ISSN: 2688-9323

Analysis of GARCH Effect of Shanghai Composite Index based on R Language

Mengxiang Deng

Anhui University of Finance and Economics, Anhui, China

Abstract

The Shanghai stock index, also known as the Shanghai Composite Index, can reflect the fluctuation of stock prices listed on the Shanghai Stock Exchange, so as to intuitively show the trend of the market. Based on the data of Shanghai Composite Index from January 2, 1991 to December 25, 2020, this paper calculates the daily rate of return. Based on the results of LM Test in R language fints package, it is confirmed that the calculated daily rate of return series has Arch effect and meets the fitting conditions of GARCH model. The established GARCH model is used to predict the index price and stock return of Shanghai stock index for 5 days. The data are compared with the actual data. The empirical results are realistic. The GARCH model established by us is reliable for the short-term prediction of Shanghai stock index.

Keywords

Stock Index; Arch Effect; GARCH Model.

1. Introduction

The stock market is an important part of China's market economy and has played a great role in promoting the development of the national economy. At the same time, risk appetite people have invested capital in the stock market, hoping to have a good return in the sea of stocks. Obviously, in the process of investment, investors expect the higher the stock yield, the better, and the lower the risk, the better. Therefore, if we can predict the future return of stocks through relevant tools, it can not only enable investors to avoid risks, but also provide reference for the government to formulate various macroeconomic policies. In recent decades, scholars have proposed many different prediction methods of stock points or returns, such as analysis of technical indicators, including trading volume curve, exponential smoothing line, K-line chart, moving average and random index, simplest chart method, Aram model, ARFIMA model, Kalman filtering method and neural network model. These methods have played a good reference role for investors in stock selection. GARCH model is an effective method to describe the volatility of financial data. It is the most commonly used and convenient heteroscedasticity series fitting model. Shanghai Composite Index has the function of price disclosure as a representation of the fluctuation of stock price in the market. It is an important index to reflect the overall trend of the market, so it has attracted extensive attention. If the Shanghai composite index is analyzed and predicted to a certain extent, it can provide important reference value for investors to make trading decisions.

2. Literature Review

Foreign scholars have studied the prediction of stock return earlier. Black (1977) and Christie (1982) proposed that the response of financial time series to positive and negative shocks is quite different[1,2]. Phichhang ou (2010) compared the three models and concluded that the hybrid model performs better in predicting the volatility of leverage effect. The research on the Volatility Analysis of stock market return in China is relatively late, but with the continuous

ISSN: 2688-9323

development and improvement of China's financial market, the research on this aspect is also gradually increasing. In 2006, Kong Huaqiang fitted the volatility of Shanghai 180 and Shenzhen 100 index by establishing EGARCH (1,1) - M model[3]. Zhang Hao (2015) calculated the annual volatility of individual stocks by using GARCH (1,1,) model, estimated the return range of a certain period in combination with the normality of stock price, and put forward investment opinions according to the characteristics of domestic stock market vulnerable to national policies[4]. Li Xiongying and Chen Xiaoling (2018) compared ARMA model, GARCH model and arma-garch model and concluded that the prediction effect of the combined model was the best[5]. Huang Xuan and Zhang Qinglong (2018) proved that arima-garch comprehensive model has great advantages in short-term prediction of Shanghai stock index[6].

3. Organization of the Text

3.1. Model Introduction

Let the stationary time series $\{y_t\}$ be an autoregressive moving average model ARMA (p, q), then:

$$y_t = c_a + \sum_{i=1}^{p} \varphi_i y_{t-i} + \sum_{j=1}^{q} \theta_j e_{i-j} + e_i \tag{1}$$

Where, c_a is a constant, p is a nonnegative integer, the order of autoregression, q is a nonnegative integer, the order of moving average term, and φ_i is the coefficient of autoregression term, θ_j is the coefficient of moving average term and e_i is the residual term. Subscript t represents a specific time, and the stock index return changes with the change of time.

