Univariate analysis of forecasting asset prices of Copper

[1]Tanu Shivnani(tanus1march@gmail.com), [2]Jasmine Gambhir (dharinigambhir@gmail.com)

Abstract: This study goes beyond being a mere academic exercise; its findings carry substantial implications for managerial practices and policy considerations. The insights gained from this study offer actionable steps for managers, particularly who are undertaking forecasting exercise of various time series analysis. They can leverage this model designed above for strategic analysis. This modelling exercise will help managers forecast various management and economic variables. It marks as a vital tool for decision making and transforms the way business operations work. The findings of this study pave the way for various industries to forecast various indicators. This modelling exercise is valuable for leaders and managers to forecast various indicators of concern to the organization.

1. Introduction

Copper is a natural element which has been proven to be indispensable to human progress for several centuries. It has wide applications in various industries and is very essential for the growth of many industries. It is the oldest known commodity which impacts the world economy and ranks third in world consumption after steel and aluminum.[1]

India is not one of the copper producing countries and the demand for copper is steadily growing. Since, India is not a major copper producing country, its consumption heavily depends on imports of copper ore.[2]

Most of the studies that I have come across dealing with this issue focus on the demand of copper. Reports by International Copper Study Group (ICSG)[3], International Copper Promotion Council, India (ICPCI) and the Indian Copper Development Centre (ICDC) give the forecasted demand of copper, consumption of copper, rate of increase of demand, production etc for the world and according to regions. These studies however

---

[1] CRN India, http://www.crnindia.com/commodity/copper.html

[2] CRN India, http://www.crnindia.com/commodity/copper.html

[3] For instance, the *ICSG Forecast 2010-11 (October 2010)*; *ICSG Monthly Press Release March 2011*; *World Copper Factbook 2010.* http://www.icsg.org/

do not provide a forecast for prices. Hence, this motivated me to carry out this study for India.

The next section gives literature review. Section 3 describes the data that has been study and the source from where it is obtained. Section 4 discusses the methodology followed by the paper and the model estimated whose results are given in section 5. Following these sections, we have the conclusion.

2. Literature Review

India is not a giant market for copper. There has been dearth of copper mines and the productivity of copper from the mines is extremely low, therefore its production levels are extremely low and it depends on imports to meet its demand. Also, only a handful of companies indulge in the extraction and refining of copper from its alloys and ores. Also, due to dearth of copper mines, India produces copper from imported ore, which is nearly 6 lakh tons of production. Its production accounts for four percent share in the world's total copper production. The countries from which India imports copper are Chile, Indonesia, Australia and Canada.[4]

Indian market is broadly classified into two parts ,namely primary and secondary. Primary segment includes producers which transform copper ore into refined copper. Secondary segment consists of producers which manufacture value added products from copper (for example wires, foil etc.)

The domestic demand of copper is around 5.5 lakh tons in the country. It is interesting to note that a large percentage of total consumption is from tele-communication providers namely MTNL and BSNL. The remaining demand comes from the construction and automobile sector.[5]

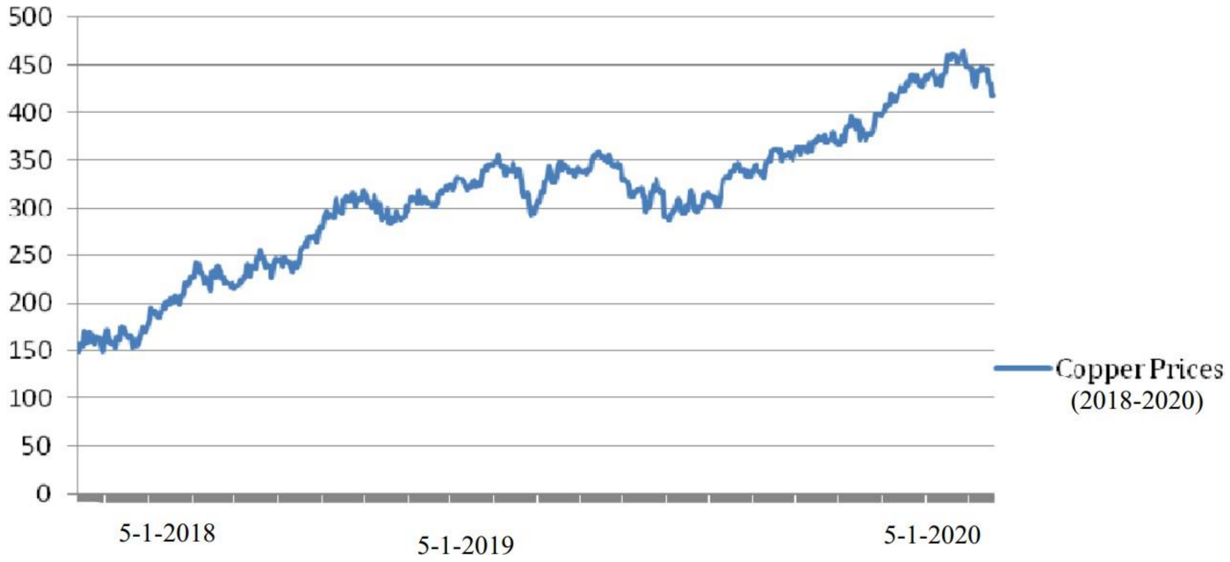
3. Material and Methods

3.1 Description of Data:

Metals are traded in various forms such as wires, sheets, bars, scrap etc and from

