ISSN 2616-5805 Vol. 3, Issue 1: 73-78, DOI: 10.25236/AJMS.2022.030110

Empirical Analysis of Fund Index Volatility Based on Conditional Heteroscedasticity Model

Xiayi Zhang

School of Mathematical Sciences, South China Normal University, Guangdong, China

Abstract: *In order to study the changes of China's fund market, this paper carries out time series modeling and fitting prediction on the series based on the monthly series data of Shanghai Securities Fund Index from January 2010 to December 2019. EGARCH (1,1) model has a good fitting effect on Shanghai Securities Fund Index series, all parameters are not 0, and the residual series of the model is tested to obey the standard normal distribution. Finally, the fitted model is used to predict the Shanghai Securities Fund Index from January to May 2020, and compare it with the real value to test the accuracy of the model. The results show that the actual values are within the prediction interval of 95% confidence coefficient, and the fitting effect of the model is superior.*

Keywords: *Shanghai Securities Fund Index, Residual Autoregressive model, EGARCH model, Cointegration Test*

1. Introduction

In recent years, the purchase of funds has gradually become the main way for Chinese people to invest and manage money matters. A securities investment fund refers to an investment tool that forms an independent fund property by raising funds through the sale of fund shares, which is managed by the fund manager and held in trust by the fund trustee. Securities investment is made in the form of asset portfolio, and fund share holders earn profits and bear risks according to their shares [1]. The huge returns of funds have brought huge development space for the fund industry. However, high returns are inevitably accompanied by high risks, such as its strong liquidity, unknown subscription and redemption prices, the management level of the parties involved in fund operation and other factors will bring risks to investors [2]. In order to reduce investment risk, it is necessary to make an empirical analysis of the volatility of the fund market to judge the trend of the fund market.

The fund index can effectively reflect the changes in the fund market. In order to study the volatility characteristics of China's securities investment funds, this paper adopts 125 monthly data of Shanghai Securities Fund Index (closing) from January 2010 to May 2020. The time series model is used to fit and predict the fund index series (Shanghai Securities Fund Index is for the closed-end fund index in the Shanghai and Shenzhen stock exchanges).

2. Time series model of Shanghai Securities fund index data

2.1. Description of data and symbols of model

In this paper, we use the data of 125 monthly index of Shanghai stock fund (data from EPS data platform https://www.epsnet.com.cn/), between January 2010 and May 2020, to analysis. This paper does not consider the data after May 2020, because the occurrence of COVID-19 in 2020 has a great impact on the stock and financial market, and the uncertainty of data is too strong and the volatility is violent, which will affect the fitting effect of the model. In order to test the correctness of the model, this paper uses the data from January 2010 to December 2019 for modeling, and uses the data from January 2020 to may 2020 to test the fitting effect of the model.

Let the Shanghai securities fund index sequence from January 2010 to December 2019 be x_t (t=1,2... 200), and the first-order lag sequence of this sequence be $lagX_t$ (t=1,2... 199).

ISSN 2616-5805 Vol. 3, Issue 1: 73-78, DOI: 10.25236/AJMS.2022.030110

2.2. Observe the trend and periodicity of sequence xt

First, SAS software is used to draw the sequence diagram of Shanghai Securities fund index sequence, as shown in Figure 1.

