
SKEW NORMAL AND SKEW STUDENT-*T* DISTRIBUTIONS ON GARCH(1,1) MODEL

Didit Budi Nugroho^{1,2}, Agus Priyono¹, Bambang Susanto¹

¹ Department of Mathematics and Data Science, Universitas Kristen Satya Wacana

² Study Center for Multidisciplinary Applied Research and Technology (SeMARTy)

e-mail: didit.budinugroho@uksw.edu ; didit.budinugroho@staff.uksw.edu

DOI: 10.14710/medstat.14.1.21-32

Article Info:

Received: 20 November 2019

Accepted: 16 April 2021

Available Online: 30 June 2021

Keywords:

Skew Distribution; GARCH, Excel's Solver; Volatility

Abstract. The Generalized AutoRegressive Conditional Heteroskedasticity (GARCH) type models have become important tools in financial application since their ability to estimate the volatility of financial time series data. In the empirical financial literature, the presence of skewness and heavy-tails have impacts on how well the GARCH-type models able to capture the financial market volatility sufficiently. This study estimates the volatility of financial asset returns based on the GARCH(1,1) model assuming Skew Normal and Skew Student-*t* distributions for the returns errors. The models are applied to daily returns of FTSE100 and IBEX35 stock indices from January 2000 to December 2017. The model parameters are estimated by using the Generalized Reduced Gradient Non-Linear method in Excel's Solver and also the Adaptive Random Walk Metropolis method implemented in Matlab. The estimation results from fitting the models to real data demonstrate that Excel's Solver is a promising way for estimating the parameters of the GARCH(1,1) models with non-Normal distribution, indicated by the accuracy of the estimation of Excel's Solver. The fitting performance of models is evaluated by using log-likelihood ratio test and it indicates that the GARCH(1,1) model with Skew Student-*t* distribution provides the best fitting, followed by Student-*t*, Skew-Normal, and Normal distributions.

1. INTRODUCTION

In the world of financial markets, volatility is an important indicator needed by market players to anticipate potential losses on financial assets. This is because volatility can be interpreted as a measurement of rising or falling asset prices, thus reflecting investment risk level. Meanwhile, in statistics, volatility is defined as distribution of financial returns (changes in asset prices) and can be measured using standard deviation or variance of returns (Abdalla & Winker, 2012). The higher the volatility, the greater the financial asset risk.

The return value of financial assets, such as stock prices and exchange rates, is generally heteroscedastic, which varies over time, so that its volatility can also be heteroscedastic. A well-known model that can be used to describe heteroscedastic volatility is Generalized Autoregressive Conditional Heteroskedasticity (GARCH) model proposed by

Bollerslev (1986). Many financial studies have shown that financial returns also have skewness characteristics, for example see Iqbal and Triantafyllopoulos (2019) and Cerqueti *et al.* (2020). Since both characteristics of tail thickness and skewness provide a better fit for returns than Normal distribution, it suggests the use of more flexible and superior distributions than Normal and Student- t distributions. Therefore, this study focuses on the application of Skew-Normal distribution (abbreviated as SN) proposed by Azzalini (2011), which generalizes Normal distribution, and Skew Student- t (abbreviated as ST) distribution introduced by Fernandez & Steel (1998), which generalizes Student- t distribution.

A common approach for estimating GARCH models is based on the maximum likelihood estimation (for example, see Tyas *et al.*, 2019 and Sartika *et al.*, 2019). When a model has a complex likelihood function with many parameters, an attractive alternative method for estimating the model is Markov Chain Monte Carlo (MCMC) method. For financial practitioners, this method becomes unattractive since it requires advanced mathematical/statistical knowledge and computer programming. Therefore, a number of tools or tool packages, such as Excel, Matlab, R, and WinBugs, can be used to estimate special cases of GARCH models. For example, Solver in Microsoft Excel was studied by Nugroho *et al.* (2018) for GARCH(1,1) model cases with Normal distribution. Recently, Nugroho, Kurniawati, *et al.* (2019), Nugroho, Susanto, *et al.* (2019), and Kusumawati *et al.* (2020) successfully employed Excel's Solver to estimate GARCH type models, such as GARCH-in-Mean, GJR-GARCH, log-GARCH, and EGARCH, with return errors in the models following Normal and Student- t distributions.

Motivated by the above results, first, this study contributes to the use of Excel's Solver to estimate GARCH models with return errors in the models following SN and ST distributions. In particular, we chose Generalized Reduced Gradient (GRG) Non-Linear method in Excel's Solver since this method has contributed to solving non-linear problems, is a popular linear programming method, is quite efficient, and has faster computation than the other two methods available in Excel's Solver (Maia *et al.*, 2017). To observe the accuracy of the GRG Non-Linear's Excel's Solver method, the estimation results were compared with Adaptive Random Walk Metropolis (ARWM) method in the MCMC scheme implemented in Matlab. This method was successfully employed and demonstrated its statistical efficiency by Nugroho & Susanto (2017) and Nugroho (2018) in estimating APARCH and GARCH models, respectively, where the return errors have Normal and Student- t distributions. Therefore, the first objective of this study is to determine the ability of GRG Non-Linear's Excel's Solver method in estimating GARCH(1,1) models with SN and ST distributions.