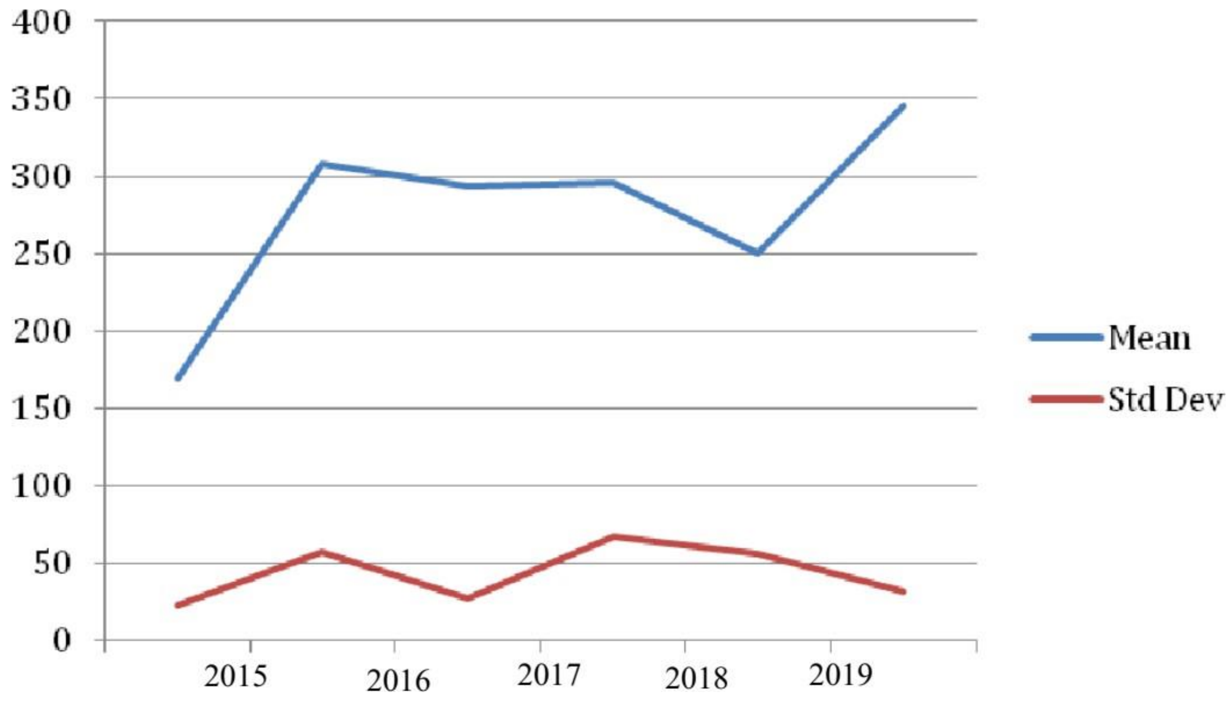
[4] CRN India, http://www.crnindia.com/commodity/copper.html#I_copper_market
[5] CRN India, http://www.crnindia.com/commodity/copper.html#I_copper_market

various locations across the country. The data that has been studied are daily spot market prices for the period January 2018 to March 2020 and has been gathered from Multi Commodity Exchange of India Ltd. (MCX).[6] These prices have been obtained from the trading centre of Mumbai, India. The figure 1 plots this data for each day.



Source: Author
Figure 1

The figure 2 plots the mean and standard deviation of copper prices for the years 2018-2020.



[6] http://www.mcxindia.com/

Source: Author

Figure 2

As it is evident from the data, prices have risen over time. This also reflects the increasing demand and consumption of copper that had been discussed earlier.

3.2 Methodology:

In this paper, we would use the Box-Jenkins Methodology (for univariate time-series analysis). In what follows, we will briefly discuss this methodology step-by-step. We have used STATA10 to obtain the results which are mentioned in the next section. In this section we will only discuss the procedure which was followed.

In the process of forecasting we would be parsimonious, or very few parameters. The larger the number of parameters to estimated, the higher is the probability of getting wrong results. Although, the larger models tend to fit the data well but it performs poorly for out-of-sample forecasts. Therefore, we use smaller values of *p* and *q* in the our analysis to get better forecasts, which are more robust.[7]

The equation 1 gives an *ARMA*(*p*,*q*) process :

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \cdots + \phi_p Y_{t-p} + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \cdots + \theta_q \varepsilon_{t-q} + \varepsilon_t \quad (1)$$

The Box-Jenkins Approach can be broken into four steps:

(1.) Transforming the data such that covariance-stationarity is satisfied.

(2.) Guess estimating lower values for *p* and *q* for an *ARMA* (*p*,*q*) model which describes the transformed series.

[7] Such a problem was also encountered during this analysis and I will mention it later.

(3.) Estimating the value of parameters $\phi(L)$ and $\theta(L)$.

(4.) Performing diagnostic checks to verify if the the model is consistent with the observed features of the data.

Following the Box-Jenkins methodology, we first check the stationarity of the given series. Just by viewing the plot of the series that was presented earlier in the paper, one can say that the series is non-stationary and needs to be transformed to a stationary series. To check this in Stata, the Augmented Dickey-Fuller Test was used which confirmed non-stationarity. To obtain a stationary series, I calculated the percentage change of prices given by the formula (see equation 2)

$$\frac{P_t - P_{t-1}}{P_{t-1}} \times 100. \quad (2)$$

[8]

This relative change is equivalent to a log transformation for daily prices. By running the Augmented Dickey-Fuller Test on the transformed series, it was confirmed to be stationary. The test results are summarised in the table 1.

	Test Statistic	1% Critical Value	5% Critical Value	10% Critical Value
Price Series	-2.715	-3.960	-3.410	-3.120
Percentage Change Series	-29.846	-3.960	-3.410	-3.120

Source: Author

Table 1

The above table confirms that after transformation, we have successfully obtained a stationary series. For the transformed series, another regression was run and it was found that the time trend component was insignificant at 5% level. This is shown in table 2 below.

	Coefficients	Standard. Error	t	P>t	[95% Confidence Interval]
Percentage Change Series (var2)					

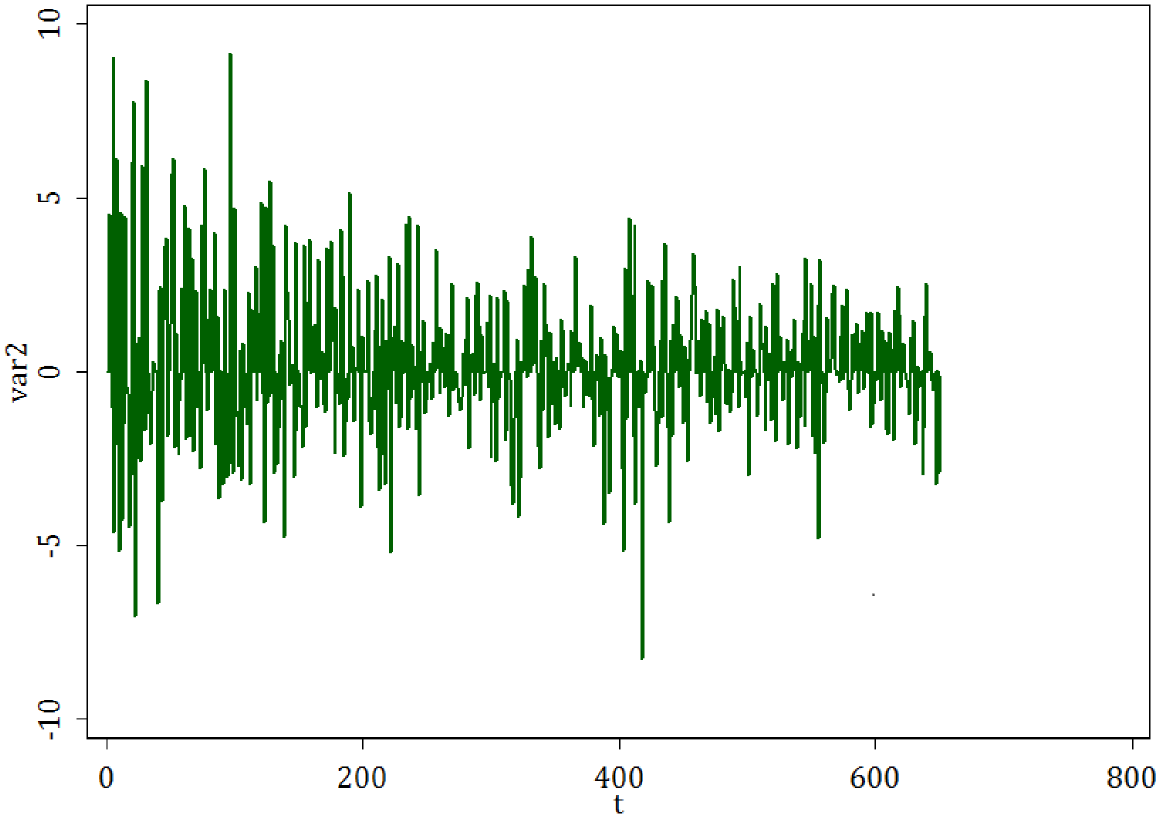
[8] These series was named as var2 during the analysis.

