ISSN 2229-5518

Study of Inflation in Pakistan Using Time Series ARMA, ARFIMA and GARCH Models

Arfa Maqsood, Syed Muhammad Aqil Burney,Suboohi Safdar

Abstract— This paper presents the statistical approach to index numbers such as CPI (Consumer Price Index) to measure the inflation rate in Pakistan. We use various statistical time series models such as Autoregressive moving average (ARMA) model, Autoregressive fractionally integrated moving average (ARFIMA) model, and Generalized autoregressive conditional heteroscedasticity (GARCH) model based on consumer price index number computed by the Laspayer's formula. Based on some accuracy measures the suitable model is estimated in order to estimate and forecast the rate of inflation in Pakistan.

Index Terms— Consumer Price Index (CPI), Inflation Rates, ARMA model, ARFIMA model, GARCH model.

1 Introduction

The term Inflation is used to represent an increase in the general price level, measured against a standard level of purchasing power. Wilson [1] explains the inflation as the condition of continued rising prices in which each unit of money buys fewer goods and services, hence results to an erosion in the purchasing power of money. Inflation rate is measured as the percentage change in the average level of prices. Inflation rate rises and falls over the years but it rarely becomes negative. If it is negative, indicating a fall in the average price level.

A large number of empirical studies are available on inflation and monetary factors influencing the rate of inflation such as growth in money supply, credit to the private sector, exchange rate, interest rate, 6-months Treasury bill rates, out gap etc. However these studies do not concur on the relative importance of each of these factors as determinants of inflation.

Several methods have been used to estimate inflation rate based on both univariate and multivariate approaches. Nasim [2] and Hossain [3] find money supply as the principal factor underlying the rising inflation in Pakistan. Khan and Qasim [4] consider a model of inflation as the linear combination of weighted average of price of tradable goods and price of non-tradable goods. These methods basically depends on searching the economic factors (variables) that influence the underlying rate of inflation and study the relationship between them using a suitable model. These models will explain deterministic and non-deterministic relation between the variables under consideration. Schimmelpfennig and Madhavi [5] present three approaches Leading indicator, univariate approach, and vector autoregressive (VAR) model to estimate and forecast inflation in Pakistan. See Schimmelpfennig and Madhavi [5] for more information on study of measurement of inflation with reference to Pakistan.

The objective of this paper is to estimate an overall inflation rate along with its two broad components i.e. CPI food and Non-food price inflation using univariate time series models.

Four different indices are published in Pakistan i.e. the consumer price index (CPI), the wholesale price index (WPI), the sensitive price index (SPI), and GDP deflator. Based on 2000-01 base period, the CPI covers the retail prices of 374 items in 35 major cities and reflects roughly the cost of living in the urban areas. The WPI covers the wholesale prices of 425 items prevailing in the city of origin of the commodities. The SPI covers prices of 53 essential items consumed by those households whose monthly income ranges from Rs. 3000 to Rs. 12000 per month. The details are documented in table 1.

TABLE 1
YEARLY CPI INFLATION, FOOD INFLATION, AND NON-FOOD INFLATION

Features	CPI	WPI	SPI
Cities covered	35	18	17
Markets covered	71	16	51
Items covered	374	425	53
No. of commodity group	10	5	-
Income groups	4	-	3000/month
Reporting frequency	Monthly	Monthly	Weekly

In Pakistan the main focus for assessing inflationary trends is placed on CPI because it closely represents the cost of living, hence called cost of living index number. The consumer price inflation is computed by the following Laspeyre's formula

$$CPI = \sum_j w_{0j} \frac{p_{tj}}{p_{0j}}$$

Where p_{tj} and p_{0j} are the prices of commodity j in current period t and base period 0 respectively.

The rest of the paper is organized as follows. Section 2 discusses the recent inflationary trends in Pakistan over the period 1997-98 to 2012-13. Section 3 presents brief review of the statistical models to univariate approach. The analysis of monthly inflation rates over the period of July 1997 to June 2012 is presented in section 4. Lastly, concluding remarks is provided in section 5.

