14th National Convention on Statistics (NCS)
Crowne Plaza Manila Galleria, Quezon City October 1-3, 2019

BAYESIAN ESTIMATION OF A-PARCH MODEL: AN APPLICATION TO JOLLIBEE FOOD CORPORATION STOCK MARKET

by

Shane Marigold L. Oliveros, MSc and Arnulfo P. Supe, PhD

For additional information, please contact:

Author's name	Shane Marigold L. Oliveros, MSc
Designation	Statistician II
Affiliation	Department of Social Welfare and Development – Field Office VII
Address	M.J. Cuenco Ave. cor Gen. Maxilom Ave., Carreta, Cebu City 6000
E-mail	shanemarigold0513@gmail.com
Author's name	Arnulfo P. Supe, PhD
Designation	Retired Professor
Affiliation	MSU – Iligan Institute of Technology
Address	Andres Bonifacio Ave., Tibanga, Iligan City 9200
E-mail	arnulfosupe@gmail.com

BAYESIAN ESTIMATION OF A-PARCH MODEL: AN APPLICATION TO JOLLIBEE FOOD CORPORATION STOCK MARKET

by

Shane Marigold L. Oliveros, MSc[1], and Arnulfo P. Supe, PhD

ABSTRACT

This study derives the posterior densities and computes the estimates of the parameters of the Asymmetric Power Autoregressive Conditional Heteroscedasticity (A-PARCH(p,q)) model using the Bayesian approach. The Markov Chain Monte Carlo (MCMC) method with Metropolis-Hastings (M-H) algorithm is used in estimating the parameters of the model. The procedure is applied to model the returns of Jollibee Food Corporation (JFC) stock market. The best fit model for the JFC stock market returns is ARMA(1,1) – A-PARCH(1,1) model with Student's t-distributed innovations with a Mean Square Error (MSE) value of 1.75107.

Index Terms— Bayesian inference, time series, estimation, heteroscedasticity, A-PARCH model, stock market, returns, volatility, innovations

1. INTRODUCTION

Most investors dislike taking risk and require a premium for holding assets with risky payoffs. The fact that market participants may forecast volatility has important implications. There are periods where the investor has forecasted prices to be very volatile, he/she should either exit the market or require a large premium as compensation for bearing an unusual high risk. Accurate forecasting will help the investors in dealing with risk and expected return, two basic elements in the decision-making process.

Among the most difficult data to model are financial data, because they are replete with volatility clusters. Small inaccuracies in forecasting will lead to large losses, thus it is important that models should be as accurate as possible. This study will attempt to model market and financial trends, and at the same time, will also model financial volatility. This will provide a very useful tool for effective forecasting that will have a large impact on a country's economy.

One of the main assumptions of the standard linear regression analysis and time series models is homoscedasticity which means that the variance of the error terms is constant across all observations. However, this assumption is rarely satisfied in most problems, especially concerning financial time series, since the presence of volatility in financial time series data causes the variance of the error terms to be nonconstant or said to be heteroscedastic.

Engle [1] in 1982 first introduced the Autoregressive Conditional Heteroscedasticity (ARCH) model which was the first model to model stock market's volatility that varies over time. However, one disadvantage of the ARCH model is that it requires a long lag length or a large number of parameters to approximately model the data. This drawback of the ARCH model is addressed by Tim Bollerslev [2] in 1986 when he developed the Generalized Autoregressive Conditional Heteroscedasticity (GARCH) model. GARCH model allows both the squares of the past errors and the past conditional variances in the current conditional variance equation. It has the same properties as the ARCH model but requires less parameters to precisely model heteroscedasticity.

When using these models, there is an imposed restriction on the parameters to ensure that the variance is positive. For this reason, variations of the ARCH model are developed to allow asymmetric effect of positive and negative stock returns such as the Exponential GARCH (EGARCH) model, GJR-GARCH model and Threshold GARCH (TGARCH) model. Another model which takes into account the asymmetric response of volatility to positive and negative shocks also known as the leverage effect of the stock market returns is the Asymmetric Power ARCH (A-PARCH) model, which was introduced by Ding, Granger and Engle [4] in 1993.

The distribution usually chosen for the error terms of these models is the normal distribution. However, according to Engle [1], the normality assumptions of the error terms may not be appropriate in some applications since heavy tails are commonly observed in economic and financial time series data. Some studies show that the Student's t-distributed errors interpret better the features of the series.