L1.	-1.15852	0.038816	-29.85	0	-1.23474	-1.0823
_trend	-0.00068	0.000418	-1.63	0.104	-0.0015	0.00014
_cons	0.426925	0.157465	2.71	0.007	0.117721	0.736129

Source: Author

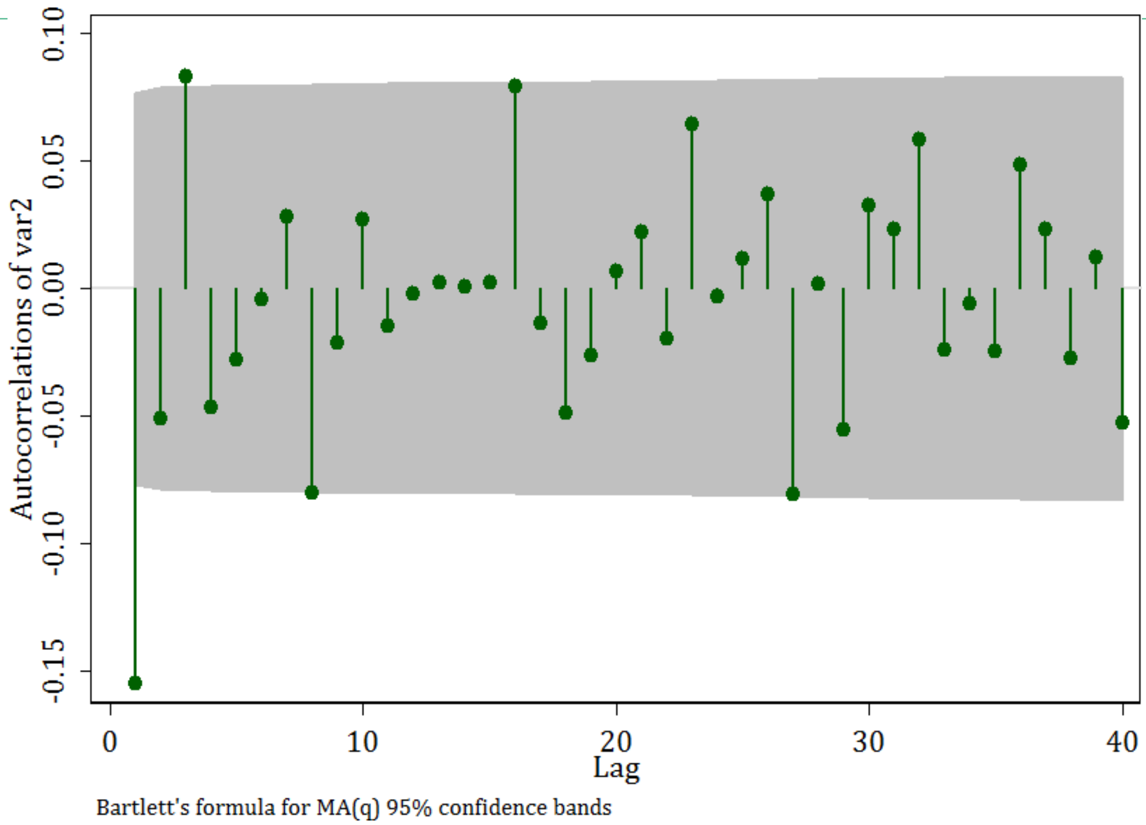
Table 2

This series is plotted below(see figure 3).



