Monetary Conditions & Core Inflation:

An Application of Neural Networks

Prudence Serju
(Summer Intern)

Macromodelling & Real Sector Analysis Unit
Research Services Department,
March 27, 2000

TABLE OF CONTENTS

INTRODUCTION

Correction Model (VECM) (Robinson 1997) and a model that decomposes inflation into its local and foreign cost components. These model forecasts are currently complemented

above noted models utilize monthly data, which effectively restrict their use to that of short-term forecasts (e.g. one or two months).

employed at the Bank of Jamaica. In essence, the paper will attempt to develop (a) model(s) capable of forecasting inflation over the long-term, with the underlying premise

One of the much-debated topics among economists is whether policy should be formulated so as to respond to inflation in all goods and services, or whether it should

Price Index (CPI). According to Roger (1995), shocks to the general price level which are perceived as one-time events should not have a lasting effect on the inflation rate, and as

respond. Similarly, Brian Motley (1997) argues that temporary price shocks are often due to supply shocks, such as unusual weather which, affects harvests. These supply shocks

if the goal of policy makers is to control inflation, core inflation should be selected in preference to headline, as the former inherently avoids shocks or disturbances that add

The Bank of Jamaica has developed a measure of core inflation applicable to the Jamaican situation using the trimmed mean method (Allen, 1997). This paper recognises

proposed models for forecasting inflation.

In what follows, section 2 gives a brief discussion on the history of core inflation in Jamaica. Section 3 briefly describes selected economic theories that form the basis for the variables used in the various models. Section 4 describes methodological issues relating to three time series models; namely a neural network model, an autoregressive moving average (ARMA) model and a vector autoregression model (VAR) model. Section 5 presents the results and evaluates the competing model forecasts. The conclusion is presented in the final section.

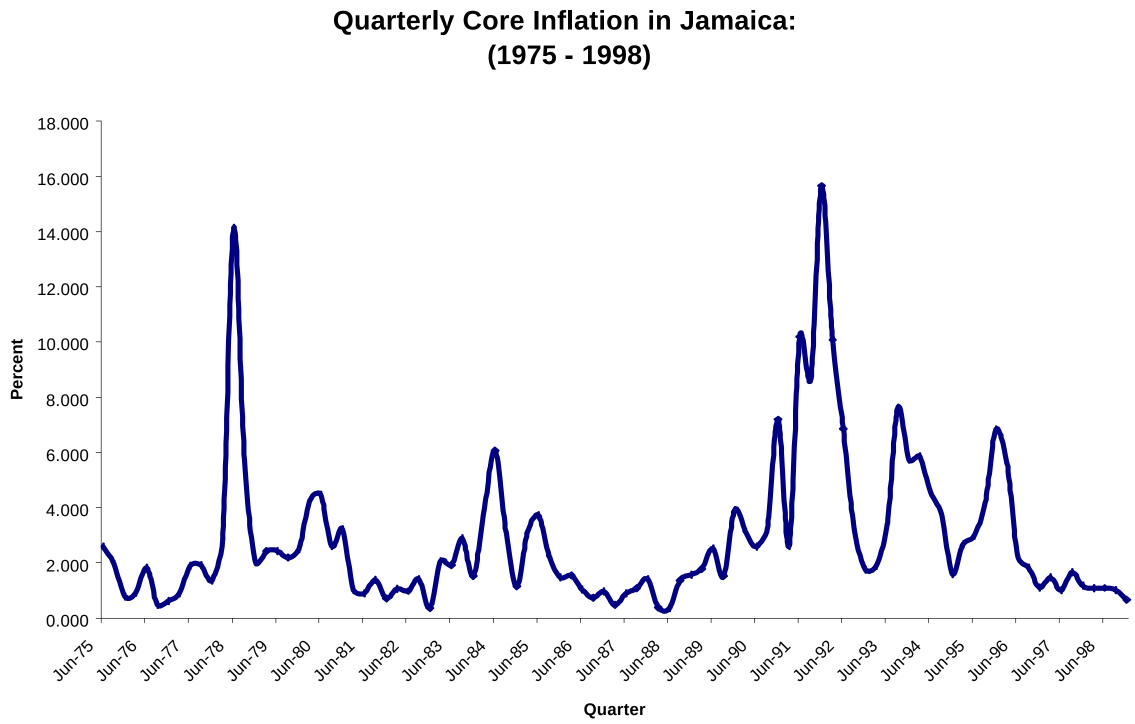
INFLATION IN JAMAICA

Figure 1 plots the graph of quarterly core inflation between 1975 and 1998. The average quarterly core inflation over the period was 2.9 percent. With the exception of 1978 and 1984, core inflation was fairly stable over the period 1975 and 1990. The 1978 hyper-inflationary episode resulted from the devaluation of the exchange rate due to the abolition of the dual exchange rate regime. This was carried out as part of the structural adjustment programme put in train by the government, under the direction of the multilateral lending agencies. Similarly, the pick-up in inflationary pressures in 1984 resulted from the introduction of the foreign exchange auction in March of that year, causing the exchange rate to devalue. In 1991, quarterly core inflation reached its peak. This resulted from the liberalization programme implemented by the government in 1990, and the subsequent freeing of the capital account in 1991. The effects from these events were felt in subsequent periods until the latter half of 1996.

It is important to recognise that the incidence of inflation over the review period has been buoyed by strong money growth. The average quarterly expansion of the monetary base between 1975 and 1998 was 5.6 per cent, with particularly strong growth being recorded between 1982 and 1983.

Since 1997, core inflation has been fairly stable, reflecting the renewed determination of the Bank of Jamaica to control inflation. This path of core inflation has occurred alongside major changes in the Jamaican economy, including substantial reductions in tariffs, partial elimination of price controls, subsidies and quantitative restrictions on commodity trade.

Figure 1



THEORY AND VARIABLE SELECTION

There are many different theories as to the causes of inflation, but there is no universally accepted theory that explains inflation in all countries. For our purposes, our variable selection process will be guided by two of these economic theories; monetarism and the purchasing power parity theory.

The monetarist school is of the view that inflation is a monetary phenomenon, which means that ".... long continued inflation is always and everywhere a monetary phenomenon that arises from a more rapid expansion in the quantity of money than in total output..." (Freedman, 1977).

Fishers' equation expressing the quantity theory of money can be used to demonstrate this proposition:

$$MV = Py \qquad 1$$

M represents the nominal money stock, V represents the income velocity of circulation, P represents the general price level, and y represents real output. Now the money stock is the product of a multiplier (K) and the monetary base (H)

$$M = KH$$

Substituting the money supply equation into 1, log linearising and differencing gives

$$\Delta p = \Delta h + \Delta v + \Delta k - \Delta y. \qquad \mathbf{2}$$

On the crude assumption that V and K are constants (Δv & $\Delta k = 0$), equation 2 becomes;

$$\Delta p = \Delta h - \Delta y \qquad \mathbf{3}$$

This view of the inflationary process in Jamaica is support by Allen (1997) where he found base money to be the most appropriate measure for monetary targeting. Real GDP enters into the process with a negative sign, in that strong output growth results in lower inflation. This result holds unambiguously to the extent that output growth is generated from expansions in aggregate supply. For the purpose of our investigation however, we are constrained by the unavailability of real GDP data.

