The Empirical Analysis of Shanghai Composite Index based on GARCH Model

Fei Ding

No.20,Chengzhong Road, Jiading District, Shanghai, China

fendi.sp@foxmail.com

Keywords: Shanghai stock index; GARCH model; VAR value

Abstract. Because of the financial field, the GARCH model has a wide range of applications in the simulation of financial time series volatility and the measurement of financial risk. Therefore, based on the GARCH model, the assumption of normal distribution is used to measure the accuracy of VAR values, and the failure rate of VAR values is detected. The results show that the daily yield of Shanghai stock index has the characteristics of "peak thick tail". The correlation test of the daily yield series of Shanghai stock index shows that the Eviews index has a strong correlation. Through the GARCH model to eliminate relevance, to yield modeling, and effectively predict the Shanghai Composite Index VAR value.

Introduction

With the rapid development of China's financial market in the past decade, but domestic experts have not formed a consensus on the measurement of financial risks, which cannot meet the needs of investors.

Foreign stock markets have a long history and mature development, and considerable achievements have been made in the research of stock market volatility. It is observed that the changes of the price of investment products and the changes of the rate of return in the securities market have stable and volatile periods, that is to say, the price fluctuations have clustering characteristics. For the volatility of stock market, clustering and persistence are the most prominent characteristics. The volatility of the stock market is time-Varying. It changes with the change of time. And in a certain period of time, there will be a consistently high or low situation, and it also has a long memory, that is, big fluctuations tend to have big fluctuations after big fluctuations, and small fluctuations tend to have small fluctuations after small fluctuations. Based on these characteristics of stock market price volatility.

Engle (1982) presented a breakthrough research result based on the volatility characteristics of data -- autoregressive conditional heteroscedasticity model (ARCH model). After he used the ARCH model to analyze the clustering of British inflation index fluctuations, various forms of the ARCH model have been widely used in theoretical research and practical operations. Bollerslev (1986) first extended the ARCH model to the generalized ARCH model, namely the GARCH model. Many empirical results show that GARCH can already reflect the price volatility of financial products in most cases. With the passage of time, it has been extended to IGARCH, EGARCH, and TGARCH model and so on. In foreign countries, researchers have used the GARCH model to carry out a large number of empirical analysis. In recent years, there are more and more researches on stock index volatility in China, but the empirical researches are still relatively few. Therefore, this article selects the Shanghai index data, based on the accurate calculation of VAR value GARCH model, from the perspective of volatility, GARCH model, calculate the VAR value, using the financial time series itself rush fat-tailed features to track changes in yield Variance method, provide a decision-making reference for preventing financial risks, hope to get the attention of the researchers. To provide investors with a certain portfolio.

Analysis of the Daily Yield of the Shanghai Composite Index. Select the closing price of 2750 stocks from January 4, 2005 to April 27, 2017 as the research object, and select the GARCH-VAR

  DOI: 10.25236/sser.2019.026

model to calculate the VAR value. Selection of samples. The peer rate of return on an asset, defined as a simple rate of return, taking natural logarithms, i.e.: Geometric yield

$$r_t = \ln(\frac{P_t}{P_{t-1}}) = \ln P_t - \ln P_{t-1} \tag{1}$$

Empirical Analysis of the Daily Rate of Return of the Shanghai Composite Index. First, the log yield represents the daily rate of return of the Shanghai composite index, and then the trend chart of the gnome yield of the Shanghai composite index is drawn through the software, see the figure below

