Asian Journal of Business and Management (ISSN: 2321 - 2802)
Volume 03 – Issue 03, June 2015

Estimation of Monthly Gold Prices using Non-Gaussian Innovations

Ajanthy Mahalingam and T. S. G. Peiris*

Department of Mathematics
University of Moratuwa, Sri Lanka

**Corresponding author's email: sarathp [AT] uom.lk*

ABSTRACT---- ***The global gold market has recently attracted a lot of attention and the price of gold has been fluctuating over the years with high volatility. The frequent fluctuation of gold prices has affected many related sectors and stock market indices and the development of a model to estimate gold prices is of much economic importance to many stake holders. Thus this study was carried out to develop a GARCH model using monthly gold price series (in US$) for the period from December 1978 to December 2012. GARCH models under normal, non-normal innovations: t-distribution and generalized error distribution were selected for the data. After several levels of statistical screening the $AR(1)-PARCH(1,1)$ with t distribution was found to be the best model for the gold price data. The model was trained using December 1978 to December 2010 and validated using January 2011 to December 2012. There was a very strong correlation (r = 0.98, p = 0.000) between observed data and the fitted data for both training and validation data sets.***

Keywords--- GARCH Modeling, Gold Prices, Time Series Analysis, Volatility Clustering.

1. INTRODUCTION

Gold has been used throughout the history as money and has been a relative standard for currency equivalents specific to economic regions or countries, until recent times. Of all the precious metals, gold is the most popular as an investment and its price is mostly volatile in the commodity market (Worthington and Pahlavani, 2007). Investors generally buy gold as a hedge against economic, political, or social fiat currency crises. The gold market is subject to speculation like other markets, especially through the use of futures contracts and derivatives (Wikipedia, 2013). The demand for gold is ever increasing, not only for jewellery, coins, and bars but also for many industries, such as electronics, space, as well as in medical technology.

Since the breakdown of the Bretton Woods agreement in 1971-1973 the official role of gold in the international monetary system has ended. After the de-linking of the US Dollar to gold there has been a wide fluctuation in gold prices. Globally, the gold that is produced is consumed 50% as jewellery, 40% in the form of investments and 10% by industry (Goodman, 1956). With the investment demand for gold being on the rise, and a complex set of factors influencing the investment demand for gold, forecasting the price of gold has been an essential activity and decision making in various sectors, particularly in banking sector. According to Abken (1980) the major determinants for the gold prices are: Supply and demand, US dollar exchange rate, Central banks and other official organizations activity, Oil price, Inflation rate and Political turmoil.

In recent years, demand for gold has been increasing rapidly, due to the world economic recession, high inflation, depreciation of the US dollar, and reduction in world gold production. These may be the reasons causing high volatility on stock exchanges, as investors tend to reconstruct their investment portfolios by replacing part of their stock shares with gold to hedge their risks (Tully and Lucey, 2007).

These reasons have awakened business statisticians and financial economists towards forecasting time series data which are characterized by various forms of volatility (Bollerslev *et al*, 1992). In many scientific and managerial applications, short term predictions for the next few values in a given variable over time can be extremely useful for planning, preparing, or controlling the system under study (George *et al*, 2007). Price forecasting is an integral part of economic decision making. For instance, speculators may use forecasts to try to earn income from investments (Ismail *et al*, 2009).

In view of the above, the objective of this study was to develop a Generalized Autoregressive Conditionally Heteroscedastic (GARCH) model with normal and non-normal distribution innovations for forecasting gold prices.

2. MATERIALS AND METHODS

2.1 Secondary Data

The month end gold prices quoted in US dollars per troy ounce for the period 31st December 1978 to 31st December 2012 were obtained from the statistics division of the World Gold Council website (WGC, 2013). The gold price series was divided in to two parts. The in-sample period for estimating the models, were from 31st December 1978 to 30th December 2010. The out of sample period for validating the fitted models, were from 31st January 2011 to 31st December 2012.

