Comparison of GARCH Family Models for Shanghai Stock Exchange Index Prediction

Review of Integrative Business & Economics Research

Jiawei Tong*
UNSW Business School, The University of New South Wales

Chenqi Wu
Farmer School of Business, Miami University - Oxford

Luxi Liang
School of Accounting, Nanjing Audit University

ABSTRACT

This paper verifies that stock markets in Mainland China have leverage effects based on various GARCH models. Through model comparison, we found that the EGARCH-t model well captures the log-return characteristics of the Shanghai Stock Exchange Index (SSEI). By studying the VaR of each model and implementing the Kupiec test, we discovered that the TGARCH model under the t-distribution is most effective in risk measurement. Moreover, the AIC and SC of the EGARCH model are lower than that of the TGARCH model while the Log-likelihood of the EGARCH model is higher than that of the TGARCH model; thus, we conclude that the EGARCH model is relatively more effective in fitting the log-return series of SSEI. In addition, the Beta-Skew-t-EGARCH model is able to simulate log-returns more comprehensively because it takes skewness and kurtosis into consideration. The actual failure rate of predicting the SSEI series using the TGARCH model is as low as 2.58%. The estimated VaR implies that domestic and overseas investors might suffer potential losses if they employ these models to predict the stock market direction and that they can optimize their investment portfolios by comparing the estimates and VaR from different GARCH models.

Keywords: Time series models; Shanghai Stock Exchange Index; Value at Risk; GARCH models.

Received 8 August 2019 | Revised 7 November 2019 | Accepted 6 January 2020.

1. INTRODUCTION

The Shanghai Stock Exchange Index (SSEI) has attracted great attention from domestic and international investors due to its enormous trading volumes and increasingly significant role in the global economy. Capital investors are able to enjoy greater economic benefits at lower risks if they are able to make more accurate and timely predictions about the long-term trends of SSEI.

Since the 1970s, past studies have been documenting shortcomings (e.g., in handling asymmetric periodic data) of linear time series models and proposing various kinds of non-linear time series models. For instance, H. Tong proposed the threshold

ISSN: 2304-1013 (Online); 2304-1269 (CDROM); 2414-6722 (Print)

autoregressive (TAR) model in 1978; Robert F. Engle proposed the Auto-Regressive Conditional Heteroskedasticity (ARCH) model in 1982, and; T. Bollerslev proposed the Generalized Auto-Regressive Conditional Heteroskedasticity (GARCH) model based on the ARCH model in 1986. The introduction of GARCH greatly inspired subsequent research on related models such as IGARCH, GJR, EGARCH, and Beta-t-EGARCH models, etc. GARCH models are also instructive for research practices. For instance, Muhammad Idrees Ahmad *et al.* (2013) used GARCH models to test and predict inflation volatility in Oman. These models were applied in economic and financial fields such as forecasting the prices of crude, gold, options, coal and other commodities according to their risk factors and values. Time series fitting and forecasting vary with the types of model used and underlying hypotheses.

This study analyzes the volatility and risks of SSEI by comparing different GARCH models, which provides statistically and economically useful methods and findings. The rest of this paper is divided into four sections as follows: Section 2 documents related theories as well as reviewing relevant results from past studies. Section 3 describes the methods of analysis used in this study, including the data, theories and estimation methods of different kinds of GARCH models, VaR, and Kupiec tests. Section 4 presents empirical results obtained from this study, including basic statistical analysis on the stationarity and ARCH effects of SSEI, estimation of the parameters of the MA-GARCH model under t-distribution, analysis of the fitting degree of different GARCH models, estimating the information impact curve of SSEI, and computation and comparing the effectiveness of VaR estimates using Kupiec tests. The findings indicate that the TGARCH-t model has the highest fitting degree. Chapter 5 summarizes the findings from this study.

2. RELATED WORK

China's stock markets commenced in 1978 and still have shortcomings after many years of development. For instance, corporations' share prices cannot accurately reflect their actual fundamental values because the shares prices are heavily affected by governmental policies, anthropic factors, and high turnover rates. With the development of the financial industry, experts and scholars have been using various theories to analyze and explore the stock market situations in China.

S. H. Yu and Z. T. Wang (2005) conducted an empirical study on China's stock markets by an ARCH model and found that the fluctuations of stock prices were not sensitive enough to related information released by firms. For example, they found that the returns to SSE stocks were consistent with ARCH effects but the kurtosis coefficient of the Shenzhen stock market was relatively high. These findings imply that China's stock markets are characterized by a large degree of speculation.

C. L. Yue *et al.* (2001) conducted a study on SSEI from September 23, 1997 to December 31, 1999 by GARCH, IGARCH and EGARCH models. The study summarized the conditional heteroscedasticity of stock returns in China's stock markets. Moreover, L. X. Wu and L. B. Xu (2002) adopted the stationary distribution theory to analyze characteristics of stock returns in China's stock market. Their empirical findings showed that the time series of China's stock returns are characterized by leptokurtic, heavy tails, and a certain degree of stationarity.

ISSN: 2304-1013 (Online); 2304-1269 (CDROM); 2414-6722 (Print)

Past studies generally adopted time series techniques for empirical analysis on stationarity, stochasticity, and fluctuations of SSEI. It was pointed out that using GARCH to forecast stock returns is merely a theoretical possibility. However, there has been scant research from the perspective of building a systematic GARCH model, obtaining a specifically fitted model to make predictions, and selecting the best model by model comparison.

3. METHODOLOGY

This section explains the research methodology used in this study. This study used data on stock indices published by Wind Financial Database to develop various GARCH models. EViews version 9 was the software used for implementing the estimations.

3.1 Data Input

The data used in this study contain historical daily closing SSEI from January 2, 2012 to January 9, 2019 (i.e., seven years). SSEI is used because it has practical research value and sufficiently represents the volatility and risks of stocks listed on SSE. Moreover, the data on SSEI are enormous and frequently updated.

Assume that P_t represents the closing price at time t, the log-return is expressed by $r_t=\ln P_t-\ln P_{t-1}$ that can be calculated in 1,942 groups of log-return.

3.2 Theory of GARCH (p, q) Model Development

3.2.1 GARCH (p, q) and IGARCH (p, q) Models

Bollerslev *et al.* (1986) generalized ARCH (p, q) model to GARCH (p, q) model. Consider the following GARCH (p, q)

$$\begin{cases} a_t = \delta_t \varepsilon_t \\ \sigma_t^2 = w + \beta_1 \sigma_{t-1}^2 + \cdots + \beta_q \sigma_{t-q}^2 + \alpha_1 a_{t-1}^2 + \cdots + \alpha_p a_{t-p}^2 \end{cases} \tag{1}$$

where σ_t is the conditional variance for the sequence $\{a_t\}$, ε_t is an independent and identically distributed random variable with a zero mean and a variance of 1 and satisfies w≥0, α_i≥0, β_i≥0, $\sum_{i=1}^{p} \alpha_i + \sum_{j}^{q} \beta_j < 1$. The requirements above ensure that the sequence $\{a_t\}$ is finite, i.e., the GARCH model is strictly stationary. When $\sum_{i=1}^{p} \alpha_i + \sum_{j}^{q} \beta_j = 1$, the model turns into an IGARCH (p, q) model.

3.2.2 EGARCH (p, q) Model

Nelson *et al.* (1991) proposed the EGARCH (p, q) model targeting to make up the weakness of the GARCH model in financial risk forecasting. The model considers weighted information which reflects the positive and negative asymmetries in the rate of return. The model can be expressed as:

$$ln(\sigma_t^2) = w + \sum_{i=1}^{q} \left(\alpha_i \left| \frac{\varepsilon_{t-i}}{\sigma_{t-i}} \right| + \gamma_i \frac{\varepsilon_{t-i}}{\sigma_{t-i}} \right) + \sum_{j=1}^{p} \beta_i \, ln\, \sigma_{t-j}^2 \tag{2}$$

If $\gamma \neq 0$, external shocks have asymmetric impacts on share price volatility. If $\gamma < 0$, negative external shocks outweigh positive external shocks in affecting stock price fluctuations, that is called leverage effects.

