International Journal of Mathematics and Physical Sciences Research ISSN 2348-5736 (Online)
Vol. 10, Issue 2, pp: (57-74), Month: October 2022 - March 2023, Available at: www.researchpublish.com

Improve the forecasting accuracy of a GARCH model using a decomposition method

Dioubi Fatene, Loudahi Lamia

College of Mathematics and Computer Science, Zhejiang Normal University

Jinhua, 321004, China.

DOI: https://doi.org/10.5281/zenodo.7649004

Published Date: 17-February-2023

Abstract: **In recent years, there has been a greater emphasis on the forecasting accuracy of heteroscedastic models. Instead of estimating the returns volatility using a generalised autoregressive conditional heteroscedastic model ($GARCH$ model), this study separates the returns internal components from the external trend first using a decomposition method called "external trend and internal components analysis method" (ETICA), then estimates the returns volatility using a $GARCH(1,1)$. The study's goal is to determine whether this separation has an effect on the prediction accuracy of the volatility of S&P 500, NASDAQ and Dow Jones stock indices. To explore the ETICA method effect, the root mean squared error has been used to compare the prediction accuracy before and after decomposition. The findings show that on average, the RMSE results were found to be lower before decomposition which means that stock returns had a higher prediction accuracy.**

Keywords: **GARCH model, decomposition method, S&P 500, NASDAQ, Dow Jones, RMSE.**

1. INTRODUCTION

In spite of the fact that fluctuations cause stock prices to rise and fall on a minute-to-minute basis, there isn't a clear equation that can predict how a stock price will behave. Nevertheless, little is known about the dynamics that affect stock prices, which can be broken down into a variety of highly connected economic, political and even psychological elements. Concerns about stock market fluctuations are widespread among scholars in stock values and the accuracy of price predictions. Stock price volatility frequently takes the form of market turbulence. Because it might more accurately reflect the volatility of the stock market than the stock price series, modeling and predicting the stock return rate has greater research value than predicting the stock price. There are several theoretical forecasting techniques. Support vector machine prediction models, and artificial neural network prediction models have all been investigated by some scholars (Guresen et al. 2011; Hu, Zhu and Tse, 2013; Lahmiri, 2016; Zhong and Enke, 2019).

Financial asset movements are typically assessed by volatility (the return on the underlying asset's conditional standard deviation), and can also be viewed as asset risk. The more the asset moves, the more likely its value will drop. The significant improvements brought about by time-series models are due to the addition of conditional variances and conditional means, which change over time. In the case of time series data, heteroscedasticity corrections should be considered. The autoregressive conditional heteroscedasticity $ARCH$ model was developed by Engle (1982). He suggested using conditional densities to combine lagged endogenous and exogenous variables with the information set's vectors of unknown parameters. One of the most difficult aspects of modeling volatility is that it has periods with low movements followed by times with significant movements. $ARCH$ is the first model that assumes that volatility is not constant.

In empirical applications, the conditional variance equation's negative variance parameter and relatively long lag seem to be problems for the model. Bollerslev (1986) created the procedure (Generalized Autoregressive Conditional Heteroskedasticity), which broadened the class of models to permit more flexible lag structures and longer memory. Similarly, Engle (1995) developed a theory that treated the terms for the lagged conditional variance as autoregressive terms.

As opposed to being an autoregression specification, the specification had the drawback of looking more like a moving average specification. Engle's fundamental model has evolved into more complicated models such as, $IGARCH$, $TGARCH$, $EGARCH$, and $GARCH-M$. Many of the models have unique characteristics that improve forecast accuracy. The ARCH and GARCH models have helped to develop financial econometric modeling. These exemplars are well-known for their ability to capture financial time series volatility clustering (Cheteni, 2016).

In a previous study, we demonstrated how separating the internal components from the external trend for stock market indices can increase the degree of predictability of financial time series (Dioubi and Khurshid, 2022). In order to verify this finding, we used a generalized autoregressive conditional heteroscedasticity model in this study. $GARCH(1,1)$ which has been shown by Javed and Mantalos (2013), that its performance is satisfactory in various research, and the initial lag is sufficient to capture the volatility's changes. This study applies a basic $GARCH$ model to assess volatility clustering in S&P 500, NASDAQ and Dow Jones indices, differently from other studies we also estimate it for both internal components and external trend. To verify the efficiency of the decomposition method in improving the predictability of the times series. To compare the accuracy of the prediction we used the RMSE results for the stock market indices before and after decomposition, and our findings support the efficiency of the separation in the enhancement of the predictability degree of the financial time series.

The paper is arranged as follows: Section 2 the data and technique utilized, Section 3 and 4 the empirical analysis and results, and Section 5 the study's conclusion.

2. METHODOLOGY

As mentioned in previous section, we tested the decomposition method which has been proved to be efficient in improving the predictability degree (Dioubi and Khurshid, 2022). Since the GARCH model was discovered to be the most straightforward and reliable of the family of volatility models, we used it to accomplish this (Engle, 2001). Different from our previous study dealing with the return's predictability, this paper explores the decomposition method effect on returns volatility. In this section we have explained briefly the decomposition method we used and which is called "The external trend and internal components analysis method". The "ETICA" method has been proposed by Barthélemy, Nadal and Berestycki (2010), to address a problem of a separation method that has been previously presented by de Menezes et Barabasi (2004) where they suggested a technique to systematically distinguish between internal and external contributions for each time series. By minimizing the impact of the external changes on the system's activity, they tested it on model systems where the size of external perturbations could be explicitly adjusted. By doing so, they were able to gain insights into the internal dynamics of a variety of systems, from Internet traffic to bit flow on a microprocessor.

