ISSN 2616-5902 Vol. 4, Issue 7: 123-128, DOI: 10.25236/AJBM.2022.040720

Explore the impact of the stock market crash on the logarithmic return of *ST Beautiful based on the ARMA-GARCH model

Zhiqin Yang, Yuanyuan Pan*

School of Science, Minzu University of China, Beijing, 100081, China
**Corresponding author*

Abstract: *The analysis and prediction of stock prices have always been the focus of our stock market data analysis. The impact of stock crashes on stock prices as well as people's purchasing psychology is huge. This paper explores the changes in the log returns of the stock *ST Beautiful (000010. SZ) before and after the stock market crash for the two major stock market crashes in 2008 and 2015, and the results show that the whole stock is divided into four time periods with conditional heteroskedasticity, which can be fitted with GARCH(1,1). The impact of the crash on the conditional heteroskedasticity of the whole stock is not significant. However, the mean information component of the log-returns before and after the crash will change relatively significantly.*

Keywords: *ARIMA model, ARMA-GARCH model, conditional heteroskedasticity*

1. Introduction

Exploring and tracking the changes in stock prices is the goal of analyzing the stock market. Our stock market has gone through many ups and downs. Severe ups and downs can lead to stock market crashes, including the 530 stock market crashes that occurred from 2007 to 2008 and the stock market crash that occurred after June 18, 2015, which had a significant impact on the stock market. So is it possible to analyze the changes before and after a stock crash from the perspective of quantitative trading based on the data available for stocks? In this paper, based on the mean-variance model in the embryonic stage of the development of financial capital asset pricing theory, the closing prices of stocks are divided into four segments according to the time points of the two stock crashes: before the 2007 crash, the 2007 crash and after, before the 2015 crash, and the 2015 crash and after. In this paper, we will analyze the closing price of the stock *ST Beautiful (000010. SZ) over the years and finally come up with the movement of the price of this stock before and after the crash.

2. Literature Review

2.1. ARMA model

The full name of the ARMA model is the autoregressive moving regression model [1], ARMA model is for smooth non-white noise series to extract meaningful information, in the current view is the most widely used smooth series model in the smooth series. It has autocorrelation coefficients and partial autocorrelation coefficients without truncated tails. The model ARMA(p,q) can be fitted to it after determining the order of the model, and a well-fitted model can be predicted for the entire data in the specific form of:

$$X_t = \sum_{i=1}^{p} a_i * X_{t-i} + \sum_{j=0}^{q} b_j * \varepsilon_{t-j} + c_0, b_0 = 1 \tag{1}$$

where ε_t is the error of the model at time *t*, X_t is the specific value at time *t*, and c_0 is a constant.

2.2. GARCH model

The GARCH model was proposed by Bollerslev in 1986, which not only contains the ARCH model but also adds the conditional variance δ_t^2 of ε_t with lagged values in each period, making the model

ISSN 2616-5902 Vol. 4, Issue 7: 123-128, DOI: 10.25236/AJBM.2022.040720

have long memory for the conditional variance [2]. In which it is modeled for a smooth white noise series, mainly to extract the fluctuation information in the data. the specific form of GRACH(p,q) is:

$$X_t = \varepsilon_t \delta_t, \delta_t^2 = C_0 + \sum_{i=1}^{p} b_i X_{t-i}^2 + \sum_{j=1}^{q} a_j \delta_{t-j}^2 \tag{2}$$

where ε_t is subject to the standard normal distribution, $C_0 \geq 0, b_i \geq 0, a_j \geq 0$ are constants and for all t, ε_t and $\{X_{t-k}, k \geq 1\}$ are independent.

3. Model Building

3.1. Pre-analysis of data

A preliminary analysis of *ST Beauty (000010. SZ) from 19951027 to 20200724 trading day closing price.