ISSN: 2304-1013 (Online); 2304-1269 (CDROM); 2414-6722 (Print)

3.2.3 TGARCH Model

Glosten *et al.* (1993) formulated a kind of expanded asymmetric GJR model and Zakoian proposed a similar TGARCH model, both are designated to deal with leverage effects of the sequence. The model could be interpreted as:

$$\delta_t^2 = w + \sum_{i=1}^{q}(\alpha_i + \gamma_i N_{t-i})\alpha_{t-i}^2 + \sum_{j=1}^{p}\beta_i \delta_{t-j}^2 \quad (3)$$

subject to $$N_{t-i} = \begin{cases} 1, \alpha_{t-i} < 0 \\ 0, \alpha_{t-i} \geq 0 \end{cases} \quad (4)$$

where parameters α_i, γ_i, β_j are nonnegative and satisfy the same restrictions as those in the GARCH model.

3.2.4 Beta-Skew-t-EGARCH Model

Harvey and Sucarrat *et al.* (2011) discussed some characteristics of the EGARCH model and further generalized the EGARCH model to the Beta-Skew-t-EGARCH model. As it turned out, the Beta-Skew-t-EGARCH model more effectively fitted the financial market after considering the characteristics of leverage effects, heavy tails, skewness and other factors. This model has two forms that are known as the single-component form and double-component form. This study mainly employs the single-component form which can be described as:

$$\begin{cases} a_t = exp(\lambda_t)\, \varepsilon_t = \delta_t \varepsilon_t \sim st(0, \sigma_\varepsilon^2, \upsilon, \gamma) \\ \lambda_t = w + \lambda_{t-1}^{+} \\ \lambda_t^{+} = \alpha_1 \lambda_{t-1}^{+} + \beta_1 u_{t-1} + k^* sgn(-y_{t-1})(1 + u_{t-1}), |\alpha_1| < 1 \end{cases} \quad (5)$$

where, $st(0, \sigma_\varepsilon^2, \upsilon, \gamma)$ denotes the Skewed-t distribution, δ_t is the volatility rate, $\varepsilon_t = \varepsilon_t^* - u_\varepsilon$, σ_ε is the standard deviation of ε_t, υ is the degree of freedom of the model, γ is the coefficient of the skewness. In addition, ε_t^* obeys the distribution with mean u_ε, degree of freedom υ, and skewness γ. When $\gamma > 1$, ε_t is a right-skewed t-distribution; when $\gamma < 1$, ε_t is a left-skewed t-distribution. In this equation, w is the intercept, α_1 is the persistence, β_1 is the ARCH parameter, k^* is the leverage parameter, and u_{t-1} is the conditional scoring item. The greater the absolute value of β_1, the more significant the fluctuations of the shock of volatility.

3.3 ARCH Effective Test

To test for ARCH effects, the Lagrange Multiple test (LM test) is used to determine whether there is a set of time series satisfying ARCH effects. The approach is to establish an *n*-order autoregressive model for the squared sequence, r_t^2, of the time series $\{r_t\}$, which can be expressed as:

$$r_t^2 = \alpha_0 + \alpha_1 r_{t-1}^2 + \cdots + \alpha_n r_{t-n}^2 + e_t \quad (6)$$

where, e_t is an independently and identically distributed random variable with a zero mean and a variance of one. Under the null hypothesis that there are no n-order ARCH effects in the sequence, the test statistics is Lagrange Multiples:

$$LM \approx N \cdot R^2 \sim \chi^2(n) \quad (7)$$

ISSN: 2304-1013 (Online); 2304-1269 (CDROM); 2414-6722 (Print)

In the above expression, N is the length of the sequence, and R^2 is the R-squared figure of Equation (6). Given a significance level of α, we reject the original hypothesis if the statistics calculated satisfies $LM > \chi^2_{1-\alpha}(n)$ or $\mathrm{P-value} < \alpha$. We believe that the time series has n-order ARCH effects, otherwise the original hypothesis should be rejected.

3.4 Akaike Information Criterion and Bayesian Information Criterion

The Akaike Information Criterion (AIC) is created by Akaike, a Japanese statistician, in a study on determining the order of a time series model. It is widely used for evaluating the effectiveness of a statistical model in fitting the data. The AIC can be defined as follows:

$$AIC = -2\,log(L) + 2k \tag{8}$$

where, L is the maximum likelihood functions of the model and k represents the number of independent parameters of the model. The model is optimized when the AIC value is minimized.

Hurvich and Tsai (1989) optimized the AIC criterion by adding a nonrandom penalty term, which is defined as:

$$AIC_c = AIC + \frac{2(k+1)(k+2)}{n-k-2} \tag{9}$$

In Equation (9), k is the number of all parameters except the noise variance; n denotes the sample size.

Bayesian Information Criterion, also known as BIC, offsets some defects of AIC and improves its functions. The expression is as follows:

$$BIC = -2\,log(L) + k\,log(n) \tag{10}$$

where, L represents the maximum likelihood function, k is the total number of the model's parameters, n is the total sample size, which is the same as that in AIC.

In real-life questions, AIC and BIC can be applied in the following ways:

(1) Select suitable models based on real-life questions;

(2) Use maximum likelihood estimation (MLE) method to estimate the parameters of the model;

(3) Choose the ultimate model with the lowest values of AIC and BIC.

3.5 VaR Theory and Computation Effects Test

VaR stands for value at risk proposed by JP Morgan in the 1990s and is widely applied to risk evaluation nowadays. VaR is mainly applied to estimate the maximum possible loss of a portfolio during a given continuous period under a normal market situation and specific confidence criteria. The expression is given as follows:

$$P(\Delta p \leq VaR) = \alpha \tag{11}$$

In the above expression, Δp denotes the loss suffered during a particularly continuous period; α denotes a specific confidence criterion.

Jorion *et al.* (1996) developed a computation of VaR denoted as $VaR = E(P) - P^*$. In this equation, $E(P)$ is the expectation of the portfolio, P is the closing price of the assets, and P^* represents the closing price of the assets under α

ISSN: 2304-1013 (Online); 2304-1269 (CDROM); 2414-6722 (Print)

conditions. The computation of VaR can be deduced on the following basis:

$$VaR = P_0[E(R) - R^*] \tag{12}$$

In the above expression, P_0 is the opening price of the asset, R^* is the rate of return based on the closing price of the asset under α conditions, R is the rate of return based on the closing price of the asset, $E(R)$ is the expectation of the rate of return or the rate of return of the portfolio, σ_t is the volatility rate of the portfolio.

To summarize, the computation of VaR under normal distributions is as follows:

$$VaR = P_0 \cdot z_\alpha \cdot \sigma_t \cdot \sqrt{\Delta t} \tag{13}$$

where, P_0 denotes the opening price of the assets, z_α denotes the α-quantile under the specific distribution, σ_t denotes the volatility rate and Δt denotes the holding period of the assets. In terms of the computation of VaR given above, we should calculate the volatility rate σ_t first if we want to calculate the value of VaR, and then implement an accuracy test on the two values. The Kupiec test indicates that this calculation is a failure event if the test result is larger than the value of VaR, otherwise the calculation is a successful event.

When VaR equals α, the corresponding expectation of the possibility of a failure event is P^* and the whole curve approximately follows the Bernoulli distribution, i.e., $P^* = 1 - \alpha$. When the test period is T days and that of the failure event is N days, the failure rate can be written as $p = \frac{N}{T}$. As such, the accuracy test can be conducted to check whether P^* keeps in pace with P.

Kupiec also proposed the concept of likelihood ratio test for the null hypothesis $P^* = P$, which is defined as follows:

$$LR = 2\,ln[(1-P)^{T-N}P^N] - 2\,ln[(1-P^*)^{T-N}P^{*N}] \tag{14}$$

Under the null hypothesis, $P^* = P$. The statistic LR has a Chi-square distribution with the degree of freedom equal to 1.