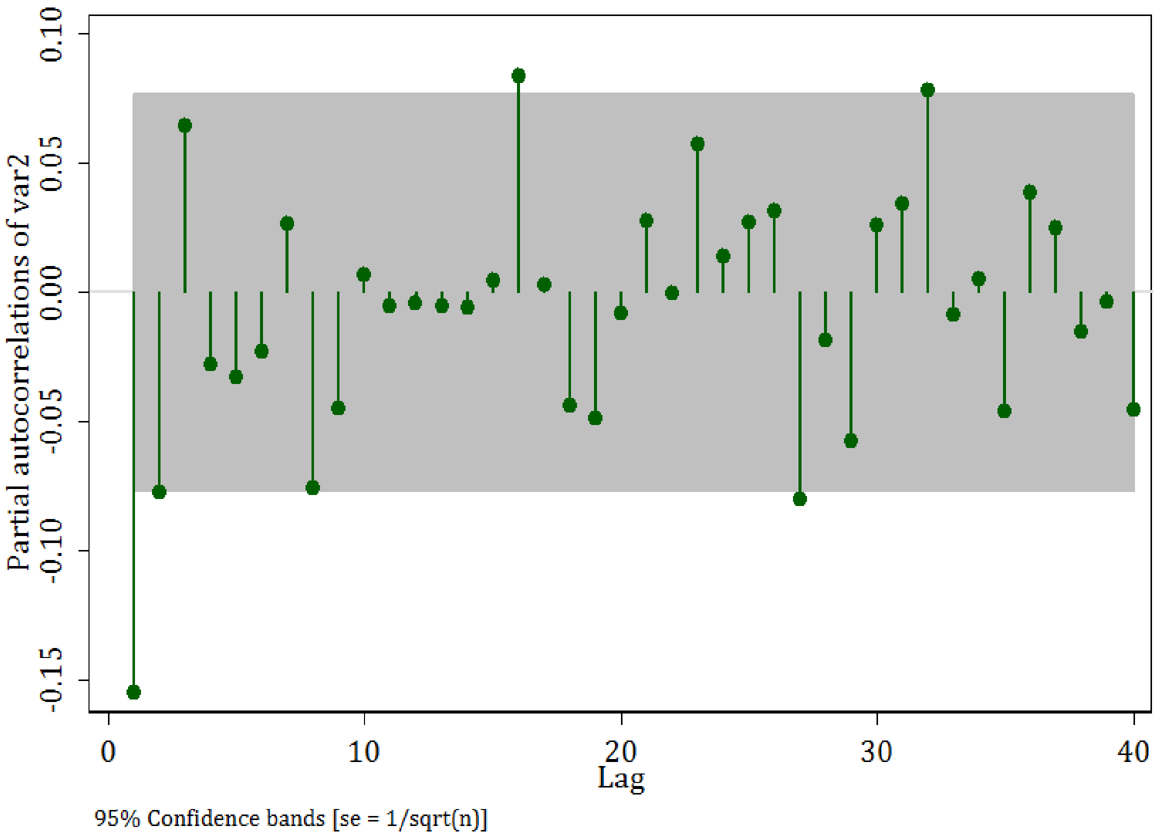
Source: Author
Figure 3

Now we come to the second step. To guess the values of p and q, we first look at the autocorrelations (AC) and partial autocorrelations (PAC) of the transformed series. The following two figures show the AC and PAC of the new series.(see figure 4& figure 5)



Source: Author

Figure 4



Source: Author

Figure 5 Partial Autocorrelation

Notice here that the first autocorrelation and partial autocorrelation are significant. Other possibilities of significance are at lags 2, 3, 8, 27 and 32. Keeping this in mind, I estimated several models out of which significant white noise was generated by the following processes(see Table 3):[9]

Process	Portmanteau Q Statistic	Portmanteau Prob>Chi2(40)	AIC	BIC
ARIMA(8,0,8)	29.3526	0.8925	2746.725	2818.381
ARIMA(16,0,16)*	13.3612	1.0000	2755.443	2898.755
ARIMA(8,0,5)	25.9771	0.9576	2752.406	2819.583
ARIMA(4,0,5)	33.0690	0.7731	2757.186	2806.449
ARIMA(4,0,8)	28.0485	0.9225	2744.966	2803.187
ARIMA(8,0,4)	26.2830	0.9534	2750.432	2813.131

*This model is without a constant, rest are inclusive of a constant term.

Source: Author

Table 3

Initially, when I tried *ARIMA*(16,0,16), I obtained a white noise process with Prob>Chi2(40) equal to 1.[10] I took it to mean that I have obtained a perfect fit, however, later I realised that I am facing the same problem that was discussed – carrying forward with this model would entail a loss of parsimony. Almost all of the coefficients of this model came out to be insignificant and hence I dropped this model.[11]

Estimating models in a similar fashion, amongst all the models estimated, *ARIMA*(4,0,8) turned out to be better fit than others with more coefficients coming to be significant. The results of this estimation will be given in the next section and for some other processes the results will be available in the appendix. The model I want to estimate is (equation3)

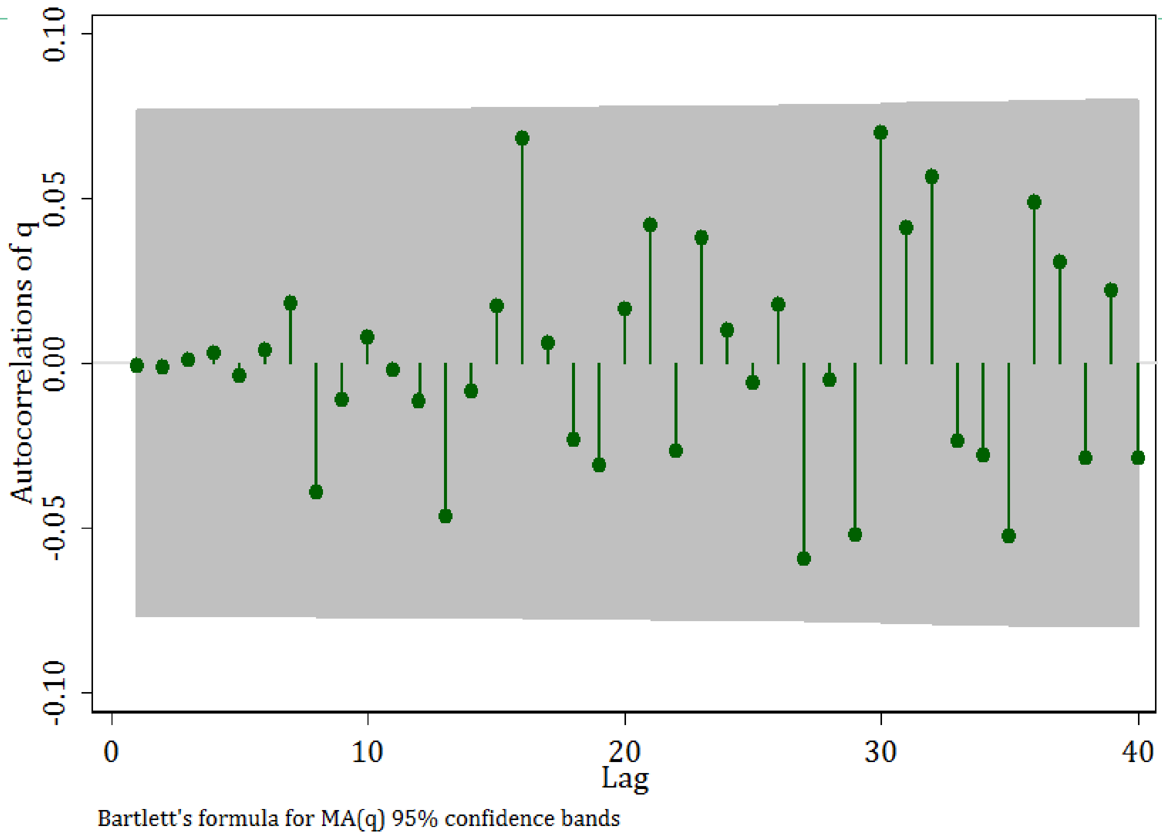
$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \phi_3 Y_{t-3} + \phi_4 Y_{t-4} + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \theta_3 \varepsilon_{t-3} + \theta_4 \varepsilon_{t-4} + \theta_5 \varepsilon_{t-5} + \theta_6 \varepsilon_{t-6} + \theta_7 \varepsilon_{t-7} + \theta_8 \varepsilon_{t-8} + \varepsilon_t \quad (3)$$

Following figures(figure 6&7) give the autocorrelations and the partial autocorrelations for the residuals (named q) of this process.