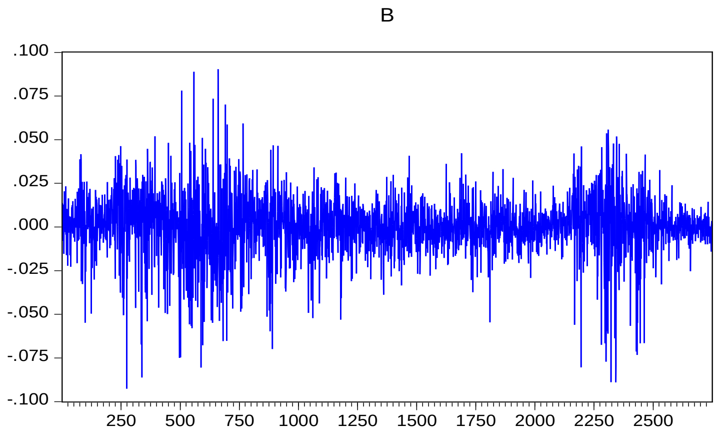

Figure 1. logarithmic yield of Shanghai composite index

From Fig. 1, we can see that the daily yield sequence fluctuates around 0, although the data fluctuates frequently at some points in time, but from the point of view of random process from the perspective of the whole process, we can basically think that belongs to the random process sequence.

The Positive Test of the Day's Logarithmic Return Rate in the Shanghai Composite Index. The most commonly used indicator to test whether a distribution is normal is bias and kurtosis, which is described in more detail below. Bias measures the degree of deviation from the left and right of a set of data, its value has 3 kinds of conditions, the bias of 0 means a symmetrical distribution of left and right, the bias is not 0 means a left and right asymmetry distribution, greater than 0 means positive, less than 0 is called negative bias, the standard positive distribution of the bias is 0. The kurtosis measures the degree to which a set of data peaks are above or below a normal distribution, and the kurtosis of any normal distribution is 3. This is the criterion to judge whether a distribution is consistent with the normal distribution, if the peak is greater than 3 called the peak state, less than 3 is called low peak state. The following table shows that the yield sequence has a bias of -0.6 and a peak of 6.9, and the preliminary judgment that the reassignment sequence is not a normal distribution.

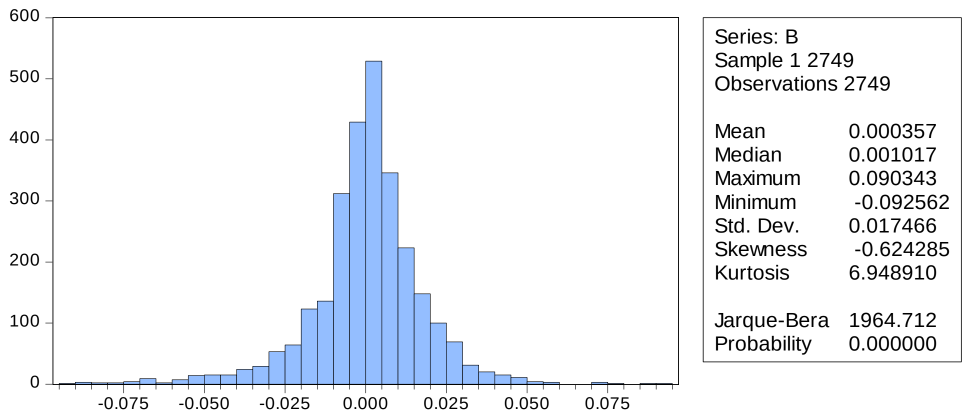

Figure 2. Distribution of the Shanghai Rate of Return

As can be seen from Fig. 2, the daily yield of the Shanghai Index shows that the asymmetric left-biased bias coefficient is less than 0, and the Peak coefficient is much greater than 3, indicating that his distribution is characterized by thick tail spikes. The J-B statistic is 1964.712, and the probability of obeying a normal distribution is 0.It can also be seen from the value of the test value that the distribution of the Shanghai composite index yield is non-normal distribution.

Smoothness Test. If it is necessary to model the volatility of the financial time series data, we should first ensure the stability of the time series data, otherwise there will be a pseudo-regression phenomenon, resulting in the parameter test results obtained on this basis have no practical value. The smoothness test is one of the indispensable test methods for analyzing time series. One of the prerequisites for modeling time series is that the sample data is stable.

The common method of smoothness test is that the ADF test is also called unit root test, and the test sequence has a unit root to determine whether the time sequence is smooth. The yield of the Shanghai Composite Index is tested at the root of the unit, and the results of the inspection are shown in Table 1.

