Research on the Volatility Characteristics of Shanghai Stock Market Based on ARCH Model Family

Jie Wang
School of Finance, Anhui University of Finance and Economics, Bengbu 233030, Anhui, China

Abstract: In order to study the volatility characteristics of the Shanghai stock market, on the basis of reviewing the relevant literature, based on the ARCH model family, this paper selects 5797 data from the daily closing price of the Shanghai Stock Exchange from December 27, 1996 to November 30, 2020 to empirically analyze the volatility characteristics of the Shanghai stock market.

Keywords: Shanghai stock market, ARCH model family, Volatility, Information shock curve.

1. Introduction

All classical econometric models assume that the random disturbance terms have homoscedasticity, but the fluctuations of many financial time series have the phenomenon of "clustering", that is, the variance fluctuates less in one period of time and fluctuates more in another period of time. Obviously, for financial time series, the assumption that random disturbance terms have homoscedasticity is inappropriate. The basic idea of the ARCH model is that the random error term $u_t$ depends on its previous values $u_{t-1}, u_{t-2}, ...$, and its variance is conditional heteroscedasticity over time. And this conditional heteroscedasticity is a linear combination of the squares of the random error lags. However, in practical applications, high-order moving averages often appear, and there are higher restrictions on parameters, which increases the difficulty of model estimation. GARCH models reduce the requirements for parameters. Based on the ARCH model and the GARCH model, many derived GARCH models have emerged one after another, such as GARCH-M, TGARCH, EGARCH models, etc., forming the ARCH model family.

China's stock market has a history of 30 years since the establishment of the Shanghai and Shenzhen stock markets. During this period, there have been many sharp rises and falls, especially before and after the stock market crash in 2007 and 2015. The Chinese stock market has experienced huge fluctuations. The volatility characteristics of the stock market, as its basic characteristics and research starting point, have always been the hot and key issues of theoretical research and empirical analysis in the field of financial investment. At present, the methods and angles of studying the volatility of the stock market tend to be diversified. This paper studies the volatility characteristics of the Shanghai stock market based on the ARCH model family.

2. Literature Review

In recent years, the ARCH model family has been widely used in the financial field, especially a large number of domestic and foreign literatures study stock market volatility and stock market forecasting.

In terms of domestic related research, Tang and Chen (2001) analyzed the ARCH effect of the Shanghai and Shenzhen stock indexes and found that there is a leverage effect, the volatility is relatively concentrated, and it has a certain impact on the future. [1] Li et al. (2003) selected the Shanghai Stock Exchange 30 Index, the Shanghai Stock Exchange, the Shenzhen Stock Exchange and the Hong Kong Hang Seng Index to predict the volatility of the Chinese stock market. They pointed out that the E-GARCH model's ability to predict the volatility of the first three indexes is greater than that of the GARCH model and the TGARCH model, but the three models' ability to predict the volatility of the Hong Kong Hang Seng Index is almost the same. [2] Yan and Li (2008) found that the volatility of my country's stock market has no leverage effect, and the GARCH (1,1) model with Student's-t distribution has the best prediction effect. [3] Luo and Yang (2013) studied the asymmetry and risk premium of the Shanghai Composite Index, and found that the GARCH model can capture the conditional heteroscedasticity, the Shanghai stock market has a positive risk premium, and its volatility is asymmetric. [4] Wu (2014) studied the volatility of the returns of the Shanghai Stock Exchange and found that the EGARCH model is more suitable for describing the volatility of the stock market. [5] Zheng and Shang (2014) further use the mixed sampling GARCH-MIDAS model to describe the volatility of Chinese stock markets. [6] Zhu et al. (2015) combined the VaR model with the TGARCH model and the EGARCH model, and believed that the TGARCH (1,1) model and the EGARCH (1,1) model had better fitting effects on the Shenzhen Component Index and the Shanghai Composite Index. The volatility of the Shanghai stock market and the Shenzhen stock market has a leverage effect. [7]

