Analysis of the Relationship Between ETF Volatility and Liquidity Based on ARMA-GARCH Model

Qingyuan Feng

Email: hwhqy@163.com

Canvard College, Beijing Technology and Business University, Beijing, 101118, China

Abstract: Based on ARMA-GARCH model, this paper takes Amihud's non liquidity ratio as an indicator to measure ETF liquidity, and makes an empirical analysis on the impact of ETF liquidity on return volatility in China. The analysis results show that the weak liquidity of ETFs has a positive impact on the volatility of returns, and it also has a certain explanatory effect on the risk premium of ETFs. However, the explanatory power of liquidity is limited, and there are other factors that affect the volatility of ETFs' returns. Finally, some policy suggestions are given.

Keywords: Volatility, ETF, ARMA-GARCH Model.

1 INTRODUCTION

Exchange Traded Fund (ETF), as an important asset allocation tool for market investors, has been favored by investors for many years since it was launched in the United States in 1993. In 2004, the first ETF product in China, Huaxia CSI 50ETF, was established. Since then, the ETF market in China has shown a vigorous development trend [7]. According to the Shanghai Stock Exchange ETF Industry Development Report (2022) released by the Shanghai Stock Exchange, by the end of 2021, there were 635 ETFs listed for trading in China, with the total assets reaching 1405.2 billion yuan, and the annual turnover of ETFs in Shanghai Stock Exchange alone exceeded 15 trillion yuan.

However, because ETFs have the characteristics of both stocks and index funds, liquidity problems may arise when trading in the secondary market of the exchange. For ETFs, liquidity is very important. Only better liquidity can attract customers, because their products are highly homogeneous. Even though the overall ETF is growing, in fact, except for the head ETF, the liquidity and trading volume of other parts are not good. The lack of liquidity of ETFs that are unpopular with investors may cause market makers to be limited in developing appropriate markets, and then increase the transaction costs and risks borne by such ETF investors [3]. Therefore, it is significant for investors to avoid unnecessary additional risks by studying the impact of ETF liquidity on its volatility [8].

BDEDM 2023, January 06-08, Changsha, People's Republic of China

DOI 10.4108/eai.6-1-2023.2330355

2 THEORY INTRODUCTION

2.1 Selection of Liquidity Indicators

The Amihud measure is selected as an indicator to measure liquidity. Since the yield data used in this paper are daily data, the Amihud measure is calculated according to the following formula:

$$\text{Amihud}_{i,t} = \frac{|Ri,t|}{Vi,t}$$

$R_{i,t}$ are the yields of securities i on trading day, $V_{i,t}$ are the trading volumes of securities i on trading day t [1]. After the ADF test of Amihud measure, we can find that the Amihud index series of ETF funds from January 1, 2015 to December 31, 2021 is stable.

2.2 ARMA-GARCH Model Introduction

ARCH model is also called "autoregressive conditional heteroscedasticity model". This model is proposed by Engle to solve the conditional heteroscedasticity of data and applied to the study of volatility. However, the lag term of the variance equation is sometimes large, and the ideal volatility equation cannot be obtained by using the ARCH model. Therefore, in 1986, Bollerslev proposed the GARCH (p, q) model to improve the defects of the ARCH model. Because most of the financial time series data have the phenomenon of volatility aggregation, showing the characteristics of thick tail distribution, the use of this model can effectively eliminate the excessive peak problem caused by data [2]. GARCH (1,1) model is often used as a modeling tool in practical problems [2]. This paper adopts the widely used GARCH (1,1) model, and adds the weak liquidity index represented by Amihud measure to the conditional variance equation of GARCH (1,1) model, and removes the impact of trading volume by extracting the residual of Amihud measure on the linear regression of trading volume. Thus, GARCH (1,1) model with weak liquidity index is obtained:

$$r_t = c+ARMA(p,q)+u_t \tag{1}$$

$$\sigma^2{}_t = \alpha_0+\alpha_1 u^2{}_{t-1}+\alpha_2\sigma^2{}_{t-1}+\beta\varepsilon_t \tag{2}$$

$$\text{Amihud}_t=\delta_0+\delta_1 V_t+\varepsilon_t \tag{3}$$

Wherein, equation (1) is called mean value equation and equation (2) is conditional variance equation. u_t follows an unknown distribution with mean value of 0 and variance of $\sigma^2{}_t$. $\sigma^2{}_t$ represents the volatility of the yield, namely the conditional variance. The model has constraints $0<\alpha_1+\alpha_2<1$, which is used to ensure that the unconditional variance of u satisfying the model is limited and unchanged, while the conditional variance $\sigma^2{}_t$ can change over time [4].

