Empirical Study on the Investment Risk of SSE 50 Index Based on GARCH-VaR Model

Yanqun Gao

gaoyanqun@gdust.edu.cn

College of Finance and Economics, Guangdong University Of Science&Technology, Dongguan, Guangdong, China

Abstract. By testing the daily logarithmic return series of SSE 50 index in the past 10 years, the result satisfies that the residuals of the series obey the ARCH distribution, and the volatility of VaR can be estimated by using the GARCH-type models, and then estimated and analyzed with the help of the t-distribution and GED-distribution of the GARCH-type models, and finally using the data from July 1, 2013 to June 30, 2023, the VaR value of each trading day can be quantitatively measured accordingly. Finally, the VaR value of each trading day is obtained by using the data from July 1, 2013 to June 30, 2023, according to which the stock market risk is quantitatively measured, and the results of the study can provide useful references for stock investors as well as the risk management of investment institutions.

Keywords: VaR calculation; GARCH-type models; Risk measure; Kupiec test.

1. Introduction

In recent years, the results of various experts and scholars in analyzing the time series of financial market returns show that most of them are characterized by sharp peaks and tails, and there is a volatility aggregation in the data. In the process of empirical research, although there are many methods for predicting stock volatility and for correlation estimation, GARCH-type models have a more important application in the financial field.After Engel (1982) proposed the ARCH model, Bollerslev (1986) optimized the model, and then proposed the GARCH model.Later generations have made many extensions on this basis, and a series of GARCH-like models have appeared. Therefore, how to select the most suitable stock market return series GARCH model among many GARCH class models becomes the key to VaR calculation. Therefore, the empirical study of the investment risk of SSE 50 index based on GARCH-VaR model is suitable for quantitative analysis of the return of SSE 50 index stocks, which provides useful reference for investors and investment institutions.

ICEMME 2023, November 17-19, Beijing, People's Republic of China

DOI 10.4108/eai.17-11-2023.2342664

2. Model Introduction

2.1. VaR method

The VaR method is often used in finance to measure the market risk to which a financial asset or portfolio of assets is exposed.VaR (Value at Risk), refers to a measure of the maximum value of loss for an asset or portfolio of assets over a specific holding period and a certain level of confidence when the market is in a state of normal volatility.

The formula is expressed as in Eq. (1).

$$Pro\left(\Delta V_{\Delta t} < VaR\right) = 1 - \alpha \tag{1}$$

where Δt denotes the holding period, ΔV denotes the amount of loss in value of a given financial asset or portfolio of assets over the holding period Δt, $1-\alpha$ denotes the given confidence level, VaR is the maximum value of loss at the given confidence level, and Pro denotes the probability that the loss of an asset will be less than the maximum value of loss.

2.2. GARCH model

For the calculation of volatility in the VaR model, in order to more accurately fit the volatility of the time series variables as well as to better solve the problems posed by the large number of estimated parameters in the regression forecasting, the economist Bollerslev (1986) proposed the Generalized Autoregressive Conditional Heteroskedasticity model (GARCH model), which is based on the original expression of ARCH model. The main idea of this model is to supplement the original expression of the ARCH model with an autoregressive part of σ_t^2 , which replaces the lagged values of multiple ε_t^2 . The advantage of this model is that on the one hand, there are fewer parameters to be estimated, and on the other hand, the future conditional variance can be predicted more accurately. Thus, it can better deal with time series data in the financial field.

The GARCH (p, q) model is set to(2):

$$\sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \cdots + \alpha_q \varepsilon_{t-q}^2 + \gamma_1 \sigma_{t-1}^2 + \cdots + \gamma_p \sigma_{t-p}^2 \tag{2}$$

(where $\alpha_0 > 0, \alpha_1, ..., \alpha_q \geq 0, \beta_1, ..., \beta_q \geq 0$)

p is the autoregressive order of σ_t^2 , which corresponds to the lagged order of the historical returns, while q is the lagged order of ε_t^2 , which corresponds to the lagged order of the prior period series variance.[1]

GARCH model can not only reflect the volatility of financial assets to a certain extent, but also better explain the phenomenon of "volatility agglomeration", the most commonly used GARCH model is GARCH (1, 1), this model can more accurately reflect the volatility of financial assets. For the asymmetric impact of the return on volatility of some assets, Jean-Michel[2], Zakoian (1990)and Nelson (1991)have proposed TGARCH, EGARCH and asymmetric GARCH models by further adjusting the GARCH (p, q).[3]

3. Empirical studies

3.1. Data selection

In this paper, the closing price of SSE 50 index is analyzed empirically using Stata/MP 17.0 software. The closing prices of SSE 50 index for a total of 2435 trading days from July 1, 2013 to June 30, 2023 were selected with the help of Tom Tom software and necessary mathematical treatments were performed. In order to effectively remove the heteroskedasticity of the data, the daily logarithmic return of the SSE 50 index is used to express its real return rt, which is expressed by the mathematical formula(3):

$$r_t = Ln\left(\frac{p_t}{p_{t-1}}\right) \tag{3}$$

where Pt is the closing price of the SSE 50 index on day t and P_{t-1} is the closing price of the SSE 50 index on day t-1 .[4][5] Using this formula, 2434 returns can be calculated.