[9] Other processes estimated during this exercise are summarised in the appendix.

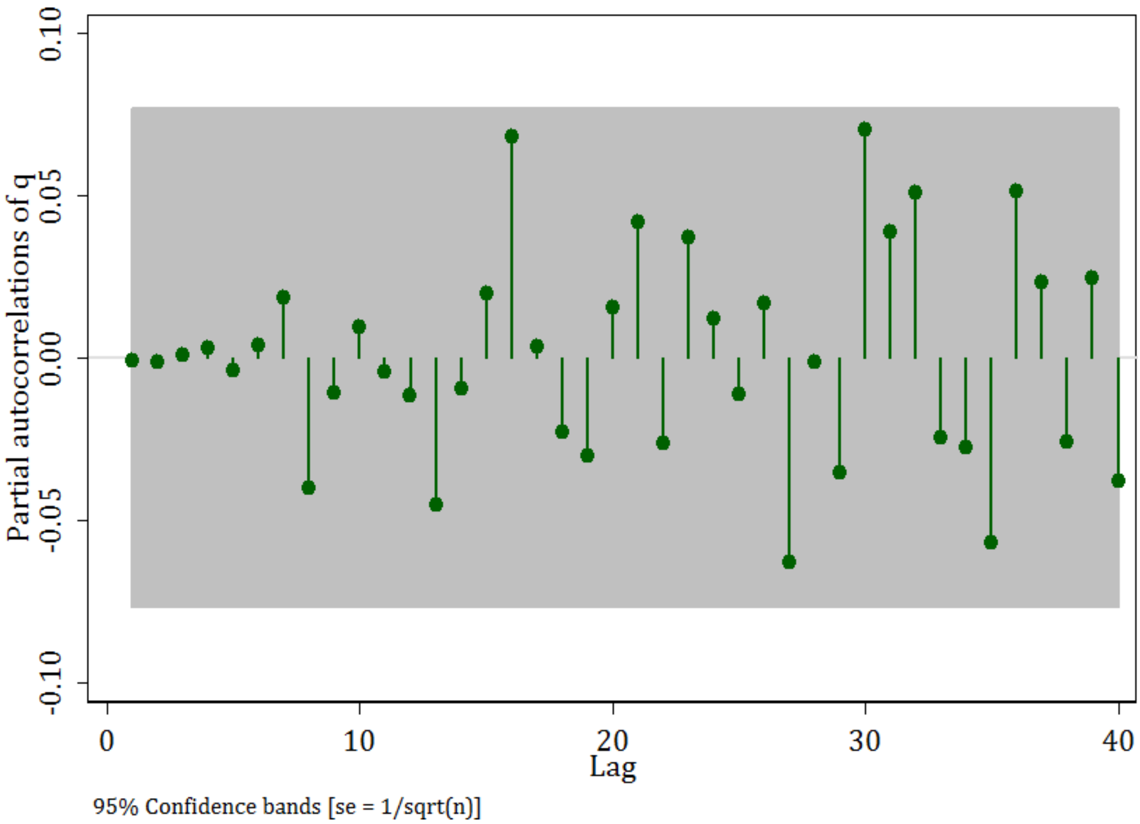
[10] This model was taken without a constant term. Optimization of an ARIMA(16,0,16) model with a constant was not feasible as computed by Stata. In all other models, the constant came out to be significant.

[11] The results of this model are given in the appendix.



Source: Author

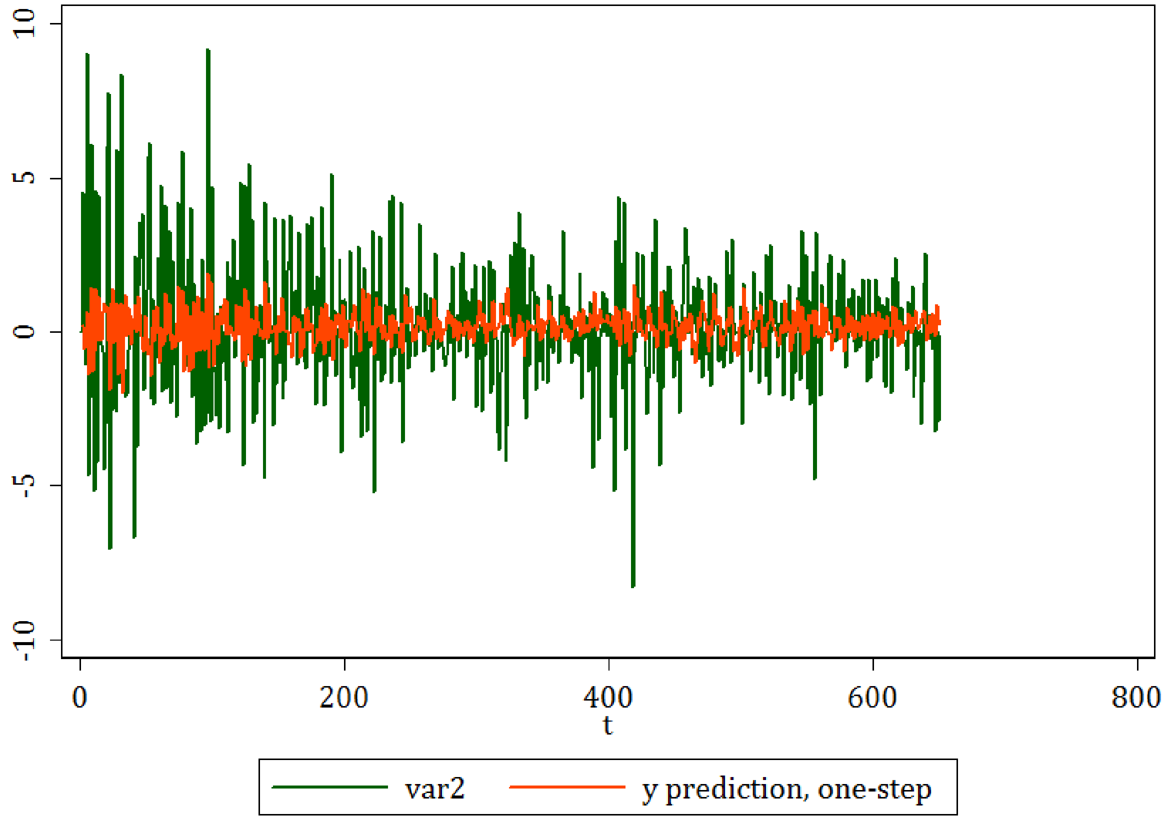
Figure 6



Source: Author

Figure 7

These estimations were done using the Maximum Likelihood Estimation (using the numerical optimization technique) and the root mean square error was calculated for each to see the goodness of fit. For *ARIMA*(4,0,8) the root mean square error was as high as 1.9587348. This process is not a very close fit to the original values as it will be shown in the next figure(see figure 8). Some factors could have been missed out of consideration which can be taken up in future work. For very large values away from the mean, this estimation does not provide very good results.



