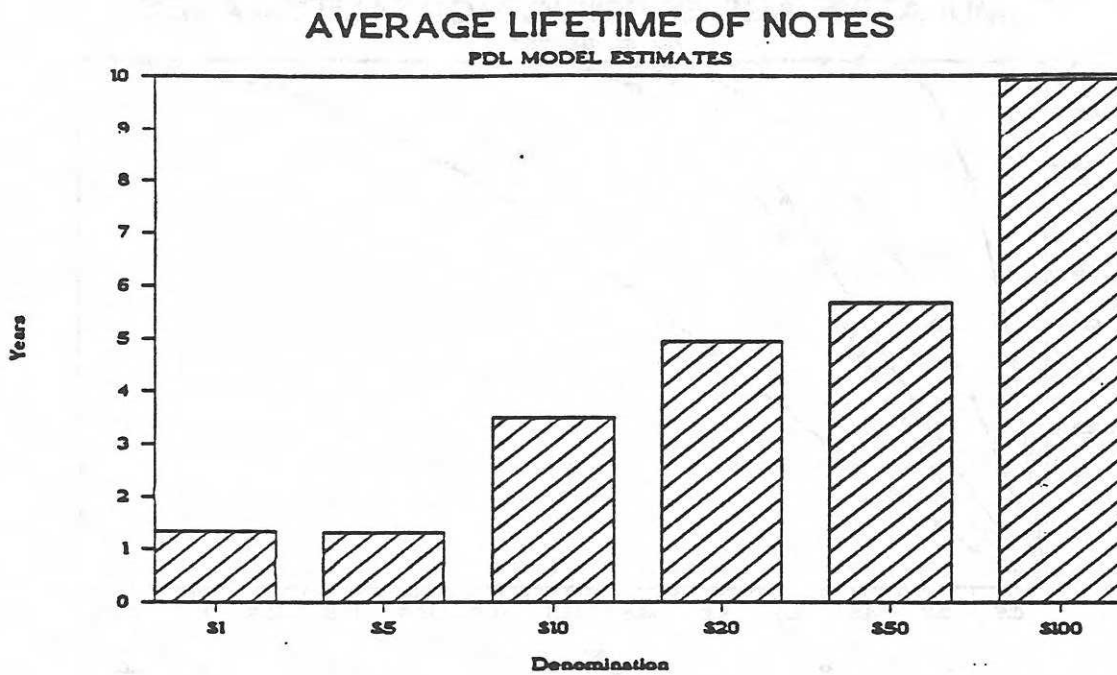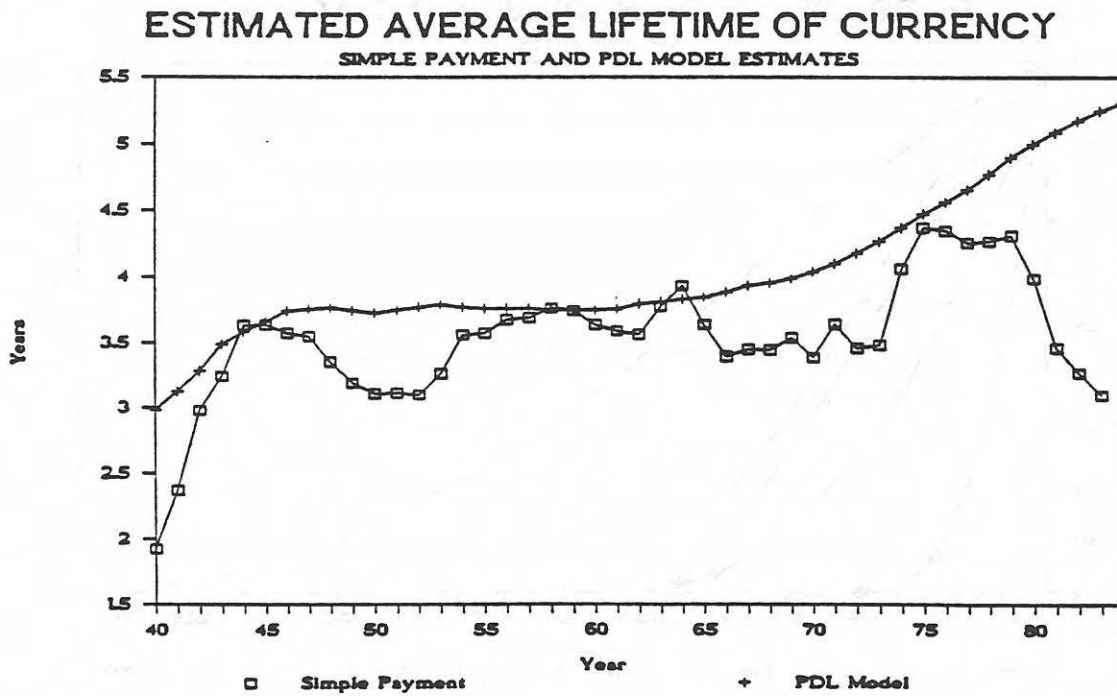
The normalized age specific death rates representing the estimated lag coefficients of the polynomial distributed lag function are displayed in Figures (9) and (10) for small and large denomination bills respectively. The estimated probabilities suggest that $1 and $5 notes are fully redeemed before reaching an age of five and a half years, with the highest death probability occurring at six months. As note denomination size increases, so does the total number of years in circulation. $20 and $50 denominations have similar mortality patterns and these denominations typically remain in circulation for eleven to thirteen years whereas $100 bills remain in circulation for as long as twenty five years.

Figure (11) displays the cumulative age specific mortality rates derived from the estimated PDL model and Figure (12) displays the corresponding estimated survival probabilities. Figure (13) displays the denomination specific estimated average lifetimes derived from equation (23). On the basis of the PDL model, the smallest denominations have an estimated average lifetime of approximately 1.3 years whereas the largest denomination note has an average expected lifetime of 9.9 years. Figure (14) displays the estimated average lifetime of the entire currency stock for both the simple payment model and the PDL model. These average lifetime estimates of the currency stock are derived as a number weighted average of the lifetimes of individual denominations. The relatively short lifetime of the entire currency stock is due to the fact that short lived small denomination notes are given large weights as they comprise a large fraction of the notes in circulation. The PDL estimates suggest that the average lifetime of the currency stock remained relatively constant between 1946 - 1960 and then began to increase smoothly. The simple payment method on the other hand suggests considerable variation in the temporal pattern of average lifetimes. The estimated average lifetimes from the polynomial distributed lag model are in general longer than those estimated by the simple circulation formula.

The polynomial distributed lag equations estimated for each denomination for the period 1940 - 1974, were employed to obtain post sample forecasts of redemptions. The residuals from the estimated equations are displayed in Figures (15) and (16) for small and large denomination notes respectively. These calculated residuals are interpreted as estimates of departures from Federal Reserve's long term quality target standards. Figures (15 - 16) suggest that currency quality standards were relaxed during the second world war and tightened in the immediate post war period. The most significant recent changes in quality standards appear to have taken place during the past decade. According to the estimates of quality departures presented in Figures (15) and (16), currency quality appears to have significantly deteriorated during the last half of the