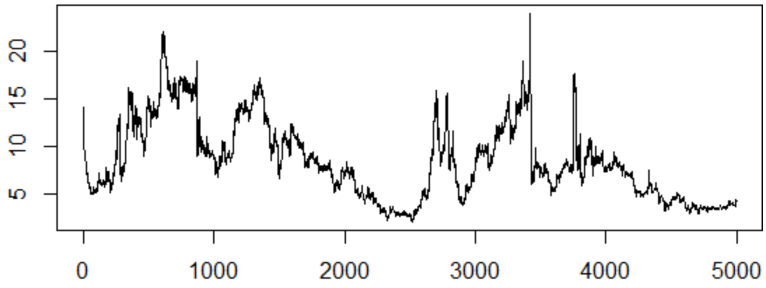

Figure 1: Time series chart of the closing price of 000010

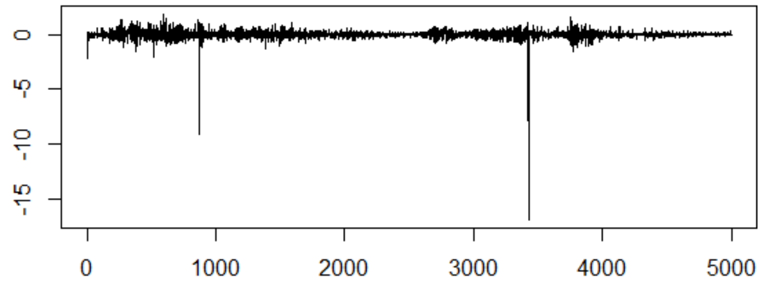

Figure 2: Time series chart of the closing price of 000010.SZ after differencing

The closing price of 000010.SZ has a maximum value of 23.873, a minimum value of 2.09, and a mean value of 8.419. The volatility of the data is still relatively high, as can be seen in Figure 1. After differencing the data, it is found that the data fluctuates around the value of 0, indicating that it is still more in line with the mean pricing theory and that the price of a stock will fluctuate up and down around its value in the long run. Although the test of smoothness and white noise of the original data shows that the data are smooth and not white noise, we use logarithmic returns for the next stage of the analysis because of the need to segregate the data for analysis afterward. ($\log r_t - \log r_{t-1} = V_t$)

3.2. Phase Analysis

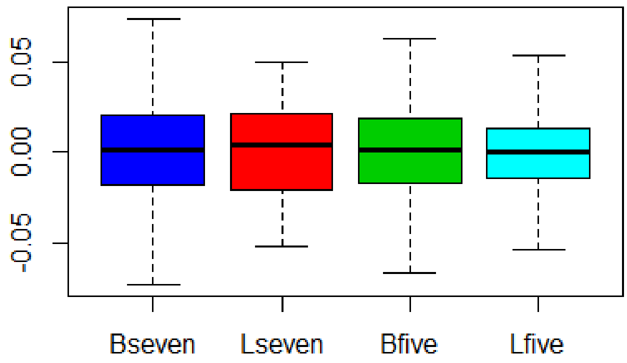

Figure 3: Box line diagram of the four stages

ISSN 2616-5902 Vol. 4, Issue 7: 123-128, DOI: 10.25236/AJBM.2022.040720

The data are divided into four periods, 20021111 to 20071008, 20080402 to 20100421, 20120316 to 20141225, and 20150519 to 20200724, and they correspond to the 2008 pre-crash log return, the 2008 post-crash log return, the 2015 pre-crash log returns, and log returns after the 2015 crash.

In the log-returns of the four stages, the mean values are basically the same, but there are some differences in the range of fluctuations of the data and the most values are also very different. The prediction of the next analysis of the data can be done with a specific analysis of the fluctuations of the data.

3.2.1. Phase I: 20021111 to 20071008

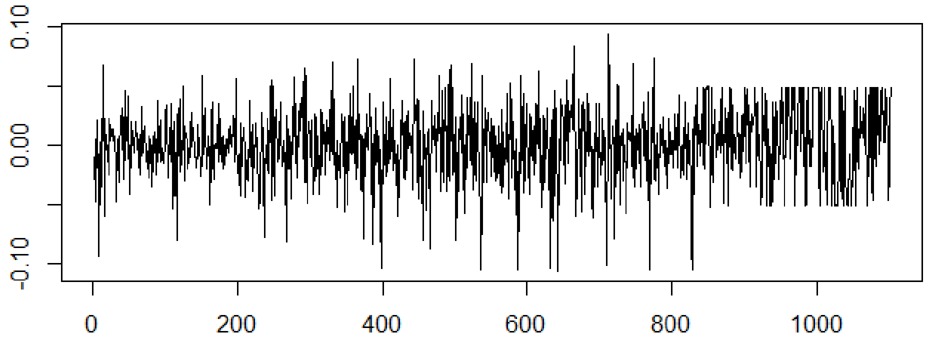

Figure 4: Logarithmic returns for the first stage

In the time series plot of log returns (Figure 4), the data fluctuate around the value of 0. The fluctuations are relatively smooth and not very large in magnitude. At the significant level of 5%, it is known that the data are satisfying smoothness and the data are not white noise series, thus, the ARIMA model can be established for the data.