4. EXPERIMENTAL RESULTS

4.1 Statistical Analysis on the Log-Return Series of SSEI

4.1.1 A Normal Distribution Test for the Log-Return Series

We used the statistical software EViews9 to implement a basic statistical analysis on the log-return series of SSEI. The first step is to analyze the data which is presented in Figure 1.

(1) The skewness is -1.056439. The data as shown in Figure 1 skews to the left as a whole. As we know, the skewness is 0 for a normal distribution and thus the null hypothesis is rejected. In conclusion, the log-return series is not in a normal distribution and there exists a heavy tail.

(2) The kurtosis is 9.927555, which is much larger than 3 as in the normal distribution. The log-return series therefore has leptokurtic features.

(3) The results of a Jarque-Bera (JB) test on the log-return series is 3730.881 (i.e.,

ISSN: 2304-1013 (Online); 2304-1269 (CDROM); 2414-6722 (Print)

the JB statistics) with a probability equal to 0.000000. However, in a Chi-square distribution, the degree of freedom = 2 has a critical value of 5.991465 at the 5% level of significance. The null hypothesis is also rejected because the JB statistics is much larger than the critical value. In conclusion, the log-return series is not in a Chi-square distribution.

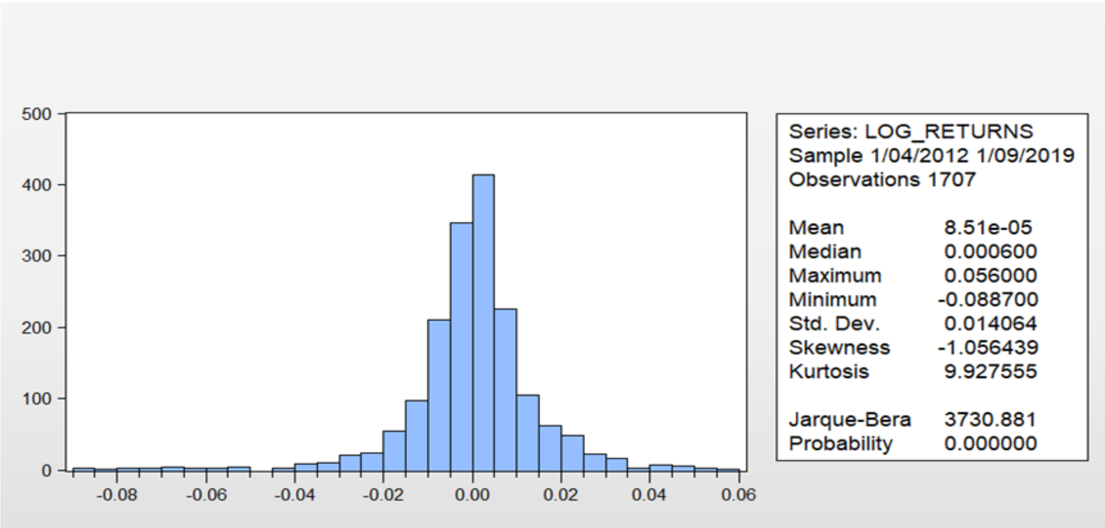


FIGURE 1: Histogram of the log-return of SSEI

Figure 2 displays the time series of the log-returns of SSEI. From the figure we can conclude that the volatility of the time series tends to be intensive initially and then slow-down over time. As a whole, the time series is characterized by volatility clustering, leptokurtic, and heavy tail.

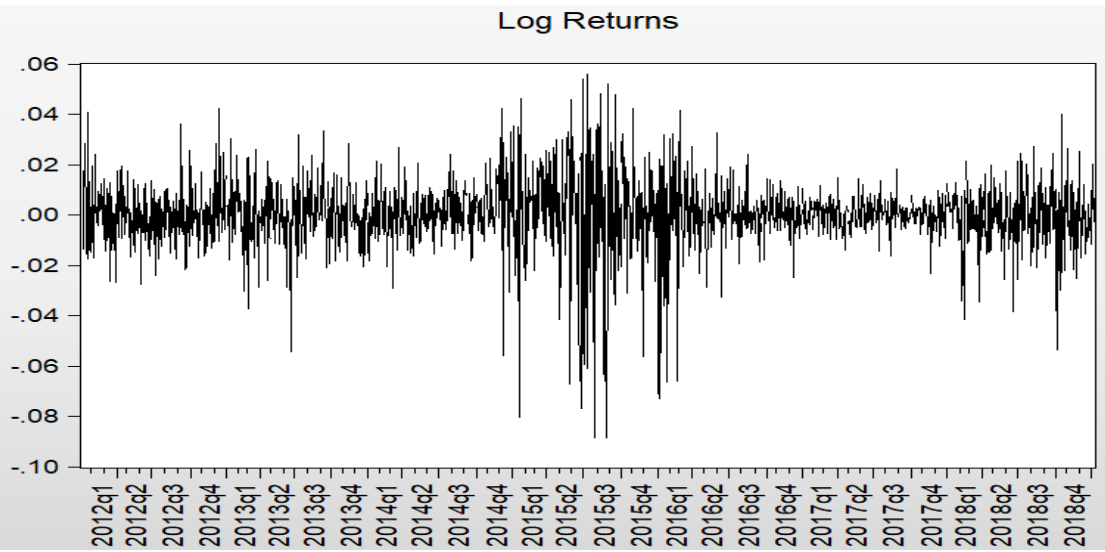


FIGURE 2: Graph of the time series of log-return of SSEI

According to the Quantile-Quantile (Q-Q) Plot of the log-return time series as presented in Figure 3, there are quantities of points outside the normal (red) line as well as a swing at the lower part of the line, which indicates that the log-return series has the

ISSN: 2304-1013 (Online); 2304-1269 (CDROM); 2414-6722 (Print)

heavy tail characteristic.

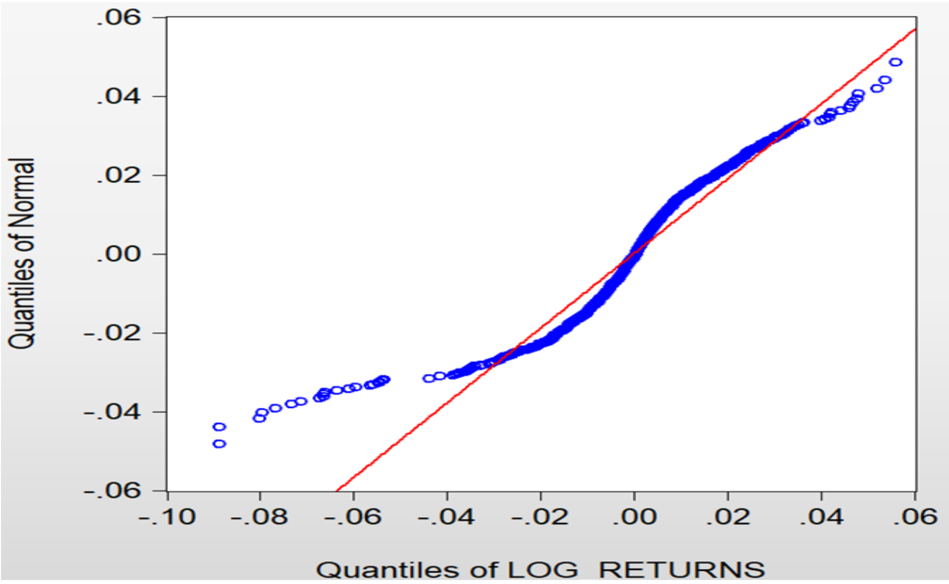


FIGURE 3: Q-Q Plot for the time series of log-return of SSEI

4.1.2 Stationarity Test for the Log-Return Series

Implementing the Phillips-Perron (P-P) unit root test on the time series using EViews9 gives the results as shown in Table 1.