FIGURE 15

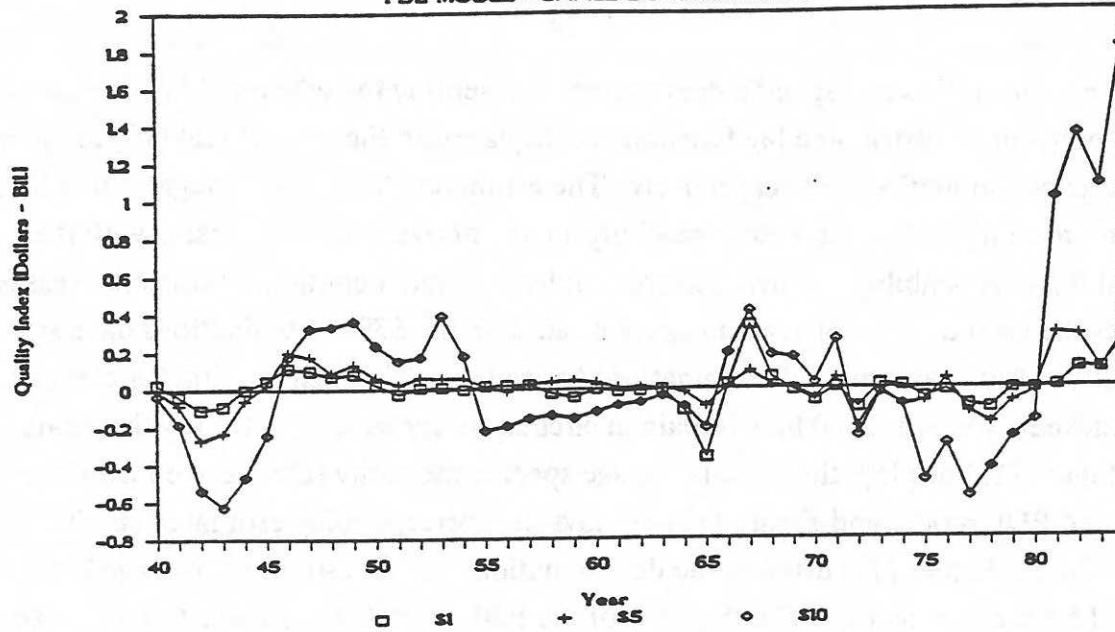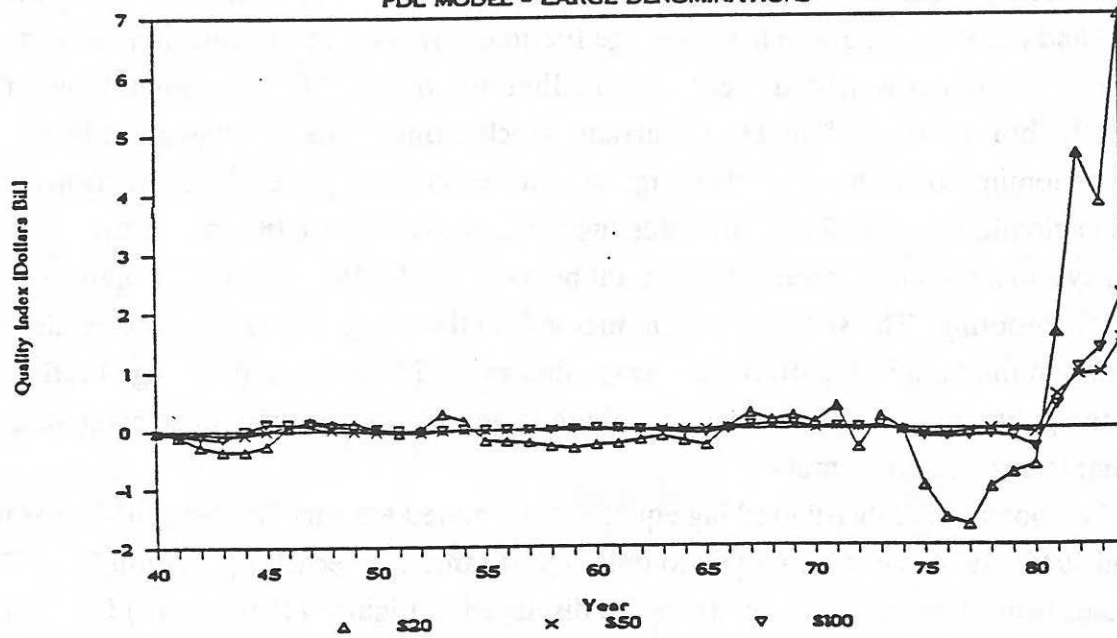# ESTIMATED QUALITY INDEX
PDL MODEL – SMALL DENOMINATIONS

FIGURE 16

ESTIMATED QUALITY INDEX
PDL MODEL – LARGE DENOMINATIONS

1970's and has improved significantly during the first half of the 1980's.[36] The pattern of forecast errors from the PDL model closely conform to institutional information concerning changes in effective quality standards. In response to widespread complaints of currency quality deterioration during the latter half of the 1970's, the Board of Governors made a concerted effort to restore the quality of circulating currency. To this end, the Board of Governors significantly raised quality standards during the first half of the 1980's. As a result of this policy, redemptions rates rose well above the normal levels dictated by currency usage. The fact that virtually all currency sorting is now undertaken with high speed sorting machines means that a systematic policy of maintaining system wide currency standards will in future eliminate the dramatic fluctuations in currency standards that appear to have characterized the last decade.

VII. Estimates of Total Lifetime Turnovers, Note Turnovers and Velocity

Given the estimates of the average lifetime of notes of each denomination we now wish to estimate the total number of lifetime turnovers (G) that notes of given physical characteristics can sustain. Since notes of all denominations are printed on the same paper and have the same physical wearing characteristics, it is assumed that (G) is the same for all denominations. As shown in equation (24), (G) is simply the product of the annual number of times notes of a particular denomination turn over and the average lifetime of the denomination. Since the denomination specific turnover of notes is itself unknown, we require some additional source of information in order to calculate (G). The only available source of information is the estimate of currency velocity provided by the Federal Reserve Board's currency usage survey. The Survey estimate of currency velocity $[V_c]$ differs from $[Z_i]$ in two significant ways. First, the estimate of $[Vc]$ is an aggregate measure comprising all currency denominations rather than a denomination specific measure. Second, it is a conventional estimate of currency velocity, namely, it measures the number of times per year the dollar value of the currency stock turns over rather than the number of times per year the average note turns over.

The number of turnovers per note in the currency population $[Z(t)]$ depends upon

