FORECASTING MONEY SUPPY AND EVALUATING ALTERNATIVE FORECAST PERFORMANCE USING BARBADIAN DATA

by

Kevin Greenidge
University of the West Indies
Cave Hill Campus

December 1995

Time present and time past are both perhaps present in time future and time future contained in time past.

T.S. Eliot

Introduction

This paper is concerned with applying different forecasting techniques to the money supply of the Barbadian economy and evaluating the efficiency/performance of those techniques. Time-series forecasting consists essentially of fitting a time-series model to historical data and then using the model to extrapolate into the future.

Time-series forecasting methods may be classified into (a) univariate (or extrapolated or projection) methods where forecasts are based only on past observations of the given series; (b) multivariate (or causal) methods where forecasts are based partly on observations of other explanatory variables.

In dealing with univariate techniques for forecasting, there exist a number of approaches, of varying degrees of complexity, whereby this end can be achieved. The more complex procedures do not produce forecasts as quickly as do the simple ones, but it is expected that a pay-off in terms of increased accuracy would be obtained through their use, since they allow for a more detailed investigation of the properties of the particular series under consideration.

Of course, one frequently possesses qualitative and quantitative information of potential relevance in addition to current and past values of the series under study and, where practicable such information ought to be incorporated into the forecasting mechanism. Nevertheless, we feel that univariate time series forecasting methods deserve detail consideration for a number of reasons. First, such methods are quick and inexpensive to operate, and may well produce forecasts of sufficient accuracy for the purposes at hand. Again, relevant extraneous information may be unavailable or only obtainable at a prohibitively high cost. Univariate forecasting procedures can be useful as a yard-stick against which the success or otherwise of a more elaborate forecasting exercise can be judged. They may also be usefully combined with other forecasting methods in the production of an overall forecast. Such combining can be achieved either on a formal or an informal basis. Finally, one can assess how much of the variation in a quantity can be explained in terms of its own past behaviour and so form a clearer understanding of what particular behaviour patterns require consideration of extraneous factors for their explanation. Thus, for example, the potential usefulness of other time series in the forecasting of a series of interest can be assessed by examining their ability to predict the univariate forecast errors.

The art and science of selecting an ARIMA model may give forecast results which are inferior to those produced by simpler exponential smoothing strategies. This can be due to a number of reasons such as model misspecification, simple error, insufficient data etc. Makridakis et al (1982) experienced mixed results in which no clear-cut "winner" emerged between smoothing and ARIMA models. Relative accuracy depended upon the presence of seasonality and trend, the forecast horizon, and other factors.

For many years, much time-series forecasting relied on the classical decomposition of a time-series into trend, seasonal, other cyclic variation and irregular variation. The trend was often assumed to be constant from year to year. Nowadays, local, rather than global assumptions are more widely favoured so that for example, the local mean at time t could be assumed to follow a random walk. Two large-scale empirical studies have found little difference in forecast accuracy between exponential smoothing and ARIMA models identified by the Box-Jenkins (1976) methodology - see Makridakis and Hibon (1979) and

Makridakis *et al.* Granger and Newbold (1974) found that Box-Jenkins produce superior forecast to the various exponential smoothing procedures and that the gains were significant.

Regression models are probably the most widely used class of multivariate models, and it should be realized that the phrase " econometric model " often means a single equation regression model. Their record is also patchy [e.g. Armstrong (1985, chapter 15)] although Fiddes (1985) finds that they perform better on average than univariate methods. More attention to the "error" structure is still needed (the errors are likely to be independent) and the use of diagnostic method [Belsey (1986)] seem likely to be further developed. The continuing difficulty of fitting multiple regression equations to economic data arises partly because there may be high correlations between the predictor variables, also the possibility of spurious co-movement between variables. Jevon (1884, p 3) alluded to the problem and Yule (1926) conducted the first formal analysis (Hendry 1986). Since, forecasting is preeminently a time-oriented exercise, it is only natural to expect that inadequate attention to time series concepts would lead to unnecessarily poor forecasts.

In our study, we review four variants of the exponential smoothing procedure, the Box-Jenkins methodology for ARIMA models, and regression models base on cointegration. We especially showed that combining individual forecasts is justifiable, and in all cases of the univariant techniques did produce forecast results which were superior to the individual forecasts, but in the case of the regression model, combining fail to improve the individual regression forecast. Bates and Granger (1969), Newbold and Granger (1974), Makridakis et al (1982) and Makridakis and Winkler (1983) have all found similar results in their studies. Our results also indicated that the Box-Jenkins methodology was not significantly superior to the other univariate procedures and at times there were equivalencies between them. A thorough review of these equivalencies is given by Grencher (1985), who argue that such occasions cannot be used as an argument in favour of the over densely populated class of ARIMA models.

Perhaps, one of the major ways in which this study differs from past studies is in the utilisation of cointegration theory and error correction models (ECMs) to specified a dynamic formulation of the money demand function. Past studies including Bourne (1974) , Howard (1981) and Ramsaran and Maraj (1985) are plagued with serial correlation in the residuals and as Granger and Newbold (1974) pointed out, serial correlation in the residuals constitutes evidence of misspecification . Moreover, most of these studies have followed the ad hoc practice of including as one of the regressors in the money demand function, a lagged dependent variable to deal with the problem of incomplete adjustment of money demand in the short run. Such actions are justified by the partial adjustment mechanism (Nerlove, 1972) or the adaptive hypothesis propounded by Cagan (1956) . However, it is well known that the partial adjustment mechanism constraints the adjustment pattern in the regressand to be the same regardless of the source of the initial disturbance (Laidler, 1982), whilst the adaptive expectation model , as noted by Muth (1960), is only optimal when the data generating process follows an autoregressive integrated moving average process of degree (0,1,1) . Recently, Hendry and Mizon (1978) and Hendry (1979), show that ECMs, which encompass models in both levels and differences, and are compatible with proportional long-run equilibrium behaviour, have less restrictive lag structures and nest the partial adjustment and adaptive expectations models. The ECM therefore circumvent the fundamental spurious regression problem discussed by Granger and Newbold (1974) through the use of appropriate difference variables in the model, but without losing long-run information due to difference data only. More support for the use of ECMs comes from the recently developed theory of cointegration. This theory allows estimation and inferences to be possible when economic variables do not satisfy the classical assumptions of constant mean and variance. If there exists a linear combination of thes non-stationary variables that is stationary, then the variables are said to be cointegrated. Engle and Granger (1987) prove that if a set of variables are cointegrated then there exists a corresponding error correction form of these variables.

Section I presents the exponential smoothing technique, section 2 the ARIMA technique, section 3 compares and evaluates the univariant procedures, section 4 develops a regression money demand function, and section 5 compares the regression with the univariant forecast models and concludes.

1. Exponential Smoothing

1.1 *Theory and Procedure*

Exponential smoothing was developed in the 1950's and is associated particularly with the names of R.G. Brown, and P.R. Winters. Many variations of this approach are now available [Gardner (1985)] including the Holt-Winter seasonal methods. Perhaps the most important reason for the popularity of exponential smoothing is the surprising accuracy that can be obtained with minimal effort in model identification. Exponential Smoothing is a method of forecasting based on a single statistical model of time series. It does not make use of information from series other than the one being forecast.

Exponential Smoothing in its various versions was developed in the last three decades as a practical tool for short-term forecasting. The procedure is relatively easy to apply and is often useful when a large number of forecasts are routinely needed; e.g. for inventory, control, production planning and marketing.

Suppose that the available data consist of a series of observations $x_1, x_2, \ldots, x_n$ and that it is required to forecast x_{n+h}, denoted by $F_{n,h}$. The earliest version of exponential smoothing, called "simple exponential smoothing" regards a time series as being made up locally of its level and residual element. The aim for such a series is to estimate the current level. This level estimate is then used as the forecast of all future values of the series. Hence it is appropriate for nonseasonal time series with no predictable upward or downward trend. The current level is estimated as a weighted average of previous observations. Specifically, geometrically (exponentially) declining weights, so that the level of the series at line t is estimated by,

$$L_t=\alpha X_t+\alpha(1-\alpha)X_{t-1}+(1-\alpha)^2X_{t-2}+\alpha(1-\alpha)^3X_{t-3}+\ldots \quad 0<\alpha<1 \qquad (1)$$

substituting t-1 for t in this expression and multiplying through by 1-α yields

$$(1-\alpha)L_{t-1}=\alpha(1-\alpha)X_{t-1}+\alpha(1-\alpha)^2X_{t-2}+\alpha(1-\alpha)_3X_{t-3}+\ldots \qquad (2)$$

subtracting this from (1) then yields

$$L_t=\alpha X_t+(1-\alpha)L_{t-1} \quad 0<\alpha<1 \qquad (3)$$

The original series x_t is replaced by a "smoothed" series L_t.

In order to employ this algorithm, it is necessary to make some "starting up" assumption, the simplest being $X_t = L_t$. The quantity α is termed the "smoothing constant", the value of which can be chosen either subjectively or objectively. A judgmental choice of the smoothing constant can be done through visual inspection of the observed data. In general, the more random the series the lower the values of the optimal smoothing constants. A more objective approach, proposed by Holt and Winters, is to select those values that would have best "forecast" the given observations. The difference between the actual value and the forecast is the forecast error, e_t. Hence, it is useful to choose a grid of values for the smoothing constant, say α=0.1, 0.2, ..., 0.9, calculate the corresponding error series, and then choose for subsequent use the value of α for which the sum squared one-step forecast errors

$$S = \sum_{t=3}^{n} e_t^2$$

is smallest. The forecast of all future values of the series are given by the latest available smooth value, so that

$$F_{n,h} = L_n \qquad (4)$$

The estimation procedure implied by (3) can be regarded as an updating mechanism, so that at time t the previous estimate of level L_{t-1} is updated in light of the new observation x_t. The new estimate of level L_t is then a weighted average of x_t and L_{t-1}.

The simple exponential smoothing algorithm yields constant forecast for all future rates of a time series. In some series, there exist information that allows the anticipation of future upward or downward movements. Hence, rather than a constant forecast function, some trending function would be preferable. The simplest possibility of this sort is a linear trend forecast function. It is not necessary that the time series exhibit a fixed linear trend; this will be rarely if ever, be the case for business and economic series. Instead, consider the possibility of evolving local linear trend over time. Holt (1957) and Winters (1960) developed an exponential smoothing algorithm that allows for local linear trend in a time series.

As before, an estimate of the current level of the time series is required. Also, an estimate of the current slope, or change in level of the series is needed. Holt's linear trend algorithm provides estimates of level and slope that adapt over time as new observations become available. The recurrence form of this algorithm is

$$L_t = \alpha X_t + (1-\alpha)(L_{t-1} + T_{t-1}) \quad 0 < \alpha < 1 \qquad (5)$$

$$T_t = \beta(L_t - L_{t-1}) + (1-\beta)T_{t-1} \quad 0 < \beta < 1 \qquad (6)$$

Equation(5) adjusts L_t directly for the trend of the previous period, T_{t-1}, by adding it to the last smoothed value S_{t-1}. This eliminates the lag and bring L_t to the approximate base of the current data value. Equation (6) then updates the trend, which is expressed as the difference between the two last smoothed values. This is appropriate because if there is a trend in the data, new value will be higher or lower than the previous ones. Since there may be some randomness remaining, it is eliminated by smoothing with β the trend in the last period (L_t - L_{t-1}), and adding that to the previous estimate of the trend multiplied by (1-β). Thus, (6) is similar to the basic forms of single smoothing, given by (3) but applies to the updating of the trend. Forecasts of future observations then follow from an assumption of a continued period-by-period estimate from a base provided by the latest level estimate. Thus, the forecast of X_{t+h} made at time t is

$$F_{t+h} = L_t + hT_t, \quad h = 1, 2, 3, \ldots \qquad (7)$$

A series may also have a seasonal component, a pattern which occurs from period to period. Consider a seasonal time series with period s (so that s=4 for quarterly data and s=12 for monthly data). The most commonly employed variant of the Holt-Winters method regards the seasonal factor I_t as being multiplicative (while trend remains additive) so that this quantity is estimated as

$$I_t = \gamma(X_t / L_t) + (1-\gamma)I_{t-s} \quad 0 < \gamma < 1 \qquad (8)$$

(It is assumed that $X_t > 0$ for all t.) The level L_t which can be thought of as a "deseasonalised" level, is estimated now by

$$L_t = \alpha(X_t/I_{t-s})+(1-\alpha)(L_{t-1}+T_{t-1}) \tag{9}$$

The trend component is again estimated using equation (6). In order to employ equations (6), (8), (9) it is again necessary to specify "starting up" values. A very simple way to accomplish this is to take

$$I_j=X_j/\left(\frac{1}{s}\sum_{k=1}^{s}X_k\right),\, j=1,2,\ldots s,$$
$$L_s=\frac{1}{s}\sum_{k=1}^{s}X_k,\ \text{and}\ T_s=0$$

The three updating equations are then used recursively for t=s+1, s+2, ..., n. Since trend is taken to be additive and seasonality multiplicative, forecasts of future values are given by

$$\begin{aligned} F_{n,h} &= (L_t+hT_t)I_{t-h-s} && h=1,2,\ldots s \\ &= (L_t+hT_t)I_{t-h-2s} && h=s+1,s+2,\ldots,2s \\ &\ \vdots \end{aligned} \tag{10}$$

One can easily modify the Holt-Winter approach to deal with situations in which the seasonal factor is thought to be additive rather than multiplicative. In this case equations (8) and (9) are replaced by

$$I_t=\alpha(X_t-L_t)+(1-\alpha)I_{t-s},\quad 0<\alpha<1 \tag{11}$$

and

$$L_t=\alpha(X_t-I_{t-s})+(1-\alpha)(L_{t-1}+T_{t-1}) \quad 0<\beta<1 \tag{12}$$

The forecasting equation (10) is now replaced by

$$\begin{aligned} F_{t,h} &= L_t+hT_t+F_{t+h-s} && h=1,2,\ldots.s \\ &= L_t+hT_t+F_{t+h-2s} && h=s+1,s+2,\ldots 2s \\ &\vdots \end{aligned} \tag{13}$$

1.2 *Evaluation of Forecasts*

Objective measures on forecast performance are based on aggregate quality over some period of time, the appropriate statistical aggregates depending on assured cost of errors function where the one-step ahead error is

$$E_{it}=X_t-f_{it} \quad t=1,2,\ldots,20$$

Two measures of forecast quality based on forecast errors, are computed in this study.