2 Inflationary Trends (1998-2013)

Pakistan has experienced sustained inflation between 7.8 to 4.6 percent ranges during the six years from 1997 to 2003. It reaches at the increase of 9.3 percent in the year 2004-05 as against 4.6 percent in the corresponding period of last year. In

2005-06, inflation has climbed down to 8.0 percent and 7.8 percent over the next period 2006-07. After remaining low and persistence single digit inflation over the ten years period from 1997-98 to 2006-07, it accelerated in double digit in the period of 2007-08 and in next fiscal year reached at its highest value of 21.0 percent, due to mainly the rise in the prices of food items, so it becomes a burden borne disproportionately by the fixed income group and poor of the country. This increase shows 75 percent increase of inflation in the previous year. Afterwards, it decreases gradually and again reduces to single digit in 2012-13. Figure 1 shows the inflationary trends in overall and its major components food and non-food groups.

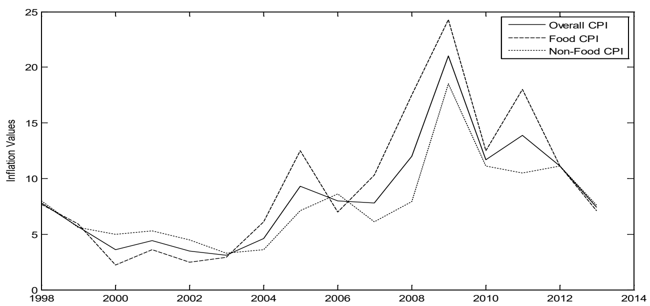

Fig 1. Yearly inflationary trends

The single largest component of CPI is the food group which showed increase of 12.5 percent in 2004-05 as more than twice of the increase 6.1 percent in same period last year. This higher level of increase in food inflation is attributed to an increase in prices of wheat, wheat flour, meat, rice, vegetable ghee, and cooking oil. The highest food inflation is 24.3 percent in the period 2008-09 that is more than twice of the rate in the second preceding year 2006-07. However, food inflation reduced to only 7.1 percent in the last fiscal year 2012-13. On the other hand, the non-food inflation grew relatively at a faster pace. The non-food inflation in 2008-09 is 18.5 percent, while it was only at 6.1 percent in period 2006-07. In the period of last four years non-food inflation rate has a significant deceleration and averaged at 10.07 percent. It is due to mainly because of tight monetary policy, careful fiscal management and improved supply of food items in the country. Although the adjustment in exchange rate and the rise in international prices of petrol, oil, lubricants (POL) products have put upward pressures on inflation but these pressures were countered by tight monetary policy. Table 1 presents the yearly overall inflation, food inflation, and non-food inflation from 1997-98 to 2012-13.

3 Univariate Time Series Models

Mostly, we are dealing with the dilemma of estimation and forecasting of univariate time series. Many financial econometrics time series such as exchange rate, share index, inflation rate, are essential to study the behavior and pattern that the series exhibits in order to have the economic scenario of a country. For this purpose, various time series models have been established and widely used. We study three popular methods to model the rate of inflation, ARIMA the simplest of all, ARFIMA based on fractional differencing, and GARCH which is suitable in the presence of volatility. We briefly define all of them to give a review to readers in the following sub-sections.

TABLE 2
YEARLY CPI INFLATION, FOOD INFLATION, AND NON-FOOD INFLATION

	CPI Inflation		
Years	Overall	Food	Non-Food
1997-98	7.8	7.7	8.0
1998-99	5.7	5.9	5.6
1999-00	3.6	2.2	5.0
2000-01	4.4	3.6	5.3
2001-02	3.5	2.5	4.5
2002-03	3.1	2.9	3.3
2003-04	4.6	6.1	3.6
2004-05	9.3	12.5	7.1
2005-06	8.0	7.0	8.6
2006-07	7.8	10.3	6.1
2007-08	12.0	17.5	7.9
2008-09	21.0	24.3	18.5
2009-10	11.7	12.5	11.1
2010-11	13.9	18.0	10.5
2011-12	11.1	11.1	11.1
2012-13	7.4	7.1	7.6