---

[36] Information from Federal Reserve officials suggests that during the late 1970's the Board of Governors received an increasing number of complaints of currency deterioration. Investigation revealed that district banks had undertaken efforts to minimize the expense of new currency during this period. The newly introduced REI sorting machines were often recalibrated to lower currency standards. The effect of these actions was to permit soiled notes to reenter the currency circulation system. In response to this finding the Board of Governors established a system wide Currency Quality Committee charged with the responsibility of improving the quality of circulating currency and reestablishing uniform currency standards throughout the system.
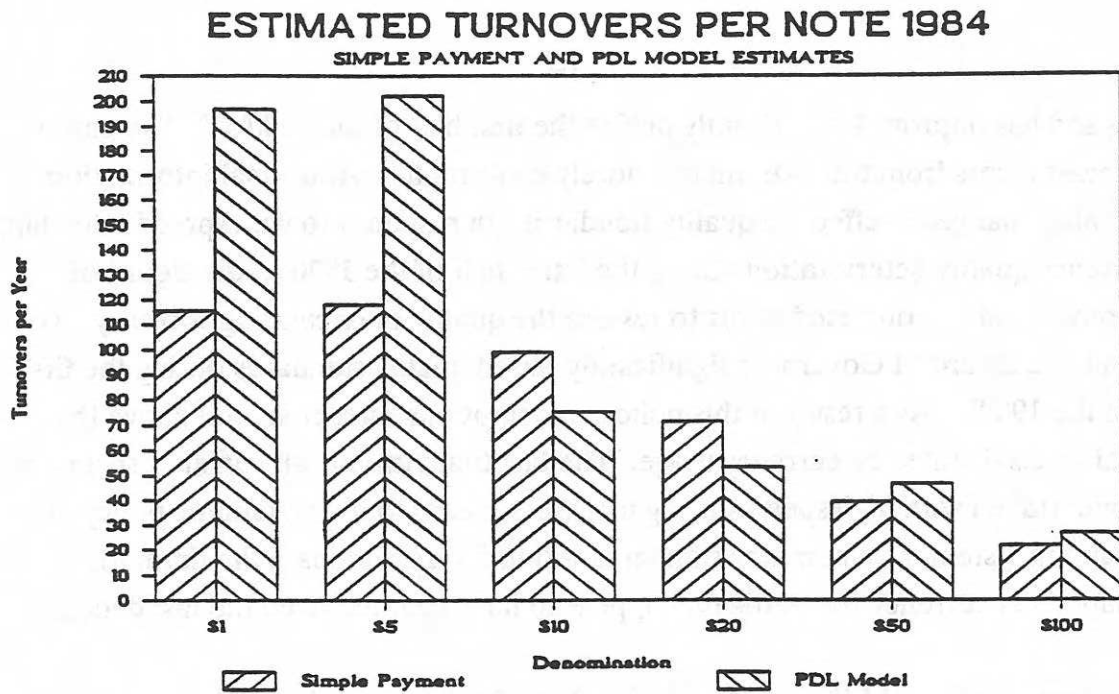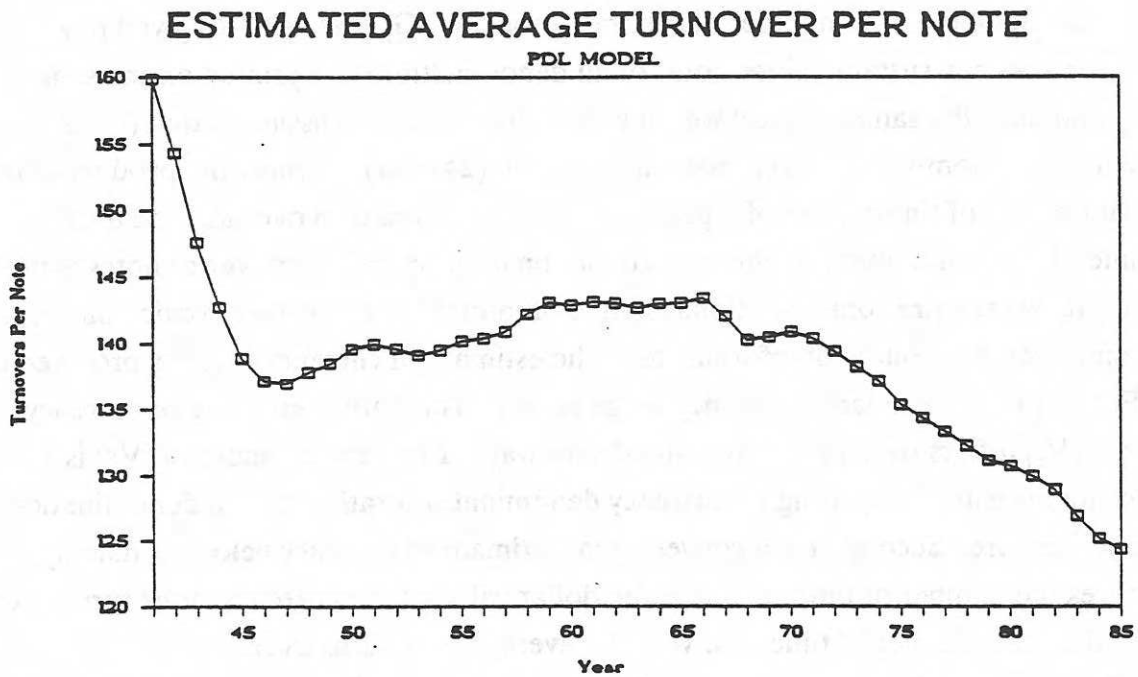
FIGURE 17

ESTIMATED TURNOVERS PER NOTE 1984
SIMPLE PAYMENT AND PDL MODEL ESTIMATES

Simple Payment   PDL Model

FIGURE 18

ESTIMATED AVERAGE TURNOVER PER NOTE
PDL MODEL

both the total number of lifetime turnovers that each note can sustain [G] and upon number share weighted average of the reciprocal of note lifetimes [$L_x(t)$].

25) $Z(t) = G \cdot L_x(t)$

Similarly, the total number of times the average dollar of currency turns over [$Vc(t)$] is;

26) $Vc(t) = G \cdot L_w(t)$

where [$L_w(t)$] is the value share weighted average of the reciprocal of note lifetimes. $L_w(t)$ is calculated from the PDL model estimates of the denomination specific note lifetimes presented in Figure (13). Given the 1984 Federal Reserve survey estimate of Vc, equation (26) is employed to obtain an estimate of G = 264. [37] This estimate of total lifetime turnovers is used in equation (24) to obtain estimates of the turnovers per note sustained during the year 1984. Figure (17) displays the estimates of denomination specific turnovers per note [Z_i] obtained from equation (24). The estimates suggest that the smallest denomination notes turnover over approximately 200 times per year compared with the estimate of 115 obtained from the simple payment model. The $100 denomination note turns over about 27 times per year compared with an estimated turnover of 21 times per year from the simple payment model.

Figure (18) displays the estimated average turnover per note [$Z(t)$] of the entire currency stock for the period 1941 - 1986. The average turnover per note appears to have declined during the second world war as individuals shifted their currency portfolios toward higher denomination notes. Thereafter, note turnover rose gradually and reached a peak in 1966. This rise represents a portfolio denomination shift away from the high denomination notes accumulated during the war years.[38] Turnovers per note have

---

[37] The Federal Reserve Board's estimate of currency turnover is based on a sample of U.S households. The calculated value of G obtained by using this estimate of turnover assumes that U.S. households hold the proportion of denomination specific shares reflected in $L_w(84)$. The Federal Reserve's survey results suggest that the specific currency denominations requested by respondents conforms closely to the existing denomination distribution of currency in circulation.