Classical parameter estimation of the ARCH-type models can be done using the maximum likelihood estimation (MLE), quasi-maximum likelihood estimation (QMLE) or generalized method of moments (MoM). On the other hand, the Bayesian analysis of this type of models is being studied by the statisticians and econometricians which believe to offer advantages over the classical approach.

This paper focuses on the Bayesian estimation of A-PARCH(*p*,*q*) model. This paper is organized as follows. Section 2 is the summary of the theoretical results of the Bayesian estimation of the A-PARCH(*p*,*q*) model with Student's t-distributed innovations. Section 3 gives the application of the procedure to real data. Finally, section 4 gives the concluding statements for this paper.

2. BAYESIAN ESTIMATION OF A-PARCH MODEL WITH STUDENT'S T-DISTRIBUTED INNOVATIONS

The Asymmetric Power Autoregressive Heteroscedasticity (A-PARCH(*p*,*q*)) model of the error terms y_t with Student's t-distributed innovations ϵ_t can be written using data augmentation as

$$y_t = \epsilon_t(\sigma h_t)^{\frac{1}{2}}$$
$$\epsilon_t \sim N(0,1)$$
$$\omega_t \sim IG\left(\frac{\upsilon}{2}, \frac{\upsilon}{2}\right)$$
$$\sigma = \frac{\upsilon - 2}{\upsilon}$$
$$h_t^{\frac{\delta}{2}} = \alpha_0 + \sum_{i=1}^{p} \alpha_i(|y_{t-i}| - \gamma_i y_{t-i})^{\delta} + \sum_{j=1}^{q} \beta_j h_{t-j}^{\frac{\delta}{2}} \qquad (1)$$

where $\alpha_0 > 0$, $\alpha_i \geq 0$, $\beta_j \geq 0$, $\delta > 0$ and $|\gamma_i| < 1$ to ensure a positive conditional variance, $h_0 = y_0 = 0$ for convenience, ϵ_t is an independent and identically distributed innovation, ω_t is an independent and identically distributed latent variable, σ is the scaling factor, $N(0,1)$ is the standard normal density, $IG\left(\frac{\upsilon}{2}, \frac{\upsilon}{2}\right)$ is the inverse gamma density with parameters $\frac{\upsilon}{2}$ and $\frac{\upsilon}{2}$, υ is the degrees of freedom parameter, and h_t is the conditional variance of y_t given $\{y_{t-1}, y_{t-2}, \ldots\}$ when $\delta = 2$.

Define the vectors $\boldsymbol{y} = (y_1, \ldots, y_t)'$, $\boldsymbol{\alpha} = (\alpha_0, \alpha_1, \ldots, \alpha_p)'$, $\boldsymbol{\beta} = (\beta_1, \beta_2, \ldots, \beta_q)'$, $\boldsymbol{\gamma} = (\gamma_1, \gamma_2, \ldots, \gamma_p)'$, $\boldsymbol{\omega} = (\omega_1, \omega_2, \ldots, \omega_T)'$, $\boldsymbol{\theta} = (\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma}, \upsilon)$, and $\Xi = (\boldsymbol{\theta}, \boldsymbol{\omega})$. Also, define a *T* x *T* diagonal matrix

$$\Sigma = \Sigma(\Xi) = diag(\{\omega_t \sigma h_t\}_{t=1}^{T}) = \begin{bmatrix} \omega_1 \sigma h_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \omega_T \sigma h_T \end{bmatrix}$$

where

$$h_t(\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\gamma}) = \alpha_0 + \sum_{i=1}^{p} \alpha_i(|y_{t-i}| - \gamma_i y_{t-i}) + \sum_{j=1}^{q} \beta_j h_{t-j}(\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\gamma})$$

when $\delta = 2$. Then, the likelihood function of Ξ is

$$L(\Xi|\boldsymbol{y}) \propto |\Sigma|^{-\frac{1}{2}} exp\left\{-\frac{1}{2}\boldsymbol{y}\Sigma^{-1}\boldsymbol{y}\right\}.$$

The following are the proposed priors for $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$, and $\boldsymbol{\gamma}$:

$$p(\boldsymbol{\alpha}) \propto N_{p+1}(\boldsymbol{\alpha}|\boldsymbol{\mu}_{\alpha},\Sigma_{\boldsymbol{\alpha}})I_{(\boldsymbol{\alpha}>0)}$$
$$p(\boldsymbol{\beta}) \propto N_q(\boldsymbol{\beta}|\boldsymbol{\mu}_{\boldsymbol{\beta}},\Sigma_{\boldsymbol{\beta}})I_{(\boldsymbol{\beta}\geq 0)}$$
$$p(\boldsymbol{\gamma}) \propto N_p(\boldsymbol{\gamma}|\boldsymbol{\mu}_{\gamma},\Sigma_{\boldsymbol{\gamma}})I_{(|\boldsymbol{\gamma}|<1)}.$$

