MPRA
Munich Personal RePEc Archive

Forecasting methods: a comparative analysis

Iqbal, Javed

Karachi University

2001

Online at https://mpra.ub.uni-muenchen.de/23856/
MPRA Paper No. 23856, posted 29 Aug 2013 14:31 UTC

Proc. Eighth Stat. Sem. 2001, K.U.
ISBN-969-8397-07-8-(189-197)

FORECASTING METHODS: A COMPARATIVE ANALYSIS

Javed Iqbal
Department of Statistics, University of Karachi

ABSTRACT

Forecasting is an important tool for management, planning and administration in various fields. In this paper forecasting performance of different methods is considered using time series data of Pakistan's export to United Sates and money supply. It is found that, like other studies of this nature, no single forecasting method provides better forecast for both the series. The techniques considered are ARIMA, Regression Analysis, Vector Autoregression (VAR), Error Correction Model (ECM) and ARCH/GARCH models.

1. INTRODUCTION

Forecasting is a probabilistic estimate or a description of a future value or condition, which includes a mean, range and probability estimate of that range. There is a considerable literature on forecasting in business and economics. Some are aimed at forecasting a particular variable of interest e.g. stock prices, money demand, exchange rates etc. for example Chan and Lee (1997), Kumar (1992), Bhawnani and Kadiyala (1997), Bleaney (1998). Others focus on a comparison of different models or techniques of forecasting, LeSage (1990), Stock and Watson (1996), Fair (1973). This study is of later style. A distinguishing feature of this study is the presentation of forecast confidence intervals besides mean forecasting. The models considered are univariate as well as multivariate. To broaden the scope, the time series considered are quarterly as well as annual. The annual data (1972-1999) for forecasting State Bank's money supply (M2) is used. Forecasting this variable is important for potential investors since money supply affects interest rate and consequently investment. It also has an impact on prices. An estimate of this variable is important for Central Bank itself so that to achieve its target of stabilizing the economy, some alteration in historical pattern of money supply could be made. Secondly quarterly data (1988:1 –1999:4) for forecasting Pakistan's export to U.S., a major trading partner of Pakistan, is considered. The export function involves a variable of economic activity of trading partner (U.S. in this case for which the data are available) instead of the home country. This variable is chosen keeping in view the availability of quarterly data. Since macroeconomic models usually involve a variable of real economic activity such as GDP, or its component. Unfortunately quarterly data for Pakistan on these real variables are not available. The data are collected from various issues of State Bank's monthly bulletin and IMF's International Financial Statistics. The results of almost all studies of this nature indicates that no single model camouflage as the best for out-sample forecasting of different series. This study also provides such an evidence.

<u>Export Function:</u>

$$E = f\,(YA, RP, V) \tag{1.1}$$

Where E: Pakistan's export to U.S (millions of U.S $), YA: United State's GDP (Billions of U.S.$ seasonally adjusted), RP: Relative prices computed as dividing unit value Index of export to U.S.

consumer price index. V: Exchange rate volatility computed as in Ariz and Shwiff (1998) as an eight quarter moving standard deviation i.e.

$$V_t = \left[\sum_{i=0}^{7}(R_{t-i} - \hat{R}_{t-i})^2 / 8\right]^{1/2} \quad (1.2)$$

Where R_t is the exchange rate (Rs per U.S.$) and $\hat{R}_t$ is its forecasted value (ARIMA model has been used to forecast it)

Expected signs of the coefficients of the explanatory variables are as follows:

$$\frac{\partial E}{\partial YA} > 0, \quad \frac{\partial E}{\partial RP} < 0, \quad \frac{\partial E}{\partial V} < 0,$$

Money Demand function:

$$M = f(Y, I) \quad (1.3)$$

M: Money balances (M2) (millions of 1981 RS), Y: GDP (millions of RS) at 1981 prices, I: Nominal interest rate on 2-3 years deposits.

Expected signs are as follows:

$$\frac{\partial M}{\partial Y} > 0, \quad \frac{\partial M}{\partial I} < 0 \quad .$$

Following this introduction, section 2 describes forecasting models considered in this study. Out-sample forecasting is compared in section 3 and section 5 provides conclusion and comments.