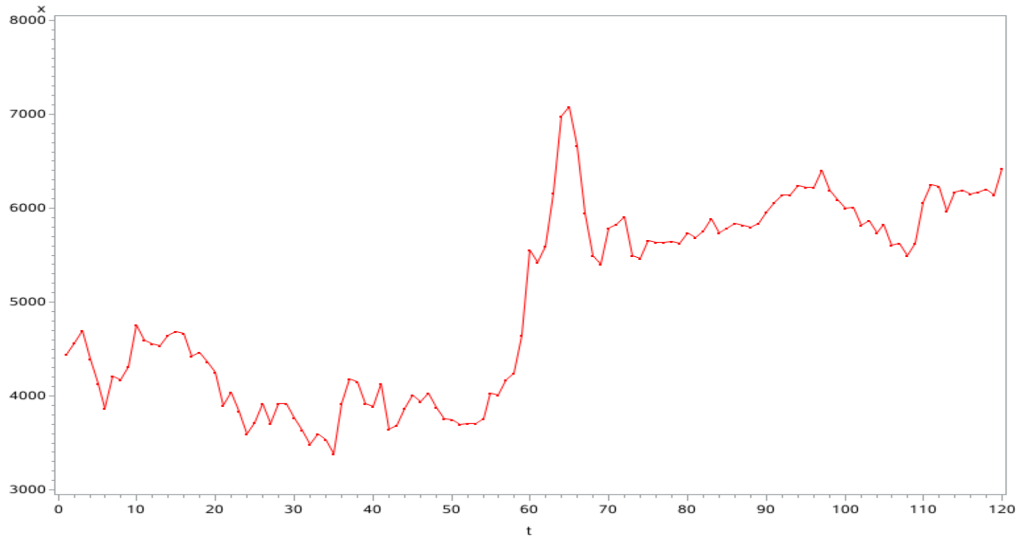

Figure 1: Sequence diagram of the sequence x_t

It can be clearly seen that the sequence is non-stationary and has a certain increasing trend, but there is no periodicity. According to the SAS calculation results, it is concluded that the sequence x_t is not a stationary sequence[3]. Combined with the characteristics of increasing trend of sequence x_t, this paper adopts the residual auto regressive model. The first-order lag sequence $lagX_t$ of the sequence is used to regress the sequence x_t to explain the deterministic trend of the sequence x_t, and then the residual series is further processed. According to the results of sequence analysis, both time series x_t and $lagX_t$ are non-stationary sequences. If you want to do regression, you should first carry out EG test to test the cointegration relationship between the two time series. If regression is required, EG test should be performed first to test the co-integration relationship between the two time series.

2.3. lagXt and xt were tested for EG cointegration relationship

First, check the single integer order of the two sequences. If they are single integer sequences of the same order, the cointegration regression model of sequence x_t with respect to $lagX_t$ can be further established. Then test the stationarity of residual sequence in the co-integration regression model. SAS program is used to perform EG test operation.

(1) Make a first-order difference between x_t and $lagX_t$, and study the stationarity of the first-order difference sequence x_t and $lagX_t$,. The results of ADF unit root test are shown in Table 1.

Table 1: ADF unit root test of first-order difference sequence x_t

Type	Lags	Rho	Pr<Rho	Tau	Pr<Tau	F	Pr>F
Zero Mean	0	-91.3232	<.0001	-8.55	<.0001	-	-
	1	-113.949	0.0001	-7.46	<.0001	-	-
	2	-133.779	0.0001	-6.42	<.0001	-	-
Single Mean	0	-91.7352	0.0001	-8.54	<.0001	36.51	0.0010
	1	-115.131	0.0001	-7.47	<.0001	27.88	0.0010
	2	-137.332	0.0001	-6.46	<.0001	20.86	0.0010
Trend	0	-91.9803	0.0004	-8.54	<.0001	36.47	0.0010
	1	-115.056	0.0001	-7.48	<.0001	28.01	0.0010
	2	-139.057	0.0001	-6.45	<.0001	20.83	0.0010

It can be judged from Table 1 that x_t and $lagX_t$ are first-order single integer sequences, and they can be tested for EG cointegration relationship in the next step.

(2) Establish the cointegration regression equation of x_t on $lagX_t$, observe the stationarity of the

ISSN 2616-5805 Vol. 3, Issue 1: 73-78, DOI: 10.25236/AJMS.2022.030110

cointegration regression residual sequence, and use SAS program to get the following results(Table 2-5):

Table 2: Cointegration regression results of x_t on $lagX_t$

Source	DF	Sum of Squares	Mean Square	F Value	Pr>F
Model	1	3110494985	3110491985	59600.5	<0001
Error	118	6158304	52189	-	
Uncorrected Total	119	3116650288	-	-	-