3 EMPIRICAL ANALYSIS

3.1 Basic Statistical Characteristics of ETF Fund Yield

The research object of this paper is China's Exchange Traded Fund (ETFs). The data use the data set from January 1, 2015 to December 31, 2021 in Guotai An CSMAR database. Excluding the missing data of holidays or individual dates, the daily returns of 1303 ETF funds considering cash dividends in 1705 trading days were obtained.

After cross sectional average of the returns of these 1303 ETFs, we can get a series of returns. Figure 1 shows the general trend of the ETF return series in 1705 trading days. From the time series diagram of ETF yield, it can be seen intuitively that the large and small fluctuations of ETF yield tend to gather in different periods. Based on this, it can be preliminarily inferred that the volatility of ETF yield has aggregation, and because the amplitude of volatility is inconsistent, it can be judged that there may be heteroscedasticity in the yield series.

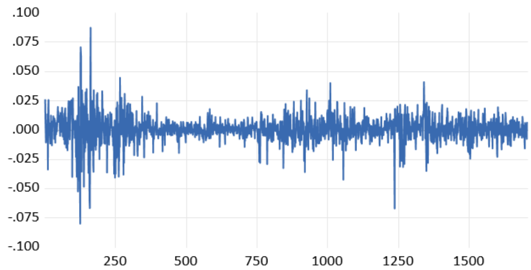

Figure 1 ETF Fund Yield Time Series

Descriptive statistics of ETF yield series are shown in Table 1 and Figure 2. It can be seen from Table 1 that the mean and standard deviation of ETF returns are small, the skewness is negative, and the kurtosis is far greater than 3, indicating that this distribution has a long left tail and a fatter tail compared with the normal distribution. The Jarque Brea statistic is 5416.71, which is significant at the 1% level. This data also shows that the ETF yield series does not obey the normal distribution. Figure 2 is a Q-Q scatter chart of the yield series. It can be found that the tail of the yield series deviates from the diagonal seriously, which indicates that the distribution of ETF returns has the characteristics of a thick tail, which again indicates that the yield series does not obey the normal distribution. The unit root test of the yield series shows that the t statistic obtained is significant at the 1% level, indicating that the yield series is stable [9].

Table 1 Descriptive Statistics of ETF yield

Mean	Standard Deviations	skewnes	kurtosis	J-B	ADF
0.00062	0.011258	-0.311215	11.70974	5416.710**	-34.29044**

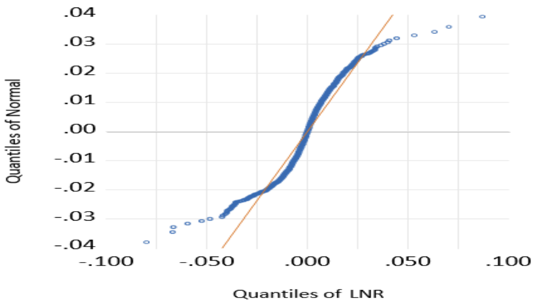

Figure 2 Q-Q Scatter Chart of ETF Yield

Next, we use the autocorrelation and partial autocorrelation coefficient graph of the yield series and Ljung Box Q statistics to determine whether the series has serial correlation. Figure 3 shows the autocorrelation and partial autocorrelation coefficients of ETF yield series. It can be seen from the figure that the autocorrelation and partial autocorrelation values of each order lag are very small, and the corresponding P value is also less than 0.001, so it is judged that there is significant autocorrelation in the yield series [10]