GARCH is an autoregressive conditional heteroscedasticity time series model that uses past variance and its predicted value to predict future variance. Heteroscedasticity means that the variance changes with time, that is, it has heteroscedasticity; Conditionality indicates the dependence on past near observation information; Autoregression describes the feedback mechanism of the relationship between predicted values and past observations. The GARCH model in the field of economics is used to characterize the heteroscedasticity information contained in the residual sequence $\{e_i\}$. Specifically, a GARCH (m, n) model can be expressed as:

$$\sigma_i = e_i \varepsilon_t \tag{2}$$

$$\sigma_i^{\ 2} = c_g + \sum_{i=1}^{m} \alpha_i e_{t-i}^2 + \sum_{j=1}^{n} \beta_j \sigma_{t-j}^2 \tag{3}$$

Where c_g is a constant, ε_t is the normalized residual sequence, $\sigma_i^{\ 2}$ is the conditional heteroscedasticity, and m and n are nonnegative integers, α_i, β_j is the coefficient. The equation can get the conditional expectation of the model. The information investors get in the transaction depends on the return at the past time and the error between the expected return and the actual return at the past time; According to the equation, the conditional variance of the model can also be described. It is not only a linear function of the square of the lag random disturbance term, but also a linear function of the conditional variance of the lag term. It shows that the fluctuation of the past time has a positive mitigation effect on the future price fluctuation, so as to simulate the volatility aggregation.

3.2. Empirical Analysis

The research object of this paper is the daily return series of Shanghai and Shenzhen 300 index. The Shanghai composite index data from January 2, 1991 to December 25, 2020 are obtained by using r-packet quantmod, and the daily closing price sequence is obtained. Based on these data, we preprocess the data and calculate the return series, model the return series with R software, analyze and predict the time series, and analyze the fluctuation of stocks.

3.2.1. Data Preprocessing

Because the financial data analysis generally studies the return on assets rather than the price of assets, and the return on assets is easier to deal with and more meaningful than the price series. Therefore, the logarithmic yield is calculated for the daily closing price, that is, the daily closing price of the Shanghai Composite Index from January 2, 1991 to December 25, 2020 is analyzed by taking the logarithmic difference. The expression of daily rate of return is:

$$r_t = ln(\frac{y_t}{y_{t-1}}) \tag{4}$$

Where r_t represents the yield of Shanghai Composite Index on day t and y_t represents the closing price on day t. The calculated index return series is the sample studied in this paper, with a total of 7329 data, of which five data from December 21 to December 25, 2020 are used as the test set to test the accuracy of the finally established GARCH model.

3.2.2. Test of Stationarity and Normality of Sequences

ADF test and JB test are used to test the stationarity and normality of the return series of Shanghai Composite Index. The results are shown in Table 1:

Table 1. The Result

Name	P_value
ADF Test	0.01
JB Test	-0.00000000000000022

In the test table, P < 0.05. It can be seen that the original hypothesis is rejected at the significance level of 0.05, that is, the daily yield series of Shanghai composite index is a stable series. According to the JB normality test, the p value is close to 0, indicating that the sequence is not normally distributed, which is also in line with the characteristics that financial data are generally peak and thick tail.

3.3. Descriptive Analysis of Yield Series

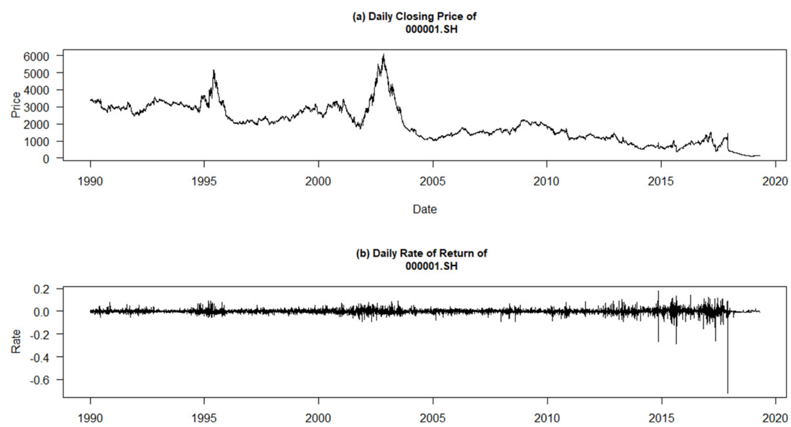

Figure 1. Sequence diagram of closing price and yield

ISSN: 2688-9323

The closing price sequence diagram can clearly show the daily price change of the stock index, and the yield sequence diagram can show the daily income of the stock index. Therefore, use the closing price and yield of Shanghai Composite Index in the observation period to draw a broken line chart to observe the changes of daily closing price and yield, as shown in Figure 1.