2.2 GARCH Models and Distribution Assumptions

Forecasting ARMA / GARCH process is in one way similar to forecasting ARMA process. However, the difference between ARMA / GARCh and ARMA process is the behavior of the prediction intervals. In series having high volatility, prediction using ARMA / GARCH model will widen to take into account the higher amount of uncertainty. Similarly, the prediction interval will narrow in times of lower volatility. Prediction interval using an ARMA model without conditional heteroskedasticity cannot adapt in this way. Thus ARMA / GARCH models were investigated to gold price series.

With in a class of autoregressive processes with white noises having conditional heteroscedastic variances, it is possible to find reasonable models of $\{y_t\}$. Us if the mean equation is considered as AR(k) process and the variance equation by GARCH(p,q) process then if $\{y_t\}$ is an autoregressive process of order k with a GARCH noise of order p and q, it is represented as:

$$y_t = \delta + \emptyset_1 y_{t-1} + \ldots\ldots + \emptyset_k y_{t-k} + u_t \quad (1)$$

$$u_t = \sigma_t e_t \quad (2)$$

$$\sigma_t^2 = \alpha_0 + \alpha_1 u_{t-1}^2 + \cdots + \alpha_p u_{t-p}^2 + \beta_1 \sigma_{t-1}^2 + \cdots + \beta_q \sigma_{t-q}^2 \quad (3)$$

Where white noise, $e_t \sim iid\ N(0,1)$ and $\{\sigma_t^2\}$ satisfies the recurrence equation (3). This model is normally denoted by AR(k)-GARCH(p,q).

3. RESULTS AND DISCUSSIONS

Developing a forecasting model should always begin with graphical display and analysis of the available data (Montgomery and Jennings, 2011). Thus gold price data from 31st December 1978 to 30th December 2010 were plotted (Fig.1) to gain a visual understanding of the behavior of the series.

Figure 1: The Time Series Plot for the Gold Price Series.

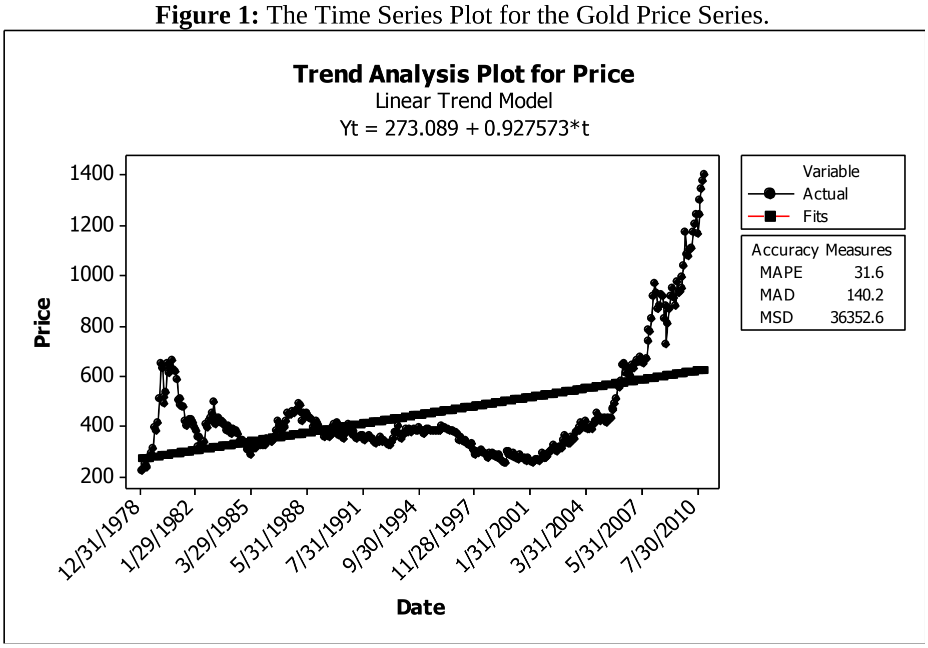
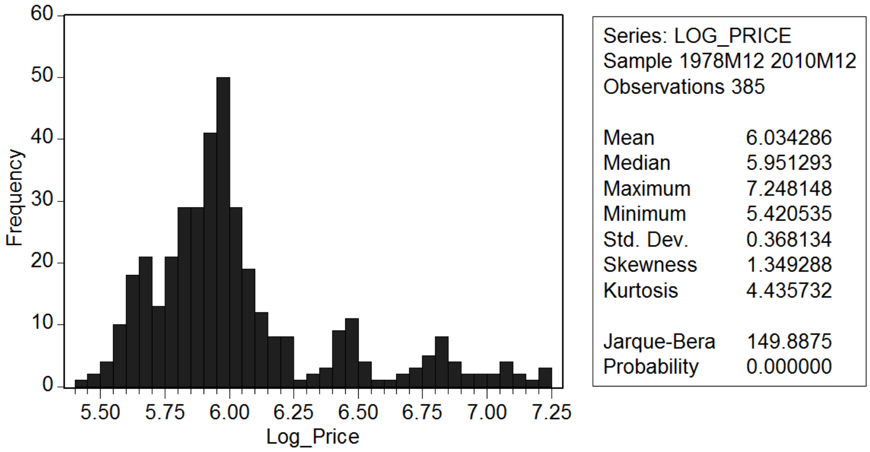