3.1 ARMA / ARIMA Models

Box and Jenkins [6] combined the idea of autoregressive (AR) and moving average (MA) term, so called ARMA model. If a series has to be differenced d times to make it stationary then apply ARMA (p, q) model to the differenced stationary series becomes autoregressive integrated moving average ARIMA (p, d, q) model. Where p is the number of autoregressive terms, d is integer differencing parameter and q is the number of moving average terms. For detailed description on this process see Abraham and Ledolter [7], and Brockwell and Davis ([8], [9]). An ARIMA (p, d, q) model is written by

$$\phi(B)(1-B)^d (X_t - \mu) = \theta(B) Z_t \qquad (1)$$

There are methods to estimate the ARIMA model, for example Burney [10] computed exact likelihood function for the special case of this process. Many statistical software have made the computation process very easy to estimate and forecast the ARIMA model.

3.2 ARFIMA Models

The acronym ARFIMA stands for "Autoregressive Fractionally Integrated Moving Average", first introduced by Granger [11], and Granger and Joyeux [12]. A fractionally integrated model aims to capture the long run behavior that appears in a time series. It is the process where the influence of shocks in a stationary [I(0)] model disappears after a limited number of periods (depending on the short memory parameters in the autoregressive and moving average terms), and where the effect of a shock lasts forever in a unit root [I(1)] process. The fractionally integrated model [FI(d)] with $d \in (0,1)$ takes up an intermediate position, see Hosking [13] and Baillie [14]. ARFIMA model is defined as

$$\phi(B)(1-B)^d (X_t - \mu) = \theta(B) Z_t \quad 0 \le d \le 1 \qquad (2)$$

Where $\phi(B) = 1 - \phi_1 B - \ldots\ldots - \phi_p B^p$ autoregressive polynomial, $\theta(B) = 1 - \theta_1 B - \ldots\ldots - \theta_q B^q$ moving average polynomial, and Z_t assumed to be Gaussian with expectation zero and variance σ_Z^2. The long memory behavior is governed by the factor $(1-B)^d$.

The population characteristics of ARFIMA process have been extensively studied by Hosking [13], and later on summarized by Tsay [15]. For -0.5<d<0.5, the process X_t is both stationary and invertible and moving average coefficients decay hyperbolically, rather than showing exponential decay characteristic of an ARIMA (p, 0, q) process. To estimate an ARFIMA model, start with the estimation of the fractionally differenced parameter d (see Maqsood and Burney [16] for estimation of d). One of the methods of estimation of d is the regression approach proposed by Geweke and Porter [17], based on the spectral density given below;

$$f(\lambda) = \left|1 - e^{-i\lambda}\right|^{-2d} f_u(\lambda) \qquad (3)$$

Where $f_u(\lambda)$ is the spectral density of the ARMA (p, q) process. Setting $U_t = \phi(B) X_t = \theta(B) Z_t$, we get the following expression

$$f_u(\lambda) = \left|\phi(e^{-i\lambda})\right|^2 f_x(\lambda) = \left|\theta(e^{-i\lambda})\right|^2 f_Z(\lambda) \qquad (4)$$

After having some algebraic manipulation, we have the simple linear regression equation.

$$Y_j = a + bx_j + \varepsilon_j \; ; \; \mathrm{j} = 1, \ldots, \mathrm{m} \qquad (5)$$

Where, $Y_i = ln I_n(\omega_i)$, $x_i = ln\left|1 - e^{-i\omega_j}\right|^2$, $\varepsilon_i = \ln(I_n(\omega_i)/f(\omega_j))$, a= $ln f_u(0)$, and $b = -d$. See Brockwell and Davis [9] for more details about the derivation. The method of least squares gives estimate of d as follows;

$$d = -\frac{\sum_{j=1}^{m} (x_j - \bar{x})(y_j - \bar{y})}{\sum_{j=1}^{m} (x_j - \bar{x})^2} \qquad (6)$$

3.3 GARCH Models

Many Financial time series often exhibit the phenomenon of volatility clustering, that is periods in which their values show wide swings for an extended time period followed by the period in which there is relative calm in rates, such as stock prices, inflation rates, exchange rates etc.