[38] The slight discontinuity in the series observed during the period 1964 - 1967 is probably due to the Act of Congress approved on June 4, 1963 which amended Section 16 of the Federal Reserve Act so as to authorize the issuance of $1 and $2 Federal Reserve notes. The purpose of this legislation was to make monetary silver available for coinage by substituting Federal Reserve notes for silver certificates. The
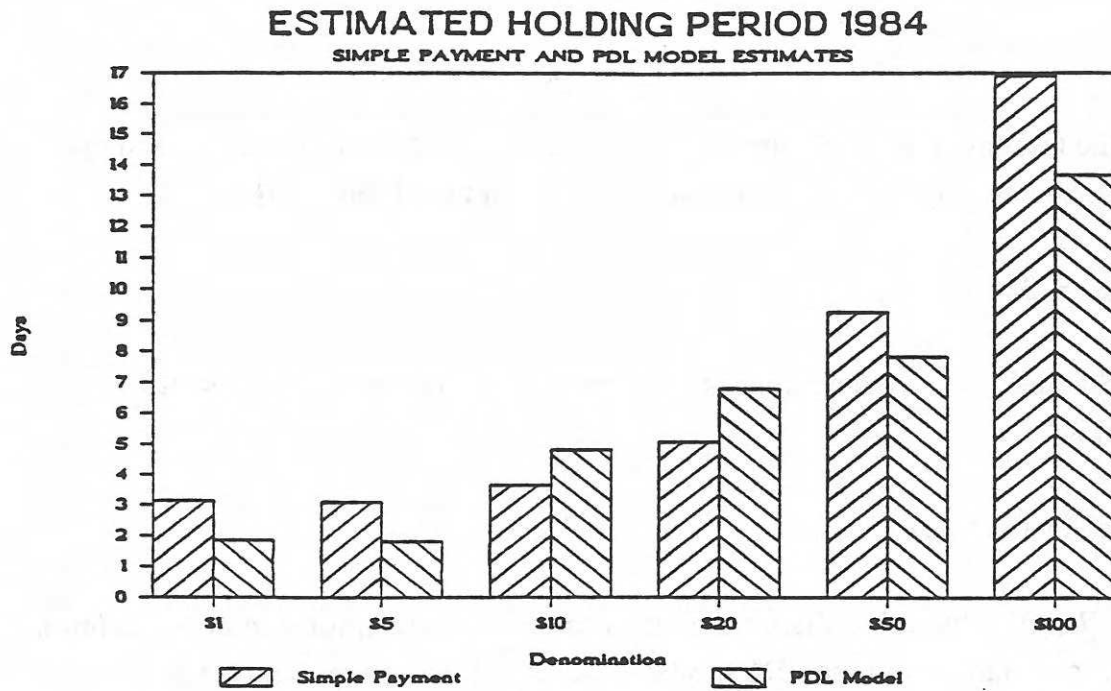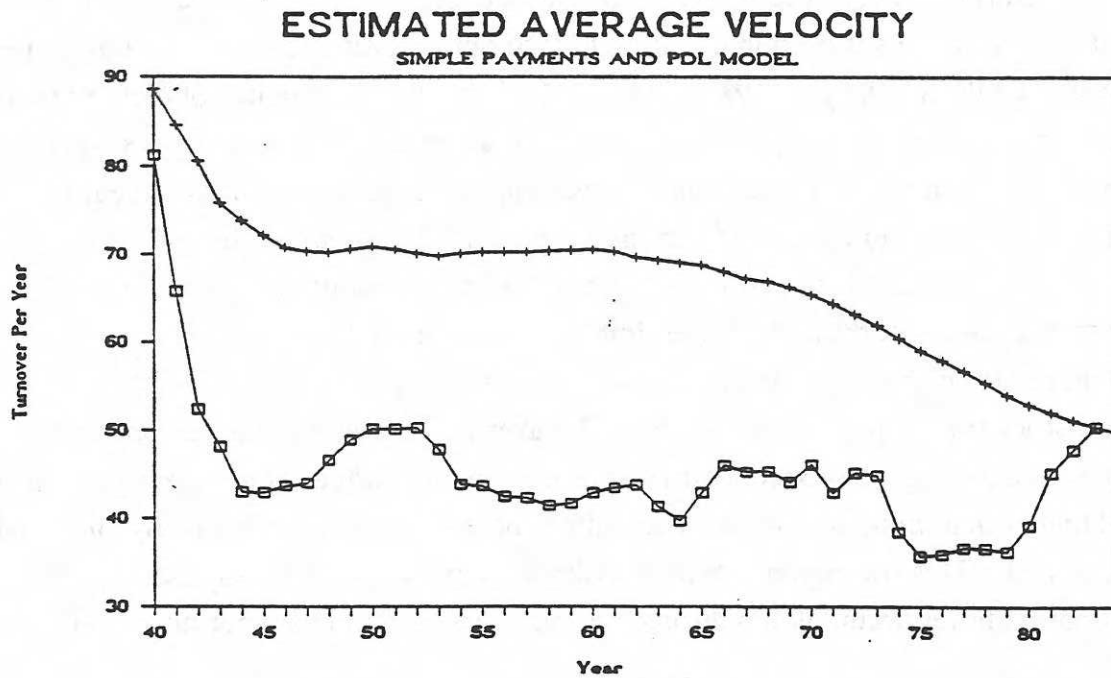
FIGURE 19

ESTIMATED HOLDING PERIOD 1984
SIMPLE PAYMENT AND PDL MODEL ESTIMATES

Days

Denomination

Simple Payment   PDL Model

FIGURE 20

ESTIMATED AVERAGE VELOCITY
SIMPLE PAYMENTS AND PDL MODEL

Turnover Per Year

Year

declined dramatically since the early 1970's reflecting a shift in currency portfolios toward large denomination bills.

Figure (19) compares estimates of the denomination specific holding periods obtained the simple payments model and the PDL model. The PDL estimates suggest that the smallest denomination notes are held for less than two days before being used to purchase goods and services whereas the $100 notes are held for approximately two weeks between expenditures.

Figure (20) displays the average velocity of the currency stock as estimated by the simple payment model and the PDL model. The PDL estimates suggest that currency velocity changes very gradually, showing a decline during the second world war and again during the past twenty years. The simple payments estimates are below those derived from the demographic model [PDL] and exhibit a pattern of cyclical fluctuations. The difference between the two series is due to the different assumptions employed by the two methods. The simple payments method assumes that the Federal Reserve maintains fixed currency quality standards and therefore attributes all fluctuations in issues and redemptions to changes in payment behavior, whereas the demographic model takes account of fluctuations in quality standards. The resulting estimates of the velocity of currency are employed to obtain provisional estimates of cash payments.

VIII. Cash Payments and the "Currency Enigma"

An initial estimate of the total volume of cash payments made with U.S. currency is obtained as the product of estimated currency velocity and the amount of currency in circulation. Figure (21) displays two provisional estimates of total cash payments based on the simple payments model and the PDL model. Based on the PDL model estimates of average velocity in 1985, the total stock of circulating currency could have sustained cash payments in excess of 8.7 trillion dollars. This huge volume of estimated cash payments gives rise to what may be described as the "currency enigma". The "currency enigma" subsumes three interrelated economic anomalies: 1) the problem of "missing currency"; 2) the problem of "missing payments"; and 3) the problem of "missing expenditures and income".

---

unusually large number of redemptions of old $1 bills by the Federal Reserve and the subsequent issues of new notes during this period is most likely responsible for the discontinuity in the estimated note turnover series.

A. The problem of "missing currency".

The problem of "missing currency" arises from the fact that a major portion of the U.S. currency supply does not appear to be held by U.S. households. This observation was most recently brought to light by the Federal Reserve Board's survey of currency usage by the adult population of the U.S. Admitted holdings of U.S. currency by the adult population accounted for only 10 - 12% of the nation's aggregate stock of currency in circulation.[39] Of the $178 billion of U.S. currency in circulation in 1986, less than $20 billion was estimated to be held by adult U.S. residents. The Federal Reserve survey concluded:

> "that more than 88 percent of the U.S. currency stock outside the vaults of depository institutions was held - apart from some that may be lost and unaccounted for - by other agents such as business enterprises, persons in other countries, and persons less than 18 years of age. Because children or legitimate businesses are not likely to have held more cash in the aggregate than all adults, this finding suggests that a large percentage of the U.S. currency stock was held for purposes not directly related to measured domestic activity." [40]

The problem of "missing currency" is apparently not of recent origin. A 1944 survey undertaken by the Department of Agriculture for the Board of Governors of the Federal Reserve System found that U.S. households held only 12.8% of the outstanding U.S. currency stock.[41] The Federal Reserve asserts the " while inconclusive, the evidence suggests that the percentage of cash unaccounted for, although apparently large, may at least have been stable" [42] Is it possible that a large and stable proportion of the nation's currency supply has been "missing" for more than four decades?

---

[39] The original 1984 survey was repeated in 1986 with substantially the same result, namely that approximately 12% of aggregate U.S. currency is held by U.S. individuals. The Federal Reserve survey covers both coin and paper currency whereas the present study is based exclusively on paper currency of $1 ; $5; $10; $20; $50; $100 denominations. Coins are assumed to be primarily held in the U.S. For illustrative purposes we use a figure of 12% as an approximation of paper currency held by U.S. households. The stock of currency in circulation refers to the stock of currency outside the vaults of depository institutions.

[40] Federal Reserve Bulletin, March 1987. p.190.

[41] Arthur Kennickell of the Federal Reserve has brought to my attention two surveys of household currency holdings that were undertaken in 1944 and 1945. The earlier and more primitive survey was conducted by the Department of Agriculture for the Board of Governors of the Federal Reserve System. That survey found that only 12.8% of outstanding U.S. currency was held by U.S. households. Indirect evidence from the 1945 Liquid Asset Study reinforces the view that only a small percentage of the stock of currency was held by U.S. households.

[42] Federal Reserve Bulletin, (1987) p. 191.