Figure 1 shows, an upward trend and also fluctuations about the trend line. The prices gradually increases since 2007 and there after all prices are above the trend line. Thus it confirm that the varaince of the series increases with time. This complex temporal pattern of monthly gold price series suggests that a trend line is not suitable. In such instances, it is customary to transform the series to log in order to stabilize the variance and remove the effect of the trend. The log transformation is used as a variance stabilizing method as well as to avoid negative values (Aczel and Sounderpandian, 2008).

Figure 2: Histogram and Normality Test for the Log (Gold Price) Series.



The log gold price series displays the properties of volatility clustering and leptokurtosis (Fig 2). Statistics shown in Figure 2 indicates the needs to consider the possibility of developing a GARCH model.

Identification of the Mean Equation

Three types of AR models having AR(1) parameter only, AR(2) parameter only and both AR(1) and AR(2) parameters only were considered for the mean model (Table 1).

Table 1: Comparison of Different Diagnostic Measure for the Identified Three Mean Equations for the Log (Gold Price) Series.

P Value of the AR Parameter	AIC Value	SIC Value	Durbin-Watson Statistic	R^2value
AR(1) p = 0.0000	-3.008188	-2.987612	2.021643	0.9787
AR(2) p = 0.0000	-2.324279	-2.303663	1.055098	0.9576
AR(1) p = 0.0000 AR(2) p = 0.8174	-3.001233	-2.970308	2.000355	0.9786

Results in Table 1 clearly indicates that the model with both AR(1) and AR(2) is not suitable as the AR(2) parameter is not significant (p = 0.814). The AR (1) model was selected on the basis of minimum AIC and SIC values. Further the DW value close to two (2.02) indicates that there is no serial correlation in the residuals of the AR (1) model. The highest R^2 value was also obtained for the AR(1) model. Thus, AR(1) is considered as the best mean equation for the log gold price series. In order to test, the randomness of the error series, correlogram of the residual squared was plotted at different lags and the results are shown in Figure 3.

Figure 3: Correlogram of Residual Squared for the Mean Equation of AR(1).