1. Assuming a quadratic cost error function, we calculate the sum of the square residuals/forecast errors. If a forecast error e_t are available, then

$$SSR = \sum_{t=1}^{n} e$$

2. Similar, a measure of average forecast quality is the Root mean squared error

$$R.M.S.E = \sqrt{\sum_{t=1}^{n} e_t^2 / n}$$

Another method used here to evaluate the forecast performance is to see how the forecast preforms as we increase the steps ahead. To do this we truncate the series early in its historical data and make n forecasts ahead, n = 1, 2, ..., n, and then compare these forecasts with the realised values, analysing whether the forecast errors grow as we move from one step ahead to two steps ahead and so on. This is out of sample forecasting. We are then able to compare the best in sample forecast technique with the out of sample best technique to see if they are the same.

1.3 Results

Using Micro TSP version 7.0, we applied the four variants of exponential smoothing described above to the quaterly series of the money supply collected from the Central Bank of Barbados annual statistical Digest. Looking at table 1.1, Our results indicates that Holt-Winters-additive Seasonal out perform the other three variants of the exponential smoothing procedures for in-sample forecasting.

Table 1.1

EXPONENTIAL SMOOTHING				
	Simple	no Seasonal	Additive Seasonal	Multiplicative Seasonal
Sum of Square Residual	6.79E+10	4.48E+10	3.73E+10	5.60E+10
Root Mean Square Error	30501.76	24767.88	22605.13	27104.92
End of Period levels: Mean	768675	771144.3	764596.2	767951.3
Trend		6792.945	8753.300	8753.300

The closest to Holt-Winters - additive seasonal is the Holt-Winters - no seasonal, while simple smoothing produced the worst results of the four variants. Chart A indicates that all the techniques do produce good fits. Chart B show that with out of sample forecasting, the forecast error increases as the lead periods increases, this is indicated by the divergency of the actual and various fitted values in the out of sample period.

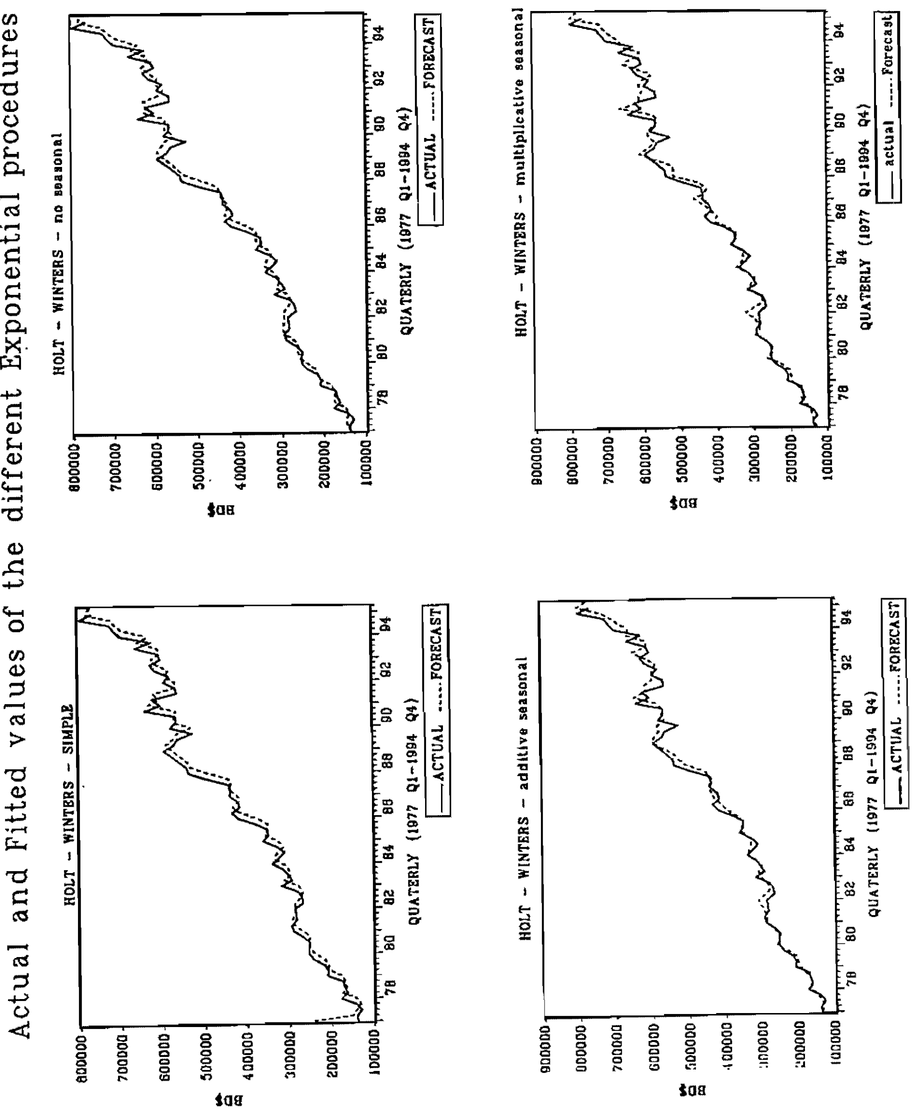
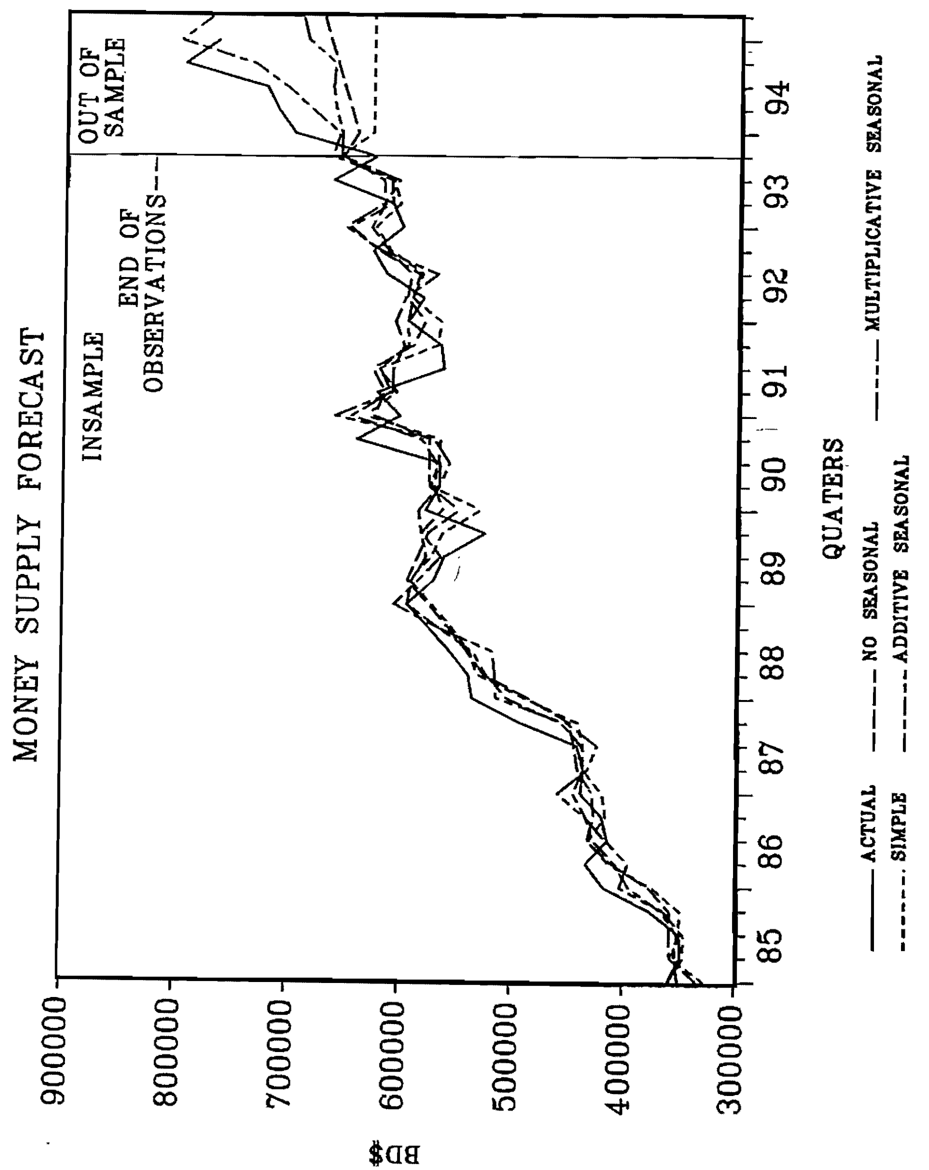


CHART A FORECAST 1
Actual and Fitted values of the different Exponential procedures

CHART B



2. ARIMA - Model (Box-Jenkins)

2.1 *Theory*

In this section we look at time series forecasting techniques based on autoregressive and moving averages processes. The autoregressive model had been proposed by G.U. Yule as early as the 1920s but it made little real impact on forecasting until the pioneer work of Box and Jenkins (1970) on ARIMA models. These techniques sometimes go by the names ARIMA or Box-Jenkins. Autoregressive integrated moving average processes (ARIMA) - consist of three components, an autoregressive component, an integrated component and an moving average component. We begin with an overview of these individual components. A time series is said to be governed by a first-order autoregressive process if the current value of the time series, X_t, can be expressed as linear function of the previous value of the series and a random shock a_t. This generating process of the data can be written as:

$$X_t = c + \phi X_{t-1} + a_t \tag{2.1}$$

where ϕ is the autoregressive parameter and a_t is assumed to have a zero mean and constant variance at all time periods t. It is further assumed that a_t is not autocorrelated, so that there is no correlation between a_t and the error in any other time period. Thus the a_t are taken to have the properties of the error terms in a regression equation where the standard assumptions for the regression model holds. In addition, the parameter C is included in the model to allow for the fact that the time series x_t can have a non-zero mean.

Hence, given the assumption about the first-order autoregressive model 2.1, the value x_t observed at time t, can be viewed as the sum of the two parts, a quantity $(c+\phi_1 X_{t-1})$ that will be known at time (t-1), and an unpredictable random term a_t, uncorrelated with anything known previously. Then, standing at time (t-1), since the random variable a_t has mean zero, the best forecast if x_t will be $(c+\phi_1 X_{t-1})$.

