

GARCH Models under Power Transformed Returns: Empirical Evidence from International Stock Indices

Didit B. Nugroho
Satya Wacana
Christian University

Tundjung Mahatma
Satya Wacana
Christian University

Yulius Pratomo
Satya Wacana
Christian University

Abstract

This study evaluates the empirical performance of four power transformation families: extended Tukey, Modulus, Exponential, and Yeo–Johnson, in modeling the return in the context of GARCH(1,1) models with two error distributions: Gaussian (normal) and Student- t . We employ an Adaptive Random Walk Metropolis method in Markov Chain Monte Carlo scheme to draw parameters. Using 19 international stock indices from the Oxford-Man Institute and basing on the log likelihood, Akaike Information Criterion, Bayesian Information Criterion, and Deviance Information Criterion, the use of power transformation families to the return series clearly improves the fit of the normal GARCH(1,1) model. In particular, the Modulus transformation family provides the best fit. Under Student's t -error distribution assumption, the GARCH(1,1) models under power transformed returns perform better in few cases.

Keywords: adaptive random walk Metropolis, GARCH(1,1), power transformations, volatility.

1. Introduction

Normality is assumed in various statistical methods such as hypothesis testing, linear regression analysis, econometric modeling, and quality control problems. Many processes, however, follow non-normal distributions perhaps due to the absence of exactly normal distributed characteristics. Transformation can be employed as an alternative approach for dealing with non-normal data. One such popular transformation is the Box-Cox family, introduced by [Box and Cox \(1964\)](#). For example, in the case of econometric models—especially Autoregressive Conditional Heteroscedastic (ARCH) model of [Engle \(1982\)](#) which was extended by [Bollerslev \(1986\)](#) to become Generalized ARCH (GARCH)—[Sarkar \(2000\)](#) and [Utami and Subanar \(2013\)](#) applied the transformation of extended Box–Cox ([Bickel and Doksum 1981](#)) and shifted Box–Cox ([Box and Cox 1964](#)), respectively, on returns data with volatilities following an ARCH(1,1)-type model. Through an empirical study, [Sarkar \(2000\)](#) found that the residual of the proposed model was closer to normality than that of the basic one. The proposed model is also able to capture the strong nonlinear dependence in financial time series data. In the study of realized variance (RV) model, [Gonçalves and Meddahi \(2011\)](#) and [Nugroho and Morimoto \(2016\)](#) applied the Box-Cox transformation on RV data, and showed

that the transformation outperforms the logarithmic transformation.

As in [Sarkar \(2000\)](#), our preliminary study showed that the parameter estimation of the extended Box–Cox transformation, i.e. $T(x, \lambda) = [\text{sign}(x)|x|^\lambda - 1] / \lambda$ for $\lambda > 0$, is significant and larger than 1, but is outperformed by the basic model. When the simple version of extended Box-Cox was applied—which is none but the extended Tukey transformation ([Tukey 1957](#)), i.e. $T(x, \lambda) = \text{sign}(x)|x|^\lambda$ for $\lambda > 0$ —the parameter estimation is significant and the extended model outperforms the basic model in terms of likelihood ratio test. This result confirms the findings of [Nugroho, Susanto, Prasetya, and Rorimpandey \(2019\)](#). With the absence of the transformation scarcity to improve statistical models, the first objective of this study is examining the performance of the power transformation families for returns in the context of GARCH model. Findings are illustrated by real examples. We give special attention to the transformation families of extended Tukey ([Tukey 1957](#)), the Exponential ([Manly 1976](#)), the Modulus ([John and Draper 1980](#)), and the Yeo–Johnson ([Yeo and Johnson 2000](#)), as all of them accommodate negative values of returns. The last three transformations were successfully applied by [Tsiotas \(2009\)](#) and [Nugroho and Morimoto \(2014\)](#) for lagged-volatility process in the context of stochastic volatility. Recently, [Nugroho \(2019\)](#) used the Yeo–Johnson function to transform the returns data that is applied to estimate a normal GARCH(1,1) model. The proposed model gives a better fit than the basic model. To the best of our knowledge, no study has been conducted on the employment of the four transformations in the context of GARCH model.

Furthermore, previous studies show evidence that financial time series often yield the leptokurtic characteristics (which means the return distribution is heavy-tailed), and even asymmetric. Financial literature accordingly suggests replacing the assumption of error normality with flexible distribution functions which are capable of accommodating heavy tails or asymmetry, such as the Student- t , generalized error, and skewed Student- t ; see, e.g., [Angelidis, Benos, and Degiannakis \(2004\)](#), [Braione and Scholtes \(2016\)](#), [Nugroho, Mahatma, and Pratomo \(2018\)](#), and [Nugroho and Morimoto \(2019\)](#). This study focuses only on two different distributions on error terms: normal and Student- t distributions, and compares the performance with that of the the power transformation families. The second goal of this study is thus extending the study done by [Sarkar \(2000\)](#), assuming the error term with Student- t distribution.

In contrast with [Sarkar \(2000\)](#) and [Utami and Subanar \(2013\)](#), who respectively suggest the maximum likelihood (ML) and second-order least square (SLS) methods for estimating their models, we use the Adaptive Random Walk Metropolis (ARWM) method of [Atchade and Rosenthal \(2005\)](#) in the Markov Chain Monte Carlo (MCMC) algorithm to estimate the model parameters. In some cases of GARCH-type models it is difficult to apply an ML-based method because of the positive and stationary conditions for variance and also the complexity of the models (see, e.g., [Nakatsuma \(2000\)](#), [Ardia \(2009\)](#), [Henneke, Rachev, Fabozzi, and Nikolov \(2011\)](#), and [Boonyakunakorn, Pastpipatkul, and Sriboonchitta \(2019\)](#)). The MCMC algorithm offers a flexible and attractive method, and has been widely used in statistics and other scientific fields.

The rest of this article is arranged as follows: Section 2 provides the methodology including the proposed model, power transformation family, and estimation method. Section 3 presents the data analysis and empirical application of the model on real financial data. The last section provides the conclusion and future work.

2. Proposed model and estimation

This section extends the GARCH(1,1) model with normal and Student- t distributed error terms, by applying a power transformation family to return series. We give attention to four power transformation families that work on real values. Firstly, GARCH model is constructed based on the assumption of normally returns errors, which are not able to accommodate the leptokurtosis commonly found in financial time series. In the Value-at-Risk (VaR) estima-

tion using GARCH model, the general finding is that the use of normal assumption might underestimate the VaR. Angelidis *et al.* (2004) and Braione and Scholtes (2016) showed the importance of allowing for heavy-tails in the distributional assumption, such as Student- t , since this distribution is able to produce better accuracy of the VaR forecasts modeled by using GARCH models. Meanwhile Satoyoshi and Mitsui (2011) and Liu, Li, and Ng (2015) considered the use of GARCH-type models with Student- t distributed errors to price options. Their empirical illustrations suggest that the proposed pricing models provide more accurate price estimates than the normal GARCH-type models.

We realize that the ML method is the most popular method to estimate GARCH parameters, but this study employs the powerful MCMC technique to deal with inference Bayesian problems. In the context of normal GARCH model, Nakatsuma (2000) and Haitao (2010) compared the MCMC method to the ML method by using real financial data and showed that the MCMC method is more precise and reliable. In particular, Nakatsuma (2000) employed the Metropolis–Hastings algorithm in the MCMC procedure and obtained that the numerical standard error of the MCMC is much less than the standard error of the ML for each parameter in the ARMA-GARCH model. In the context of ARCH models, Andrade (2011) also recommended the Metropolis–Hastings MCMC algorithm since it provides more accurate estimates and more robust than the ML method in estimating the credibility intervals of parameters.

2.1. GARCH(1,1) models with power transformed returns

The GARCH model is one of econometric models widely employed on modeling and forecasting of time-varied variance. A member of the family is the GARCH(1,1), implying an interpretation that a variance of tomorrow is simply a function of long-term average variance, squared return of today, and today's variance. This comes in the form:

$$\left. \begin{aligned} R_i &= \epsilon_i, & \epsilon_i &\sim N(0, \sigma_i^2), & i &= 1, \dots, n, \\ \sigma_i^2 &= \omega + \alpha \epsilon_{i-1}^2 + \beta \sigma_{i-1}^2, & i &= 2, \dots, n, \end{aligned} \right\} \quad (1)$$

where R_i denotes the asset return of an asset at time i .