Figure (22) displays the allocation of per capita currency holdings based on the assumption that over the period 1940-1986 only 12% of the U.S. currency stock was held by U.S. adult residents. In 1986, U.S. adults were estimated to have held an average of approximately $100 in cash, leaving some $850 per capita unaccounted for. If we assume that another 10% of the nation's currency is held by business enterprises[43] this accounts for another $17 billion in cash holdings. In addition, one must take account of unreported currency held for the purpose of making cash payments in the "underground" economy. The IRS has estimated that "unreported legal source income" of filers and non filers of tax returns amounted to some $250 billion in 1981 and that another $35 billion was unreported from "illegal activities".[44] Extrapolation of these IRS estimates to 1984 suggests that perhaps as much as $320 billion dollars was expended in the "underground economy". Other estimates suggest that "unreported" income in 1984 may have been as high as $650 billion.[45] Roughly 70% of these "underground" expenditures are believed to have been effected by cash payments suggesting a total of between $210 - $455 billion of "underground" cash payments. Applying the estimated velocity of currency to "underground" cash payments implies that between $4.4 and $9.0 billion of missing cash may be held by participants in the "underground economy". Given these estimates, Figure (23) displays a provisional allocation of U.S. currency holdings among individuals, firms and the "underground" economy. Taking the higher estimate of "underground" currency holdings, Figure (23) suggests that in 1984, U.S. adults held $19.4 billion in cash; firms held $16.1 billion and "underground" holdings amounted to some $9.0 billion leaving approximately $117 billion or 73% of the nation's currency supply unaccounted for.

B. The problem of "missing payments".

The currency enigma deepens when efforts are made to obtain provisional estimates of cash payments and their implications. Provisional estimates of domestic currency payments are derived by applying the estimated currency velocity series to the provisional allocation of currency holdings among individuals, firms and the

---

[43] It is unlikely that firms would hold an amount of currency equal to that held by all households since firms are very adept at economizing on cash holdings by the use of cash management methods.

[44] Income Tax Compliance Research, Department of the Treasury July 1983.

[45] Feige, (1987).

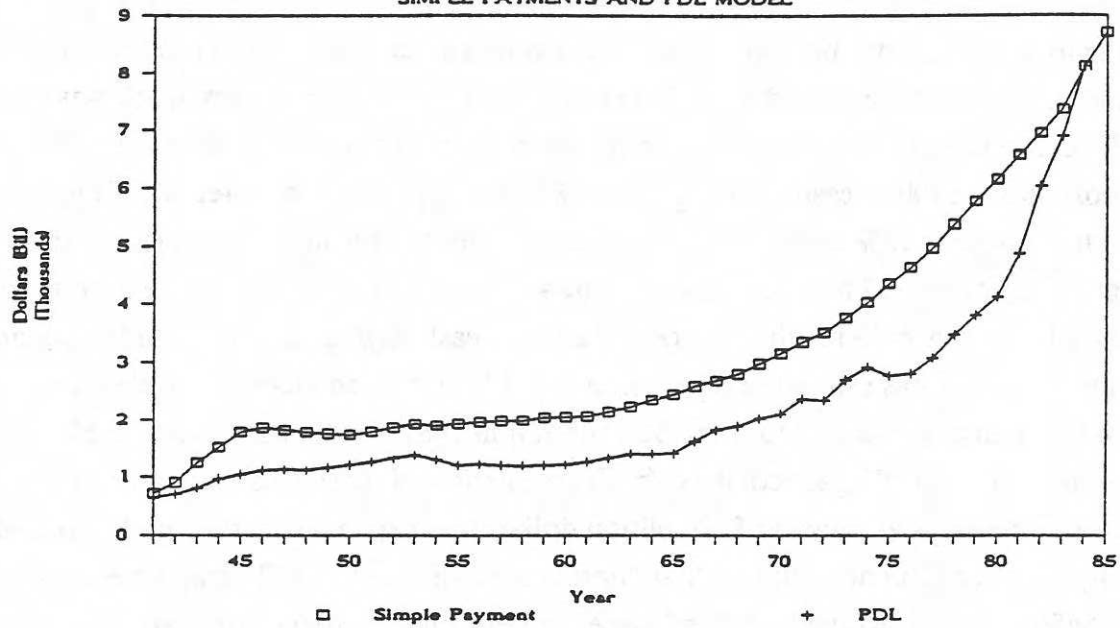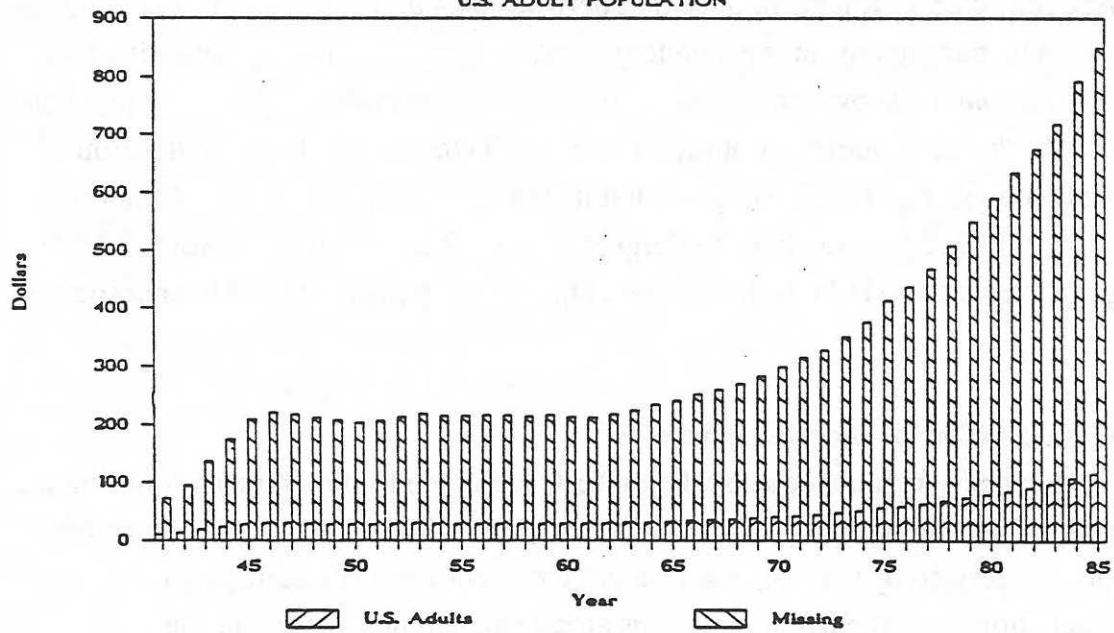FIGURE 21

ESTIMATED CASH PAYMENTS
SIMPLE PAYMENTS AND PDL MODEL

□   Simple Payment   +   PDL

FIGURE 22

PER CAPITA CURRENCY HOLDINGS
U.S. ADULT POPULATION

▨   U.S. Adults   ◹   Missing

"underground"[46]. Figure (24) displays the estimated reported per capita cash payments of U.S. adults and Figure (25) displays the ratio of these cash payments to total personal consumption expenditures. By 1985, per capita reported cash payments of U.S. adults are estimated to be $5776 suggesting that 40% of total personal consumption expenditures were effected by cash payments undertaken with only 12% of the nation's currency supply.[47] The time series estimates displayed in Figure (25) reveal that the ratio of these cash payments to consumption expenditures have steadily declined since the end of World War II. Assuming that the proportion of "missing cash" remained stable since World War II implies that 12% of the nation's currency supply accounted for cash payments that averaged 93% of personal consumption expenditures during the decade of the 50's; 67% during the 60's; 52% during the 70's and 41% during the first half of the 1980's.[48]

Figure (26) displays the ratio of estimated total domestic cash payments of individuals, firms and "underground" to GNP.[49] The calculations suggest that during the 1940's and 1950's estimated domestic currency payments exceeded measured GNP. The decade averages of the ratio of estimated currency payments to GNP for the 1950's - 1980's are respectively 1.18; .86; .70 and .61. These anomalous findings suggest that

---

[46] These estimates are derived by assuming that firms and individuals held stable proportions of the currency supply over time and that 70% of "unreported income" series estimated by Feige (1986) is paid with cash.