Assume that the priors of the parameters $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$, and $\boldsymbol{\gamma}$ are independent implying that $p(\boldsymbol{\theta}) = p(\boldsymbol{\alpha})p(\boldsymbol{\beta})p(\boldsymbol{\gamma})$. Then, the joint posterior density where the construction is based on the Bayes' theorem is given by

$$\pi(\boldsymbol{\theta}|\boldsymbol{y}) \propto L(\boldsymbol{\theta}|\boldsymbol{y})p(\boldsymbol{\theta}).$$

Before deriving the posterior densities of the parameters $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$, and $\boldsymbol{\gamma}$, first transform the conditional variance in (1) by defining $l_t = v_t - h_t$ where $v_t = y_t^2$. So, we obtain

$$v_t = \alpha_0 + \sum_{i=1}^{p} \alpha_i(|y_{t-i}| - \gamma_i y_{t-i})^2 + \sum_{j=1}^{q} \beta_j v_{t-j} - \sum_{j=1}^{q} \beta_j l_{t-j} + l_t.$$

Following Nakatsuma [4], the variable l_t is approximated by a variable $z_t \sim N(0,2h_t^2)$. Then, the auxiliary model is given by

$$z_t = v_t - \alpha_0 - \sum_{i=1}^{p} \alpha_i(|y_{t-i}| - \gamma_i y_{t-i})^2 - \sum_{j=1}^{q} \beta_j v_{t-j} + \sum_{j=1}^{q} \beta_j z_{t-j}$$

when $\delta = 2$ where z_t is a function of $(\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\gamma})$ and $z_t = v_t = 0$ for $t \leq 0$. Now, define a *T* x 1 vector $\boldsymbol{z} = (z_1, \dots, z_T)'$ and a *T* x *T* diagonal matrix

$$\Lambda = diag(\{2h_t^2\}_{t=1}^{T}) = \begin{bmatrix} 2h_1^2 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 2h_T^2 \end{bmatrix}.$$

The approximate likelihood function of $(\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\gamma})$ is

$$L(\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\gamma}|\boldsymbol{y}) = |\Lambda|^{-\frac{1}{2}} exp\left\{-\frac{1}{2}\boldsymbol{z}'\Lambda\boldsymbol{z}\right\}.$$

The construction of the posterior (proposal) densities for $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$, and $\boldsymbol{\gamma}$ will be based on this likelihood function.

Based on the approximate likelihood function of $(\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\gamma})$ and the prior densities of the parameters $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$, and $\boldsymbol{\gamma}$, and using the Bayes' theorem, the following are the posterior (proposal) densities of the parameters $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$, and $\boldsymbol{\gamma}$**:**

For $\boldsymbol{\alpha}$:

$$\pi(\boldsymbol{\alpha}|\boldsymbol{y}) \propto exp\left\{-\frac{1}{2}(\boldsymbol{\alpha} - \widehat{\boldsymbol{\mu}}_{\boldsymbol{\alpha}})'\widehat{\Sigma}_{\boldsymbol{\alpha}}^{-1}(\boldsymbol{\alpha} - \widehat{\boldsymbol{\mu}}_{\boldsymbol{\alpha}})\right\}$$
$$\propto N(\widehat{\boldsymbol{\mu}}_{\boldsymbol{\alpha}},\widehat{\Sigma}_{\boldsymbol{\alpha}})I_{(\boldsymbol{\alpha}>0)}$$

with $\widehat{\Sigma}_{\boldsymbol{\alpha}}^{-1} = \boldsymbol{C}'\widetilde{\Lambda}^{-1}\boldsymbol{C} + \Sigma_{\boldsymbol{\alpha}}^{-1}$ and $\widehat{\boldsymbol{\mu}}_{\boldsymbol{\alpha}} = \widehat{\Sigma}_{\boldsymbol{\alpha}}(\boldsymbol{C}'\widetilde{\Lambda}^{-1}v + \Sigma_{\boldsymbol{\alpha}}^{-1}\boldsymbol{\mu}_{\boldsymbol{\alpha}})$ where $\boldsymbol{C}$ is a *T* x $(p+1)$ matrix whose *t*th row is $\boldsymbol{c}_t$, the *T* x *T* diagonal matrix $\widetilde{\Lambda} = diag(\{2h_t^2(\widetilde{\boldsymbol{\alpha}},\boldsymbol{\beta},\boldsymbol{\gamma})\}_{t=1}^{T})$ and $\widetilde{\boldsymbol{\alpha}}$ is the previous draw of $\boldsymbol{\alpha}$ in the Metropolis-Hastings (M-H) sampler.