By looking at a dynamical system that enables them to give a time series $r(t)$ to each component "i" and capture the time-dependent behaviour of N components, where $t=1,...,T$ and $i=1,...,N$. Because each time series signifies the sum of the contributions from the system's internal dynamics and outside perturbations, they assumed that they could separate the two components by writing down separate times series for each:

$$r_i = r_i^{\text{int}}(t) + r_i^{ext}(t) \quad (1)$$

Additionally, they define A_i as the proportion of all traffic passing though the component "i" in time period $t=1,...,T$ to all traffic passing through all components observed in the same time period:

$$A_i = \frac{\sum_{i=1}^{T} r_i(t)}{\sum_{i=1}^{T}\sum_{i=1}^{N} r_i(t)} \quad (2)$$

Then :

$$r_i^{ext}(t) = A_i \sum_{i=1}^{N} r_i(t) \qquad (3)$$

And :

$$r_i^{int}(t) = r_i(t) - \left(\frac{\sum_{i=1}^{N} r_i(t)}{\sum_{i=1}^{T} \sum_{i=1}^{N} r_i(t)} \right) \sum_{i=1}^{N} r_i(t) \qquad (4)$$

And which, by definition, has a zero average because it accounts for variations in traffic from that which is anticipated to pass through component "i". This assumption has only been proven to be true in specific circumstances, so the Barthelemy et al. (2010) method ignored it and proposed their method under various scenarios. They assumed that the global trend was independent of internal contributions, which were assumed also to independents, as well as the external parts were thus characterized as follows:

$$r_i^{ext}(t) = a_i \omega(t) \qquad (5)$$

Where $\omega(t)$ is the overall trend shared by all equities that are responding to it with the prefactor a_i. According to the authors:

$$r_i(t) = a_i \omega(t) + r_i^{\text{int}}(t) \qquad (6)$$

This method has been borrowed and showed that it helps in improving the returns predictability of different stock markets from America and China. In this study we tested the efficiency of this to improve and enhance the predictability of stock market volatility.

2.1 Data collection

In this study, we used the closing prices of stocks from the USA following stock markets:

- S&P 500
- NASDAQ
- Dow Jones

Matei's 2009 study found that increasing the number of observations to over a thousand will help the GARCH model deliver more accurate results. Granger (1992) has shown that a comprehensive investigation over a longer time period can be utilized to assess the predictability of stock prices or returns. As a result, the websites of yahoo finance were used to collect the 1989 daily stock values, which covered the period from January 8, 2015 to November 30, 2022.

The following data transformation is necessary for the decomposition method employed in this study:

$$r_i(t) = \frac{P_i(t) - P_i(t-1)}{P_i(t-1)} * 100 \qquad (7)$$

Where: the closing prices at instants t and $t-1$, respectively, are denoted by $P_i(t)$ and $P_i(t-1)$, and: $i = 1,\ldots,3$ (Because we have three stock indices).

2.2 Empirical analysis

2.2.1 Descriptive statistical analysis

Figure 1, represents the daily closing prices of our stock indices. It demonstrates that there are no periodic fluctuations in the three cases: S&P 500 stocks, NASADQ and the Dow Jones. Those times series are initially regarded as the non-stationary series:

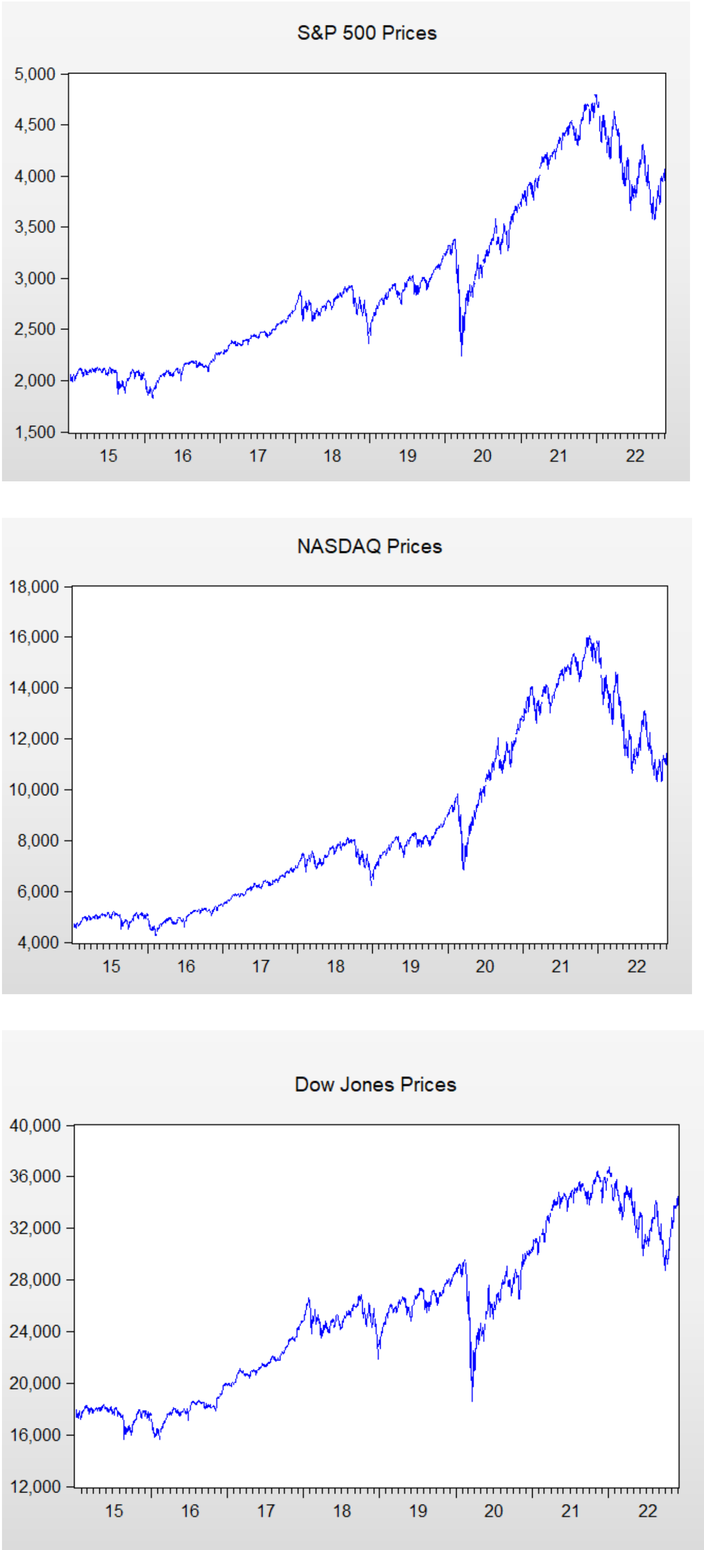


Figure 1: Daily closing prices for S&P 500, NASDAQ and Dow Jones indices

As mentioned in the precedent section, the closing prices has been transformed following the equation 7:

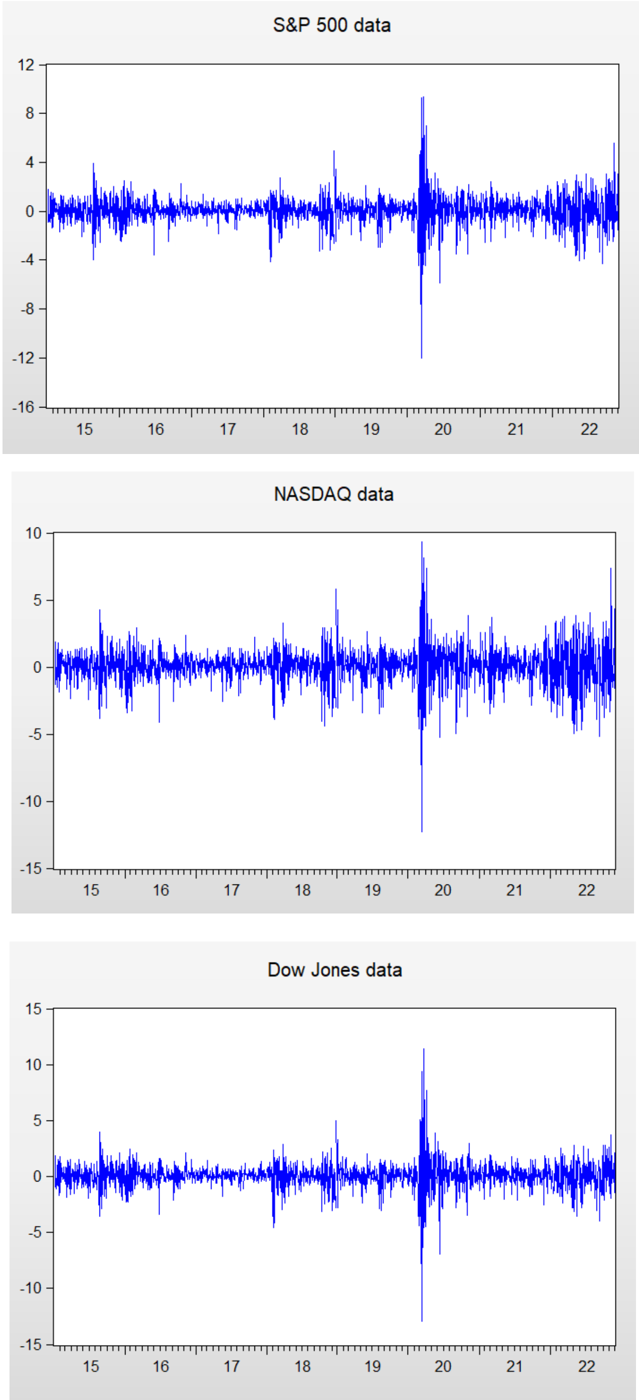


Figure 2: Transformed data from S&P 500, NASDAQ and Dow Jones indices

An examination of the fundamental statistical characteristics of the data series is required prior to data processing. The daily transformed data of the S&P 500, NASDAQ and Dow Jones indices are depicted in Figure 2. In addition, each disposal in this essay is subjected to statistical analysis using the program Eviews 12.

Figure 20 show that there is a tiny time trend present in the date series. Additionally, it exhibits characteristics of time-varying variation and clustering. To fit the volatility of the S&P 500, NASDAQ, and Dow Jones Indices, classic conditional variance models with the assumption of homoscedasticity are therefore no longer appropriate. Instead, since models can deal with time series that exhibit heteroscedasticity and clustering, they might successfully finish this task. Additionally, as we can see from Figure 2, the times series had higher volatility around the corona virus (cov.19) crisis and which is also a good candidate for $GARCH$ process.

In this study, as we previously mentioned, instead of using the returns we have separated the internal components from the external trend and fit the data with the $GARCH$ model, hence the following graphs show both of the quantities changes, and as we can see from the graphs we can conclude the same thing which is that the $GARCH$ models seems to be the heteroscedasticity model fitting the data in these cases.

2.2.2 Normality test:

Figure 3 display descriptive statistics for stock returns for the investigated indices. As we can see from this figure, the normality assumption is disproved by the Jarque-Bera statistics. Thus, validating the general norm that stock returns are not regularly distributed in the case of using the returns, and also in the case of using their internal components and external trend and the figures are in the appendix.