Model (1) has some constraints including $\omega > 0$, $\alpha, \beta \geq 0$ for positive variance, and $\alpha + \beta < 1$ for the covariance stationarity. In most cases, we get a model with a small α and large β . The $\alpha + \beta$ magnitude is called variance persistence, measuring the endurance of the innovation effects on future volatility (or, how quickly the volatility to decay towards its averages). When the sum is 1.0, we have a long memory in volatility, then the IGARCH model should be employed. The high persistence (approaching 1.0) indicates slow decay, while low persistence indicates fast decay.

This study generalizes the GARCH(1,1) model by applying four power transformation families for return series. Given a transformation function indicated by T and involving a transformation parameter λ_T , the return equation in Model (1) is now expressed as

$$T(R_i, \lambda_T) = \epsilon_i, \quad \epsilon_i \sim N(0, \sigma_i^2),$$

where $T(R_i, \lambda_T)$ represents a power transformation function upon data R_i . For any family T with the parameter λ_T , the likelihood function of original data R_i , expressed by $L(R_i | \sigma_i^2, \lambda_T, T)$, is fully determined through an inverse transformation $T(R_i, \lambda_T) \rightarrow R_i$. Thus this consists of transformed likelihood data function multiplied by Jacobian of involves transformation, i.e.

$$L_i(R_i | \omega, \alpha, \beta, \lambda_T, T) = L_i(T(R_i, \lambda_T) | \omega, \alpha, \beta, \lambda_T, T) \times J(R_i, \lambda_T | T), \quad (2)$$

with the Jacobian term is expressed by $J(R_i, \lambda_T | T) = \frac{\partial T(R_i, \lambda_T)}{\partial R_i}$. Let us denote the observed data by $\mathbf{R} = (R_1, R_2, \dots, R_n)$. For instance, by taking the natural logarithm for the likelihood

function L_i di Eq. (2), the log-likelihood function for the model is given by

$$\begin{aligned} \log L(\mathbf{R}|\omega, \alpha, \beta, \lambda_T, T) = & -\frac{n}{2} \log(2\pi) - \frac{1}{2} \sum_{i=1}^n \log(\sigma_i^2) - \frac{1}{2} \sum_{i=1}^n \frac{[T(R_i, \lambda_T)]^2}{\sigma_i^2} \\ & + \sum_{i=1}^n \log(J(R_i, \lambda_T|T)) \end{aligned}$$

When the conditional distribution of R_i is the standardised t -distribution with zero mean, the variance σ_i^2 , and degrees of freedom ν , following the definition of t -distribution in [Bollerslev \(1987\)](#), the log-likelihood equation for the model is expressed as

$$\begin{aligned} \log L(\mathbf{R}|\omega, \alpha, \beta, \nu, \lambda_T, T) = & n \log \Gamma\left(\frac{\nu+1}{2}\right) - n \log \Gamma\left(\frac{\nu}{2}\right) - \frac{1}{2} \sum_{i=1}^n \log(\pi(\nu-2)\sigma_i^2) \\ & - \frac{\nu+1}{2} \sum_{i=1}^n \log\left(1 + \frac{[T(R_i, \lambda_T)]^2}{(\nu-2)\sigma_i^2}\right) + \sum_{i=1}^n \log(J(R_i, \lambda_T|T)) \end{aligned}$$

2.2. Transformation families

A conventional power transformation was firstly proposed by [Tukey \(1957, 1977\)](#) to simplify data analysis, addressing on three characteristics: (1) effects are additive; (2) the error variability is constant; and (3) the error distribution is symmetrical and possibly nearly normal. The transformation was then modified by [Box and Cox \(1964\)](#) to improve additivity, normality, and homoscedasticity of a set of observed data. The transformations of [Tukey](#) and [Box and Cox](#) works only on positive real values. To accommodate all real numbers, several modifications were already proposed. This study discusses and compares four transformation families, namely the Extended Tukey (ET), Eksponensial (Exp), Modulus (Mod), and Yeo–Johnson (YJ), that are applied to returns series whose volatility follows the GARCH model.

[Manly \(1976\)](#) proposed an exponential transformation allowing negative values. This was reported as effective to transform skew unimodal distribution to normal-like, almost symmetrical distribution, but not good enough for bimodal or U-shaped distribution.

[John and Draper \(1980\)](#) proposed a modification called the Modulus transformation allowing negative values. They reported that the transformation works most effectively on a distribution which is nearly symmetrical to a centre point, and change the long-tailed distribution to become more normal. The basic idea is applying the same power transformation to two tailed distribution which is symmetrical to zero.

[Yeo and Johnson \(2000\)](#) proposed a power transformation family suitable for reducing skewness to normality and having many good characteristics such as the Box–Cox power transformation family. While estimating the transformation parameters, they found a parameter value which minimise the Kullback-Leibler distance between the normal and transformed distribution.

Thus while the BC, Exp, and YJ transformations are used to make skewed distribution to become more symmetrical and normal, the modulus transformation family proposed by [John and Draper \(1980\)](#) can be used to eliminate the positive kurtosis.

The formulas for the transformation families compared in this study are shown in [Table 1](#), with associated Jacobian added. The signum function is defined as

$$\text{sign}(y) = \begin{cases} +1, & y > 0, \\ 0, & y = 0, \\ -1, & y < 0. \end{cases}$$

Table 1: The four transformation families and their associated Jacobian

Family T	$T(R_i, \lambda_T)$	$J(R_i, \lambda_T T)$
Tukey	$= \begin{cases} \text{sign}(R_i) R_i ^{\lambda_T}, & \lambda_T > 0 \\ \text{sign}(R_i) \log(R_i), & \lambda_T = 0 \end{cases}$	$= \begin{cases} \lambda_T R_i ^{\lambda_T-1}, & \lambda_T > 0 \\ \text{sign}(R_i) R_i ^{-1}, & \lambda_T = 0 \end{cases}$
Exp.	$= \begin{cases} \frac{\exp\{\lambda_T R_i\} - 1}{\lambda_T}, & \lambda_T \neq 0 \\ R_i, & \lambda_T = 0 \end{cases}$	$= \exp\{\lambda_T R_i\}$
Mod.	$= \begin{cases} \frac{\text{sign}(R_i)[(R_i + 1)^{\lambda_T} - 1]}{\lambda_T}, & \lambda_T \neq 0 \\ \text{sign}(R_i) \log(R_i + 1), & \lambda_T = 0 \end{cases}$	$= (R_i + 1)^{\lambda_T-1}$
YJ	$= \begin{cases} \frac{(R_i + 1)^{\lambda_T} - 1}{\lambda_T}, & R_i \geq 0, \lambda_T \neq 0 \\ \log(R_i + 1), & R_i \geq 0, \lambda_T = 0 \\ \frac{(1 - R_i)^{2-\lambda_T} - 1}{\lambda_T - 2}, & R_i < 0, \lambda_T \neq 2 \\ -\log(-R_i + 1), & R_i < 0, \lambda_T = 2 \end{cases}$	$= \begin{cases} (R_i + 1)^{\lambda_T-1}, & R_i \geq 0 \\ (-R_i + 1)^{1-\lambda_T}, & R_i < 0 \end{cases}$

The values $\lambda_{ET} = 1$, $\lambda_{Exp} = 0$, $\lambda_{Mod} = 1$, and $\lambda_{YJ} = 1$ correspond with no transformation.

2.3. Parameter estimation

The MCMC algorithm involves two steps; see [Robert and Casella \(2011\)](#) and [Geyer \(2011\)](#). The first step is the construction of Markov chain having a posterior distribution as its stationary distribution. The second step involves the use of the output of Markov chain to calculate some posterior estimation such as the mean, standard deviation, and credible set, using the Monte Carlo approximation. Previous papers have applied the MCMC algorithm in modeling the volatility, and shown its applicability. Further explanation on MCMC can be found on [Robert and Casella \(2011\)](#).