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.203	0.203	70.114	0.000
		2	-0.017	-0.060	70.589	0.000
		3	-0.009	0.007	70.724	0.000
		4	0.045	0.046	74.182	0.000
		5	0.016	-0.003	74.646	0.000
		6	-0.036	-0.038	76.917	0.000
		7	0.027	0.046	78.147	0.000
		8	0.057	0.039	83.668	0.000
		9	-0.002	-0.022	83.673	0.000
		10	-0.088	-0.080	97.071	0.000
		11	-0.005	0.030	97.119	0.000
		12	0.038	0.022	99.606	0.000
		13	0.076	0.068	109.67	0.000
		14	-0.054	-0.076	114.66	0.000
		15	-0.010	0.021	114.84	0.000
		16	0.017	0.002	115.36	0.000
		17	0.024	0.021	116.34	0.000
		18	0.005	0.009	116.38	0.000
		19	0.013	0.017	116.68	0.000
		20	0.051	0.027	121.23	0.000

Figure 3 Autocorrelation and Partial Autocorrelation Coefficient of ETF Yield

To sum up, the ETF yield series shows these statistical characteristics: peak and fat tail, autocorrelation and bias. For the autocorrelation of the yield series, this paper plans to introduce the conditional mean model - ARMA model to eliminate the series correlation. However, the index GARCH (EGARCH) model proposed by Nelson (1991), the asymmetric power ARCH (APARCH) model proposed by Ding, Granger and Engle (1993) all allow positive and negative asset returns to have an asymmetric impact on volatility, which can show better results in biased description and analysis [5].

3.2 Construction of mean value equation by ARMA model

First, we need to determine the form of the mean value equation. According to the autocorrelation and partial autocorrelation coefficient graph of the yield (Figure 3), we can roughly see the maximum values of the two parameters p and q in the ARMA (p, q) model. Taking max (p)=5 and max (q)=5, after comparing AIC, SC and HQC values of 36 models, the optimal model was determined to be ARMA (4,3) using the principle of minimizing information criteria. Table 2 shows the regression results of ARMA (4,3) model.

Table 2 Regression Results of ARMA (4,3) Model

Term	coefficient	T	P
AR(1)	-0.599912	-27.12507	0.0000
AR(2)	-0.614815	-25.51083	0.0000
AR(3)	-0.746230	-41.04747	0.0000
AR(4)	0.213533	16.05401	0.0000
MA(1)	0.819291	42.15887	0.0000
MA(2)	0.741298	28.70628	0.0000
MA(3)	0.909103	51.90958	0.0000

The mean value equation established by ARMA (4,3) is given by the following equation:

$$r_t=-0.600r_{t-1}-0.615r_{t-2}-0.746r_{t-3}+0.214r_{t-4}+u_t+0.819u_{t-1}+0.741u_{t-2}+0.909u_{t-3}$$

Figure 4 shows the residual sequence diagram of the mean value equation. It can be preliminarily judged from the observation chart that there may be conditional heteroscedasticity effect in the residual sequence.

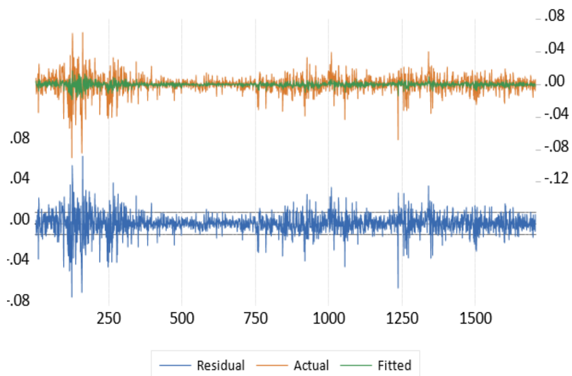

Figure 4 Residual Sequence Diagram of Mean Value Equation

Next, test whether the residual of the equation has the ARCH effect according to the Lagrange multiplier (LM). By using the information criterion minimization principle, the lag order is determined to be 5, and the ARCH-LM test results are shown in Table 3.

Table 3 ARCH-LM test results of ARMA (4,3) residuals

LM	P
408.8875	0.0000

The test results show that the LM statistic is very significant at the 1% level, that is, the residual sequence has an obvious ARCH effect.