Phillips-Perron Unit Root Test		
Null Hypothesis: D (Log-returns) has a unit root		
	Adj. t-Stat	Prob.*
Phillips-Perron test statistic	-420.6198	0.0001
Test critical values:		
1% level	-3.963434	
5% level	-3.412447	
10% level	-3.128171	

TABLE 1: Results of the P-P unit root test for the log-return of SSEI

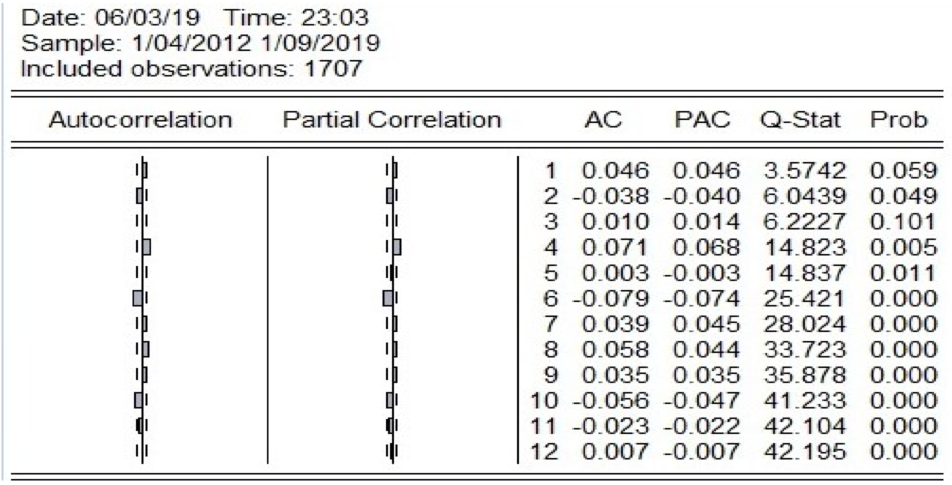
In terms of the results from Table 1, we find that the corresponding critical values of the test statistics at the levels of 1%, 5% and 10% are all larger than the P-P statistics. In other words, the null hypothesis of a unit root is rejected. In other words, the time series is stationary.

4.1.3 Autocorrelation Test

We used EViews9 to test for the existence of autocorrelation in the time series and set the lag operator equal to 12. From Figure 4, the time series is stationary with autocorrelation, especially at the sixth order where the correlation is significant. Hence, we established the mean square equation with AR (1), AR (2), AR (3), AR (4), AR (5),

ISSN: 2304-1013 (Online); 2304-1269 (CDROM); 2414-6722 (Print)

AR (6), MA (1), MA (2), MA (3), MA (4), MA (5) and MA (6).


Date: 06/03/19 Time: 23:03
Sample: 1/04/2012 1/09/2019
Included observations: 1707

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.046	0.046	3.5742	0.059
		2	-0.038	-0.040	6.0439	0.049
		3	0.010	0.014	6.2227	0.101
		4	0.071	0.068	14.823	0.005
		5	0.003	-0.003	14.837	0.011
		6	-0.079	-0.074	25.421	0.000
		7	0.039	0.045	28.024	0.000
		8	0.058	0.044	33.723	0.000
		9	0.035	0.035	35.878	0.000
		10	-0.056	-0.047	41.233	0.000
		11	-0.023	-0.022	42.104	0.000
		12	0.007	-0.007	42.195	0.000

FIGURE 4: Graph of the correlogram of the log-return series of SSEI

First, we tried to determine the order of each model. The experimental results are shown in Figures 5.1 and 5.2. According to the adjoint probability presented in the figures, we find that AR (1), AR (4), AR (6), MA (1), MA (4) and MA (6) are all highly effective.

Variable	Coefficient	Std. Error	t-Statistic	Prob.
AR (1)	0.047132	0.014987	3.144821	0.0017
AR (2)	-0.032825	0.013980	-2.347959	0.0190
AR (3)	0.011505	0.013542	0.849562	0.3957
AR (4)	0.066069	0.014577	4.532283	0.0000
AR (5)	0.000645	0.015366	0.041983	0.9665
AR (6)	-0.074367	0.017256	-4.309650	0.0000

FIGURE 5.1: Results of significance test for AR models

Variable	Coefficient	Std. Error	t-Statistic	Prob.
MA (1)	0.055003	0.014895	3.692585	0.0002
MA (2)	-0.025329	0.014044	-1.803513	0.0715
MA (3)	0.008487	0.014091	0.602284	0.5471
MA (4)	0.056597	0.015000	3.773178	0.0002
MA (5)	0.001628	0.015399	0.105705	0.9158
MA (6)	-0.079018	0.016912	-4.672312	0.0000

FIGURE 5.2: Results of significance test for MA models

Based on the above-mentioned results, we built various ARMA models and discovered that ARMA (4,4), ARMA (4,6) and ARMA (6,4) are more effective than the

ISSN: 2304-1013 (Online); 2304-1269 (CDROM); 2414-6722 (Print)

other models. Comparing the values of AIC and SC, we ultimately found that ARMA (4,6) is the optimal model. The following Tables 2.1, 2.2 and 2.3 are the test results of the three models.

Variable	Coefficient	Std. Error	t-Statistic	Prob.
AR (4)	0.0773318	0.077160	10.02223	0.0000
MA (4)	-0.710318	0.086288	-8.231956	0.0000
AIC			-5.697354	
SC			-5.687789	

Table 2.1: Results of significance test for ARMA (4,4)

Variable	Coefficient	Std. Error	t-Statistic	Prob.
AR (4)	0.065559	0.014240	4.603741	0.0000
MA (6)	-0.073750	0.016893	-4.365611	0.0000
AIC			-5.697915	
SC			-5.688350	

Table 2.2: Results of significance test for ARMA (4,6)

Variable	Coefficient	Std. Error	t-Statistic	Prob.
AR (6)	-0.074416	0.016994	-4.379006	0.0000
MA (4)	0.059433	0.014609	4.068396	0.0000
AIC			-5.697534	
SC			-5.687969	

Table 2.3: Results of significance test for ARMA (4,6)

According to results as shown in Table 2, the coefficients of three models are significant, the residuals are not correlated, and the values of the AIC and SC are small, which indicates that all of the above three models fit the data well. Additionally, the AIC and SC of ARMA (4,6) are -5.697915 and -5.688350, which are the lowest among these models.

4.1.4 ARCH Effects Tests

After creating the mean square equation of ARMA (4,6), the residual series could be formed. In accordance with the fluctuations as shown in Figure 6.1, the volatility clustering of the time series is obvious, thus ARMA (4,6) has significant ARCH effects.

ISSN: 2304-1013 (Online); 2304-1269 (CDROM); 2414-6722 (Print)