There is a long history of controversy over the validity of the monetarist proposition. The assumptions of a stable velocity of circulation and the exogeniety of the money stock have been challenged by many. It has been argued that the velocity of circulation responds to changes in the rate of inflation, and that fiscal deficits and hence the money stock in developing countries increases with increases in the rate of inflation (Barnes

1996). Finally, monetarist tends to concentrate on long-run scenarios, and on economies where changes in the money growth are the primary sources of disturbances.

The determination of domestic prices from import prices can be extracted from the theory of purchasing power parity (PPP) of exchange rate determination, which in one sense assumes a single price, and in another asserts that the exchange rate is merely the ratio of domestic prices to foreign prices:

$$P^D = \xi P^F \qquad \textbf{4}$$

P^D represents domestic price, ξ represents the exchange rate, and P^F represents foreign prices. The equation purports that given a fixed exchange rate regime, domestic prices will adjust proportionately to changes in the price level of trading partners. On the other hand, the equation suggests that, ceteris paribus, increases in the rate of exchange will trigger an upward movement in the domestic price level.

According to Bond (1980), developing countries tend to have a high import content in domestic production and consumption. Consequently, increased foreign prices or currency devaluation will inevitably cause domestic costs and prices to rise. Barnes (1996), for example, found import prices to be a major source of inflation in Jamaica.

Log linearising and differencing equation 4 gives:

$$)p =)e +)p^f \qquad \textbf{5}$$

Combining 5 and 3 and adding behavioural parameters results in the following model of the inflation process in Jamaica:

$$)p = a_0 + a_1)e + a_2)p^f + a_3)h - a_4)\mathbf{y} \qquad \textbf{6}$$

Based on equation 6, base money, the exchange rate and foreign prices were therefore chosen as the variables of interest for this model. Given the unavailability of quarterly

GDP, the last term was suppressed. Oil prices were used as a proxy for foreign prices. In addition to these variables, we find it important to include a proxy for the Bank of Jamaica's signal rate. This rate has traditionally reacted to volatility in the exchange rate and has been identified as an important element of the monetary transmission process in Jamaica by Robinson (1998). The 30-days treasury bill rate was selected as a proxy for the signal rate given its high degree of correlation with the reverse repurchase rate. We anticipate that this variable will therefore act as a damper on inflation.

Except for oil prices, which were taken from West Texas Intermediate Crude Oil Price listings, all the variables were taken from BOJ's database. For the core series, initial work had been done for the period 1992 to 1999. For the purpose of this paper, the CPI was collected from STATIN for the period 1975 to 1998 and the index created using the same methodology currently employed by BOJ in reporting core inflation.

For the purpose of modelling and estimation, the EVIEWS and RATS statistical packages were employed.

MODELS AND METHODOLOGICAL ISSUES

Neural Network Models

Over the last two decades there has been increased research on conditions under which various forecasting methods perform best (Makridakis et al. 1982, 1986). It has been found that no single method dominate the forecasting landscape. However, it has been established that simple and parsimonious models are robust under a wide range of conditions.

Recently a new forecasting methodology, Artificial Neural Networks (ANNs), has emerged which is well suited to the task of prediction and forecasting. Remus & O'Connor (1998) indicate that ANNs excel in pattern recognition and forecasting from pattern clusters. ANNs have two advantages when compared with other traditional methods of forecasting. Firstly, they are universal approximators of functions in that they can approximate whatever functional form best characterizes the time series. In this context, they are inherently non-linear, but can overcome the limitations of linear forecasting models. Secondly, ANNs have been proven to be better than traditional forecasting methods for ***long term*** forecast horizons, but are often as good as traditional forecasting methods with shorter forecast horizons.

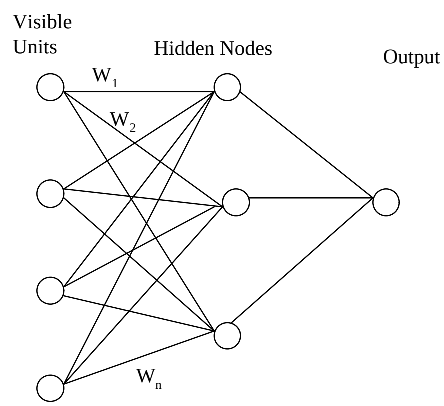
McCullock's (1949) paper laid the theoretical basis for the development of artificial neutral network (ANN). Minsky (1951) developed the computing platform on which current ANN models are processed. The break though for applied ANN occurred with the development of the back-propagation technique in 1986 (Rumelhart & McClelland, 1986).

ANN models are inspired from biological neural networks. They are developed on software that attempt to mimic the human brain's ability to classify patterns or to make predictions or decisions based on past experience (Gately, 1996). While the human brain relies on inputs from the five senses, ANN uses inputs from data sets.

Architecture of ANN Models

ANNs typically contain three or more layers consisting of input (visible units), hidden, and output units.

Figure 1.
Typical Neural Network Architecture



Input layers receive input patterns directly, while hidden layers neither receive inputs directly nor are given direct feedback. Hidden units are the stock of units from which new features and new internal representations can be created. In a neural network model, the user may specify a pattern of inputs to the visible units, but by assumption the user is not allowed to specify external inputs to the hidden unit. Their net input is based only on the outputs from other units to which they are connected.

Interactive Activation

The units in a neural network takes on continuous activation values between a maximum and minimum value, though their output – the signal they transmit to other units – is not necessarily identical to their activation. In the ANN model, output = $[a_j]^+$. Here, a_j refers to the activation of unit j, and the expression $[a_j]^+$ has value a_j $\forall$ $a_j>0$; otherwise its value is zero. The index j ranges over all units with connections to unit i.

Units change their activation based on the current activation of the unit and the net input to the unit from other units, or from outside the network. The net input to a particular unit (say unit i) is the weighted summation of all the output from other units plus any external input:

$$net_i = \sum_j w_{ij} output_j + extinput_i \qquad 7$$

In general the weights (w_{ij}) can be positive or negative.

Once the net input into a unit has been computed, the resulting change in the activation of the unit is as follows:

If ($net_i > 0$)

$$\Delta a_i = (max - a_i)net_i - decay(a_i - rest) \qquad 8$$

Otherwise

$$\Delta a_i = (a_i - min)net_i - decay(a_i - rest)$$

where max, min, rest and decay are all parameters. In particular, we choose max = 1, min ≤ rest ≤ 0, and decay between 0 and 1. Note also that a_i is assumed to start, and to stay within the interval [max, min].

The optimal activation of the unit occurs when the incremental activation of the unit is zero. Setting $\Delta a_i = 0$ and rearranging expression 8 results in the following equilibrium condition for the activation of the unit:

$$a_i = \frac{(max)(net_i) + (rest)(decay)}{net_i + decay}. \qquad 9$$

Using max = 1 and rest = 0, this simplifies to

$$a_i = \frac{(net_i)}{net_i + decay} \qquad 10$$

Equation 10 indicates that the equilibrium activation of a unit will always increase as the net input increases; however, it can never exceed 1 (or, in the general case, max). The decay term acts as a kind of restoring force that tends to bring the activation of the unit back to zero (or to rest in the general case). Decreasing the value of this decay parameter increases the equilibrium activation of the unit.