There are several ways to construct the Markov chain. One of the most popular classes is a method based on the random walk Metropolis (RWM). The theoretical explanation of RWM is provided by [Roberts and Rosenthal \(2001\)](#) on optimal scaling and posterior shape, and by [Roberts, Robert, and Frigesso \(2003\)](#) on convergence. In the efforts to improve the efficiency of RWM in terms of convergence, [Atchade and Rosenthal \(2005\)](#) proposed an adaptive model of the method, named the adaptive RWM (ARWM). The model sequentially sets the parameter proposal to achieve optimal acceptance rate. [Nugroho and Susanto \(2017\)](#) and [Nugroho \(2018\)](#) applied the method in the context of GARCH-type models. The following is the ARWM algorithm to update a parameter value of x :

- (i) Initialise the parameter x_0 and step size s_0 .
- (ii) At the $(j + 1)$ iteration, given x_j and s_j .
 - (a) Draw the proposal $x^* = x_j + s_j \cdot q_j$, where $q_j \sim N(0, 1)$.
 - (b) Calculate the Metropolis ratio: $r(x_j, x^*) = \frac{p(x^*|\cdot)}{p(x_j|\cdot)}$, where $p(x|\cdot)$ represents a posterior distribution, and $\delta = \min\{1, r(x_j, x^*)\}$ is defined.
 - (c) If $\delta > u$ for $u \sim \text{Uniform}(u; 0, 1)$, then the proposal is accepted and $x_{j+1} = x^*$; otherwise, the proposal is rejected and $x_{j+1} = x_j$.
 - (d) Calculate: $s^* = \max \left\{ s_{min}, s_j + \frac{\frac{m(x^*)}{(j+1)} - \tau}{(j+1)^\eta} \right\}$, where $m(x^*)$ is the frequency of proposal acceptance x^* and τ is the expected acceptance probability. If $s^* > s_{max}$,

then $s_{j+1} = s_{max}$; if $s^* < s_{max}$, then $s_{j+1} = s_j$.

Following Roberts and Rosenthal (2009), we set $s_{min} = 10^{-5}$ and $s_{max} = 10$ as such so that the asymptotic acceptance rate of the algorithm is approximately 0.44, and $\lambda = 0.6$. The choice of step size scale of s_i has big impacts on the changes of proposal values. Intuitively, when s_i^2 is very small, the algorithm resulting movement is also very small, which further resulting a poor mixing time. Contrarily, if s_i^2 is very large, then the movement is also large, resulting in the rejection of the proposal, and thus a poor mixing algorithm.

Furthermore, here we accept the fact that the proposal distribution for the Metropolis ratio in the ARWM method can be calculated using the posterior distribution $p(\theta|\mathbf{R})$ for a parameter $\theta = (\omega, \alpha, \beta, \nu, \lambda_T)$. In the Bayesian framework, the posterior distribution is the product of likelihood function and prior distribution:

$$p(\theta|\mathbf{R}) \propto L(\mathbf{R}|\theta) \times p(\theta),$$

where $p(\theta)$ is the prior distribution capturing prior uncertainty on the parameter θ . In this study we adopt the prior distribution:

$$\omega, \alpha, \beta \sim N(0, 1000) \mathbf{1}_{[\omega, \alpha, \beta > 0]},$$

where $\mathbf{1}_{[x]} = 1$ if condition x holds and 0 otherwise, as in Ardia and Hoogerheide (2010),

$$\nu \sim \text{Exp}(0.01) \mathbf{1}_{[\nu > 2]}$$

as in Deschamps (2006) and $\lambda_T \sim N(0, 1000)$.

3. Empirical application

3.1. Data and preliminary findings

Before estimation, we shall look at the sample data, which—in this case—is the daily returns. It consists of 19 international stock indices, freely available at the Oxford-Man Institute's Realised Library (<https://realized.oxford-man.ox.ac.uk/data/download>). The study involves data within the period of January 2000 to December 2007. The full sample are the following: AEX Index, Bovespa, CAC 40, DAX, DJIA, Euro STOXX 50, FTSE 100, FTSE MIB, FT Straits Times, Hang Seng, IBEX 35, IPC Mexico, KOSPI Composite, Nasdaq 100, Nikkei 225, Russel 2000, S&P 500, S&P CNX Nifty, and Swiss Market.

Table 2 reports standard summary statistics of the 19 stock indices. The skewness and kurtosis of the return series result in some deviation from a normal standard distribution as indicated by Jarque-Bera (JB) normality test with critical value of 5.98, at the 5% significance level. Therefore, assuming normality for the errors when estimating the models should be rejected. It is suggested that the error distribution of returns is assumed to be not normal. Specifically the S&P CNX Nifty data is highly skewed, as the skewness is less than -1 ; Nikkei 225 data is negatively skewed, as the skewness is between -1 and -0.5 ; the other ones are fairly symmetrical, as the skewness is between -0.5 and 0.5 . For this characteristic, it is probable to transform the return data so that its distribution can approximate the symmetric distribution (Sarkar 2000), a part from its ability to capture the presence of nonlinear dependence in time series data. Meanwhile the kurtosis for all series is greater than 3 which indicates a leptokurtic distribution (i.e. tails are heavier than the normal distribution). Therefore our presumption is that the Modulus transformation introduced by John and Draper (1980) is better than the others.

3.2. Estimation results

For parameter estimation, this study employs the MCMC simulation with 6000 iterations for the Markov chain, where the initial 1000 samples are discarded (the practice of so-called burn-in) and the remaining samples are stored for Monte Carlo calculation. The burn-in is done

Table 2: Skewness, kurtosis, and JB Statistics for untransformed returns

Stick Index	Mean	SD	Skewness	Kurtosis	JB Stats
AEX Index	-0.042	1.156	-0.203	9.77	8660.8
Bovespa	0.003	1.717	-0.173	8.10	4748.1
CAC 40	-0.037	1.205	-0.147	7.29	3477.9
DAX	-0.030	1.305	-0.099	7.98	4642.7
DJIA	0.023	1.121	-0.009	11.70	13969.6
Euro STOXX 50	-0.026	1.332	-0.200	7.76	4281.0
FTSE 100	-0.036	0.936	-0.143	7.42	3629.5
FTSE MIB	-0.019	0.976	-0.305	11.50	13434.6
FT Straits Times	-0.034	0.915	-0.374	11.94	12823.3
Hang Seng	-0.048	1.000	-0.072	15.93	28524.6
IBEX 35	-0.051	1.260	-0.035	8.41	5485.6
IPC Mexico	0.037	1.284	-0.001	8.22	5060.1
KOSPI Composite	-0.044	1.183	-0.342	9.02	6676.0
Nasdaq 100	-0.018	1.359	0.105	10.60	10703.0
Nikkei 225	-0.033	1.164	-0.553	13.87	21361.5
Russel 2000	0.010	1.405	-0.263	7.41	3653.1
S&P 500	0.010	1.164	-0.174	11.06	12051.2
S&P CNX Nifty	0.023	1.199	-1.007	13.88	19736.0
Swiss Market	-0.019	0.976	-0.305	11.49	13434.6

to eliminate the worst early samples and its implementation was successful in our MCMC calculations although not reported (see, e.g., Figure 1). It means that our Markov chains (after the burn-in period) have reached stationarity. For our data set, we set the plausible default initial values:

$$\omega_0 = 0.01, \alpha_0 = 0.2, \beta_0 = 0.7, \lambda_{T,0} = 0.5, \nu_0 = 10.$$

Notice that although the initial values vary (even for the very poor initial values $\lambda_{T,0} = 4$ and $\nu_0 = 40$) in our experiments, the posterior results are almost equal to the results obtained for the above values. This result is consistent to [Ardia and Hoogerheide \(2010\)](#). This is one of the advantages of the MCMC method compared to the ML method which is often sensitive to initial values ([Takaishi 2010](#)).

Furthermore, the efficiency of Markov chain might be measured by Integrated Autocorrelation Time (IACT). In this study, IACT was estimated using adaptive truncated periodogram estimator of [Sokal \(1997\)](#). Although not reported, we particularly notice that the Markov chains of ν and λ_T are well mixed, which are indicated by the values—less than 50—of the IACT. The IACT quantity can be interpreted as the number of MCMC samples required for an independent sample to be drawn, or equivalently, the length of Markov Chain (after the burn-in) divided by the Effective Sample Size (ESS). Therefore, we have ESSs larger than 100. According to [Givens and Hoeting \(2013\)](#), it indicates that the estimation method is efficient to estimate the parameters of degrees of freedom and power transformation, which are not simple, and recommends to make reliable inferences on the interest parameters. In a visual way, for example, Figure 1 displays the traceplots of the posterior samples for the NLR-GARCHt(1,1) models adopting FTSE 100 returns and shows that the chains fluctuate around their means.

Due to space limitation, this study only presents the posterior mean for λ_T in the GARCHn(1,1) model as in Table 3 and for (ν, λ_T) in the GARCHt(1,1) model as in Table 4. The asterisk symbol indicates a significant estimated value in terms of 95% Highest Posterior Density (HPD) interval, meaning that the interval does not contain a value corresponding to no transformation. Here the HPD interval is calculated using the algorithm proposed by [Chen and Shao \(1999\)](#).

