ISSN: 2688-9323

Analysis of Stock Price Volatility based on GARCH Family Models

Xuewen Lu

Statistics and Applied Mathematics Institute, Anhui University of Finance and Economics, Bengbu 233000, China

Abstract

This article uses the GARCH family model to analyze the volatility of the stock market based on the Shanghai and Shenzhen 300 Index, making investor decision-making strategies more accurate and providing guidance. Multiple GARCH family models were established and the significance test results were compared. The results showed that the asymmetric ARIMA-EGARCH model had a better fitting effect on the Shanghai and Shenzhen 300 Index, indicating that the volatility of the Shanghai and Shenzhen 300 Index exhibited sharp peaks, thick tails, and asymmetric characteristics, indicating the existence of leverage effect in China's stock market.

Keywords

CSI 300 Index; GARCH Family Model; Volatility Asymmetry.

1. Introduction

In the more than 20 years of development of China's stock market, the volatility of the stock market has been significantly influenced by policies. Since the equity split reform in May 2005, China's Shanghai and Shenzhen stock markets have experienced unprecedented periods of bull and bear markets. Therefore, this article selects the daily closing prices of the Shanghai and Shenzhen stock markets from January 1, 2000 to May 21, 2022 as the research sample, takes the daily returns of the Shanghai and Shenzhen stock markets as the research object, and uses GARCH family models for empirical analysis to study models that are more suitable for fitting the volatility asymmetry of the Shanghai and Shenzhen stock markets.

Due to the heteroscedasticity and volatility clustering effects of financial product yield sequences, Engle[1] proposed the Autoregressive Conditional Heteroscedasticity (ARCH) model. Bollerslev[2] improved the ARCH model by proposing the Generalized Autoregressive Conditional Heteroscedasticity (GARCH) model, which has been widely used. Xu Lixia[3] applied the GARCH family model to the study of volatility in the Chinese stock market, using GARCH, TGRCH, and EGARCH models to fit the volatility of the Chinese stock market. The results showed that the stock return series had the characteristics of sharp peaks, thick tails, and volatility clustering, and the GARCH family model could well fit the volatility of the stock market.

2. Empirical Analysis

Establishing GARCH family models for financial time series should ensure the stationarity of the time series. Therefore, when establishing the mean equation for the return series of the Shanghai and Shenzhen stock indices, the first step is to conduct a stationarity test; Secondly, make model judgments based on the results of automatic recognition by the mean model; Once again, perform conditional heteroscedasticity tests on the residual sequences of the established mean equations for the Shanghai and Shenzhen stock indices. If the test results show the presence of ARCH effects, an ARCH class model can be established to fit the stock index return series and reflect the volatility of the stock index return series. If the model fits well, the ARCH effect of the error term can be eliminated.

ISSN: 2688-9323

2.1. Data Sources and Descriptive Analysis

This article selects the daily closing prices of the Shanghai and Shenzhen 300 Index from January 1, 2002 to May 21, 2022, totaling 4942 trading days. The data is sourced from the NetEase Finance platform. The yield is calculated using the "closing price closing price" method and logarithmically, in the form of $r_t = \ln P_t - \ln P_{t-1}$.

Among them, P_t is current closing price, P_{t-1} is previous closing price.

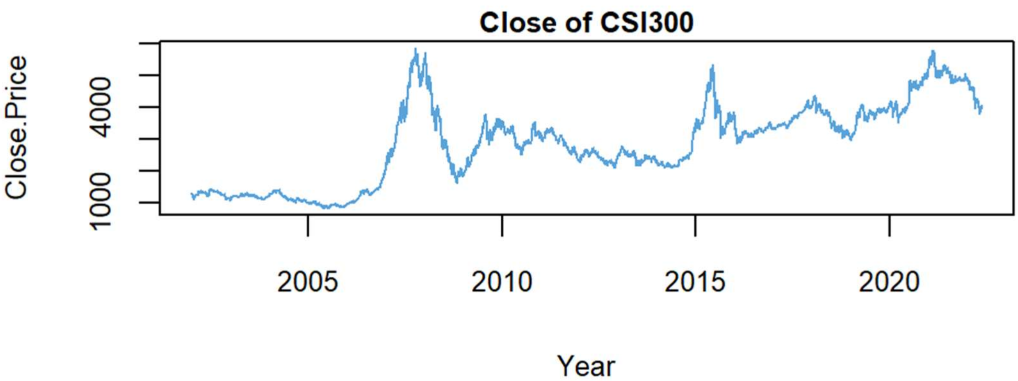

Figure 1. Closing Price Time Series Chart

From Figure 1, it can be seen that the closing price experienced two significant fluctuations around 2008 and 2015. From Figure 2, it can be seen that there is a pattern of clustering in the return rate, where there is a large fluctuation followed by a small one. Moreover, the yield is mostly close to 0, and from a long-term perspective, the rise and fall offset each other.