Correlogram of Residuals Squared

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.152	0.152	8.8870	
		2	0.203	0.185	24.934	0.000
		3	0.345	0.309	71.268	0.000
		4	0.195	0.112	86.139	0.000
		5	0.105	-0.029	90.474	0.000
		6	0.305	0.179	127.05	0.000
		7	0.088	-0.044	130.13	0.000
		8	0.009	-0.119	130.16	0.000
		9	0.115	-0.034	135.41	0.000
		10	0.157	0.119	145.16	0.000
		11	0.008	0.002	145.18	0.000
		12	0.075	-0.021	147.40	0.000
		13	0.111	0.050	152.34	0.000
		14	0.020	0.018	152.50	0.000
		15	0.046	-0.007	153.33	0.000
		16	0.068	-0.047	155.22	0.000
		17	0.040	0.030	155.87	0.000
		18	0.030	0.025	156.22	0.000
		19	0.030	-0.037	156.60	0.000
		20	0.026	-0.013	156.87	0.000
		21	0.045	0.053	157.69	0.000
		22	0.021	0.007	157.86	0.000
		23	0.036	-0.011	158.39	0.000
		24	0.050	0.029	159.40	0.000
		25	0.133	0.159	166.74	0.000
		26	0.128	0.116	173.57	0.000
		27	0.046	-0.068	174.45	0.000
		28	0.080	-0.037	177.12	0.000
		29	0.118	0.054	182.94	0.000
		30	0.034	-0.045	183.43	0.000
		31	0.163	0.060	194.62	0.000
		32	0.163	0.114	205.81	0.000
		33	0.001	-0.025	205.81	0.000
		34	0.035	-0.098	206.34	0.000
		35	0.226	0.124	227.98	0.000
		36	0.024	0.011	228.22	0.000

The correlogram (Figure 3) of the residual squared for the mean equation of AR(1) confirms that the errors of the mean equation are not having constant variance. In other words the low probability value ($p < 0.05$) rejects the homoscedasticity hypothesis of the residuals of the mean equations.

The ARCH test results shown in Table 2 strongly suggest the presence of ARCH effect in the residuals.

Table 2: The ARCH-LM Test for the Mean Equation of AR (1)

ARCH Test:

F-statistic	8.999732	Probability	0.002878
Obs*R-squared	8.838205	Probability	0.002950

Identification of the Variance Equation

In order to identify some possible parsimonious models for variance equation the plot of ACF and PACF of the squared residuals are used (Figure 3). It was found that the autos and partials of residual squares are significant irrespective of lags. Thus it is difficult to identify a suitable model for the variance equation. However, it is not recommended to go for higher order of GARCH models for the variance equation. Thus, it was decided to consider GARCH(1,1), GARCH(0,1), GARCH (1,0), GARCH(2,0) and GARCH(0,3). The comparison of parameters and its significance of the above five models are shown in Table 3.

Table 3: Parameter Estimation of the Parsimonious GARCH Models

Model	Parameter	Estimate	Standard Error	P - Value
AR(1) − GARCH(1,1)	ARCH(1) GARCH(1)	0.090231 0.884103	0.026459 0.032818	0.0006 0.0000
AR(1) − GARCH(0,1)	ARCH(1)	0.458588	0.008828	0.0000
AR(1) − GARCH(1,0)	GARCH(1)	-0.440719	3.534087	0.9008
AR(1) − GARCH(0,2)	ARCH(1) ARCH(2)	0.450368 0.162354	0.083942 0.057346	0.0000 0.0046
AR(1) − GARCH(0,3)	ARCH(1) ARCH(2) ARCH(3)	0.340861 0.094631 0.152771	0.077098 0.054494 0.062145	0.0000 0.0825 0.0140

Based on the significance of the parameters the GARCH(1,1), GARCH(0,1) and GARCH(0,2) models can be considered for the variance equations.