For $\boldsymbol{\beta}$:

$$\pi(\boldsymbol{\beta}|\boldsymbol{y}) \propto exp\left\{-\frac{1}{2}\left(\boldsymbol{\beta} - \widehat{\boldsymbol{\mu}}_{\beta}\right)'\widehat{\Sigma}_{\boldsymbol{\beta}}^{-1}\left(\boldsymbol{\beta} - \widehat{\boldsymbol{\mu}}_{\beta}\right)\right\}$$

$$\propto N(\widehat{\boldsymbol{\mu}}_{\boldsymbol{\beta}}, \hat{\Sigma}_{\boldsymbol{\beta}}) I_{(\boldsymbol{\beta} \geq 0)}$$

with $\hat{\Sigma}_{\boldsymbol{\alpha}}^{-1} = \boldsymbol{\Phi}'\widetilde{\Lambda}^{-1}\boldsymbol{\Phi} + \Sigma_{\boldsymbol{\beta}}^{-1}$ and $\widehat{\boldsymbol{\mu}}_{\boldsymbol{\beta}} = \hat{\Sigma}_{\boldsymbol{\beta}}(\boldsymbol{\Phi}'\widetilde{\Lambda}^{-1}s + \Sigma_{\boldsymbol{\beta}}^{-1}\boldsymbol{\mu}_{\boldsymbol{\beta}})$ where $\boldsymbol{\phi} = (\phi_1, \dots, \phi_T)'$, $\boldsymbol{s} = (s_1, \dots, s_T)'$, the *T* x *T* diagonal matrix $\widetilde{\Lambda} = diag\left(\left\{2h_t^2(\boldsymbol{\alpha}, \widetilde{\boldsymbol{\beta}}, \boldsymbol{\gamma})\right\}_{t=1}^{T}\right)$ and $\widetilde{\boldsymbol{\beta}}$ is the previous draw of $\boldsymbol{\beta}$ in the Metropolis-Hastings (M-H) sampler.

For $\boldsymbol{\gamma}$:

$$\pi(\boldsymbol{\gamma}|\boldsymbol{y}) \propto exp\left\{-\frac{1}{2}(\boldsymbol{\gamma} - \widehat{\boldsymbol{\mu}}_{\boldsymbol{\gamma}})'\hat{\Sigma}_{\boldsymbol{\gamma}}^{-1}(\boldsymbol{\gamma} - \widehat{\boldsymbol{\mu}}_{\boldsymbol{\gamma}})\right\}$$
$$\propto N(\widehat{\boldsymbol{\mu}}_{\boldsymbol{\gamma}}, \hat{\Sigma}_{\boldsymbol{\gamma}}) I_{(|\boldsymbol{\gamma}|>0)}$$

with $\hat{\Sigma}_{\boldsymbol{\gamma}}^{-1} = \boldsymbol{C}'\widetilde{\Lambda}^{-1}\boldsymbol{C} + \Sigma_{\boldsymbol{\gamma}}^{-1}$ and $\widehat{\boldsymbol{\mu}}_{\boldsymbol{\gamma}} = \hat{\Sigma}_{\boldsymbol{\gamma}}(\boldsymbol{C}'\widetilde{\Lambda}^{-1}v + \Sigma_{\boldsymbol{\gamma}}^{-1}\boldsymbol{\mu}_{\boldsymbol{\gamma}})$ where $\boldsymbol{C}$ is a *T* x $(p+1)$ matrix whose *t*th row is $\boldsymbol{c}_t$, the *T* x *T* diagonal matrix $\widetilde{\Lambda} = diag(\{2h_t^2(\boldsymbol{\alpha}, \boldsymbol{\beta}, \widetilde{\boldsymbol{\gamma}})\}_{t=1}^{T})$ and $\widetilde{\boldsymbol{\alpha}}$ is the previous draw of $\boldsymbol{\gamma}$ in the Metropolis-Hastings (M-H) sampler.