3.3 Establishment of benchmark GARCH (1,1) model

When constructing the GARCH (1,1) model of benchmark income, as the tail distribution is unknown, first compare the GARCH (1,1) models under different tail distribution assumptions to determine the optimal model. Here, three different tail distributions are compared, namely, normal distribution, student-t distribution and generalized error distribution (GED). Table 4 lists the AIC, SC and HQC values of the model under the three tail distributions.

Table 4 Comparison of models under different tail distributions

Tail distribution	AIC	SC	HQC
Normal	-6.597068	-6.587494	-6.593524
Student-t	-6.672364	-6.659599	-6.667639
GED	-6.667961	-6.655195	-6.663236

It is not difficult to see that GARCH (1,1) model based on Student-t distribution is optimal.

Table 5 Regression Results of GARCH Benchmark(1,1) Model

	Coefficient	Z	P
α_0	8.45E-07**	2.553372	0.0107
α_1	0.079742***	5.921469	0.0000
α_2	0.915373***	74.21357	0.0000
$\alpha_1 + \alpha_2$	0.995115		

Table 5 shows the estimation results of the GARCH (1,1) model of benchmark returns when the tail distribution follows the t distribution. It can be seen that in the benchmark model, α_1 and α_2 are significant at 1% or higher significance level. The sum of α_1 and α_2 is highly close to 1, which indicates that the return sequence of ETF has a high sustainability feature; α_2 is greater than 0.9 and highly significant, indicating that the market memory is strong and the impact of conditional variance is more lasting.

In order to test whether GARCH (1,1) model eliminates ARCH effect, this paper conducts autocorrelation test again, this time it is the residual sequence of the model. The following table shows the ARCH-LM test results of the residual sequence, as shown in Table 6.

Table 6 ARCH-LM Results of GARCH (1,1) residual

LM	P
7.189669	0.2069 0.2069

The test results show that the model effectively eliminates the ARCH effect of residual sequence.

3.4 GARCH (1,1) model with liquidity indicators

In order to study the impact of liquidity on the volatility of ETF returns, we will add liquidity indicators to the benchmark GARCH (1,1) model of ETF returns. Amihud measure is selected as the liquidity indicator here to measure the illiquidity of securities [6].

First, the stability of Amihud measure and trading volume series is tested. The results are shown in Table 7. At the 5% significance level, it can be considered that the two sequences are stable and single integer of the same order.

Table 7 ADF Test Results of Amihud Measure and Transaction Volume Series

	T	P
Amihud	-7.034456***	0.0000
Transaction volume	-3.843325***	0.0026

Use the Engle Granger two-step method to carry out cointegration test for two variable sequences. The OLS is used to regress the two sequences, extract the residual sequence, and test. Here, ADF test is selected. The results show that the T value is -12.05119, and the significance level is 1%, which is significantly stable, the residual sequence is stable, and there is cointegration between Amihud measure and trading volume.

We carry out OLS regression in equation (3) and extract residual sequence. The estimated results of parameters of GARCH (1,1) model including liquidity indicators obtained by adding residual series to GARCH (1,1) model are shown in Table (8).

The results show that the coefficient of weak current term β is positive and significant at the 5% significance level; After adding weak liquidity indicators, α_1 and α_2 in the model is still highly significant, but its corresponding z values have decreased, and the sum of α_1 and α_2 used to reflect the persistence of volatility has also decreased to a certain extent.. The value of α_2 is still large and significant, which indicates that the aggregation and persistence of volatility are still very obvious, and there are still unknown factors that can be used to explain the volatility of ETF fund returns.

Table 8 Regression Results of GARCH (1,1) Model with Current Items

	Coefficient	Z	P
α_0	1.11E-06***	3.026290	0.0025
α_1	0.075679***	5.667164	0.0000
α_2	0.915487***	73.43036	0.0000
β	0.061765**	2.193103	0.0283
$\alpha_1 + \alpha_2$	0.991166		

To sum up, the liquidity of ETF can be used to explain the volatility of its yield. Weak liquidity will increase the volatility of ETF's yield, while when liquidity increases, the volatility of ETF's yield will weaken; However, liquidity has limited ability to explain ETF yield volatility. The introduction of weak liquidity indicators cannot significantly reduce the persistence and aggregation of ETF yield volatility. There are still other factors behind it that can explain the volatility of ETF yield; There is an obvious risk premium in China's ETFs, and due to the leverage effect, when the ETFs have insufficient liquidity, it is more likely to cause large fluctuations in their yield, which will lead to investors taking greater risks and expanding transaction costs.