Table 3: Variance analysis of cointegration regression

Roost MSE	Dependent Mean	Coeff Var	R-Square	Adj R-Sq
228.491	5018.00815	4.5559	0.9980	0.9980

Table 4: Parameter estimates of cointegration regression

Variable	DF	Parameter	Standard Error	t Value	Pr>|t|
$lagX_t$	1	1.00247	0.00411	244.13	<.0001

Table 5: Unit root test of cointegration regression residuals

Type	Lags	Rho	Pr<Rho	Tau	Pr<Tau	F	Pr>F
Zero Mean	0	-91.7661	<.0001	-8.58	<.0001	-	-
	1	-115.246	0.0001	-7.50	<.0001	-	-
	2	-137.507	0.0001	-6.49	<.0001	-	-
Single Mean	0	-91.7797	0.0001	-8.55	<.0001	36.55	0.0010
	1	-115.280	0.0001	-7.47	<.0001	27.92	0.0010
	2	-137.667	0.0001	-6.46	<.0001	20.90	0.0010
Trend	0	-91.9468	0.0004	-8.53	<.0001	36.45	0.0010
	1	-115.940	0.0001	-7.47	<.0001	27.97	0.0010
	2	-138.836	0.0001	-6.45	<.0001	20.81	0.0010

There is a significant linear correlation between x_t and $lagX_t$ (Table 2), it can be seen that the residual order can be regarded as a stationary sequence (Table 5). According to the EG test, there is a co-integration relationship between x_t and $lagX_t$, so it is reasonable to use $lagX_t$ sequence to perform regression on x_t, and $lagX_t$ sequence can be used to explain the deterministic trend of x_t in the residual autoregressive model.

2.4. Conditional heteroscedasticity model

After verifying the co-integration relationship between sequence x_t and sequence $lagX_t$, sequence $lagX_t$ can be used to fit the deterministic part of x_t sequence in residual autoregressive model, and the heteroscedasticity and autocorrelation of residual sequence can be tested.

The SAS software is used to fit the model and test the heteroscedasticity and autocorrelation. The results are as follows (Table 6,7,8):

Table 6: Parameter estimation of deterministic part in residual autoregressive model

Variable	DF	Estimate	Standard Error	t Value	Approx Pr>|t|	Variable Label
Intercept	1	110.6346	107.0001	1.03	0.3033	
lagx	1	0.9821	0.0210	46.77	<.0001	lagx

Table 7: Autocorrelation estimation of residual sequence

Lag	Covariance	Correlation
0	51281.9	1.000000
1	11707.6	0.228300
2	-2471.0	-0.048185
3	-4585.4	-0.089416

ISSN 2616-5805 Vol. 3, Issue 1: 73-78, DOI: 10.25236/AJMS.2022.030110

Table 8: Conditional heteroscedasticity test of residual sequence

Order	Q	Pr>Q	LM	Pr> LM
1	5.9006	0.0151	5.5967	0.0180
2	5.9006	0.0523	5.8427	0.0539
3	22.9484	<.0001	23.2702	<.0001
4	49.0603	<.0001	34.2702	<.0001
5	49.2178	<.0001	34.8654	<.0001
6	50.7301	<.0001	35.0624	<.0001
7	54.9177	<.0001	36.9358	<.0001
8	54.9507	<.0001	39.5395	<.0001
9	55.0126	<.0001	39.6579	<.0001
10	55.0931	<.0001	41.5663	<.0001
11	56.0184	<.0001	41.8875	<.0001
12	57.1491	<.0001	41.8882	<.0001

According to the parameter estimation of the deterministic part of the residual autoregressive model (Table 6), it can be seen that $lagX_t$ coefficient is significantly not 0, sequence $lagX_t$ has a strong explanatory ability for the deterministic part of x_t, and there is a significant linear correlation between the two sequences, and the influence of $lagX_t$ on x_t is positively correlated.

The residual sequence has obvious conditional heteroscedasticity (Table 7,8), so it is necessary to further explain the autocorrelation and conditional heteroscedasticity of the residual sequence, fully extract relevant information. Use AR-GARCH and other conditional heteroscedasticity models to fit and explain the residual sequence.