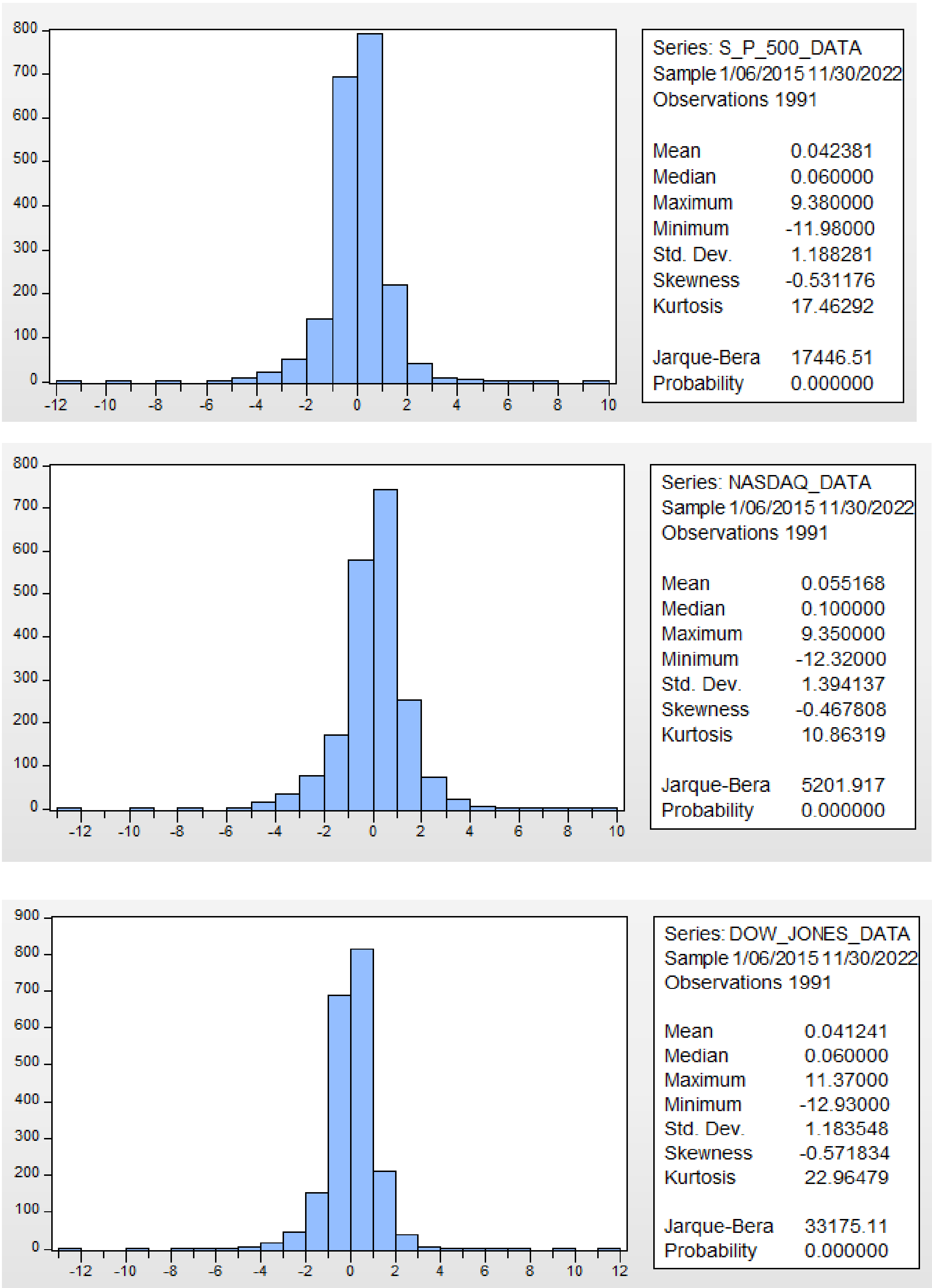


Figure 3: Normality test for the S&P 500, NASDAQ and Dow Jones indices

2.2.3 Stationarity test

The series must be stationary for the volatility to be captured without the use of $ARMA$ extensions on the $ARCH$ and $GARCH$ models. The original time series were not stationary which are the closing prices (as we can see from the figure 1). In this study, the data has been transformed in accordance with equation (8) to be stationary. To the stationarity, we run an Augmented Dickey Fuller test for the 3 stock indices' returns used in this work, as well as their internal and components and external trend.

Augmented Dickey-Fuller Unit Root Test on S_P_500_DATA

Null Hypothesis: S_P_500_DATA has a unit root
Exogenous: Constant
Lag Length: 8 (Automatic - based on SIC, maxlag=25)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-14.23636	0.0000
Test critical values:	1% level	-3.433450	
	5% level	-2.862796	
	10% level	-2.567485	

Augmented Dickey-Fuller Unit Root Test on NASDAQ_DATA

Null Hypothesis: NASDAQ_DATA has a unit root
Exogenous: Constant
Lag Length: 8 (Automatic - based on SIC, maxlag=25)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-14.88305	0.0000
Test critical values:	1% level	-3.433450	
	5% level	-2.862796	
	10% level	-2.567485	

Augmented Dickey-Fuller Unit Root Test on DOW_JONES_INT

Null Hypothesis: DOW_JONES_INT has a unit root
Exogenous: Constant
Lag Length: 0 (Automatic - based on SIC, maxlag=25)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-44.82779	0.0001
Test critical values:	1% level	-3.433437	
	5% level	-2.862790	
	10% level	-2.567482	

Figure 4: Stationarity test for the S&P 500, NASDAQ and Dow Jones indices

The test's findings demonstrate the stationary nature of the time series data before and after decomposition as can be seen from the graph 4. The results of the ADF test disprove the null hypothesis that a unit root exists in the returns series from the data before decomposition. Compared to the crucial values, the ADF statistic is lower. If the test statistics are less than the crucial value, the null hypothesis is rejected in favour of the one-sided alternative. At all three levels of significance, the test in this study rejects the null hypothesis that a time series has a unit root. Therefore, we disprove the hypothesis that the time series are non-stationary (the internal components and external trend stationarity test results are in the appendix).

2.2.4 ARCH test

For investigating the time dynamics of the second moments, the ARCH test is an essential instrument (i.e., conditional variance). Contrary to popular belief, time-varying conditional volatility, volatility clustering, and, ultimately, the presence of a fat-tailed distribution is not necessarily indicated by the presence of a considerable excess kurtosis.

The ARCH test recommends several model types by assisting in the detection of a time-varying phenomenon in the conditional volatility. An LM test for autoregressive conditional heteroscedasticity effect was done, and the figures in the appendix show the output of the test done using the S&P 500, NASDAQ and Dow Jones indices returns, as well as their internal components and external trends, which have produced similar results. The null hypothesis was rejected, the p-value for all the time series was 0.0000, and as a result, the series have an $ARCH$ effect.

2.3 Empirical results

The $GARCH$ model's findings are presented in this section. Numerous research that looks at $GARCH$ model selection has found, according to Javed and Mantalos (2013), that the "performance of the $GARCH(1,1)$ model is satisfactory." The first lag, is enough to capture the volatility's fluctuations. Bollerslev developed the $GARCH-t$ ($GARCH$ with student t distribution) model as a special modification of the GARCH in 1987 since it has been demonstrated that $\varepsilon_t = \eta_t \sqrt{h_t}$ occasionally appears to have thicker tails compared to the normal distribution.

The tables 1, 2 and 3 represents the GARCH $GARCH(1,1)$ model where we ca see that all the parameters are significant at 1% level. The sum $(\alpha + \beta)$ coefficients of the S&P 500, NASDAQ, and Dow Jones Indices are also extremely near to one. This demonstrates how volatile shocks are relatively persistent across all stock markets. High frequency financial data typically exhibits this characteristic.