Furthermore, analysis of model and estimation method was based on empirical application to real daily data of FTSE100 and IBEX35 stock indices from January 2000 to December 2017. Therefore, the second contribution of this study is to provide empirical evidence on performance comparison between SN and ST distributions against Normal and Student- t distributions in GARCH model contexts. In particular, it aims to obtain the best distribution in GARCH models in fitting two observation data. To the best of our knowledge, there has not been any study related to the two objectives mentioned in this study.

2. LITERATURE REVIEW

GRG Non-Linear method was first developed by Abadie & Carpenter (1969) to expand Reduced Gradient (RG) method so that it accommodates non-linear inequality constraints in optimization problems. The GRG Non-Linear method was proven by Lee *et al.*

(2004) as an effective method for non-linear problems with non-linear constraints. Based on our study, few researches study the use of GRG Non-Linear method to estimate GARCH type models, even though the method is available in Excel's Solver tool which makes it easier for financial practitioners.

Nugroho *et al.* (2018) provided fairly clear steps on how to estimate GARCH models using the GRG Non-Linear' Excel's Solver method. As a simple framework, they focused on GARCH(1,1) models with Normally distributed return errors, and noted a critical issue in selecting initial values, namely that the initial values should be close to the possible solutions. For more complex models, the GRG Non-Linear's Excel's Solver method was used by Nugroho, Kurniawati, *et al.* (2019) to estimate GARCH-in-Mean, GJR-GARCH, and log-GARCH models with Normally distributed return errors, and by Nugroho, Susanto, *et al.* (2019) to estimate non-linear GARCH models with return errors that have Normal and Student-*t* distributions. Both studies found that Excel's Solver produces estimation values for several parameters that violate the model constraints even though it has no effect on other parameter values and objective functions. The violation happened because the estimation values are extremely close to zero or one. Recently, Kusumawati *et al.* (2020) employed the GRG Non-Linear's Excel's Solver method to estimate EGARCH models with return errors following Normal and Student-*t* distributions. In particular, they did not find any violation of the estimation results against the model constraints. In general, the above studies have concluded that GRG Non-Linear's Excel's Solver method has a good ability in estimating the other types which are more complicated than GARCH models since their estimation values are relatively close to the estimation results of ARWM method.

ARWM method was developed by Atchade & Rosenthal (2005) as one of the methods to improve the performance of Random Walk Metropolis (RWM) method, which was first applied by Metropolis *et al.* (1953). This improvement is done by handling changes in parameter proposal values automatically. To the best of our knowledge, only few researches study the use of ARWM method to estimate GARCH-type models, for example see Salim *et al.* (2016), Nugroho & Susanto (2017), Nugroho (2018), Nugroho, Kurniawati, *et al.* (2019), Nugroho, Susanto, *et al.* (2019), and Kusumawati *et al.* (2020). In particular, Nugroho (2018) reported that the ARWM method was statistically efficient (in terms of correlation) to estimate GARCH(1,1) models with Normal and Student-*t* distributions. In all these studies, none focuses on models with return errors that have SN and ST distributions as observed in this study.

SN distribution was developed by Azzalini (2011) to extend Normal distribution so that it can accommodate distribution's skewness. The performance of SN distribution in GARCH models were compared with Normal, Student-*t*, Skew-*t*, Generalized Error Distribution (GED), and Skewed GED distributions by Altun *et al.* (2018), and the study showed that SN distribution is only better than Normal distribution.

In contrast to Altun *et al.* (2018) that applied Skew Student-*t* distribution of Azzalini & Capitanio (2003), this study applied Skew Student-*t* distribution of Fernandez & Steel (1998). Compared to Normal and Student-*t* distributions, Cifter (2012) provided evidence that Skew Student-*t* distribution of Fernandez & Steel (1998) is superior when applied to GARCH-type models.

3. METHODOLOGY

3.1 GARCH(1,1) Model

The GARCH model introduced by Bollerslev (1986) illustrates that today's conditional variance is not only based on the past returns but also on the past variances. One of the popular and commonly used GARCH type models in practice is GARCH(1,1) model. Hansen & Lunde (2005) compared 330 ARCH type models and found no empirical evidence that GARCH(1,1) models outperformed the others. The GARCH(1,1) model of Bollerslev (1986) is expressed as follows

$$R_t = \varepsilon_t, \text{ where } \varepsilon_t \sim N(0, \sigma_t^2) \quad (1)$$

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

where N represents Normal distribution and R_t represents return for asset at time t . The above model has requirements for parameters: $\omega, \alpha, \beta > 0$ for ensuring positivity of conditional variance and $\alpha + \beta < 1$ to ensure the process stationarity. In most financial case studies, return is defined as difference in the natural logarithm of asset values. It is based on the assumption that asset value follows geometric Brownian motion.