3.2. Characterization of logarithmic yield series

3.2.1. Normality test

The most intuitive way to test normality is to draw graphs, through histograms, kernel density plots, and QQ plots. Although the graphical method can roughly determine the normality of the perturbation term, rigorous statistical testing is also essential. Skewness and kurtosis indicators are often used as the basis for testing the normality of the series.

As can be seen from the statistical characteristics of the logarithmic returns and the histogram of the logarithmic returns, the mean of the logarithmic return series of the SSE50 is 0.001953, with more observations on both sides of the mean, the standard deviation is 0.0143723, and the skewness is -0.4526097<0, which indicates that the return series of the SSE50 has a long left trailing tail. The kurtosis is greater than 3, which is higher than that of the normal distribution, indicating that the return series has steeper peaks than the normal distribution, and overall shows the characteristics of sharp peaks and thick tails. The results of the above analysis do not conform to the nature and characteristics of the normal distribution, thus the logarithmic return series of the SSE 50 index is asymmetrically distributed. It is further concluded that the return series of SSE 50 index does not obey normal distribution and the assumption of normal distribution is rejected.

As can be seen from the logarithmic return of QQ Chart, the quantile scatter of the logarithmic return series of the SSE 50 index is not concentrated around the regression line marked in red, which does not conform to the quantile plot of the normal distribution. Thus, it can also be concluded that the series does not obey the normal distribution.

3.2.2. Stability tests

The ADF test is the most commonly used unit root test to determine whether a time series is smooth by testing for the presence of a unit root in the series.[6] The test is a left one-sided test, its rejection domain is only in the leftmost part of the distribution, the test statistic is the t-statistic.

Table 1 shows that the t-statistic value of ADF test is -13.854, which is smaller than the critical values of -3.430, -2.860, -2.570 for the t-statistic at the test level of 1%, 5%, 10%, respectively, so the original hypothesis of the existence of a unit root can be rejected, and it is concluded that the logarithmic return series of SSE 50 is smooth. The logarithmic return series of SSE 50 is smooth.

Table 1. Unit root test results of SSE 50 logarithmic return series

		Dickey-Fuller critical value		
	Test statistic	1%	5%	10%
Z(t)	-13.854	-3.430	-2.860	-2.570

3.2.3. Correlation test

The methods of correlation test include drawing, BG test, Box-PierceQ test, and DW test, and the most commonly used is to test the partial autocorrelation function and autocorrelation function using the Ljung-Box Q method.[7] By testing whether there is autocorrelation in the residuals to determine whether the variables can feedback effective information, and then conclude whether the model is valid.

As can be seen in Autocorrelation of the logarithmic returns and logarithmic return bias correlation chart, the statistics of partial autocorrelation and autocorrelation of the logarithmic return series of the SSE50 fall overwhelmingly within the 95% shaded confidence interval.

As can be seen from Table 2, the p-value of Ljung-Box Q-test is 0.0052, which can be considered that there is no autocorrelation and the Q-statistic is insignificant, i.e., the hypothesis that the autocorrelation statistic is zero can be not rejected at the 5% significance level, so that the series of the logarithmic returns of the SSE 50 index is significantly not autocorrelated.

Table 2. Q-statistics of the logarithmic return of the SSE 50 index

Portmanteau test for white noise
Portmanteau (Q) statistic = 66.6098
Prob > chi2(40) =0.0052

3.2.4. Existence test for ARCH effect

The GARCH model can be estimated only when the time series satisfies the existence of ARCH effect in the residuals of the series. In order to determine whether there is conditional heteroskedasticity in the disturbance term, a 5th order LM test is conducted to determine whether there is an ARCH effect in the log return series of the SSE 50 index. The results are shown in Table 3:

Table 3. ARCH-LM test of the logarithmic return of SSE 50 index

lags(p)	chi2	df	prob>chi2
1	96.881	1	0.0000
2	185.019	2	0.0000
3	239.173	3	0.0000
4	253.524	4	0.0000
5	258.983	5	0.0000