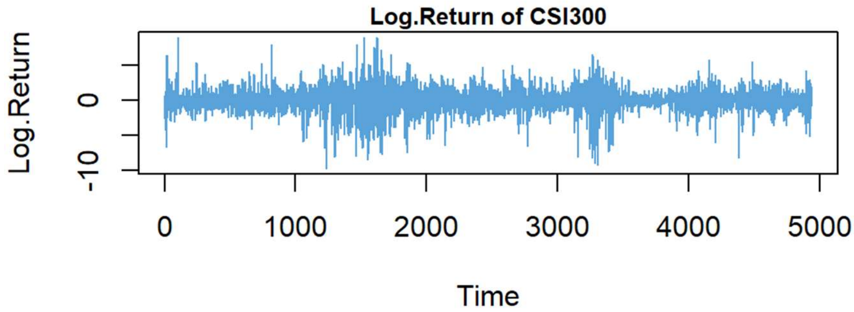

Figure 2. Logarithmic Return Time Series Chart

2.2. Stationarity Test

The Augmented Dickey Fuller Test (ADF) is an extended form of the DF test, which can perform unit root tests on sequences with high-order lag. The original assumption was that there is a unit root, which means the sequence is non-stationary. This article uses the adf. test() function for unit root testing, and the test results are shown in Figure 3.

```
> show(adf.test(r.data))

        Augmented Dickey-Fuller Test

data:  r.data
Dickey-Fuller = -14.965, Lag order = 17, p-value = 0.01
alternative hypothesis: stationary
```

Figure 3. Stability Test Results

From Figure 3, it can be seen that the ADF test result is p-value equal to 0.01, less than 0.05, indicating that the null hypothesis can be rejected, that is, the sequence is stationary.

2.3. Mean Model Recognition

After the sequence is tested as a stationary sequence, the auto.arima() function is first used to automatically identify the mean model of the sequence. The model results are shown in Figure 4.

```
> md1 <- auto.arima(r.data)
> md1
Series: r.data
ARIMA(4,0,4) with zero mean

Coefficients:
         ar1      ar2      ar3     ar4      ma1     ma2     ma3      ma4
      0.2464  -0.0749  -0.0151  0.8026  -0.2337  0.0449  0.0558  -0.7978
s.e.  0.0895   0.0991   0.0825  0.0777   0.0863  0.0978  0.0781   0.0750

sigma^2 = 2.655:  log likelihood = -9419.59
AIC=18857.18   AICc=18857.22   BIC=18915.73
```

Figure 4. Mean Model Results

From Figure 4, it can be seen that the identified mean model is ARMA (4,4), and then the parameter significance test is performed on the ARMA (4,4) model. The test results are shown in Figure 5.

```
Table: Result of Coef. Test

|   |        ar1|        ar2|        ar3| ar4|        ma1|        ma2|       ma3| ma4|
|:--|---------:|---------:|---------:|---:|---------:|---------:|--------:|---:|
|p  | 0.0059107| 0.4497863| 0.8546559|   0| 0.0067783| 0.6460047| 0.475365|   0|
```

Figure 5. Results of Parameter Significance Test

From Figure 5, it can be seen that only the coefficients ar1, ar4, ma1, and ma4 passed the significance test. Therefore, a mean model of sparse coefficients was established, and the coefficients of ar2, ar3, ma2, and ma3 were set to 0.