For the GARCHn(1,1) model, it can be seen that almost all sample data provide significant

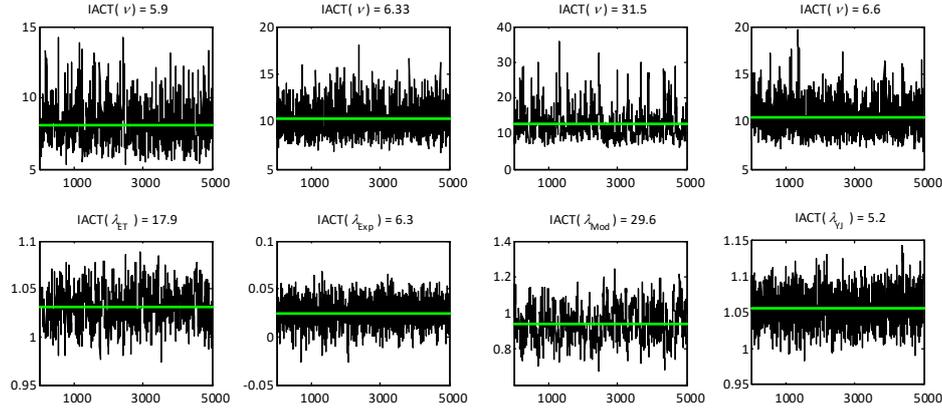


Figure 1: Trace plots of posterior samples for the NLR-GARCH $t(1,1)$ models adopting FTSE 100 returns by applying, from left to right: ET, Exp, Mod, and YJ transformations. The green line represents a posterior mean

evidence for all considered power transformations. Modulus transformation provides a significant estimated value for λ_T in all data cases. ET transformation is not significant for Russell 2000 data only, Exp transformation is not significant for CAC 40, FT Straits Times, and IBEX 35, while YJ transformation is not significant for CAC 40 and FT Straits Times data. These results indicate the success of power transformation families to transform the most of financial time series returns data in the context of GARCH models with normally distributed errors. In particular, Mod transformation is the most successful for capturing nonlinearities of the return time series.

Table 3: Parameter estimates of λ_T for GARCH $n(1,1)$ models. The values in parentheses are the corresponding standard deviations

Data	ET	Exp	Mod	YJ
AEX Index	0.9115* (0.0103)	0.0212* (0.0078)	0.5861* (0.0384)	1.0418* (0.0147)
Bovespa	0.9539* (0.0111)	0.0137* (0.0059)	0.8033* (0.0315)	1.0251* (0.0123)
CAC 40	0.9033* (0.0110)	0.0114 (0.0076)	0.5904* (0.0375)	1.0223 (0.0144)
DAX	0.9077* (0.9073)	0.0194* (0.0072)	0.6455* (0.0343)	1.0412* (0.0138)
DJIA	0.8967* (0.0106)	0.0199* (0.0081)	0.5702* (0.0390)	1.0416* (0.0157)
Euro STOXX 50	0.9037* (0.0106)	0.0168* (0.0067)	0.6237* (0.0347)	1.0322* (0.0131)
FTSE 100	0.9386* (0.0111)	0.0334* (0.0102)	0.6508* (0.0451)	1.0712* (0.0174)
FTSE MIB	0.9079* (0.0103)	0.0532* (0.0079)	0.5130* (0.0392)	1.0925* (0.0150)
FT Straits Times	0.9520* (0.0122)	-0.0025 (0.0111)	0.6602* (0.0465)	1.0045 (0.0188)
Hang Seng	0.9177* (0.0118)	0.0156* (0.0083)	0.5836* (0.0417)	1.0312* (0.0160)
IBEX 35	0.9100* (0.0107)	0.0076 (0.0062)	0.6136* (0.0342)	1.0315* (0.0131)
IPC Mexico	0.9117* (0.0104)	0.0139* (0.0067)	0.6148* (0.343)	1.0289* (0.0137)
KOSPI Composite	0.9344* (0.0113)	0.0513* (0.0082)	0.6768* (0.0372)	1.0951* (0.0147)
Nasdaq 100	0.9544* (0.0110)	0.0114* (0.0067)	0.7909* (0.0349)	1.0514* (0.0142)
Nikkei 225	0.8743* (0.0103)	0.0655* (0.0066)	0.4684* (0.0360)	1.1230* (0.0136)
Russel 2000	0.9834 (0.0120)	0.0204* (0.0075)	0.9038* (0.0363)	1.0408* (0.0141)
S&P 500	0.8935* (0.0103)	0.0268* (0.0082)	0.5648* (0.0395)	1.0570* (0.0153)
S&P CNX Nifty	0.8738* (0.0105)	0.0593* (0.0073)	0.4613* (0.0383)	1.1006* (0.0151)
Swiss Market	0.9086* (0.0106)	0.0529* (0.0081)	0.5157* (0.0395)	1.0927* (0.0153)

Although not reported, this study has also studied and compared the residuals of competing models to determine the transformations that are successful in transforming returns so that the distribution of residuals approaches normal distribution. For all cases, we find that—although still deviate from normality—all power transformations successfully reduce the coefficients of

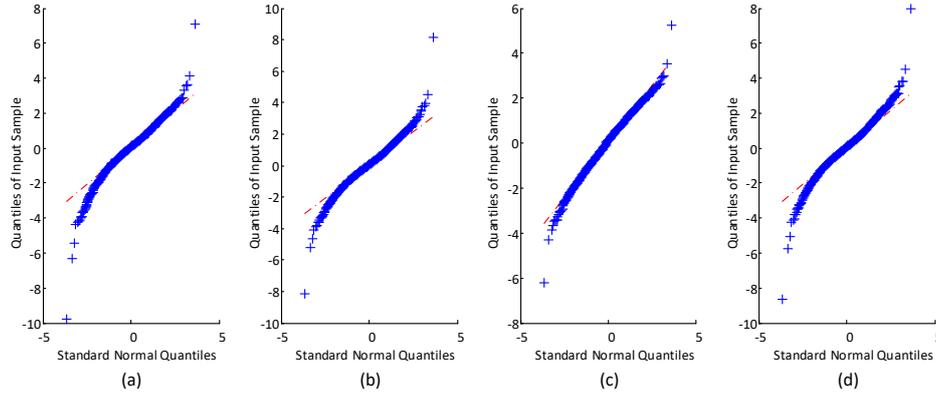


Figure 2: QQ plots of S&P CNX Nifty returns (a) without transformation, (b) with ET transformation, (c) with Exp transformation, (d) with Mod transformation, and (e) with YJ transformation against the standard normal distribution

skewness and kurtosis as well as the JB statistics, even when the estimation of λ_T is not significant. The Exp, Mod, and YJ transformations provide the best performance to reduce the skewness coefficient in the cases of six, five, and eight data, respectively. When comparing the coefficient of kurtosis, this study finds that Mod transformation is the most effective for reducing the coefficient of kurtosis in all cases of data. Furthermore, the results of the normality test confirm the previous assumption that the Mod transformation is most suitable for all cases of return data, which is indicated by the smallest JB statistical value. Moreover, the quantile-quantile (QQ) plots in Figure 2, for example, demonstrate that the estimated residuals from fitting the S&P CNX Nifty daily returns to the NLR-GARCH(1,1) models are closer to normal distribution. In particular, the panel (d) suggests that the specification of Mod transformation provides the best fit.

Move to the model with the Student- t distribution specification for return errors, Table 4 shows that the parameter of power transformation is no longer significant in most cases when the normal distribution is replaced by the Student- t distribution. It is very clear because the use of different distributions for return errors will affect the treatment of returns. In this case, the Student- t distribution causes a weak support for the power transformation of return. The parameter estimates are significant for six, five, nine, and six data when applying the ET, Exp, Mod, and YJ transformations, respectively. Regarding the degrees of freedom parameters ν , the ET, Exp, and YJ transformations do not affect the estimate of ν . It means that the estimates of ν from the two model classes are close each other. In contrast, Mod transformation increases the degrees of freedom in all data cases, except the FT Straits Times data. It means that the tail-thickness of Student- t distribution is reduced, which is in accordance with the objective of Mod transformation which changes long-tailed distribution to be more normal. Although the power transformations are supported by several data, the next section discusses whether the significance of λ_T is sufficient to provide evidence against the basic GARCH t (1,1) model.