Learning

Neural network models are of interest because they learn, naturally and incrementally, in the course of processing. One classical procedure for learning is the error correcting or delta learning rule as studied by Widrow and Hoff (1960) and by Rosenblatt (1959). The delta rule in its simplest form, can be written as

$$\Delta w_{ij} = \varepsilon \delta_i a_j$$

where ε is the value of the learning parameter and $e_{i,}$ the error for unit *i*, is the difference between its teaching input (t_i) and its obtained activation (a_i)

$$\delta_i = t_i - a_i$$

Note that if $\delta_{pi} < 0$, the adjustment to weight w_{ij} will be negative so that the influence of input *i* is reduced. When the weights are changed according to this rule, each weight is moved towards its own minimum and the system moves downhill in *weight-space* until it reaches its minimum error value. When all the weights have reached their minimum points, the system has reached equilibrium. This procedure of finding the set of weights that minimises the error function is called *gradient descent*. If the system is able to solve the problem entirely, the system will reach zero errors and the weights will no longer be modified. On the other hand, if the network is unable to get the problem exactly right, it will find a set of weights that produces as small an error as possible.

For a simple network with say two input units and a single output unit, learning occurs by activating each unit, preparing an output, and then comparing this output with the teaching input. The error between the teaching input and the output of the network is then used to adjust the weights through the fixed learning parameter. The correct set of weights is approached asymptotically if the training procedure is continued through several sweeps, each of these sweeps being referred to as a ***training epoch***. Each epoch results in, theoretically, a set of weights that is closer to the perfect solution. To get a measure of the closeness of the approximation to a perfect solution, we can calculate the total sum of squared errors that result on each epoch. This measure of the state of learning of the network gets smaller over each epoch, as do the changes in the strength of the connections. Minsky and Papert (1969) has shown that the error correcting rule will find a set of weights that drives the error as close to zero as desired, provided that such a set of weights exist.

It should be noted that such a set of weights as described above exists only if for each input-pattern – target-pair, the target can be predicted from a linear combination of the activation units. That is the weights must satisfy the linear predictability constraint:

$$t_{ip} = \sum_j w_{ij} a_{jp}$$

for output unit *i* in all patterns *p*. This constraint can be overcome by the use of hidden units, which in turn introduces problems relating to the training of the network.

Training Hidden Units: Back Propagation

The application of the back propagation rule involves two phases. During the first phase the input is presented and propagated forward through the network to compute the output value a_{pj} for each unit (we will assume a single output unit for simplicity). This output is then compared with the target, resulting in a δ term for the output unit.

$$\delta_{pi} = (t_{pi} - a_{pi})\, f'_i\,(net_{pi})$$

where $net_{pi} = \sum_j w_{ij} a_{pj} + bias_i$ is the activation function, and $f'_i(net_{pi})$ is the first derivative of the activation function with respect to a change in the net input to the unit.

The second phase involves a backward pass through the network (analogous to the initial forward pass) during which the δ term is computed for each unit in the network. In the case of the hidden units, there is no specified target so that δ is determined recursively in terms of the δ terms of the units to which they directly connect and the weights of those connections. That is

$$\delta_{pi} = f'_i(net_{pi}) \sum_k \delta_{pk} w_{ki}$$

Once these two phases are complete, we can compute the weighted error derivative for each weight. These weighted error derivatives can then be used to compute actual weight changes on a pattern by pattern basis, or they may be accumulated over the ensemble of patterns.

ARMA & VAR Models

This class of linear stochastic difference equation underlies most of the theory of time-series econometrics. Assuming stationarity, the general form of the ARMA(p, q) model is as follows:

$$y_t = a_o + \sum_{i=0}^{p} a_1 y_{t-1} + x_t$$

where $x_t = \sum_{j=0}^{q} \beta_j \varepsilon_{t-j}$ and $\beta_0 = 1$.

Note that the model for the forcing process x_t contains q lags and that the homogeneous part of the difference equation contains p lags.

Of importance is the Box-Jenkins (1976) methodology for estimating time-series models. One of the most commonly used model selection criteria for determining P and Q in the Box-Jenkins methodology is the Schwartz Bayesian Criteria (SBC). This criterion is defined as follows:

$$SBC = T \times \log(RSS) + n \ln (T),$$

where RSS is the residual sum of squares, n is the number of parameters estimated and T is the number of usable observations. The model with the smallest SBC is the best.

Vector autoregression models are multivariate extension of ARMA models. In the VAR framework, the x_t sequence is endogenized, thereby forming a system of equations. The VAR model in matrix form is given as follows:

$$Y_t = A\, Y_{t-1} + e_t$$

Where Y_t is a vector of variables, A is a matrix of polynomials in lag operator and e_t, is a vector of random errors. VAR models are simple multivariate models in which each variable is explained by its own past values and the current and past values of all other variables in the system (Robinson, 1996).

The disadvantage associated with using a VAR model is that the number of variables and lags may lead to over-parametization. This results in poor out-of-sample forecasts (although in-sample forecasts may be good) and multicollinearity among the different lagged variables.

To avoid this dilemma, a likelihood ratio test is employed to determine the most appropriate lag length to be included in the VAR. This statistic is given by;

$$(T-c)\,(\log |\Sigma_r| - \log |\Sigma_u|\,)$$

where T is the number of observations, c is the number of parameters estimated in each equation of the unrestricted system, and $|\sum_r|$ and $|\sum_u|$ are the variance/covariance matrices of the restricted and unrestricted system respectively. This can be used to test the null hypothesis that the restrictions are not binding. The restrictions are the number of reduced lag length. It has a χ^2 distribution with degrees of freedom equal to the number of restrictions in the system.

The test is employed iteratively (reducing the number of lag length each time) until the most appropriate lag length is reached. It may also be used to test for the presence of seasonal dummy variables. However, the null hypothesis would then be that the seasonal dummy variables are equal to zero.

RESULTS

Unit Root Test

The usual practice is to pre-test the variables for the presence of unit roots as detailed in Appendix II. Table A (Appendix I) shows the results of the unit root tests. Core inflation (CORE), base money (LNBAS), exchange rate (LNEXR), oil prices (OIL), and treasury bill (TBILL) all exhibit unit root properties. It is interesting to note that the SBC chose a lag length of 9 for treasury bill.

Neural Network

The network was trained in RATS using data over the period 1975:1 to 1996:4. One hidden layer with three nodes was specified along with an R-square of 0.95 as the convergence criteria. The network converged after 15,899 epochs. Various attempts to obtain a better fit indicated that the exchange rate and base money should enter the model with four lags, the treasury bill variable should be excluded, and three seasonal factors, the pulse dummies, and that an autoregressive term should be included.

The Llung-Box Q-statistic indicates the presence of mild serial correlation in the error term of the neural forecast. Perhaps, one way to remedy this would be the specification of a more complex network or to use alternative learning algorithms such as the feedforward training techniques. RATS unfortunately does not cover multi-layer network design, nor does it contain alternative algorithms to the back-propogation technique. For more in-depth work on neural nets, software such as "Matrix Backpropagation", "WinNN" or "The Brain" would be required[1].

ARMA

Two pulse dummies were created to take account of the two one - off shocks to core inflation. The first pulse dummy was used to capture the effect of the 1978 shock, equaling one in June 1978 and zero elsewhere. The second is a combination of a gradually changing and prolonged pulse dummy, equaling one in December 1991, and a half (0.5) in June (1991), September (1991), March (1992) and June (1992). Three seasonal dummies were also created to capture the seasonal influences in the data. The core series was smooth using Hodrick-Prescot filter[2], and the long run component included in the model.