Table 1 Unit root test results of the daily rate of return of the Shanghai Composite Index

		t-statistic	Prob.
Augmented Dick-Fuller test statistic		-23.93713	0.0000
Test critical values	1%level	-3.432541	
	5%level	-2.862394	
	10%level	-2.567269	

Table 1 shows that the value of the ADF test t statistic is -23.93713, and the corresponding probability value is 0.0000. The test level is divided into 1%, 5%, 10%, the threshold values of the t-statistic are -3.432541, -2.862394, and -2.567269. The ADF test t statistics are all smaller than the threshold value of the test level, so reject the original assumption that there is no unit root, so the yield sequence is stable.

Relevance Test. After the smoothness test, the data should be tested for correlation. Check for the presence of residuals for self-correlation. If the established model residual sequence is self-correlated, the model is invalid and there is no interpretation force. Correlation testing is generally tested by LQ method.

Table 2 Correlation Test of The Daily Rate of Return of the Shanghai Composite Index

Lagging Number	AC	PAC	Q-Stat	Prob.
1	0.023	0.023	1.4501	0.229
2	-0.026	-0.026	3.2743	0.195
3	0.027	0.028	5.2471	0.155
4	0.078	0.076	21.836	0.000
5	0.007	0.005	21.990	0.001
6	-0.064	-0.061	33.148	0.000
7	0.026	0.026	35.048	0.000
8	0.012	0.001	35.454	0.000
9	0.006	0.009	35.552	0.000
10	-0.010	-0.002	35.820	0.000
11	0.023	0.020	37.236	0.000

AR (1), MA (1), AR (2), MA (2), AR (3), MA (3), ARMA (1,1), ARMA (2,2), ARMA (3,3) can be established through graph analysis. After many tests and in accordance with the AIC guidelines and SC guidelines, it was found that it was better to model ARMA (1,1).The result is table 3.

Table 3 ARMA (1,1) Results are as follows

Variance Equation				
Variable	Coefficient	STD.Error	T-Statistic	Prob
AR(1)	-0.911287	0.051525	-17.68623	0.0000
MA(1)	0.930710	0.046742	19.91179	0.0000
R-squared	0.001810	Mean dependent VAR		0.000357
Adjusted R	0.001083	S.D. Dependent VAR		0.017466
S.E. of regression	0.017456	Akaike into criterion		-5.257119
Sum squared resid	0.836781	Schwarz criterion		-5.250660
Log likelihood	7228.910	Hannan-Quinn criter		-5.254785
Durbin-Watson stat	1.987981			

The results in Table 3 show that the two significance levels of the ARMA model are 0 and 0, indicating that the parameters have passed the significance test at the 5% significance level, and other indicators have passed. Therefore, we should accept ARMA (1,1) as a model for modeling the number of days in the Shanghai Composite Index. The value of the Durbin Watson test statistic is 1.987981, which is very close to 2, which means that there is no sequence self-correlation in the residual sequence.

Next, we use the ARCH effect to carry out heteroscedasticity test on residual sequence.

Existential Test of ARCH Effect. Engel proposed in 1982 whether the residual sequence had a Lagrange multiplier test with an ARCH effect, or ARCH-LM test. ARCH effect testing is primarily used to analyze the presence of conditional heteroscedasticity in time series. In order to verify the existence of conditional heteroscedasticity, the yield sequence is tested arch-LM.

Table 4 Results of the heteroscedasticity test of the daily yield of the Shanghai Composite Index

Heteroskedasticity Test:ARCH			
F-statistic	90.21982	Prob.F(1,2746)	0.0000
Obs*R-squared	87.41356	Prob.Chi-Square(1)	0.0000

As shown in table 4, the statistical value is 90.2198, the P value is almost 0, indicating that the model is significant, the observation value is 87.41356, the P value is almost 0, indicating that there is no heterotheoristic hypothesis of the rejection model residuals. There is a clear heteroscedasticity in the selected SSS composite index yield sample.

At the same time, looking at the residual stakes of the mean equation, it can be found that the residual fluctuations fluctuate "aggregation" phenomenon: larger amplitudes occur in some time periods, and smaller amplitudes occur in others, indicating that the residuals have a higher-order ARCH effect.

Based on the above analysis and test of the effect of the sample series of the Shanghai index yield (stability, self-correlation, heteroscedasticity), it is reasonable to think that the GARCH model is used to describe the volatility of the yield.