3.2 Skew Normal Distribution

According to Azzalini (2011), for random variable X which has Normal Probability Density Function (abbreviated as PDF) $f(x)$ and Normal Cumulative Density Function (abbreviated as CDF) $F(x)$, the SN distribution function for X can be expressed as $g(x|\lambda) = 2f(x)F(\lambda x)$, where $\lambda \in \mathbb{R}$ represents skewness parameter. If $X \sim N(0, \sigma^2)$, then PDF and CDF for Normal distribution are respectively

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{\left(-\frac{x^2}{2\sigma^2}\right)} \quad (3)$$

$$F(x) = \frac{1}{2} \left(1 + \text{Erf} \left(\frac{x}{\sqrt{2\sigma^2}} \right) \right) \quad (4)$$

where Erf represents error function as follows:

$$\text{Erf}(y) = \frac{2}{\sqrt{\pi}} \int_0^y e^{-z^2} dz.$$

Therefore, the PDF for the SN distribution can be expressed as follows:

$$SN(\lambda) = g(x|\lambda) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{\left(-\frac{x^2}{2\sigma^2}\right)} \left(1 + \text{Erf} \left(\frac{\lambda x}{\sqrt{2\sigma^2}} \right) \right) \quad (5)$$

where $\sigma > 0$. The distribution is going to lose its symmetrical property at $\lambda \neq 0$, in which it is going to be left-skewed (negative skewness) if $\lambda < 0$, and right-skewed (positive skewness) if $\lambda > 0$.

3.3 Skew Student- t Distribution

In contrast to SN distribution which only accommodates skewness, there is a distribution that allows skewness as well as kurtosis, namely ST distribution, proposed by Fernandez & Steel (1998). The PDF for random variable X following ST distribution with a mean of 0 and a variance of σ^2 is expressed as follows:

$$ST(\gamma, \nu) = f(x|\gamma, \nu) \quad (6)$$

$$= \frac{2\Gamma\left(\frac{\nu+1}{2}\right)}{(\gamma + \gamma^{-1})\Gamma\left(\frac{\nu}{2}\right)\sqrt{\pi(\nu-2)\sigma^2}} \left(1 + \frac{x^2}{(\nu-2)\sigma^2} (\gamma^2 1_{(x \leq 0)} + \gamma^{-2} 1_{(x > 0)})\right)^{-\frac{\nu+1}{2}}$$

where $\gamma > 0$ represents skewness parameter and $\nu > 2$ represents degrees of freedom that controls the distribution's tail thickness. The distribution is going to lose its symmetrical property when $\gamma \neq 1$. In other words, when $\gamma = 1$, the ST distribution is reduced to Student- t distribution as introduced by Bollerslev (1987). In particular, the distribution is going to be left-skewed if $\gamma \in (0,1)$ and right-skewed if $\gamma > 1$.

3.4 Log-likelihood Function for GARCH Models

For parametric statistical inference objective, the model parameter estimations are based on the likelihood functions formed from probability distributions depending on a fixed set of parameters. For computation problems, it is preferable to use natural logarithm of likelihood (in short, log-likelihood) and that does not change the objective, meaning that maximizing the log-likelihood function also maximizes the likelihood function. This method helps numerically increase the computer's numerical accuracy by calculating the number of probability logarithms rather than multiplying small value probabilities (Solomon, 2015).

$\mathbf{R} = \{R_t\}_{t=1}^T$ represents return series and $\sigma_1^2, \sigma_2^2, \dots, \sigma_T^2$ represents conditional variance series associated with returns. When return error ε_t follows SN distribution, the log-likelihood function for returns with conditional variance following the GARCH(1,1) models, is expressed as follows:

$$L(\mathbf{R}|\theta_1) = -\frac{T}{2}\log(2\pi) - \sum_{t=1}^T \left[\log(\sigma_t) - \frac{R_t^2}{2\sigma_t^2} + \log\left(1 + \text{Erf}\left(\frac{\lambda R_t}{\sqrt{2\sigma_t^2}}\right)\right) \right] \quad (7)$$

where $\theta_1 = (\omega, \alpha, \beta, \lambda)$. Meanwhile, the return series with error ε_t following ST distribution and its conditional variance following the GARCH(1,1) models has a log-likelihood function expressed as follows:

$$L(\mathbf{R}|\theta_2) = T \left[\log(2) + \log(\gamma + \gamma^{-1}) + \log \Gamma\left(\frac{\nu+1}{2}\right) - \log \Gamma\left(\frac{\nu}{2}\right) - \frac{1}{2}\log(\pi(\nu-2)) \right] - \sum_{t=1}^T \left[\log(\sigma_t) - \frac{\nu+1}{2}\log\left(1 + \frac{R_t^2}{(\nu-2)\sigma_t^2} s\right) \right] \quad (8)$$

where $\theta_2 = (\omega, \alpha, \beta, \gamma, \nu)$ and s uses the value of γ^2 if $R_t \leq 0$ and γ^{-2} if $R_t > 0$.