According to relevant foreign studies, Pagan and Schwert (1990) study on the volatility of US stock returns shows that stock volatility is non-stationary, and the GARCH model has good fitting effect and prediction accuracy. But traditional ARCH or GARCH models cannot capture the nonlinear characteristics of stock returns. [8] Danielson (1994) used the maximum likelihood method to study the daily data of the S&P500 index and concluded that the EGARCH(2,1) model was better than the ARCH(5) and GARCH(1,2) models in fitting the data. [9] Awartani and Corradi (2005) used the GARCH model family to predict the S&P500 index, and found that the asymmetric GARCH model under asymmetric information is better than the GARCH (1,1) model. When there is no information asymmetry, the GARCH(2,1) model fits better. [10] In addition, Sabiruzzaman et al. (2010) studied the fluctuation law of the daily trading volume index of the
Hong Kong Stock Exchange and found that in the presence of asymmetric effects, the TGARCH model can describe the influence of positive and negative information respectively, and the fitting results are better than the GARCH model. The trading volume has obvious leverage effect, indicating that the TGARCH model is more suitable for describing the volatility of trading volume.[11] Girardin and Joyeux (2013), based on the GARCH-MIDAS model, proved that both Chinese A and B share markets are speculative. And pointed out that the macroeconomic fundamentals are increasingly affected by the volatility of the stock market, especially the CPI inflation.[12] Sharma and Vipul (2016) found that the Realized GARCH model had better predictive ability for stock market volatility.[13]

To sum up, in the past 20 years, a large number of scholars have used the ARCH model family to study the volatility characteristics of the stock market and to predict its volatility. Most studies show that the volatility of China's stock market is generally leveraged, clustered and persistent, but few authors conduct research from the perspective of risk premium, which is the new research perspective of this paper.

China's stock market is known as the "barometer" of the national economy. With the improvement of China's stock market, especially in recent years, a series of major reforms have been implemented, such as improving the stock delisting system, setting up the Science and Technology Innovation Board, and implementing the registration system on a trial basis, whether the volatility characteristics of China's stock market are consistent with the research results of a large number of scholars on the volatility characteristics of China's stock market in the early years is worth exploring. From the perspective of China's stock market, the listed companies on the Shanghai Stock Exchange have large market capitalization and large trading volume. The Shanghai Stock Exchange relatively more accurately reflects China's stock market conditions and is more sensitive to various shocks. Therefore, on the basis of reviewing the relevant literatures using the ARCH model family to study the stock market, this paper uses the ARCH model family to study and analyze the volatility characteristics of the Shanghai stock market, and increase the research perspective of risk premium. Use ordinary ARCH model, GARCH model, GARCH-M model, TGARCH model and EGARCH model to study the clustering, persistence, risk premium level and leverage effect of Shanghai stock market volatility step by step.

3. ARCH Model Family

3.1. ARCH Model

The expression of the ARCH model is:

\[
\begin{aligned}
\{y_t &= \beta_0 + \beta_1 y_{t-1} + \beta_2 y_{t-2} + \ldots + \beta_p y_{t-p} + u_t \\
\sigma_t^2 &= \alpha_0 + \alpha_1 u_{t-1}^2 + \alpha_2 u_{t-2}^2 + \ldots + \alpha_q u_{t-q}^2
\end{aligned}
\]

(1)

where \(\alpha_0 > 0, \alpha_1, \alpha_2, \ldots, \alpha_q \geq 0; \sum_{i=1}^q \alpha_i < 1, i = 1, 2, \ldots, q\)

The first formula is the mean value equation, and the second formula is the variance equation, and various ARCH models that appear later are extended on this basis. (1) is the ARCH(q) model.

The limitation of the basic ARCH model is that for the conditional variance \(\sigma_t^2\) to be always positive, \(\alpha_i\) must be non-negative. Since the variance equation often has a high lag order in practical applications, these constraints may not be met when estimating parameters. Since the variance equation in formula (1) is a distributed lag model of \(\sigma_t^2\), it is considered to replace the lag value of the random error square term with one or more lag orders of \(\sigma_t^2\) to reduce the number of parameters and relax the coefficient constraints. That is, the GARCH model is obtained.