It can be determined that the order of the model is ARIMA (1,0,1), and according to eACF, it can be established: ARIMA (0,0,1). Models ARIMA (1,0,1) and ARIMA (0,0,1) are established and compared. According to the AIC criterion of smaller AIC values, it is known that the model ARIMA (0,0,1) is more appropriate.

Then the specific model fitted to this data is:

$$v_t = 0.1107\varepsilon_{t-1} + 4 * 10^{-3} + \varepsilon_t \tag{3}$$

The residuals of the model ARIMA (0,0,1) fluctuate from -3 to 3 with no correlation, and the p-values are greater than the significance level of 0.05. The residuals are more normal, but there are some deviations from the normal values at the first end. In the white noise test of the residuals of the model, the original hypothesis is accepted as the p-value is greater than the significance level of 5% and the data are white noise. In the ARCH effect test, at the significance level of 5%, the residuals of the model are by the ARCH effect since the P-values of the data are less than 5%.

The next step is to do an ARCH effect fixed order on the residuals of the model only, followed immediately by ARMA-GARCH modeling. The orders of the GARCH models selected after the AIC and BIC criteria are GARCH(1,1).

By analyzing the *R* results, the function of the model takes the specific form of:

$$\begin{cases} v_t = \delta_t \varepsilon_t + 8.4 * 10^{-2} \varepsilon_{t-1} \\ {\delta_t}^2 = 1.133 * 10^{-5} + 4.136 * 10^{-2} v_{t-1} + 0.9487 {\delta_{t-1}}^2 \end{cases} \tag{4}$$

For the residual analysis of the model, the residuals of the modelδ_t fluctuate between 0.02 and 0.035 and the residuals ε_t fluctuate uniformly around the value of 0. The p-value in the normality test of the model (Jarque-Bera and Shapiro-Wilk) is less than the significance level of 0.05, so the residuals of the model do not obey the normal distribution, and its Q-Q plot shows that some of the data deviate from normality at both ends, especially at the front end. In the white noise test for the residuals, the p-values (Ljung-Box) are all much greater than the significant level of 0.05, indicating that it is white noise. It can also be seen specifically that the residuals of the model are white noise, and the ACF plot of the residuals also shows that the series is not correlated. In the white noise test of the squared residuals, which is also white noise, and in the ACF plot of the squared residuals, it is also seen that it is also largely within the significance level, i.e. it is largely uncorrelated. In the ARCH effect test of the residuals (LM Arch Test), the p-value is greater than 0.05, and the model residuals have no ARCH effect anymore. The model is

ISSN 2616-5902 Vol. 4, Issue 7: 123-128, DOI: 10.25236/AJBM.2022.040720

quite good except for the normality, which is not very good, so the model fits better.

3.2.2. Phase II: 20080402 to 20100421

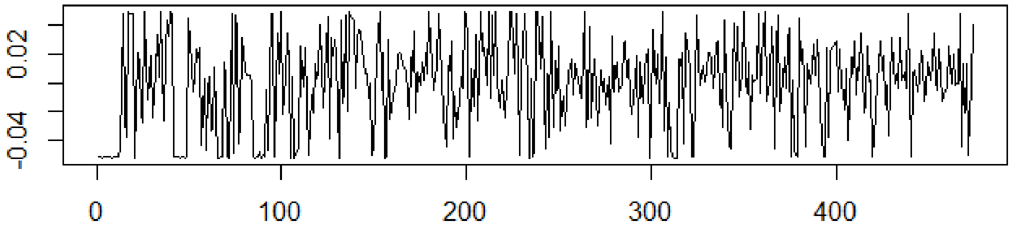

Figure 5: Time-series diagram of the logarithmic rate of return in the second stage

In the time series plot of log returns, the data fluctuate around the value of 0. The fluctuations are relatively smooth and not very large in magnitude. At the significant level of 5%, it is known that the data are satisfying the smoothness and the data are not white noise series, thus, the ARIMA model can be established for the data.