3.5 Estimation Method

There are various methods for estimating the GARCH model parameters. This study focuses on two estimation methods, namely GRG Non-Linear's Excel's Solver and ARWM methods. GRG Non-Linear method was first introduced by Abadie & Carpenter (1969) to solve non-linear optimization problems. The basic idea of this method is to transform constraint inequality into an equality form by adding a non-negative slack variable. The optimization procedure of the GRG Non-Linear method can be seen in more detail in Abadie

& Carpenter (1969) and Lasdon *et al.* (1974) since this study directly employed this method in Excel's Solver to observe its practicality for financial practitioners. Meanwhile, Powell & Batt (2008) explained the basic principles of the GRG Non-Linear method in Excel's Solver as follows. Solving an optimization problem begins with the initial values of the model parameters. These values are going to change little by little so that the values of the objective function become optimal. It means that the values of the objective function are expected to gradually "increase" if the objective is to maximize, and "decrease" if the objective is to minimize.

This study follows the same steps as in Nugroho *et al.* (2018) to estimate the parameters of the considered models using the GRG Non-Linear's Excel's Solver method. The objective of the optimization problem is to maximize the log-likelihood function in Equation (7) and (8) with their constraints following the model parameter requirements. Unfortunately, Excel's Solver does not provide strict conditions for inequality sign so the parameter estimations may not fulfill the requirements. Therefore, this study investigates whether Excel Solver is able to estimate the models using the GRG Non-Linear method without breaking the constraints.

Furthermore, the estimation accuracy of the GRG Non-Linear's Excel's Solver method is evaluated by comparing them with the estimations produced by the ARWM method implemented in MATLAB. Choi & Lam (2017) noted that the two estimation methods give extremely similar results when the difference was relatively close to zero. In this case, the estimation results from the ARWM method were assumed to be the true values because this was proven to be statistically efficient by Nugroho (2018) and Nugroho *et al.* (2017) in empirical applications. The ARWM method was proposed by Atchade & Rosenthal (2005) to improve the performance of RWM method by handling proposal values automatically. This method is employed in the MCMC algorithm which is known to be extremely effective in handling log-likelihood function with many parameters and complex forms. A brief history of the MCMC algorithm can be read in Robert & Casella (2011).

To estimate parameter θ , the ARWM method follows these steps (see Salim *et al.*, 2016):

- (i) Parameter θ_n and step size s_n are given.
- (ii) Generating proposal $\theta^* = \theta_n + \eta_n$, where $\eta_n \sim N(0, s_n)$.
- (iii) Calculating Metropolis ratio:

$$r(\theta_n, \theta^*) = L(\theta^* | \cdot) + p(\theta^*) - L(\theta_n | \cdot) - p(\theta_n),$$
 where $p(x)$ is the log-prior of x .
- (iv) Generating $x \sim U(0,1)$. If $x < \exp\{r(\theta_n, \theta^*)\}$, then θ^* is accepted.
- (v) Updating step size: $s \in [s_{\min}, s_{\max}]$ and then calculating:

$$s^* = \max \left\{ s_{\min}, s_n + \frac{\frac{m(\theta^*)}{n+1} - \bar{\tau}}{(n+1)^{0.6}} \right\},$$

where $m(\theta^*)$ represents frequency of proposal acceptance θ^* . If $s^* > s_{\max}$, then s_{\max} is accepted, otherwise s^* is accepted. This study sets

$$s_{\min} = 10^{-5}, s_{\max} = 10, \bar{\tau} = 0.44,$$

where $\bar{\tau}$ is chosen so that the acceptance rate is closer to 0.44 (Roberts & Rosenthal (2009)).

The ARWM method is employed in the first step of the MCMC algorithm to estimate each model parameter. In summary, for GARCH(1,1) models with SN distribution, the first step of the MCMC scheme has the following stages:

1. (α, β, λ) is known, generating parameter ω using the ARWM method.

2. (ω, β, λ) is known, generating parameter α using the ARWM method.
3. $(\omega, \alpha, \lambda)$ is known, generating parameter β using the ARWM method.
4. (ω, α, β) is known, generating parameter λ using the ARWM method.

