

The Volatility Research in CSI 300 Index Futures by Using High Frequency Data based on GARCH Model

Junbo Wang^a, Susheng Wang, Yongbo Kang

School of economics and management, Harbin Institute of Technology, Shenzhen, China

^a407995864@qq.com

Abstract: By using the high frequency data in CSI 300 index futures, we research the pattern of CSI 300 index futures, and find that the returns of high frequency CSI 300 index futures have obvious volatility clustering effect, and there is peak and fat-tailed phenomenon. After that, the ARCH effect test is carried out on the residual error of the high frequency data, and the results show that the residuals have obvious ARCH effect. After eliminating the autocorrelation, the optimal GARCH model is established, and the adequacy of the model fit was verified. The fitting results show that GARCH can well describe the characteristics of high frequency volatility of CSI 300 index futures, the impact on the conditional variance, has a strong persistence, long memory effect is found in the volatility.

Keywords: High frequency data, GARCH model, CSI 300 index futures, volatility characteristics.

1. Introduction

Volatility is an important indicator of the capital market. The study of volatility has always been an important part of the micro market research. In order to fully describe the volatility of the return on assets, in 1982, Engle [1] proposed autoregressive conditional heteroscedasticity (ARCH) model, which provides a systematic framework for the study of volatility. The statistical characteristics of the ARCH model and the accurate description of the volatility are widely used in the regression analysis and prediction of financial time series data. ARCH model, has experienced two development, one is T.Bollerslev (1986) [2] generalized vector autoregressive model (GARCH) is proposed. And then, based on the GARCH model, ARCH model has made new achievements more; another one is that the long memory model has made a breakthrough in economics. Here, we will focus on the research of GARCH model. The GARCH model provides a more flexible structure of the lag, the ARCH model of correlation coefficient of fading and usually only GARCH (1, 1) model, compared with the ARCH model, which greatly reduces the parameter estimation. Bollerslev and Engle (1986) [3] based on GARCH model is proposed in a short memory between GARCH model and GARCH model IGARCH (long discussion unit root GARCH model) model. Chou, R.F (1988) [4] establishes a GARCH-M model for the phenomenon that the rate of return depends on its volatility, Curto (2009) [5] with the error distribution as normal, stationary and t distribution of the ARMA-GARCH model for the stock market research, GARCH model of a variety of development, but the foundation is inseparable from the basic model of GRACH.

Domestic aspects, most of the articles about vector autoregressive model are focused on the empirical test and application. Hui Jun, Zhu Cui (2010) [6] use the ARMA-ARCH model for empirical analysis on the fund market to solve the problem of financial volatility time series. From the previous studies, the research and application of GARCH model in the CSI 300 index futures is relatively mature, can really describe the volatility of the yield.

2. Method

A process is called a GATCH (p,q) process if its first two conditional moment exist and satisfy:

$$E(\epsilon_t | \epsilon_{\mu}, \mu < t), t \in Z$$

There exist constant $\omega, \alpha_i, i = 1, \dots, q$ and $\beta_j, j = 1, \dots, p$ such that:

$$\sigma_t^2 = \text{Var}(\epsilon_t | \epsilon_{\mu}, \mu < t) = \omega + \sum_{i=1}^q \alpha_i \epsilon_{t-i}^2 + \sum_{j=1}^p \beta_j \sigma_{t-j}^2, t \in Z \quad (1)$$

It can be written in a more compact way as:

$$\sigma_t^2 = \omega + \alpha(B)\epsilon_{t-i}^2 + \beta(B)\sigma_{t-j}^2, t \in Z \tag{2}$$

Where B is the standard backshift operator, and α and β are polynomials of degrees q and p, respectively:

$$\alpha(B) = \sum_{i=1}^q \alpha_i B^i, \beta(B) = \sum_{j=1}^p \beta_j B^j \tag{3}$$

If $\beta(z) = 0$ we can get ARCH (q) process.