Table 1: GARCH (1,1) results for S&P 500 index

S&P index			S&P 500 index (internal components)		S&P 500 index (external trend)	
Mean equation						
Variable	Coefficient	Prob	Coefficient	Prob	Coefficient	Prob
C	0.094292	0.0000	1.050727	0.0000	0.950114	0.0000
Variance equation						
C	0.019243	0.0002	0.001041	0.0118	0.021322	0.0002
α	0.198201	0.0000	0.082149	0.0000	0.188012	0.0000
β	0.800587	0.0000	0.914884	0.0000	0.810656	0.0000
$\alpha+\beta$	0.998788		0.997033		0.998668	

Table 2: GARCH (1,1) results for NASDAQ index

NASDAQ index			NASDAQ index (internal components)		NASDAQ index (external trend)	
Mean equation						
Variable	Coefficient	Prob	Coefficient	Prob	Coefficient	Prob
C	0.133464	0.0000	1.050727	0.0000	0.969613	0.0000
Variance equation						
C	0.026492	0.0007	0.001041	0.0118	0.022240	0.0002
α	0.156832	0.0000	0.082149	0.0000	0.183145	0.0000
β	0.833263	0.0000	0.914884	0.0000	0.816383	0.0000
$\alpha+\beta$	0.990095		0.997033		0.999528	

Table 3: GARCH (1,1) results for Dow Jones index

	Dow Jones index		Dow Jones index (internal components)		Dow Jones index (external trend)	
Mean equation						
Variable	Coefficient	Prob	Coefficient	Prob	Coefficient	Prob
C	0.085481	0.0000	0.943131	0.0000	0.900586	0.0000
Variance equation						
C	0.023720	0.0001	0.000607	0.0171	0.019177	0.0002
α	0.196575	0.0000	0.086280	0.0000	0.187880	0.0000
β	0.802229	0.0000	0.912079	0.0000	0.810734	0.0000
$\alpha+\beta$	0.998804		0.998614		0.998614	

To check whether have improved the predictability of the stock returns by decomposing them, to compare forecasting methods for accuracy under quadratic loss, we used a $GARCH$ model to calculate the mean squared error. Due to the fact that $RMSE$ penalizes large forecast errors more severely than other commonly used accuracy statistics, it tends to demonstrate which method prevents large errors the best (Thompson, 1990). The table 4 represents the RMSE of the 3 stock indices before and after decomposition.

Table 4: RMSE values obtained using the GARCH (1,1) model

Stock index		RMSE
S&P 500	Returns	1.1790
	Internal components	0.1151
	External trend	1.2201
NASDAQ	Returns	1.3871
	Internal components	0.3859
	External trend	1.2502
Dow Jones	Returns	1.1776
	Internal components	0.3165
	External trend	1.1613

Using $GARCH(1,1)$ model, we got:

$$RMSE(returns) > RMSE(Int.c) \quad (8)$$

And:

$$RMSE(returns) > RMSE(Ext.t) \quad (9)$$

By summing both equations:

$$2*RMSE(returns) > RMSE(Int.c) + RMSE(Ext.t) \quad (10)$$

Then:

$$RMSE(returns) > \frac{1}{2}\left(RMSE(Int.c) + RMSE(Ext.t)\right) \qquad (11)$$

We can conclude that on average, the $RMSE(returns)$ is higher than $RMSE(Int.c)$ and $RMSE(Ext.t)$ under the same prediction model. Which means that using the $GARCH(1,1)$ model, it's better to predict the returns external parts and the internal parts separately since they lead to lower mean square error values, which confirms the effectiveness of the decomposition method in the predictability improvement. To confirm this result, we estimated the coefficients for the first 1927 observations (covering the period from January 6th 2015 to August 30th 2022), then the conditional variance for the last 64 observations, and then generated the error for each stock. The following is the formula for the $RMSE$:

$$RMSE = \sqrt{MSE} = \sqrt{\frac{1}{64} * \sum_{i=1927}^{1991} \left(r(t)_i^2 - \hat{\varepsilon}(t)_i^2\right)^2}$$

Where $\hat{\varepsilon}(t)_i^2$ is the conditional estimated variance and $r(t)_i^2$ is the squared continuously compounded rate of return for each of the three stocks at time t. The table below present the $RMSE$ of the returns, the internal components and the external trend. The results show that the return error, is grater on average: 1.6635>0.8856 (S&P 500); 2.0541>1.1094 (NASDAQ 100) and 1.3930>1.0232 (Dow Jones). We conclude than that the decomposition method leads to better forecasts with $GARCH(1,1)$.

Stock index		RMSE
S&P 500	Returns	1.6635
	Internal components	0.0876
	External trend	1.6837
NASDAQ	Returns	2.0541
	Internal components	0.5011
	External trend	1.7189
Dow Jones	Returns	1.3930
	Internal components	0.4497
	External trend	1.5968

Figure 5: RMSE for the forecasted observations

3. CONCLUSION

Numerous studies have demonstrated that the GARCH method is excellent for modeling stock data time series. Generally, scholars used it with other prediction models and compare the results and find the best fitting model. In this study, we first decomposed the returns from S&P 500, NASDAQ and Dow Jones stock indices into internal components and external parts via the external trend and internal components analysis method. Then using the $GARCH(1,1)$ model, we predicted the volatility of the returns before and after decomposition in order to explore the effect of the ETICA method and to test its efficiency in improving the predictability of the returns. Our empirical analysis showed that on average, the root mean squared error of the returns before decomposition was higher, which means that the prediction error has decreased using the decomposition method. On other words, the forecast error after decomposition is lowest on average. Because there is no such thing as a perfect model and small degrees of error will always exist, it is worthwhile to minimize this risk due to the fact that strong models always give investors solid orientations.

REFERENCES

[1] Barthélemy, M., Nadal, J. P., & Berestycki, H. (2010). Disentangling collective trends from local dynamics. Proceedings of the National Academy of Sciences, 107(17), 7629-7634.

[2] Cheteni, P. (2016). Stock market volatility using GARCH models: Evidence from South Africa and China stock markets.

[3] de Menezes, M. A., & Barabási, A. L. (2004). Separating internal and external dynamics of complex systems. Physical review letters, 93(6), 068701.

[4] Ding, Z., Granger, C. W., & Engle, R. F. (1993). A long memory property of stock market returns and a new model. *Journal of empirical finance*, *1*(1), 83-106.

[5] Dioubi, F., & Khurshid, A. (2022). How External Trends and Internal Components Decomposition Method Improve the Predictability of Financial Time Series? *Journal of Financial Risk Management*, *11*(3), 621-633.

[6] Engle, R. (1995). ARCH: selected readings. Oxford University Press.

[7] Engle, R. (2001). GARCH 101: The use of ARCH/GARCH models in applied econometrics. Journal of economic perspectives, 15(4), 157-168.

[8] Engle, R. F., & Bollerslev, T. (1986). Modelling the persistence of conditional variances. *Econometric reviews*, *5*(1), 1-50.