[47] It is of course possible that the survey methods employed by the Federal Reserve considerably understate the actual currency holdings of U.S. households. If the distribution of currency holdings is highly skewed, a very small fraction of households may hold very large currency balances. These households are unlikely to be adequately represented within the sample survey. Smith (1986) estimated that the richest one half of one percent of households own 27 percent of the total net worth of all American households. The under representation of these ultra wealthy households would help to account for some of the "missing currency". However, the unaccounted for cash payments made with these currency holdings only deepen the problem of "missing payments" and "missing expenditures".

[48] The estimated fraction of consumption expenditures effected by cash payments of U.S. adults would be higher if the adult population held greater cash balances than those reported in the Federal Reserve survey. The problem arises from self selection bias that manifests itself in either high rates of non response or in false reporting of actual cash holdings. Currency is a superior medium of exchange for transactions that individuals wish to keep hidden from government authorities. Cash is therefore widely believed to be the preferred means of payment in illegal activities and for income payments that are unreported for tax purposes. Respondents engaged in such activities are therefore unlikely to accurately report their true currency holdings.

[49] A substantial portion of GNP represents government purchases of goods and services and investment undertaken by business enterprises. Payment for these final goods and services is typically effected by check transactions. These observations only serve to deepen the problem of missing cash payments.

FIGURE 23

PROVISIONAL ALLOCATION OF U.S. CURRENCY

1984

Figure 23: Provisional Allocation of U.S. Currency, 1984. Bar chart with y-axis "Dollars (Bil)" ranging from 0 to 120. Categories: U.S. ADULTS (~20), U.S. FIRMS (~16), UNDERGROUND (~7), MISSING (~119).

FIGURE 24

PER CAPITA CASH PAYMENTS: U.S. ADULTS

Figure 24: Per Capita Cash Payments: U.S. Adults. Bar chart with y-axis "Dollars (Thousands)" ranging from 0 to 6, x-axis "Year" from approximately 42 to 85.

either cash payments are grossly overestimated,[50] or alternatively, that official estimates of recorded consumption expenditures and GNP are grossly underestimated.[51] This latter possibility gives rise to the problem of "missing income and expenditures".

C. Missing Income and Expenditures.

The apparent disparity between estimated cash payments and official estimates of income and expenditures may in part be due to a systematic understatement of income and expenditures in the National Income and Product Accounts. The National Income and Product Accounts have recently been substantially revised to reflect earlier omissions from the accounts due to estimated under-reporting of income on Federal Tax returns.[52] It is likely that these revisions represent only the first step in recognizing the full extent of understated GNP. It is possible that a substantial fraction of final expenditures for goods and services as well as income payments effected by cash elude current procedures for estimating NIPA accounts. However this issue is finally resolved, the current disparity between estimated cash payments and estimated expenditures on goods and services and the related disparity between the total amount of currency in circulation and the amount estimated to be held by individuals in the U.S. can no longer be ignored.

D. The implications of "missing cash" and "missing payments".

Accepting for the moment the FED's surprisingly low estimates of the amount of U.S. currency held by Americans, we still must examine the implications of the corollary finding that perhaps as much as 75-80% of the nation's currency supply is held and spent abroad. If currency held abroad turned over at only one fifth the rate of domestic velocity, the volume of cash payments thus generated would be sufficient to purchase almost 10% of the officially recorded GNP of the rest of the world. If this missing currency turned over at the same rate as that estimated for domestic currency the cash

---

[50] It should be noted that these estimates of cash payments exclude all cash payments made with "missing currency".

[51] The possibility that official government statistics on income and expenditures are systematically downward biased was suggested by Feige (1979).

[52] In 1985 the Bureau of Economic Analysis was forced to include a upward revision of $147 billion to reflect previous understatement of NNP. A similar upward revision of $100 billion was made for Personal Income.
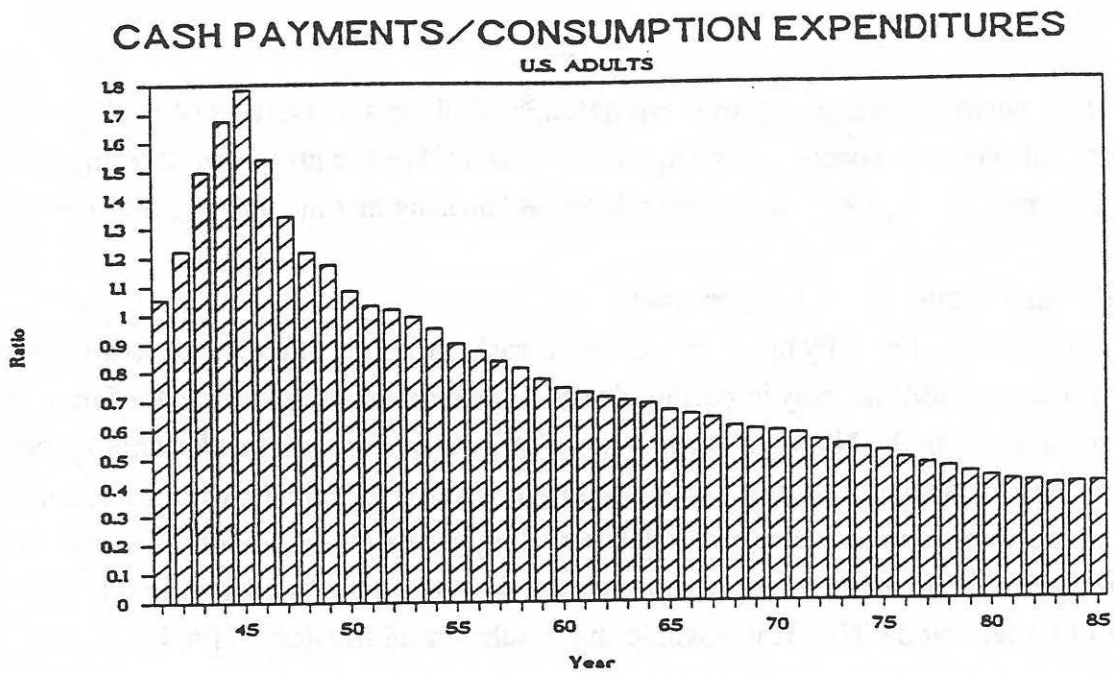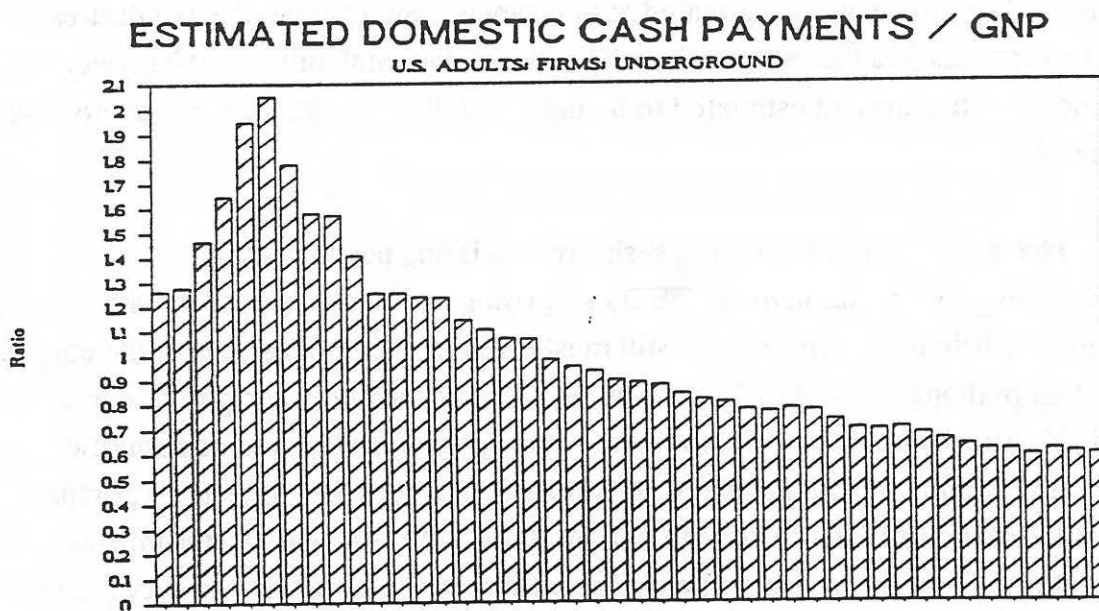
FIGURE 25

CASH PAYMENTS/CONSUMPTION EXPENDITURES
U.S. ADULTS

FIGURE 26

ESTIMATED DOMESTIC CASH PAYMENTS / GNP
U.S. ADULTS: FIRMS: UNDERGROUND

payments thus generated could buy roughly 50% of the rest of the world's GNP.[53] One is left with the uncomfortable choice of concluding that either world GNP is grossly underestimated or that a major fraction of U.S. currency is held abroad in idle hoards.