Source: Author

Figure 8

However, since the above model is the best model according to the BIC criterion and also has the least mean square error, the forecasting was done on the basis of this model only.[12] Ten data points were initially kept aside to be forecasted so that we can compare the observed and the forecasted value and obtain the forecasting error.

[12] ARIMA(8,0,4) produces results very close to ARIMA(4,0,8). The results are given in the appendix.

4. Results and Discussions:

The following results were obtained, after estimating ARIMA(4,0,8)

model(see table 4):

Sample: 1 - 651

Number of obs = 651

Wald chi2(11) = 16248.66

Log likelihood = -1359.483

Prob > chi2 = 0.0000

var2	Coefficients	Standard Error	Z	P>z	[95% Confidence Interval]	
_cons	0.1776	0.0629	2.82	0.005	0.0543	0.3009
ARMA AR						
L1.	0.7929	0.0174	45.53	0	0.7588	0.8271
L2.	-0.582	0.0136	-42.71	0	-0.609	-0.556
L3.	0.8092	0.0131	61.72	0	0.7835	0.8349
L4.	-0.965	0.0174	-55.49	0	-0.999	-0.931
MA						
L1.	-0.965	0.0537	-17.96	0	-1.07	-0.859
L2.	0.6808	0.0494	13.78	0	0.584	0.7776
L3.	-0.815	0.0438	-18.63	0	-0.901	-0.729
L4.	0.986	.	.	.	.	.
L5.	-0.052	0.0596	-0.88	0.379	-0.169	0.0644
L6.	-0.133	0.0668	-1.99	0.046	-0.264	-0.002
L7.	0.1136	0.0522	2.17	0.03	0.0112	0.216
L8.	-0.089	0.039	-2.29	0.022	-0.166	-0.013
/sigma	1.9445	0.0701	27.73	0	1.807	2.0819

Source: Author

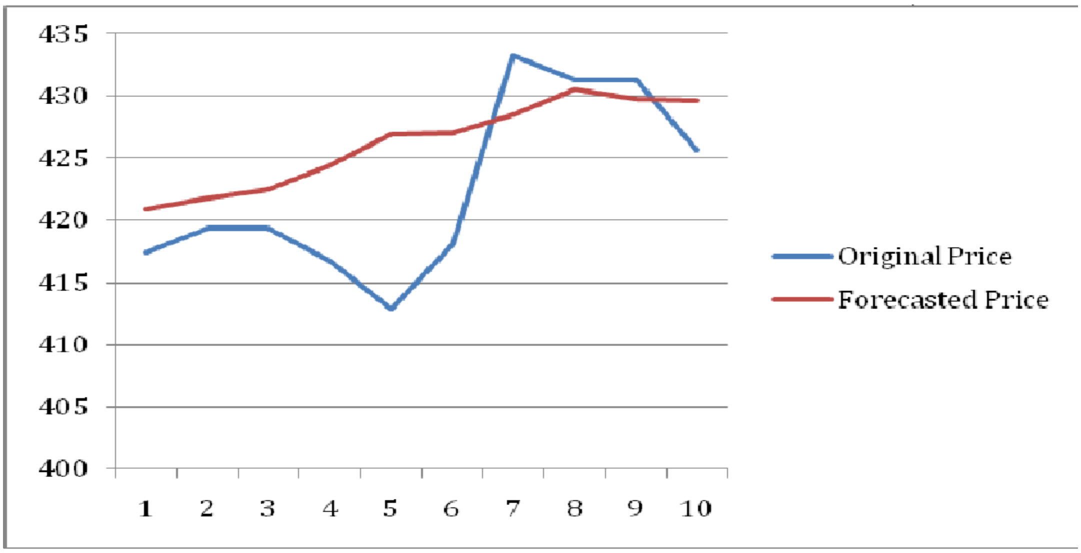
Table 4

In the above model, except for the coefficient for MA Lag 5, rest of the coefficients are significant.[19]

The percentage change series was forecasted ten periods ahead and then inverted to get the forecasted values of prices. These values are given in the table below (table 5) and plotted in the graph that follows(figure 9).

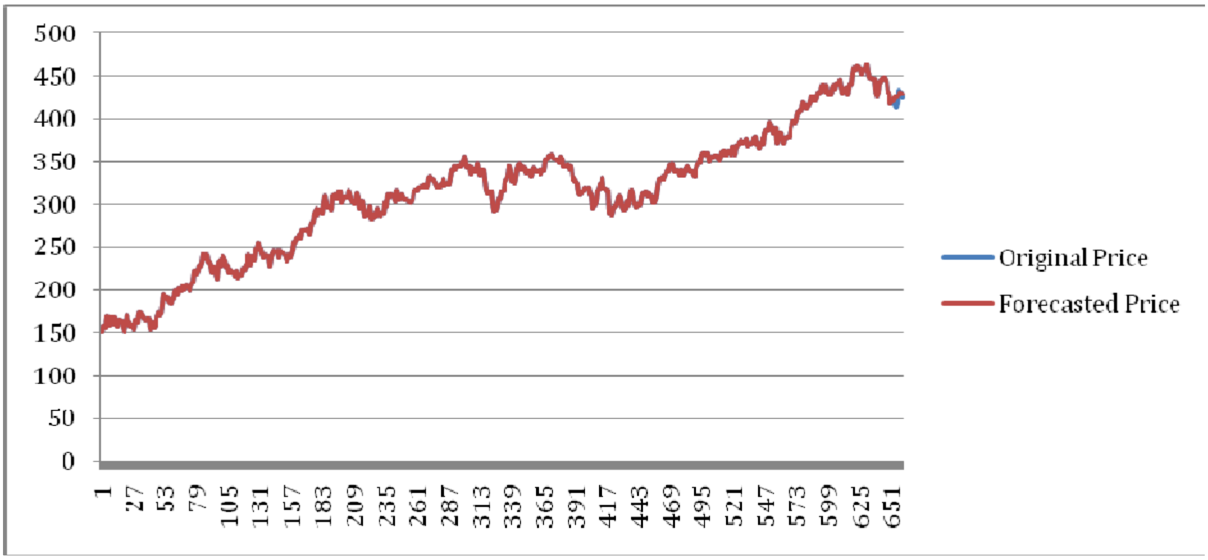
Date	Original Price	Forecasted Price
Mar 11 2020	417.45	420.9539941
Mar 12 2020	419.35	421.8346884
Mar 14 2020	419.35	422.4554607
Mar 15 2020	416.7	424.3729763
Mar 16 2020	412.9	426.9021102
Mar 17 2020	418.1	427.0664987
Mar 18 2020	433.25	428.4786582
Mar 19 2020	431.35	430.5567462
Mar 21 2020	431.35	429.7645046
Mar 22 2020	425.7	429.635384