2.4. Establish Variance Model

2.4.1. ARCH Effect Test

After establishing the above model, perform ARCH effect testing on the residuals. The Ljung Box statistic Q (m) can perform autocorrelation tests on residual sequences. The original assumption is that there is no autocorrelation in the sequence, and conditional heteroscedasticity can be tested in the squared residual sequence. Use the archTest() function in the MTS package for verification, and the verification results are shown in Figure 6.

```
> print(paste0('m = 10'))
[1] "m = 10"
> archTest(mean_md_1$res, lag = 10)
Q(m) of squared series(LM test):
Test statistic:  1142.435  p-value:  0
Rank-based Test:
Test statistic:  876.0903  p-value:  0
> print(paste0('m = 20'))
[1] "m = 20"
> archTest(mean_md_1$res, lag = 20)
Q(m) of squared series(LM test):
Test statistic:  1857.586  p-value:  0
Rank-based Test:
Test statistic:  1569.97  p-value:  0
```

Figure 6. Results of parameter significance test

ISSN: 2688-9323

The test results in Figure 6 show that there is autocorrelation in the residual sequences with a lag of 10 and 20 orders, thus rejecting the null hypothesis, indicating the presence of ARCH effect in the residual sequences.

2.4.2. Establish a Standard GARCH Model

The above ARCH effect indicates that conditional variance depends on past values. Therefore, the GARCH model can be considered for parameter estimation of the variance equation. The fitting results using the GARCH model are shown in Figure 7.

From the model results in Figure 7, it can be seen that the fitted GARCH model parameters are all significant, and the P-value in the Box Ljung test result is greater than significance. Therefore, it can be considered that the residual of the model is not sequence correlated, indicating that the model has a good fitting effect. The Jarque Bera Test is used to test the normality of regression residuals. The original assumption was that the residual sequence follows a normal distribution, but in reality, the p-value of the test result is very small, indicating that the residual sequence does not follow a normal distribution. Therefore, the model can be optimized to consider other GARCH models.

```
Call:
garch(x = r.data, order = c(1, 1))

Model:
GARCH(1,1)

Residuals:
     Min       1Q   Median       3Q      Max 
-6.57064 -0.53403  0.04164  0.58668  5.54217 

Coefficient(s):
    Estimate  Std. Error  t value Pr(>|t|)    
a0  0.016773    0.002949    5.688 1.28e-08 ***
a1  0.073365    0.003815   19.228  < 2e-16 ***
b1  0.923517    0.003599  256.615  < 2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

      Diagnostic Tests:
              Jarque Bera Test

      data:  Residuals
      X-squared = 941.87, df = 2, p-value < 2.2e-16

              Box-Ljung test

      data:  Squared.Residuals
      X-squared = 0.10455, df = 1, p-value = 0.7464
```

Figure 7. GARCH Model Results

2.5. Model Optimization

2.5.1. Fitting

The above model considers the standard GARCH (1,1) model of ARIMA (0,0,0), where the parameters of the mean model are all set to 0. From the analysis of the mean model, it can be fitted with the mean model of ARIMA (4,0,4) and the variance model of GARCH (1,1). The normality test results in Figure 7 indicate that the distribution of residuals is not suitable for standard normal distribution, and other types of distributions should be considered.

For the mean model, consider ARIMA without intercept terms (4,0,4); For the variance model, the order is set to 1st order ARCH and 1st order GARCH, considering five types of models: standard GARCH (sGARCH), exponential GARCH (e-GARCH), GJR GARCH, threshold GARCH (TGARCH), and nonlinear asymmetric GARCH (NAGARCH); For residual distribution types, consider five types of distributions: standard normal distribution (norm), standard t-distribution (std), partial t-distribution (sstd), generalized error distribution (ged), and Johnson's SU distribution (jsu). The fitting results of each model are shown in Figure 8.

ISSN: 2688-9323

From the model fitting results in Figure 8, it can be seen that the asymmetric exponential GARCH model, i.e. the eGARCH model, can achieve the maximum likelihood estimate while minimizing both AIC and BIC, indicating that the eGARCH model has a good fitting effect. From the fitting of the residual distribution of the model, it can be seen that the AIC and BIC of the non normal distribution are significantly lower than those of the normal distribution, indicating that the residual follows a heavy tailed distribution.