Equation (2.1) can be broadened to include more lagged variables. Suppose that we want to base forecasts on the two most recent observations. This suggests that specifications of the model

$$X_t = c+\phi_1 X_{t-1} + \phi_2 X_{t-2} + a_t \qquad (2.2)$$

where the past value X_{t-2}, with associated parameter ϕ_2, has been incorporated into the model. The formulation 2.2 is called a second-order autoregressive model. In general, a p^{th} order autoregressive model, AR(P), is written as

$$X_t = c + \phi_1 X_{t-1} + \phi_2 X_{t-2} + \ldots . \phi_p X_{t-p} + a_t \qquad (2.3)$$

Now, the time series x_t is said to be generated by an autoregressive model of order P[AR(P)], if

$$\dot{X}_t = x_t - \mu$$

and

$$\dot{x}_t = \phi \dot{X}_{t-1} + \phi_2 \dot{X}_{t-2} + \ldots . \phi_p \dot{X}_{t-p} + a_t \qquad (2.4)$$

where a_t is white noise. If we assume that the series is stationary, a requirement which imposes certain restrictions on the values that can be jointly taken by the autoregressive parameter, ϕ_1, ϕ_2, ... ϕ_p. In that case, it can be shown that x_t has a mean μ, and that the autocorrelations obey

$$\rho_k = \phi_1 \rho_{k-1} + \phi_2 \rho_{k-2} + \ldots + \phi_p \rho_{k-p} \quad k=1, 2, 3 \ldots \qquad (2.5)$$

Hence for a stationary autoregressive process, the plot of the autocorrelations decay towards zero as the lag k increases. The shape of this decay, which can be exponential, sinusoidal, or a mixture of the two, is indicated by equation 2.5

At this point, we introduce a back-shift operator B for convenience, where

$$BX_t = X_{t-1}$$

hence the autoregressive model of 2.4 can be written as

$$\dot{X}_t = \phi_1 B\dot{X}_t + \phi_2 B^2 \dot{X}_t + \ldots + \phi_p B^p \dot{X}_t + a_t$$

or

$$(1 - \phi_1 B - \phi_2 B^2 \ldots . . \phi_p B^p)\, \dot{X}_t = a_t$$

from equation 2.5 we can derive the potential autocorrelations of the autoregressive process. Suppose that the generating model of the time series is AR(k), and let ϕ_{k1},

ϕ_{k2}, ...,ϕ_{kk} denote the autoregressive parameters. Then it follows from equation 2.5 that

$$\rho_j = \phi_{k1}\ \rho_{j-1} + \phi_{k2}\ \rho_{j-2} + \dots\ \phi_{kk}\ \rho_{j-k} \quad j=1, 2, 3.\dots k \tag{2.6}$$

Now, for given values of the autocorrelations, equation 2.6 constitutes a set of k simultaneous linear equations in the k autoregressive parameters. These equations can therefore be solved for those parameters. For any process, the solution of equation 2.6 for the last parameter ϕ_{kk} yields the partial autocorrelation of order k of the process. Assume now that the true generating model is AR(p), where $P \leq k$. Clearly, if $P = k$, ϕ_{kk} will be the last autoregressive parameter, and will differ from zero. On the other hand, if $p<k$, we can think of the AR(p) model as an AR(k) model in which the last k-p parameters are zero. In particular, then, the partial autocorrelation ϕ_{kk} will be zero. This establishes the important conclusion that, for an AR(p) model, the partial autocorrelation of order p is non-zero, but all partial correlation of orders higher than p will be zero, that is

$$\phi_{kk}=0 \quad k=p+1,\ p+2, \dots.$$

Therefore knowledge of the partial autocorrelation of a process would provide a very easy way to recognise pure autoregressive behaviour of any order. A plot of ϕ_{kk} against the lag k has a very distinctive shape, involving an abrupt cut-off, with all values equal to zero at lags higher than the true autoregressive order p. In practice, we use the sample partial autocorrelation for this plot. Furthermore, it is known that, for moderately large sample size n, the sample partial autocorrelation of order greater than P for an AR(P) process have a distribution that is approximately normal with mean zero and standard error $n^{-1/2}$. To assess their significance, the partial autocorrelations are typically compared with $\pm 2n^{-1/2}$. This gives the analyst a good opportunity to detect pure autoregressive behaviour in an observed time series.

A time series X_t is said to be generated by a first-order moving average model if

$$X_t - \mu = X_{t} = a_t - \theta\ a_{t-1} \tag{2.7}$$

where θ_1 is a fixed parameter and a_t is white noise.

Since the white noise process has mean zero, it follows immediate from equation 2.7 that x_t has mean zero, and hence that x_t has mean μ. The variance of the process x_t is

$$Var(X_t) = E\ (X_t^2) = E\ (a_t^2) + \theta_2^2\ E\ (a_{t-1}^2) - 2\theta_1\ E\ (a_t a_{t-1})$$

since a_t is white noise, if r_a^2 denotes the variance of the white noise process

$$Var\ (X_t) = (1-\phi_1^2)\sigma_a^2$$

The first autocovariance of this process if $\Upsilon_1 = E\ (x_t x_{t-1}) = -\theta_1\ r_a^2$ and dividing this autocovariance by the variance it follows that the first autocorrelation is $\rho_1 = \dfrac{-\theta_1}{1+\theta_1^2}$

The other autocovariance follow in the same manner where $\Upsilon_k = E(x_t\ x_{t-k})$.

Now, for any $k>1$, it follows since a_t is white noise that each of the expectations on the right-hand side of this expression is zero, and hence that the autocovariance and the corresponding autocorrelation are zero. Hence for an MA(1) process, the autocorrelation of order one is non-zero, but all autocorrelations of order greater than one will be zero that is

$$\rho_k = 0 \quad k=2, 3 \ldots .$$

The correlogram for the MA(1) process will then be quite distinctive, with only zero values beyond lag one.

The MA(1) model can be generalised by allowing further lagged white noise terms. The time series x_t is said to be generated by a moving average model of order q, MA(1), if

$$x_t - \mu = \dot{x}_t = \mu_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} \;\ldots\ldots - \theta_q a_{t-q} \tag{2.8}$$

where $\theta_1, \theta_2, \ldots, \theta_q$ are moving average parameters and a_t is white noise. Clearly this process has mean μ. Also, arguing in exactly the same way as for the MA(1) process, it is quite straight-forward to show that, for an MA(q) process, the autocorrelation of order q is non-zero, but all autocorrelations of order higher than q will be zero, that is

$$\rho_k = 0 \quad k=q+1, q+2$$

The correlogram for a moving average process is quite distinctive, exhibiting an abrupt cut-off after lag q. This fact helps an analyst to recognize pure moving average behaviour and to determine the order of a moving average generating model.

For moving average processes, the partial autocorrelations do not cut off, but rather decay towards zero with increasing lag length. The contrasts in the behaviour of autocorrelations and partial autocorrelations for pure autoregressive and pure moving averages are very important, and worth repeating:

1. The autocorrelation, ρ_k of a pure autoregressive process decay towards zero with increasing lag length K. By contrast, for a pure moving average process of order q, the autocorrelation cut off abruptly - they are all zero for k bigger than q.

2. The partial autocorrelation, ϕ_{kk}, of a pure moving average process decay towards zero with increasing lag length k. By contrast, for a pure autoregressive process of order P, the partial autocorrelations cut off abruptly- they are all zero for k bigger than P.

The levels of many businesses and economic time series do not exhibit stationarity. Stationarity requires, among other things, that the series has a fixed mean over time, but often such a concept lacks credibility. Yet stationarity is an assumption on which AR and MA models are built. Fortunately, it turns out that these stationary models very often are appropriate for the representation of a simple transformation of business and economic times series. Although the levels of a times series may not be stationary. Frequently period-to-period changes, or first differences, of the series will be stationary. Thus, if the observed series is x_t, the series of changes

$$w_t = x_t - x_{t-1} = (1-B)X_t$$

will be stationary. It may therefore be reasonable to fit a stationary model to the first differences of an observed time series. The simplest case of this sort is where the first differences of a series are white noise, so that

$$X_t - X_{t-1} = a_t$$

This is known as a random walk model, where the time series does not appear to oscillate about a fixed mean, to whose neighbourhood it tends to return.

Sometimes, the series may require differencing a second time or even a third time. In general, a series may require differencing some number of times to induce stationarity, so that

$$w_t = (1-B)^d X_t \qquad (2.9)$$

is a stationary series. A process for which some differencing is required to achieve stationarity is called an integrated process. Suppose that an observed time series x_t has been differenced sufficiently to yield a stationary series w_t. We can then consider the possibility of fitting to w_t a stationary autoregressive moving average model. Often, after differencing, it is reasonable to assume that the differenced series has mean zero, so that an ARMA (p, q) model can be written

$$\phi(B)w_t = \theta(B)a_t$$

where

$$\phi(B) = 1-\phi_1 B-\ldots\ldots-\phi_p B^p$$

and

$$\phi(B) = 1-\phi_1 B-\ldots\ldots-\phi_q B^p$$

combining equations 2.9 and 2.10, the model of the original series x_t is

$$\phi(B)\ (1-B)^d X_t = \theta(B)\ a_t$$

This is an autoregressive integrated moving average model of order (p, d, q) i.e. ARIMA(p, d, q). Hence, p is the number of autoregressive terms, d the degree of differencing and q the number of moving average terms in the model.

Recognition that the data-generating process is integrated can be achieved through examination of the sample autocorrelation of a time series. For integrated processes, the simple autocorrelation r_k exhibit very smooth behaviour, rather than decaying quickly towards zero, for moderately high lags k.

We now introduce an extension of the above ARIMA class of models to take into account seasonality in times series.

Let x_t denote the time series of interest, observed with s periods per year, so that s=4 for quarterly data and s=12 for monthly data. Suppose that it is possible to find a linear transformation of the series yielding a non-seasonal series z_t. This transformation may be of the general ARIMA form, but with relationships between observations s periods apart. This transformation may involve seasonal differencing, so that we could compute

$$x_t - x_{t-s} = (1-B^s)\, x_t = Z_t$$

Thus, for monthly series, the seasonal differences will be made up of the change from one January to the next, the change from one February to the next and so on. Now, it is possible that this single seasonal differencing will be sufficient to yield a series z_t that is free from seasonality. However, generally this will not be the case. It may be that further seasonal differencing, or seasonal autoregressive or seasonal moving average terms will also be required. In general, we consider the possibility of D seasonal differences, P seasonal autoregressive terms and Q seasonal moving average terms, so that the nonseasonal series z_t is defined from

$$\begin{aligned}&(1-\Phi_1 B^s-\dots-\Phi_p B^{ps})\,(1-B^s)^D x_t\\ &=(1-\psi_1 B^s-\dots-\psi_Q B^{Qs})\; z_t\end{aligned} \tag{2.11}$$

where the Φ_i and ψ_i are fixed seasonal autoregressive and moving average parameters.

If the series z_t of equation 2.11 is non-seasonal, it could be represented by the regular ARIMA(p, d, q) model.

$$(1-\phi_1 B-\dots-\phi_p B^p)(1-B)^d Z_t = (1-\theta_1 B-\dots-\theta_q B^q)\; a_t \tag{2.12}$$

Amalgamating equation 2.11 and 2.12 then yields the model

$$\begin{aligned}&(1-\phi_1 B-\dots-\phi_p B^p)(1-\Phi_1 B^s-\dots-\Phi_p B^{ps})(1-B)^d\,(1-B^s)^D x_t\\ &=(1-\theta_1 B-\dots-\theta_q B^q)(1-\psi_1 B^s-\dots-\psi_q B^{qs})\,a_t\end{aligned} \tag{2.13}$$

This class of models are called multiplicative seasonal ARIMA (p, d, q,)(P, D, Q)s models.

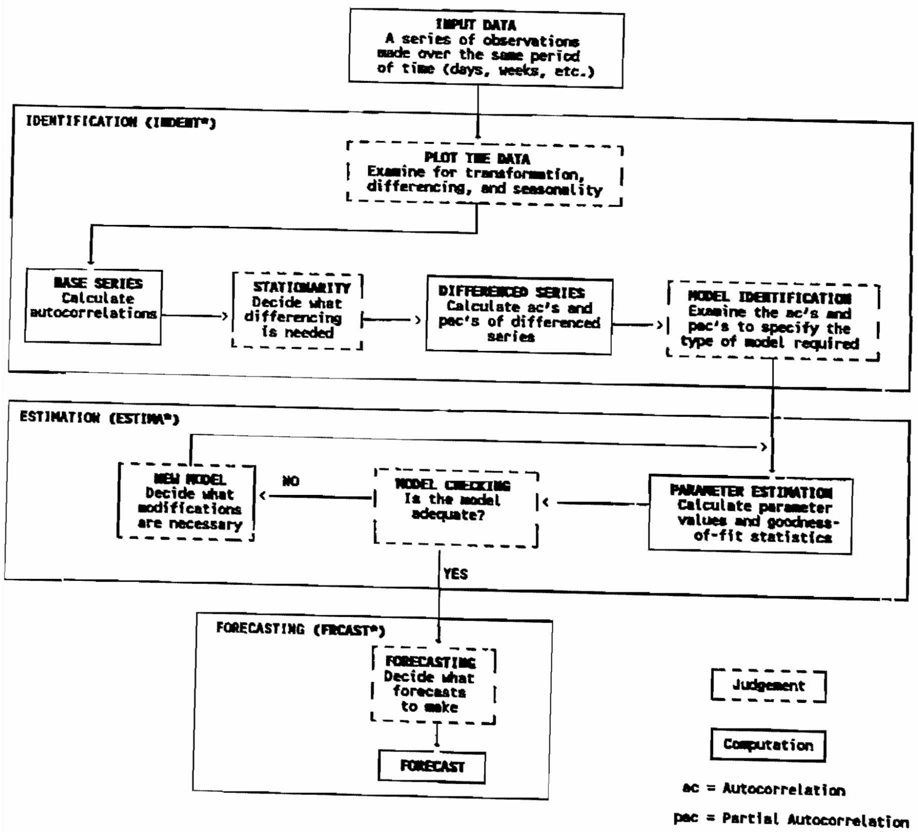
2.2 *Procedure*

We now present the Box-Jenkins (1970) iterative approach for constructing time series models. This approach basically consist of four steps:

a. *identification* of preliminary specifications of the model;
b. *estimation* of the parameters of the model;
c. *diagnostic* checking of model adequacy; and
d. *forecasting* future realizations.