Table B in Appendix I gives the SBC results of the various Box-Jenkins models, which indicate that an AR(1) model was favoured. The results of the ARMA model are contained in table 1. The two pulse functions were found to be statistically significant, as well as the filtered series and the AR component. The Llung-Box Q-Statistics (Table C,

---

[1] WinNN, for example, is a Neural Networks package for windows 3.1 and above. WinNN incorporates a very user-friendly interface with a powerful computational engine. It provides an alternative to using more expensive and hard to use packages. WinNN can implement feed forward, multi-layered NN and uses a modified fast back-propagation for training. It also has various neuron functions. It allows testing of the network performance and generalization. All training parameters can be easily modified while WinNN is training. Results can be saved on disk or copied to the clipboard. It also supports plotting of the outputs and weight distribution.

[2] This is a smoothing method that is used to obtain an estimate of the long-term trend component of a series. Technically, the Hodrick-Prescott (HP) filter is a two-sided linear filter that computes the smoothed series s of y by minimizing the variance of y around s, subject to a penalty that constrains the second difference of s. That is, the HP filter chooses δ_t to minimize:

$$\Sigma^{T}_{t=1}(y - \delta_t)^2 + \lambda\Sigma^{T-1}_{t=2}(\delta_{t+1} - \delta_t) - (\delta_t - \delta_{t-1}))^2$$

The penalty parameter λ controls the smoothness of the series δ_t, the larger the λ, the smoother the δ_t. As λ approach infinity, δ_t approaches a linear trend.

Appendix I) indicated the absence of serial correlation, while the White's test (Table D, Appendix I) suggested that the error term was free from heteroskedasticity.

Table 1
Result of ARMA (1,0) model
Sample 1975:1, 1996:4

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.002	0.006	0.252	0.802
Dummy 1	0.104	0.012	8.544	0.000
Dummy 2	0.110	0.014	8.125	0.000
Q1	0.001	0.003	0.305	0.761
Q2	0.001	0.004	0.421	0.675
Q3	0.002	0.003	0.723	0.472
Filter Core	0.740	0.192	3.847	0.000
Core(1)	0.425	0.097	4.371	0.000

R-squared 0.80
SBC -5.56

VAR

The Johansen Cointegration test (Table E, Appendix I) on the variables revealed one cointegrating equation at the 5% significance level[3]. Based on these results a Vector Error Correction (VEC) model was estimated. The likelihood ratio test (Table F, Appendix I) indicated that the variables should enter the model at two lags.

The long-run equilibrium results (Table 2) of the VEC indicate the nature of the relationships between core inflation and its specified determinants. The exchange rate, base money, and foreign prices all have positive effects on core inflation, with the exchange rate having the greatest impact. A one-unit shock to the exchange rate causes core inflation to increase by 38% over the long run, while shocks to the treasury bill causes core inflation to decline marginally by 1%. The effects of oil prices in the long run appear to be relatively strong.

[3] A brief dissertation on cointegration theory is provided in Appendix II

Table 2
Long Run Equilibrium Relationship: VEC Model

Core Index	Ex-rate	Tbill	Base Mon	Oil Prices
1.000	0.3850	-0.0905	0.1803	0.2886

The impulse responses from the VEC model are shown in figure 2 (Appendix I). A unit shock to the exchange rate has an immediate positive permanent impact on the core index. This impact is highly significant for the first six quarters, after which it dies out. A unit shock to base money, surprisingly, has an immediate positive impact on the core index, which lasts for as much as seven quarters. Likewise, a unit shock to foreign prices has an immediate positive impact on the core index, peaking in the second quarter before settling to its equilibrium level. A unit shock to the treasury bill rate has a short lived, marginal positive impact on the index up to the second quarter, after which it has a long lived negative impact. The core index has the most influence on itself, suggesting that the inflationary process in Jamaica carry significant inertia. A unit shock to the index has immediate, positive, and significant effects over the first four quarters.

Based on the impulse responses in figures 3 and 4 (Appendix I), base money has a significant impact on the exchange rate, while the exchange rate has a marginal impact on base money. By deduction, the causation appears to run from base money to the exchange rate, and finally to prices.

To determine the contribution of each variable to the core inflation process, we look at the variance decomposition for the sixty-step-ahead forecasts. The results (Table G, Appendix I) show that most of the variability in core inflation was caused mainly by shocks to itself, base money and the exchange rate, respectively throughout the period. Surprisingly, shocks associated with foreign prices have a marginal effect on core inflation over the forecast horizon.

Forecast Evaluation

Table H (Appendix I) provides a comparison of the forecasting accuracy of the three models under consideration. Based on these statistics, the ANN model has the greatest predictive power. It has the lowest mean squared error (MSE), root mean squared error, and mean absolute error. Of importance to note, the Theil-U coefficient of the ANN model was approximately five times smaller than the Theil-U coefficient associated with the ARMA and VEC models.

Figures 5, 6, and 7 show the (in sample) forecasts compared with the actual series for the three models. All the models captured the shock to core inflation in 1978. The ANN model captured the major turning points in core inflation well, although there are indications that between 1986 and 1989 it over-estimated the series. The other two models appear to have done a fairly good job over this period. Also, the liberalization effect was best capture by the ANN model, followed by the ARMA model. The VEC model failed to capture the full effect of this policy, and apparently did a poor job at estimating the second to last spike during the period. Overall, it is easy to conclude from the visual examination that the ANN and ARMA models did a better job at forecasting core inflation when compared with the VEC model.