Source: Author
Table 5



Source: Author
Figure 9

[19] The dot that appears in the results is due to variance becoming zero. The next best model after ARIMA(4,0,8) is ARIMA(8,0,4).



Source: Author

Figure 10

From figure above(figure 10), we can see that original price and forecasted price are nearly the same with scope of small margin of errors.

5. Conclusion:

On the basis of above analysis, it was found that the forecasted values are extremely close to the observed values and tend to follow the trends of the observed prices. From the above table, one can observe that the forecasted price values are more or less rising (or falling) as the original prices are rising (or falling). For the values which are away from the mean, there is a huge difference between the observed values and the forecasted values. It can be found that the volatility had disappeared in the forecast which was originally present in the original series which could be because of heteroscedasticity which we have not taken into account and might require to use ARCH/GARCH models to take care of this problem, which can be taken up in the future work as it would help the organization for undertaking multivariate forecasting analysis.

6. Managerial Implications

This study goes beyond being a mere academic exercise; its findings carry substantial implications for managerial practices and policy considerations. The insights gained from this study offer actionable steps for managers, particularly who are undertaking forecasting exercise of various time series analysis. They can leverage this model designed above for strategic analysis. This modelling exercise will help managers forecast various management and economic variables. It marks as a vital tool for decision making and transforms the way business operations work. The findings of this study pave the way for various industries to forecast various indicators. This modelling exercise is valuable for leaders and managers to forecast various indicators of concern to the organization.

References:

- *Business Line* (2010), 'Copper Consumption in India Set to Grow', March 13, 2010, *Business Line, The Hindu Group.* http://www.thehindubusinessline.in/2010/03/13/stories/2010031351381600.htm
- CRN India http://www.crnindia.com/commodity/copper.html#I_copper_market
- Hamilton, J. D (1994), *Time Series Analysis*, Princeton University Press, New Jersey.
- India Report (2011), 'Copper, Tin Strengthen on Increased Demand', April 09, 2011. http://www.indiareport.com/India-usa-uk-news/latest-news/1026793/NCR/4/20/4
- International Copper Promotion Council, India (ICPCI):
    - Advani, A, 'Shining Prospects of Copper in the SME Sector', ICPCI
    - http://www.copperindia.org/cu/wcms/en/home/about-copper/copper-facts/index.html
    - http://www.copperindia.org/cu/wcms/en/home/about-copper/insights/index.html
- International Copper Study Group (ICSG):
    - *ICSG Forecast 2010-11 (October 2010)*
    - *ICSG Monthly Press Release March 2011*
    - *World Copper Factbook 2010*
    - http://www.icsg.org/
- Metalmarkets (2011), 'Copper Prices Lower On Stockpiles, Manufacturing Data' http://www.metalmarkets.org.uk/2011/04/01/copper-prices-lower-on-stockpiles-manufacturing-data/
- Metalnewsnet (2010), 'India Copper Demand Likely To Rise 8%-10% This Fiscal Year', 22nd June, 2010. http://www.metalnewsnet.com/Metallurgy/lA0khdzDJQy8.html
- Multi Commodity Exchange Ltd. http://www.mcxindia.com/

Appendix:

The other models that were estimated during this exercise and did not yield a white noise process are summarized below(see table 6):

Process	Portmanteau Q Statistic	Portmanteau Prob>Chi2(40)	AIC	BIC
ARIMA(0,0,1)	39.6964	0.4838	2750.648	2764.084
ARIMA(0,0,3)	38.8414	0.5223	2751.116	2773.509
ARIMA(0,0,8)	30.6366	0.8567	2753.835	2798.62
ARIMA(0,0,16)	26.0674	0.9564	2764.137	2844.751
ARIMA(1,0,0)	41.5476	0.4031	2752.511	2765.947
ARIMA(2,0,0)	38.1268	0.5548	2750.593	2768.507
ARIMA(0,0,2)	38.6558	0.5308	2752.007	2769.921
ARIMA(3,0,0)	38.0081	0.5602	2749.876	2772.268
ARIMA(8,0,0)	30.7139	0.8544	2753.977	2798.762
ARIMA(16,0,0)	25.7385	0.9608	2763.753	2844.366
ARIMA(1,0,1)	39.0508	0.5128	2752.301	2770.215
ARIMA(3,0,3)	37.6571	0.5762	2750.618	2781.967
ARIMA(3,0,8)	30.1406	0.8713	2758.807	2817.027
ARIMA(8,0,3)	30.0399	0.8741	2758.637	2816.858
ARIMA(4,0,4)	31.9523	0.8139	2755.852	2800.637
ARIMA(8,0,16)	19.5951	0.9972	2755.18	2871.621
ARIMA(16,0,8)	23.5588	0.9821	2752.109	2864.072
ARIMA(5,0,4)	30.0686	0.8733	2754.155	2803.418

Source: Author

Table 6

ARIMA(16,0,16)