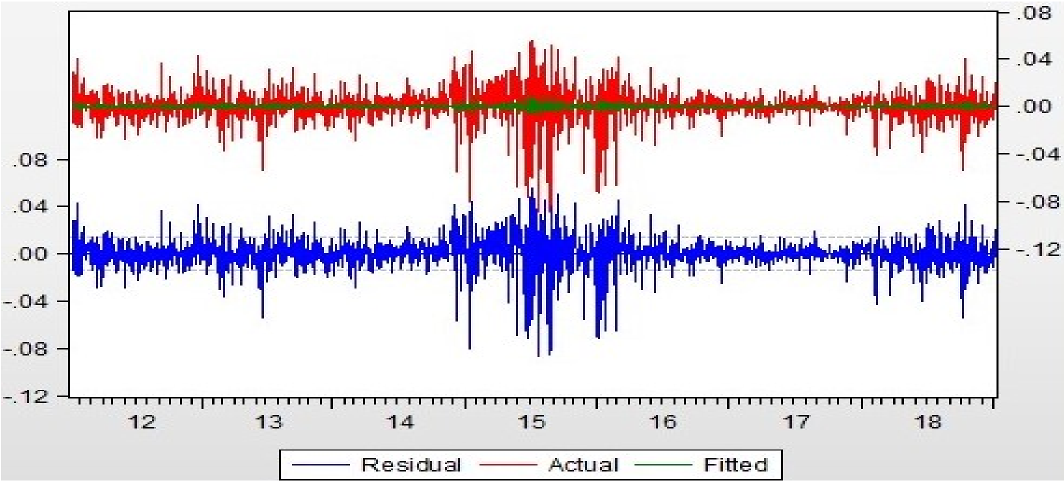
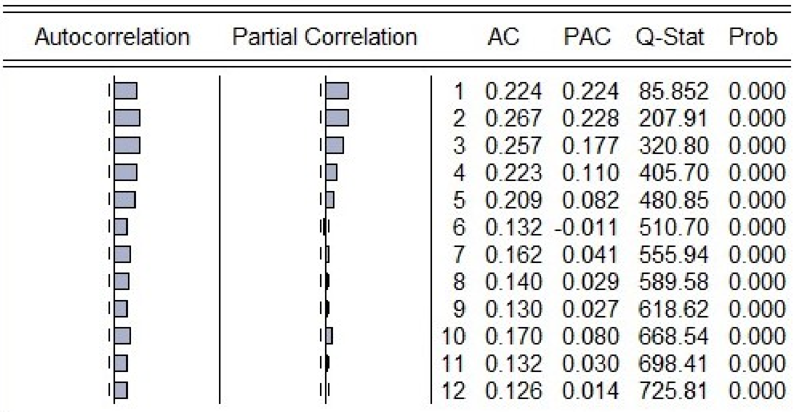


FIGURE 6.1: Graph of residuals of the log-return series

We then implemented the ARCH-LM test as shown below. Based on Figure 6.2 showing the correlogram of the squared residuals, we found that the partial correlation coefficient is truncating in the fifth order. Thus, we set the lag operator equal to 5.



Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.224	0.224	85.852	0.000
		2	0.267	0.228	207.91	0.000
		3	0.257	0.177	320.80	0.000
		4	0.223	0.110	405.70	0.000
		5	0.209	0.082	480.85	0.000
		6	0.132	-0.011	510.70	0.000
		7	0.162	0.041	555.94	0.000
		8	0.140	0.029	589.58	0.000
		9	0.130	0.027	618.62	0.000
		10	0.170	0.080	668.54	0.000
		11	0.132	0.030	698.41	0.000
		12	0.126	0.014	725.81	0.000

FIGURE 6.2: Graph of the correlogram of squared residuals

The results of the ARCH-LM test are shown in Table 3. As the table presented, the probabilities of the F-statistics and the Obvious*R-squared are both 0.0000. In addition, the values of the two statistics are 57.35580 and 246.1686, respectively, which are much larger than the critical value. Hence, the null hypothesis is rejected, i.e., the ARCH effects in the time series are significant and the GARCH can be used for modelling.

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	7.81E-05	1.46E-05	5.331066	0.0000
RESID^2 (-1)	0.103304	0.024193	4.270012	0.0000
RESID^2 (-2)	0.162911	0.024191	6.734340	0.0000
RESID^2 (-3)	0.148217	0.024246	6.113072	0.0000

ISSN: 2304-1013 (Online); 2304-1269 (CDROM); 2414-6722 (Print)

RESID^2 (-4)	0.100949	0.024192	4.172762	0.0000
RESID^2 (-5)	0.082378	0.024185	3.406128	0.0007
Heteroskedasticity Test: ARCH				
F-statistic	57.35580	Prob (F-statistic)		0.0000
Obs*R-squared	246.1686	Prob. Chi-Square		0.0000

TABLE 3: Results of ARCH-LM test (Heteroskedasticity test: ARCH)

4.2 Estimation and Analysis of the GARCH Models' Parameters

We used EViews9 to estimate the parameters of the first three models and used the betategarch package of R Software to estimate the parameters of the fourth model (i.e., the Beta-Skew-t-EGARCH model). The method is explained in detail in the following sections.

4.2.1 Creating MA (6)-GARCH (1,1)-t Model

Table 4 presents the estimation results for the model parameters, where υ is the degree of freedom. As a whole, the estimated parameters conform to the application conditions of the model. Besides, ω is significant at the 5% level and other parameters are all significant at 1% level. Moreover, the AIC and SC are -6.194115 and -6.178173, respectively, which indicates that the model is well-fitted. Furthermore, $\alpha_1 + \beta_1$=0.998056<1 suggesting that the model is stationary.

Variable	Coefficient	Std. Error	z-Statistic	Prob.
MA (6)	-0.054200	0.023501	-2.306261	0.0211
ω	9.38E-07	4.09E-07	2.293277	0.0218
α_1	0.060220	0.011390	5.287097	0.0000
β_1	0.937836	0.009883	94.89864	0.0000
υ	4.356378	0.531822	8.191426	0.0000
AIC			-6.194115	
SC			-6.178173	

TABLE 4: Estimation results of parameters of MA (6)-GARCH (1,1)-t model

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.970248	0.059474	16.31390	0.0000
WGT_RESID^2 (-1)	-0.030491	0.024234	-1.258220	0.2085
WGT_RESID^2 (-2)	0.021248	0.024235	0.876743	0.3807
Heteroskedasticity Test: ARCH				
F-statistic	1.211497	Prob (F-statistic)		0.298008
Obs*R-squared	2.423813	Prob. Chi-Square		0.2976

TABLE 5: Results of ARCH LM test for MA (6)-GARCH (1,1)-t model

ISSN: 2304-1013 (Online); 2304-1269 (CDROM); 2414-6722 (Print)

Afterwards, we tested for the ARCH effects in the model. Detailed results are presented in Table 5. From Table 5 with the lag operator equal to 2, we can conclude that the probabilities of the F-statistics and Obs*R-squared are 0.2980 and 0.2976, respectively, which are larger than the critical level of 0.05. In addition, the values of the two statistics are both smaller than the given critical value of 5.991465. Thus, we accept the null hypothesis that the ARCH effects have been eliminated and it is reasonable to use MA (6)-GARCH (1,1) for forecasting.

4.2.2 Establishing ARMA (3,3)-EGARCH (1,1)-t Model

As shown in Table 6.1, majority of the parameters are significant at the 1% level and the leverage coefficient γ equals -0.017641<0, which means that there are leverage effects in the time series of SSEI and that the fluctuations caused by negative news outweigh those by positive news.

In addition, $\alpha_1+\beta_1$=0.143599+0.992045=1.135644>1, indicating that the current volatilities cause continuous influences towards subsequent predictions. The AIC and SC are -6.196461 and -6.174142, respectively, suggesting that the model is well-fitted.

Variable	Coefficient	Std. Error	z-Statistic	Prob.
AR (3)	-0.785902	0.149264	-5.265162	0.0000
MA (3)	0.818362	0.138324	5.916262	0.0000
ω	-0.174041	0.037496	-4.641549	0.0000
α_1	0.143599	0.022210	6.465666	0.0000
γ	-0.017641	0.013421	-1.314481	0.1887
β_1	0.992045	0.003466	286.1950	0.0000
υ	4.406151	0.541142	8.142316	0.0000
AIC			-6.196461	
SC			-6.174142	

TABLE 6.1: Estimation results of parameters of ARMA (3,3)-EGARCH (1,1)-t model

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.962729	0.059507	16.17840	0.0000
WGT_RESID^2 (-1)	-0.029895	0.024228	-1.233892	0.2174
WGT_RESID^2 (-2)	0.030326	0.024230	1.251617	0.2109
Heteroskedasticity Test: ARCH				
F-statistic	1.593749	Prob (F-statistic)		0.203466
Obs*R-squared	3.187148	Prob. Chi-Square		0.2032

TABLE 6.2: Results of ARCH LM test for ARMA (3,3)-EGARCH (1,1)-t model

We then conducted ARCH LM tests on the model with the lag operator equal to 2. The results are shown in Table 6.2. The F-statistics and Obs*R-squared are 1.593749

ISSN: 2304-1013 (Online); 2304-1269 (CDROM); 2414-6722 (Print)

and 3.187148, which are both smaller than the given critical value of 5.991465. The corresponding probabilities of F-statistics and Chi-Square are 0.2035 and 0.2032 respectively, which means that the null hypothesis is accepted. That is to say, the ARMA (3,3)-EGARCH (1,1)-t model has no ARCH effects and can be used for prediction.