Figure 5

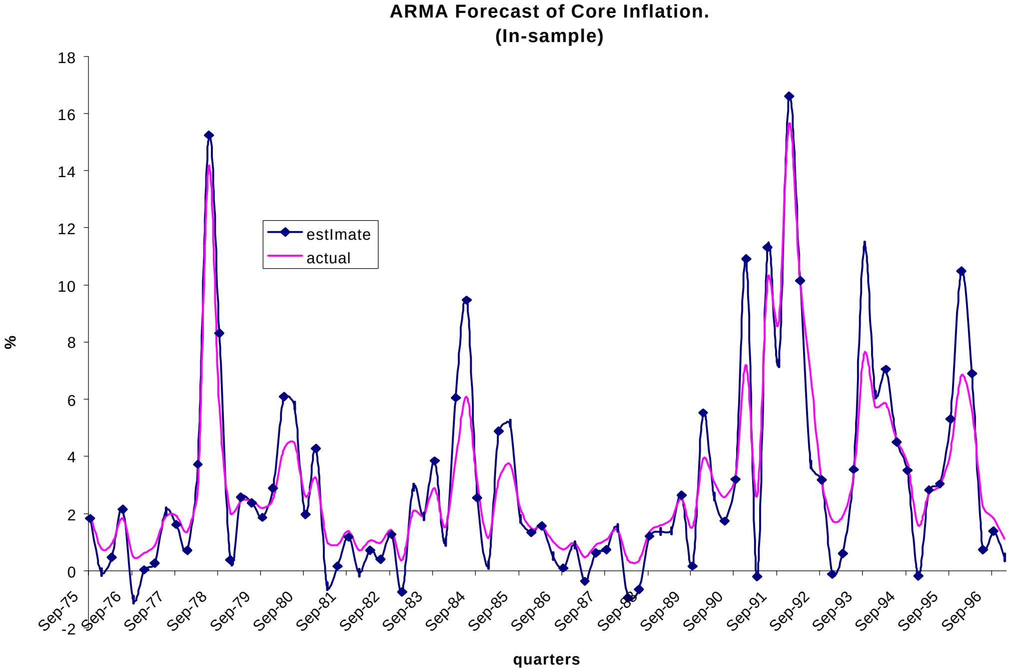


Figure 6

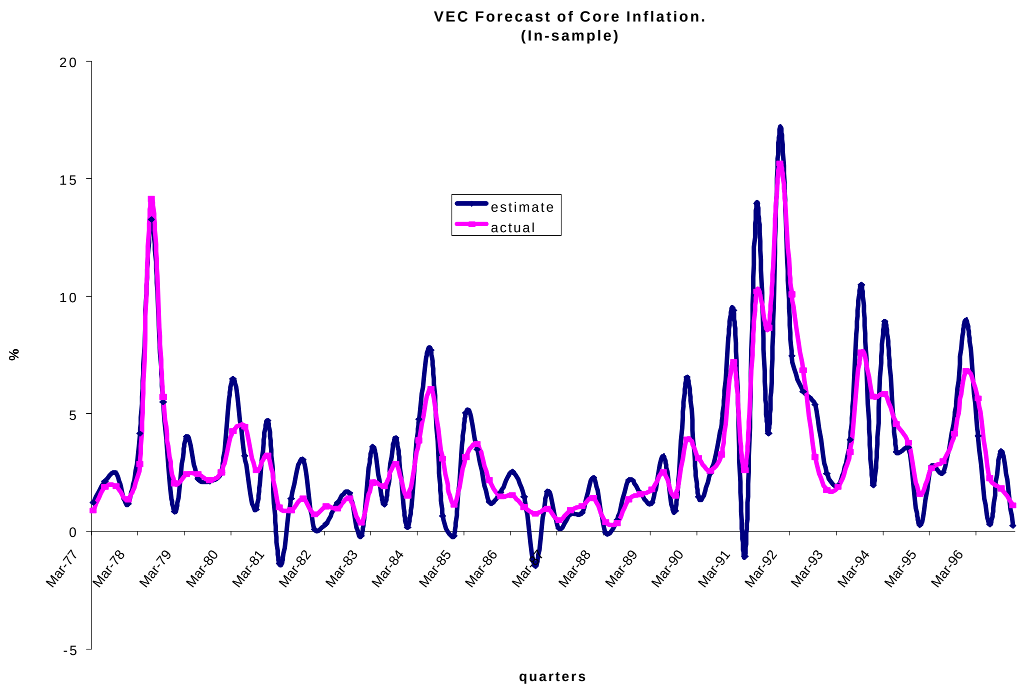
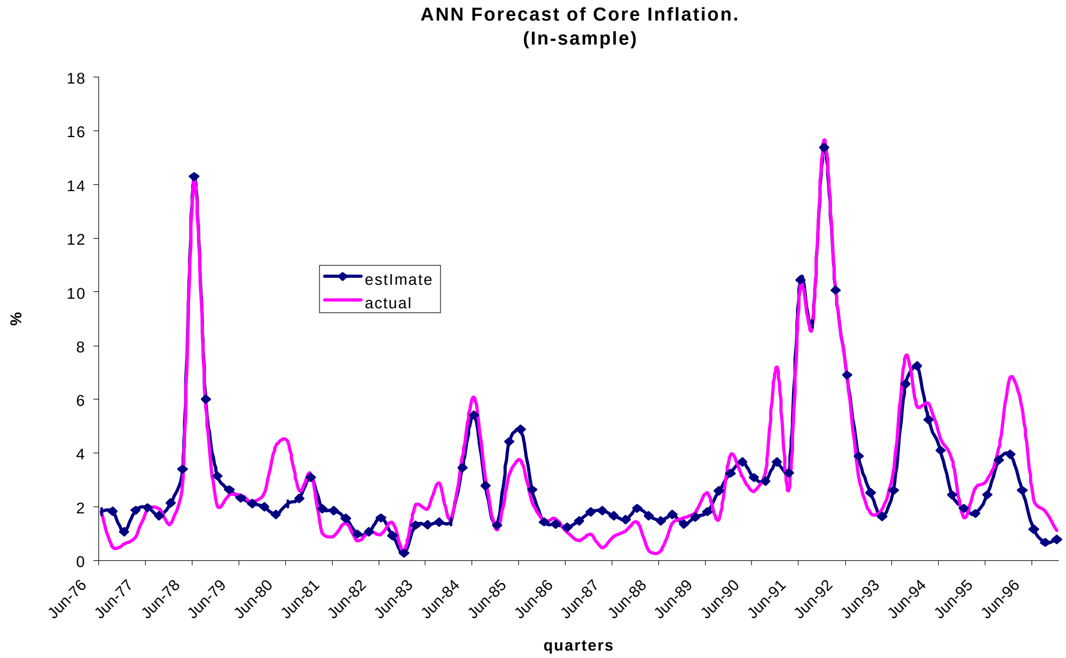


Figure 7



A forecast encompassing test was also used to assess the ARMA and ANN forecasting abilities. If one model forecast encompass another, that model's forecast is said to be unbiased, contains all the information present in the other, but contains more useful information. Failure of one model's forecasts to forecast encompass another indicates that it is possible to gain by combining the forecasts. The results in table I (Appendix I) indicated that the ANN model forecast encompass the ARMA model.

Despite the fact that the ANN model dominated the in-sample forecast performance, the ARMA model did a better job with the out-of-sample forecast. Significantly, the Janus quotient (J-Quotient) for the ARMA model was approximately seven times smaller than the quotient for the other models. It should be noted that the Janus quotient is a more powerful tool than the Theil-U coefficient when dealing with out-of-sample forecast evaluations.

Figure 8

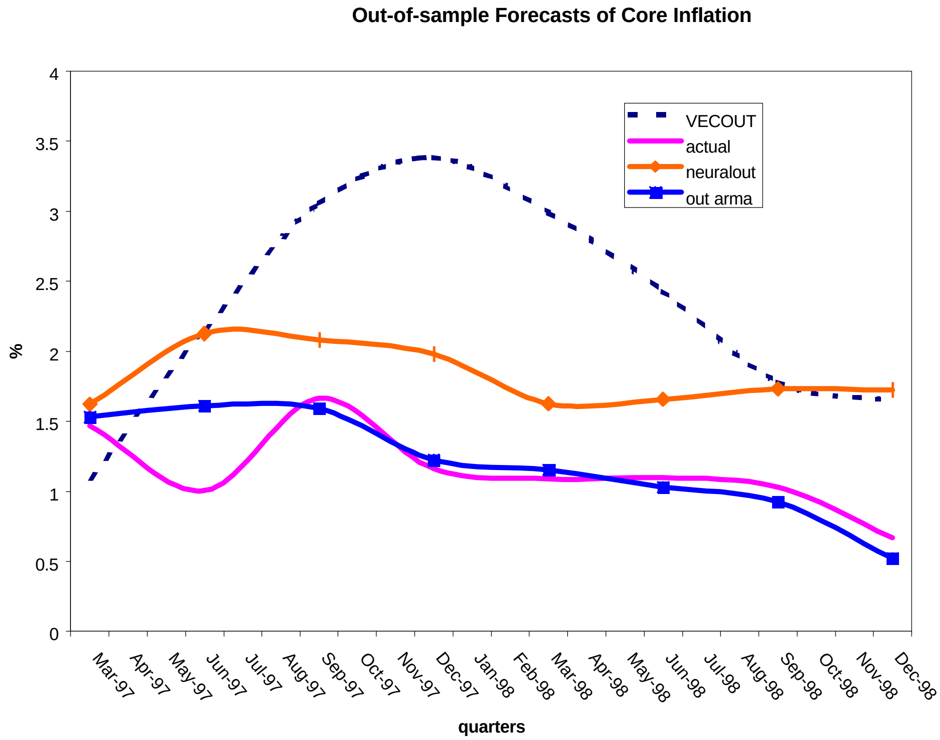


Figure 8 shows the out-of-sample forecast. The graph indicates that the ARMA model remains closest to the actual core inflation out-turn between the first quarter of 1997 and the last quarter of 1998. The VEC model predicts an upsurge in core inflation over the period, while the ANN suggests an inflation out-turn which was higher than actual. These results are not heartening. What they would suggest is that the simple specification of the ARMA model is more robust for forecasting core inflation over the long term. Another interpretation may very well be that the relationship between core inflation and its primary determinants may not be holding in a predictable manner since 1996. The omission of quarterly GDP from the estimation process may also provide another significant bias to both models' forecasting abilities.