3.3. Model selection

In modeling case, one of the important parts of statistical analysis is the selection of the best model from a set of candidate models, given a set of data. Since the NLR-GARCH models nest a basic GARCH model, Log-likelihood Ratio Test (LRT) can be used to compare those two models. The LRT statistic for the general model M_g against the simpler model M_s is calculated as

$$\text{LRT} = 2 \left(\log L(\mathbf{R}|M_g, \hat{\theta}) - \log L(\mathbf{R}|M_s, \hat{\theta}) \right),$$

Table 4: Parameter estimates of (ν, λ_T) for GARCHt(1,1) models

Data	(ν, λ_T)				
	NoT	ET	Exp	Mod	YJ
AEX Index	7.56	(6.91,1.0168)	(7.47,0.0002)	(12.08,0.8348*)	(7.50,1.0102)
Bovespa	12.45	(10.36,1.0184)	(12.22,0.0055)	(19.76,0.9082)	(12.04,1.0098)
CAC 40	7.77	(8.32,0.9902)	(7.75,-0.0006)	(16.90,0.7767)	(7.81,1.0057)
DAX	8.36	(9.25,0.9855)	(8.27,0.0065)	(13.67,0.8557*)	(8.32,1.0223)
DJIA	7.80	(9.58,0.9684*)	(7.62,-0.0027)	(14.43,0.7839*)	(7.69,0.9997)
Euro STOXX 50	8.10	(8.89,0.9846)	(7.93,0.0028)	(15.34,0.8151*)	(7.96,1.0074)
FTSE 100	9.90	(8.06,1.0316*)	(10.24,0.0249*)	(12.52,0.9403)	(10.49,1.0558*)
FTSE MIB	7.33	(5.60,1.0517*)	(7.31,0.0171)	(8.12,0.9541)	(7.25,1.0336)
FT Straits Times	9.63	(6.21,1.0750*)	(9.64,-0.0126)	(9.23,1.0463)	(9.69,0.9925)
Hang Seng	8.23	(8.53,0.9983)	(8.25,0.0011)	(20.16,0.7590*)	(8.35,1.0105)
IBEX 35	7.77	(7.28,1.0099)	(7.69,0.0046)	(12.48,0.8403*)	(7.69,1.0266)
IPC Mexico	7.63	(6.98,1.0180)	(7.62,-0.0029)	(11.22,0.8854)	(7.69,0.9979)
KOSPI Composite	9.38	(7.55,1.0362*)	(10.19,0.0385*)	(11.65,0.9422)	(10.13,1.0731*)
Nasdaq 100	11.65	(9.27,1.0290)	(11.41,0.0072)	(12.08,0.9995)	(11.87,1.0496*)
Nikkei 225	6.66	(6.87,0.9916)	(6.87,0.0288*)	(10.94,0.7719*)	(6.78,1.0511*)
Russel 2000	23.39	(19.38,1.0169)	(25.38,0.0169*)	(26.05,0.9848)	(25.91,1.0366*)
S&P 500	7.53	(8.92,0.9724)	(7.52,0.0037)	(12.09,0.8130*)	(7.54,1.0153)
S&P CNX Nifty	5.84	(5.89,1.0007)	(6.13,0.0421*)	(10.02,0.7503*)	(5.97,1.0645*)
Swiss Market	7.32	(5.55,1.0528*)	(7.31,0.0172)	(7.94,0.9635)	(7.29,1.0339)

where $\hat{\theta}$ denotes the vector of parameter estimates. The criterion is based on the ratio likelihood having an asymptotic chi-squared distribution, with degrees of freedom equal to the number of additional parameters in the general model. For our study, notice that at the 1%, 5%, and 10% level of significance with 1 degree of freedom, we reject M_s when the LRTs are greater than 6.64, 3.84, and 2.71, respectively. In these cases, the LRT tells us to favor M_g over M_s .

Furthermore, the models in a class of NLR-GARCH models are non-nested, so we can not compare them by conducting an LRT. We therefore computed the value of three selection criteria which do not assume nested models, namely Akaike Information Criterion of [Akaike \(1976\)](#), Bayesian Information Criterion (BIC) of [Schwarz \(1978\)](#), and Deviance Information Criterion (DIC) of [Spiegelhalter, Best, Carlin, and van der Linde \(2002\)](#). The AIC and BIC are respectively calculated using:

$$\text{AIC} = 2k - 2 \log L(\mathbf{R}|\theta_{max}) \quad \text{and} \quad \text{BIC} = k \log(s) - 2 \log L(\mathbf{R}|\theta_{max}),$$

with k is the number of model parameters, s is the number of data points, and θ_{max} is the parameter set that maximizes the observed data likelihood $L(\mathbf{R}|\theta)$. Meanwhile, DIC is a Bayesian generalization of AIC and is based on the posterior distribution of the deviance $D(\theta) = -2 \log p(\mathbf{R}|\theta)$. The formula of DIC is expressed by

$$\text{DIC} = 2\overline{D(\theta)} - D(\bar{\theta}) = -4E_{\theta|\mathbf{R}}[\log L(\mathbf{R}|\theta)] + 2 \log L(\mathbf{R}|\bar{\theta}),$$

with $\bar{\theta}$ is the parameter mean. Notice that the models associated with the lower AIC, BIC, and DIC are more adequate.

Table 5 reports the log-likelihood values for all models that are fitted to the observed data. It can be seen in the case of NLR-GARCHn(1,1) models with a significant parameter for λ_T , the LRT at the 5% and 10% levels of significance favors the models and provides an evidence against the basic model. This finding is confirmed by AIC and DIC, where Table 6 shows that the DIC values for the NLR-GARCHn(1,1) models are smaller than the basic model in almost all cases. (The AIC estimates are not presented here because of space limitation and its similarity to DIC.) The AIC and DIC results show that the basic model outperforms

the following NLR-GARCH $n(1,1)$ models in a few cases: model with ET transformed return for Russel 2000 data series, the model with Exp transformed return for CAC 40, FT Straits Times, and IBEX 35 data series, and the model with YJ transformed returns for CAC 40 and FT Straits Times data series. In particular, we found that the GARCH $n(1,1)$ model with Mod transformed return best fits data, indicated by smallest AIC and DIC values, followed by ET specification (except for KOSPI Composite and Russel 2000 data series). This result is similar to those of the BIC presented in Table 7, which the only exception is the Russel 2000 data series. The third and fourth best ranks are very competitive between the Exp and YJ transformations. In this case, the BIC selected fewer data than the AIC and DIC. Both AIC and DIC provide evidence for 16 data, whereas the BIC provides evidence for seven data only. Those findings therefore confirm the previous findings that the financial return series should be transformed using power transformations, specifically Mod transformation, in the context of GARCH $(1,1)$ model with normally distributed errors.

After we compare the models with Student- t specification based on the LRT statistics, there are seven data adopting the ETR-GARCH $t(1,1)$ model, five data adopting the ExpR-GARCH $t(1,1)$ model, five data adopting the ModR-GARCH $t(1,1)$ model, and six data adopting the YJR-GARCH $t(1,1)$ model which provide evidence against the GARCH $t(1,1)$ model. On that result, both AIC and DIC indicate that each of ET, Mod, and YJ transformations performs the best fit on three sample data. A similar result is produced by BIC value (except for CAC 40 and IPC Mexico in the Mod specification). These results seem to weakly support the specification of the power transformation families for return series in most cases when the return errors are Student- t distributed.

Furthermore, both AIC and DIC indicate that the superiority of NLR-GARCH $t(1,1)$ model is supported by two sample data with insignificant in the ET parameter and three sample data with insignificant in the Exp and Mod parameters. Meanwhile, in the case of data providing significant evidence in the transformation parameter, the ETR-GARCH $t(1,1)$, ExpR-GARCH $t(1,1)$, ModR-GARCH $t(1,1)$, and YJR-GARCH $t(1,1)$ models outperform the GARCH $t(1,1)$ model on 5 of 6, 2 of 5, 1 of 9, and 3 of 6 sample data, respectively. The results on Mod and YJ specifications are similar to the BIC results. According to BIC the ETR-GARCH $t(1,1)$ and ExpR-GARCH $t(1,1)$ models outperform the GARCH $t(1,1)$ model on 3 of 6 and 1 of 5 sample data with a significant transformation parameter. Those findings show that the statistical significance of the transformation parameter in the NLR-GARCH $t(1,1)$ models can not guarantee whether the model performs better than the basic GARCH $t(1,1)$. On the basis of DIC estimates, the FTSE 100, FT Straits Times, Hangseng, IPC Mexico, and KOSPI Composite data series support all transformation specifications, where the best performance is produced by different transformations.