4.2.3 Creating ARMA (3,3)-TGARCH (1,1)-t Model

As presented in Table 7.1, all of the estimated parameters satisfy the requirements of the model. All of the estimated parameters are significant at the 1% level except for the leverage coefficient. The values of the AIC and SC are -6.192227 and -6.169908, which indicates that the model is effective. However, the leverage coefficient, γ, is larger than 0 and is significant at the 10% level, which means that it is reasonable to believe that the SSEI series has leverage effects.

Variable	Coefficient	Std. Error	z-Statistic	Prob.
AR (3)	-0.758254	0.161242	-4.702573	0.0000
MA (3)	0.794105	0.150075	5.291390	0.0000
ω	1.03E-06	4.27E-07	2.420603	0.0155
α_1	0.054054	0.013903	3.888054	0.0001
γ	0.014626	0.016839	0.868562	0.3851
β_1	0.992045	0.003466	286.1950	0.0000
υ	4.360087	0.532778	8.183680	0.0000
AIC			-6.192227	
SC			-6.169908	

TABLE 7.1: Estimation results of parameters of ARMA (3,3)-TGARCH (1,1)-t model

Then we again adopted the ARCH LM tests to examine the model with the lag operator equal to 2. As Table 7.2 shows, the F-statistics and Obs*R-squared are 1.245167 and 2.491080, which are both smaller than the given critical value of 5.991465. The corresponding probabilities of the F-statistics and Chi-Square are 0.2882 and 0.2878, respectively, so we accept the null hypothesis that the ARMA (3,3)-TGARCH (1,1)-t model has no ARCH effects and is effective in forecasting.

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.974225	0.059442	16.38959	0.0000
WGT_RESID^2 (-1)	-0.032343	0.024235	-1.334572	0.1822
WGT_RESID^2 (-2)	0.019331	0.024236	0.797619	0.4252
Heteroskedasticity Test: ARCH				
F-statistic	1.245167	Prob (F-statistic)		0.288155
Obs*R-squared	2.491080	Prob. Chi-Square		0.2878

TABLE 7.2: Results of ARCH LM test for ARMA (3,3)-TGARCH (1,1)-t model

ISSN: 2304-1013 (Online); 2304-1269 (CDROM); 2414-6722 (Print)

4.2.4 Beta-Skew-t-EGARCH (1,1) Model

We adopted the betategarch-package of R Software to implement the parameter estimation. The estimation results presented in Table 8 show that all of the estimated parameters are significant at the 1% level. Note that df is the degree of freedom. $kappal \neq 0$ represents the leverage coefficient, which means that the SSEI series has leverage effects. $skew = 1.00347902 > 1$ indicates that the distribution is right-skewed rather than left-skewed.

Moreover, $phli + kappastar =$ 0.96055416+0.001710015 = 0.96226417 is smaller than 1, suggesting that the time series is stationary. In conclusion, the model is capable of predicting the time series.

4.2.5 Comparison of Different Models

In all of the models we have formulated in previous sections, the estimated parameters are all significant, so it is reasonable to believe that these models are effective. The asymmetric coefficients of the latter three asymmetric GARCH models are also significant, which indicates that the SSEI series has leverage effects. In fact, stock markets are known to have leverage effects, and thus asymmetric models should be more effective than a simple GARCH model in predicting the series. Moreover, the asymmetric coefficient of the EGARCH model is more sensitive and significant than that of the TGARCH model. Furthermore, the AIC and SC of the EGARCH model are smaller than those of the TGARCH model while the value of Log-likelihood of the EGARCH model is larger than that TGARCH model, thus we conclude that the EGARCH model is relatively more effective in fitting the log-returns of SSEI. The Beta-Skew-t-EGARCH model simulates the log-returns more comprehensively because this model takes the skewness and kurtosis into consideration.

```
Date: Wed Jun 12 19:06:28 2019
Message (nlminb): relative convergence (4)

Coefficients:
                 omega       phil      kappal   kappastar        df
Estimate:   0.05391117 0.96055416 0.016146509 0.001710015 12.405100
Std. Error: 0.03776637 0.02696598 0.006586151 0.004326747  3.441445
                  skew
Estimate:   1.00347902
Std. Error: 0.03399201

Log-likelihood: -2654.556599
BIC:                3.136361
```

TABLE 8: Estimation results of parameters of Beta-Skew-t-EGARCH (1,1) model

4.3 Information Impact Curve of SSEI

To analyze the log-return series of SSEI more intuitively, we plot a graph to describe the volatility of the time series influenced by the information impact based on the ARMA (3,3)-EGARCH (1,1)-t model as shown in Figure 7. In the figure, the left-hand-side and right-hand-side represented the shocks from unfavorable and favorable information, respectively. We can easily find that the curve is asymmetric and that the left-hand-side is steeper than the right-hand-side. Therefore, the SSEI series has leverage effects – the impact of bad news outweighs that of good news on the log-return series of SSEI.

ISSN: 2304-1013 (Online); 2304-1269 (CDROM); 2414-6722 (Print)

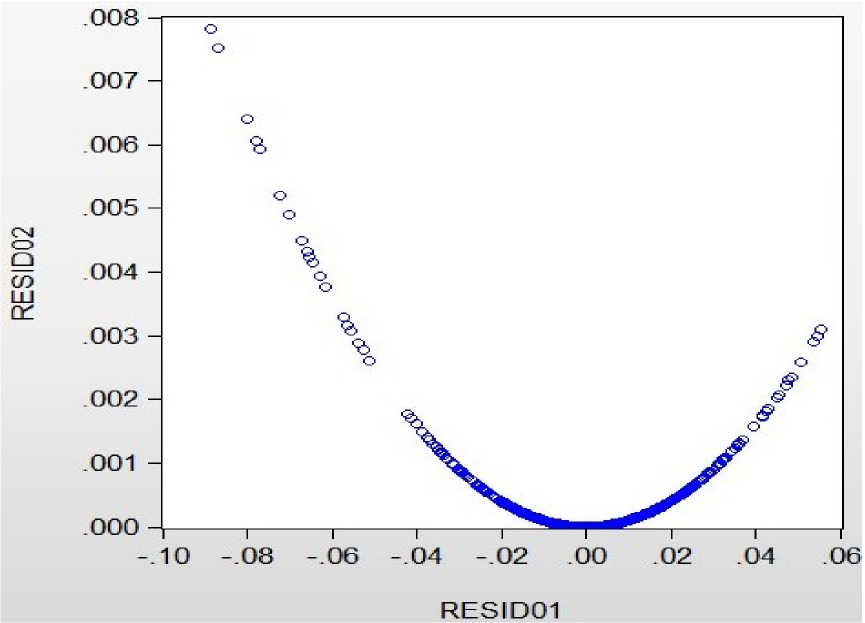


FIGURE 7: Information impact curve for log-returns series

4.4 Analysis and Testing of VaR Computation for the Log-returns of SSEI

Since the GARCH models under t-distribution as established in previous sections can fit the changes of the log-return series of SSEI, we use these GARCH models under t-distribution to estimate the value of VaR. The calculation steps are as follows:

(1) Use EViews9 software to calculate the parameters of each model;

(2) Compute conditional variance based on the parameters, and then obtain the volatility rate series δ_t by calculating the square root of the conditional variance;

(3) Substitute δ_t into $VaR = P_{t-1} \cdot Z_\alpha \cdot \delta_t$ to obtain the daily value of VaR.

where, Z_α is the α quantile of the t-distribution. EViews9 was used to obtain the results presented in Table 9.