Autoregressive Conditional Heteroscedasticity ARCH (p) models capture this volatility clustering, originally developed by Engle [18]. The ARCH (p) model is defined as follows

$$\sigma_t^2 = \alpha_0 + \alpha_1 Z_{t-1}^2 + \ldots + \alpha_p Z_{t-p}^2 \qquad (7)$$

The conditional variance of Z_t at time t depends on p lagged of the squared error terms. If the conditional variance of Z_t at time t depends not only on the squared error terms in the p previous time periods but also its conditional variance in q previous time periods. This model can be generalized to a GARCH (p, q) model, originally proposed by Bollerslev [19], written as

$$\sigma_t^2 = \alpha_0 + \alpha_1 Z_{t-1}^2 + \ldots + \alpha_p Z_{t-p}^2 + \beta_1 \sigma_{t-1}^2 + \ldots + \beta_q \sigma_{t-q}^2 \qquad (8)$$

Where $Z_{t-p}^2 \Rightarrow$ p lagged squared error term, and $\sigma_{t-q}^2 \Rightarrow$ q lagged conditional variance.

4 Analysis of Inflation Rate

This section presents an analysis to estimate and forecast monthly inflation rate (general, food, and non-food) using the three time series models that have been briefly discussed in section 3. The data consists of 192 observations in all from the period July 1997 to June 2013, which are recorded in inflation monitor, published by State Bank of Pakistan (SBP).

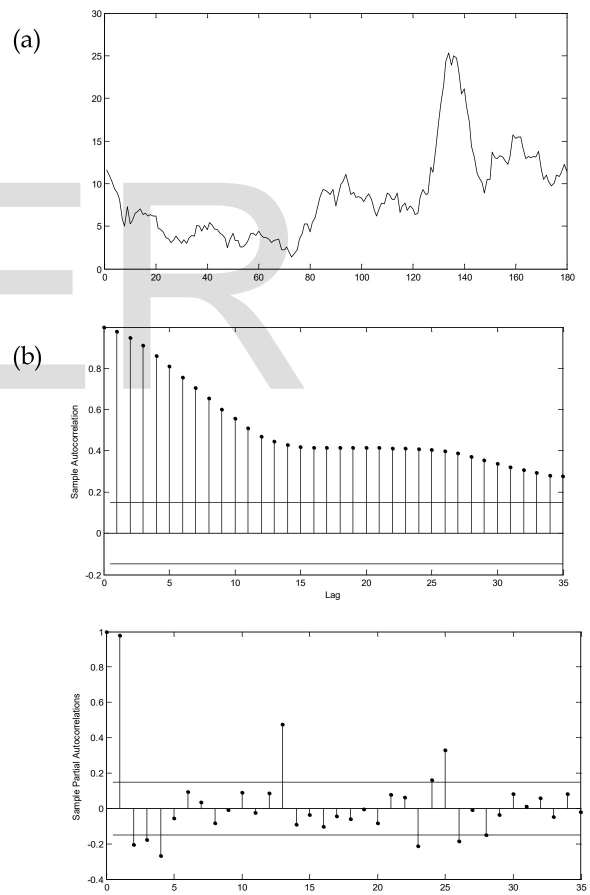

Fig. 2(a). Plot of monthly inflation series (b). ACF/PACF plot

We consider 180 observations from July 1997 to June 2012 are in-sample for the purpose of constructing models and rest for July 2012 to June 2013 are used to compare with forecasting results. The plot of original series X_t versus observation numbers along with its autocorrelation function (ACF) and partial autocorrelation function (PACF) are shown in figure 2 (a-b). It exhibits an increasing pattern until the end of year

2008 and then decreasing up to the period under consideration. Next, we observe the ACF decays at relatively slow rate, exhibiting a significant dependence between the observations. While, PACF displays six prominent spikes that are out of 95% confidence bands, confirming the contribution of autoregressive component in estimation of X_t.