Similar stages are also carried out for the parameter estimation cases of GARCH(1,1) models with ST distribution. The above process is started by setting the initial parameter values: $\alpha_0 = 0.2$, $\beta_0 = 0.7$, $\lambda_0 = 1$, $\nu_0 = 1$, and $\gamma_0 = 1$. Meanwhile, for the prior distribution of each parameter, Normal distribution with a mean of 0 and a variance of 10 is used. The first step of the MCMC algorithm is carried out for 6000 iterations, where the first 1000 iterations are discarded as “burn-in” period. Using the remaining 5000 estimations, the second step of the MCMC is to calculate the posterior mean, standard deviation, and 95% HPD (highest posterior density) interval. In this step, the HPD interval is estimated by using Chen–Shao’s approach in Le et al. (2020). The HPD interval is known as the shortest interval among all Bayesian confidence intervals.

4. RESULT AND DISCUSSION

4.1 Data Description

To investigate the ability of GRG Non-Linear’s Excel’s Solver method as well as to evaluate the performances of the studied models, this study uses stock index data of FTSE100 and IBEX35. FTSE100 stock index is a stock index of the top 100 companies listed on the London Stock Market. This market is the fourth largest stock market in the world and the largest one in Europe so it typically reflects the financial situation in Europe. Meanwhile, IBEX35 is an index managed by the Sociedad de Bolsas and consists of the 35 most liquid Spanish securities traded on the Madrid Stock Market. Return data of both indices were downloaded from the Oxford-Man Institute of Quantitative Finance (see web <https://realized.oxford-man.ox.ac.uk>) which is available to public for free. This study observes the daily return data from January 2000 to December 2017. There are 4503 data for FTSE100 and 4532 data for IBEX35. Table 1 presents a summary of the statistical description and normality test for the two indices’ returns.

The statistical description table shows that the average of the two returns is close to 0, so it is appropriate to model returns with an average equal to 0. The skewness values for the two returns are not equal to 0, meaning that the return distributions are not symmetrical. Meanwhile, the kurtosis values for both returns are greater than 3, meaning that the return distribution tails are thicker than Normal distribution tail. Thus, the two statistical values indicate that the returns for both observation data are not Normally distributed. This is confirmed using the Jarque–Bera (JB) normality test which rejects the Normal distribution for both returns, with the JB statistics being greater than the critical value.

Table 1. Statistics for Returns of the FTSE100 and IBEX35 indices

FTSE100		IBEX35	
Mean	-0.0354	Mean	-0.0504
Standard Deviation	0.9296	Standard Deviation	1.2554
Kurtosis	7.04	Kurtosis	8.48
Skewness	-0.1459	Skewness	-0.0359
Minimum	-5.76	Minimum	-7.59
Maximum	7.04	Maximum	13.04
Stat. of JB	3869.6	Stat. of JB	5667.7
Critical Value of JB	5.99	Critical Value of JB	5.98

4.2 Result on FTSE100 and IBEX

The estimation results of GRG Non-Linear's Excel's Solver and ARWM methods in Matlab for GARCH(1,1) models with the return errors following SN and ST distributions, are respectively reported in Tables 2 and 3, along with GARCH(1,1) models with the return errors following Normal and Student-*t* distributions as comparison. First, the estimation results of the two methods are compared. The superscript “*” indicates that the values are significant in terms of the 95% HPD interval. It is observed that each data case shows estimation values (both parameters and log-likelihood) that are relatively close to each other, produced by the two estimation methods when the same model is applied. This indicates the good numerical accuracy of the GRG Non-Linear's Excel's Solver method. Moreover, the results also show that none of the estimation deviates from the model constraints even though Excel's Solver does not provide strict conditions for inequality. This means that the result of $\omega, \alpha, \beta > 0$ or $\alpha + \beta = 1$, as in Nugroho, Kurniawati, *et al.* (2019) and Nugroho, Susanto, *et al.* (2019), is not obtained. Therefore, it can be said that GRG Non-Linear's Excel's Solver method is reliable for estimating GARCH(1,1) models with SN and ST distributions, even though both have complex likelihood functions.

Table 2. The Estimation Results on The GARCH(1,1) Model with Normal (N) and SN Distributions Adopting the FTSE100 and IBEX35 Data

Data	Distribution	ω	α	β	λ	$\alpha + \beta$	L
GRG Non Linear's Solver's Excel							
FTSE100	N	0.0063	0.0965	0.8982	-	0.9947	-5104.62
	SN	0.0062	0.0969	0.8979	-0.064	0.9948	-5098.88
IBEX35	N	0.0129	0.0915	0.9040	-	0.9955	-6776.41
	SN	0.0130	0.0915	0.9039	-0.046	0.9954	-6773.08
ARWM							
FTSE100	N	0.0074	0.1038	0.8898	-	0.9936	-5106.17
	SN	0.0068	0.1008	0.8934	-0.065*	0.9942	-5100.84
IBEX35	N	0.0148	0.0966	0.8979	-	0.9945	-6777.67
	SN	0.0142	0.0942	0.9005	-0.049*	0.9947	-6774.19

Table 3. The Estimation Results on the GARCH(1,1) Model With Student-*T* (T) and ST Distributions Adopting the FTSE100 and IBEX35 Data.