Empirical Study of the Volatility of the Shanghai Index based on the GARCH Model. Empirical research design: Using GARCH model to make an empirical analysis of the daily yield sequence of the Shanghai Composite Index.

From the statistical characteristics of the daily rate of return of the Shanghai Composite Index, the yield sequence is stable and has a high-order ARCH effect, so the GARCH model needs to be modeled. The GARCH model is basically a low-order GARCH model. In most applications, only low-level GARCH models are applied, such as GARCH (1, 1), GARCH (2, 1) models, and GARCH (1, 2).

In many cases, the GARCH (1, 1) model has been able to meet the requirements. Consider using the GARCH model to describe the peaking and tailings of the stock market. After many experiments with Eviews software, the ARMA (1, 1)-GARCH (1, 1) model is obtained, and its parameters are estimated as table 5.

Table 5 Results of the ARMA (1, 1)-GARCH (1, 1) model

Variable	Coefficient	STD.Error	Z-Statistic	Prob
AR(1)	-0.829171	0.158529	-5.230407	0.0000
MA(1)	0.849849	0.149802	5.673157	0.0000
		Variance Equation		
C	8.78E-07	2.75E-07	3.195406	0.0014
RESD(-1)^2	0.055897	0.004318	12.94476	0.0000
GARCH(-1)	0.942970	0.003941	239.2942	0.0000
R-squared	0.001280	Mean dependent VAR		0.000357
Adjusted R	0.000917	S.D. Dependent VAR		0.017466
S.E. of regression	0.017458	Akaike into criterion		-5.574716
Sum squared resid	0.837225	Schwarz criterion		-5.563951
Log likelihood	7667.447	Hannan-Quinn criter		-5.570827
Durbin-Watson stat	1.991593			

Based on the results shown in Table 5, the yield and Variance equations of ARMA (2, 2)-GARCH (1, 1) are estimated as:

$$\mathrm{r_t} = 0.849849\mathrm{u_{t-1}} - 0.82917r_{t-1} + u \qquad (2)$$

$$\phi^2 = 8.78 \times 10^{-7} + 0.0558973\varepsilon_{t-1}^2 + 0.942970\phi_{t-1}^2 \qquad (3)$$

It can be seen that a1+b1 is almost equal to 1, which has financial significance: in the stock market, the impact of earnings at a certain moment will have a sustained utility, volatility attenuation comparison (B1=0.942907), showing the cluster characteristics of volatility. a1=0.055897 indicates whether the current day's fluctuations still have a certain impact on the overall volatility.

It can be seen that the yield sequence of the Shanghai Index can be better fitted by the minimum information criterion AIC= -5.574716 in the table.

The following is a test of ARMA (1, 1)-GARCH (1, 1) (1, 1) by the residual sequence to determine whether the characteristics of stock market volatility can be characterized.

Table 6 ARMA (1, 1)-GARCH (1, 1) Heteroscedasticity Test Results

Heteroskedasticity Test:ARCH			
F-statistic	0.319975	Prob.F(1,2746)	0.5717
Obs*R-squared	0.320171	Prob.Chi-Square(1)	0.5715

From Table 6, we can see that the statistics F and R2 are large, and the P values are also large, indicating that there is no heteroscedasticity in the residual sequence.

It is shown that ARMA (1, 1)-GARCH (1, 1) can better capture the heteroscedasticity characteristics of the daily rate of return of the Shanghai Composite Index.

VAR's Calculation and Back Testing. The VAR value is calculated using the normal distribution and GARCH model to fit the yield fluctuation equation.

Empirical Analysis based on ARMA (1, 1)-GARCH (1, 1). Yield stake takes advantage of the FORECAST feature in Eviews to easily get the daily rate of return. The Variance is analyzed by the ARMA (1,1)-GARCH (1,1) model, and the Variance sequence can be easily obtained using the PROCESS-Make GARCH Variance Services feature in Eviews. At this point, it is easy to calculate the VAR value based on the ARMA (1,1)-GARCH (1,1) model.

$$\mathrm{VAR} = \mathrm{r} - \mathrm{z_a}\phi \quad (4)$$

when the confidence level is 95%:

$z_\alpha = 1.65$

The VAR value based on the ARMA (1,1)-GARCH (1,1) model is derived.