[9] Glosten, L. R., Jagannathan, R., & Runkle, D. E. (1993). On the relation between the expected value and the volatility of the nominal excess return on stocks. *The journal of finance*, *48*(5), 1779-1801.

[10] Granger, C. W. (1992). Forecasting stock market prices: Lessons for forecasters. *International Journal of Forecasting*, *8*(1), 3-13.

[11] Guresen E, Kayakutlu G, Daim TU (2011) Using artificial neural network models in stock market index prediction. Expert Syst Appl 38 :10389-10397.

[12] Hu, Z., Zhu, J., & Tse, K. (2013, November). Stocks market prediction using support vector machine. In 2013 6th International Conference on Information Management, Innovation Management and Industrial Engineering (Vol. 2, pp. 115-118). IEEE.

[13] Javed, F., & Mantalos, P. (2013). GARCH-type models and performance of information criteria. Communications in Statistics-Simulation and Computation, 42(8), 1917-1933.

[14] Lahmiri S (2016) A variational mode decomposition approach for analysis and forecasting of economic and financial time series. Expert Syst Appl 55 :268–273.

[15] Lee, G. G., & Engle, R. F. (1993). A permanent and transitory component model of stock return volatility. *Available at SSRN 5848*.

[16] Matei, M. (2009). Assessing volatility forecasting models: why GARCH models take the lead. *Romanian Journal of Economic Forecasting*, *12*(4), 42-65.

[17] Nelson, D. B. (1991). Conditional heteroskedasticity in asset returns: A new approach. *Econometrica: Journal of the econometric society*, 347-370.

[18] Thompson, P. A. (1990). An MSE statistic for comparing forecast accuracy across series. *International Journal of Forecasting*, *6*(2), 219-227.

[19] Zakoian, J. M. (1994). Threshold heteroskedastic models. *Journal of Economic Dynamics and control*, *18*(5), 931-955.

[20] Zhong X, Enke D (2019) Predicting the daily return direction of the stock market using hybrid machine learning algorithms. Financial Innovation 5(4) :1–20.

International Journal of Mathematics and Physical Sciences Research ISSN 2348-5736 (Online)
Vol. 10, Issue 2, pp: (57-74), Month: October 2022 - March 2023, Available at: www.researchpublish.com

APPENDIX:

Normality tests:

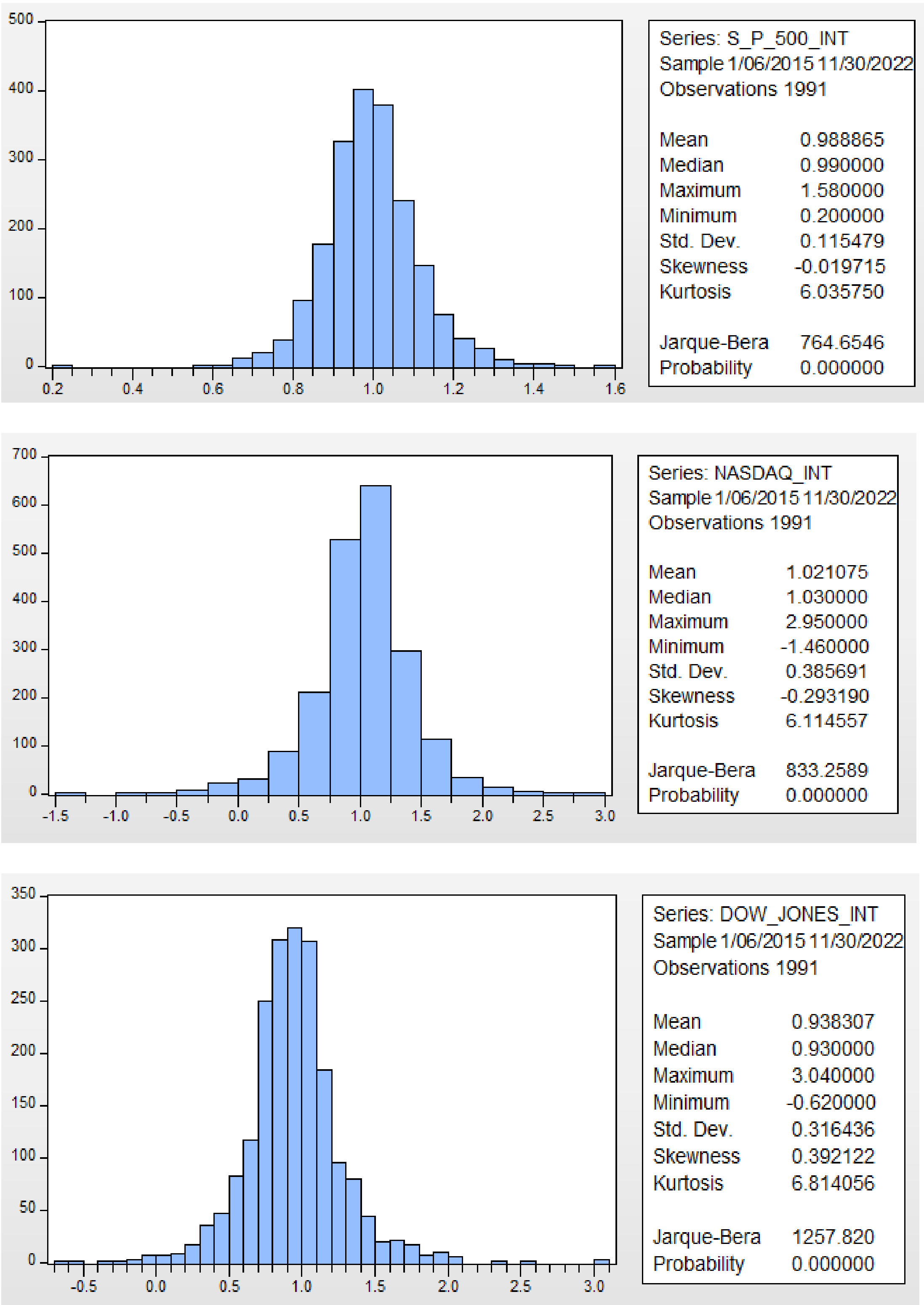
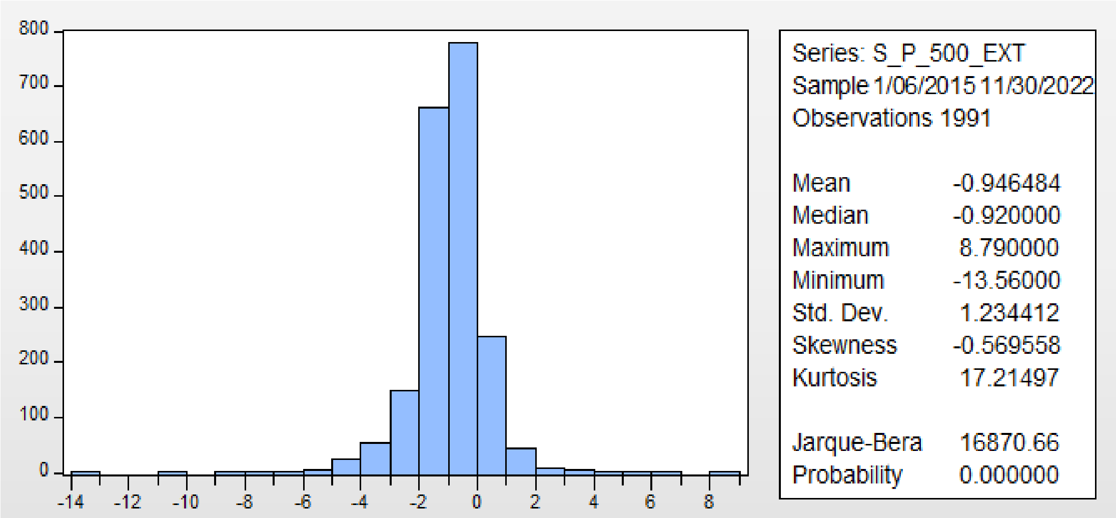
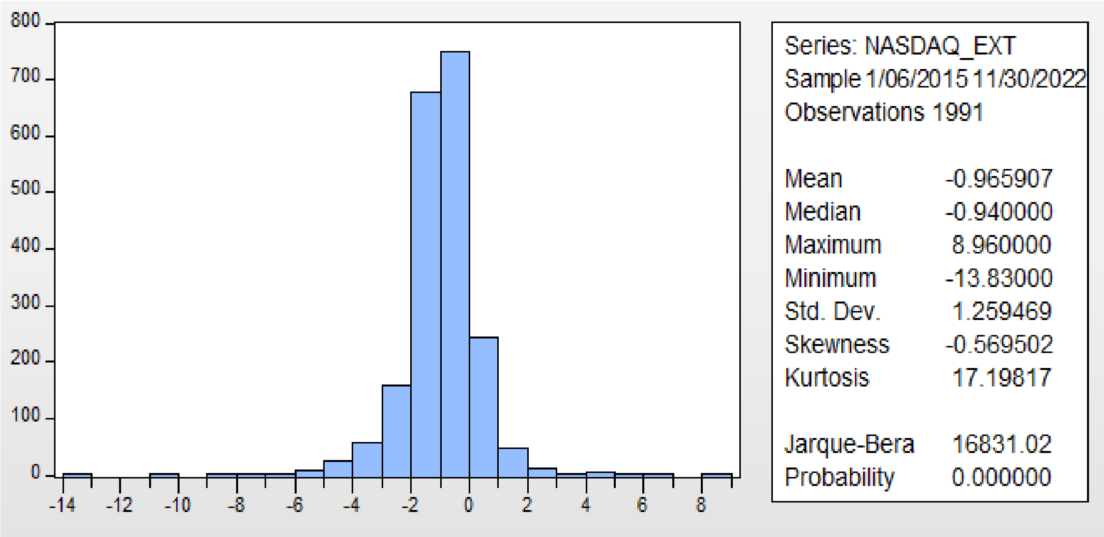