According to the ACF and PACF diagrams, the order of the model can be determined as ARIMA (1,0,3), and according to eACF, it can be established as ARIMA (0,0,3) or ARIMA (1,0,1), model building and comparison.

According to the AIC criterion that the smaller the AIC value is, the model ARIMA (1,0,1) is appropriate.

The corresponding model function takes the form:

$$v_t = 0.7103v_{t-1} - 0.5414\varepsilon_{t-1} - 0.0001 + \varepsilon_t \tag{5}$$

The residuals of the model ARIMA (1,0,1) fluctuate from -2 to 2, with no correlation and p-values greater than the significance level of 0.05. The residuals are more normal, but there are some deviations from the normal values at both ends. In the white noise test of the residuals of the model, the original hypothesis is accepted as the p-value is greater than the significance level of 5% and the data are white noise. In the ARCH effect test, at the significance level of 5%, the residuals of the model are by the ARCH effect since the P-values of the data are less than 5%.

The next step is to do an ARCH effect fixed order on the residuals of the model only, followed immediately by ARMA-GARCH modeling. The orders of the GARCH models selected after the AIC and BIC criteria are GARCH(1,1).

$$\begin{cases} v_t = \delta_t\varepsilon_t + 0.9338v_{t-1} - 0.8874\varepsilon_{t-1} \\ {\delta_t}^2 = 1.294 * 10^{-4} + 0.1561v_{t-1} + 0.7242{\delta_{t-1}}^2 \end{cases} \tag{6}$$

For the residual analysis of the model, the residuals of the model δ_t fluctuate between 0.025 and 0.004 and the residualsε_t fluctuate uniformly around the value of 0. The p-value in the normality test of the model (Jarque-Bera and Shapiro-Wilk) is less than the significance level of 0.05, so the residuals of the model do not obey the normal distribution, and its Q-Q plot shows that some of the data deviate from normality at both ends, especially at the front end. In the white noise test for the residuals, the p-values (Ljung-Box) are all greater than the significant level of 0.05, indicating that it is white noise. It can also be seen specifically that the residuals of the model are white noise, but the ACF plot of the residuals also shows that the series correlates with about 3rd order. The white noise test of the squared residuals, which is also white noise, is seen in the ACF plot of the squared residuals, which shows that it is also partially correlated. In the ARCH effect test of the residuals (LM Arch Test), the p-value is greater than 0.05, and the model residuals have no ARCH effect anymore. The model fits well, except for the normality and partial correlation, which are not very good.

3.2.3. Phase III: 20120316 to 20141225

The data set (Figure 6) is very volatile over a certain period, with the difference between the minimum and maximum values being close to 1.3. For the rest of the time, the log-returns are stable around the value of 0 and the price fluctuations are not significant.

When the smoothness test was performed on this group of data, it was found that the data were smooth at the significance level of 5%; the data were white noise because the P-value was greater than the significance level of 0.05 when the data were subjected to the white noise test. the P-value of the ARCH effect test reached 1 and the data were not ARCH effect. These indicate that the stock price fluctuations

ISSN 2616-5902 Vol. 4, Issue 7: 123-128, DOI: 10.25236/AJBM.2022.040720

during the period belong to the consolidation state and there is not much profit margin if the stock is held for a long time during this period.

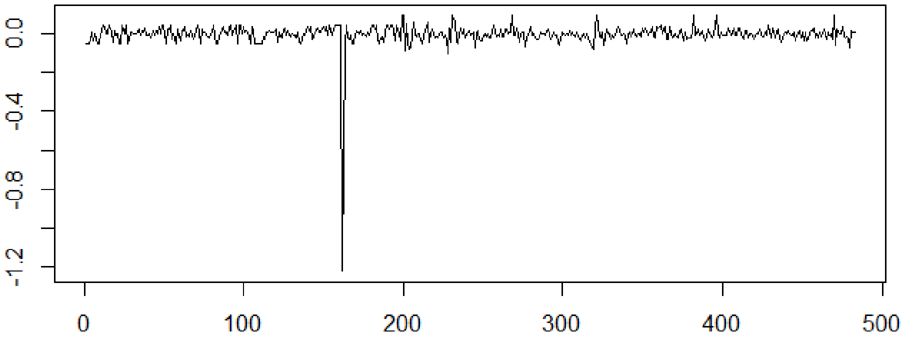

Figure 6: Time-series diagram of the logarithmic rate of return in the third stage