Table 3 shows that the test results for ARCH(1)-ARCH(5) all indicate that there is a significant ARCH effect. Therefore, there is an ARCH effect in the residual series of the logarithmic return of the SSE 50 index.[8]

3.3. GARCH-like model estimation and analysis

From the above test results, it can be seen that the logarithmic return series of SSE50 can be estimated and analyzed by GARCH model. Based on the characteristics and properties of the logarithmic return series of SSE 50 with "sharp peaks and thick tails", this paper adopts the GED distribution and t-distribution of the GARCH model for estimation and analysis. According to the minimum criteria of AIC and SBIC, it is determined that the best fitting effect is achieved when the order of the GARCH model is p=1, q=1. Then, the best model among four models, GARCH(1, 1), TARCH(1, 1), ARCH-M(1, 1) and EGARCH(1, 1), was selected to calculate the VaR value.

From the estimation results of Stata for each model and the collation results in Table 4, it can be seen that the TARCH term, ARCHM term and EARCH term (asymmetric effect) are not significant. Moreover, the AIC and SBIC values of the GARCH(1, 1) model under the GED distribution are smaller, and it can be further concluded that the estimation results of the GARCH(1, 1)-GED model are (4):

$$\sigma_t^2 = 2.24 \times 10^{-6} + 0.0748547\ \varepsilon_{t-1}^2 + 0.9144665\ \sigma_{t-1}^2 \tag{4}$$

Table 4. Estimates of the parameters of each model for the logarithmic return series of SSE 50 index (t-distribution, GED)

Model	α_0	α_1	λ_1	β_1
GARCH(1,1)-t	2.34×10^{-6} (3.24)	0.0756005 (6.64)		0.9165963 (83.74)
GARCH(1,1)-ged	2.34×10^{-6} (3.26)	0.0748547 (7.47)		0.9144665 (89.62)
TARCH(1,1)-t	2.34×10^{-6} (3.36)	0.0860392 (5.59)	-0.0176214 (-1.02)	0.9140458 (81.93)

TARCH(1,1)-ged	2.34×10^{-6} (3.34)	0.0805802 (6.33)	-0.0099539 (-0.66)	0.9131949 (88.77)
ARCHM(1,1))-t	0.0001885 (8.01)	0.2628158 (4.67)		
ARCHM(1,1))-ged	0.0001518 (18.68)	0.237214 (5.54)		
EGARCH(1,1)-t	-0.1185863 (-2.90)	0.0090145 (-0.69)	0.1708181 (7.84)	0.9849329 (208.64)
EGARCH(1,1)-ged	-0.1189676 (-3.05)	-0.0052643 (-0.45)	0.1698404 (8.55)	0.9852708 (220.36)

By performing the heteroskedasticity test on the residuals of the model, it was found that the P-value was not significant at 0, and there was no significant ARCH effect, indicating that the GARCH(1,1)-GED model was built effectively.[9]

3.4. Calculation and testing of VaR values

Based on the above definition of VaR, the formula for VaR can be obtained using the conditional variance calculated from the GARCH model(5):

$$VaR = P_{t-1} G_{\alpha} \sigma_t \tag{5}$$

where P_{t-1} is the closing price of the previous trading day, G_α is the quantile of a significance level under the GED distribution, and σ_t is the conditional standard deviation obtained from the GARCH model.[10][11]

3.4.1. In-sample calculation and out-of-sample prediction of VaR values

Through comparative analysis, the GARCH(1,1) model under the GED distribution is selected among the GARCH models, and the conditional variance series of returns can be obtained by using Stata, and the conditional standard deviation series is further calculated, and then the shape parameter under the GED distribution is found to be 1.184008, and then the quartiles of G0.1 with the confidence levels of 90%, 95%, and 99% are found, G0.05 and G0.01, which are -1.186251, -1.645393 and -2.652884, respectively, and then substitute σt, Pt-1 and the corresponding quartiles at the respective confidence levels into Equation (5) to calculate the corresponding VaR values.

By comparing the actual loss values in Figures 1, 2, and3 with the VaR estimates at each confidence level, it is found that most of the actual loss values of the closing price of the SSE 50 Index at the three confidence levels are above the VaR curve, indicating that the prediction range of VaR includes most of the actual loss values.