Table: Result of GARCH Model

ModelName	LogL	AIC	BIC	RMSE
ARMA(4,4)-sGARCH-norm	-8766.405	3.552886	3.567369	1.631544
ARMA(4,4)-sGARCH-std	-8598.392	3.485283	3.501082	1.632356
ARMA(4,4)-sGARCH-sstd	-8595.592	3.484555	3.501670	1.632005
ARMA(4,4)-sGARCH-ged	-8591.445	3.482471	3.498270	1.632787
ARMA(4,4)-sGARCH-jsu	-8592.432	3.483275	3.500391	1.632515
ARMA(4,4)-eGARCH-norm	-8752.048	3.547479	3.563279	1.630043
ARMA(4,4)-eGARCH-std	-8645.518	3.504763	3.521879	1.641222
ARMA(4,4)-eGARCH-sstd	-8582.698	3.479740	3.498173	1.633551
ARMA(4,4)-eGARCH-ged	-8577.664	3.477298	3.494414	1.633429
ARMA(4,4)-eGARCH-jsu	-8577.046	3.477452	3.495885	1.632426
ARMA(4,4)-gjrGARCH-norm	-8763.240	3.552010	3.567809	1.631199
ARMA(4,4)-gjrGARCH-std	-8592.096	3.483139	3.500255	1.632771
ARMA(4,4)-gjrGARCH-sstd	-8589.786	3.482609	3.501041	1.632229
ARMA(4,4)-gjrGARCH-ged	-8586.670	3.480943	3.498059	1.632961
ARMA(4,4)-gjrGARCH-jsu	-8586.271	3.481186	3.499619	1.632296
ARMA(4,4)-TGARCH-norm	-8760.024	3.550708	3.566507	1.630899
ARMA(4,4)-TGARCH-std	-8586.685	3.480949	3.498065	1.632991
ARMA(4,4)-TGARCH-sstd	-8584.250	3.480368	3.498801	1.632423
ARMA(4,4)-TGARCH-ged	-8582.256	3.479156	3.496272	1.633009
ARMA(4,4)-TGARCH-jsu	-8580.847	3.478991	3.497423	1.632513
ARMA(4,4)-NAGARCH-norm	-8763.848	3.552256	3.568055	1.630683
ARMA(4,4)-NAGARCH-std	-8591.530	3.482910	3.500026	1.633007
ARMA(4,4)-NAGARCH-sstd	-8589.093	3.482329	3.500761	1.632431
ARMA(4,4)-NAGARCH-ged	-8587.392	3.481235	3.498351	1.631752
ARMA(4,4)-NAGARCH-jsu	-8587.138	3.481537	3.499970	1.632835

Figure 8. GARCH Family Model Fitting Results

2.5.2. Model Comparison

Table 1. Significance Test Results of eGARCH Model Coefficients

	norm	snorm	std	sstd	ged	sged	jsu
ar1	0.0035	0.0000	0.0000	0.0134	0.0000	0.0000	0.0000
ar2	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
ar3	0.0000	0.0000	0.0000	0.0033	0.0000	0.0001	0.0000
ar4	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
ma1	0.3627	0.0008	0.0000	0.0000	0.0000	0.0000	0.0000
ma2	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
ma3	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
ma4	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
omega	0.0000	0.0000	0.0000	0.0119	0.0000	0.0000	0.0000
alpha1	0.0999	0.0865	0.0162	0.1293	0.0053	0.0076	0.0038
beta1	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
gamma1	0.0000	0.0000	0.0000	0.2643	0.0000	0.0000	0.0000
skew	NA	0.0000	NA	0.0000	NA	0.0000	0.0175
shape	NA	NA	0.0000	0.0000	0.0000	0.0000	0.0000

ISSN: 2688-9323

Select the eGARCH model and compare the parameter significance, test results, and model performance of different distributions. Here, the distributions consider normal distribution, t-distribution, generalized error distribution, skewed distribution corresponding to the three distributions, and Johnson's SU distribution. The results are shown in Table 1, Table 2, and Figures 9 to 11.

Table 2. Individual Test Results for Coefficient Stability of eGARCH Model

	norm	snorm	std	sstd	ged	sged	jsu
ar1	0.0998	0.0952	0.1063	0.5034	0.1915	0.0877	0.2373
ar2	0.0361	0.0354	0.1099	0.4957	0.1030	0.5536	0.0425
ar3	0.1148	0.1132	0.1086	0.4575	0.0388	0.0732	0.0820
ar4	0.1802	0.1775	0.1059	0.5869	0.1141	0.7737	0.0890
ma1	0.0964	0.0904	0.1677	0.5926	0.1097	0.1036	0.1327
ma2	0.0318	0.0312	0.1741	0.6109	0.0538	0.6281	0.0381
ma3	0.1338	0.1317	0.1765	0.5533	0.0603	0.0744	0.0587
ma4	0.1857	0.1846	0.1649	0.7701	0.0514	0.8677	0.0756
omega	0.1707	0.1926	0.7501	0.3472	0.3056	0.3014	0.3375
alpha1	0.0846	0.1059	0.1517	0.0929	0.0877	0.0978	0.0911
beta1	0.1690	0.1725	0.7427	0.3310	0.2876	0.2935	0.3214
gamma1	0.3755	0.3475	0.3544	0.1721	0.1963	0.1842	0.1590
skew	NA	0.7982	NA	0.4472	NA	0.2461	0.5897
shape	NA	NA	2.6254	0.1144	0.0821	0.1025	0.1053
10%	0.3530	0.3530	0.3530	0.3530	0.3530	0.3530	0.3530
5%	0.4700	0.4700	0.4700	0.4700	0.4700	0.4700	0.4700
1%	0.7480	0.7480	0.7480	0.7480	0.7480	0.7480	0.7480