From Figure 1, the closing price of the Shanghai Composite Index shows a downward trend as a whole, which is related to the change of the calculation method of the Shanghai Composite Index. However, there were two large increases around 1996 and 2003. After consulting the data, it is found that this is related to the market environment and national policies at that time. On the contrary, the yield sequence diagram of the Shanghai Composite Index has almost no major changes and fluctuates around 0, which is related to the coexistence of returns and risks of the stock market itself. From the results of graphic analysis, there are obvious fluctuations and aggregation in 2008 and 2015. In addition, some descriptive statistical analysis is made on the yield series of Shanghai Composite Index. The results are shown in Table 2:

Table 2. Descriptive statistical variables of yield series

Variable	value
Mean	-0.000446
Standard deviation	0.02238
Skewness	-5.3386
Kurtosis	163.1857

From table 2, in the adjusted Sina yield data, the average value is -0.000446, which is very close to 0, indicating that the yield series of Shanghai Composite Index has a significant trend of concentration to 0; The standard deviation is 0.02238, which is close to 0, indicating that the dispersion of the yield of Shanghai Composite Index during this period is relatively small, which can also be said to be non dispersive. It can also be seen from Figure 1 that the dispersion of the yield series from 1991 to 2015 is very small, and the change after 2015 begins to become larger; The skewness is -5.3386, which is significantly not equal to 0, indicating that the yield distribution of Shanghai composite index is asymmetric; The kurtosis is equal to 163.1857, which is significantly greater than 3, indicating that there is an obvious peak thick tail phenomenon in the yield of Shanghai Composite Index, which is consistent with the previous JB test results.

In addition, the autocorrelation function (ACF) and partial autocorrelation function (PACF) can describe the autocorrelation of the return series, as shown in Figure 2:

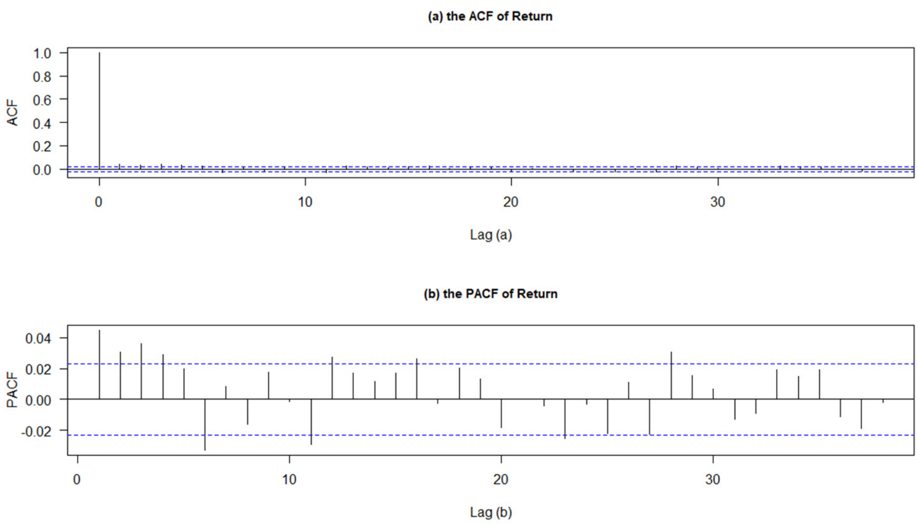

Figure 2. ACF chart and PACF chart of yield series

ISSN: 2688-9323

It can be seen from Figure 3 that most of the function values of the two graphs jump up and down within the confidence interval (the blue dotted line area in the figure), so the autocorrelation of the return series is very low, or has a very weak autocorrelation. Therefore, it is not necessary to introduce the autocorrelation component into the conditional expectation model to meet the mean equation in the GARCH model. The return is composed of a constant term and a random disturbance term.

3.4. Arch Effect Test

The timing chart of the yield shows that there may be arch effect in the daily yield data. If there is arch effect, the GARCH model can be fitted. On the contrary, GARCH model cannot be used to fit the equation. The LM Test in the fints package can be used to test the arch effect. The original hypothesis of the test is that there is no arch effect. The results obtained are shown in Table 3:

Table 3. ARCH Effect test results

Test statistic	value
Chi-squared	41.863
P_value	0.00003512

According to LM Test, the value of chi square statistic is 41.863, and its corresponding p value is almost 0, less than 0.05, that is, reject the original hypothesis at the significance level of 5%, which shows that there is arch effect in the yield series, so GARCH model can be fitted.