CONCLUSION

Of the three models estimated the ANN model was the most appropriate in making in-sample forecast of core inflation. The ARMA model performed marginally better than the VEC model for in-sample forecasts. In addition, the ARMA model performed better than the two more sophisticated models when making out-of-sample forecast. In this context, the recommended models for estimating and forecasting core inflation are the ANN and ARMA models.

Based on the results from the models, volatility in core inflation was due mainly to innovations to itself, to base money, the exchange rate, and to a lesser extent foreign price. The results of the impulse responses indicated that core inflation responded immediately to a shock to base money.

It appears that further work needs to be done on the ANN model. The model's potential for producing long-term reliable forecasts is great, but in the context of software limitation, this could not be immediately exploited.

APPENDIX I

Table A

Augmented Dickey-Fuller Test			
	T statistics		
Variables	Levels	First Difference	lag
CORE	-1.85	-3.80	1
BASE MONEY	-0.46	-3.65	4
EXCHANGE RATE	-2.58	-5.23	1
TBILL	-1.44	-4.60	9
OIL PRICES	-2.25	-8.09	1
5% critical value	-3.45	-3.45	N/A
1% critical value	-4.04	-4.04	N/A

Table B
Model Selection using AIC and SBC criteria.
Box-Jenkins Models

Models	***AIC***	***SBC***
*arma(1,1)	-316.84	-300.25
arma(1,1)	-307.24	-290.66
*arma(1,2)	-314.85	-295.89
arma(1,2)	-308.60	-292.01
*ar(1)	**-318.70**	**-304.49**
*arma(1,4)	-314.22	-290.53
*ma(4)	-313.92	-290.22
*ar(||1,6||)	-281.90	-300.31
ar(1)	-280.60	-290.37

Table C

Llung-Box Q-Statistics

Period	**Q-Statistics**	**Probability**
5	0.32	0.955
10	8.51	0.648
15	13.81	0.454
20	15.58	0.685

Table D

Whites Heteroskedascity Test

F Statistic	R Squared	Probability
1.45	11.32	0.184

Table E

JOHANSEN COINTEGRATION TEST

Sample 1975:2 to 1998:4

Null Hypothesis	Eigenvalve	Likelihood Ratio	5% Critical Valve	1% Critical Valve
r = 0	0.061	116.71	68.52	76.07
r < 1	0.193	32.80	47.21	54.46
r < 2	0.106	13.75	29.68	35.65
r < 3	0.039	3.77	15.41	20.04

Likelihood ratio test indicates one cointegrating equation at the 5% significance level.

Table F

Results of Likelihood Ratio Test

Lag Lengths	Chi-square ratio	Significance Level
8 vs 4	90.39	0.74
4 vs 2	57.64	0.21
2 vs 1	76.03	0.00

Figure 2

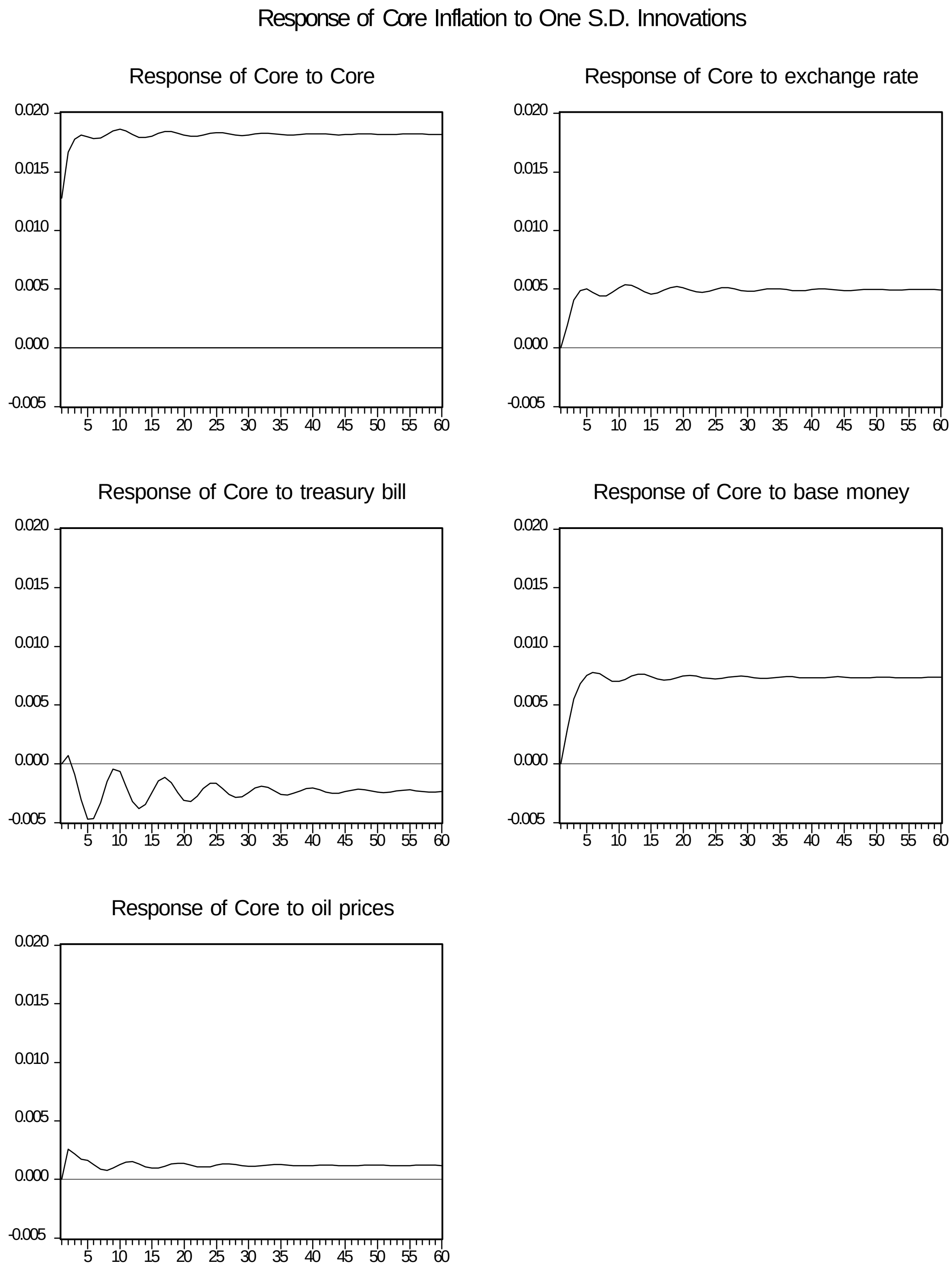

Response of Core Inflation to One S.D. Innovations
Response of Core to Core
Response of Core to exchange rate
Response of Core to treasury bill
Response of Core to base money
Response of Core to oil prices
0.020
0.015
0.010
0.005
0.000
-0.005
5 10 15 20 25 30 35 40 45 50 55 60