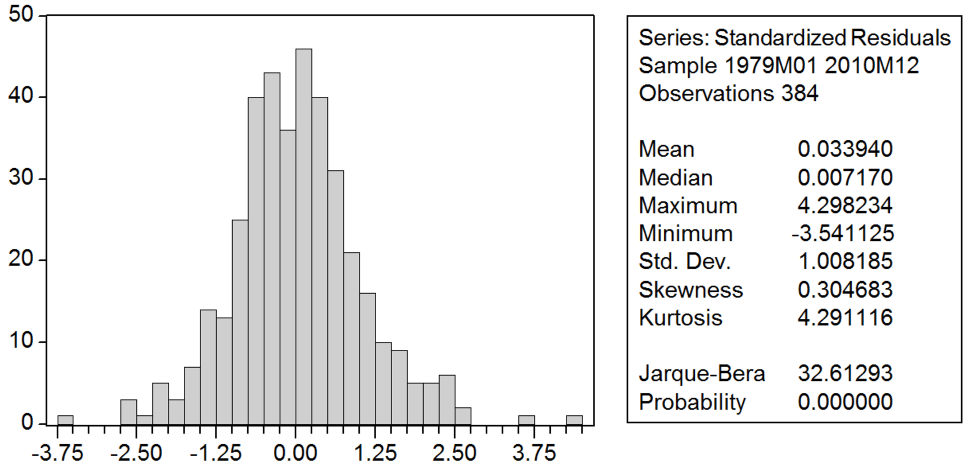
Identification of the Best Fitted GARCH Model.

Table 4: Comparison of Diagnostic Stat for the Selected Three GARCH Models

Model	AIC Value	SIC Value	Log Likelihood
AR(1)-GARCH(1,1)	-3.239062	-3.187621	626.8998
AR(1)-GARCH(0,1)	-3.079733	-3.038580	595.3087
AR(1)-GARCH(0,2)	-3.135490	-3.084049	607.0141

Based on the minimum values of AIC, SIC and log likelihood values (Table 4) it can be concluded that the GARCH(1,1) is most suitable for the variance equation.

Figure 4: Normality Test for the AR(1) −GARCH(1,1) Model.



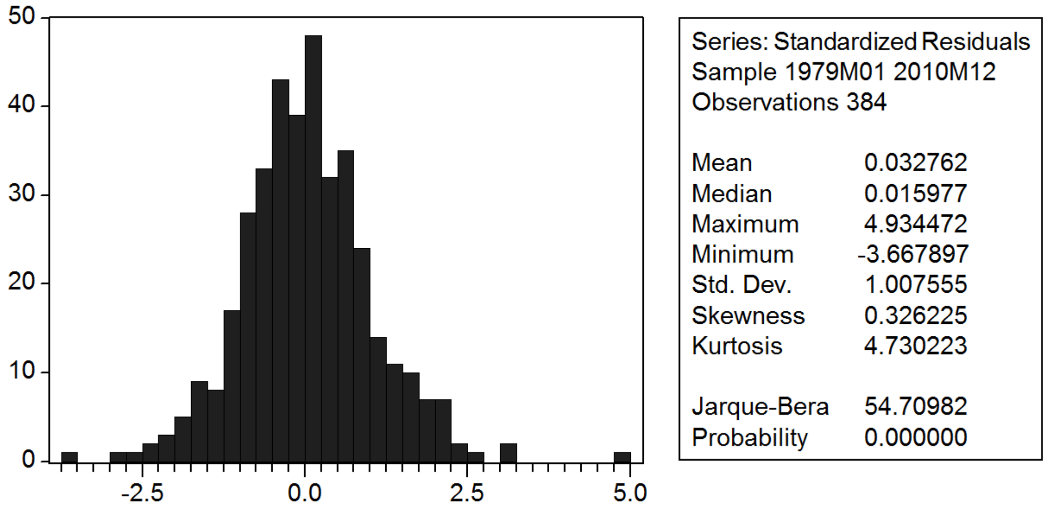
The normality assumptions of errors do not satisfy for the models and it shows thick tails (Figure 4). Thus in addition to normality, the standardized student t distribution and Generalized Error Distribution (GED) were also considered to find the best fitted GARCH model. Variants of GARCH models for both symmetric and asymmetric types were considered with normal and non-normal distribution innovations in modeling log gold price series. The corresponding results are shown in Table 5.

Table 5: GARCH Model Identification of Generalized GARCH Models.