3.2. GARCH Model

Based on the limitations of the ARCH model on parameter constraints, the ARCH model is improved, and the GARCH model after the generalization of the ARCH model reduces the requirements for parameters, and its expression is:

\[
\begin{aligned}
\{y_t &= \beta_0 + \beta_1 y_{t-1} + \beta_2 y_{t-2} + \ldots + \beta_p y_{t-p} + u_t \\
\sigma_t^2 &= \omega + \sum_{i=1}^p \varphi_i \sigma_{t-i}^2 + \sum_{j=1}^q \gamma_j u_{t-j}^2
\end{aligned}
\]

(2)

where \(\omega > 0 ; \varphi_i \geq 0 (i = 1, 2, \ldots, p)\) and \(\gamma_j \geq 0 (j = 1, 2, \ldots, q)\) m \(\sum_i \varphi_i + \sum_j \gamma_j < 1\).

This model is called the GARCH(P,q) model. In the variance equation, the \(u_{t-j}^2\) term is the ARCH term, reflecting the lag effect of previous information; the \(\sigma_{t-i}^2\) term is the GARCH term, that is, the previous conditional variance. The lag order of the GARCH model is reduced, the number of parameters is reduced, and the coefficient requirements are reduced.

3.3. GARCH-M Model

Financial markets generally believe that the risk of financial assets is proportional to the return. Risk premium refers to the high return required by investors to offset the high risk they take, and the variance or standard deviation of the index series can be used to characterize their risk level. Taking into account the conditional variance of returns to express expected risk and as an additional regressor yields an ARCH-M model (or an ARCH mean model). A more general GARCH-M model is used to describe the relationship between risk and return. Its expression is:

\[
\begin{aligned}
\{y_t &= \beta_0 + \beta_1 y_{t-1} + \beta_2 y_{t-2} + \ldots + \beta_p y_{t-p} + c \sigma_t^2 + u_t \\
\sigma_t^2 &= \omega + \sum_{i=1}^p \varphi_i \sigma_{t-i}^2 + \sum_{j=1}^q \gamma_j u_{t-j}^2
\end{aligned}
\]

(3)

where \(c \geq 0\); \(\varphi_i \geq 0 (i = 1, 2, \ldots, p)\) and \(\gamma_j \geq 0 (j = 1, 2, \ldots, q)\).\(\sum_i \varphi_i + \sum_j \gamma_j < 1\).

Where \(c\) represents the risk premium parameter. Generally, when \(c > 0\), the return of financial assets is proportional to its volatility. The typical GARCH(1,1)-M model is widely used to describe the relationship between the stock market's return and its volatility. Of course, the conditional variance \(\sigma_t^2\) in the model can also be replaced by its standard deviation \(\sigma_t\) or \(\ln(\sigma_t^2)\) after taking the logarithm.

3.4. TGARCH Model and EGARCH Model

Different types of events in the capital market have different degrees of impact on asset price changes. "Good news" and the same amount of "bad news" make assets rise and fall differently, that is, there is asymmetry in the impact of information, and this asymmetry is the leverage effect. The asymmetric information shock curve can test the leverage effect of capital market. In view of this situation, the GARCH model needs to be improved.
3.4.1. TGARCH Model
The general expression of the TGARCH model (or threshold GARCH model) is:

\[
\begin{align*}
    y_t &= \beta_0 + \beta_1 y_{t-1} + \beta_2 y_{t-2} + \cdots + \beta_p y_{t-p} + u_t \\
    \sigma_t^2 &= \omega + \sum_{i=1}^p \varphi_i u_{t-i}^2 + \sum_{k=1}^q \gamma_k \sigma_{t-k}^2 + \delta \sigma_{t-j}^2
\end{align*}
\] (4)

The general TGARCH(1,1) model variance equation is:

\[
\sigma_t^2 = \omega + \varphi_1 u_{t-1}^2 + \gamma \sigma_{t-1}^2 + \delta \sigma_{t-j}^2
\]

Among them, \( y \) is the leverage effect coefficient; \( d_{t-1} \) is a dummy variable, and the \( \gamma u_{t-1}^2d_{t-1} \) term is the leverage effect TGARCH term. When \( u_{t-1} > 0 \), \( d_{t-1} = 0 \), which will bring \( \alpha \) times the impact; when \( u_{t-1} < 0 \), \( d_{t-1} = 1 \). At this time, an impact of \( (\alpha + \gamma) \) times will be brought. When \( \gamma \neq 0 \), there is an asymmetric effect.