2. FORECASTING MODELS

a) **ARIMA** One of the most popular univariate forecasting model proposed by Box and Jenkins (1970). For a stationary time series y_t, an ARMA(p,q) model is expressed as

$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + ... + \phi_p y_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + ... + \theta_q \varepsilon_{t-q} \quad (2.1)$$

where ε_t is a white noise disturbance term normally and independently distributed with mean 0 and variance σ^2.

This model can be expressed as weighted sum of disturbances ε_t as

$$y_t = \varepsilon_t + \psi_1 \varepsilon_{t-1} + \psi_2 \varepsilon_{t-2} + ... + \psi_k \varepsilon_{t-k} + ... \quad (2.2)$$

Where ψ weights are functions of the modal parameters $\phi's$ and $\theta's$.

An h-step ahead forecast error variance FEV (h) for y is given by

$$\text{FEV (h)} = (1 + \psi_1^2 + \psi_2^2 + ... + \psi_{h-1}^2)\sigma^2 \quad (2.3)$$

A 95% forecast confidence interval for h-step ahead forecast is given by

$$\hat{y}_{t+h} \pm 1.96\sqrt{FEV(h)} \quad (2.4)$$

b) **Regression Analysis** A general linear regression model is given by:

$$y_t = \beta_1 + \beta_2 x_{2t} + \beta_3 x_{3t} + \dots\dots\dots\dots\dots + \beta_k x_{kt} + \varepsilon_t \quad (2.5)$$

Which in matrix form is given by

$y = x\beta + \varepsilon$. Where the dimensions of y is n × 1, that of x is n × k, of β is k × 1 and ε is an n × 1 vector of white noise errors with covariance matrix $\sum = \sigma^2 I$

Let $x_0 = [1 \ x_{02} \ x_{03} \dots x_{0k}]'$ be the vector of values for which forecast is required, then $\hat{y}_{0/x0} = x_0'\hat{\beta}$. Forecast error variance FEV (h) is given by

$$FEV(h) = \sigma^2[1 + x_0'(x'x)x_0] \quad (2.6)$$

c) **ARCH/GARCH model** The autoregressive conditional hetroskedasticity (ARCH) model introduced by Engle (1982) is extensively used to model financial time series which are believed to have varying conditional variance or volatility e.g. stock prices, interest rate, inflation etc. In this method, mean and variance are modeled simultaneously. An ARCH model of order q is expressed as

$y_t = x_t\gamma + \varepsilon_t$ Where ε_t ~N $(0, \sigma_t^2)$ and

$$\sigma_t^2 = \omega + \alpha_1\varepsilon_{t-1}^2 + \alpha_2\varepsilon_{t-2}^2 + \dots\dots + \alpha_q\varepsilon_{t-q}^2 \quad (2.7)$$

ARCH model were generalized by Bollerslev (1986) as GARCH (Generalized ARCH), in which the conditional variance depends on past value of itself. The conditional variance in a GARCH (p,q) model is expressed as

$$\sigma_t^2 = \omega + \alpha_1\varepsilon_{t-1}^2 + \alpha_2\varepsilon_{t-2}^2 + \dots\dots + \alpha_q\varepsilon_{t-q}^2 + \beta_1\sigma_{t-1}^2 + \beta_2\sigma_{t-2}^2 + \dots\dots\dots + \beta_p\sigma_{t-p}^2 \quad (2.8)$$

The most widely used of these models is GARCH (1,1) for which the conditional variance is given by

$$\sigma_t^2 = \omega + \alpha\varepsilon_{t-1}^2 + \beta\sigma_{t-1}^2 \quad (2.9)$$

The forecast error variance in GARCH models is given by σ_t^2

d) **Vector Autoregression (VAR)** VAR introduced by Sim (1980), is used to model interrelated time series simultaneously. This model, unlike structural multi-equation econometric models, relies little on economic theory. Each variable in the model is expressed as function of the lagged values of all the endogenous variables in the system. A VAR of order p is given by

$$y_t = A_1 y_{t-1} + A_2 y_{t-2} + ... + A_p y_{t-p} + \varepsilon_t \tag{2.10}$$

Where ε_t ~N (0,Ω) is a vector of white noise errors with covariance matrix Ω. The errors are assumed to be serially independent in each equation but they may be contemporaneously correlated across equations. A_i's (i =1 ,2,…,p) are the matrices of the coefficients to be estimated. y_t is a vector endogenous variables. For our export model y_t = [E, YA, RP, V]′. A VAR model, unlike structural multi equation modes, does not require the knowledge of explanatory variable for forecast period. An h-step ahead forecast of y_t is

$$\hat{y}_{t+h} = A_1 \hat{y}_{t+h-1} + A_2 \hat{y}_{t+h-2} + ... + A_p y_{t+h-p} \tag{2.11}$$