4. Conclusion and future work

This study investigated the performance of four power transformation families, i.e. the extended Tukey, exponential, Modulus, and Yeo–Johnson, in transforming the returns of GARCH $(1,1)$ models. The proposed models, namely NLR-GARCH $(1,1)$ models, are extensions of the basic GARCH model of Bollerslev (1986, 1987) and provide the alternative models to Sarkar (2000) and Nugroho *et al.* (2019). We performed the Bayesian inference of the proposed model by the ARWM sampler in the MCMC method. Firstly, in terms of autocorrelation time, we observed that the ARWM sampler has high efficiency for sampling the transformation parameter in the nonlinear function. It is concluded that our MCMC algorithm works well and the ARWM sampler can reliably estimate the key parameters in the NLR-GARCH $(1,1)$ models. The reader can easily extend and employ the computing algorithm for the other specifications of NLR-GARCH models. Secondly, all observed return series were found to support all considered power transformation families in the context of normal GARCH $(1,1)$ model where the Modulus transformation appears to be the best fitting model, and this outperformance is always statistically significant. Therefore this study concludes

that the power transformation families can be applied successfully to transform the returns for GARCH(1,1) model with normally distributed errors. Meanwhile, the results of Student- t NLR-GARCH(1,1) models indicate that the performance of transformations does not depend on the significance of transformation parameter.

Areas for further research would include investigating the skew version of error distribution in the returns. It would also be interesting to apply the power transformation families to the volatility process in order to extend the results of Nugroho *et al.* (2019).

Acknowledgement

This work was supported by: (1) Higher Education Excellent Fundamental Research Grant of The Ministry of Research, Technology and Higher Education of the Republic of Indonesia for the 2019 fiscal year (Contract Number 001/L6/AK/SP2H.1/PENELITIAN/2019), and (2) Universitas Kristen Satya Wacana's Internal Funding for the 2018–2019 research (Contract Number 037/Penel./Rek./1/V/2019).

References

- Akaike H (1976). "A New Look at the Statistical Model Identification." *IEEE Transactions on Automatic Control*, **19**(6), 716–723. doi:10.1109/TAC.1974.1100705.
- Andrade MG (2011). "A Comparative Study of Bayesian and Maximum Likelihood Approaches for ARCH Models with Evidence from Brazilian Financial Series." *New Mathematics and Natural Computation*, **7**(2), 347–361. doi:10.1142/S1793005711001974.
- Angelidis T, Benos A, Degiannakis S (2004). "The Use of GARCH Models in VaR Estimation." *Statistical Methodology*, **1**(1–2), 105–128. doi:10.1016/j.stamet.2004.08.004.
- Ardia D (2009). "Bayesian Estimation of a Markov-switching Threshold Asymmetric GARCH Model with Student- t Innovations." *Econometrics Journal*, **12**(1), 105–126. doi:10.1111/j.1368-423X.2008.00253.x.
- Ardia D, Hoogerheide LF (2010). "Bayesian Estimation of the GARCH(1,1) Model with Student- t Innovations." *The R Journal*, **2**(2), 41–47. doi:10.32614/RJ-2010-014.
- Atchade YF, Rosenthal JS (2005). "On Adaptive Markov Chain Monte Carlo Algorithms." *Bernoulli*, **11**(5), 815–828. doi:10.3150/bj/1130077595.
- Bickel PJ, Doksum KA (1981). "An Analysis of Transformations Revisited." *Journal of American Statistical Society*, **76**(374), 296–311. doi:10.2307/2287831.
- Bollerslev T (1986). "Generalized Autoregressive Conditional Heteroskedasticity." *Journal of Econometrics*, **31**(3), 307–327. doi:10.1016/0304-4076(86)90063-1.
- Bollerslev T (1987). "A Conditional Heteroskedastic Time Series Model for Speculative Prices and Rates of Return." *The Review of Economics and Statistics*, **69**(3), 542–547. doi:10.2307/1925546.
- Boonyakunakorn P, Pastpipatkul P, Sriboonchitta S (2019). "Value at Risk of SET Returns Based on Bayesian Markov-Switching GARCH Approach." In V Kreinovich, S Sriboonchitta (eds.), *Structural Changes and their Econometric Modeling*, pp. 329–341. Cham: Springer Nature Switzerland AG. doi:10.1007/978-3-030-04263-9_25.
- Box GEP, Cox DR (1964). "An Analysis of Transformations." *Journal of the Royal Statistical Society, Series B (Methodological)*, **26**(2), 211–252. doi:10.1111/j.2517-6161.1964.tb00553.x.

- Braione M, Scholtes NK (2016). “Forecasting Value-at-Risk under Different Distributional Assumptions.” *Econometrics*, **4**(3), 1–27. doi:10.3390/econometrics4010003.
- Chen MH, Shao QM (1999). “Monte Carlo Estimation of Bayesian Credible and HPD Intervals.” *Journal of Computational and Graphical Statistics*, **8**(1), 69–92. doi:10.2307/1390921.
- Deschamps PJ (2006). “A Flexible Prior Distribution for Markov Switching Autoregressions with Student-*t* Errors.” *Journal of Econometrics*, **133**(1), 153–190. doi:10.1016/j.jeconom.2005.03.012.
- Engle RF (1982). “Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of the United Kingdom Inflation.” *Econometrica*, **50**(4), 987–1007. doi:10.2307/1912773.
- Geyer CJ (2011). “Introduction to Markov Chain Monte Carlo.” In SP Brooks, A Gelman, GL Jones, XL Meng (eds.), *Handbook of Markov Chain Monte Carlo*, pp. 3–48. Chapman & Hall/CRC.
- Givens GH, Hoeting JA (2013). *Computational Statistics*. 2nd edition. John Wiley & Sons. doi:10.1002/9781118555552.
- Gonçalves S, Meddahi N (2011). “Box–Cox Transforms for Realized Volatility.” *Journal of Econometrics*, **160**(1), 129–144. doi:10.1016/j.jeconom.2010.03.026.
- Haitao P (2010). “The Application of MCMC Arithmetic in Time Series.” In *The Eighth Annual Meeting of Uncertain System in China*, p. 12.
- Henneke JS, Rachev ST, Fabozzi FJ, Nikolov M (2011). “MCMC-Based Estimation of Markov Switching ARMA-GARCH Models.” *Applied Economics*, **43**(3), 259–271. doi:10.1080/00036840802552379.
- John JA, Draper NR (1980). “An Alternative Family of Transformations.” *Applied Statistics*, **29**(2), 190–197. doi:10.2307/2986305.
- Liu Y, Li JSH, Ng ACY (2015). “Option Pricing under GARCH Models with Hansen’s Skewed-*t* Distributed Innovations.” *North American Journal of Economics and Finance*, **31**, 108–125. doi:10.1016/j.najef.2014.10.007.
- Manly BF (1976). “Exponential Data Transformation.” *The Statistician*, **25**(1), 37–42. doi:10.2307/2988129.
- Nakatsuma T (2000). “Bayesian Analysis of ARMA-GARCH Models: A Markov Chain Sampling Approach.” *Journal of Econometrics*, **95**(1), 57–69. doi:10.1016/S0304-4076(99)00029-9.
- Nugroho DB (2018). “Comparative Analysis of Three MCMC Methods for Estimating GARCH Models.” *IOP Conference Series: Materials Science and Engineering*, **403**(1), 012061. doi:10.1088/1757-899X/403/1/012061.
- Nugroho DB (2019). “GARCH(1,1) Model with the Yeo–Johnson Transformed Returns.” *IOP Conference Series: Journal of Physics Conference Series*, **1320**(1), 012013. doi:10.1088/1742-6596/1320/1/012013.
- Nugroho DB, Mahatma T, Pratomo Y (2018). “Modeling of Stochastic Volatility to Validate IDR Anchor Currency.” **20**(2), 165–185. doi:10.22146/gamaijb.26006.
- Nugroho DB, Morimoto T (2014). “Realized Non-Linear Stochastic Volatility Models with Asymmetric Effects and Generalized Student’s *t*-Distributions.” *Journal of The Japan Statistical Society*, **44**(1), 83–118. doi:10.14490/jjss.44.83.