Figure 2.1 contains a flow diagram of this approach.

Figure 2.1 Functional Diagram of the Box-Jenkins Approach



Model Identification

It is necessary to first decide on the amount of differencing required, then on the autoregressive and moving average orders of a model to be fitted to the differenced series.

1. To decide on the amount of differencing, the sample autocorrelation of the original series and some differenced versions are examined. It is usual practice to look at the sample autocorrelation of X_t, $(1-B)X_t$, $(1-B^s)X_t$, *and* $(1-B^s)X_t$, though on occasions further differencing of one type or another might be indicated. The necessity for further regular differencing will be suggested by smooth behaviour in the sample autocorrelation at moderate high lags. High sample autocorrelation at multiples of the seasonal periods indicate that further seasonal differencing may be desirable. An alternative method is to perform an Augmented Dickey-Fuller (ADF) test on the three differenced series and the undifferenced series to test for stationarity. The ADF test consist in runing a regression of the first difference of the series against the series lagged once, the lagged differenced terms, and optionally, a constant and a time trend . With the lagged difference terms, the regression is

$$\Delta y_t = \beta_1 y_{t-1} + \beta_2 \Delta y_{t-1} + \beta_3 \Delta y_{t-2} + \beta_4 + \beta_5 t$$

The output of the ADF test consist of the t-statistics on the coefficient of the lagged test variable and critical values for the test of a zero coefficient. If the t-statiatics is smaller than the reported critical values, the null hypothesis of nonstationarity is reiected and the series is accepted to be statationary.

2. The next step is to examine the sample autocorrelation and sample partial autocorrelations of the difference series -

$$W_t = (1-B)^d\,(1-B)^D X_t$$

As indicated in Figure 2.2.

These sample qualities are compared with two standard error limits, and a decision on p, q, P and Q can be made, i.e the type of model.

Figure 2.2

Autocorrelation and Partial Autocorrelation Functions

Model	**Autocorrelation Function**	**Partial Autocorrelation Function**
AR(P)	Tails Off	Cuts off after lag p
MA(q)	Cuts off after lag q	Tails off
ARMA(p, q)	Tails off	Tails off
AR(p), Seasonal. AR(P)	Tails off	Cuts off after lag p+sP
MA(q), Seasonal. MA(Q)	Cuts off after lag q+sQ	Tails off
Mixed Models	Tails off	Tails off

It may be best for the analyst to carry forward several models for comparison. Before proceeding, however, it is important to stress here the principle of parsimony. The search is not for a very elaborate model, but for the simplest model that appears to adequately represent the data.

Model Estimation

Assume now that a particular model from the general ARIMA class has been selected for fuller analysis. The parameters of the model can be estimated using least squares or maximum likelihood. This can readily be done on any statistically software packages. We use micro TSP in our analysis.

Model Checking

Two major approaches to model checking have traditionally been used. First, a more elaborate model - that is, one containing additional parameters - can be estimated, and the statistical significance of the extra parameters estimates checked. Thus, if an ARMA (P, q) model has been fitted, an alternative model with either one or two extra autoregressive terms or one or two extra moving average terms can be tried.

A second approach to model-checking is based on the fact that, if the model is correctly specified, the error terms a_t will be white noise; that is, all autocorrelations of these errors will be zero. In practice, the true errors a_t will be unknown, but they are estimated by the residuals from the fitted model. The autocorrelations of the errors are then estimated by the residuals autocorrelations.

$$r_K = \frac{\sum_{t=k+1}^{n} \hat{a}_t\,\hat{a}_{t-k}}{\sum_{t=1}^{n} \hat{a}_t^2} \qquad K=1, 2, \ldots$$

If the assumed model is adequate, their residual autocorrelations should be close to zero. They are typically compared with limits $\pm$ 2 [(n-k) / n(n+2)]½ or $\pm$ 2n -½ to assess their significance. The Ljung and Box (1978) statistic can also be used to test on the squares of the first M residuals autocorrelations, where M is a moderately large number - typically at lease ten; for economic data, M = 12 and M = 24 have proven to be useful.

The Box - Ljung statistic, which is a modification of an earlier proposal of Box and Pierce (1970), is

$$Q = n(n+2)\sum_{k=1}^{M} (n-k)^{-1} \hat{r}_k^2 \qquad (2.14)$$

It is known that, if an ARMA (p, q) specification is correct, the statistic 2.14 has a distribution close to the chi-square with (M-p-q) degrees of freedom. Model inadequacy would be indicated by large absolute values for the residual autocorrelation, and consequently large values for the statistic 2.14. Accordingly, the hypothesis that the assumed specification is correct would be rejected for a value of Q exceeding tabulated upper fail points of the chi-square distribution.

Forecasting

The general ARIMA (p, d, q) model can be written as

$$(1-\phi_1 B-\ldots-\phi-\rho B^p)[(1-B)^d X_t-\mu] = (1-\theta_1 B-\ldots\theta_q B^2)a_t$$

However, for forecasting, the autoregressive and differencing terms can be amalgamated, and the model is expressed as

$$(1-\Phi_1 B-\ldots-\Phi_{p+d} B^{p+d})X_t = C+(1-\theta_1 B-\ldots\theta_q B^q)a_t \qquad (2.15)$$

where

$$1-\Phi B-\ldots-\Phi_{p+d} B^{p+d} = (1-\phi_1 B-\ldots-\phi_p B^p)\ (1-B)^d$$

and

$$C = (1-\phi_1-\ldots-\phi_p)\mu$$

Hence, Form 2.15 of the model is then,

$$X_t = C+\Phi_1 X_{t-1}+\ldots+\Phi_{p+d} X_{t-p-d}+a_t-\theta_1 a_{t-1}-\ldots-\theta_q a_{t-q} \qquad (2.16)$$

Therefore, the future value X_{n+h} can be obtained by setting t = n+h in equation 2.16 the quantity to be predicted is then

$$X_{n+h} = C+\Phi_1 X_{n+h-1}+\ldots+\Phi_{p+d}X_{n+h-p-d}+a_{n+h}-\theta_1 a_{n+h-1}-\ldots-\theta_q a_{n+h-q} \qquad (2.17)$$

These calculations are done using Micro TSP, which also allows the computation of interval forecasts. The confidence intervals are based on an assumption that the white noise error terms a_t are normally distributed.

2.3 Empirical work and Results

We begin by determining the amount of differencing, if needed, through an examination of the sample autocorrelations. Exhibit 2.1 shows part of the output for the sample autocorrelation of Money Supply [M1], 1st differences of Money Supply [D(M1)], seasonal differences of Money Supply [D(M1, 0, 4)] and 1st difference and seasonal differences of Money Supply [D(M1, 1, 4)]. The sample autocorrelation of the undifferenced series behave quite smoothly at high lags, and, are very high for short lags and dies off very gradually for longer lags, suggesting the necessity for some differencing. Furthermore, the first partial autocorrelation is close to one, but the second is negative, and the rest are quite small. This pattern is sbsolutely typical of a series containing an integrated component.

Using the augmented Dicky-Fuller (ADF) test for a unit root or nonstationarity indicates that only D(M1) achieved stationarity at the 1% level (table2.1). Hence, in the case of Money Supply, first differences gives the kind of autocorrelations that can be fitted with an ARMA model.

Table 2.1 **TEST FOR STATIONARITY**

SERIES	**DF**		**ADF**	
	NO TREND	TREND	NO TREND	TREND
M1	-0.0793	-3.3028	0.3349	-2.3604
D(M1)	-11.0363	-11.0212	-7.4343	-7.4305
D(M1,0,4)	-3.4807	-3.5196	-2.0203	-2.0542
D(M1,1,4)	-3.5798	-3.7723	-3.0224	-3.0562
Mckinnon critical values:	-3.5226 -2.9017 -2.5879	-4.0890 -3.4720 -3.1629	-3.5239 -2.9023 -2.5882	-4.0909 -3.4730 -3.1635

The next step is to decide what kind of ARMA model to use. If the autocorrelation function dies off smoothly at a geometric rate, and the partial autocorrelations were zero after one lag, then a first-order autoregressive model would be suggested. Alternatively, if the autocorrelation were zero after one lag and the partial autocorrelations declined geometrically, a first-order moving average process would come to mind. Since neither of these cases covers our findings for D(M1), we have to consider more complicated models. We needed a model that will explain the significantly high lags in the autocorrelations and partial autocorrelation at lag 1,6,8,9 and 10.

An equation with MA and AR terms at these lags was estimated. Insignificant terms were eliminated one at a time and the model re-estimated until only significant terms remained. Hence, we obtain a parsimonious representation of D(M1),Exhibit 2.2, with MA terms lagged at the 6th,8th,9th and 10th quaters, and AR terms at the 1st and 6th quaters. Next we check the residuals to ensure that there are no remaining autocorrelation that our model has not accounted for. Exhibit 2.3 gives a plot of the autocorrelation and partial autocorrelation of the residuals. The Ljung-Box statistic indicates that we should accept the null-hypotheses of white noise and our model is correctly specified at the 96.97% confidence level. A further check of the residuals was done using the Augmented Dickey-Fuller test for non-stationarity. This indicates stationarity. This is therefore a reasonable stopping point in the search for an ARIMA model of the Money Supply series.

Chart C indicates that our model do give a good account of the data generation process of the money supply.

CHART C

ARIMA FORECAST OF M1 (1977Q1 - 1995Q1)