Where $y_{t+h-1}, ..., y_{t+1}$ are forecasted using similar scheme. For a VAR (1), we have $\hat{y}_{t+h} = A\hat{y}_{t+h-1}$ and forecast error variance (as given in Clement and Hendry (1993)) is

$FEV(h) = \sum_{j=0}^{h-i} A^j \Omega (A^j)'$. With normality assumption, we can derive simple expression for the forecast intervals for individual time series by considering the diagonal elements of FEV (h) matrix.

e) **<u>Error Correction Model (ECM)</u>** It is widespread in modern econometric literature that a regression involving non-stationary time series variables may be spurious or misspecified unless the variables are Co-integrated. The time series x_{1t} and x_{2t} each integrated of order 1, are said to be co-integrated if any linear combination if these e.g. ($z_t = \alpha_1 x_{1t} + \alpha_2 x_{2t}$) is stationary. The concept of co-integration in statistics is equivalent to stable long run relationship in economics. The short run disequilibrium can be modeled as an error correction model.

To get the idea lets consider a time series regression

$$y_t = \alpha + \beta' x_t + \varepsilon_t \tag{2.12}$$

Where y_t is the dependent variable and x_t is a vector of independent variables. If y_t and x_t are in equilibrium, then the error $y_t - \alpha - \beta' x_t$ will equal zero. However in disequilibrium it will be non-zero. This quantity measures the extent of disequilibrium between y_t and x_t and hence is called disequilibrium error. In disequilibrium y_t can be assumed to be related with x_t and the lagged values of y_t and x_t, one typical form of which is

$$y_t = \gamma + \delta_0 x_t + \delta_1 x_{t-1} + \pi y_{t-1} + \varepsilon_t \tag{2.13}$$

Subtracting y_{t-1} from each side of eq (2.13) above and reparametezing we get

$$\Delta y_t = \delta_0 \Delta x_t - \mu (y_{t-1} - \alpha - \beta x_{t-1}) + \varepsilon_t \tag{2.14}$$

Where the parameter μ depends on parameters in eq (2). The term in bracket is called the error correction term, which incorporates past period's disequilibrium. Eq (2.14) is an example of error correction model. The estimation of ECM involves two steps (Engle and Granger (1987)). First the long run equilibrium relationship between non-stationary variables is estimated. In our present case this would be the regression E = f (YA, RP, V) for export function. In the second step ECM model eq (2.14) is estimated using residuals lagged one period from the long run regression. Forecasting of y can then be performed using the ECM model and the forecast error variance is given as eq (2.6).

3- FORECAST COMPARISON

The comparison of forecasting models is based on Root Mean Square Error (RMS)

$$\text{RMS} = \sqrt{\sum_{i=1}^{h} (y_{t+i} - \hat{y}_{t+i})^2 / h} \tag{3.1}$$

and Theil's Inequality Coefficient (U)

$$U = \frac{\sqrt{\sum_{i=1}^{h} (y_{t+i} - \hat{y}_{t+i})^2 / h}}{\sqrt{\sum_{i=1}^{h} y_{t+i}^2 / h} + \sqrt{\sum_{i=1}^{h} \hat{y}_{t+i}^2 / h}} \tag{3.2}$$

The former depends on the unit of measurement of the variable to be forecasted and later is unit free lying between 0 and 1 with the forecast getting better as the U is closer to 0.

These measures are widely used for out-sample forecasting e.g. in Chan and Lee (1997) and Kumar (1992) among others.

Table-I: Forecast Evaluation for Export and Money Supply Series

Models	Export		Money Supply	
	RMS	Theil Coefficient	RMS	Theil Coefficient
Regression	39.975	0.0417	5866.57	0.0102
ARIMA	28.527	0.0304	27853.21	0.0468
ARCH	36.330	0.0381	5414.98	0.0094
VAR	35.421	0.0367	22119.91	0.0374
ECM	96.092	0.0948	26948.88	0.0452

Table-I presents RMS and U- measures for the models forecasted. It can be seen that ARIMA model outperforms the others for exports and a GARCH (1,1) model appears to be the best for money supply. These results are consistent with other studies, for example Delurgio(1998) and Bleaney(1998), which indicate that a univariate model usually perform better for monthly or quarterly data and annual data usually are best forecasted by a form of regression model. The ARCH model requires Knowledge of future values of the explanatory variables for mean estimation, however main use of these modes has been for forecasting volatility which, fortunately, does not require forecast of explanatory variables.