$$\sigma_t^2 = \omega + \sum_{i=1}^q \alpha_i \epsilon_{t-i}^2 \tag{4}$$

By definition, the innovation of the process ϵ_t^2 is the variable $v_t = \epsilon_t^2 - \sigma_t^2$, Substituting in (1), we get the representation:

$$\epsilon_t^2 = \omega + \sum_{i=1}^r (\alpha_i + \beta_j) \epsilon_{t-i}^2 + v_t - \sum_{j=1}^p \beta_j v_{t-j}, t \in Z \tag{5}$$

Where r is max (p,q), with $\alpha_i = 0$ ($\beta_j = 0$) if $i > q$ ($j > p$). Under additional assumptions, we can state that if (ϵ_t) is GARCH (p, q), then (ϵ_t^2) is an ARMA(r, p) process. In particular, the square of an ARCH (q) process admits, if it is stationary, an AR (q) representation. The ARMA representation will be useful for the estimation and identification of GARCH processes

3. Empirical

In this paper, one minute high frequency data from 2013-1-4 to 2013-4-9 is selected as the study sample, these samples select one minute high frequency data to describe the pattern of CSI 300 index future. After excluding the turnover no change data, we get a total of 16470 sample records. All the data come from the China Financial Futures Exchange. Suppose p_t is the CSI 300 index futures price at time t, the return fluctuation can be defined as follows:

$$r_t = 100 * (\ln p_t - \ln p_{t-1})$$

Table 1: The statistical characteristics of the CSI 300 index futures return fluctuation

Date	Number	Average	Std.Err	Median	Min	Max	Range	Skewness	Kurtosis
Data	16470	0	0.08	0	-1.93	0.89	2.82	-1.07	30.71

Remark: The result is the data of return Multiply 100.

The table shows that t skewness does exist in some of the observation sample; they do not obey symmetric distribution generally. The kurtoses of observational samples are significantly more than 3, which mean there is significant summit and fat tail phenomenon. The sample does not conform to normal distribution. Next, the return fluctuation and volume of the observed samples was shown in Figure 1.

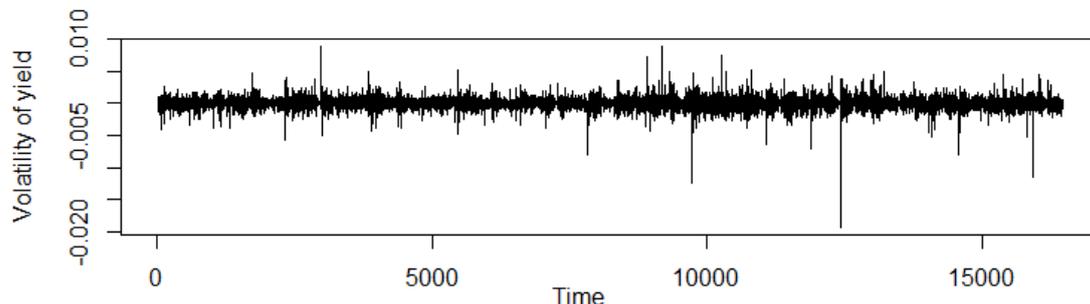


Figure 1: The fluctuation of CSI 300 index futures return

From the over charts, the CSI 300 index futures return fluctuation show that the equences of yield show obvious volatility gathered, fat-tailed, information asymmetry features.

The main purpose of this paper is to study the volatility of the CSI 300 stock index futures. Due to the large amount of sample data, it accord with the assumption that the large sample size has a weak stationarity, therefore, the data has a weak stationarity, we do not need to do a smooth test samples. Next step, the correlation between the residual sequences of the return series is tested to confirm whether the residual sequence has the conditional heteroscedasticity, that is, ARCH effect. We use ARCH-LM test to get the following results: the Chi-squared result is 41.936, and the p-value is 3.414e-05, which reject original hypothesis, there is heteroscedasticity on residuals obviously.