Table: P-value Table of Nyblom Joint Stability Test

	JoinStat	10%	5%	1%
ARMA(4,4)-eGARCH-norm	1.3582	2.69	2.96	3.51
ARMA(4,4)-eGARCH-snorm	2.1712	2.89	3.15	3.69
ARMA(4,4)-eGARCH-std	28.0517	2.89	3.15	3.69
ARMA(4,4)-eGARCH-sstd	4.9015	3.08	3.34	3.90
ARMA(4,4)-eGARCH-ged	2.6540	2.89	3.15	3.69
ARMA(4,4)-eGARCH-sged	2.4870	3.08	3.34	3.90
ARMA(4,4)-eGARCH-jsu	3.0064	3.08	3.34	3.90

Figure 9. Joint test results for the stability of eGARCH model numbers

Table: P-value Table of Sign Bias Test

	Sign Bias	Negative Sign Bias	Positive Sign Bias	Joint Effect
ARMA(4,4)-eGARCH-norm	0.2316	0.6830	0.0276	0.1688
ARMA(4,4)-eGARCH-snorm	0.2277	0.6522	0.0295	0.1743
ARMA(4,4)-eGARCH-std	0.1190	0.4032	0.0300	0.1437
ARMA(4,4)-eGARCH-sstd	0.1225	0.4840	0.0546	0.2306
ARMA(4,4)-eGARCH-ged	0.4945	0.7910	0.0825	0.3519
ARMA(4,4)-eGARCH-sged	0.0843	0.4390	0.0309	0.1441
ARMA(4,4)-eGARCH-jsu	0.3876	0.6496	0.1032	0.4045

Figure 10. EGARCH Model Symbol Bias Test Results

Table: P-value Table of Goodness-of-fit

	20	30	40	50
ARMA(4,4)-eGARCH-norm	0.0000	0.0000	0.0000	0.0000
ARMA(4,4)-eGARCH-snorm	0.0000	0.0000	0.0000	0.0000
ARMA(4,4)-eGARCH-std	0.0001	0.0002	0.0005	0.0018
ARMA(4,4)-eGARCH-sstd	0.0014	0.0034	0.0013	0.1502
ARMA(4,4)-eGARCH-ged	0.3229	0.4790	0.2252	0.2313
ARMA(4,4)-eGARCH-sged	0.3308	0.3307	0.7597	0.4388
ARMA(4,4)-eGARCH-jsu	0.0052	0.0098	0.0061	0.0226

Figure 11. The Goodness of Fit Test Results of the eGARCH Family Model

From the significance test table of eGARCH model coefficients in Table 1, it can be seen that the fitted coefficients are generally significant, with only a few parameters not significant. From the individual tests for coefficient stability in Table 2 and the joint tests for coefficient stability in Figure 9, it can be seen that at a significance level of 5%, normal distribution, partial normal distribution, generalized error distribution, and partial generalized error distribution all accept the null hypothesis that the parameters are stable. From the results of the symbol bias test in Figure 10, it can be seen that the difference between the positive and negative residuals of the eGARCH model under impact is not significant, indicating that the asymmetric model effectively eliminates the leverage effect. From the results of Pearson's goodness of fit test in Figure 11, it can be seen that the original hypothesis is that the residual distribution is not different from the theoretical distribution. The results indicate that the generalized error distribution and the partial generalized error distribution cannot reject the original hypothesis, indicating that these two distributions are well suited to the model.

2.5.3. Model Selection

The information criteria table of the eGARCH model is shown in Figure 12.