Figure 6: Normality test for stock indices internal components





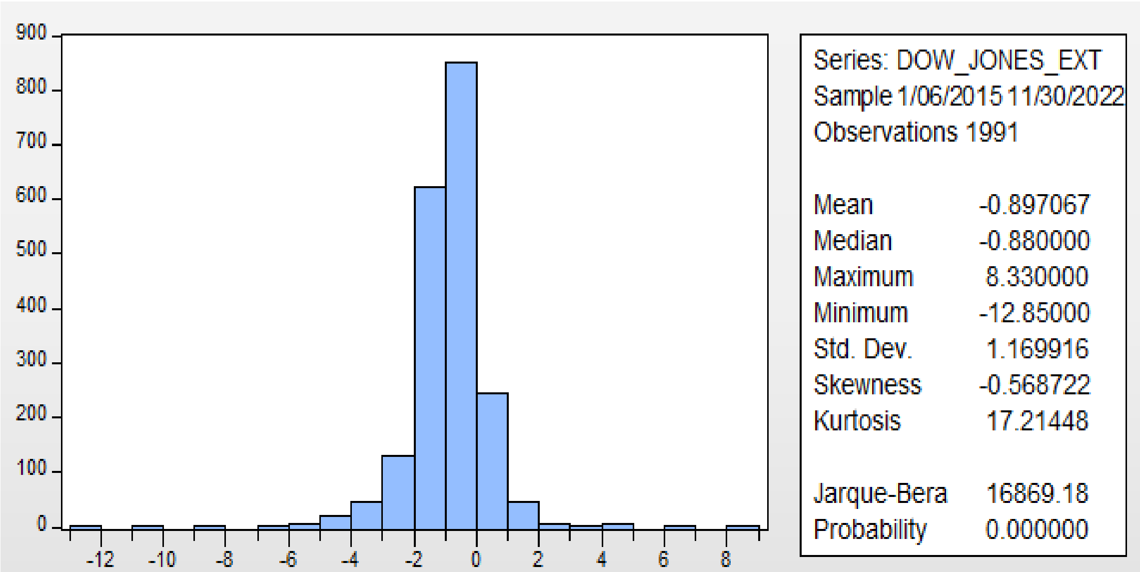


Figure 7: Normality test for stock indices external trends

Stationarity test:

Augmented Dickey-Fuller Unit Root Test on S_P_500_INT

Null Hypothesis: S_P_500_INT has a unit root
Exogenous: Constant
Lag Length: 0 (Automatic - based on SIC, maxlag=25)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-44.48524	0.0001
Test critical values:	1% level	-3.433437	
	5% level	-2.862790	
	10% level	-2.567482	

*MacKinnon (1996) one-sided p-values.