The contention that more than three quarters of U.S. currency is held abroad in idle hoards is highly questionable. The fragmentary survey evidence suggests that the problem of "missing currency" may have been with us for some forty five years. If the overwhelming percentage of U.S. currency was held in idle hoards for such an extended period of time, we would expect that average note lifetimes would be much longer than those calculated from observations on redemptions and issues of currency. Moreover, the significant inter-temporal swings in interest rates, inflation and exchange rates would have created considerable economic incentives for individuals to adjust their portfolio stores of value over time. Such shifts are not however reflected in the estimates of average note lifetimes. Moreover, evidence on the amount of currency that returns to the Federal Reserve System each year suggests that the outstanding currency does in fact circulate relatively freely throughout the economic system. Federal Reserve Statistics on the amount of currency received from circulation in 1986 suggests that the that the stock of smaller denomination bills in circulation [$1 - $20] returns to the Federal Reserve approximately every six months. The stock of large denomination bills [$50 and $100] returns to the Federal Reserve within a period of approximately twenty seven months. These figures suggest that currency held abroad is either spent on goods and services or is exchanged for interest earning assets.[54] In either case, this currency will find its way to foreign fiduciary institutions which have an incentive to exchange U.S. currency for interest earning assets. A significant fraction of the U.S. currency used abroad would therefore be expected to return to the U.S. and reenter the circulation system within the U.S.[55] Thus, it is difficult to accept the suggestion that the "missing currency" is predominantly held in idle hoards.

---

[53] Per capita currency holdings in the U.S. are in fact significantly lower than per capita currency holdings in many other nations. It is likely that the problem of missing currency and missing payments is an international phenomenon.

[54] U.S. currency held abroad can be exchanged for dollar denominated bearer certificates that are marketable and yield interest. Some Eastern European countries provide incentives for their citizens to exchange U.S. currency for dollar denominated local currency. In Poland for example, dollar denominated local currency is issued and required for the purchase of luxury items such as automobiles.

[55] In order to examine this hypothesis more directly, it would be useful to sample U.S. currency notes held in foreign banks or returning from abroad in order to observe whether the age-quality distribution of such notes is significantly different from the age-quality distribution of notes circulating in the U.S.

If huge quantities of U.S. currency are held abroad, this represent both advantages and perils for the U.S. economy. Foreign holdings of U.S. currency are in effect an interest free loan to the U.S. The seigniorage profits arising from this interest free loan made by foreign holders of U.S. currency could be in excess of ten billion dollars annually. Such a loan serves to reduce both the domestic deficit and the balance of payments deficit.

On the other hand, massive holdings of U.S. currency abroad, particularly in the form of large denomination bills, pose a threat to the integrity of the U.S. dollar. Rapid advances in both printing and copying technologies have made U.S. currency very vulnerable to forgery. The counterfeit threat is heightened by the fact that U.S. currency design has remained essentially unchanged for almost half a century.[56] Despite the vigilance of the Secret Service in minimizing the amount of domestic counterfeiting, the major counterfeit threat comes from abroad where the power of the Secret Service to deter, detect, and eliminate counterfeit U.S. currency is severely circumscribed.[57] Gresham's law suggests that a loss of confidence in the integrity of U.S. currency could precipitate a serious upheaval in international currency markets.[58]

IX. Summary, Conclusions and Recommendations

Models of the currency circulation system have generated some preliminary answers

---

[56] It is ironic that the political pressures to maintain the present design of our currency as a symbol of the strength of the American dollar may only serve to weaken it. As distinct from many other currencies, U.S. currency design is very simple. It is not multicolored, contains no watermarks or threads and has little in the way of micro printing or other counterfeit deterrent features. High speed sorting machines presently rely on a magnetic trace in the signature on notes as the sole means of detecting counterfeit notes. Even much more highly elaborated notes of other countries have been shown to be vulnerable to counterfeiting. The Central Bank of the Netherlands has withdrawn seventeen notes from circulation during the past forty years. In 1985, the Netherlands Central Bank found it necessary to recall the entire stock of one hundred guilder "De Ruyter" notes which have been in circulation for only thirteen years.

[57] Foreign production of counterfeit U.S. currency is difficult to monitor. Detection and elimination of counterfeit notes in circulation abroad is hampered by the lower level of familiarity with U.S. currency by foreigners and by the apparently huge supplies of legitimate notes held abroad. The seriousness of the foreign counterfeit problem is exemplified by the growing number of European firms that refuse to accept $50 and $100 dollar denomination notes in payment for goods.

[58] Improvement in the design of U.S. currency as a counterfeit deterrent strategy has been under consideration for some time. Recent proposals to modify U.S. notes include the addition of a thread and additional micro printing to the current design. Even if implemented, these relatively minor changes are unlikely to significantly reduce the counterfeit threat. Moreover, unless the government is willing and able to undertake a major recall of large denomination notes, it is likely that older issued notes and their facsimiles will remain in circulation for many years.

and some challenging new questions concerning the whereabouts of currency and the total magnitude of cash payments. Disaggregation of the currency stock into denomination specific components permits an examination of the role played by each currency denomination in the currency payments system. The smallest denomination notes [$1 -$10] account for 61% of the number of notes in circulation but make up only 12% of the value of notes in circulation [Table 2]. These denominations account for approximately 30.9% of total cash payments. Since the cost of note provision and note maintenance depends upon the physical number of notes in circulation and the frequency with which these notes require replacement, it is these smallest denomination notes that place the greatest cost burden on the Bureau of Engraving and Printing and on the Federal Reserve Banks. Based on our estimates of age specific mortality rates, $1 and $5 denomination notes have an average lifetime of approximately 1.4 years and $10 notes have an average lifetime of approximately 3.5 years.

TABLE 2
Denomination Specific Shares of Number of Notes, Value and Payments 1985

	Denomination					
	$1	$5	$10	$20	$50	$100
Percent of:						
Number of Notes	38.82	10.11	11.80	27.24	4.37	7.66
Value of Notes	2.24	2.91	6.80	31.38	12.58	44.10
Cash Payments	8.83	11.81	10.30	33.69	11.77	23.61

The largest denomination notes [$50 and $100] represent only 12% of the number of notes in circulation but account for 57% of the value of the nation's currency supply. These denominations are estimated to effect 35% of the value of cash payments. Our estimates suggest that the average lifetime of a $100 bill is approximately ten years and that such notes may remain in circulation as long as twenty five years before final redemption.

The denomination specific shares presented in Table (2) also serve to highlight the depth of the problems of "missing currency" and "missing payments". The Federal Reserve Board estimates that in 1986 U.S. household's held only $20 billion of the nation's $178 billion of total currency in circulation. It therefore appears that " a large percentage", perhaps as much as 80 % of the U.S. currency stock "was apparently held

in unreported hoards, "underground" for illegitimate purposes, or off shore."[59] It is unlikely that hoarded currency or currency used for illegitimate purposes will be held in small denomination notes. Yet, if most of the smallest denomination notes [$1 - $10] that make up 12% of the currency supply are assumed to be used by U.S. households for legitimate transactions purposes, the Federal Reserve's findings require that virtually all of the nation's $20, $50 and $100 dollar bills are "underground" or overseas. This hypothesis can only be evaluated by research designed to estimate the stock, size distribution and age-quality distribution of notes held in overseas deposits.