- Nugroho DB, Morimoto T (2016). “Box–Cox Realized Asymmetric Stochastic Volatility Models with Generalized Student’s t -Error Distributions.” *Journal of Applied Statistics*, **43**(10), 1906–1927. doi:10.1080/02664763.2015.1125862.
- Nugroho DB, Morimoto T (2019). “Incorporating Realized Quarticity into a Realized Stochastic Volatility Model.” *Asia-Pacific Financial Markets*, **26**(4), 495–528. doi:10.1007/s10690-019-09276-2.
- Nugroho DB, Susanto B (2017). “Volatility Modeling for IDR Exchange Rate Through APARCH Model with Student- t Distribution.” *AIP Conference Proceedings*, **1868**, 040005–1–040005–8. doi:10.1063/1.4995120.
- Nugroho DB, Susanto B, Prasetya KNP, Rorimpandey R (2019). “Modeling of Returns Volatility Using GARCH(1,1) Model under Tukey Transformation.” *Jurnal Akuntansi dan Keuangan*, **21**(1), 12–20. doi:10.9744/jak.20.1.12-20.
- Robert C, Casella G (2011). “A Short History of Markov Chain Monte Carlo: Subjective Recollections from Incomplete Data.” *Statistical Science*, **26**(1), 102–115. doi:10.1214/10-STS351.
- Roberts GO, Robert C, Frigesso A (2003). “Linking Theory and Practice of MCMC.” In PJ Green, NL Hjort, S Richardson (eds.), *Highly Structured Stochastic Systems*, Oxford statistical science series, 27 edition, pp. 145–178. Oxford University Press, Oxford.
- Roberts GO, Rosenthal JS (2001). “Optimal Scaling for Various Metropolis–Hastings Algorithms.” *Statistical Science*, **16**(4), 351–367. doi:10.1214/ss/1015346320.
- Roberts GO, Rosenthal JS (2009). “Examples of Adaptive MCMC.” *Journal of Computational and Graphical Statistics*, **18**(2), 349–367. doi:10.1198/jcgs.2009.06134.
- Sarkar N (2000). “ARCH Model with Box–Cox Transformed Dependent Variable.” *Statistics & Probability Letters*, **50**(1), 365–374. doi:10.1016/S0167-7152(00)00117-6.
- Satoyoshi K, Mitsui H (2011). “Empirical Study of Nikkei 225 Options with the Markov Switching GARCH Model.” *Asia-Pacific Financial Markets*, **18**(1), 55–68. doi:10.1007/s10690-010-9120-6.
- Schwarz G (1978). “Estimating the Dimension of a Model.” *The Annals of Statistics*, **6**(2), 461—464. doi:10.1214/aos/1176344136.
- Sokal A (1997). “Monte Carlo Methods in Statistical Mechanics: Foundations and New Algorithms.” In C DeWitt-Morette, P Cartier, A Folacci (eds.), *Functional Integration*, pp. 131–192. Springer Science+Business Media, New York. doi:10.1007/978-1-4899-0319-8_6.
- Spiegelhalter DJ, Best NG, Carlin BP, van der Linde A (2002). “Bayesian Measures of Model Complexity and Fit (with Discussion).” *Journal of the Royal Statistical Society: Series B*, **64**(4), 583–639. doi:10.1111/1467-9868.02022.
- Takaishi T (2010). “Bayesian Inference with an Adaptive Proposal Density for GARCH Models.” In *Seventh International Conference on Applications of Physics in Financial Analysis*, p. 01211. IEEE.
- Tsiotas G (2009). “On the Use of Non-Linear Transformations in Stochastic Volatility Models.” *Statistical Methods and Applications*, **18**(4), 555–583. doi:10.1007/s10260-008-0113-9.
- Tukey JW (1957). “On the Comparative Anatomy of Transformations.” *The Annals of Mathematical Statistics*, **28**(3), 602–632. doi:10.1214/aoms/1177706875.
- Tukey JW (1977). *Exploratory Data Analysis*. Addison-Wesley.

- Utami H, Subanar (2013). “Second Order Least Square Estimation on ARCH(1) Model with Box–Cox Transformed Dependent Variable.” *Journal of the Indonesian Mathematical Society*, **19**(1), 99–110. doi:10.22342/jims.19.2.166.99–110.
- Yeo IK, Johnson R (2000). “A New Family of Power Transformations to Improve Normality or Symmetry.” *Biometrika*, **87**(4), 954–959. doi:10.1093/biomet/87.4.954.

Affiliation:

Didit B. Nugroho
Department of Mathematics
Satya Wacana Christian University
Jl. Diponegoro 52–60 Salatiga 50711 Jawa Tengah, Indonesia
E-mail: didit.budinugroho@uksw.edu, didit.budinugroho@staff.uksw.edu
URL: <https://fsm.uksw.edu/pages/profile-of-dbn>
ORCID iD: <https://orcid.org/0000-0002-3693-0950>

Tundjung Mahatma
Department of Mathematics
Satya Wacana Christian University
Jl. Diponegoro 52–60 Salatiga 50711 Jawa Tengah, Indonesia
E-mail: tundjung.mahatma@uksw.edu

Yulius Pratomo
Department of Economics
Satya Wacana Christian University
Jl. Diponegoro 52–60 Salatiga 50711 Jawa Tengah, Indonesia
E-mail: yulius.pratomo@uksw.edu

Table 5: Log likelihood estimates for NLR-GARCH(1,1) models

Data	GARCHn(1,1) model				
	Transformation for returns				
	NoT	ET	Exp	Mod	YJ
AEX	-6035.55	-6003.20	-6032.34	-5972.43	-6031.95
Bovespa	-8044.89	-8037.70	-8042.64	-8025.28	-8043.09
CAC 40	-6512.75	-6474.70	-6512.34	-6446.61	-6512.12
DAX Index	-6726.05	-6691.16	-6722.66	-6673.12	-6722.15
DJIA	-5734.21	-5690.74	-5731.83	-5673.41	-5731.00
Euro STOXX 50	-6868.32	-6829.91	-6865.27	-6807.07	-6865.62
FTSE 100	-5106.10	-5092.07	-5101.06	-5073.35	-5098.34
FTSE MIB	-5388.42	-5352.07	-5366.04	-5303.05	-5371.17
FT Straits Times	-4221.13	-4214.19	-4221.40	-4195.61	-4221.65
Hang Seng	-5251.24	-5226.82	<i>-5249.41</i>	-5200.76	<i>-5249.45</i>
IBEX 35	-6777.18	-6743.97	-6776.91	-6711.60	-6774.96
IPC Mexico	-6715.06	-6683.08	-6713.07	-6652.03	-6713.11
KOSPI Composite	-6031.79	-6016.23	-6013.27	-5994.26	-6013.26
Nasdaq 100	-6485.12	-6477.55	<i>-6483.98</i>	-6468.36	-6478.70
Nikkei 225	-6093.54	-6020.82	-6044.37	-5965.71	-6055.04
Russel 2000	-7093.31	-7092.79	-7090.34	-7090.11	-7089.49
S&P 500	-5869.48	-5823.38	-5864.74	-5806.53	-5862.89
S&P CNX Nifty	-5424.15	-5359.13	-5392.22	-5309.28	-5402.54
Swiss Market	-5388.44	-5352.40	-5352.40	-5302.73	-5371.29
GARCHt(1,1) model					
AEX	-5955.88	-5962.06	-5962.31	-5961.00	-5962.50
Bovespa	-8014.23	-8023.98	-8024.50	-8023.33	-8024.25
CAC 40	-6440.42	-6442.56	-6442.90	<i>-6438.71</i>	-6442.26
DAX Index	-6662.80	-6663.68	-6663.67	-6662.27	-6663.08
DJIA	-5662.54	-5665.04	-5666.15	-5662.83	-5666.21
Euro STOXX 50	-6797.72	-6800.81	-6801.00	-6798.41	-6800.85
FTSE 100	-5063.69	-5059.69	-5059.82	-5061.59	-5058.24
FTSE MIB	-5277.62	-5274.50	-5277.93	-5278.97	-5277.89
FT Straits Times	-4183.72	-4172.26	-4179.86	-4180.24	-4179.91
Hang Seng	-5199.29	-5197.19	-5196.87	-5192.39	-5197.31
IBEX 35	-6696.76	-6702.99	-6702.93	-6700.94	-6702.05
IPC Mexico	-6641.36	-6638.73	<i>-6639.63</i>	-6638.53	-6639.37
KOSPI Composite	-5978.88	-5975.04	-5971.42	-5977.62	-5970.35
Nasdaq 100	-6454.09	-6454.63	-6455.97	-6456.62	<i>-6452.38</i>
Nikkei 225	-5943.50	-5951.89	-5947.73	-5945.44	-5948.38
Russel 2000	-7083.82	-7088.07	-7087.90	-7086.18	-7086.82
S&P 500	-5791.20	-5793.55	-5794.74	-5791.73	-5794.31
S&P CNX Nifty	-5276.30	-5299.25	-5291.16	-5293.34	-5294.78
Swiss Market	-5277.38	-5274.68	-5277.99	-5278.76	-5277.88

Note: **Bold** and *italic* values in each specification of error distribution indicate acceptance of proposed models, respectively, at the **5%** and *10%* significance levels of LRT.