Next, we computed the value of VaR and the results are presented in detail in Table 10. From Table 10, we can conclude that the value of VaR is a little bit different between the four models under t-distribution. We analyzed the VaR that we had calculated by using R=C to evaluate the actual profit and loss because we were able to observe the number of failure days and compute the corresponding failure rates. In order to visualize the changes of VaR and the log-return series, we use a positive value to represent an actual loss and a negative value to represent an actual profit. The results of the predicted VaR are plotted in Figure 8.

ISSN: 2304-1013 (Online); 2304-1269 (CDROM); 2414-6722 (Print)

Confidence level	GARCH model	EGARCH model	TGARCH model	Beta-Skew-t-EGARCH model
95%	2.082712	2.076619	2.082252	1.777451

TABLE 9: Quantiles of different GARCH models under t-distribution

Models	GARCH model	EGARCH model	TGARCH model	Beta-Skew-t-EGARCH model
Mean	77.5295868148	76.6110523452	77.5466345175	83.9403
Std. Dev.	53.4941012	52.3868275	53.88683447	59.3013
Maximum	316.5740863803	336.8295446819	320.7757886582	323.7642
Minimum	29.0848046340	28.6220594053	29.4778174260	31.4115
Median	55.4903872124	54.6948929949	54.9834935056	56.2053

TABLE 10: Results of Value at Risk under t-distribution

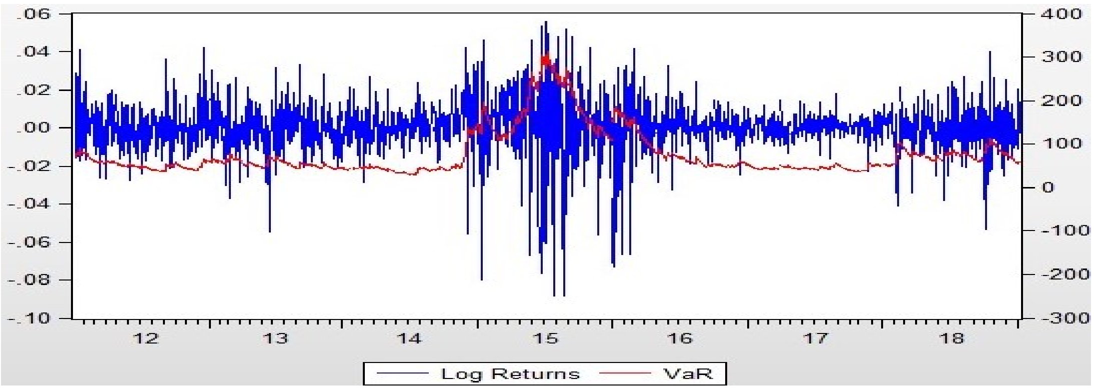


FIGURE 8.1: Graph of the VaR Test for SSEI fitting with MA (6)-GARCH (1,1) model

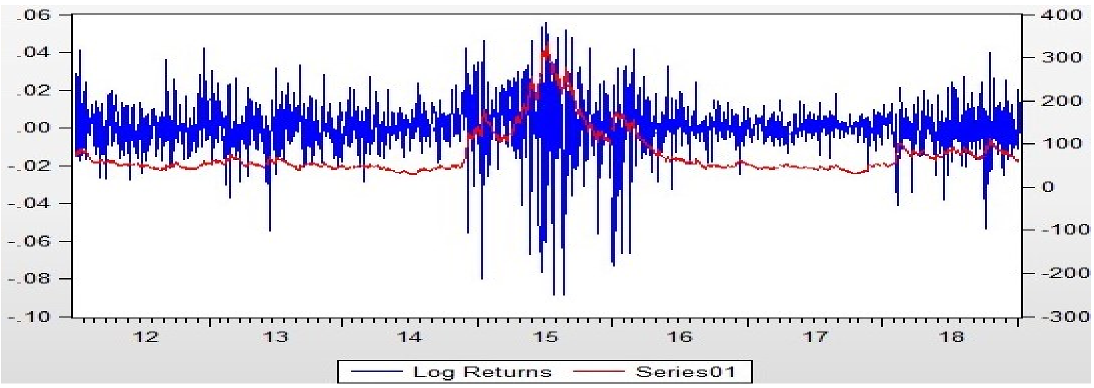


FIGURE 8.2: Graph of the VaR Test for SSEI fitting with ARMA (3,3)-EGARCH (1,1) model

ISSN: 2304-1013 (Online); 2304-1269 (CDROM); 2414-6722 (Print)

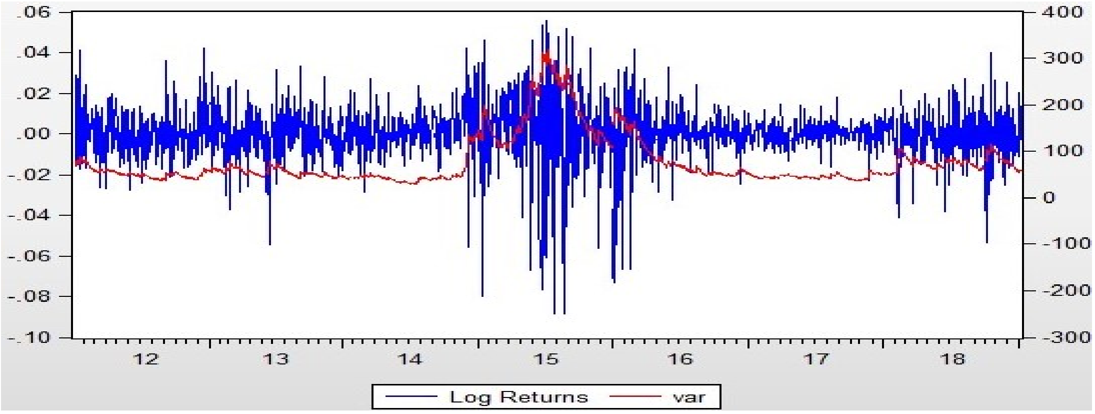


FIGURE 8.3: Graph of the VaR Test for SSEI fitting with ARMA (3,3)-TGARCH (1,1) model

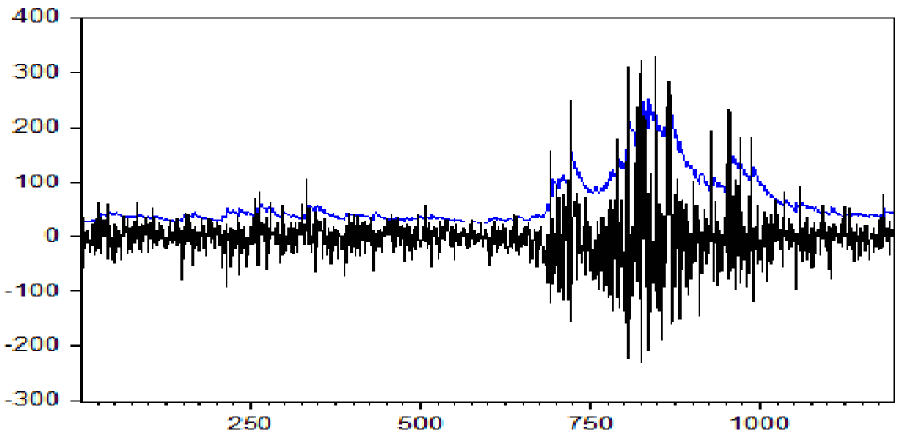


FIGURE 8.4: Graph of VaR Test for SSEI fitting with Beta-skew-t-EGARCH (1,1) model

The results in Table 11 demonstrate the risk measurement effectiveness and the failure rate of each model at the 95% confidence level.