Model	**AIC**	**SIC**	**Log Likelihood**	**JB**
AR(1)-GARCH(1,1) Normal	-3.239062	-3.187621	626.8998	32.612
AR(1)-EGARCH(1,1) Normal	-3.248796	-3.187067	629.7688	28.690
AR(1)-PARCH(1,1) Normal	-3.254058	-3.192329	630.7791	34.584
AR(1) CGARCH(1,1) Normal	-3.245176	-3.173159	630.0737	22.302
AR(1)-GARCH(1,1) Student's t	-3.273821	-3.212092	634.5736	38.138
AR(1)-EGARCH(1,1) Student's t	-3.276672	-3.204655	636.1210	35.057
AR(1)-PARCH(1,1) Student's t	-3.286513	-3.214497	638.0106	54.709
AR(1) CGARCH(1,1) Student's t	-3.240266	-3.157961	630.1311	24.455
AR(1)-GARCH(1,1) Generalized Error	-3.269669	-3.207940	633.7764	38.641
AR(1)-EGARCH(1,1) Generalized Error	-3.276679	-3.204662	636.1223	44.483
AR(1)-PARCH(1,1) Generalized Error	-3.274779	-3.202762	635.7575	34.082
AR(1) CGARCH(1,1) Generalized Error	-3.255564	-3.173259	633.0683	42.010

From Table 5, the AR(1) −PARCH(1,1) model with t-distribution was considered as the best model for the log gold price series based on the minimum AIC and SIC values. The GARCH residuals of this model also presents longest tail (kurtosis = 4.3029) among the other models (Figure 5).

Figure 5: Histogram and Normality Test for Standardized Residuals for the AR(1) −PARCH(1,1) model.



The standardized residuals are leptokurtic and the significant Jarque-Bera test indicates that the standardized residuals are not normally distributed.

Parameter Estimation of $\mathbf{AR(1) - PARCH\ (1,1)}$Model with t-Distribution

Parameter estimation was carried out using EViews. The parameter coefficients on the dependent variable of the log (gold price) series were obtained and the results are tabulated in Table 6.

Table 6: Parameter Estimation of $\mathrm{AR(1) - PARCH\ (1,1)}$ Model with t-Distribution.

Dependent Variable: LOG_PRICE
Method: ML - ARCH (Marquardt) - Student's t distribution
Date: 05/13/13 Time: 03:52
Sample (adjusted): 1979M01 2010M12
Included observations: 384 after adjustments
Convergence not achieved after 500 iterations
Variance backcast: ON
@SQRT(GARCH) = C(3) + C(4)*(ABS(RESID(-1)) - C(5)*RESID(-1)) + C(6)*@SQRT(GARCH(-1))

	Coefficient	Std. Error	z-Statistic	Prob.
C	5.769571	0.220361	26.18234	0.0000
AR(1)	1.011019	0.006260	161.4965	0.0000
	Variance Equation			
C(3)	0.001184	0.000626	1.890836	0.0586
C(4)	0.073426	0.030083	2.440780	0.0147
C(5)	-0.404172	0.286676	-1.409856	0.1586
C(6)	0.917048	0.031129	29.46007	0.0000
T-DIST. DOF	6.790580	2.381940	2.850861	0.0044

R-squared	0.978588	Mean dependent var	6.035884
Adjusted R-squared	0.978247	S.D. dependent var	0.367274
S.E. of regression	0.054169	Akaike info criterion	-3.286513
Sum squared resid	1.106219	Schwarz criterion	-3.214497
Log likelihood	638.0106	F-statistic	2871.631
Durbin-Watson stat	2.030452	Prob(F-statistic)	0.000000