3. EGARCH model is used to fit residual sequence and its fitting effect

After several attempts, it is found that when the residual sequence is fitted by AR-GARCH model, the AR term in the final fitting result is not significantly 0, so it is considered that there is no autocorrelation in the residual sequence[4].Therefore, after removing the autoregression term (AR term) of residual, the conditional heteroscedasticity model was fitted again[5].Through experiments, it is found that among many models (ARCH model, EGARCH model etc.), EGARCH(1,1) model has the best fitting effect.

Table 9: Statistical indicators of EGARCH (1,1) model

Exponential GARCH Estimates			
SSE	6114670.2	Observations	119
MSE	51384	Uncond Var	-
Log Likelihood	-797.12717	Total R-Square	0.9491
SBC	1622.92909	AIC	1606.25434
MAE	158.761967	AICC	1607.00434
MAPE	3.28327552	HQC	1613.02543
Normality Test	3.0714	Pr>ChiSq	0.2212

Table 10: Statistical indicators of EGARCH(1,1) model

Variable	DF	Estimate	Standard Error	T Value	Approx Pr>|t|	Variable Label
Intercept	1	94.8942	80.9775	1.17	0.2413	
lagx	1	0.9523	0.0160	61.26	<.0001	lagx
EARCH0	1	14.6044	1.5241	9.58	<.0001	
EARCH1	1	0.8233	0.2542	3.23	0.0012	
EARCH2	1	-0.3793	0.1420	-2.67	0.0075	
THETA	1	0.3193	0.1581	2.02	0.0034	

Table 9 shows that the R square of the model is 0.95, and the model is valid.

Table 10 shows that at the significance level of 0.05, EGARCH model parameters are not 0. Each variable of EGARCH model is effective.

According to the normal distribution test in Table 7, the residual sequence of EGARCH model does not negate the null hypothesis of standard normal distribution, that is, the new residual sequence obtained after fitting the EGARCH model is considered to obey the standard normal distribution.

It can be seen from the coefficients and tests in Table 7 and 8 that EGARCH (1,1) model has a good fitting effect on sequence x_t, and can well explain the trend change of sequence x_t and the heteroscedasticity of residuals.

To sum up, the EGARCH regression model fitted is:

$$\begin{gathered} x_t = 94.8942 + 0.9823x_{t-1} + u_t \\ u_t = \sqrt{h_t}e_t\,,\ \ e_t \sim NID(0,1) \\ \ln(h_t) = 14.6044 - 0.3793\ln(h_{t-1}) + 0.8223\mathrm{g}(e_{t-1}) \\ \mathrm{g}(e_t) = 0.3193e_t + [|e_t| - E|e_t|] \end{gathered} \tag{1}$$

4. Test the prediction effect of the model

The EGARCH (1,1) model is used to evaluate and predict the Shanghai Securities Fund index data from January 2010 to May 2020. The fitting effect diagram is shown in Figure 2, where the two blue lines are the upper and lower limits of the 95% confidence interval, and the red line is the point estimation of the Shanghai Securities Fund index in each month. The black dot is the true value of the Shanghai Stock Fund index each month.