TABLE 3
DESCRIPTIVE STATISTICS

Descriptive Statistics	Overall CPI	CPI Food	CPI Non-Food
Mean	8.5	9.60	7.71
Median	7.7	8.55	6.7
Max	25.3	34.1	20.2
Min	1.4	-0.6	1.9
St. Dev	5.12	7.17	4.03
Skewness	1.24	1.24	1.18
Kurtosis	4.56	4.69	4.28
N	180	180	180

Table 3 presents some descriptive statistics of the underlying series. The mean of overall inflation rate is 8.505 and the standard deviation is 5.12. One important step in analyzing a time series is to test for unit root or stationary stipulation of the series. To test the null hypothesis of unit root, we employ Augmented Dickey Fuller (ADF) test and Phillips Perron (PP) test. Results of test statistics and critical values are summarized in table 4. For the actual series X_t, both ADF and PP test statistics provide statistical evidence of acceptance of null hypothesis, implying that the series is non-stationary at level and hence need to be differenced.

TABLE 4
TEST OF STATIONARITY FOR VARIOUS TIME SERIES

Time Series	Test Statistic	P-value	5%	10%
	Augmented Dickey Fuller Test			
Overall Inflation	-2.5196	0.318	-3.436	-3.142
First Order Differenced	-12.697**	0.000	-3.435	-3.141
Fractional Differenced	-9.2752**	0.000	-3.435	-3.141
	Phillips Perron Test			
Overall Inflation	-3.107	0.107	-3.435	-3.141
First Order Differenced	-12.722**	0.000	-3.435	-3.141
Fractional Differenced	-9.0803**	0.000	-3.435	-3.141

*Significant at levels of 10%
**Significant at levels of 5%

To stabilize the variability in series, it is recommended to work on logarithm of series rather than the original series. The log series is differenced at lag1 to make it stationary. The first order differenced series, denoted by Y_t, along with ACF/PACF plot is shown in figure 3 (a-b). The jumps are shown around the centre point at zero showing the series is stationary at first order difference, which is also confirmed from the tests of stationary in table 4. By the inspection of ACF and PACF plots, we determine the observed number of autoregressive and moving average lags of underlying series.

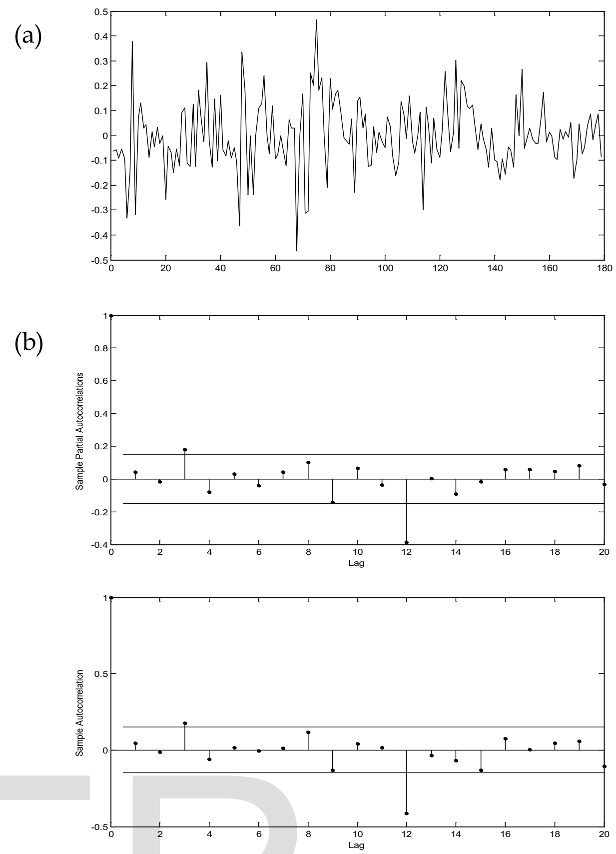

Fig. 3(a) Plot of first-order differenced series (b). ACF/PACF plot

TABLE 5
PERFORMANCE OF ARIMA, GARCH, AND ARFIMA MODELS

Models	AIC	SC	RMSE	S.E (Estimate)
Overall CPI				
ARIMA(1, 1, 1)	-1.098	-1.045	0.137	1.721
GARCH(1, 0)	-1.09	-1.02	0.140	1.70
ARFIMA(1, 0.4, 8)	-1.10	-0.93	0.130	1.18
Food CPI				
ARIMA(1, 1, 1)	1.14	1.20	0.42	1.55
GARCH(1, 0)	0.79	0.88	0.42	1.55
ARFIMA(1, 0.4, 1)	1.61	1.21	0.43	1.16
Non-Food CPI				
ARIMA(1, 1, 1)	-0.95	-0.92	0.15	2.06
GARCH(1, 0)	-1.18	-1.09	0.16	1.54
ARFIMA(1, 0.37, 4)	-0.91	-0.80	0.15	1.31