Inverted AR Roots	1.01
	Estimated AR process is nonstationary

From Table 6, for the conditional mean equation, the coefficient for the constant parameter was found as μ = 5.76957 (p = 0.000) and the coefficient for the AR (1) parameter was ϕ_1 = 1.0110 (p = 0.000) while for the conditional variance equation, the coefficients for the parameters were α_0 = 0.00118, α_1 = 0.073426, γ_1 = -0.404172 and β_1 = 0.917048. A high value of β_1 means that volatility is persistent and it takes a long time to change. A high value of α_1 means that volatility is spiky and quick to react to the market movements (Dowd, 2002). The sum of ARCH coefficients of the identified model is close to one (0.001184 + 0.073426 - 0.404172 + 0.917048) indicating that volatility shocks are quite persistent. This further implies that shocks to the conditional variance will be highly persistent indicating that large changes in price tend to be followed by large changes and small changes tend to be followed by small changes.
These results are often observed in high frequency financial data. It can also be seen that both R^2 and adjusted R^2 values are almost equal. The model is able to explain nearly 97% of the observed variability. The $\mathrm{AR(1) - PARCH\ (1,1)}$ model can be written into conditional mean and conditional variance equations as;

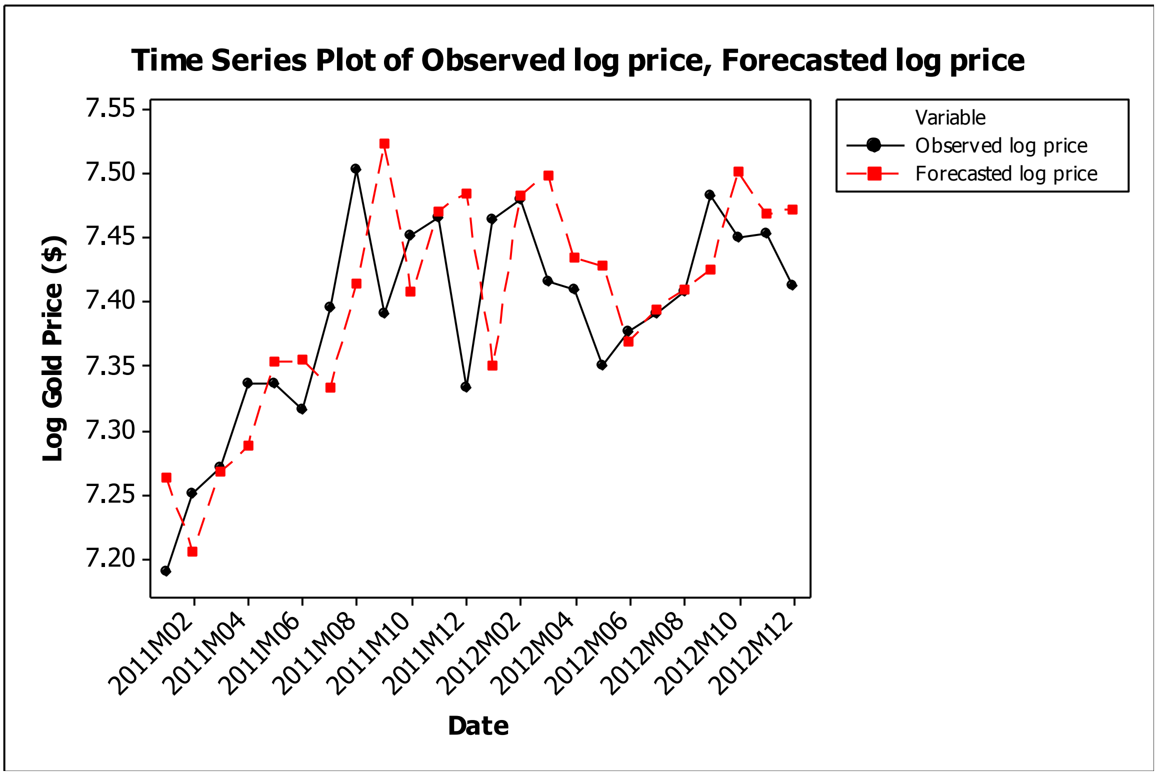
$$log\ y_t = 5.7695 - 1.011\ log\ y_{t-1} + u_t$$

$$\sigma_1 = 0.00118 + 0.07342|u_{t-1}| + 0.40417u_{t-1} + 0.9170\sigma_{t-1}$$

Validation of the Fitted Model for an Independent Data Set (Out of Sample) Forecast

The out of sample forecast was also carried out using EViews for the AR(1) − PARCH(1,1) model. The comparison of forecast values and observed values for the period from January 2011 to December 2012 (out of sample) is shown in Figure 6.