3.2.4. Phase IV: 20150519 to 20200724

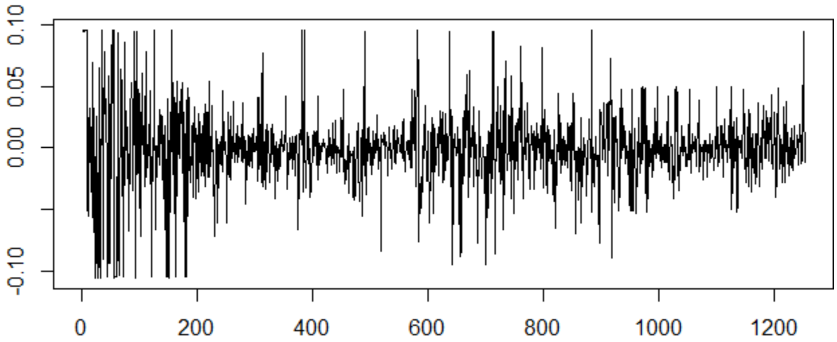

Figure 7: Time-series diagram of the logarithmic rate of return of the fourth stage

In the time series plot of log returns (Figure 7), the data fluctuate around the value of zero, with some periods being more volatile and others being less volatile within the period and may contain heteroskedasticity. At the significant level of 5%, it is known that the data are satisfying smoothness and that the data are not white noise series, thus, an ARIMA model can be built for the data.

The order of the model can be determined from the ACF and PACF plots as ARIMA (10,0,10), while according to eACF, it is possible to build: ARIMA (0,0,4), ARIMA (1,0,2), ARIMA(2,0,1), ARIMA (3,0,3), ARIMA (4,0,5), ARIMA (4,0,5) ARIMA (5,0,5), ARIMA (6,0,5), ARIMA (7,0,7) modeling and comparison. The model ARIMA (6,0,5) is more appropriate according to the AIC criterion that the smaller the AIC value is, the better.

The corresponding model function takes the form

$$v_t = -0.5346v_{t-1} - 0.0474v_{t-2} + 1.554v_{t-3} + 0.6963v_{t-4} + 0.8112v_{t-5} - 0.1715v_{t-6} + 0.0698\varepsilon_{t-1} + 0.1439\varepsilon_{t-2} - 0.1636\varepsilon_{t-3} - 0.703\varepsilon_{t-4} - 0.9719\varepsilon_{t-5} - 9 * 10^{-4} + \varepsilon_t \quad (7)$$

The residuals of the model ARIMA (6,0,5) fluctuate from -4 to 4 with no correlation, and the p-values are all greater than the significance level of 0.05. The residuals are approximately normal, but there are many deviations from the normal values at both ends. In the white noise test of the residuals of the model, the original hypothesis is accepted as the p-value is greater than the significance level of 5% and the data are white noise. In the ARCH effect test, at a significance level of 5%, the residuals of the model are by ARCH effect since the p-values of the data are all much less than 5%.

The next step is to do an ARCH effect fixed order on the residuals of the model only, followed immediately by ARMA-GARCH modeling. The orders of the GARCH models selected after the AIC and BIC criteria are GARCH(1,1).

$$v_t = \delta_t\varepsilon_t - 0.8894v_{t-1} - 0.4223v_{t-2} - 0.1911v_{t-3} + 0.4083v_{t-4} + 0.7512v_{t-5} - 0.03275v_{t-6} + 0.8711\varepsilon_{t-1} + 0.3863\varepsilon_{t-2} - 0.1319\varepsilon_{t-3} - 0.4618\varepsilon_{t-4} - 0.8014\varepsilon_{t-5} \quad (8)$$

$$\delta_t{}^2 = 1.838 * 10^{-5} + 0.1556v_{t-1}^2 + 0.8662\delta^2{}_{t-1} \quad (9)$$