Models	Expected Failure Rate	Expected Failure Days	Actual Failure Days	Actual Failure Rate	LR
GARCH model	5%	85	44	2.58%	2.5215
EGARCH model	5%	85	46	2.70%	2.9367
TGARCH model	5%	85	44	2.58%	2.5215
Beta-Skew-t-EGARCH model	5%	85	57	3.34%	3.1401

Note: A total of 1706 days

TABLE 11: Failure rate test of log-returns of SSEI

Based on Table 11, we can say that, in general, the actual failure days and failure rates of these models obey the t-distribution and are both smaller than the expected

ISSN: 2304-1013 (Online); 2304-1269 (CDROM); 2414-6722 (Print)

failure days and failure rates, respectively. The LR of these four models are smaller than the critical value of $\chi^2_{0.05}(1)$=3.84, which suggests that the models are effective in predicting the VaR of the SSEI series. Also, by comparing the goodness of fit and the VaR across the models, we found that TGARCH model is the best in analyzing the log-return series of SSEI.

5. DISCUSSION AND ANALYSIS

The academia and the industry both pointed out that it is essential to forecast stock prices in terms of time series analysis. However, it is not an easy task to precisely forecast stock prices due to the variability, complexity and other unique characteristics of stock prices. Moreover, stock prices are usually affected by many factors, with some of them quantifiable and some non-quantifiable. This creates high obstacles to research projects on this issue. With the development of modern statistical methods, researchers have started to pay more attention to how well time-series models fit the sample data and to focus on improving the existing models. For example, when we implement a complete regression analysis, the model may fit a specific set of data so well that the same model cannot be applied to other sets of data, which is known as overfitting. Under this circumstance, the significant difference in terms of fitting degree between the sample data and the out-of-sample data makes the prediction results not applicable in practice, especially when there are sudden changes that the model cannot sensitively respond to. This paper aims to identify a model that is stationary even for out-of-sample data by comparing modern and traditional statistical models. Based on a commonly examined financial time series, this paper elaborates various forecasting methods used in the financial industry as well as related algorithms and then applies them in the industry to implement the prediction.

In the process of estimating the parameters of the four models, we found that the fitting degree of the EGARCH model under t-distribution is the best for predicting the log-return series of SSEI. We also discovered that the SSEI series has leverage effects via analyzing the information impact curve. Moreover, we compared the GARCH models under different distributions to examine the SSEI series and verified that the TGARCH model under t-distribution fits well the volatility of the stock market. In addition, we implemented the Kupiec test and found that the TGARCH model follows a t-distribution and has the most effective VaR estimation. Furthermore, the distribution of interference factors can be divided into non-parametric or semi-parametric, which is more convenient for calculating quantiles. In this study, we implemented the Maximum Likelihood Estimation (MLE), the Gaussian Mixture Model (GMM), the Least Absolute Deviation (LAD) as well as the Empirical Likelihood Estimation (ELE), etc., because we believe that these estimation methods are more effective compared to the traditional MLE and Least Square Estimation (LSE). For future research, additional parameters reflecting industrial characteristics can be added to the models and the estimation methods for the parameters, which will make the forecasting of financial risks more reliable and accurate.

REFERENCES

[1] Shanghai Stock Exchange, *Introduction* [Online]. Available from:

ISSN: 2304-1013 (Online); 2304-1269 (CDROM); 2414-6722 (Print)

http://www.sse.com.cn/aboutus/sseintroduction/introduction/

[2] H. Tong. “Threshold Models in Nonlinear Time Series Analysis[J]”. *Lecture Notes in Statistics*, 1983, 21.

[3] MI Ahmad. “GARCH Models for Inflation Volatility in Oman[J]”. *Review of Integrative Business & Economics Research*, 2013, 2(2):1-7.

[4] Janacek G. “An Introduction to Bispectral Analysis and Bilinear Time Series Models. by T. Subba Rao; M. M. Gabr[J]”. *Journal of the Royal Statistical Society Series A (General)*, 1987, 150(2):326-329.

[5] Bollerslev, T. “Generalized Autoregressive Conditional Heteroskedasticity[J]”. *Journal of Econometrics*, 1986, 31:309-328.

[6] Granger C. W. J. & Morris M. J. “Time series modeling and interpretation[C]”. *Essays in Econometrics*. 2001.

[7] Hamilton J. D. “A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle[J]”. *Econometrica*, 1989, 57(2):357-384.

[8] Engle R. F. “Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation[J]”. *Econometrica*, 1982, 50(4):987-1007.

[9] Madsen H. “Time series analysis[J]”. *Chapman & Hall/crc Boca Raton Fl*, 2007, 27(1864):103-121.

[10] Olive D. J. “Prediction intervals for regression models[J]”. *Computational Statistics & Data Analysis*, 2007, 51(6):3115-3122.

[11] Diebold F. X. & Kilian L. “Unit-Root Tests Are Useful for Selecting Forecasting Models[J]”. *Journal of Business & Economic Statistics*, 2000, 18(3):265-273.

[12] Chang I, Tiao G. C. & Chen C. “Estimation of Time Series Parameters in the Presence of Outliers[J]”. *Technometrics*, 1988, 30(2):193-204.

[13] Lorraine Denby & Douglas Martin R. “Robust Estimation of the First-Order Autoregressive Parameter[J]”. *Publications of the American Statistical Association*, 1979, 74(365):7.

[14] Glosten L. R. & Jagannathan R., Runkle D. E. “On the Relation between the Expected Value and the Volatility of the Nominal Excess Return on Stocks[J]”. *The Journal of Finance*, 1993, 48(5):1779-1801.

[15] Harvey A. & Sucarrat G. “EGARCH models with fat tails, skewness and leverage[J]”. *Computational Statistics & Data Analysis*, 2014, 76(76):320-338.

[16] Gong R., Chen Z. C. & Yang D.R. “To Evaluate VaR of China Stock Marketing Comparatively by Using GARCH Family Model and Comment[J]”. *The Journal of Quantitative & Technical Economics*, 2005, 22(7):67-81.

[17] Chen S. D. & Yu S. D. “Analysis of China’s Stock Market Using VaR Method Based on GARCH Model[J]”. *Jilin University Journal Social Sciences Edition*,

ISSN: 2304-1013 (Online); 2304-1269 (CDROM); 2414-6722 (Print)

2002(4):11-17.

[18] Xu W. & Huang Y. L. “Empirical Analysis on GARCH type Models and VaR[J]”. *The Journal of Quantitative & Technical Economics*, 2008, 25(1):120-132.

[19] Ding Y. Z. “Leverage-Effect and Calculation of Value at Risk (VaR) In Shanghai Stock Market —Based on GARCH-t, EGARCH-t and TGARCH-t Models[J]”. *Statistical Thinktank*, 2009(1):9-13.

[20] Zhang H. B. (2005) “An Empirical Study on Asymmetry of Volatility in the China's Stock Market[D]”. Thesis (Master), Dongbei University of Finance and Economics.

[21] Zhou S. F. & Yuan X. X. “Asymmetry and Speculation in China Stock Market: An Empirical Study[J]”. *Contemporary Economic Management*, 2005, 27(3).

[22] Liu Z. W. “Akaike’s Information Criterion and its Significance[J]”. *Journal of Mathematics in Practice and Theory*, 1980(3):65-73.

[23] Philippe Jorion. “Value at Risk[M]”. *CITIC Press*, 2010.

[24] Shi T. X. & Qian J. Y. “Market Risk Measurement of Stock Index in China Based on VaR Models[J]”. *Journal of China University of Geosciences (Social Sciences Edition)*, 2010, 10(4):119-124.

[25] Li C. X., Zhang M. & Chen W. “The Application of ARCH Model in Shanghai Stock Market[J]”. *Journal of Mathematics in Practice and Theory*, 2008, 38(8):1-6.

ISSN: 2304-1013 (Online); 2304-1269 (CDROM); 2414-6722 (Print)