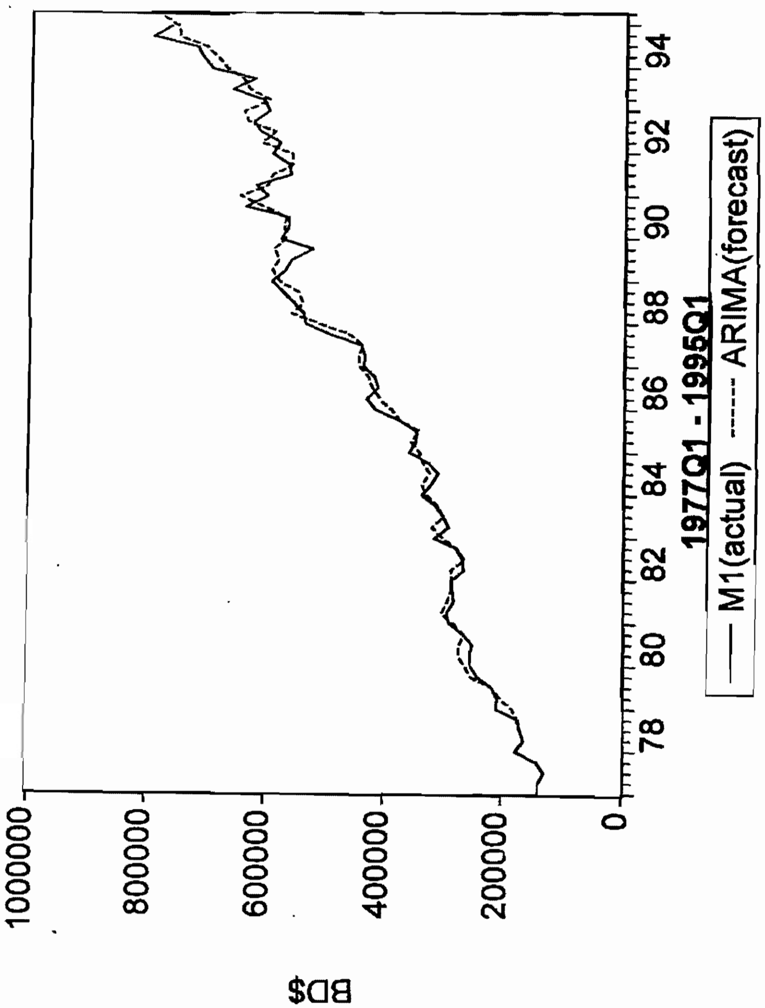


Exhibit 2.1

```
IDENT M1
Date: 9-15-1995 / Time: 14:57
SMPL range: 1977.1 - 1995.1
Number of observations: 73
==============================================================================
       Autocorrelations            Partial Autocorrelations        ac      pac
==============================================================================
|          .  |*************   |          .  |*************   |  1  0.945  0.945
|          .  |*************   |          .  |* .             |  2  0.890 -0.036
|          .  |************    |          .  |* .             |  3  0.842  0.048
|          .  |***********     |          . *|  .             |  4  0.793 -0.042
|          .  |***********     |          .  |* .             |  5  0.752  0.051
|          .  |**********      |          .  |  .             |  6  0.715  0.015
|          .  |**********      |          .  |  .             |  7  0.676 -0.033
|          .  |*********       |          .  |* .             |  8  0.644  0.047
|          .  |*********       |          .  |  .             |  9  0.611 -0.028
|          .  |********        |          . *|  .             | 10  0.570 -0.078
|          .  |********        |          .  |* .             | 11  0.537  0.044
|          .  |********        |          .  |  .             | 12  0.507 -0.001
==============================================================================
 Box-Pierce Q-Stat  453.64    Prob  0.0000        SE of Correlations  0.117
 Ljung-Box  Q-Stat  503.13    Prob  0.0000
==============================================================================
```

```
IDENT D(M1)
Date: 9-15-1995 / Time: 14:57
SMPL range: 1977.2 - 1995.1
Number of observations: 72
==============================================================================
       Autocorrelations            Partial Autocorrelations        ac      pac
==============================================================================
|         ****|  .             |        ****|  .              |  1 -0.283 -0.283
|          . *|  .             |          .**|  .             |  2 -0.040 -0.131
|          .  |* .             |          .  |* .             |  3  0.115  0.072
|          .**|  .             |          . *|  .             |  4 -0.149 -0.108
|          .  |**.             |          .  |* .             |  5  0.133  0.081
|          ***|  .             |          ***|  .             |  6 -0.228 -0.219
|          .  |  .             |          .**|  .             |  7 -0.023 -0.127
|          .  |***             |          .  |**.             |  8  0.229  0.148
|          .  |* .             |          .  |****            |  9  0.086  0.290
|         ****|  .             |        ****|  .              | 10 -0.306 -0.296
|          .  |**.             |          .  |  .             | 11  0.180  0.029
|          .  |  .             |          .  |  .             | 12 -0.033 -0.032
==============================================================================
 Box-Pierce Q-Stat   26.93    Prob  0.0079        SE of Correlations  0.118
 Ljung-Box  Q-Stat   30.48    Prob  0.0024
==============================================================================
```

```
IDENT D(M1,0,4)
Date: 9-15-1995 / Time: 14:58
SMPL range: 1978.1 - 1995.1
Number of observations: 69
==============================================================================
       Autocorrelations            Partial Autocorrelations        ac      pac
==============================================================================
|          .  |*******         |          .  |*******         |  1  0.523  0.523
|          .  |*****           |          .  |* .             |  2  0.353  0.110
|          .  |**.             |          .**|  .             |  3  0.132 -0.125
|         ****|  .             |      ******|  .              |  4 -0.279 -0.475
|          .**|  .             |          .  |****            |  5 -0.122  0.319
|          . *|  .             |          .  |***             |  6 -0.077  0.200
|          .  |  .             |          . *|  .             |  7 -0.014 -0.043
|          .  |**.             |          .**|  .             |  8  0.150 -0.179
|          .  |* .             |          .  |* .             |  9  0.097  0.076
|          .  |  .             |          .**|  .             | 10 -0.036 -0.126
|          . *|  .             |          .  |* .             | 11 -0.040  0.051
|          ***|  .             |         ****|  .             | 12 -0.254 -0.273
==============================================================================
 Box-Pierce Q-Stat   42.29    Prob  0.0000        SE of Correlations  0.120
 Ljung-Box  Q-Stat   45.89    Prob  0.0000
==============================================================================
```

IDENT D(M1,1,4)
Date: 9-15-1995 / Time: 14:58
SMPL range: 1978.2 - 1995.1
Number of observations: 68

Autocorrelations	Partial Autocorrelations		ac	pac
		1	-0.160	-0.160
		2	0.091	-0.044
		3	0.187	0.236
		4	-0.599	-0.549
		5	0.150	-0.357
		6	-0.035	-0.006
		7	-0.102	0.112
		8	0.238	-0.203
		9	0.067	0.005
		10	-0.133	-0.107
		11	0.253	0.230
		12	-0.267	-0.157

Box-Pierce Q-Stat	53.02	Prob	0.0000	SE of Correlations	0.121
Ljung-Box Q-Stat	59.35	Prob	0.0000		

Exhibit 2.2

LS // Dependent Variable is D(M1)
Date: 9-15-1995 / Time: 15:00
SMPL range: 1978.4 - 1995.1
Number of observations: 66
Convergence achieved after 13 iterations

VARIABLE	COEFFICIENT	STD. ERROR	T-STAT.	2-TAIL SIG.
C	9817.7333	1522.1256	6.4500151	0.0000
MA(6)	0.3295718	0.0694469	4.7456661	0.0000
MA(8)	0.4949121	0.0819446	6.0395931	0.0000
MA(9)	0.3648553	0.0841723	4.3346255	0.0001
MA(10)	-0.3383207	0.0744308	-4.5454390	0.0000
AR(1)	-0.3367979	0.1094247	-3.0778974	0.0032
AR(6)	-0.4929570	0.1230430	-4.0063805	0.0002

R-squared	0.365388	Mean of dependent var	9056.647
Adjusted R-squared	0.300851	S.D. of dependent var	26875.37
S.E. of regression	22471.88	Sum of squared resid	2.98E+10
Log likelihood	-751.2714	F-statistic	5.661687
Durbin-Watson stat	1.839434	Prob(F-statistic)	0.000107

Exhibit 2.3

IDENT RESID
Date: 9-15-1995 / Time: 15:01
SMPL range: 1978.4 - 1995.1
Number of observations: 66

Autocorrelations	Partial Autocorrelations		ac	pac
		1	0.074	0.074
		2	0.013	0.008
		3	0.111	0.110
		4	-0.213	-0.233
		5	0.099	0.145
		6	0.072	0.037
		7	-0.143	-0.113
		8	0.025	-0.027
		9	-0.079	-0.040
		10	-0.024	0.035
		11	0.131	0.066
		12	-0.043	-0.032

Box-Pierce Q-Stat	8.27	Prob	0.9878	SE of Correlations	0.123
Ljung-Box Q-Stat	9.36	Prob	0.9697		

2.4 *Evaluation of forecast*

At this point we evaluate the forecasts derived from the selected Arima model as compare to those derived from the previous four variants of exponential smothing. Comparing table 1.1 and exhibit 2.2, we find that the arima technique give better forecast results over the entire period, in terms of having both the smallest Sum of Square Residuals (2.98E+10) and the smallest Root Mean Square Error (21246.79).

A further investigation of the performance of the various techniques in out of sample forecasting is indicated by table 2.1 which shows the percentage of times one method outperforms another for forecasts made up to eight steps ahead. We excluded from this analysis the simple and multiplicative seasonal techniques since they both produced inferior results. Note that the Box-Jenkins approach is superior to both of Holt-Winters methods for short lead time, but some of this advantage is lost when forecasting further ahead. This is probably a reflection of the fact that for many non-seasonal time series information available at time t is a highly relevant determinant of the immediate future, but becomes less relevant for more distant time periods. This is similar to the results obtained by Newbold and Granger (1974) and Payne (1973).

Table 2.1

Comparison of Box-Jenkins (B-J), Holt-Winters-additive seasonal [HW(As)], Holt-Winters-no seasonal[HW(Ns)] forecasts: Percentage of times first named method outperformed second for various lead time. In terms of Room Mean Squared Errors.

Comparisons	Lead Times (Months)							
	1	2	3	4	5	6	7	8
B.J: HW(As)	80	79	65	60	48	45	50	40
B-J: HW (Ns)	75	75	70	55	50	50	55	50
HW (As): HW(S)	58	60	48	45	55.	60	55	70

To get some stronger feeling for the possible gains to be obtained from use of the Box-Jenkins approach, we considered the one-step-ahead forecasts in more detail. Table 2.2 presents two test statistics to indicate if one method is significantly superior to the other. The first statistic is a test of the difference of means, the null hypothesis being that the difference between the mean of the forecast error is not significantly different from zero. The second statistic is test for zero correlation between the sum of the forecasts errors (P_n) and the difference of the forecasts errors (Q_n). Given two alternative forecasts $f_{n,1}^{(1)}$ and $f_{n,1}^{(2)}$ with corresponding one step forecast errors

$$e_{n,1}^{(1)} = x_{n+1} - f_{n,1}^{(1)}$$

and

$$e_{n,1}^{(2)} = x_{n+1} - f_{n,1}^{(2)}$$

then

$$P_n = e_{n,1}^{(1)} + e_{n,1}^{(2)}$$

and

$$Q_n = e_{n,1}^{(1)} - e_{n-1}^{(2)}$$

Granger (1989) suggests that if any correlation is found between P_n and Q_n, then this implies a significant difference in RMSE values.

Table 2.2
Comparison of B-J, HW (As) and HW (S) forecasts: calculated difference of means t-Statistics and correlation value for one-step-forecast errors

	Money Supply
B-J: HW (As)	0.4253 -0.0425
B-J:HW(Ns)	0.0375 -0.0639
H-W (As): HW(Ns)	0.0467 0.1178

Both statistics when comparing the different forecasting methods indicates no significant difference in the Root Mean Squared Errors. These results along with those from Table 2.1 implies that although one forecast method might out perform another method by a certain percentage of times, it may not be significantly superior. This strongly suggests that the combined forecast may be worthwhile investigating.

3. *Regression Model*

3.1 Model

So far only methods of forecasting the money supply from its past values have been considered. We now examine forecasting the money supply given other explanatory variables in the Barbadian economy. Also, employing the recent econometrics development of co-integration that deals with the problem of spurious regression. In this study the demand for money is taken as

$$LnM_1 = \beta_o + \beta_1 LnCPI + \beta_2 LnDTODD + \beta_3 LN3TDR + \beta_4 LnPLR \qquad (3.1)$$

Where M_1 is the quantity of Nominal Money Balances demanded, CPI is the consumer price index, DTODD is Debits to Demand Deposits, used as a transaction proxy, 3TDR is the 3 months deposit rate, and PLR in the Prime Lending Rate. Prior expectations of the signs of the coefficients of all the variables are positive with the exception of the interest rate on 3 months deposits, which is expected to be negative.

3.2 Econometrics Methodology

Much of classical inference is predicted on the assumption that economic time series are co-variance stationary, i.e. the series have finite second moments, and the mean and co-variance structure of the data do not change across observations. However, as is well known, most economic time series of interest exhibit stochastic non-stationarities, thus undermining the foundations of traditional inference. Regressing one trended variable on another may lead to the erroneous inference of a significant relationship where none exists, a problem Granger and Newbold (1974) refer to as the spurious regression problem.

Following Granger (1981), define d as the order of integrability of a time series where d represents the number of times the series must be differenced to induce a stationary ARMA representation. Co-integration can be defined formally as follows: all component variables of a vector x_t are said to be cointegrated, of order d,b i.e. ($x_t \sim CI(d,b)$) if (i) all components of $x_t \sim I(d)$ and (ii) there exists a vector of $a \neq 0$, so that $Z_t = a' x_t \sim I(d-b)$, $b>0$; the vector a is then called the cointegrating vector (Engle and Granger, 1987)[1]. When the dimensions of $x_t > 2$, a is not necessarily unique, as there are possibly several cointegration vectors, some of which may be linearly dependent. In general for an N-vector x_t, there could exist a matrix A of dimension Nxr such that the r-vector $z_t = A' x_t \sim I(d-b)$.

The co-integrating vector describes a static, long-run relation. The problem of modelling short-run dynamics is solved by the Granger representation theorem which states that if a set of variables are co-integrated CI(1,1), then there is a corresponding, valid ECM form of those variables (Engle and Granger, 1987). This makes it possible to incorporate the notion of a steady state in a dynamic model. More formally, for N linearly independent co-integrated vectors, d=b=1, there exists an error correction representation of N stationary random variables x_t, if one can write

$$B(L)(1-L)x_t = -G Z_{t-1} + v(L)\epsilon_t \qquad (3.2)$$

where B(L) is a finite order polynomial with $B(o)=L_N$ and v(L) is a finite order lag polynominal; ϵ_t is a stationary multivariate disturbance and G is a matrix of dimension rxN. As 3.2 contains only stationary variables, there is no problem with employing

[1] The above asserts that two variables with different orders of integration cannot be cointegrated, although when there are more than three series, mixtures of different-order series are possible (Hall and Henry,1985).

the usually inference. The Granger representation theorem thus offers a sound theoretical rationale for employing ECMs when the level terms form a co-integrating set; it shows that if the data generating process is an equation such as 3.2, then x_t must be a cointegrating set of variables.