The conclusion is further supported by fig. No 1 to 4, which display graphically the realized and forecast from different models and approximately 95% forecast confidence intervals.

Table-II: Augmented Dickey Fuller Test of Unit Root

	Export Model				Money Supply Model		
Series	E	YA	RP	V	M	Y	I
Level	-4.807	-0.747	-1.258	-3.76	-2.702	-2.672	-1.493
Ist Difference	-7.46**	-5.413**	-6.48**	-7.58**	-4.426**	3.517*	-5.082**

(* indicates significant at 5% level ** indicates significance at 1% level)

It appears that all the series are stationary at the first difference

Fig.1 Realized And Forecasted Export

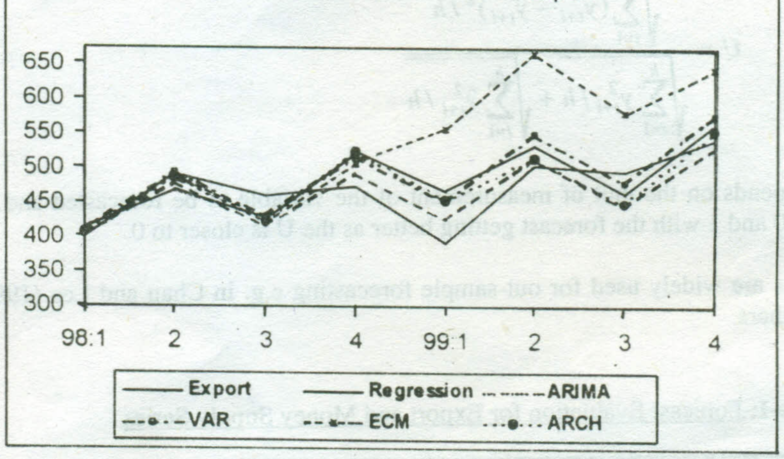


Fig.2 Realized And Forecasted Money Supply

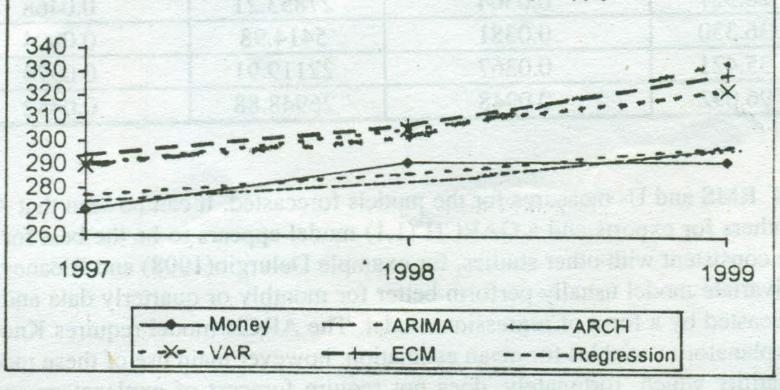


Fig. 3: 95 % Forecast Intervals From Different Models Of Export

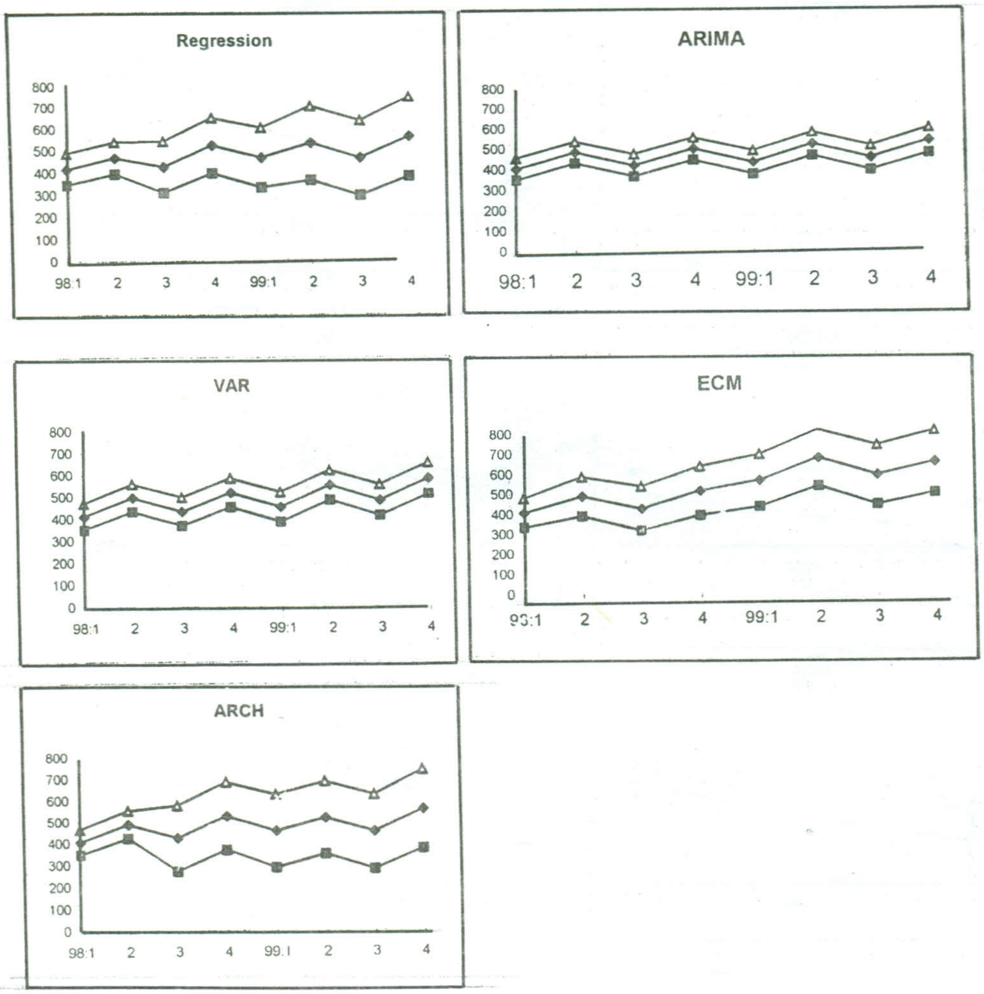
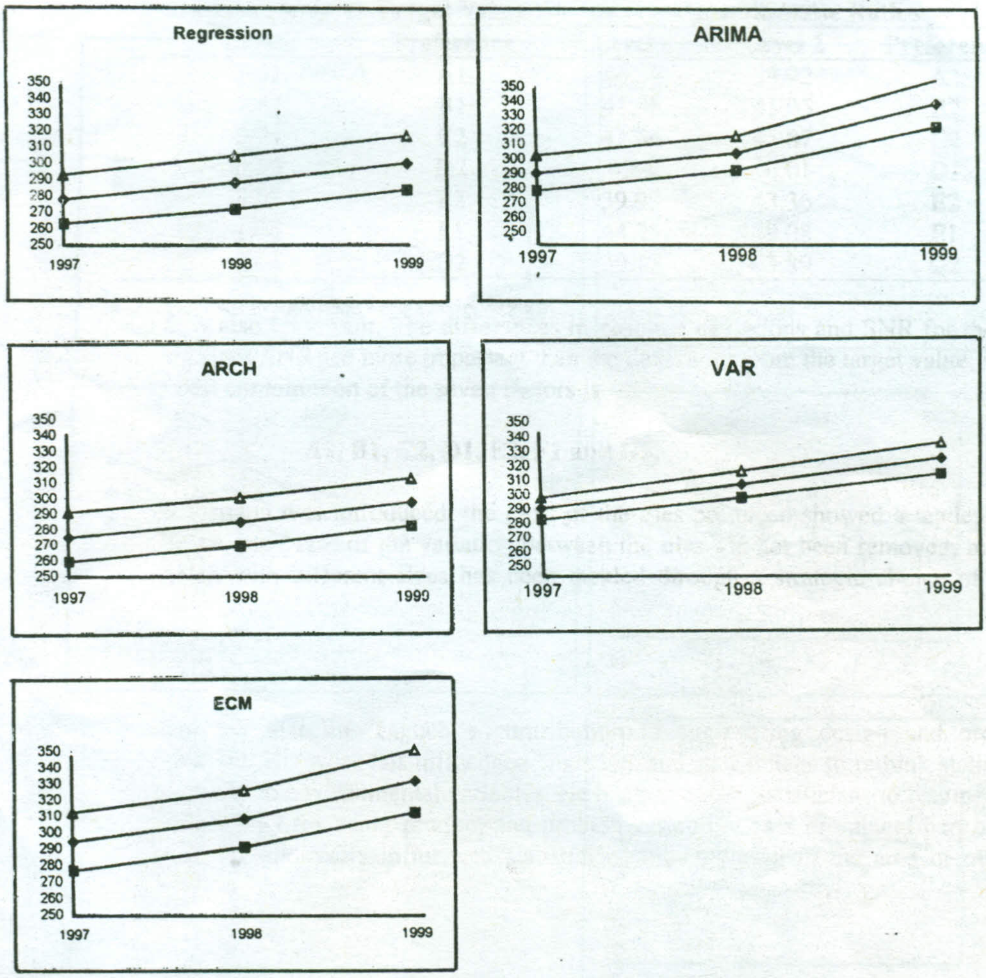