If notes turnover as rapidly as suggested by the estimates in this paper, then the volume of cash payments can not be reconciled with official Bureau of Economic Analysis estimates of aggregate economic activity. Either economic activity is underestimated or the number of transactions sustained by notes is considerably overestimated. Current estimates of economic activity depend heavily on the survey responses of firms and individuals who may have substantial incentives for false reporting. One way to circumvent this fundamental methodological problem is to design an information gathering system that is based on observations of the outcome of economic behavior rather than on self reporting of behavior. A major statistical effort to monitor the volume of check and currency payments would provide an important complement to current procedures for measuring economic activity. Independently derived estimates of total payments could provide an independent check on the veracity of our current information system detailing the volume of aggregate economic transactions.

The foregoing investigation into the currency circulation system raises as many questions as it answers and these questions comprise what we call the "currency enigma". It is now clear that huge amounts of U.S. currency in circulation are "missing". Since per capita holdings of many European currencies are well in excess of U.S per capita holdings, it appears that the problem of "missing currency" is an international rather than a purely domestic problem. Estimates of the velocity of currency suggest that payments effected by cash can not be readily reconciled with official statistics on income and expenditures.

The denomination structure of currency holdings has dramatically shifted toward the largest denomination notes. These notes now comprise 45% of the value of the nation's currency supply. The long estimated lifetime of large denomination notes, the possibility that many of these notes circulate outside of our borders, and the

---

[59] Federal Reserve Bulletin, March, 1987 (p. 191).

technological advances in copying and printing technologies, suggest that our nation's currency may be very vulnerable to extensive foreign counterfeiting.

Despite the deeply entrenched political opposition to changes in our currency, the potential threat of a counterfeit crisis, the widespread use of cash to circumvent the prohibition on drug traffic, and its use in tax evasion suggest that the time is ripe to consider the costs and benefits of fundamental changes in our currency system. Policy options include a new design for our present currency that incorporates new technologies for effective counterfeit deterrence, [60] and the recall of at least the largest denomination notes. The government might also consider measures to reduce the incentives for citizens to employ currency transactions as a means of conducting illegal activities and circumventing the public will as expressed in current tax codes.

One of the major roles of government in a market economy is to establish and maintain the institutions that promote legitimate exchange. Government accomplishes this task by providing legal tender and enforcing contracts. From this perspective, government can be viewed as a "ubiquitous broker", a tacit party to all exchanges. The government's provision and maintenance of currency provides the brokerage function of facilitating exchange.[61] Brokerage functions are typically rewarded with brokerage fees that establish spreads between purchase and sale prices. The government could also establish a brokerage fee for currency usage by selling currency at a premium and purchasing currency at a discount.[62] The revenues thus collected from a "currency brokerage fee" could used to decrease present income tax rates so as to maintain a revenue neutral system. A reduction in marginal income tax rates decreases incentives for tax evasion [63]. A "brokerage fee" on currency usage raises the cost of employing cash payments as a means of tax evasion and raises the cost of conducting other illegal activities that employ currency as a means of circumventing government detection.

However the "currency enigma" is finally resolved, the principle contribution of this

---

[60] Since currency is often identified as a symbol of national pride, a red, white and blue design may satisfy the need for a "patriotic symbol" and may also serve as a useful counterfeit deterrent.

[61] Under current institutional arrangements, the government's sole source of compensation for the provision and maintenance of currency derives from the seigniorage it earns on the outstanding stock of currency. Seigniorage returns are estimated to be somewhat in excess of ten billion dollars per annum.

[62] The government could for example charge a 2% fee for both the sale and purchase of currency. See Feige (1989b).

[63] The revenue loss resulting from federal income tax evasion is estimated to be in the neighborhood of $100 billion per annum.

paper is the estimation of a provisional time series for currency velocity. These velocity estimates can be employed to construct a historical time series of index number measures of both the money supply and payments velocity. Such money aggregate indices can then be tested against conventional money stock measures to see if they yield better predictions of economic activity. Independent currency velocity measures permit estimates of total payments which in turn can be used as a check on the reliability of a conceptually consistent aggregate of total transactions. Until such time that total transactions measures are established, independent estimates of (MV) can be used as a proxy for (PT) in money demand functions to more adequately reflect the principle motive for holding money, namely to effect exchange. As Fisher suggested at the beginning of the century, the availability of estimates of currency velocity is the necessary condition for opening "a new realm in monetary statistics".

REFERENCES

Avery, Robert B. et. al. "The Use of Cash and Transaction Accounts by American Families", Federal Reserve Bulletin, February, 1986.

Avery, Robert B. et al. "Changes in the Use of Transaction Accounts and Cash from 1984 to 1986" Federal Reserve Bulletin, March, 1987.

Boeschoten, W.C. and Fase, M. M. G. Betalingsverkeer en officieuze economie in Nederland 1965-1982 Monetaire Monografieen Nr.1, De Nederlandsche Bank n.v. 1984.

Bureau of Economic Analysis, "Revised Estimates of the National Income and Product Accounts of the United States, 1929-85: An Introduction" Survey of Current Business, Vol. 65, Number 12 December, 1985.

Cramer, J.S. "The Regular and Irregular Circulation of Money in the United States" A & E Report #11/80, University of Amsterdam. 1980 pp. 225-32.

Cramer, J. S. "The Volume of Transactions and the Circulation of Money in the United States", 1950 -1979, Journal of Business and Economic Statistics, April 1986, 2,

Corrado, Carol and Paul Spindt "T vs. Q: The Measurement and Analysis of Monetary Transactions" Mimeo, December, 1988

Dotsey, Michael "The Demand for Currency in the United States", Journal of Money Credit and Banking Vol. 20 No 1 (February 1988) pp. 22 - 40.

Feige, Edgar L. "How Big is the Irregular Economy?" Challenge 22 Nov./Dec. 1979 pp. 5-13.

Feige, Edgar L. The Swedish Payments System and the Underground Economy Industrial Institute for Economic and Social Research, Stockholm, 1985.

Feige, Edgar L. "A Re-Examination of the Underground Economy in the United States"

IMF Staff Papers Vol.33 No. 4 (December 1986)

Feige, Edgar L. "An Analysis of Internal Revenue Service Estimates of Unreported Income" University of Wisconsin-Madison. 1987.

Feige, Edgar L. (ed.) The Underground Economies: Tax Evasion and Information Distortion, Cambridge University Press, 1989.

Feige, Edgar L. "Taxing All Transactions: The Automated Payment Transaction Tax System", Presented at the International Institute of Public Finance, Buenos Aires, August, 1989.

Fisher, I. The Purchasing Power of Money Macmillan, 1911.

Fisher, I. "A New Method of Estimating the Velocity of Circulation of Money". Journal of the Royal Statistical Society, Vol. 72, 1909.

Friedman, M. "The Quantity Theory of Money- A Restatement" Studies in the Quantity Theory of Money The University of Chicago Press, 1956.

Internal Revenue Service, Income Tax Compliance Research Department of the Treasury, July, 1983.

Keynes, J. M. A Treatise on Money Macmillan, 1930.

Keynes, J. M. The General Theory of Employment, Interest, and Money Macmillan, 1936.

Laurent, R. Currency Transfers by Denomination Ph. D Thesis, University of Chicago 1970.

Selden, R. "Monetary Velocity in the United States" in Friedman, M ed. Studies in the Quantity Theory of Money University of Chicago Press, 1956.

Smith, James D. "Wealth in America" ISR Newsletter, Winter 1986-87

Spindt, Paul, A. "Money is What Money Does: Monetary Aggregation and the Equation of Exchange." J.P.E. 93 (February, 1985): 175-204.