Sample: 1 – 651

Number of obs = 651

Wald chi2(31) = 14673.68

Log likelihood = -1345.721

Prob > chi2 = 0.0000

var2	Coefficients	Standard Error	z	P>z	[95% Confidence Interval]	
ARMA Ar						
L1.	-0.57108	0.874106	-0.65	0.514	-2.28429	1.142141
L2.	0.362075	0.65177	0.56	0.579	-0.91537	1.639522
L3.	-0.17225	0.491217	-0.35	0.726	-1.13501	0.790522
L4.	-0.93561	0.122	-7.67	0	-1.17473	-0.6965
L5.	-0.31158	0.876995	-0.36	0.722	-2.03046	1.407296
L6.	0.275912	0.47036	0.59	0.557	-0.64598	1.1978
L7.	-0.02918	0.434909	-0.07	0.947	-0.88159	0.823225
L8.	-0.31388	0.094024	-3.34	0.001	-0.49817	-0.1296
L9.	0.378399	0.344666	1.1	0.272	-0.29713	1.053932
L10.	0.876566	0.413215	2.12	0.034	0.066678	1.686453
L11.	0.052091	0.834248	0.06	0.95	-1.583	1.687186
L12.	-0.27098	0.190462	-1.42	0.155	-0.64427	0.102323
L13.	0.671538	0.399143	1.68	0.092	-0.11077	1.453844
L14.	0.721332	0.614667	1.17	0.241	-0.48339	1.926057
L15.	0.066718	0.764164	0.09	0.93	-1.43102	1.564452
L16.	0.177362	0.163167	1.09	0.277	-0.14244	0.497163
Ma						
L1.	0.393571	42.28805	0.01	0.993	-82.4895	83.27662
L2.	-0.48996	124.1694	0	0.997	-243.858	242.8775
L3.	0.308634	49.28715	0.01	0.995	-96.2924	96.90967
L4.	0.926783	93.3934	0.01	0.992	-182.121	183.9745
L5.	0.064187	68.48543	0	0.999	-134.165	134.2932
L6.	-0.32985	38.54442	-0.01	0.993	-75.8755	75.21583
L7.	0.191896	.	.	.	.	.
L8.	0.291106	130.1084	0	0.998	-254.717	255.2989
L9.	-0.59958	90.15958	-0.01	0.995	-177.309	176.1099
L10.	-0.84615	224.8797	0	0.997	-441.602	439.91
L11.	0.208567	120.5156	0	0.999	-235.998	236.4148
L12.	0.257685	139.796	0	0.999	-273.737	274.2528
L13.	-0.89446	157.2975	-0.01	0.995	-309.192	307.4029
L14.	-0.58867	84.56501	-0.01	0.994	-166.333	165.1557
L15.	0.156651	31.59164	0	0.996	-61.7618	62.07513
L16.	-0.17203	28.68557	-0.01	0.995	-56.3947	56.05066
/sigma	1.878126	156.3859	0.01	0.99	-304.633	308.3889

Source: Author

Table 7

ARIMA(8,0,4)

Sample: 1 – 651 Number of obs = 651

Wald chi2(12) = 2764.97

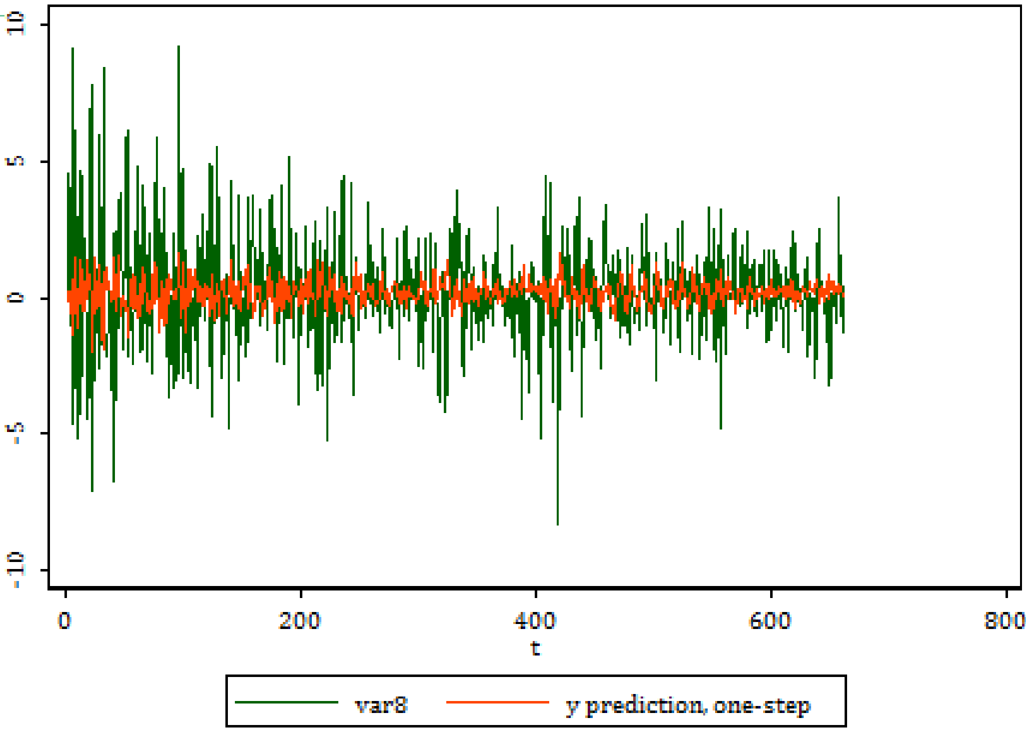
Log likelihood = -1361.216 Prob > chi2 = 0.0000

var2	Coefficients	Standard Error	z	P>z	[95% Confidence Interval]	
_cons	0.177228	0.065765	2.69	0.007	0.04833	0.306125
ARMA Ar						
L1.	0.677236	0.136001	4.98	0	0.410679	0.943792
L2.	-0.44228	0.187167	-2.36	0.018	-0.80912	-0.07544
L3.	0.580161	0.164546	3.53	0	0.257656	0.902665
L4.	-0.81311	0.098245	-8.28	0	-1.00566	-0.62055
L5.	-0.06491	0.05246	-1.24	0.216	-0.16773	0.037906
L6.	-0.08954	0.0566	-1.58	0.114	-0.20048	0.021389
L7.	0.076968	0.047094	1.63	0.102	-0.01533	0.169271
L8.	-0.08991	0.039729	-2.26	0.024	-0.16778	-0.01204
Ma						
L1.	-0.85371	0.129202	-6.61	0	-1.10694	-0.60048
L2.	0.517633	0.194181	2.67	0.008	0.137046	0.89822
L3.	-0.55196	0.185082	-2.98	0.003	-0.91472	-0.18921
L4.	0.802814	0.123674	6.49	0	0.560417	1.04521
/sigma	1.955748	0.041398	47.24	0	1.87461	2.036886

Source: Author

Table 8

The root mean square error of forecast is 1.3219383.



Source: Author
Fig 11

*var8 is the percentage change series with the next ten observed values included.