Fig. 4: 95 % Forecast Intervals From Different Models Of Money Supply



4. CONCLUSIONS

A comparison of forecasting performance of different time series econometric models is considered using data from Pakistan's macroeconomy. Like previous studies no single model appeared the best for both the variables forecasted. The message revealed by the study is that to forecast a particular variable one should not rely on a single forecasting method rather performance of different competing models should be checked, then forecast should be performed from the best models from the comparison.

REFERENCES

1. Ariz, A.C. and Shwiff, S.S. (1998). Does exchange rate volatility affects import flows in G-7 countries, evidece from Co-integration model. *Applied Economics,* 30, 1269-1276
2. Bhawnani, V. and Kadiyala K.R.(1997). Forecasting foreign exchange rates indeveloping economies. *Applied Economics*, 29, 51-62
3. Bleaney,M.F. (1998). Exchange rate forecasts at the long horizon:Are error correction models superior?. *Canadian Jounal of Economics*, 31(4), 853-864
4. Bollersev, T. (1986). Generalized Autoregressive Conditional Hetroskedasticity. *Journal of Econometrics*, 31, 307-327
5. Box, G.E.P. and Jenkins C.M. (1970). Time Series Analysis, forecasting and Control. Holden day, San Francisco.
6. Chan, H.L. and Lee, S.K. (1997). Modelling and forecasting the demand for coal inchina. *Energy Economics*, 19, 271-287
7. Clements, M.P. and Hendry, D.F. (1993). On the limitation of Comparing Mean Square forecasting error. *Journal of Forecasting*, 12, 617-637
8. Delurgio, S.A. (1998). Forecasting Principles and Application. Irvin/McGraw Hill, Singapore.
9. Enble, R.F. and Granger, C.W.G. (1987). Co-integration and Error Correction: Represenration, Estimation, and Testing. *Econometrica,* 55, 251-276
10. Engle, R.F. (1982). Autoregressive conditional Hetroskedasticity with estimation of variance of U.K. Inflation. *Econometrica,* 50 ,987-1008
11. Engle, R.F. and Yoo, B.S. (1987). Forecasting and testing in Co-inegrating system, *Journal of Economtrics*, 35, 143-159
12. Fair, R. (1973). A comparison of alternative estimation of macro models. *International Economic Review*, 14,(2), 261-273
13. Gujrati, D.N. (1995). *Basic Econometrics*. Mc-Graw Hill
14. Johansen, S. (1988). Statistical Analysis of Co-inegration Vectors. *Journal of Economic Dynamics and Control,* 12, 231-254
15. Kumar, M.D. (1992). The forecasting Accuracy of Crude Oil Futures prices. *IMF Staff Pappres*, 39(2), 432-461
16. LeSage, J.P. (1990). A comparison of the forecasting ability of ECM and VAR models. *Review of Economics and Statistics*, 72(4), 664-671
17. Sim, C.A. (1980). Macroeconomics and reality. *Econometrica,* 48, 1-48
18. Stock, J.H. and Watson, M.W. (1996). Evidence on Structural Instability in Macroeconomic relations. *Jouranal of American Statistical associatian,* 14(1), 11-30
19. Thomas, R.L. (1997). Modern Econometrics. An introduction. Addison-Wesley, England.