Engle and Granger (1987) demonstrate that if OLS is employed to estimate the parameters of the co-integrating equation, then the parameters of the ECM can be consistently estimated if the first stage estimates of the co-integrating equation are imposed on a second stage ECM. This 'control' is implemented by including the lagged error terms from the co-integrating regression in a general ECM[2]. This procedure is referred to as Granger - Engle two-step procedure. Stock (1987) shows that the order of convergence of these first stage OLS estimated is $O(T^1)$ compared with $O(T^{0.5})$ in the standard case; this faster convergence in the non-stationary case is sometimes referred to as 'super-consistency'. This fast rate of convergence of the estimator of the co-integrating parameter vector means that the estimators of the short-run parameters are asymptotically independent of the co-integrating vector. However, note that although the OLS estimates of the long-run parameters are consistent, there are not efficient (see Phillips, 1991). The second-stage standard errors are 'correct', that is, consistent for the true stantard errors, and the standard t- statistics can be used for the short-run estimates (stock, 1987:, Engle and Granger, 1987). The full ECM thus derived escapes the chance of a 'spurious regression'.

Since testing for a co-integrated vector requires the series to be integrable of order d >0, we proceed in the following way: (i) investigate the temporal properties of the variables in the static "Long-Run' equilibrium formulation of interest; (ii) test for a vector of co-integrated variables; (iii) estimate the ECM, and (iv) test the adequacy of the resulting equation using standard diagnostic tests and encompassing principles.

[2] The lagged residuals represent a level term which acts to dampen short-run deviations and return the system towards the steady-state equilibrium values; 'error correction'. An ECM can be developed from such economic considerations as incomplete information or cost of adjustment (Engle and Granger, 1987).

3.3 *Emperical Results*

Table 3.1 contains the results of the test for the order of integrability of several series proposed for inclusion in the long-run static regression or later in the ECM. The test for stationarity in the levels of variables is given by

$$\Delta X_t = \alpha + \beta_t + \delta_0 x_{t-1} + \sum_{j=1}^{J} \delta\ \Delta x_{t-j} + \epsilon_t \tag{3.3}$$

J is chosen to be sufficiently large to ensure that the error term is free of significant serial dependence. When J = o, the Dickey-Fuller test obtains and $J \neq 0$ defines the augmented Dickey-Fuller test. The null hypothesis that x_t follows a random walk is rejected if the coefficient on x_{t-1} is significantly negative.

Table 3.1

TEST FOR STATIONARITY: 1977: i - 1994: iv

Variable(X)	DF TEST		ADF TEST	
	Levels	1st - Differences	Levels	1st - Differences
LnM1	3.12	-8.564	4.153	-4.147
LnCPI	7.430	-4.115	3.039	-2.646
LnDTODD	0.443	-12.355	0.871	-6.491
LnMAX3TDR	-0.203	-8.074	-0.2111	-4.31
LnMIN3TDR	-0.4381	-7.631	-0.6011	-4.057
LnMAXPLR	-0.163	-5.837	0.096	-3.520
LnMINPLR	-0.203	-8.207	0.149	-3.897

Mackinon critical values: 1% -2.5954
5% -1.9449
10% -1.6181

Since the DF tests may lose power when the i.i.d. assumption is invalid - see Phillips (1987) - the residuals (ϵ_t) are tested for serial correlation using a Lagrange multiplier test and a variant of white's 1980 test for heteroscedasticity. For the ADF tests, we employed different lag structures as was necessary to eliminate excess serial cornelation. For the levels of the series, none rejects the null hypothesis of nonstationarity at the 5%, or even at the 10% level. After first differencing each of the series rejects the null hypothesis of nonstationarity at the 1% level. Hence the results of the unit root tests indicates that all variables considered are integrated of order 1. Consequently, first differencing is required for stationarity. Next we investigate whether equation 3.1 is a co-integrated set or not. To asertain whether the null hypothesis of no co-integration is rejected, we check to see whether the OLS residuals from Equation 3.1 are I(o). Since there exist the possibility of non-uniqueness of the co-integrated parameters, in all the co-integration regressions that follow, we examined all possible co-integration regressions and reported the one with the highest adjusted coefficient of determination. Such a procedure has been used by Hall (1986), since a high $\bar{R}^2$ minimises the potential bias in the estimate of the co-integration parameter (see Bernerjee et al. 1986). In every case, the left-hand - side conditioning variable turns out to be the natural logarithm of the money stock.

The results of the co-integration regressions are shown in table 3.2. Standard errors of coefficients are not reported for the estimated cointegration parameters since test statistics are not well developed. Further, interpretation of the magnitude of the co-integration parameters can be problematic. Benerjee et al. (1988), report that the small sample bias in estimating the co-integration parameters diminishes as the $\bar{R}^2$ value approaches one. Since the reported co-integration regressions has $\bar{R}^2$ values greater than 0.96 this may indicate that the bias is small in our case. D-W is the co-integrating regression Durbin-Watson statistic due to Sargan and Bhargava (1983). The D-W statistic can be used to give a rough indication as to whether there is co-integration. The value 0.97 would seem to indicate co-integration[3]. The more formal tests, the DF and ADF tests, both rejects the null hypothesis of nonstationary errors in the cointegration regressions equations v, vi and viii at the 5% level. The ADF test do not reject the null hypothesis of nonstationary errors in the cointegration regressions equations i, ii, iv and viii, and only rejects the null hypothesis in equations iii at the 10% level. Of the three equations which pass both the DF and ADF test for cointegration only equation vi has the correct sign for the lending rate.

Having achieved a suitable specification of the cointegration equation, we can proceed to the second stage of the Granger-Engle procedure. This involves regressing the first difference of dependent variables in the cointegration equation onto lagged values of

[3] This is at the 5% level. See Engle and Granger (1987, Table 2). It should be noted that the D-W test has low power to reject the null of no co-integration (a unit root) against alternatives close to the unit circle although an argument can be made for its use on the grounds that its distribution is invariant to nuisance parameters such as a constant (Banarjee et al, 1986).

the first-differences of all the variables plus the lagged value of the error-correction term. First we estimate an error-corection model with four lags of each variable, eliminating lags whose coefficients are insignificant, and re-estimating the simpler model. This is following the 'general-to-specific' modelling methodology (see for example Hendry and Mizon, 1978, and Miller 1991). These results are reported in table 3.3.

TABLE 3.2

COINTEGRATION REGRESSION: 1977:1 - 1994:IV

	COEFFICIENTS OF:										
Variable	CONST	LnCPI	LNDTODD	LnMAX3TDR	LnMIN3TDR	LnMAXPLR	LnMINPLR	R2	D-W	DF	ADF
i) LnM1	4.81	1.286	0.27	-0.222				0.97	1.09	-5.09	-3.5111
ii) LnM1	4.15	1.216	0.32		-0.085			0.97	1.00	-4.6	-3.3644
iii) LnM1	5	1.35	0.27			-0.35		0.97	1.05	-4.97	-4.-13
iv) LnM1	4.6	1.32	0.30				-0.306	0.97	1.04	-4.94	-3.63
v) LnM1	4.97	1.32	0.26	-0.142		-0.139		0.97	1.05	-4.99	-4.03
vi) LnM1	4.93	1.04	0.25	-0.297			0.136	0.97	1.07	-5.06	4.19
vii) LnM1	5.36	1.42	0.25		0.099	-0.58		0.97	1.07	-5.043	4.31
viii) LnM1	4.71	1.33	0.30		0.023	-0.539		0.97	1.04	-4.95	-3.64

McKinnon critical values: 1% -4.4031
5% -4.0255
10% -3.9292

Table 3.3

Dependent Variable is DLnM1
No. of observations: 6.8
SAMPLE 1977Q1 - 1995Q2

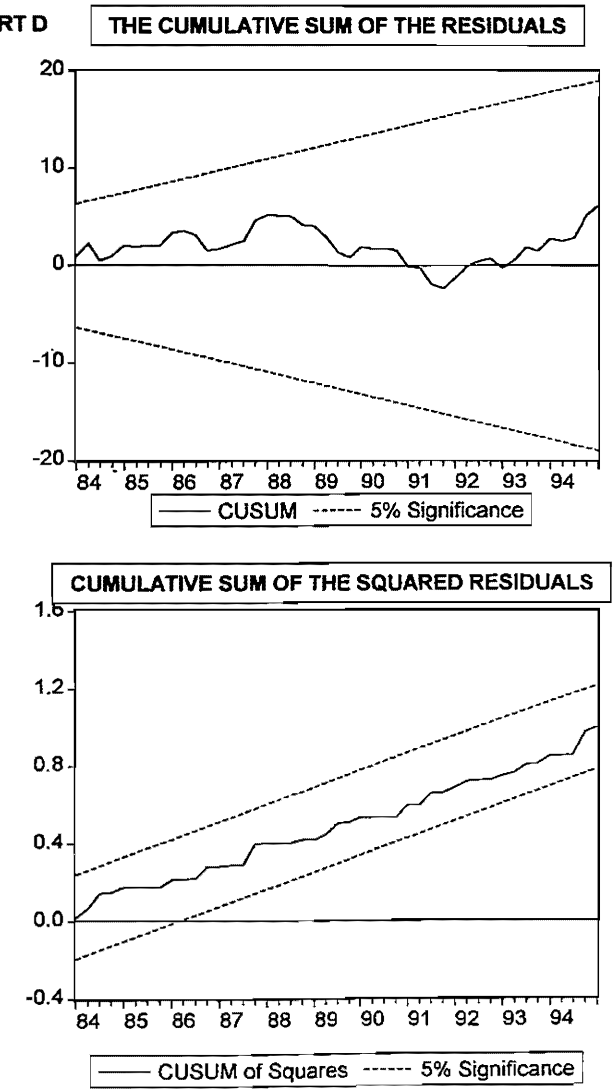
VARIABLE	COEFFICIENT	STANDARD ERROR	T-STATISTICS	2-TAIL SIG.
C	0.01631	0.00886	1.8407	0.0711
DLnTODD(-1)	0.11476	0.03371	3.2800	0.0018
DLnTODD(-4)	0.12490	0.03371	3.7046	0.0005
DLnCPI(-1)	0.77198	0.42905	1.8663	0.0829
DLnDCPI(-4)	-0.80869	0.38222	-2.1157	0.0309
DLnMAX3TDR	-0.17661	0.04845	-3.6443	0.0006
DLnMAX3TDR(-1)	0.11304	0.05859	1.9302	0.0588
DLnMAX3TR(-4)	0.23269	0.06234	3.7322	0.0006
DLnMINPLR(-1)	-0.20564	0.08581	-2.7596	0.0206
DLnMINPLR(-2)	0.19166	0.06956	2.7596	0.0078
DLnMINPLR(-4)	-0.16976	0.08853	-1.9175	0.0604
DLnM1	0.27667	0.08962	3.0872	0.0032
EC(-1)	-0.23401	0.07820	-2.9923	0.0041

R-squared	0.546462	Dubin Watson Stat:	2.0279
Adjusted R-squared	0.447509	S.S.R.	0.10453
F-Statistics	4.53321	Prob(F.Statistic)	0.0004
ARCH TEST:- F-stat	0.289190		
Prob (F-stat)	0.9930	Reset F.Stats	0.33933

The correlation matrix (not shown) reveals that the explanatory variables are nearly orthogonal variables, suggesting that multicollinearity is not problematic. The $\bar{R}^2$ is not high, but a high is not a precondition for a well specified model, an interesting example is provided by Hendry [1979, equation 4 and 20]. The normality test, Jarguq-Bera test statistic, indicates the residuals are normal but the implied marginal significance of the test is 0.114. Further investigations however showed that the movements of the scaled residuals (skewnes 0.084 and Kurtosis 2.761) are not significantly different from those of a standard normal distribution. The final data coherence check was the RESET test for general functional form. The low F-statistic leads us to accept the null hypothesis of no mis-specification of the functional form.

The stability of the model was checked over the period 1990-94 using the chow test. The low F-statistic indicates stability. An alternative test, the Cusum square plot of Brown et al (1975), suggests no evidence of significant parameter change as it does not break its 95% confidence interval (chart D). Hence we conclude that the model is stable.

CHART D



Wendel McClean (1995) results indicated that within the context of a small open economy the price, transaction variable, deposit rate and lending rate are not determined jointly with the demand for money. If otherwise, then the OLS estimates in the model will be inconsistent. The Wu-Hausman statistic is (WUT) is used to test the interpendence betwen the regessor set and the error term. The implementation of the test requires the regression of each of the current first differenced regressor terms in table 3.3 on a set of instruments such as lagged regressor from the ECM as well as lagged levels of variables, say from the static regression. The residuals from these regressions are saved and an F-test for their inclusion in the ECM is conducted. A significant test statistic implies rejection of the null hypothesis of weak exogeneity (see Engle at al; 1983 for an example). Consequently, the null hypothesis of weak exogenity is accepted as the estimated value of 2.01 lies well below the critical value of 3.25. Based on this procedure, the null hyopothesis of 'exogeneity' cannot be rejected at the 5% level of significance. Also, although they are indirect evaluations, the Chow and Cusum square statistics also test weak exogeneity, and thereby valid conditioning, which is not rejected by the data.