Failed Frequency Test Analysis. Based on the above-derived normal model and ARMA (1,1)-GARCH (1,1) model, the VAR value is compared to which of the two models is better for the Shanghai index simulation. Use the failure frequency test analysis. Assuming that the total number of days to be examined is T and the number of days of failure is N, the failure frequency is P (N/T), and the comparison of whether P is significantly different from P'. The zero-hypothesis is set to P-P, So that the test of VAR model accuracy and effectiveness is equivalent to assessing whether the frequency of P failure is significantly different from P.

Rejecting the zero hypothesis means that P is significantly different from P, that the model is not acceptable, and vice versa. At the 5% significance level, if LR is 3.8415 we reject the null hypothesis, that is, reject the model. If LR 3.8415 we cannot reject the zero assumption, that is, accept this model.

We tested them at normal and GARCH (1,1) distributions, and listed the results at the 5% significance level, such as Table 7.

Table 7 Results of two methods

Distribution	Expedition days	Failed days	LR
Normal distribution	2747	129	11.28
GARCH(1， 1)	2746	120	0.25

From the Table 7 can be seen from the test, the normal distribution at the 5% significance level LR 3.8415, indicating that the model did not pass the test, GARCH (1,1) model LR 3.8415, we cannot refuse zero assumptions, that is, accept the model. This actually shows that the Shanghai Composite Index has the characteristics of a peak and thick tail.

Therefore, it can be seen that the GARCH (1,1) model is feasible by means of the failed frequency test.

Analysis of Empirical Results and Economic Significance. The research object of this paper is to select the closing price of 2750 stocks from January 4, 2005 to April 27, 2017, this paper first uses the Red Pool Information Standard and Schwartz Information Criterion to determine the best effect of the ARMA (1,1) model of the Shanghai Composite Index Fund. GARCH (1,1), again examining the residuals, found that the information has been fully extracted, indicating that the yield sequence of the Shanghai Composite Index has been fully depicted. Based on normal distribution and ARMA (1,1)-GARCH (1,1) models to calculate the stock market risk value VAR value, so the VAR value is calculated using the model under construction and VAR's prediction test is performed using LR (likely estimate), the results show that ARMA (1,1)-GARCH (1,1)

The model has a good effect on VAR prediction, which shows that the ARMA (1,1)-GARCH (1,1) model is more suitable for studying China's financial markets. The yield distribution of the Shanghai Composite Index is far from the normal distribution. The above empirical research shows that the yield sequence is -0.6, the peak is 6.9, the yield sequence is not a normal distribution, its distribution has a large degree of bias and kurtosis, while presenting the characteristics of spikes and thick tails. Obviously, the ARMA (1, 1)-GARCH (1, 1) model is more meaningful to study market risk than the traditional normal distribution.

S**ummary**

This also provides evidence from the side that the financial time series is mostly a thick tail. Therefore, the choice of accurate VAR model can effectively describe the peak and thick tail characteristics of the return on financial assets, as well as verify the accuracy and effectiveness of Various risk measurement models, on the basis of which to develop risk management and investment management strategy, to determine the national regulatory authorities to monitor the stock market risk system.

Advanced risk measurement technology not only helps banks and non-bank financial institutions to measure and analyze more accurately, but also financial regulators rely on more scientific risk models to manage risk.

Acknowledgements

Thank you my teacher and my roommate in the roommate process to give great help, I can successfully complete this article.

References

[1]Engle. R.F: *Econometrica, Econometric Society* (United States, July, 1982). Vol. 50(4), p. 987-1007

[2Bollerslev. T: *Journal of Econometrics* (1986). 31, P.307-327

[3]Nelson. D.B: *Econometrica* (1991). 59, P.347-370

[4]Schwert. G.W: *Review of Financial* Studies 3, P.77-102.

[5]Calvet. L. and Fisher. A.J: *Review of Economics and Statistics*. 84, P.381-406

[6] G. Han and Wang Jing: *Xi'an Traffic Journal of the University* (2008). Vol. 9(4)