Methods Used in the Study

A.1. CUSUM Square Test (BDE) for model stability

Departure from model specification can be studied by making use of residuals. The plots of OLS residuals against time or the cumulative sums of residuals against time, may not fully reflect the sensitive departures in a model specification. Hence, Brown et al (1975) made use of recursive residuals to examine departures in model specification. The details of the recursive residual test, alternatively known as Brown - Durbin - Evan (BDE) test or Cusum Square test, are as follows:

Let us consider the basic regression model

\[ Y_t = X_t \beta_t + u_t, \quad t = 1, 2, \ldots, n \]

where \( Y_t \) is \( (t \times 1) \) vector of observations on the dependent variable, \( X_t \) is \( (t \times k) \) matrix of observations on \( k \) independent variables, \( \beta_t \) is \( (k \times 1) \) vector of parameters, and \( u_t \) is \( (t \times 1) \) vector of observations on error terms. The suffix "\( t \)" is used for \( \beta \) also to indicate possibility of its varying over the given "\( n \)" time periods. It is also assumed that \( u_t \) is normally distributed with mean zero and variance \( \delta_t^2 \). Now, the null hypothesis \( (H_0) \) of constancy of regression relationship may be written as:
\[ H_0 \quad \beta_1 = \beta_2 = \beta_n = \beta \text{(say)} \]
\[ \delta_1^2 = \delta_2^2 = \delta_n^2 = \delta^2 \text{(say)} \]

Here the emphasis is more on detecting differences among \( \beta \)'s than among \( \delta \)'s.

Let \( \beta_{t-1} \) be the OLS estimate of \( \beta \) based on the series of observations from period 1 through \( t-1 \). Then the prediction error for period \( t \) is given by

\[ e_t = Y_t - x_t \beta_{t-1} \]

where \( X_t \) is a row vector containing \( t \)'th observation on all the independent variables. The recursive residuals, \( e_t^* \), is defined as

\[ e_t^* = e_t / \left(1 + x_t(x_{t-1}'x_{t-1})^{-1}x_t'\right)^{1/2} \]

where \( x_{t-1} \) contains only the first \( t-1 \) rows of \( x_n \). For a set of \( n \) given time series observations, there will be \((n-k) \times 1\) vector of residuals. These residuals have a number of desirable properties. They are easy to compute and straightforward to understand intuitively as a form of prediction error. They also react distinctively to various departures from \( H_0 \).

The next step is to define a statistics

\[ S_t = \sum_{k+1}^{t} e_t^{*2} / \sum_{k+1}^{n} e_t^{*2}, \quad t = k+1, ... , n \]

where \( S_t \) is the ratio of the sum of square of one-period prediction residuals from \( k-1 \) period to \( t \) period, to the sum of squares of one-period prediction error from the \( k-1 \) to \( n \)th period. The values of \( S_t \) therefore will be between zero and unity as indicated below.
\[ S_t = 0 \quad \text{if} \quad t < k+1 \]
\[ S_t = 1 \quad \text{if} \quad t = n \]

Under \( H_0 \), \( S_t \) will have a Beta distribution with approximate mean value,

\[ E(S_t) = \frac{(t+k)}{(n-k)} \]

where "E" denotes expectations. This indicates that the plots of \( S_t \) against \( t \) should be along with its mean path. For a two sided test of hypothesis, a linear confidence interval for \( S_t \) is constructed. Accept \( H_0 \) if the values of \( S_t \) lie in the interval

\[ C_0 \pm \frac{(t-k)}{(n-k)} \]

where \( C_0 \) is distributed as Pyke's modified Kolmogorov-Smirnov statistics. The values for \( C_0 \) at various significance level are available in Brown et al (1975) by entering the table at \( m = \frac{(n-k)}{2} \) when \( (n-k) \) is even and interpolating linearly between \( m \) and \( m= \frac{1}{2} (n-k) + \frac{1}{2} \) when \( (n-k) \) is odd.