Based on some criterion functions and accuracy measures such as Akaike information criterion suitable (AIC), Shewarz criterion (SC) , and root mean square error (RMSE) , the ARIMA(1,1,1) can be regarded as an appropriate model. The values are shown in table 5. The estimated ARIMA(1,1,1) model is

$$Y_t = 0.01013 + 0.9305 Y_{t-1} - 0.9877 Z_{t-1}$$

The ACF plot of original series in figure 2 shows declination at relatively slow rate might reveal the presence of long memory in series. We, therefore apply the technique of estimating an ARFIMA model to the log series to capture this pattern. At first, we estimate the differencing parameter d through regression approach used by Brockwell and Davis [9] and equal to 0.4. We denote fractionally differenced series by U_t. The methodology developed by Box and Jenkins is employed to establish an ARMA model to fractionally differenced time series U_t. The estimated ARFIMA (1, 0.4, 8) model based on accuracy measures given in table 5, is written as follows.

$$\phi(B)(1-B)^{0.4}U_t = \theta(B)Z_t$$

Where, $\phi(B) = 1 - 0.6288B$ and $\theta(B) = 1 - 0.057B - 0.039B^2 + 0.173B^3 + 0.13B^4 + 0.1B^5 + 0.226B^6 + 0.12B^7 + 0.297B^8$

Lastly, these models are compared with that of GARCH model suitable in the presence of volatility. The heteroscedasticity in original series depicted in figure 2, is observed as we have the periods in which wide swings followed by the period in which relatively short swings. We, therefore, estimate GARCH process to model the variance of error term and determine GARCH(1,0) or simply an ARCH(1) model. It indicates the error variance is not constant rather depends on one step preceding square error term. The formulation is given below

$$Var(Z_t) = 0.0157 + 0.19Z_t^2$$

The forecast values from all these three models are obtained for one-year period from July 2012 to June 2013 and presented in table 6. We observe the values obtained from ARIMA and GARCH models are close to each other and to the actual inflation rates. While, ARFIMA model provides the forecast with minimum value of standard error of estimate. One can simply observe the performance of these models based on accuracy measures which are presented in table 5. In accordance with the criterion function, GARCH model is more appropriate as having the least values of AIC and SC. Whereas, ARFIMA is regarded as an adequate process to capture the variability of inflation time series if RMSE and SE (estimate) are taken into account.

TABLE 6
FORECAST OVERALL INFLATION RATES USING ARIMA, ARFIMA, AND GARCH MODEL

Months	CPI (True)	Forecasted Values		
		ARIMA	GARCH	ARFIMA
Jul-12	9.60	11.42	11.30	10.43
Aug-12	9.10	9.70	9.60	9.73
Sep-12	8.80	9.19	9.10	9.32
Oct-12	7.70	8.89	8.80	9.04
Nov-12	6.90	7.78	7.70	8.44
Dec-12	7.90	6.97	6.90	7.87
Jan-13	8.10	7.98	7.90	8.07
Feb-13	7.40	8.18	8.10	8.13
Mar-13	6.60	7.48	7.40	7.84
Apr-13	5.80	6.67	6.60	7.39
May-13	5.10	5.86	5.80	6.86
Jun-13	5.90	5.15	5.10	6.34

We extend our analysis to two broad components of inflation rates i.e. food inflation and non-food inflation. The same process of fitting these models are applied to first order differenced and fractionally differenced food and non-food inflation series. We just show the performance of models using different accuracy measures and criterion functions in table 5. GARCH process is more robust to employ estimating food and non-food inflation rates as it has the smaller values of accuracy measures than other models.