The derivation of the full conditional posterior density of $\boldsymbol{\omega} = (\omega_1, \dots, \omega_T)'$, an independent and identically random variables from an inverse gamma density with parameters $\frac{\upsilon}{2}$ and $\frac{\upsilon}{2}$ is straightforward. The prior density of $\boldsymbol{\omega}$ is given by

$$f(\boldsymbol{\omega}|\upsilon) \propto \prod_{t=1}^{T} \omega_t^{-\frac{\upsilon}{2}-1} exp\left\{\frac{-\upsilon}{2\omega_t}\right\}$$

and its likelihood function is given by

$$L(\boldsymbol{\omega}|\boldsymbol{y}) \propto \prod_{t=1}^{T} \omega_t^{-\frac{1}{2}} exp\left\{-\frac{1}{2}\left(\frac{y_t^2}{\omega_t \sigma h_t}\right)\right\}.$$

Hence, by Bayes' theorem, the joint posterior density of $\boldsymbol{\omega}$ is given by

$$\pi(\boldsymbol{\omega}|\boldsymbol{y}) \propto \prod_{t=1}^{T} \omega_t^{-\frac{\upsilon+1}{2}-1} exp\left\{\frac{-\tilde{\upsilon}_t}{2\omega_t}\right\}$$

which is the kernel of an inverse gamma density with parameters $\frac{\upsilon+1}{2}$ and $\tilde{\upsilon}_t = \frac{y_t^2}{\sigma h_t} + \upsilon$.

The rejection sampling is used to generate draws for the degrees of freedom parameter υ. Following Deschamps' (2006) choice of the prior density of the degrees of freedom parameter υ, the translated exponential density with parameters $\mu > 0$ and $\lambda \geq 0$ and is given by

$$f(\upsilon) = \mu e^{-\mu(\upsilon - \lambda)}$$

is used. The joint density of the vector $\boldsymbol{\omega} = (\omega_1, \dots, \omega_T)'$ conditional on υ is

$$f(\boldsymbol{\omega}|\upsilon) = \left(\frac{\upsilon}{2}\right)^{\frac{T\upsilon}{2}} \left[\Gamma\left(\frac{\upsilon}{2}\right)\right]^{-T} \prod_{t=1}^{T} \omega_t^{-\frac{\upsilon}{2}-1} exp\left\{\frac{-\upsilon}{2\omega_t}\right\}.$$

Thus, the posterior (proposal) density of υ is

$$\pi(\upsilon|\boldsymbol{\omega}) \propto \left(\frac{\upsilon}{2}\right)^{\frac{T\upsilon}{2}} \left[\Gamma\left(\frac{\upsilon}{2}\right)\right]^{-T} exp\{-\varphi\upsilon\}$$

where $\varphi = \frac{1}{2}\left[\sum_{t=1}^{T}\left(\frac{1}{\omega_t} + \ln \omega_t\right)\right] + \mu$.

3. Application to Real Data

The Bayesian estimation of the A-PARCH(*p*,*q*) model is applied to real data particularly to the stock market closing price index of Jollibee Food Corporation (JFC) from May 03, 2012 to April 23, 2019. JFC also known as Jollibee after its primary fast food brand is a Filipino multinational company based in Pasig, Philippines. The data is from the Investing website (https://www.investing.com/equities/jollibee-foods-historical-data).

This paper uses the daily return of the closing price index as the variable of interest and is calculated as

$$r_t = 100[\ln p_t - \ln p_{t-1}]$$

where r_t is the rate of returns for each stock and p_t is the closing price index for each stock at time t.

TABLE I: Descriptive Statistics of the JFC Return Series

	Statistic	**p-value**
Sample Size	1692	
Minimum	-10.48796	
Maximum	9.40701	
Mean	0.05819	
Standard Deviation	1.81376	
Skewness	0.00972	
Kurtosis	2.84917	
Jarque-Bera Test for Normality	577.75	$< 2.2 \times 10^{-16}$*
Augmented Dickey-Fuller Test for Stationarity	-13.285	0.01*
Box-Pierce Test for Serial Correlation	34.083	5.282×10^{-9}*
Ljung-Box Test for Heteroscedasticity	34.143	5.121×10^{-9}*

- * on the p-value implies rejection of the null hypothesis H_0.
- Augmented Dickey-Fuller Test H_0: The series is not stationary.
- Jarque-Bera Test H_0: The series is normally distributed.
- Box-Pierce Test H_0: The series has no serial correlation.
- Ljung-Box Test H_0: There is no ARCH effect present in the series.