Data	Distribution	ω	α	β	ν	γ	$\alpha + \beta$	L
GRG Non Linear's Solver's Excel								
FTSE100	T	0.0053	0.0914	0.9047	9.42	-	0.9961	-5061.67
	ST	0.0053	0.0916	0.9037	10.42	0.958	0.9953	-5051.40
IBEX35	T	0.0080	0.0701	0.9270	7.53	-	0.9971	-6695.12
	ST	0.0084	0.0702	0.9260	7.81	0.969	0.9962	-6689.69
ARWM								
FTSE100	T	0.0059	0.0938	0.9012	9.91	-	0.9950	-5063.56
	ST	0.0065	0.0993	0.8945	11.12	0.957*	0.9938	-5053.69
IBEX35	T	0.0098	0.0749	0.9207	7.83	-	0.9956	-6696.74
	ST	0.0087	0.0678	0.9277	8.03	0.968*	0.9955	-6691.79

The second concern lies on the estimation results for skewness parameter when SN distribution is applied. This study obtains that the estimate of λ is around -0.064 for the FTSE100 data and -0.046 for the IBEX35 data. These values indicate that the distributions

of the two returns are left-skewed, meaning that most returns of the FTSE100 and IBEX35 data are negative. This confirms the pre-analysis shown in Table 1. However, Excel's Solver does not provide information on whether the estimations are significant or not, meaning whether the values actually deviate from 0 or not. This is one of the weaknesses of Excel's Solver. Furthermore, the significance of parameter λ is obtained from the MCMC results by estimating the 95% HPD interval, namely $\lambda \in [-0.083, -0.012]$ for the adoption of the FTSE100 data, and $\lambda \in [-0.100, -0.027]$ for the adoption of the IBEX35 data. These results conclude that skewness parameter needs to be included in the models.

Third, the application of ST distribution results in a value of skewness parameter between 0 and 1. It means that the return distributions are going to be left-skewed, and it is the same when SN distribution is applied. From the MCMC results, 95% HPD interval also indicates that the skewness values are significant, namely $\gamma \in [0.940, 0.974]$ for the adoption of the FTSE100 data and $\gamma \in [0.950, 0.986]$ for the adoption of the IBEX35 data. Meanwhile, the values of the degrees of freedom are around 11 for the FTSE100 data and 8 for the IBEX35 data. This indicates that the distributions of both returns have thicker tails than Normal distribution, which confirms the pre-analysis in Table 1 based on the kurtosis values. In particular, the results also confirm that the distribution tail of the IBEX35 return is thicker than the FTSE100 return, indicated by their kurtosis values. It means that the larger the kurtosis value, the thicker the distribution tail (which is indicated by smaller degree of freedom). Therefore, these results proved the importance of including skewness parameter as well as kurtosis. Thus, these empirical results provide evidence of support on the use of ST distribution for return errors.

4.3 Model Selection

For model selection that provides the best fit, this study uses log-likelihood ratio test (abbreviated as LRT). The LR test is used to compare the performance of two nested models (one of the models that is a special case of the other one) when in-sample data are applied for the model estimation process. This method is not necessarily the best test, but this test has properties that minimize test errors (Roussas, 1997). Suppose there is a basic model of M_0 and an alternative model of M_1 with log-likelihood values of L_{M_0} and L_{M_1} , respectively, the statistical value for the LRT are expressed as follows (Wu & Vos, 2018):

$$\text{LRT} = 2(L_{M_1} - L_{M_0}) \quad (9)$$

which has an asymptotic distribution with the degree of freedom equals to the difference in the number of parameters between the M_0 and M_1 models. The asymptotic distribution of the test is the χ^2 distribution. In particular, the critical values of the χ^2 distribution with the degree of freedom of 1 at 1%, 5%, and 10% significance levels are 6.64, 3.84, and 2.71, respectively. The alternative model of M_1 significantly provides a better fit than the M_0 model if the LRT is greater than the critical value.

As observed in Table 2, it can be easily calculated that the LRT statistics for the GARCH(1,1) models with SN distribution and the GARCH(1,1) models with Normal distribution are 11.48 (Solver result) and 10.66 (Matlab result), respectively, for fitting the FTSE100 data, as well as 6.66 (Solver result) and 6.96 (Matlab result) for fitting the IBEX35 data. Therefore, for each data case, the LR test rejects GARCH(1,1) models with Normal distribution at any significance level. Thus, GARCH(1,1) models with SN distribution provide a better fit than GARCH(1,1) models with Normal distribution. This confirms the previous result stating that skewness parameter needs to be included in the models.

As observed in Table 3, the GARCH(1,1) models with ST distribution provides a better fit than the GARCH(1,1) models with Student- t distribution for each data case and all significance levels. This is indicated by the LRT statistical values of 20.54 (Solver result) and 19.74 (Matlab result) for the FTSE stocks, as well as 10.86 (Solver result) and 9.9 (Matlab result) for the IBEX35 stocks. These results suggest the use of distributions that accommodate skewness and kurtosis. This also confirms the previous results stating that there is a significance of skewness and kurtosis for both data.