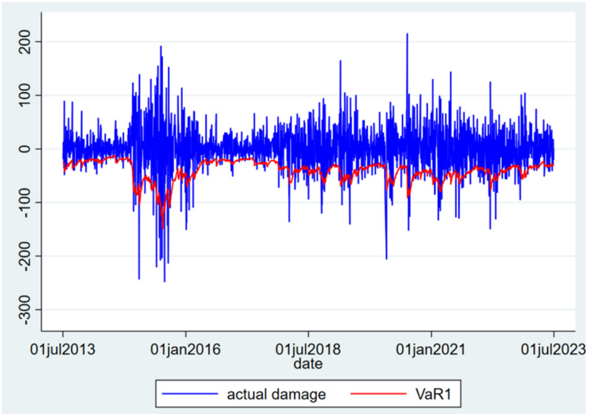

Fig. 1. Actual loss vs. VaR value with 90% confidence level

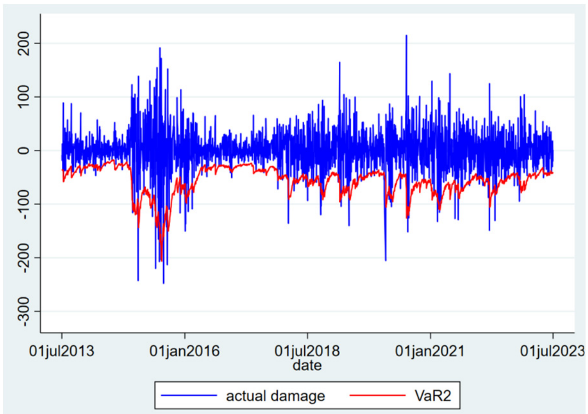

Fig. 2. Actual loss vs. VaR value with 95% confidence level

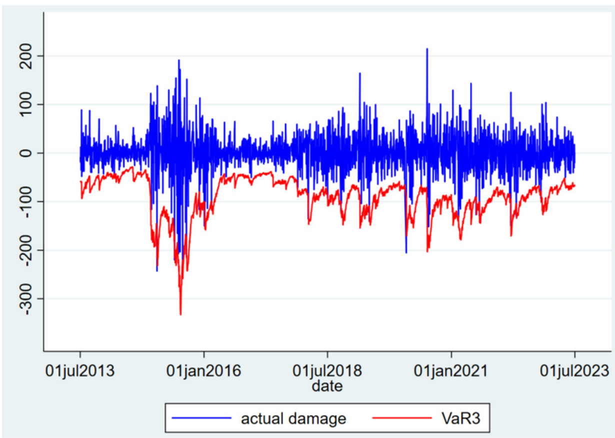

Fig. 3. Actual loss vs. VaR value with 99% confidence level

To further verify the accuracy of the model, the conditional variance of the GARCH(1,1) model is utilized for prediction, and the obtained prediction results are shown in Fig. 4. It can be seen that the conditional variance of the daily logarithmic return of the SSE 50 index, although fluctuating from time to time and sometimes even rising sharply, eventually decreases gradually and converges to the unconditional variance.

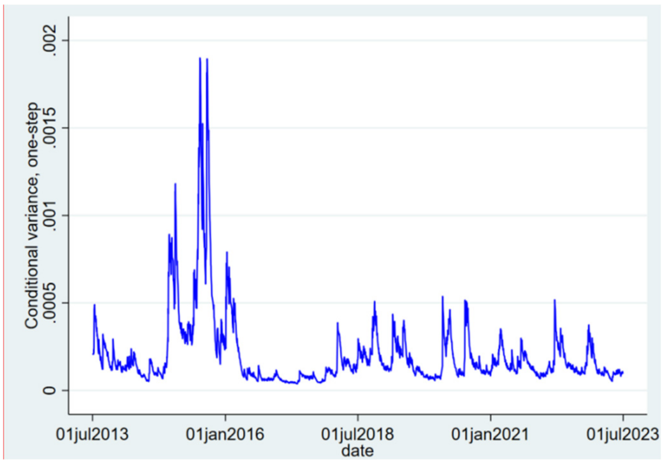

Fig. 4. Time trend of conditional variance of the log return series of the SSE 50 index

3.4.2 Test of VaR value

For the Va R estimates calculated using the GARCH-VaR method, the Kupiec failure frequency test was performed on them at 90%, 95%, and 99% confidence levels.

In Table 5, the LR statistic is the statistic of the Failure Frequency Test proposed by Kupiec, which is used to post-hoc test the model, and then to observe whether the model is effective in capturing the market risk or not. From the results, it can be seen that the mean values of VaR at confidence levels of 90%, 95% and 99% are -40.78154, -56.56616 and -91.2022, respectively, and the standard deviations are 18.90917, 26.22802, and 42.28771, respectively, with a large discrepancy between the three means and three standard deviations. As the confidence level increases, the number of actual loss values larger than the VaR values gradually decreases.