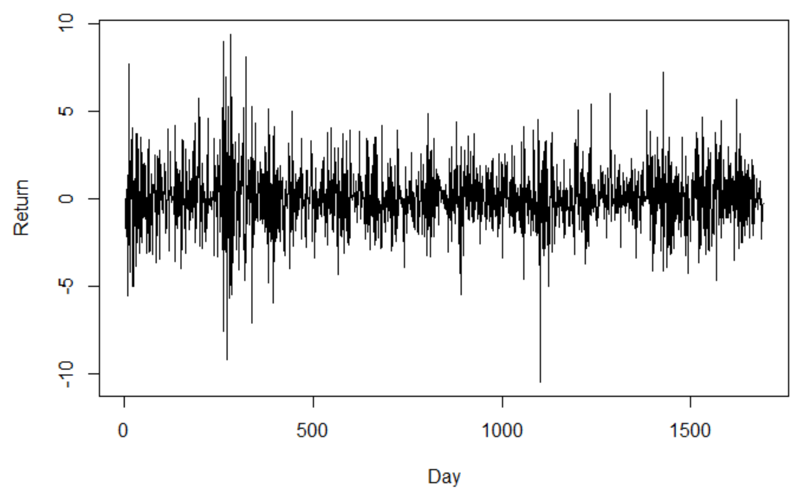

Fig. 1. Historical Plot of the JFC Return Series

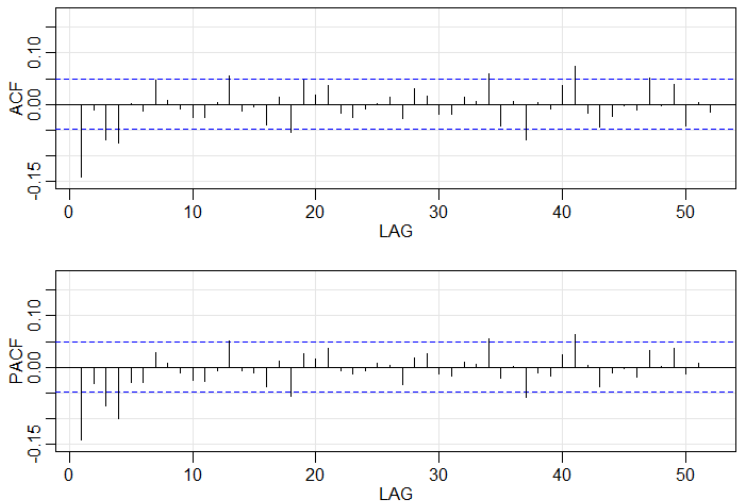

Fig. 2. Correlogram of the ACF and PACF of JFC Return Series

In Fig. 1, the time series line plot of the JFC return series shows that the series has a constant mean over time implying that the series is stationary with respect to the mean. The correlogram of the autocorrelation function (ACF) of the series shown in Fig. 2 shows that the series cuts off at certain lags suggesting that the series is stationary. Based on the result of the Augmented Dickey-Fuller test in Table I, the null hypothesis that the series is not stationary is rejected since its *p*-value (= 0.01) is less than the level of significance $\alpha = 0.05$. Hence, there is a sufficient evidence that the series is stationary at the level of significance $\alpha = 0.05$. Therefore, the JFC return series is stationary. However, it can be observed in Fig. 1 that volatility clustering is present in the series suggesting that there is a need of modelling the heteroscedasticity.

Looking at Table I, the mean (= 0.05819) of the return series is less than its standard deviation (= 1. 81376) indicating a high volatility of the return series. The skewness (= 0.00972) of the return series suggests that the series exhibits a positively skewed distribution (skewed to the right) which means that the right tail of the distribution is longer than the left. Its kurtosis (= 2.84917) implies that the return series exhibits a platykurtic distribution. Compared to a normal distribution, it has a shorter and thinner tail and its central peak is lower and broader.

Moreover, based on the Jarque-Bera test result in Table I, the null hypothesis that the return series is normally distributed is rejected since its *p*-value ($< 2.2 \times 10^{-16}$) is less than the level of significance $\alpha = 0.05$. Hence, there is a sufficient evidence that the series is not normally distributed. Additionally, the Box-Pierce test result shows that the null hypothesis (H_0: The return series has no serial correlation.) is rejected since its *p*-value ($= 5.282 \times 10^{-9}$) is less than the level of significance $\alpha = 0.05$. Thus, there is a sufficient evidence that serial correlation is present in the return series implying serial dependence. Further, based on the Ljung-Box test result, the null hypothesis that there is no ARCH effect present in the return series is rejected since its *p*-value ($= 5.121 \times 10^{-9}$) is less than the level of significance $\alpha = 0.05$. Therefore, the return series has an ARCH effect or its variance is not constant over time (heteroscedastic).

a. Selection of ARIMA(p,d,q) Model

In choosing the best fit mean model, the correlogram of the autocorrelation function (ACF) and partial autocorrelation function (PACF) are used. ACF is used to identify the extent of the lag in a moving average model while PACF is used to identify the extent of the lag in autoregressive model.