In general, even though the fitting performance of the models with SN and ST distributions cannot be compared using the LRT, the goodness of fit can still be determined. Since the goal is to maximize the log-likelihood, the largest value is better. Thus, based on Table 2 and 3, the best fitting for the FTSE100 and IBEX35 data to the GARCH(1,1) models are provided by the ST distribution specification, followed by the Student- t , SN, and Normal distributions. These results are different from Chu *et al.* (2017) in which GARCH(1,1) models with Normal distribution are better than models with SN or ST distributions when the models are fitted to seven cryptocurrencies.

5. CONCLUSION

This study used the GRG Non-Linear method available in Excel's Solver and the ARWM method in the MCMC scheme implemented in Matlab to estimate GARCH(1,1) models. The models are not only based on the Normal distribution for return errors, but also on the Skew-Normal, Student- t , and Skew Student- t distributions fitted for the daily returns of the FTSE100 and IBEX35. The GRG Non-Linear's Excel's Solver method was found to produce accurate estimations and did not violate the constraints even though Excel's Solver did not provide strict conditions for inequality. The fitting performance of models were evaluated using log-likelihood comparison test criteria. The results provided evidence that the model with Skew Student- t distribution specification is the best among other distributions considered in this study. This is due to the fact that Skew Student- t distribution can accommodate skewness and kurtosis that are typically found in the daily returns of financial assets. Therefore, the assumption that return errors have a distribution which allows both skewness and kurtosis, plays an important role in GARCH(1,1) models.

ACKNOWLEDGMENT

This study has been supported by the UKSW Vice Rector V through a mandatory research grant for the 2019/2020 odd semester (Contract Number: 436/Penel./Rek./9/V/2019).

REFERENCES

- Abadie, J., & Carpenter, J. (1969). Generalization of the Wolfe Reduced Gradient Method to the Case of Nonlinear Constraints. In R. Fletcher (Ed.), *Optimization* (pp. 37–47). Academic Press.
- Abdalla, S. Z. S., & Winker, P. (2012). Modelling Stock Market Volatility Using Univariate GARCH Models: Evidence from Sudan and Egypt. *International Journal of Economics and Finance*, 4(8), 161–176.
- Altun, E., Tatlidil, H., Ozel, G., & Nadarajah, S. (2018). Does the Assumption on Innovation Process Play An Important Role for Filtered Historical Simulation Model? *Journal of Risk Financial Management*, 11(1), 7.
- Atchade, Y. F., & Rosenthal, J. S. (2005). On Adaptive Markov Chain Monte Carlo

- Algorithms. *Bernoulli*, 11(5), 815–828.
- Azzalini, A. (2011). Skew-normal Distribution. In M. Lovric (Ed.), *International encyclopedia of statistical science*. Springer Berlin Heidelberg.
- Azzalini, A., & Capitanio, A. (2003). Distributions Generated by Perturbation of Symmetry with Emphasis on A Multivariate Skew t-Distribution. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 65(2), 367–389.
- Bollerslev, T. (1986). Generalized Autoregressive Conditional Heteroskedasticity. *Journal of Econometrics*, 31(3), 307–327.
- Bollerslev, T. (1987). A Conditionally Heteroskedastic Time Series Model for Speculative Prices and Rates of Return. *The Review of Economics and Statistics*, 69(3), 542–547.
- Cerqueti, R., Giacalone, M., & Mattera, R. (2020). Skewed Non-Gaussian GARCH Models for Cryptocurrencies Volatility Modelling. *Information Sciences*, 527, 1–26.
- Choi, S. W., & Lam, D. M. H. (2017). Comparing Two Methods-Agreeing to Disagree. *Anaesthesia*, 72(5), 651–653.
- Chu, J., Chan, S., Nadarajah, S., & Osterrieder, J. (2017). GARCH Modelling of Cryptocurrencies. *Journal of Risk and Financial Management*, 10(4), 17.
- Cifter, A. (2012). Volatility Forecasting with Asymmetric Normal Mixture GARCH Model: Evidence from South Africa. *Romanian Journal of Economic Forecasting*, 15(2), 127–142.
- Fernandez, C., & Steel, M. (1998). On Bayesian Modeling of Fat Tails and Skewness. *Journal of the American Statistical Association*, 93(441), 359–371.
- Hansen, P. R., & Lunde, A. (2005). A Forecast Comparison of Volatility Models: Does Anything Beat A GARCH(1,1)? *Journal of Applied Econometrics*, 20(7), 873–889.
- Iqbal, F., & Triantafyllopoulos, K. (2019). Bayesian Inference of Multivariate Rotated GARCH Models with Skew Returns. *Communications in Statistics: Simulation and Computation*, 0(0), 1–19.
- Kusumawati, A. M., Nugroho, D. B., & Sasongko, L. R. (2020). Estimasi volatilitas Melalui Model EGARCH(1,1) Berdistribusi Student-t dan Alpha Skew Normal. *Jurnal Ekonomi Kuantitatif Terapan*, 13(2), 259–272.
- Lasdon, L. S., Fox, R. L., & Ratner, M. W. (1974). Nonlinear Optimization Using the Generalized Reduced Gradient Method. *Recherche Opérationnelle*, 8(3), 73–103.
- Le, H., Pham, U., Nguyen, P., & Pham, T. B. (2020). Improvement on Monte Carlo Estimation of HPD Intervals. *Communications in Statistics - Simulation and Computation*, 49(9), 2164–2180.
- Lee, H., Chen, S., & Kang, H. (2004). A Study of Generalized Reduced Gradient Method with Different Search Directions. *Measurement Management Journal*, 1(1), 25–38.
- Maia, A., Ferreira, E., Oliveira, M. C., Menezes, L. F., & Andrade-Campos, A. (2017). Numerical Optimization Strategies for Springback Compensation in Sheet Metal Forming. In *Computational Methods and Production Engineering* (pp. 51–82).
- Metropolis, N., Rosenbluth, A. W., Rosenbluth, N. M., Teller, A. H., & Teller, E. (1953). Equations of State Calculations by Fast Computing Machine. *The Journal of Chemical Physics*, 21, 1087–1091.