**A.2. KPSS Unit Root Test:**

This test, due to Kwiatkowski et al (1992) provides a test statistics under the null hypothesis of stationarity and alternative of unit root to test that the series is differenced stationary. The test statistics are derived by computing the following statistics

\[ (1-T^2) \sum_{t=1}^{T} \{S_t^2/\hat{\sigma}^2(I)\} \]

where, \( T \) is sample size, \( S_t = \sum_{t=1}^{T} e_t \), \( e_t \) being the residual from a regression of say, \( Y_t \) on an intercept and a time trend. Also \( \hat{\sigma}^2(I) \) is a consistent estimator of the long run variance of \( Y_t \) and is constructed as in Kwiatkowski et al (1992). The critical values for test
statistics both with and with out trend at 5% level are presented in Kwiatkowski et al (1992)

A.3. Johansen and Juselius (JJ) Procedure:

Until recently, the empirical works involving the estimation of cointegrating vectors employed the single-equation error-correction technique proposed by Engle and Granger (1987). In the past few years, there has been a movement towards estimating cointegrating relationship in a system of equations framework to make better use of all the information available in the long and short run fluctuation of each variable. Johansen (1988) propounded a method (which was later refined by Johansen and Juselius (1990)), that allowed for the testing of more than one cointegrating vector.

The Johansen test for cointegration begins by considering the unrestricted reduced form of a system of variables which, by assumption, can be represented as a finite order vector Autoregressive model (VAR)

\[ Y_t = \mu + A(L) Y_{t-1} + U_t \]  (A 1)

where \( Y_t = [Y_{1t} \quad Y_{nt}]^T \) \( T \) being the transpose, \( A(L) = \sum_{i=1}^{k} A_i L^i \). \( E(U_t U_t^T) = \sum \)

The key requirement for this methodology is that none of the variables in the system be integrated of order zero. This approach also assumes that the disturbances of the model, \( U_t \) are distributed normally.
Equation (A 1) can be presented in the form of an equivalent vector error correction model (VECM) from as the following

\[ \Delta Y_t = \mu + B(L) \Delta Y_{t-1} + \Pi Y_{t-k} + U_t \quad (A.2) \]

If there is no cointegration among the set of \( n \) variables included in \( Y \), then the rank of \( \Pi \) is zero. In this case equation (A 2) becomes a simple VAR in first differences. On the other hand, if the rank of \( \Pi \) is \( n \), there are \( n \) linearly independent combinations of \( Y_t \) that are I(0). Since any linear combination of I(0) variables must also be I(0), this means that the \( n \) variables in \( Y_t \) must themselves be I(0), and it is appropriate to estimate an unrestricted VAR in levels, in equation (A 1) (though it may be rewritten in error-correction form as in equation A 2). In the case where, some but probably fewer than \( n \), stationary linear combinations exist means that there are cross-equation restrictions among the coefficients of \( A(L) \) in equation (A 1). The advantage of representing eqn (A 1) in VECM form is that in equation (A 2), all of the cross-equation restrictions are concentrated in the matrix \( \Pi \). In the Johansen approach, the number of such restrictions is determined by, first estimating \( \Pi \) without restrictions from equations (2) and then determining the rank of \( \Pi \). The restrictions on \( \Pi \) are then imposed and a restricted VECM is estimated.

In practice, this process is achieved in two steps. First \( \Delta Y_t \) and \( Y_{t-k} \) are each regressed on a constant and lagged \( \Delta Y \), yielding residuals.

\[ R_{0t} = \Delta Y_t - \hat{a} - \tilde{b}(L) \Delta Y_{t-1} \quad (A.3) \]

\[ R_{kt} = Y_{t-k} - \hat{c} - \tilde{d}(L) \Delta Y_{t-1} \quad (A.4) \]
Equation (A 1) can be presented in the form of an equivalent vector error correction model (VECM) from as the following

\[ \Delta Y_t = \mu + B(L) \Delta Y_{t-1} + \Pi Y_{t-k} + U_t \]  
(A.2)

If there is no cointegration among the set of n variables included in Y, then the rank of \( \Pi \) is zero. In this case equation (A 2) becomes a simple VAR in first differences. On the other hand, if the rank of \( \Pi \) is n, there are n linearly independent combinations of \( Y_t \) that are I(0). Since any linear combination of I(0) variables must also be I(0), this means that the n variables in \( Y_t \) must themselves be I(0), and it is appropriate to estimate an unrestricted VAR in levels, in equation (A 1) (though it may be rewritten in error-correction form as in equation A 2).

In the case where, some but probably fewer than n, stationary linear combinations exist means that there are cross-equation restrictions among the coefficients of \( A(L) \) in equation (A 1). The advantage of representing eqn (A 1) in VECM form is that in equation (A 2), all of the cross-equation restrictions are concentrated in the matrix \( \Pi \). In the Johansen approach, the number of such restrictions is determined by, first estimating \( \Pi \) without restrictions from equations (2) and then determining the rank of \( \Pi \). The restrictions on \( \Pi \) are then imposed and a restricted VECM is estimated.