Then, the Akaike Information Criterion (AIC) of the candidate models are compared. The lower the value of the AIC, the more accurate the model fits the series.

Looking at Fig. 2, the ACF spikes at lag 1 suggesting that the possible model for the return series may include MA(1). Also, the PACF shows a spike at lag 1 indicating that the possible model may include AR(1). Thus, the candidate models for the returns series are AR(1), MA(1) and ARMA(1,1).

TABLE II: Comparison of the Candidate Mean Models

Model	**AIC**
ARMA(1,0)	6787.09
ARMA(0,1)	6784.51
ARMA(1,1)	**6770.72**

Now, upon comparing the three candidate mean models of the return series given in Table II, ARMA(1,1) has the lowest value of AIC. Therefore, the best fit mean model for the JFC return series is ARMA(1,1).

After choosing the mean model that best fits the return series, diagnostics are performed on the residuals.

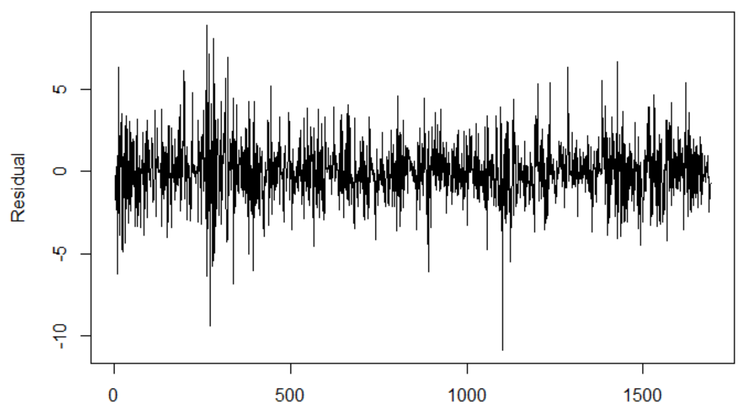

Fig. 3. Residual Plot

TABLE III: Diagnostics in the Mean Model Residuals

	Statistic	**p-value**
White Noise Test	15.883	0.02621*
Ljung-Box Test for Heteroscedasticity	57.335	3.675×10^{-14}*
Jarque-Bera Test for Normality	537.88	$< 2.2 \times 10^{-16}$*

- * on the p-value implies rejection of the null hypothesis H_0.
- Ljung-Box test is performed on the squared residuals of the mean model.
- White Noise Test H_0: The residuals follow a white noise behavior.
- Ljung-Box Test H_0: There is no ARCH effect present in the series.
- Jarque-Bera Test H_0: The series is normally distributed.

It can be observed in Fig. 3 that the residuals of the mean model are stationary with respect to the mean since its mean is constant over time. However, the residuals are not stationary with respect to the variance since the fluctuation is not constant over time. Hence, the residuals of the mean model are not stationary. This claim is supported by the white noise test result. Based on the white noise test result, the null hypothesis that the residuals follow a white noise behavior is rejected since its *p*-value (= 0.02621) is less than the level of significance $\alpha = 0.05$.Thus, there is no sufficient evidence that the residuals follow a white noise behavior implying that the residuals do not follow a white noise behavior.

Furthermore, it is evident in the residual plot given in Fig. 3 that volatility clustering is present on the residuals suggesting that the residuals can be modelled by a heteroscedasticity model. This observation is supported by the Ljung-Box test result. Based on the result, the null hypothesis that there is no ARCH effect present in the squared residuals of the mean model is rejected since its *p*-value (3.675×10^{-14}) is less than the level of significance $\alpha = 0.05$. Therefore, the squared residuals have an ARCH effect or its variance is not constant over time (heteroscedastic) implying that the residuals of the mean model are can be modelled by a heteroscedasticity model. In addition, the Jarque-Bera test result shows that the null hypothesis that the residuals are normally distributed is rejected since its *p*-value ($< 2.2 \times 10^{-16}$) is less than the level of significance $\alpha = 0.05$. Hence, the residuals of the mean model do not follow a normal distribution which suggests that the Student's t distribution must be used as the distribution of the innovations.

b. Selection of A-PARCH(p,q) Model

After performing the diagnostics which are the white noise test, Ljung-Box test and Jarque-Bera test for the residuals of the mean model of the return series, the results suggest that the residuals can be modelled by a heteroscedasticity model.