Due to the high order GARCH model leads to excessive GARCH and ARCH coefficients, the model is unstable. Therefore, we establish the GARCH (1, 1) model to fit the residual model. We found in the fitting of GARCH model in the process of establishing the model of the residual does not fully conform to the normal distribution, but there are some skewness, so in this paper us the partial normal distribution of GARCH model estimation. The following table is the best result of the GARCH model which is established by using the rug arch package in the R language (table 2).

Table 2: The optimal estimation of GARCH model parameters

GARCH model	Parameters	Value	Error	T test	Pr(> t)	Significance
GARCH(1, 1))	Mu	0.000020	0.000006	3.642762	0.00027	***
	MA1	-0.022641	0.234885	-0.096393	0.92321	-
	MA2	-0.029000	0.233629	-0.124128	0.90121	-
	omega	0.000000	0.000000	0.168071	0.86653	-
	alpha1	0.038781	0.002163	17.925993	0.000000	***
	beta1	0.948130	0.002195	431.9823	0.000000	***

Note: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

The fitting results from the model show that the coefficient is close to zero. Other coefficients were significantly above the 0.001 level. The regression model of alpha and beta values were less than 1, and the alpha + beta <1, regression model and convergence, and both are very close to 1, according to the results of Engle and Bollerslev, if alpha and beta and close to 1, the impact on conditional variance has strong persistence. The model fitting results show that the beta1 of models is very large, which close to 0.9. It indicates that the CSI 300 stock index futures market volatility has long memory characteristics. After fitting the GARCH model, we use Weighted ARCH LM Tests to test the model fitting residuals in order to verify whether the model is fully fit, the specific results of the table as shown in table 3.

Table 3: Weighted ARCH LM Tests

	Statistic	Shape	Scale	P-Value	Significance
ARCH Lag[3]	0.03944	0.500	2.000	0.8426	-
ARCH Lag[5]	0.07648	1.440	1.667	0.9910	-
ARCH Lag[7]	0.12571	2.315	1.543	0.9990	-

Note: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

From the results of ARCH LM Tests fitting, we can see that there is no ARCH effect after fitting the model, which shows that our model can reflect the actual volatility characteristics of the stock index futures, and the fitting is very good in Weighted. In this way, we get the fitting equation, which shows that our model is robust.

4. Conclusion

This paper makes an empirical analysis on the model of GARCH logarithm yield continuous CSI 300 stock index futures trading day rate based on the study of the CSI 300 index futures volatility, and get the optimal GARCH model and full fitting of the model is tested, this paper obtained the following conclusions in the process of research:

- (a). The high frequency data of stock index futures have the characteristics of peak and thick tail.
- (b). GARCH (1, 1) model can well fit the characteristics of the high frequency volatility of Shanghai and Shenzhen 300 stock index futures;
- (c). The impact of conditional variance is very persistent: according to the model regression results we find that the sum of alpha and beta in the GARCH model is close to 1;
- (d). Shanghai and Shenzhen 300 stock index futures market intraday volatility has long memory characteristics: the model fitting results show that the beta value of the model is very large, more than 0.9.

References

- [1] Engle, R.F. Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Infl. in *Econometrica*. 1982.
- [2] Bollerslev, T., Generalized autoregressive conditional heteroskedasticity. *Eeri Research Paper*, 1986. 31(3): p. 307-327.
- [3] Engle, R.F. and T. Bollerslev, Modelling the persistence of conditional variances. *Econometric Reviews*, 1986. 5(1): p. 1-50.
- [4] Chou, R.Y., Volatility persistence and stock valuations: Some empirical evidence using garch. *Journal of Applied Econometrics*, 1988. 3(3): p. 279-94.
- [5] Curto, J.D., J.C. Pinto and G.A.N. Tavares, Modeling stock markets' volatility using GARCH models with Normal, Student's and stable Paretian distributions. *Statistical Papers*, 2009. 50(2): p. 311-321.
- [6] Hui Jun, Zhu Cui, ARMA-ARCH model analysis of the securities investment fund market. *Journal of HeFei University of Technology: Natural Science Edition*, 2010. 33(7): Page 1108-1112.