5 CONCLUSION

This paper presented an empirical application of financial time series processes to the inflation rates in Pakistan. On the basis of 180 observations, we analyzed the inflation series using three univariate approaches. Based on some decision measures, presented in table 5, the GARCH (or simply ARCH) model is preferable for underlying series, that is, the conditional variance of error Z_t depends on one previous squared error term. ARFIMA process provided a model with relatively lower values of standard error of the estimate implying towards the precision of estimate. We used these estimated models for making 12-steps ahead forecast and observed the values obtained from ARIMA and GARCH models close to each other and that of actual values.

We have applied these methods to two major components food and non-food inflation and the performance measures are shown in table 5. Due to volatility in series GARCH has been found a healthier model in forecasting food and non-food inflation than other two models.

REFERENCES

[1] J. P. Wilson, *Inflation Deflation Reflation*, Business Books, London, 1980.

[2] A. Nasim, *Determinants of Inflation in Pakistan*, Karachi, State Bank of Pakistan Press, 1995.

[3] A. Hossain, "The Monetarism versus the Neo-Keynesian Views on the Acceleration of Inflation: Some Evidence from South Asian Countries (with Special Emphasis on Pakistan)", *The Pakistan Development Review*, vol. 29, no. 1, pp. 19-32, 1990.

[4] A. H. Khan and M. A. Qasim, "Inflation in Pakistan Revisited", *The Pakistan Development Review*, vol. 35, no. 4, pp. 747–759, 1996.

[5] A. Schimmelpfennig and B. Madhavi, "Three Attempts at Inflation Forecasting in Pakistan", *Inflation Monetary Fund*, WP/05/105, 2005.

[6] G. E. P. Box and G. M. Jenkins, *Time Series Analysis: Forecasting and Control*, 2nd ed., San Francisco: Holden-Day, 1976.

[7] B. Abraham and L. Ledolter, *Statistical Methods for Forecasting*, Wiley, New York, 1983.

[8] P. J. Brockwell and R. A. Davis, *Time Series: Theory and Methods*, Springer-Verlang, 1991.

[9] P. J. Brockwell and R. A. Davis, R. A., *Introduction to Time Series and Forecasting*, 2nd ed., Springer-Verlang, 2002.

[10] S. M. A. Burney, "Computing Exact Likelihood Function of Vector MA(1) Model", *Pakistan Journal of Statistics*, vol. 8, no. 1, pp. 51-59, 1992.

[11] C. W. J. Granger, "Long Memory Relationships and the Aggregation of Dynamic Models", *Journal of Econometrics*, vol. 14, pp. 227-238, 1980.

[12] C. W. J. Granger and R. Joyeux, "An Introduction to Long-Range Time Series Models and Fractional Differencing", *Journal of Time Series Analysis*, vol. 1, pp. 15-30, 1980.

[13] J. R. M. Hosking, "Fractional Differencing", *Biometrika*, vol. 68, pp. 165-176, 1981.

[14] R. T. Baillie, "Long Memory Processes and Fractional Integration in Econometrics", *Journal of Econometrics*, vol. 73, pp. 5-59, 1996.

[15] R. S. Tsay, *Analysis of Financial Time Series*, John Wiley & Sons, Inc., 2002.

[16] A. Maqsood and S. M. A. Burney, "Long Memory Forecasting: An Application to KSE Share Index", *Pakistan Journal of Statistics*, vol. 30, no. 3, pp. 333-344, 2014.

[17] Geweke, J. and Porte, H., The Estimation and Applicat6ion of Long Memory Time Series Models, Journal of Time Series Analysis, 1983, 4, 221-238.

[18] R. Engle, "Autoregressive Conditional Heteroscedasticity With Estimates of the Variance of United Kingdom Inflation", *Econometrica*, vol. 50, no. 1, pp. 987-1007, 1982.

[19] T. Bollerslev, "Generalized Autoregressive Conditional Heteroscedasticity", *Journal of Econometrics*, vol. 31, pp. 307-327, 1986.

Authors Affiliations

[1] Department of Statistics, University of Karachi

[2] Institute of Business Management, Karachi

Corresponding Author: Dr. Arfa Maqood amaqsood@uok.edu.pk