3.4.2. EGARCH Model
Compared with the TGARCH model whose leverage effect is quadratic, the EGARCH model has an exponential leverage effect. \( \ln(\sigma_t^2) \) is used to replace \( \sigma_t^2 \) on the left side of the equation of the variance equation of the model. Therefore, no matter whether the right side of the equation is positive or not, \( \sigma_t^2 \) can be guaranteed to be always positive, so there is no restriction on the model parameters. It makes parameter estimation more convenient and stable. The general expression for this model is:

\[
\begin{align*}
    y_t &= \beta_0 + \beta_1 y_{t-1} + \beta_2 y_{t-2} + \cdots + \beta_p y_{t-p} + u_t \\
    \ln(\sigma_t^2) &= \omega + \sum_{i=1}^p \varphi_i \ln(u_{t-i}^2) + \sum_{i=1}^q \gamma_i \ln(\sigma_{t-j}^2) + \sum_{j=1}^q \delta_j \ln(\sigma_{t-j}^2)
\end{align*}
\] (5)

The general EGARCH(1,1) model variance equation is:

\[
\ln(\sigma_t^2) = \omega + a \ln(\sigma_{t-1}^2) + \varphi \frac{u_{t-1}}{\sigma_{t-1}} + \gamma \frac{u_{t-1}}{\sigma_{t-1}}
\]

where \( \varphi \) is the leverage factor. When \( \varphi \neq 0 \), the impact of the shock has a leverage effect. When \( u_{t-1} > 0 \), the impact will be \( (\varphi + \gamma) \) times; when \( u_{t-1} < 0 \), \( (\varphi - \gamma) \) times the impact will be generated.

4. Empirical Analysis

4.1. Data Selection and Initial Model Setting
There is a strong correlation between the Shanghai and Shenzhen stock markets in China's stock market. The Shanghai Stock Exchange has a high stock market value and large trading volume, which more systematically and accurately reflects the state of the Chinese stock market. This paper selects 5,797 data of the daily closing price index of the Shanghai Stock Exchange and Shanghai Composite Index from December 27, 1996 to November 30, 2020 as sample observations. The data comes from the WIND database, and the measurement software is Eviews8.0. Use \( sp_t \) to represent the closing price index of the Shanghai Stock Exchange. In order to reduce the rounding error, the natural logarithm of \( sp_t \) is processed as \( \ln sp_t \). Since the stock price index sequence often follows a random walk process with drift, the following model is first established:

\[
\ln sp_t = \alpha + \rho \ln sp_{t-1} + \mu_t
\]

The OLS estimation results are as follows:

The statistics of the estimated results are all significant, and the fitting effect is very good, which further proves that the closing price index sequence of the Shanghai Stock Exchange conforms to the random walk model setting. The stationary ADF test was performed on the residuals of the fitted model, and the statistic was -12.1247, which were all less than the critical values at the 10%, 5%, and 1% significance levels, indicating that the residual sequence was stationary. Make a residual statistical plot (Figure 1) and a residual series plot (Figure 2).

Figure 1. Residual Statistical Chart

Figure 2. Residual sequence diagram
It can be seen from the residual statistics chart that the mean of the residual series is 0.0004 and the standard deviation is 0.0158, indicating that the returns of the Shanghai stock market are generally positive, but the volatility is relatively large. The skewness is -0.4049 and the kurtosis is 7.9930. Compared with the normal distribution, there is a left skewed peak and thick tail. It can be seen from the residual sequence diagram that the fluctuations are "clustered", with large fluctuations for a period of time and small fluctuations for a period of time, and there may be conditional heteroscedasticity. The ARCH-LM test was carried out, and the lag order was 21 by the AIC and SC criteria, and F=30.4993, T*R2=578.5274, and the null hypothesis was rejected, indicating that the residual sequence had an ARCH effect. The correlation coefficient and Q statistic are significant in the residual squared correlation test, which also indicates that the residual series has an ARCH effect.

Due to the high lag order of the ARCH model, it will lead to too many parameters, and it is difficult to meet the parameter constraints and the guarantee of significance, so the GARCH model is used to correct it.