Table 5. Results of kupiec test for VaR

confidence level	VaR mean	VaR std	No.of failures[a]	No.of standards	failure rate[b]	LR
90%	**-40.78154**	18.90917	210	2435	8.62%	10.58
95%	**-56.56616**	26.22802	109	2435	4.48%	2.46
99%	*-91.2022*	42.28771	24	*2435*	0.98%	0.27

Note: [a] denotes the number of actual loss values that exceed the VaR value at the corresponding confidence level. [b]is the ratio of the number of failures to the number of criteria.

As can be seen from the number of failures in Table 5, the VaR values derived from the GED distribution based GARCH(1,1) model are less than the standardized level at all three levels of confidence. The last column of Table 5 shows that the value of the statistic is 10.58 greater than the chi-square value of 2.07 for a degree of freedom of 1 at the 90% confidence level, so the original hypothesis is rejected. The statistics of 2.46 and 0.09 at 95% and 99% confidence level respectively are less than the chi-square value of 3.84 for a degree of freedom of 1and the chi-square value of 6.63 for a degree of freedom of 1 ,so the hypothesis of test failure cannot be rejected. Therefore, the model fits the characteristics of daily logarithmic return of SSE 50 index better, estimates the value of VaR more accurately at 95% and 99% confidence level, and effectively measures the maximum loss that may be caused by market risk.

4. Conclusion

This paper draws the following conclusions by selecting the closing prices of SSE 50 index from July 1, 2013 to June 30, 2023 as samples and using Stata software to conduct empirical analysis:

First, the GARCH-like model can fit the nonlinear dynamic fluctuation of the daily logarithmic return of the SSE 50 index very well, and the series has smoothness and non-normality, and by testing and comparing and analyzing the different models, it is concluded that the most desirable estimation model is the GARCH(1,1) model under the GED distribution, which indicates that the model under the distribution can reflect more accurately the daily logarithmic return of the SSE 50 index This indicates that the model under this distribution can more accurately reflect the characteristics of the daily logarithmic return of the SSE 50 index with sharp peaks and thick tails.

Second, by testing the VaR estimates of the daily log returns of the SSE 50 index at three different confidence levels of 90%, 95% and 99%, it is concluded that the GARCH(1,1) model under the GED distribution passes the failure rate test at the confidence levels of 95% and 99%, indicating that the model is able to more accurately measure the maximum loss that may be incurred by the market risk.

Third, by analyzing the data characteristics of multiple distributions, the GED distribution is more applicable to the stock market of the SSE 50 index than the t distribution and the normal distribution, and is more effective in measuring the investment risk of the SSE 50 index.

In summary, the investment risk of SSE 50 Index has been relatively stable, but it is still necessary to be alert to its market volatility and invest with caution.

References

[1] Liu Qian.Wang Hedi,Investment Risk Analysis of Commercial Real Estate REITs in China Based on GARCHVaR Model [J]. Modern Marketing (Business Edition),2022(02):60-62

[2] Jean-Michel,Zakoian, Threshold heteroskedastic models[J].Journal of Economic Dynamics & Control,1994,18(5):931-955

[3] Nelson,D.B.Conditional Heteroskedasticity in Asset Returns:A New Approach [J]. Econometrica, 1991,59(2):347-370

[4] Qiu Jianze, The Impact of SSE 50 ETF Option Launch on SSE 50 Index - Based on Marko [J]. Financial News,2020(03):53-56
[5] Yang Haikun. Stock Market Asset Dynamic Allocation-A Study Based on DCC-GARCH and Efficient Frontier [D]. Fuzhou: Minjiang College,2019
[6] Yang Xiangjun,The economic value of patents: assessment model and fluctuation cycle [D]. Dalian: Dalian University of Technology, 2021
[7] Zhang Jingwen,Analysis of Shanghai and Shenzhen index risk measure based on GARCH-VaR model [D]. Daqing: Northeast Petroleum University, 2016
[8] Li Quanzhou. Research on Volatility Arbitrage Strategy of SSE 50ETF Options Based on Delta Neutrality[D].Changchun: Jilin University of Finance and Economics, 2020
[9] LI Zhuwei,LU Xuejiao,YANG Qianqian,WANG Xiaoshan. R&D and design of carbon option products in China--Based on carbon emission quota [J]. Investment Research,2022(04):48-51
[10] Yao Aiping,Ding Xiaowen. The effectiveness of risk measurement in China's iron ore futures market based on GARCH-VaR model[J]. Business Research,2021(03):36-39
[11] Ye Jingting. Market Risk Measurement and Control of Bank of China Fund [D]. Shanghai: East China Normal University, 2019