For the residual analysis of the model, the residuals of the model δ_t fluctuate between 0.02 and 0.08, and the residuals of ε_t fluctuate almost uniformly around the value of 0. The p-value in the normality

ISSN 2616-5902 Vol. 4, Issue 7: 123-128, DOI: 10.25236/AJBM.2022.040720

test of the model (Jarque-Bera and Shapiro-Wilk) is less than the significance level of 0.05, so the residuals of the model do not obey the normal distribution, and its Q-Q plot shows that some of the data deviate from normality at both ends. In the white noise test for the residuals, the p-values (Ljung-Box) are all greater than the significant level of 0.05, indicating that it is white noise, but the ACF plot of the residuals also shows that the series correlates with about 3rd order. The white noise test of the squared residuals, which is also white noise, in the ACF plot of the squared residuals, shows that it is also partially correlated. The ARCH effect test of the residuals (LM Arch Test) has a p-value greater than 0.05, and the model residuals have no ARCH effect anymore. The model fits well except for normality and partial correlation which is not very good, so the model fits well in general.

4. Conclusion

Table 1: Specific models for each stage

Stage	Models
20021111 to 20071008	ARMA(0,1)+GARCH(1,1)
20080402 to 20100421	ARMA(1,1)+GARCH(1,1)
20120316 to 20141225	White noise sequence
20150519 to 20200724	ARMA(6,5)+GARCH(1,1)

The log-returns of this stock have been smooth overall, still fluctuating around the value. Before the 2015 crash, the log-returns were shown to be white noise series, which failed to reflect the conditional heteroskedasticity, probably due to the small size of the data, 484 data in this section. The other three stages of the fitted model can be seen in the mean information part of each stage when there is a difference, that is, the value of the stock is different at each stage, and we should take into account the change in the value of the stock itself when predicting the price of the stock. The GARCH part of the model is GARCH1(1,1), (Figure 8) which shows that the log-returns of this stock had approximately the same heteroskedasticity of volatility before and after the 2008 crash; and was still affected in the 2015 crash. Overall, the stock itself has heteroskedasticity, it has a stable heteroskedasticity to the stock market changes, and the fluctuations are less affected by the crash than the mean, which means that the value of the stock itself may be more affected by the crash. After several models were built, it was found that the model fit was quite good, but the only shortcoming was that the residuals of the model were not quite in line with the normal distribution, indicating that the model could not fully extract the information contained in the data, which will be a place for future improvement.

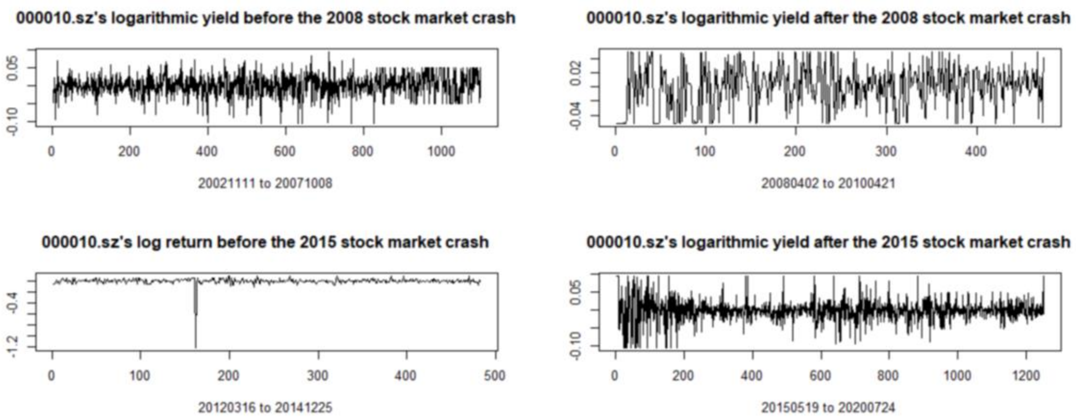

Figure 8: Time series of stock log returns for each stage

References

[1] Zhuang Jingpu. Research on talent demand and training matching prediction based on neural network SARIMA model [D]. Nanjing University of Posts and Telecommunications, 2020.
[2] Bollerslev Tim. Generalized autoregressive conditional heteroskedasticity [J]. North-Holland, 1986, 31(3).