4.2. Establish a GARCH Model

Build a GARCH(1,1) model and perform OLS estimation to get:

\[
\ln \sigma_t^2 = 2.23 \times 10^{-6} + 0.0863\sigma_{t-1}^2 + 0.9995\sigma_{t-1}^2 \\
Z = 8.5837 \quad 23.6857 \quad 279.7006 \\
R^2=0.9984 \quad DW=1.9773 \quad \text{log likelihood}=16603.00 \\
AIC=-5.7274 \quad SC=-5.7216
\]

The estimation results showed that each parameter was significant, the log-likelihood value increased, and the AIC and SC values decreased. According to the ARCH-LM test (the lag order is 1), F=2.6797, T*R2=2.6794, accepting the null hypothesis, and the residual series has no ARCH effect. The correlation coefficient and Q statistic in the residual squared correlation test chart are not significant, which also indicates that there is no ARCH effect in the residual series. The GARCH(1,1) model can eliminate the heteroscedasticity of the residuals, but the correlation coefficient and Q statistic in the residual correlation plot are significant, indicating that the residual series has autocorrelation.

The AR(4)-GARCH(1,1) model is established to get:

\[
\ln \sigma_t^2 = 0.0031 + 0.9996\ln \sigma_{t-1} + \mu_t \\
Z = 1.2366 \quad 3001.771 \\
\]

The AR(4)-GARCH(1,1) model without constant term is established to get:

\[
\ln \sigma_t^2 = 0.0004 + 1.0225\ln \sigma_{t-1} - 0.0491\ln \sigma_{t-2} \\
+ 0.0605\ln \sigma_{t-3} - 0.0344 \\
+ \ln \sigma_{t-4} + \mu_t \\
Z = 1.4701 \quad 77.9503 \quad -2.6326 \quad 3.0642 \quad -2.5927 \\
\sigma_t^2 = 2.11 \times 10^{-6} + 0.0863\sigma_{t-1}^2 + 0.9122\sigma_{t-1}^2 \\
Z = 8.4153 \quad 24.0560 \quad 300.4243 \\
R^2=0.9984 \quad DW=2.0188 \quad \text{log likelihood}=16607.06 \\
AIC=-5.7304 \quad SC=-5.7200
\]

Therefore, the model is adjusted, and the AR(4)-GARCH(1,1)-M model without constant term is established to obtain the result:

\[
\ln \sigma_t^2 = 1.0330\ln \sigma_{t-1} - 0.0690\ln \sigma_{t-2} \\
+ 0.0731\ln \sigma_{t-3} - 0.0372 \\
+ \ln \sigma_{t-4} + 3.9987\sigma_{t-1}^2 + \mu_t \\
Z = 86.1668 \quad -170.4723 \quad 6.0670 \quad -425.9912 \quad 0.0190 \\
\sigma_t^2 = 2.96 \times 10^{-6} + 0.0904\sigma_{t-1}^2 + 0.9010\sigma_{t-1}^2 \\
Z = 9.0911 \quad 23.7920 \quad 254.3556 \\
R^2=0.9984 \quad DW=2.0316 \quad \text{log likelihood}=16597.54 \\
AIC=-5.7274 \quad SC=-5.7182
\]

The estimated coefficient of \( \sigma_t^2 \) is 3.9987 and the Z statistic is significant, that is, when one more unit of risk is taken, 3.9987 more units of return will be obtained, indicating that the return and risk are positively correlated, which is in line with economic significance.

4.4. Establish TGARCH model and EGARCH Model

4.4.1. The Results of Establishing Ar(4)-Tgarch(1,1) Model Are as Follows:

\[
\ln \sigma_t = 0.0041 + 1.0201\ln \sigma_{t-1} - 0.0409 \\
+ \ln \sigma_{t-2} + 0.0555\ln \sigma_{t-3} \\
- 0.0351\ln \sigma_{t-4} + \mu_t \\
Z = 1.4362 \quad 77.0970 \quad -2.1575 \quad 82.7774 \quad -2.6392 \\
\sigma_t^2 = 2.23 \times 10^{-6} + 0.0707\sigma_{t-1}^2 + 0.9112\sigma_{t-1}^2 \\
+ 0.0259\sigma_{t-1}^2 \\
Z = 8.7437 \quad 15.5398 \quad 300.5271 \quad 4.6368 \\
R^2=0.9984 \quad DW=2.0188 \quad \text{log likelihood}=16607.06 \\
AIC=-5.7304 \quad SC=-5.7200
\]