4 CONCLUSIONS AND RECOMMENDATIONS

The following conclusions are drawn from the analysis of modeling results:

(1) The volatility of China's ETF returns has obvious characteristics of aggregation and persistence, that is, larger volatility tends to be accompanied by larger volatility, smaller volatility tends to be accompanied by smaller volatility, and periods of high volatility and low volatility will alternate;

(2) Liquidity is an important factor to explain ETF yield volatility. Weak liquidity has a positive impact on ETF yield volatility. Weakening ETF liquidity will lead to an increase in ETF yield volatility. Accordingly, when liquidity is enhanced, ETF yield volatility will weaken;

(3) There is leverage effect in the volatility of ETF yield in China, and the negative impact on the volatility is often greater than the positive impact on the volatility. Therefore, when the liquidity is weakened, it is more likely to cause a large range of yield volatility, making investors bear greater risks;

(4) ETFs in China have an obvious risk premium, and the liquidity of ETFs is one of the factors that explain the risk premium. Therefore, if the ETF has insufficient liquidity, the weakening of liquidity will lead to the increase of yield volatility and risk, which will lead to the expansion of transaction costs.

(5) Liquidity has limited ability to explain ETF yield volatility, and the aggregation and persistence of volatility cannot be fully explained by liquidity. There are still unknown other explanatory factors behind it that can be used to explain ETF yield volatility;

Based on the research results, relevant policy recommendations are proposed:

(1) In order to further improve the liquidity of ETFs, some fund companies concerned are expected to do something, mainly in a series of practices such as increasing market makers or adjusting the minimum subscription and redemption units. So as to improve the trading volume of ETF and expand the scale of ETF, so as to promote the better and faster development of ETF;

(2) Due to the head effect and first mover advantage of China's ETF market, in the context of the increase in the size of China's ETF market in recent years and the cooling of new funds, focusing on liquidity will help build investor confidence and the recovery of ETF market issuance; For funds with seriously insufficient liquidity, relevant regulations should be formulated to clear them from the market in a timely manner.

REFERENCES

[1] Amihud Y, Mcndelson H. Asset Pricing and the Bid-ask Spread[J]. Journal of Financial Economics, 1986,17(2):223-249.

[2] Bae K, Kim D. Liquidity risk and ETF return, variance and tracking error [J]. Journal of Financial Economics, 2020:138.

[3] Chi Jun, Yang Chunpeng. Investor Trading Behavior and Electronic Transfer Function Price Fluctuation[J]. Investment Research. 2019, 38 (03): 133-143

[4] Li Mingxuan, Yu Hanjun. Research on RMB Exchange Rate Volatility Based on Alma Gale Model[J]. Times Finance 2020 (33): 1-8

[5] Jonathan D. Cryer, Kung Sik Chan. Time Series Analysis and Application R Language (2nd Edition) [M]. Translated by Pan Hongyu et al. Beijing: China Machine Press, 2011:40-58

[6] Jiang Fuwei, Ning Wei, Xue Hao. Institutional investment and financial stability. manage the world from the perspective of A-share ETF arbitrage trading 2022,38(04)：25-46

[7] Li Hongxiang. Research on the relationship between investor sentiment and stock return -- based on the empirical score of Shanghai Stock Exchange 50ETF volatility index. Chinese prices 2020,(08):36-37

[8] Liao Xiaoting, Zhu Jianhua, Li Rongqiang. Research on Investor Behavior of Shanghai ETF. China's management informatization 2021,24(05):149-150

[9] Wu miao. nestled in Changfengyuan ETF, stock liquidity and liquidity synchronization Economics (Quarterly) 2021,21(02)：645-670

[10] Zhang Yimin, Wang Xiandong. Research on Shanghai Stock Exchange's 50 square feet volatility[J]. Modern Marketing 2021 (08): 34-35