Based on the battery of diagnostic test, we are prepared to accept Table 3.3 as a well specified error correction model in the statistical sense.

Having being satisfied as to the existence of a stable error-correction model we now relaxed the restriction imposed by the ECM, this is mainly for forecasting purposes. This involves us first replacing the lag residuals by the logarithms of the arguments of the money demand function and the level of the dependent variable lagged one period. We then reestimate the general unrestricted model (GUM) with four lags as before for the period 1977Q1 to 1995Q1, leaving the remaining years for forecast comparisons. We then reduced the generalized unrestricted model by the variable elimination process outlined above. Table 3.4 reports estimates for the final and fully data accepted model.

Table 3.4
Dependent Variable is DlnM1
SAMPLE 1977Q1 - 1989Q4:

Variable	0	(-1)	(-2)	(-3)	(-4)
C	1.45214				
LnM1		-0.24688 (0.10873)			
LnCPI		0.42467 (0.16271)			
LnDTODD		0.14753 (0.03933)			
LnMAX3TDR		0.27195 (0.08922)			
LnMIMPLR		-0.93351 (0.16049)			
DLnDTODD			-0.07567 (0.03331)	-0.14222 (0.03573)	
DLnCPI		1.08058 (0.45258)	1.17294 (0.38506)	0.75007 (0.39695)	
DLnMAX3TDR	-0.22368 (0.05986)	-0.2625 (0.08607)	-0.38339 (0.07203)	-0.22414 (0.06848)	
DLnMINPLR	-0.18454 (0.09961)	0.42141 (0.11606)	0.73061 (0.11247)	0.44579 (0.10975)	0.23350 (0.0729)
DLnM1		-0.48055 (0.12966)	-0.4618 (0.11399)	-0.29257 (0.10592)	

R-Squared	0.72932	Durbin Watson	1.83606
Adjusted R-squared	0.596989	S.S.R	0.062389
Log-Likelihood	141.30	Prob(F-Stat.)	0.000001
F-Statistic	5.511296	Reset F-stat:	1.33382
ARCH Test(F-Stst.)	0.65890	Chow-Forecast Test F-Stat:	0.44632
Serial Cor. LM Test: (F-Stat.)	0.42561		

This model is also well accepted by every test and showed a much improved R-squared.

3.4 Forecasting

In comparing the regression model with the models which were based on univariant techniques, we confine our analysis to a smaller and more recent sub-sample, this being after 1990Q1. This is because we believe that stretching our analysis over the entire sample period (1977Q1 - 1995Q2) will produce results which are bias against the regression model.

First we reestimate the univariant models over the shorter sample period (1977Q1 - 1989Q4). This involves constructing an ARIMA model for this period since the pervious model may not be optimal in thie case. Table 3.5 shows the new optimal ARIMA model.

Table 3.5
Dependent Variable is D(LM1)
Sample: 1979Q2 1989Q4
Included observations: 43 after adjusting endpoints
Convergence achieved after 9 iterations

Variable	Coefficient	Std. Error	T-Statistic	Prob.
C	0.018368	0.004478	4.101692	0.0002
AR(6)	-0.332388	0.104306	-3.186658	0.0028
AR(8)	0.528243	0.105841	4.990907	0.0000
MA(8)	-0.885139	0.035906	-24.65177	0.0000

R-squared	0.545163	Mean dependent var	0.021405
Adjusted R-squared	0.510175	S.D. dependent var	0.052306
S.E. of regression	0.036608	Akaike info criterion	-6.526586
Sum squared resid	0.052265	Schwartz criterion	-6.362753
Log likelihood	83.30724	F-statistic	15.58166
Durbin-Watson stat	2.062166	Prob(F-statistic)	0.000001

Table 3.6

Forecast Evaluation (1990Q1 - 1995Q2)

Technique	Root Mean Squared Error	Sum Squared Error
H-W(S)	0.056134	0.056134
H-W(AS)	0.050122	0.052766
H-W(MS)	0.050112	0.052736
H-W(NS)	0.052422	0.057709
B-J	0.049289	0.039501
Regression	0.029732	0.029732

Table 3.6 shows that the regression forecast results are superior to those of the univariant models in terms of the Root Mean Squared Error and the Sum Squared Error statistics. the Box - Jenkins forecast ranked second in terms of forecast performance. There were not any significant difference between Holt-Winters-additive seasonal and Holt-winters-multiplicative seasonal.

Chart E compares the ARIMA forecast and the Regression forecast against the actual level of the money supply. The chart indicates that the regression model is forecasting with more accuracy the turning points of the money supply than the ARIMA model, for example, between 1993Q1 and 1994Q3 the ARIMA forecast continue to show an upward trend while the actual money supply fluctuated, this fluctuation was to a greater extent captured by the regression forecast. Hence, we concluded that the regression forecast did indeed produce forecast results which are superior to the different univariant procedures.

Table 3.7 gives the forecasts results of both the univariant procedures and the regression model for the period 1990Q1 - 1995Q2. Next explored whether combining the forecasts will improved their performance.

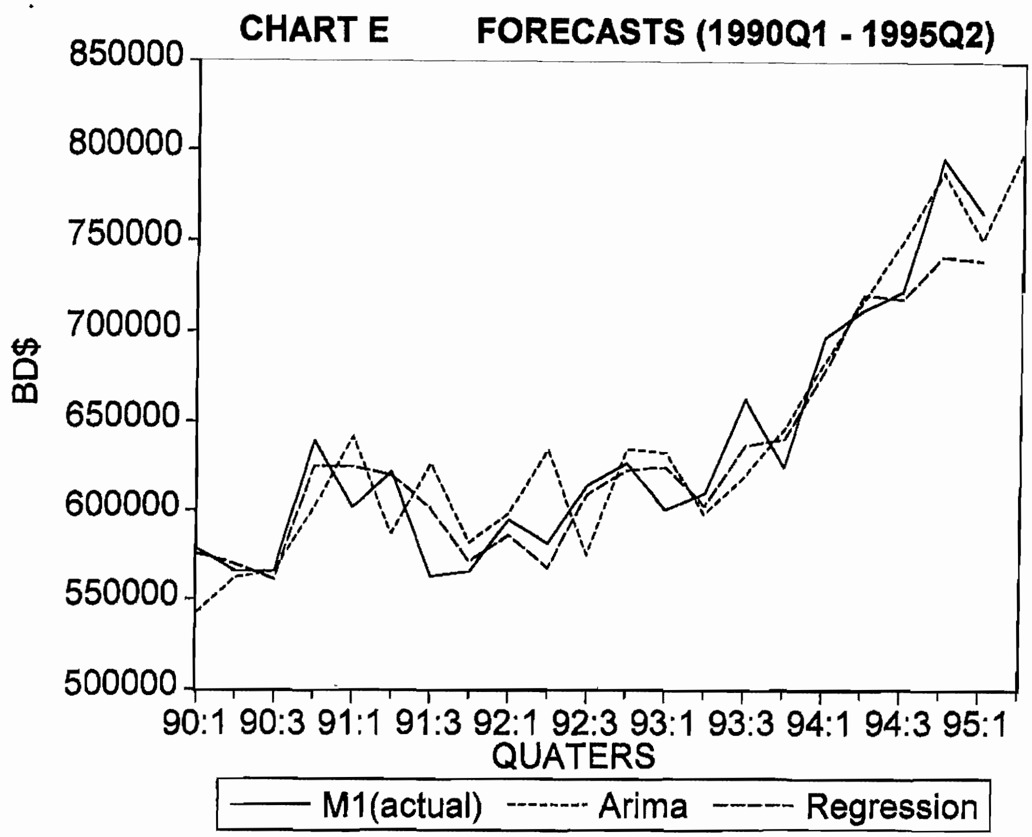


TABLE 3.7

FORECASTS OF THE MONEY SUPPLY (1990Q1 - 1995Q4)

Date	Money Supply	Holt-Winters AS		Holt-Winters NS		Holt-Winters Single		Holt-Winters MS		Arima		Regression	
		Forecast	Errors	Forecast	Errors	Forecast	Errors	Forecast	Errors	Forecast	Errors	Forecast	Errors
1990:1	578179.8	575730	2449.8	549850.8	28328.94	589503.2	-11323.5	575791.1	2388.676	5425[illegible].6	35583.14	575247.273	2932.494
1990:2	565697.8	580027.9	-14330.1	566620.1	-922.314	581350.7	-15653	580047.2	-14349.4	561947.5	3750.33	569400.808	-3703.02
1990:3	565408.2	578053.4	-12645.3	583900.8	-18492.7	570068.7	-4660.54	578015.2	-12607.1	565601.2	-193.061	561011.436	4396.723
1990:4	639355.9	592470.3	46885.6	601708.6	37647.33	566718.5	72637.34	592306.3	47049.63	600912.1	38443.79	624721.733	14634.15
1991:1	601702.8	619829.8	-18127	620059.4	-18356.6	617978	-16275.2	619938.1	-18235.3	641505.3	-39802.5	624375.416	-22672.8
1991:2	621784.1	612704.4	9079.701	638969.9	-17185.8	606248.5	15535.64	612721.1	9063.042	586774	35010.1	619921.873	1862.251
1991:3	562660	624355.7	-61695.6	658457.1	-95797.1	617363.3	-54703.3	624300.9	-61640.9	626462.4	-63802.4	600840.367	-38180.4
1991:4	565402.8	611230.8	-45828.1	678538.6	-113136	577575.8	-12173.1	611088.8	-45686	581501.3	-16098.5	571009.634	-5606.87
1992:1	594633.8	586587.8	8045.965	699232.8	-104599	568809.3	25824.48	586728.4	7905.366	598092.6	-3458.81	585667.816	8965.98
1992:2	581270.9	594078.4	-12807.6	720557.7	-139287	587234.9	-5964.09	594123.6	-12852.8	634215.4	-52944.5	567337.114	13933.74
1992:3	614246	593125.2	21120.73	742533.2	-128287	582946.5	31299.44	593093.5	21152.5	574837	39408.98	608999.397	5246.585
1992:4	626475	627674.4	-1199.37	765178.9	-138704	605253.3	21221.69	627566.8	-1091.82	634494.6	-8019.6	622514.548	3960.454
1993:1	600426.8	628567.6	-28140.8	788515.3	-188088	620416.3	-19989.5	628699.3	-28272.5	632603.8	-32177	624451.321	-24024.5
1993:2	609625.7	615757.4	-6131.62	812563.3	-202938	605997.7	3628.052	615781.5	-6155.78	597957.6	11668.16	602581.535	7044.207
1993:3	662633.5	618840.8	43792.77	837344.7	-174711	608600.5	54033.09	618791.1	43842.42	619509.6	43123.93	636252.513	26381.03
1993:4	624444.6	667246.5	-42802	862882	-238437	646928.3	-22483.7	667195	-42750.5	645716.6	-21272	640329.72	-15885.2
1994:1	696898	644532.7	52365.33	889198.1	-192300	630704.6	66193.41	644681	52216.97	682714.5	14183.52	677614.624	19283.38
1994:2	712270.1	676768.9	35501.19	916316.7	-204047	677558.2	34711.95	676759.5	35510.61	717871.7	-5601.54	720834.309	-8564.19
1994:3	722941.9	703847.3	19094.53	944262.4	-221321	702305.3	20636.6	703752.1	19189.77	750380.3	-27438.4	718612.273	4329.598
1994:4	796561.3	741424	55137.35	973060.4	-176499	717061.8	79499.55	741484.2	55077.15	789105.9	7455.45	742066.632	54494.71
1995:1	766375.8	773740.8	-7364.71	1002737	-236361	773290.2	-6914.39	773855.8	-7479.98	751110.2	15265.65	739900.818	26475.03
1995:2		773511.3		10[illegible]18		768318.4		7734[illegible]		798[illegible]			

4. THE COMBINATION OF FORECASTS

4.1 *Combination of Forecasts*

The idea of combining individual forecasts in the production of an overall forecast was originally proposed by Bates and Granger (1969). In their paper the combination of pairs of forecasts only was discussed, but the methodology can easily be extended to the combination of several forecasts (see Reid, 1969). Suppose one has K forecasts f_{1t}, f_2t, ...,f_{kt} of some quantity, X_t, and these individual forecasts are unbiased. Then the linear combinations

$$f_{ct} = w_1 f_{1t} + w_2 f_{2t} + \ldots\ldots + w_k f_{kt} \qquad o \leq w \leq \quad \sum_{i=1}^{k} w_i = 1 \qquad (4.1)$$

will also be unbiased.

It is straightfoward to show this unbiasness.