Figure 3

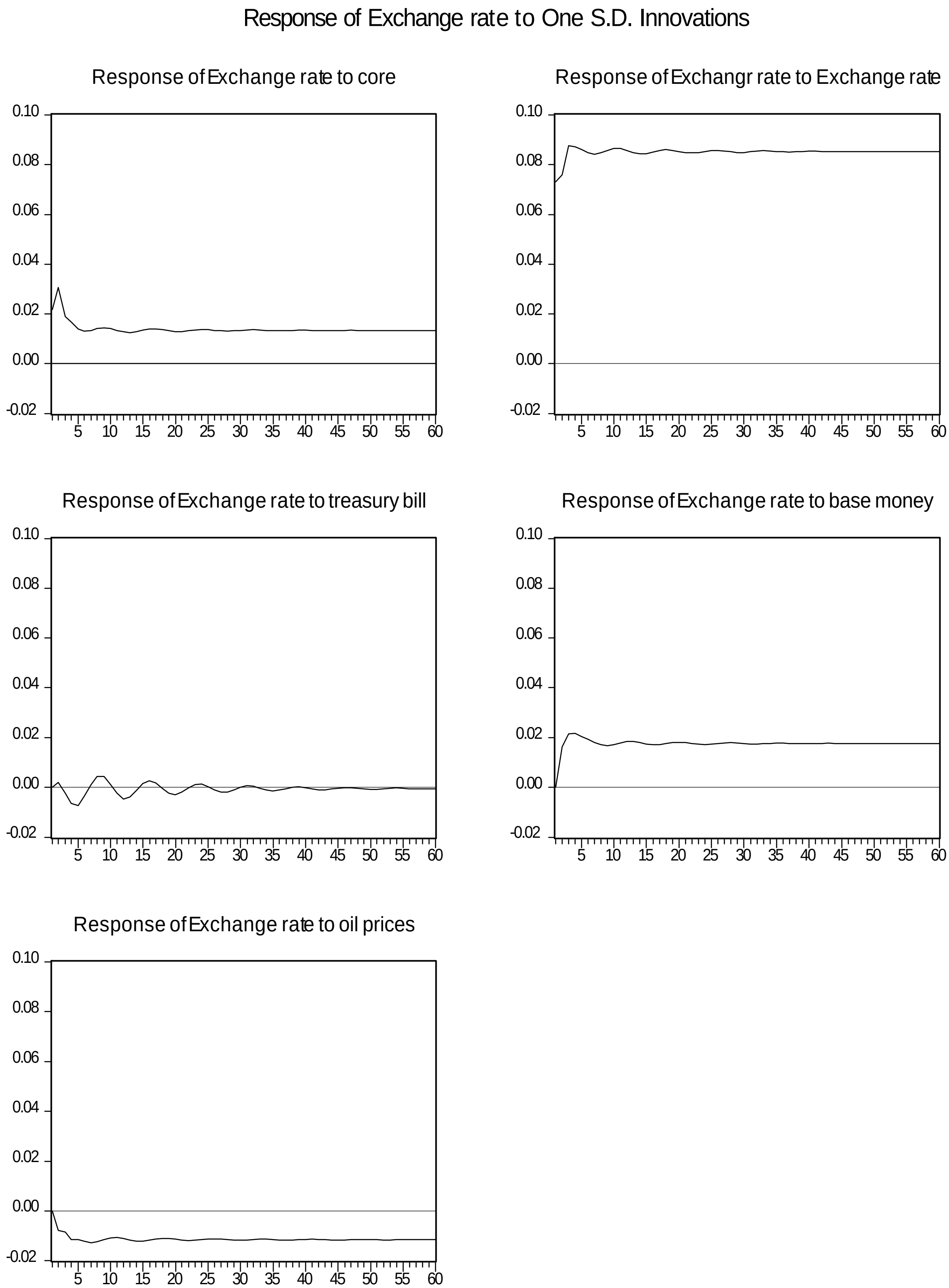

Response of Exchange rate to One S.D. Innovations
Response of Exchange rate to core
Response of Exchangr rate to Exchange rate
Response of Exchange rate to treasury bill
Response of Exchange rate to base money
Response of Exchange rate to oil prices
0.10
0.08
0.06
0.04
0.02
0.00
-0.02
5 10 15 20 25 30 35 40 45 50 55 60