Table: Table of Information Criteria

ModelName	Akaike	Bayes	Shibata	HQ	LLH
ARMA(4,4)-eGARCH-norm	3.5475	3.5633	3.5475	3.553	-8752.0479
ARMA(4,4)-eGARCH-snorm	3.5452	3.5623	3.5452	3.5512	-8745.3951
ARMA(4,4)-eGARCH-std	3.5048	3.5219	3.5047	3.5108	-8645.518
ARMA(4,4)-eGARCH-sstd	3.4797	3.4982	3.4797	3.4862	-8582.6981
ARMA(4,4)-eGARCH-ged	3.4773	3.4944	3.4773	3.4833	-8577.6641
ARMA(4,4)-eGARCH-sged	3.4756	3.494	3.4755	3.482	-8572.3474
ARMA(4,4)-eGARCH-jsu	3.4775	3.4959	3.4774	3.4839	-8577.046

Figure 12. GARCH Model Information Criteria Table

From Figure 12, it can be seen that the LLH (-8572.3474) of the partial generalized error distribution reaches its maximum while its HQ (3.482) reaches its minimum, indicating that the ARMA (4,4) - eGARCH (1,1) - SGED model is optimal. Therefore, the ARMA (4,4) - EGARCH (1,1) - SGED model can be fitted to the Shanghai and Shenzhen 300 index data selected in this article. The theoretical model and model parameter estimation and significance are shown below.

Theoretical model:

$$\begin{cases} y_t = x_t\rho + \mu_t, \ \ \mu_t \sim SGED(0,1,shape,skew) \\ z_t = \dfrac{\mu}{\sigma_t} \\ \ln\left(\sigma_t^{\ 2}\right) = \omega_0 + \alpha z_{t-1} + \gamma\left(\dfrac{|\mu_{t-1}|}{\sigma_{t-1}} - E|z_{t-1}|\right) + \beta \ln \sigma^2_{\ t-1} \end{cases}$$

The estimation and significance of model parameters are shown in Figure 13.

Table: Table of Final Model Coef.

	Estimate	Std. Error	t value	Pr(>|t|)
ar1	-0.1760084	0.0321576	-5.473307	0.0000000
ar2	0.3269869	0.0254178	12.864488	0.0000000
ar3	0.1945664	0.0487647	3.989899	0.0000661
ar4	0.6507676	0.1064380	6.114051	0.0000000
ma1	0.1743615	0.0331571	5.258649	0.0000001
ma2	-0.3425918	0.0661205	-5.181325	0.0000002
ma3	-0.1703474	0.0365798	-4.656868	0.0000032
ma4	-0.6240696	0.0006391	-976.483873	0.0000000
omega	0.0084267	0.0018795	4.483539	0.0000073
alpha1	-0.0265205	0.0099384	-2.668487	0.0076194
beta1	0.9879399	0.0005464	1807.953295	0.0000000
gamma1	0.1539949	0.0179114	8.597598	0.0000000
skew	0.9478213	0.0500413	18.940763	0.0000000
shape	1.2083808	0.0359432	33.619181	0.0000000

Figure 13. Results of ARMA (4,4) - EGARCH (1,1) - SGED Model

3. Conclusion

This article analyzes the volatility of the Shanghai and Shenzhen 300 Index and finds that there have been two periods of significant volatility in China's stock market. The first fluctuation occurred around 2008, during which the global financial crisis occurred. It is evident that the closing price time series chart showed a steep peak with a longer duration; The second wave occurred around 2015, which was due to the addition of leveraged funds and policy tightening, resulting in brief bull and bear markets. The level of volatility was no less than that of the 2008 financial crisis, but the duration was relatively short.

This article compares multiple GARCH models and finds that the asymmetric ARIMA-EGARCH model has the best fitting effect on the logarithmic return of the Shanghai and Shenzhen 300 Index. At the same time, the partial generalized error distribution is close to the theoretical distribution, indicating that the volatility of the Shanghai and Shenzhen 300 Index exhibits sharp peaks, thick tails, and asymmetric characteristics, indicating the existence of leverage effect in China's stock market.

References

[1] ENGLE R: Autoregressive conditional heteroskedasticity with estimates of the variance of U.K. inflation, Econometrica, vol.50(1982)No.4,p.987-1008.

[2] BOLLERSLEV T: Generalized autoregressive conditional heteroskedasticity, Journal of Eccnometrics, vol.31(1986)No.3,p.307-327.

[3] Xu Lixia: Research on the Application of Volatility Models in the Chinese Stock Market, Statistics and Decision Making, vol.12(2010)p.168-170.