Let e_{it} denote the error resulting when the forecast f_{it} is used to predict x_t, so that

$$e_{it} = x_t - f_{it} \qquad (i = 1, 2, \ldots, k)$$

Since the K't individual forecasts are all unbiased, the associated errors all have expectation zero; that is

$$E(e_{it}) = 0 \qquad (i = 1, 2, \ldots\ldots, k)$$

If the composite forecast is formed as the weighted average (3.1), the error of the forecast will be

$$\begin{aligned} e_{ct} &= x_t - f_{ct} = x_t - (w_1 f_{1t} + w_2 f_{2t} + \ldots + w_k f_{kt}) \\ &= w_1(x_t - f_{1t}) + w_2(x_t - f_{2t}) + \ldots + w_k(x_t - f_{kt}) \\ &= w_1 e_{1t} + w_2 e_{2t} + \ldots + w_k e_{kt} \end{aligned}$$

Then, if each of the individual forecasts is unbiased, each e_{it} has expectation zero, from which it follows that $E(e_{ct}) = 0$ so that the combined forecast is also unbiased. It remains to determine the weights to be allocated to the constituent forecasts in forming the composite. One of the most common procedures used is regression-based weights. Suppose there are two forecasts to combine, with weights, w_1 and w_2 summing to one, so that $w_2 = 1 - w_1$

We can think of the actual observation as being the weighted average of the two forecasts plus an error term e_{ct} made in using the combined forecast to preduct x_t. Then.

$$x_t = w_1 f_{1t} + (1-w_1) f_{2t} + e_{ct}$$

or equivalently

$$(x_t - f_{2t}) = w_1(f_{1t} - f_{2t}) + e_{ct}$$

This can be extended to the case where there are k forecasts to be combined, using w_1, w_2, ... ,w_k. Since these weights are required to sum to one, we can write

$$w_k = 1 - w_1 - w_2 - \ldots\ldots w_{k-1}$$

The regression model is then

$$x_t = w_1 f_{1t} + w_2 f_{2t} + \ldots + w_{k-1} f_{k-1,t} + (1 - w_t - w_2 - \ldots - w_{k-1})\, f_{kt} + e_{ct}$$

The weights can then be estimated by applying least squares to

$$+ w_{k-1}(f_{k-1,t-j} - f_{k,t-j}) + e_{c,t-j} \qquad j = 1, \ldots, n$$

$$(x_{t-1} - f_{kt-j}) = w_1(f_{1,t-j} - f_{k,t-j}) + w_2(f_{2,t-j}) - f_{k,t-j} + \ldots$$

If these weight estimates turn out negative, the corresponding forecast should be dropped from further consideration and the regression re-estimated by least squares. The procedure is done until the remaining weights are positive.

The above procedure for the combination of forecasts is appropriate when the individual forecasts are unbiased. However, Granger and Romanathan (1984) have argued that this will not always be true. In such circumstances, the regression-based approach is easily appropriately modified by fitting the model

$$x_t = \alpha + \beta_1 f_{1t} + \beta_2 f_{2t} + \ldots + \beta_k f_{kt} + e_{ct} \qquad (4.2)$$

but now without restricting either α to be zero or the β, to sum to unity. This model is fitted by least squares to the available historical record, and then projected forward to derive the required composite forecast. Newbold and Boss (1990) argue that the errror terms in the regression equations will be auto-correlated beyond one-step ahead. With this in mind, in our evaluation studies on the combination of forecasts, we have considered in detail only one-step ahead forecast.

4.2 Results

Table 4.1 shows the sum Squared Forecast Errors for Box-Jenkin combined with Holt-Winters-additive seasonal B-J:HW(As), Box-Jenkin combined with Holt-Winters simple B-J:HW(S), Holt-Winter-additive seasonal combined with Holt-Winter-simple[HW(As):HW(S)] and for the individual methods. The number in parentheses indicates the weights of the individual forecasts, i.e. β_1 and β_1 in equation 4.2.

Table 4.1

One-Step-Ahead - Sum Squared Forecast Errors

BJ:HW (As)	B-J:HW (S)	HW(S):HW(As)	HW(S)	HW(As)	B-J
2.78E+10 (0.7,0.26)	2.48E+10 (0.67,0.32)	3.21E+10 (0.22,0.79)	4.48E+10	3.73E+10	2.98E+10

It can be seen from Table 4.1 that Box-Jenkins combined with any of the Holt-Winters methods give better results in terms of SSR for the series, outperforming also the Box-Jenkins individual method. It is interesting to note that the combined Holt-Winters methods produce forecast results which are superior to their individual forecasts for all series. When the univariant techniques are combined with the regression model, the regression equation, table 4.2, indicates that the individual weights of the univarant models were highly insignificant. After an elimination of the insignificant variables, the final results table 4.3 indicates that the regression model should not be combine with any of the univariant forecasts since sush a procedure will not improve the forecast of the regression model.

Table 4.2

Dependent Variable is LM1
Sample: 1990Q1 1995Q1
Included observations: 21 after adjusting endpoints

Variable	Coefficient	Std. Error	T-Statistic	Prob.
H-W(AS)	-46.23158	54.74940	-0.844422	0.4126
H-W(MS)	47.11353	54.98148	0.856898	0.4059
ARIMA	0.027071	0.220697	0.122659	0.9041
REG.	1.216603	0.238087	5.109910	0.0002
H-W(NS)	-0.587667	0.461832	-1.272469	0.2239
H-W(S)	-0.402616	0.394444	-1.020718	0.3247
C	-89237.31	58306.12	-1.530496	0.1482

R-squared	0.939594	Mean dependent var	633761.7
Adjusted R-squared	0.913706	S.D. dependent var	68136.87
S.E. of regression	20015.82	Akaike info criterion	20.06976
Sum squared resid	5.61E+09	Schwartz criterion	20.41793
Log likelihood	-233.5302	F-statistic	36.29417
Durbin-Watson stat	1.978155	Prob(F-statistic)	0.000000

Table4.3

Dependent Variable is LM1

Sample: 1990:1 1995:1

Included observations: 21 after adjusting endpoints

Variable	Coefficient	Std. Error	T-Statistic	Prob.
REG.	1.140298	0.073047	15.61056	0.0000
C	-84826.62	46214.64	-1.835492	0.0821

R-squared	0.927671	Mean dependent var	633761.7
Adjusted R-squared	0.923865	S.D. dependent var	68136.87
S.E. of regression	18800.77	Akaike info criterion	19.77370
Sum squared resid	6.72E+09	Schwartz criterion	19.87318
Log likelihood	-235.4215	F-statistic	243.6897
Durbin-Watson stat	2.162530	Prob(F-statistic)	0.000000

5. Conclusion

Our study indicates that the Holt-Winters - additive seasonal give the best forecast results of the exponential smoothing procedures for the money supply in the Barbadian economy. Also, that the ARIMA model give superior forecast results to hte exponential smoothing techniques. Furthermore, the individual univariant forecast can be improve upon by combining. Finally, the study justifies effort in constructing a multivariant model since the regression model produced superior forecast results to both the individual and combined univariant forecasts, and could not be improve upon by combining with any of the univariant techniques.

REFERENCES

Armstrong, J.S. (1985), **"Long Range Forecasting: From crystal to computer"**, 2nd ed., (Wiley, New York).

Banerjee, A., et al. (1988), **"Exporing Equilibrium Relationships in Econometrics through Static Models: Some Monte Carlo Evidence"**, ***Oxford Bullentin of Economics and Statistics,*** 48, pp. 253-278

Bates, J.M. and C.W.J. Granger (1969), **"The Combination of Forecasts"**, ***Operation Research Quarterly,*** 20, pp. 451-468.

Bourne, C., (1974) **"Dynamic Utility-Maximising Models of the Demand for MoneyCaribbean Economics"**, ***Social and Economics Studies,*** Vol.23, pp.418-85

Box, G. E. P. and G.M. Jenkins., (1970), **"Time Series Analysis, Forecasting and Control**. (San Francisco: Holden Day).

Brown, R.L., et al (1975) **"Techniques for Testing the Constancy of Regression Relations over Time",(with discussion),** ***Journal of the Royal Statistical Society,*** B, Vol.37, pp. 140-92

Cagan, P., (1956), **"The Monetary Dynamics of Hyperinflation"**, in M. Friedman, ed., Studies in the Quantity Theory of Money, (University of Chicago Press, Chicago,(IL)).

Chatfield, C. and D.L. Prothero (1976), **"Box-Jenkins Seasonal Forecasting: Problems in a Case Study"**, *Journal of the Royal Statistical Society, A, 136,* Part 3, pp 295-336.

Engle, R.F and Granger, C.W.J (1987), **"Co-integration and Error Correction: Representation, Estimation, and Testing,"** *Ecometrica, 55.*pp.251-276

Granger, C.W.J. (1986), **"Development in the study of Cointegrated Economic Variables,"** *Oxford Bullentin of Economics and Statistics, 48, 3.*

Granger, C.W.J., and P. Newbold (1987), **Forecasting Economic Time Series, (Academic Press: New York).**

Granger, C.W.J., (1989), **Forecasting in Business and Economics,** (Academic Press: New York, 2nd ed.).

Granger, C.W.J., and P. Newbold, (1974) **"Spurious Regressions in Econometrics",** *Journal of Econometrics,*Vol.2, pp. 111-20

Hall, S.G. and S.G.B. Henry, (1987) **"Wage Models",** *National Institute Economic Review*, No. 119, pp.70-75.

Hendry, David. (1986), **"Econometric Modelling with Cointegrated Variables: An Overview,"** *Oxford Bulletin of Economic and Statistics, 40, 3.*

Hendry, D.F., (1979), **"Predictive Failure and Econometrics Modelling in Macroeconomics: The Transcations Demand for Money"**, (in P.Ormerod, (ed.), Modelling the Economy, Heinemann, London).

Hendry, D.F. and G. Mizon, (1978), **"Serial Correlation as a Convenient Simplification, Not a Nuisance: a Comment on a Study of the Demand for Money by the Bank of England"**, *Economic Journal, Vol. 88, pp.549-63.*

Howard, M., (1981), **"The Demand For Money in the Trade Oriented Economy of Barbados, 1953-1977"**, *Central Statistical Office Research Papers*, Trinidad and Tobago.

Jarque, C.M., and A.K Bera, (1980), **"Efficient Test for Normality, Homoscedasticity and Serial Independence of Regression Residuals",** *Economic Letters*, Vol.6, pp. 255-9

Laidler, D., (1982), **The Demand for Money: Theories and evidence**, (New York, Harper and Row).

Makridakis, S., et al. (1982), **"The Accuracy of Extrapolation (time series) Methods: Results of a Forecasting competition"**, *Journal of forecasting, 1, 111-153.*

Makridakis, S., and M. Hibon, (1979), **"Accuracy of Forecasting: An emperical investigation (with discussion)"**, *Journal of the Royal Statistical Society, Series A, 142, 97-145.*

Makridakis, Spyros, and S.C. Wheckwright, (1978), **Forecasting Methods and Applications**, (John Wiley and Sons Publishing).

Makridakis, S., and Winkler, R.L. (1983), **"Averages of Forecasts: Some Empirical Results"**, *Management Science, 29, 987-996.*

Muth, J.F., (1960), **"Optimal Properties of Exponentially Weighted Forecasts"**, *Journal of the American Statistical Association*, Vol.55, pp. 197-206.

Nerlove, M., (1972), **"Lags in Economic Behaviour"**, *Econometrica*, Vol.40,pp.221-51.

Newbold, P. and T. Bos, (1990), **Introductory Business Forecasting**, (Southwestern Publishing Company).

Pagan, A.R. and A.D. Hall, (1983), **"Diagnostic Test on Residual Analysis"**, *Econometric Review*, vol.2, pp.159-218.

Payne, D.J., (1973), **"The determination of regression relationships using stepwise regression techniques,"** Ph.D. Thesis, Dept. of Mathematics, Univ. of Nottingham.

Phillpps, P.C.B., (1987), **"Time Series Regression with a Unit Root"**, *Econometrica*, Vol.55, pp.277-301.

Pfeffermann D, and J. Allon, (1989), **"Multivariate Exponential Smoothening: Method and Practice"**, *International Journal of Forecasting, No. 5*, pp. 83-98.

Ramsaran, r. and R.V. Maraj, (1985), **"The Demand For Money in Trinidad and Tobago -Experiments with Data From the 1970-1984 Period"**, paper [resented at the 17th Annual Conference of the Regional Programme of Monetary Studies, Bahamas.

Reid, D.J. (1969), **A comparative Study of Times Series Prediction Techniques.**

Stekler, H.O. (1991), **" Macroeconomic Forecast Evaluation Techniques",** ***International Journal of Forecasting, No. 7*****, pp. 375-384.**

Yule, G. U. (1969), **"Why do we sometimes get nonsense correlations between time series? -a study in sampling and the nature of time series",** *Journal of Royal Statistical Society, 89, 1-64.*