Figure 4

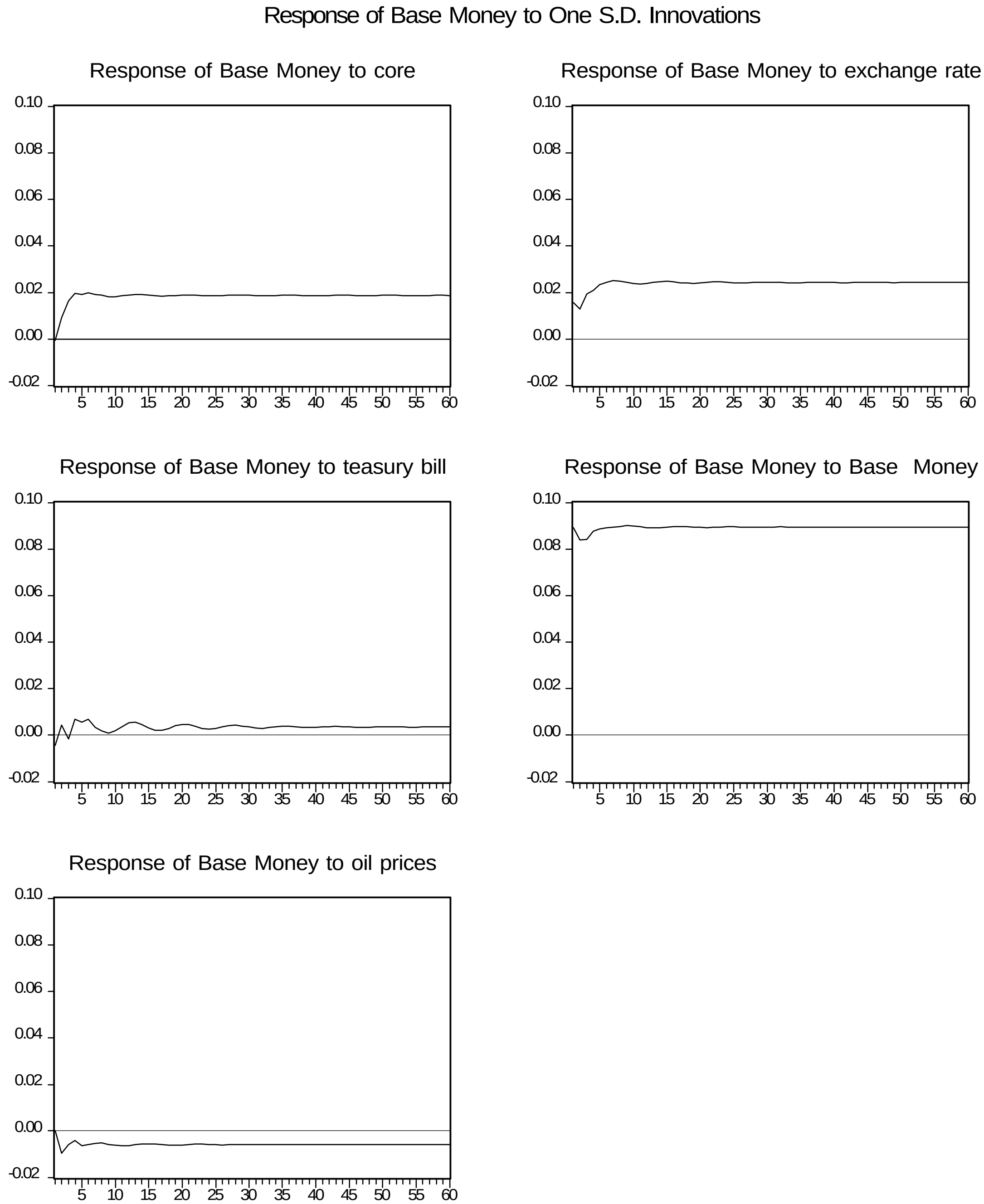

Response of Base Money to One S.D. Innovations
Response of Base Money to core
Response of Base Money to exchange rate
Response of Base Money to teasury bill
Response of Base Money to Base Money
Response of Base Money to oil prices
0.10
0.08
0.06
0.04
0.02
0.00
-0.02
5 10 15 20 25 30 35 40 45 50 55 60

Table G
Variance Decomposition of Core Inflation
(%)

Periods	S.E.	Core Inflation	Exchange Rate	Treasury Bill	Base Money	Oil Prices
5	0.04	84.36	4.14	1.97	8.50	1.04
10	0.06	81.77	4.75	1.84	11.03	0.64
15	0.08	80.43	5.21	1.98	11.84	0.54
20	0.09	80.31	5.41	1.74	12.04	0.50
25	0.10	80.16	5.46	1.66	12.26	0.46
30	0.11	79.96	5.54	1.65	12.40	0.45
35	0.12	79.93	5.59	1.58	12.46	0.43
40	0.13	79.87	5.62	1.55	12.54	0.42
45	0.14	79.81	5.65	1.54	12.59	0.41
50	0.14	79.79	5.67	1.51	12.62	0.41
55	0.15	79.76	5.68	1.49	12.66	0.40
60	0.16	79.73	5.70	1.48	12.69	0.40

Table H
Model Forecast Evaluations.
In-sample Forecast.
Sample 1975:1 1996:04

Model	**MSE**	**RMSE**	**MAE**	**Theil U**	**Janus**
ARMA	0.00017	0.013	0.0092	0.16	0.07
VEC	0.00024	0.015	0.0118	0.17	0.80
Neural	0.00009	0.010	0.0071	0.03	0.47

Table I
Forecast Encompassing test results
Sample 1975:1 1996:4

Null Hypothesis	F-Statistics	P-Value
ARMA	3.9294	0.0254
ANN	0.3828	0.6832

APPENDIX 2
Methodological Review: Unit Roots & Cointegration

Unit Roots

A process whose characteristic root is equal to or greater than one is called a unit root. Unit roots are non-stationary and shocks to such series are permanent. There is no long run mean to which the series returns, and its variance is time-dependent.

When estimating a series with unit root, the assumptions of the classical regression model (OLS) are violated. According to Granger & Newbold (1956), what results is a spurious regression, in which the model has a good fit with variables appearing statistically significant, but the results are without any economic meaning.

The standard test for the presence of a unit root is the Augmented Dickey-Fuller (ADF) test, defined as follows:

$$\Delta y_t = a_0 + a_1 y_{t-1} + \sum_{j=1}^{r} \Delta y_{t-j} + \varepsilon_t$$

If a_1 is zero, the y_t process has a unit root. If a series is found to have a unit root, it has to be difference "d" times so as to make it stationary. A series that has to be differenced "d" times to make it stationary is said to be integrated of order d; {I(d)}.

Cointegration / Error Correction

Cointegration refers to a linear combination of non-stationary variables that are stationary. A unique feature of cointegrated variables is that their time paths are influenced by the extent of any deviation from long run equilibrium (Enders 1996). If an (n×1) vector; $x_t = (x_{1t,}\ x_{2t,}\ x_{nt,})$ is cointegrated, it has an error correction representation of the form:

$$\Delta x_t = \eta_0 - \eta x_{t-1} + \eta_1 \Delta x_{t-1} + \ldots + \eta_p \Delta x_{t-p} + \epsilon_t$$

where η_0 is an (n×1) vector of intercept terms, η_i are (n×n) coefficient matrices, η is a matrix with elements η_{jk} such that one or more of the η_{jk} is not equal to zero, and ϵ_t is an (n×1) vector with elements ϵ_{it}, where ϵ_{it} may be correlated with ϵ_{jt}.

The Johansen cointegrating test is one method of testing for cointegration. This methodology estimates the coefficient matrix in an unrestricted form, then test whether we can reject the restrictions implied by the reduced rank of the coefficient matrix. It also determines how many of the characteristic roots of the coefficient matrix is less than unity. The Johansen statistic is a likelihood ratio test, which is given as follows:

$$LR = -T \sum_{i=r+1}^{p} \ln(1 - \cup_i)$$

where 'r' is the number of cointegrating vector and L is the estimate of the population parameter. The null hypothesis is that there are at least "r" cointegrating vectors. The test is an iterative search process over the range of "r", which continues until the null hypothesis is accepted.

BIBLIOGRAPHY

Allen, Courtney, Research Department, Measuring Core Inflation in Jamaica, Bank of Jamaica, 1997.

Bond, Marian, Exchange Rates, Inflation, and the Vicious Circle, Finance and Development, March 1980, pg. 27 – 31.

Claus, Iris, A Measure of Underlying Inflation in the United States, Bank of Canada Working Paper 97-20, 1997.

Diebold, Francis X., Elements of Forecasting, University of Pennsylvania, South-Western College Publishing, 1998.

Enders, Walter, Applied Econometric Time Series, Canada, John Wiley & Sons, 1995.

Enders, Walter, Rats Handbook For Econometric Time Series, United States of America, John Wiley & Sons, 1996.

Engert, Walter & Scott Hendry, Forecasting Inflation with the M1- VECM: Part 2, Department of Monetary and Financial Analysis, Bank of Canada, 1997.

Gately, Edward, Neural Networks for Financial Forecasting, United States of America, John Wiley & Sons, Inc, 1996.

Holden, Ken, Vector Autoregression Modeling and Forecasting, Journal of Forecasting, Vol. 14, 159 – 166, 1995.

Norden, Simon van, A Note on Forecast Encompassing Tests, Canada, 1992.

Quah, Danny T, & Shaun P Valley, Measuring Core Inflation, Bank of England Working Paper Series NO. 31, 1995.

Remus, William & Marcus O'Conner, Neural Networks for Time Series Forecasting, Principles of Forecasting, 1997.

Robinson, Wayne, Research Department, Forecasting Inflation using VAR Analysis, Bank of Jamaica, 1996.

Rumelhart, David & James McClelland, Parallel Distributed Processing, United States of America, Massachusetts Institute of Technology, 1988.

Stanley, Jeannette, Introduction to Neural Networks, "Computer Simulations of Biological Intelligence", Printed in the United States of America, 1990.

Stennett, Robert, Research Department, Projecting Velocity in Jamaica, Bank of Jamaica, 1998.

Warren McCullock's and Walter Pitts', 'A logical calculus of ideas immanent in Nervous Activity' (Bulletin of Mathematical Biophysics Volume 5, pp 115-133),

Webb, Roy H., Forecasts of Inflation from VAR Models, Journal of Forecasting, Vol. 14, 267 – 285, 1995.