In choosing the best fit heteroscedasticity (A-PARCH) model, the correlogram of the autocorrelation function (ACF) and partial autocorrelation function (PACF) are used to identify the order (*p*,*q*) of the model. Then, the log-likelihood and Akaike Information Criterion (AIC) of the candidate models are compared. The model that has the highest log-likelihood value and the lowest value of AIC is the heteroscedasticity model that best fits the series.

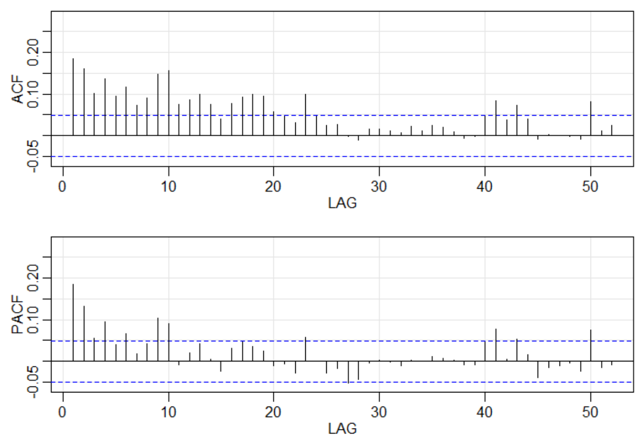

Fig. 4. Correlogram of the ACF and PACF of Squared Residuals

Looking at Fig. 4, the ACF of the squared residuals is highly significant at lag 1 indicating that the order *p* of the possible heteroscedasticity model is 1. Also, the PACF shows a highly significant spike at lag 1 indicating that the order *q* of the possible model is 1. Thus, the A-PARCH model for the residuals is A-PARCH(1,1).

Therefore, the final model for the JFC stock market returns is ARMA(1,1) – A-PARCH(1,1) with Student's t-distributed innovations.

TABLE IV: Bayesian Estimates of the ARMA(1,1) – A-PARCH(1,1) with Student's *t*-distributed Innovations Parameters

Parameter	Estimates
ϕ_1	0.29589
θ_1	-0.46822
α_0	0.18483
α_1	0.18401
γ_1	0.31214
β_1	0.74935
δ	0.79492
ν	5.01306

This model is used to compute the one-step ahead forecast of the JFC stock market returns with a Mean Square Error (MSE) value of 1.75107 shown in Fig. 5.

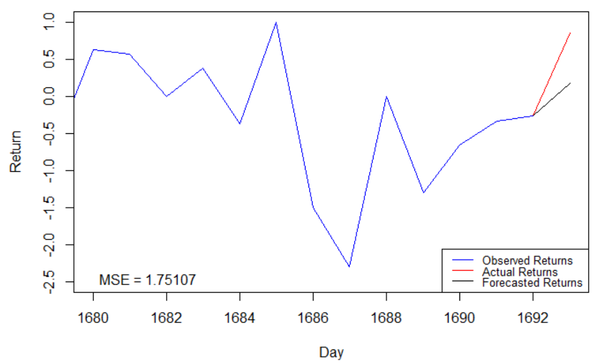

Fig. 5. One-Step Ahead Forecast of JFC Returns

4. CONCLUSION

The posterior densities for the parameters of the A-PARCH(p,q) model with Student's t-distributed innovations are derived using the Bayesian approach. It has been shown that the posterior density for the model parameters $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$, and $\boldsymbol{\gamma}$ is normal distribution given that the prior density for the model parameters is normal distribution. Furthermore, Bayesian estimation of the A-PARCH(p,q) model is applied to model the returns of JFC stock market. It has been found out that the best fit model for the JFC stock market returns is ARMA(1,1) – A-PARCH(1,1) model with Student's t-distributed innovations.

REFERENCES

[1] R. F. Engle, "Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation", *Econometrica*, vol. 50, no.4, pp. 987-1008, 1982.
[2] T. Bollerslev, "Generalized Autoregressive Conditional Heteroskedasticity", *Journal of Econometrics*, vol. 31, pp. 307-327, 1986.
[3] T. Nakatsuma, "A Markov-Chain Sampling Algorithm for GARCH Models", *Studies in Nonlinear Dynamics and Econometrics*, vol. 3, no. 2, pp. 107-117, 1998.
[4] Z. Ding, C. W. J. Granger, and R. F. Engle, "A Long Memory Property of Stock Market Returns and a New Model", *Journal of Empirical Finance*, vol. 1, pp. 83-106, 19