3.5. Estimation of GARCH Model

GARCH (1,1) is one of the most commonly used GARCH models and is also the most suitable model for financial time series modeling. Here, the garchfit function provided in the fgarch package is used to fit the GARCH model. The results of model estimation are shown in Table 4

Table 4. GARCH（1,1） Model results

coefficient	t
mu	0.000791
omega	1.61e-12
alpha1	<0.00000000002
beta1	<0.00000000002

It can be seen from table 4 that all coefficients are significantly different from zero at the significant level of 0.05, indicating that the fluctuation of Shanghai stock index return in the past has a significant impact on the current fluctuation and has a fluctuation aggregation effect. After fitting GARCH (1,1) model, in order to check whether it is the best fitting model, fit GARCH (1,2), GARCH (2,1) and GARCH (2,2) models at the same time, and use information criteria to select the best model. The results are shown in Table 5:

Table 5. The results of the four parameters of the GARCH model estimates

Model	AIC	BIC	SIC	HQIC
GARCH(1,1)	-5.4684	-5.4647	-5.4684	-5.4672
GARCH(1,2)	-5.4680	-5.4633	-5.4680	-5.4664
GARCH(2,1)	-5.4692	-5.4645	-5.4692	-5.4676
GARCH(2,2)	-5.4691	-5.4634	-5.4691	-5.4672

ISSN: 2688-9323

It can be seen from the values of the information criterion that the values of AIC, BIC, SIC, HQIC of the 4 GARCH models fitted were found to increase significantly as the parameters increased, and the values of AIC, BIC, etc. of the model did not increase significantly. Theoretically, the most concise model GARCH (1, 1) modeling is the most appropriate.

3.6. Standardized Residual Analysis of GARCH Model

For a reasonable GARCH model, the standardized residual is a sequence of independent and identically distributed random variables. Therefore, the adequacy of the fitted GARCH model can be tested by checking the residual sequence. The ACF and PACF diagrams of the residual sequence can test whether the information extracted by the model is sufficient. The ACF and PACF diagrams of the residual of GARCH (1,1) model and the ACF and PACF diagrams of the square of the residual are as follows:

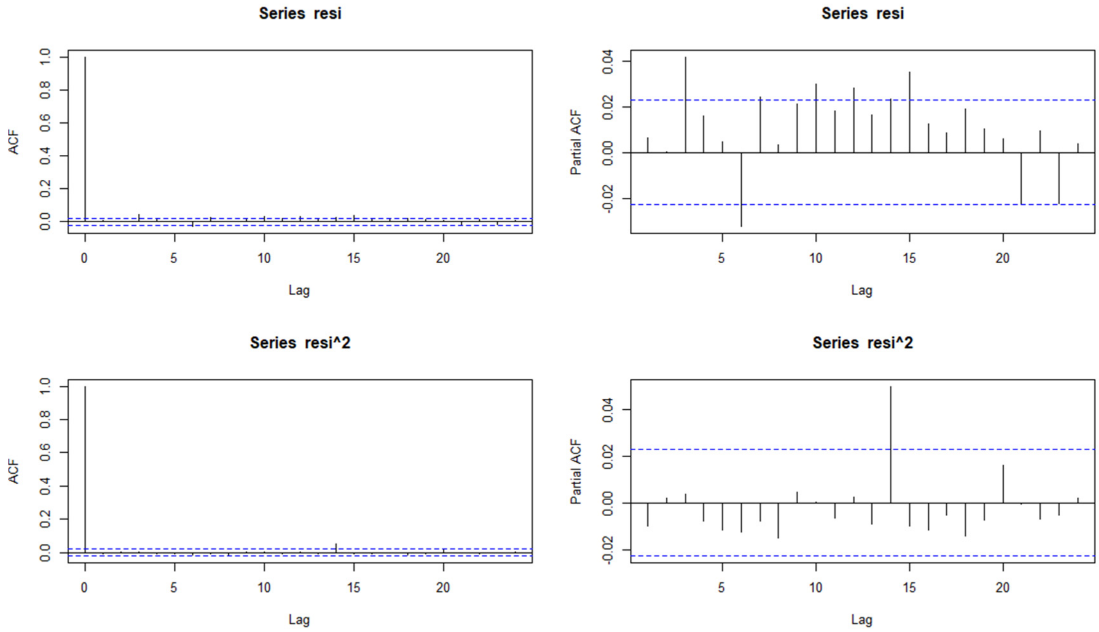

Figure 3. ACF and PACF diagram of residual and residual square sequence

From Figure 4, most of the function values of ACF and PACF diagrams of the residual sequence jump up and down within the confidence interval (blue dotted line area in the figure), so the standardized residual sequence does not have autocorrelation or has a certain weak correlation (PACF diagram result). The ACF and PACF diagrams of residual square series have no obvious tailing or truncation, and all function values are within the confidence interval, so they have no sequence correlation. Then the ACF value of standardized impact square is compared with the ACF value of return square. The results show that GARCH model can effectively explain the return series. In addition, Ljung box test is conducted for the square of standardized residuals. The original hypothesis is that there is no autocorrelation in the sequence. At the significance level of 0.05, the P values obtained when the lag terms are 10, 15 and 20 are 0.8431, 0.3602 and 0.5542 respectively, which are greater than 0.05. Therefore, the original hypothesis cannot be rejected, indicating that there is no sequence correlation in the standardized residual square series, that is, GARCH (1,1) is a significantly effective model.

3.7. Model Prediction and Verification

The established GARCH model is used to predict the index price and stock return of Shanghai stock index for a total of five days from December 21 to 25, 2020. The predicted index price is compared with the collected real price, and the relative error is calculated. The results are shown in Table 6:

ISSN: 2688-9323

Table 6. Comparison between predicted value and real value

Date	Real rate of return	Predict yields	The actual closing price	Predict the closing price	relative error
2020-12-21	0.75%	0.76%	3420.57	3467.77	1.38%
2020-12-22	1.88%	1.86%	3356.78	3313.81	1.28%
2020-12-23	0.76%	0.76%	3382.32	3391.45	0.27%
2020-12-24	0.57%	0.58%	3363.11	3405.49	1.26%
2020-12-25	0.99%	0.98%	3396.56	3351.05	1.34%

It can be seen from table 1 that although there is a deviation between the predicted value and the actual value of the model, the errors are within the range of 5%, and the overall prediction trend is consistent with the actual. The signs of the predicted rate of return and the actual rate of return are consistent, and the predicted closing price and the actual closing price are also fluctuating. It shows that GARCH (1,1) model is reliable for the Shanghai Composite Index to predict the stock fluctuation trend in the short term. Therefore, this model can be used as a good tool to analyze and predict the index.

4. Conclusion

This paper mainly analyzes and forecasts the daily return of Shanghai Composite Index, which proves the feasibility of fitting the model to predict the short-term stock volatility in the future to a certain extent. The following conclusions are drawn:

1. The closing price of the Shanghai composite index generally shows a downward trend, which is related to the change of the calculation method of the stock index. There is no obvious linear trend in its yield series, but there is a phenomenon of fluctuation aggregation;

2. The high kurtosis of the yield of Shanghai composite index indicates that there are extreme price changes in stocks. The skewness is negative, indicating that the probability of significant decline in the rate of return is great;

3. Model prediction is applicable to short-term prediction. Because long-term prediction is affected by many factors, such as industry changes, enterprise operation and national policies, the accuracy of short-term prediction is better than long-term prediction. Therefore, investors need to be more cautious if they take the long-term prediction results of the model as a reference.

References

[1] Black, F. Studies of Stock Price Volatility Changes. Proceedings of the 1976 Meeting of the Business and Economic Statistics Section, American Statistical Association,(Washington DC, American,1976). 81, p.177-181.

[2] Christie A A . The stochastic behavior of common stock variances : Value, leverage and interest rate effects. Journal of Financial Economics, Vol.10(1982) No.4, p.407-432.

[3] H.Q. Kong: Financial Market Volatility Model and Empirical Research (MS.,Capital University of economics and trade, China 2006J, p.27.

[4] H. Zhang, Y.F. Li, J. Mou. Research on the prediction of China's stock market return based on GARCH (1,1) model. Economist, Vol.30 (2015), No.8, P.97-99 + 101.

[5] X.Y. Li, X.L. Chen, K.H. Zeng.Forecast research of stock return rate of the Big-four Chinese banks based on three models. Journal of Quantitative Economics, Vol.35(2018),No.4, p.21-27.

[6] X. Huang, Q.L. Zhang. Analysis and forecasting the volatility based on ARMA-GARCH model of the CSI 300 index. China Price, Vol.31(2018) No.6, p.44-46.