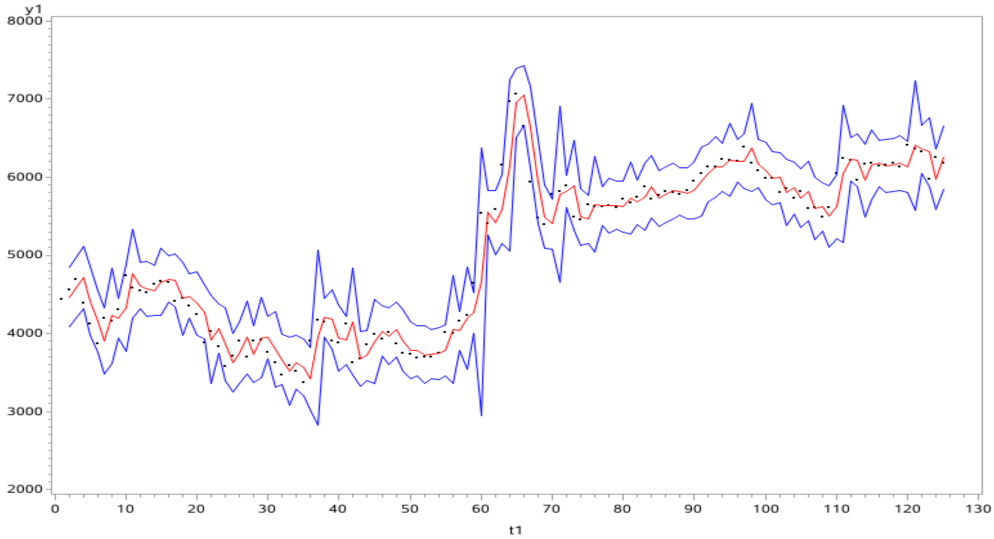

Figure 2: Fitting effect diagram of EGARCH (1,1) model for Shanghai Securities fund index from January 2010 to May 2020

According to the fitting effect diagram of the model, it can be seen that the real values of the Shanghai Securities fund index all fall within the estimated 95% confidence interval, and the red line formed by the point estimation is relatively consistent with the real value, so the fitting effect of the model fits well. Since the data from January to May 2020 were used as test data and did not participate in the model fitting, the prediction of the Shanghai Securities fund index from January to May 2020 by the model was more effective in judging the model effect. The comparison between the predicted value and the real value from January to May 2020 is shown in Table 11.

Table 11: Comparison between real value and predicted value of Shanghai Securities Fund Index from January to May in 2020

	The actual value	Predictive value	The lower end of the 95% forecast range	The top end of the 95% forecast range
2020.1	6376.69	6406.46	5575.09	7237.84
2020.2	6341.1	6358.98	6052.85	6665.11
2020.3	5985.96	6324.02	5879.54	6768.49
2020.4	6268.63	5975.15	5591.74	6358.55
2020.5	6196.77	6252.82	5846.55	6659.09

It can be seen from the Table 9 that the predicted value of the Shanghai Securities Fund Index Model from January to May 2020 is close to the actual value, and the actual value is within the 95% confidence interval. Therefore, it can also be judged that the EGARCH (1,1) model has a good fitting effect and a

ISSN 2616-5805 Vol. 3, Issue 1: 73-78, DOI: 10.25236/AJMS.2022.030110

strong explanatory ability for Shanghai Securities fund index series x_t.

5. Conclusion

The Shanghai securities fund index series referred in this paper can be explained by the autoregressive model and the EGARCH (1,1) model. It can be concluded from the model that the Shanghai securities fund index series has the following characteristics:

(1) The fund index is positive autocorrelation. The bigger of the figure last month, the bigger of the fund index will be this month.

(2) The fund index is heteroscedasticity. Under the impact of external factors on the fund index, its volatility will increase. But over time, its volatility diminishes.

(3) The value of the fund index has an asymmetric effect on its fluctuation intensity. A rise in the index has a greater impact on its range than a fall.

References

[1] Lu Haixia. Based on intervention analysis model for the forecast of Shanghai securities fund index [J]. Value engineering, 2019, 38(24): 11-13. DOI: 10.14018/j.carol carroll nki cn13-1085/n.2019.24.005.
[2] Jiang Tao, Wu Junfang. Application of ARIMA Model in fund Index Forecasting [J]. Statistical Education, 2007(07): 12-13.
[3] Li Chuner. AHCH class model of fund income volatility empirical study [J]. Journal of modern business, 2019 (17): 80-82. The DOI: 10.14097/j.carol carroll nki. 5392/2019.17.036.
[4] Wang Yan. Applied Time Series Analysis [M]. China Renmin University Press, 2005. 167-171, 170-172.
[5] Pealat Clement, Bouleux Guillaume, Cheutet Vincent. Improved Time series clustering based on new Geometric Frameworks [J]. Journal of Pattern Recognition, 2022 124.