In practice, this process is achieved in two steps. First \( \Delta Y_t \) and \( Y_{t-k} \) are each regressed on a constant and lagged \( \Delta Y \), yielding residuals

\[ R_{0t} = \Delta Y_t - \hat{a} - \hat{b}(L) \Delta Y_{t-1} \]  
(A.3)

\[ R_{kt} = Y_{t-k} - \hat{c} - \hat{d}(L) \Delta Y_{t-1} \]  
(A.4)
The ordinary least squares (OLS) estimate of $\Pi$ from equation (A.2), with no restriction on the rank of $\Pi$, would then be obtained by regressing $R0_t$ on $Rk_t$.

$\Pi$ can always be recoded in a variety of ways, as

$$\Pi = \alpha \beta \quad \text{where } \alpha, \beta \text{ are } n \times n$$

However, under the hypothesis that $\Pi$ is of reduced rank, the dimensions of $\alpha$ and $\beta$ can be reduced to $n \times r$, $r < n$. The Johansen procedure involves choosing the first column of $\beta$ as the maximum likelihood estimate of the linear combination of $Rk_t$, (that is, $\beta_1 Rk_t$) that is most correlated with $R0_t$. The second column of $\beta$ is the linear combination, orthogonal to $\beta_1$, which has the next highest correlation, and so on. A representation of the VECM which is equivalent to equation (A.2) is then given by

$$\Delta Y_t = \mu + B(L) \Delta Y_{t-1} + \alpha ECM_{t-k} \quad (A\ 5)$$

where $ECM_{t-k} = \beta Y_{t-k}$

$\beta = [\beta_1 \beta_n]^T$

Johansen and Juselius (1988) provide two test statistics, known as the trace and $\lambda$-max statistics, to determine the rank of the matrix $\Pi$. If for example, the tests indicate that two cointegrating vectors exist, maximum likelihood estimates of those vectors are given by the first two columns of $\beta$ and the VECM (equation (A.2)), is given under the restriction of two cointegrating vectors by

$$\Delta Y_t = \mu + B(L) \Delta Y_{t-1} + \alpha_1 ECM_{1t-k} + \alpha_2 ECM_{2t-k} \quad (A\ 6)$$

where $ECM_i = \beta_i Y_{t-k}$ \quad \text{i=1,2}
All the variables in (A 6) are now stationary and inference on any of the parameters in (A 6) can proceed using asymptotically valid standard distribution theory. The coefficients $\alpha_1$ and $\alpha_2$ represent the effects of the stationary linear combinations ECM1 and ECM2 on the system and are referred to as loadings (in a single equation approach, they are frequently referred to as "speeds of adjustment").

It should be noted that no structural interpretation can be placed on the error-correction terms without making further identifying assumptions. However, in the present case, if there were only one cointegrating vector in the system, and if it is included in the financial aggregate, one might wish to interpret it as a "long run" demand function. If more than one cointegration vector exists among the data, any structural interpretation is difficult to place on it, though examining the effect of each cointegrating vector on future changes of the variables in the system may provide indirect evidences to the nature of the relationship that has been identified.

A.4. Tests for Non-nested Hypotheses (The Davidson-Mackinnon J-Test)

Let us consider the problem of testing two hypotheses

\[ H_0 \; y = \beta x + u_0 \quad u_0 \sim \mathcal{N}(0, \sigma_0^2) \quad (A \; 7) \]

\[ H_1 \; y = \gamma z + u_1 \quad u_1 \sim \mathcal{N}(0, \sigma_1^2) \quad (A \; 8) \]

The hypotheses are called non nested since the explanatory variables under one of the hypotheses are not a subset of the explanatory variables in the other.
H₀ being the maintained hypothesis, the test procedure for testing H₀ against H₁ is as follows

Equation (A 8) is estimated by OLS and the predicted value of y is obtained as \( \hat{y}_1 = \hat{y} z \). Then the regression equation

\[
y = \beta x + \alpha \hat{y}_1 + u \quad (A \ 9)
\]

is estimated to test the hypothesis that \( \alpha = 0 \). If the hypothesis is rejected, then H₀ is rejected by H₁. If the hypothesis is not rejected, then H₀ is not rejected by H₁.

Similarly testing H₁ against H₀ starts with estimating equation (A 7) by OLS and generating the predicted value of y viz.

\[
\hat{y}_0 = \hat{\beta} x
\]

Then the regression equation,

\[
y = \gamma z + \delta \hat{y}_0 + v \quad (A \ 10)
\]

is estimated to test the hypothesis \( \delta = 0 \). If the hypothesis is rejected, then H₁ is rejected by H₀ and vice versa.