The estimated values of each parameter are significant, and the fitting effect is good. The coefficient of the asymmetric effect term is 0.0259>0, indicating that there is a leverage effect in the volatility of the Shanghai stock market. When there is "good news", the external shock will only bring 0.0707 times the shock to the Shanghai stock market; when there is "bad news", the external shock will bring 0.0707+0.0259=0.0966 times the shock to the Shanghai stock market.

4.4.2. The Results of Establishing the AR(4)-EGARCH(1,1) Model Are As Follows:

\[
\ln \sigma_t = 0.0019 + 1.0117\ln \sigma_{t-1} - 0.0250 \\
+ \ln \sigma_{t-2} + 0.0460\ln \sigma_{t-3} \\
- 0.0329\ln \sigma_{t-4} + \mu_t \\
Z = 0.8162 \quad 229.4455 \quad -5.2007 \quad 8.7823 \quad -8.9596 \\
\ln(\sigma_t^2) = -0.2775 + 0.1935\mu_{t-1} - 0.0243\mu_{t-1} \\
+ 0.9842ln(\sigma_{t-1}^2) \\
Z = -18.3175 \quad 27.5844 \quad -5.9184 \quad 592.8262 \\
R^2=0.9984 \quad DW=2.0012 \quad \text{log likelihood}=16633.64 \\
AIC=-5.7396 \quad SC=-5.7292
\]

The estimated values of each parameter are very significant, and the model fitting effect is very good. Since the coefficient of the asymmetric term is -0.0243<0, the impact of the shock has an asymmetric effect. When \( \mu_{t-1} > 0 \), that is, "good news", it will produce 0.1935-0.0243=0.1692 times the impact. When \( \mu_{t-1} < 0 \), that is, "bad news" occurs, there will be a shock of 0.1935+0.0243=0.2178 times.

By making a graph of information shock (Figure 3), the asymmetric impact of information shock can be seen intuitively. Make:
47

47

It can be seen that when a negative shock occurs, the curve is steeper; when a positive shock occurs, the curve is relatively flat, showing asymmetry. It shows that the negative shock of "bad news" has a greater impact on the volatility of the Shanghai stock market than the positive shock of "good news", and the Shanghai stock market has a significant leverage effect.

5. Conclusion

This paper selects 5,797 data of the daily closing price index of the Shanghai Stock Exchange and Shanghai Composite Index from December 27, 1996 to November 30, 2020 as sample observations to study the volatility characteristics of the Shanghai stock market, and draws the following conclusions:

First, there is an ARCH effect in the Shanghai stock market, and the residual sequence of the first-order random walk process model has the characteristics of sharp peaks, thick tails, non-normal distribution, and the phenomenon of "clustering" of fluctuations. The GARCH(1,1) model can capture the conditional heteroskedasticity of the Shanghai stock market, and the AR(4)-GARCH(1,1) model can eliminate autocorrelation, which is better than describing the volatility of the Shanghai stock market.

Second, the AR(4)-GARCH(1,1)-M model without constant term can better describe the positive risk premium of the Shanghai stock market. When one more unit of risk is taken in the stock market, an additional return of 3.9987 units will be obtained.

Third, the empirical analysis of AR(4)-TGARCH(1,1) model and AR(4)-EGARCH(1,1) model proves that the Shanghai stock market has leverage effect. The negative shock of "bad news" has a greater impact on the volatility of the Shanghai stock market than the positive shock of "good news".

Finally, in the results of the model estimates in the empirical part, the coefficients of the ARCH term and the GARCH term are both positive, and the sum of the two coefficients is less than 1 and close to 1, indicating that the volatility of the Shanghai stock market in the past has a continuous impact on the future.

Acknowledgment

Funding project: Anhui University of Finance and Economics Graduate Research Innovation Fund Project (ACYC2020121).

References