Table 6: Model comparison via DIC

Data	GARCHn(1,1) model				
	Transformation for returns				
	NoT	ET	Exp	Mod	YJ
AEX	12074.43	12010.79 ₂	12069.59 ₄	11948.57 ₁	12068.13 ₃
Bovespa	16092.90	16078.63 ₂	16088.65 ₃	16054.94 ₁	16089.96 ₄
CAC 40	13028.84	12953.76 ₂	13029.16	12897.44 ₁	13028.34
DAX Index	13454.93	13386.07 ₂	13449.64 ₄	13350.47 ₁	13448.62 ₃
DJIA	11471.38	11385.16 ₂	11467.89 ₃	11350.90 ₁	11465.73 ₂
Euro STOXX 50	13738.61	13663.75 ₂	13733.73 ₃	13618.22 ₁	13735.07 ₄
FTSE 100	10215.13	10189.03 ₂	10206.00 ₄	10149.50 ₁	10200.77 ₃
FTSE MIB	10780.07	10708.29 ₂	10736.15 ₃	10609.09 ₁	10746.98 ₄
FT Straits Times	8444.39	8432.68 ₂	8447.34	8394.85 ₁	8447.21
Hang Seng	10506.40	10457.60 ₂	10502.24 ₃	10406.18 ₁	10503.79 ₄
IBEX 35	13557.03	13491.14 ₂	13557.34	13427.62 ₁	13553.96 ₃
IPC Mexico	13433.71	13370.24 ₂	13430.16 ₄	13309.23 ₁	13429.53 ₃
KOSPI Composite	12067.69	12035.21 ₄	12030.34 ₃	11992.32 ₁	12029.95 ₂
Nasdaq 100	12974.20	12959.61 ₂	12971.58 ₄	12940.78 ₁	12961.00 ₃
Nikkei 225	12190.41	12045.99 ₂	12092.80 ₃	11935.56 ₁	12113.92 ₄
Russel 2000	14189.85	14190.28	14184.60 ₃	14183.35 ₁	14184.14 ₂
S&P 500	11741.75	11650.30 ₂	11733.69 ₄	11617.17 ₁	11729.27 ₃
S&P CNX Nifty	10851.58	10722.54 ₂	10789.00 ₃	10623.04 ₁	10808.79 ₄
Swiss Market	10779.23	10709.07 ₂	10736.82 ₃	10608.05 ₁	10746.75 ₄
	GARCHt(1,1) model				
AEX	11915.62	11929.25	11928.92	11926.94	11929.99
Bovespa	16032.38	16052.55	16054.08	16051.19	16052.98
CAC 40	12884.29	12889.92	12891.20	12882.46₁	12888.76
DAX Index	13329.64	13332.33	13332.24	13329.90	13331.01
DJIA	11329.57	11335.03	11337.16	11330.89	11337.28
Euro STOXX 50	13599.02	13606.30	13606.62	13602.10	13606.18
FTSE 100	10131.39	10124.01 ₂	10124.42 ₃	10127.97 ₄	10121.12₁
FTSE MIB	10559.40	10553.70₁	10560.52	10563.13	10560.68
FT Straits Times	8371.49	8349.56₁	8364.61 ₃	8365.84 ₄	8364.21 ₂
Hang Seng	10403.15	10399.66 ₃	10398.29 ₂	10389.83₁	10400.34 ₄
IBEX 35	13397.41	13410.97	13410.49	13407.30	13408.94
IPC Mexico	13286.50	13282.27 ₂	13284.60 ₄	13281.72₁	13283.46 ₃
KOSPI Composite	11961.57	11954.94 ₃	11947.81 ₂	11960.30 ₄	11945.14₁
Nasdaq 100	12912.25	12914.44	12916.43	12918.33	12909.43₁
Nikkei 225	11891.35	11908.62	11900.38	11895.88	11901.17
Russel 2000	14174.79	14186.47	14182.63	14180.51	14181.75
S&P 500	11587.11	11592.38	11595.01	11588.78	11593.78
S&P CNX Nifty	10556.84	10603.34	10587.10	10591.71	10594.61
Swiss Market	10558.43	10554.42₁	10560.79	10562.37	10560.57

Note: Subscript indicates the rank of the model in each specification of error distribution. **Bold** value indicates the **best model** in all specifications.

Table 7: Model comparison via BIC

Data	GARCHn(1,1) model				
	Transformation for returns				
	NoT	ET	Exp	Mod	YJ
AEX	12093.38	12036.31 ₂	12094.60	11974.53 ₁	12093.92
Bovespa	16111.77	16104.40 ₂	16114.74	16084.02 ₁	16115.88
CAC 40	13048.14	12978.99 ₂	13054.46	12923.23 ₁	13053.96
DAX Index	13474.10	13412.67 ₂	13474.99	13376.03 ₁	13473.69
DJIA	11490.66	11411.19 ₂	11492.96	11376.38 ₁	11491.87
Euro STOXX 50	13758.83	13689.83 ₂	13760.20	13643.87 ₁	13760.92
FTSE 100	10234.46	10213.77 ₂	10231.93 ₄	10175.86 ₁	10226.58 ₃
FTSE MIB	10799.04	10733.83 ₂	10761.53 ₃	10634.84 ₁	10772.37 ₄
FT Straits Times	8464.03	8457.69 ₂	8472.27	8420.22 ₁	8472.26
Hang Seng	10523.50	10482.61 ₂	10528.67	10430.86 ₁	10528.11
IBEX 35	13576.91	13517.49 ₂	13583.97	13452.38 ₁	13579.50
IPC Mexico	13451.74	13395.53 ₂	13455.90	13333.74 ₁	13455.73
KOSPI Composite	12085.59	12061.57 ₂	12056.02 ₄	12017.69 ₁	12055.92 ₃
Nasdaq 100	12991.91	12984.77 ₂	12997.92	12966.65 ₁	12987.11 ₃
Nikkei 225	12208.82	12071.50 ₂	12118.17 ₃	11960.69 ₁	12139.69 ₄
Russel 2000	14208.87	14215.38	14209.99	14209.73	14208.52 ₁
S&P 500	11761.36	11676.65 ₂	11758.89 ₄	11642.55 ₁	11755.92 ₃
S&P CNX Nifty	10869.79	10747.33 ₂	10812.93 ₃	10647.18 ₁	10834.41 ₄
Swiss Market	10799.02	10733.89 ₂	10761.57 ₃	10634.87 ₁	10772.32 ₄
	GARCHt(1,1) model				
AEX	11941.56	11961.12	11962.40	11959.16	11962.11
Bovespa	16058.06	16085.10	16085.66	16079.64	16086.03
CAC 40	12911.01	12922.25	12922.50	12914.50	12922.47
DAX Index	13355.22	13364.46	13364.51	13361.26	13363.44
DJIA	11354.17	11367.15	11369.31	11362.35	11369.51
Euro STOXX 50	13625.48	13638.96	13639.39	13633.57	13639.34
FTSE 100	10156.99	10156.84	10156.93	10160.28	10153.87 ₁
FTSE MIB	10586.27	10584.67 ₁	10593.16	10594.69	10592.90
FT Straits Times	8396.36	8380.78 ₁	8396.12	8396.34	8396.68
Hang Seng	10427.19	10430.79	10430.76	10421.42 ₁	10430.55
IBEX 35	13423.27	13442.99	13443.43	13438.55	13441.30
IPC Mexico	13312.58	13314.77	13315.94	13314.27	13316.08
KOSPI Composite	11987.46	11987.14	11979.82 ₂	11992.11	11978.15 ₁
Nasdaq 100	12942.09	12946.11	12949.42	12950.06	12937.72 ₁
Nikkei 225	11916.10	11940.81	11932.40	11927.56	11934.12
Russel 2000	14201.31	14208.15	14205.55	14203.57	14204.41
S&P 500	11611.11	11623.90	11625.85	11620.18	11625.39
S&P CNX Nifty	10581.41	10634.93	10618.86	10623.02	10625.80
Swiss Market	10586.25	10584.70 ₁	10593.11	10594.78	10592.89

Note: Subscript indicates the rank of the model in each specification of error distribution. **Bold** value indicates the **best model** in all specifications.