As suggested by Davidson and Mackinnon the models given above H₀ and H₁ can be combined into a single model viz.

\[
y = (1-\alpha) \beta x + \alpha (\gamma z) + u \quad (A \ 11)
\]
where one has to test $\alpha = 0$ versus $\alpha = 1$. As it is difficult to estimate $\beta$, $\gamma$, and $\alpha$ from this model we get the estimates of $(1-\alpha)\beta$ and $\alpha\gamma$. As shown by Davidson-Mackinnon one can substitute $\hat{y}_1 = \hat{y}z$ for $\gamma z$ in equation (A.11) and then test $\alpha = 0$. They also show that under $H_0$, $\hat{\alpha}$ from (A.9) is asymptotically $N(0, 1)$. Because $\alpha$ and $\beta$ are estimated jointly they call it the J-test. It is to be noted here that the J-test is one degree of freedom tests irrespective of the number of explanatory variables in $H_0$ and $H_1$. 
APPENDIX - B

DESCRIPTION AND SOURCES OF DATA

Time series data on all the variables used in the empirical analysis are from 1970-1996 for the annual sample and from 1985 04 to 1996 09 for the monthly sample. For the extended sub sample 1994 04 1996 07, an additional asset called the certificate of deposit is included in the analysis. All the monetary variables are measured in crores of rupees at current prices. The monthly data have been deseasonalised by using the additive seasonal adjustment procedure.

B.1. Monetary Assets

Four financial assets viz., currency with the public (CU), net demand deposits with all commercial and co-operative banks (DD), net time deposits with all commercial and co-operative banks (TD) and post office savings bank deposits (PD) have been considered for both annual and monthly analysis. A new asset called certificate of deposit is also included in the analysis for the sample 1994 04 1996 07. The sample size has been constrained by the availability of data.

B.2. Interest rate variables

B.2.1. Own rate of return

Regarding rate of return on individual assets, currency being the most liquid amongst all assets, the rate of return on it is assumed to be zero. For demand deposits, an
implicit competitive rate is constructed using Klein's (1974) methodology. The formula employed for constructing this rate is

\[ r_{DD} = r_T \left[ 1 - \left( \frac{BR}{DD} \right) \right] \]

where \( r_{DD} \) is the implicit rate on demand deposits, \( r_T \) is 91 day treasuring bill rate and BR is bank reserve held against demand deposits, 12 month deposit rate is used as a proxy for time deposit rate. In case of postal savings deposits, the rate of interest per annum on post office savings bank accounts with limits of investment lying between a minimum of Rs 20/- and a maximum of Rs 50,000, is used. In the absence of such a rate commercial bank 3 months savings deposit rate is used. For certificate of deposit, its own rate of return is used.

B.2.2. Bench mark rate

Theoretically R, the benchmark rate should be the rate on a particular asset which is completely illiquid and does not provide any monetary service. The rate on human capital may serve the purpose since it does not render any monetary service in a world where there is no slavery system. However, due to difficulties associated with the availability of such an asset, the present study proxies R by taking the rate of return on the highest yielding asset in period t and calculates user costs for each asset in that particular time period. In the absence of any direct measurement of a benchmark rate, one can only construct proxy measures, which need not be rate of return on only one asset. Thus, rates of return like long term government bond yield rate, company deposit rate yield on private debentures and UTI dividend rates have served this purpose for different time periods in the study. The highest available rate from among these rates in a particular year is considered as the benchmark for that year.
B.3. Income Measure

For annual data gross domestic product at factor cost and for monthly data index of industrial production have been used as income measure. The series on index of industrial production has 1980-81 as its base.

B.4. Prices

The wholesale price index (WPI) (with base=1980-81) has been chosen as a measure of prices.

B.5. Scale and Opportunity cost variables

Real gross domestic product (gross domestic product deflated by WPI) for annual data and index of industrial production for monthly data have been chosen as "scale" variables in money demand tests. The weighted average of call money rates is used as the opportunity cost variable in simple sum money demand functions whereas the respective Divisia user cost aggregates have been used as opportunity cost variables in Divisia money demand functions.

SOURCES OF DATA

G O I, Ministry of Finance, Economic Survey (Annual)
R. B I, Reserve Bank of India Bulletin (Monthly)
R. B I, Report on Currency and Finance (Annual)