Figure 6: The Plot of Actual Log Gold Price against the Forecasted Log Gold Price using AR(1) − PARCH(1,1) Model (For the "Validation data set").



From the Figure 6, it can be concluded that the trend of forecast prices follows the actual log (gold price) closely. The correlation value between the actual and the forecasted log gold price ($r = 0.9893$, $p = 0.000$) implies the strong positive significant relationship between the observed data and forecasted values.

4. CONCLUSIONS

This study was undertaken to develop a GARCH model for the monthly gold price series from December 1978 to December 2010. Different types of GARCH models for both symmetric and asymmetric types were considered under normality and non normality (Student-t and Generalized Error) GARCH distributions. Due to high volatility and in particularly when fat-tailed asymmetric densities are taken into account, the performance of asymmetric GARCH was found to be better than symmetric GARCH. Of such GARCH families AR(1) − PARCH(1,1) with t- distribution was identified as the 'best fitted' GARCH model.

The estimated parameters of the PARCH (1,1) model, the coefficients of ARCH (α_1) and GARCH (β_1) in the conditional variance equation are statistically significant. Also, as is typical of GARCH model estimates for financial time series data, the sum of the coefficients on the lagged squared error and lagged conditional variance is close to unity. This implies that shocks to the conditional variance will be highly persistent indicating that large changes in price tend to be followed by large changes and small changes tend to be followed by small changes, which means that volatility clustering is observed in the gold price series. The developed GARCH model could help speculators to devise trading strategies which could enable to various stake holders to enjoy better profits.

In this study, only the gold price series was used and no exogenous variables which affect the price of gold were included. However in reality, this is a very important consideration as factors such as global economic growth, inflation, etc affect gold prices. Thus it is suggested to consider the external variables to improve the model.

5. REFERENCES

- Abken, P. A., 1980. The Economics of Gold Price Movements. *Economics Review*, Federal Reserve Bank of Richmond 66 (April) 1980, 3-13.
- Aczel Amir D and Sounderpandian Jayavel, 2008. *Complete Business Statistics*. Sixth Edition.
- Bollerslev, T. P., Chou, R. Y. and Kroner, K. F., 1992. ARCH Modeling in Finance: A Review of the Theory and Empirical Evidence. *Journal of Econometrics* 5 (52-59).
- Dowd, K., 2002. *An Introduction to Market Risk Measurement*. Chichester: John Wiley and Sons.
- George, E.P.Box, Gwilym, M.Jenkins and Gregory, C.Reinsel., 2007. *Time series Analysis Forecasting and Control*, Google Book – Third Edition.
- Goodman, B., 1956. The Price of Gold and International Liquidity. *J. Finance*, 11(1956), 15–28.
- Ismail, Z.A. Yahya and Shabri, A, 2009. Forecasting Gold Prices Using Multiple Linear Regression Method. *American Journal of Applied Sciences.* ISSN Science Publications 1546-9239.
- Montgomery Douglas C and Jennings Cheryl L, 2011. *Introduction to Time Series Analysis and forecasting.* (Online) Available at: www.Onelinelibrary.wiley.com (Accessed on 22 January 2013).
- Tully, E. and Lucey, B.M., 2007. A Power GARCH Examination of the Gold Market *Research in International Business and Finance* 21, 316–325.
- Wikipedia, 2013. Wikipedia, The Free Encyclopedia. (Online) Available at: http://en. Wikepedia.org/ wiki. (Accessed on 20 November 2012).
- World Gold Council, www.gold.org (Accessed on 18 December 2012).
- Worthington, A.C., and Pahlavani, M., 2007. Gold Investment as an Inationary Hedge: Cointegration Evidence with Allowance for Endogenous Structural Breaks. *Applied Financial Economics Letters* 3(4), 259-262.