- Nugroho, D. B. (2018). Comparative Analysis of Three MCMC Methods for Estimating GARCH Models. *IOP Conference Series: Materials Science and Engineering*, 403(1), 012061.
- Nugroho, D. B., Kurniawati, D., Panjaitan, L. P., Kholil, Z., Susanto, B., & Sasongko, L. R. (2019). Empirical Performance of GARCH, GARCH-M, GJR-GARCH and log-GARCH Models for Returns Volatility. *Journal of Physics: Conference Series*, 1307(1), 012003.
- Nugroho, D. B., & Susanto, B. (2017). Volatility Modeling for IDR Exchange Rate Through APARCH Model with Student-t Distribution. *AIP Conference Proceedings*, 1868(1), 040005.
- Nugroho, D. B., Susanto, B., Prasetya, K. N. P., & Rorimpandey, R. (2019). Modeling of Returns Volatility Using GARCH(1,1) Model under Tukey Transformations. *Jurnal Akuntansi dan Keuangan*, 21(1), 12–20.
- Nugroho, D. B., Susanto, B., & Pratama, S. R. (2017). Estimation of Exchange Rate Volatility Using APARCH-type Models: A Case Study of Indonesia (2010–2015). *Jurnal Ekonomi dan Ekonomi Studi Pembangunan*, 9(1), 65–75.
- Nugroho, D. B., Susanto, B., & Rosely, M. M. M. (2018). Penggunaan MS Excel untuk Estimasi Model GARCH(1,1). *Jurnal Matematika Integratif*, 14(2), 71–83.
- Powell, S. G., & Batt, R. J. (2008). Modeling for Insight: A Master Class for Business Analysts. In *Modeling for Insight: A Master Class for Business Analysts*.
- Robert, C., & Casella, G. (2011). A Short History of Markov Chain Monte Carlo: Subjective Recollections from Incomplete Data. *Statistical Science*, 26(1), 102–115.
- Roberts, G. O., & Rosenthal, J. S. (2009). Examples of Adaptive MCMC. *Journal of Computational and Graphical Statistics*, 18(2), 349–367.
- Roussas, G. G. (1997). A Course in Mathematical Statistics. Second Edition. In *Academic Press* (Vol. 134, Issue 4).
- Salim, F. C., Nugroho, D. B., & Susanto, B. (2016). Model Volatilitas GARCH(1,1) dengan Error Student-t untuk Kurs Beli EUR dan JPY terhadap IDR. *Jurnal MIPA*, 39(1), 63–69.
- Sartika, Q. R., Widiyah, T., & Mukid, M. A. (2019). Value at Risk Instock Portfolio Using t-Copula. *Media Statistika*, 12(2), 175–187.
- Solomon, J. (2015). *Numerical Algorithms: Methods for Computer Vision, Machine Learning, and Graphics*. CRC Press,.
- Tyas, M. D. P., Maruddani, D. A. I., & Rahmawati, R. (2019). Perhitungan Value at Risk dengan Pendekatan Threshold Autoregressive Conditional Heteroscedasticity-Generalized Extreme Value. *Media Statistika*, 12(1), 73–85.
- Wu, Q., & Vos, P. (2018). Inference and Prediction. In V. N. Gudivada & C. R. Rao (Eds.), *Computational Analysis and Understanding of Natural Languages: Principles, Methods and Applications* (pp. 131–192). North-Holland.