Augmented Dickey-Fuller Test Equation
Dependent Variable: D(S_P_500_INT)
Method: Least Squares

Augmented Dickey-Fuller Unit Root Test on NASDAQ_INT

Null Hypothesis: NASDAQ_INT has a unit root
Exogenous: Constant
Lag Length: 0 (Automatic - based on SIC, maxlag=25)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-44.86747	0.0001
Test critical values:	1% level	-3.433437	
	5% level	-2.862790	
	10% level	-2.567482	

*MacKinnon (1996) one-sided p-values.

Augmented Dickey-Fuller Test Equation
Dependent Variable: D(NASDAQ_INT)
Method: Least Squares

Augmented Dickey-Fuller Unit Root Test on DOW_JONES_INT

Null Hypothesis: DOW_JONES_INT has a unit root
Exogenous: Constant
Lag Length: 0 (Automatic - based on SIC, maxlag=25)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-44.82779	0.0001
Test critical values:	1% level	-3.433437	
	5% level	-2.862790	
	10% level	-2.567482	

*MacKinnon (1996) one-sided p-values.

Augmented Dickey-Fuller Test Equation
Dependent Variable: D(DOW_JONES_INT)
Method: Least Squares

Figure 8: Stationarity test for stock indices internal components

Augmented Dickey-Fuller Unit Root Test on S_P_500_EXT

Null Hypothesis: S_P_500_EXT has a unit root
Exogenous: Constant
Lag Length: 8 (Automatic - based on SIC, maxlag=25)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-14.36795	0.0000
Test critical values:	1% level	-3.433450	
	5% level	-2.862796	
	10% level	-2.567485	

*MacKinnon (1996) one-sided p-values.

Augmented Dickey-Fuller Test Equation
Dependent Variable: D(S_P_500_EXT)
Method: Least Squares

Augmented Dickey-Fuller Unit Root Test on NASDAQ_EXT

Null Hypothesis: NASDAQ_EXT has a unit root
Exogenous: Constant
Lag Length: 8 (Automatic - based on SIC, maxlag=25)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-14.36693	0.0000
Test critical values:	1% level	-3.433450	
	5% level	-2.862796	
	10% level	-2.567485	

*MacKinnon (1996) one-sided p-values.

Augmented Dickey-Fuller Test Equation
Dependent Variable: D(NASDAQ_EXT)
Method: Least Squares

Augmented Dickey-Fuller Unit Root Test on DOW_JONES_EXT

Null Hypothesis: DOW_JONES_EXT has a unit root
Exogenous: Constant
Lag Length: 8 (Automatic - based on SIC, maxlag=25)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-14.37139	0.0000
Test critical values:	1% level	-3.433450	
	5% level	-2.862796	
	10% level	-2.567485	

*MacKinnon (1996) one-sided p-values.

Augmented Dickey-Fuller Test Equation
Dependent Variable: D(DOW_JONES_EXT)
Method: Least Squares

Figure 9: Stationarity test for stock indices external trends

ARCH test

Heteroskedasticity Test: ARCH

F-statistic	405.4883	Prob. F(1,1987)	0.0000
Obs*R-squared	337.1035	Prob. Chi-Square(1)	0.0000

Test Equation:
Dependent Variable: RESID^2
Method: Least Squares
Date: 01/14/23 Time: 14:32
Sample (adjusted): 1/08/2015 11/30/2022
Included observations: 1989 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.810659	0.109366	7.412366	0.0000
RESID^2(-1)	0.411914	0.020456	20.13674	0.0000

Heteroskedasticity Test: ARCH

F-statistic	22.87528	Prob. F(1,1987)	0.0000
Obs*R-squared	22.63769	Prob. Chi-Square(1)	0.0000

Test Equation:
Dependent Variable: RESID^2
Method: Least Squares
Date: 01/14/23 Time: 14:36
Sample (adjusted): 1/08/2015 11/30/2022
Included observations: 1989 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.011914	0.000731	16.30161	0.0000
RESID^2(-1)	0.106683	0.022305	4.782811	0.0000

Heteroskedasticity Test: ARCH

F-statistic	346.2224	Prob. F(1,1987)	0.0000
Obs*R-squared	295.1439	Prob. Chi-Square(1)	0.0000

Test Equation:
Dependent Variable: RESID^2
Method: Least Squares
Date: 01/14/23 Time: 14:34
Sample (adjusted): 1/08/2015 11/30/2022
Included observations: 1989 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.914264	0.118938	7.686917	0.0000
RESID^2(-1)	0.385432	0.020714	18.60705	0.0000

Figure 10: ARCH test for S&P 500 returns, the internal components and external trend respectively

Heteroskedasticity Test: ARCH

F-statistic	316.0334	Prob. F(1,1987)	0.0000
Obs*R-squared	272.9402	Prob. Chi-Square(1)	0.0000

Test Equation:
Dependent Variable: RESID^2
Method: Least Squares
Date: 01/14/23 Time: 14:27
Sample (adjusted): 1/08/2015 11/30/2022
Included observations: 1989 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	1.201959	0.123308	9.747636	0.0000
RESID^2(-1)	0.371226	0.020882	17.77733	0.0000

Heteroskedasticity Test: ARCH

F-statistic	208.3078	Prob. F(1,1987)	0.0000
Obs*R-squared	188.7317	Prob. Chi-Square(1)	0.0000

Test Equation:
Dependent Variable: RESID^2
Method: Least Squares
Date: 01/14/23 Time: 14:31
Sample (adjusted): 1/08/2015 11/30/2022
Included observations: 1989 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.103023	0.007843	13.13560	0.0000
RESID^2(-1)	0.308785	0.021395	14.43287	0.0000

Heteroskedasticity Test: ARCH

F-statistic	346.9350	Prob. F(1,1987)	0.0000
Obs*R-squared	295.6610	Prob. Chi-Square(1)	0.0000

Test Equation:
Dependent Variable: RESID^2
Method: Least Squares
Date: 01/14/23 Time: 14:29
Sample (adjusted): 1/08/2015 11/30/2022
Included observations: 1989 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.951251	0.123746	7.687138	0.0000
RESID^2(-1)	0.385771	0.020711	18.62619	0.0000

Figure 11: ARCH test for NASDAQ returns, the internal components and external trend respectively

Heteroskedasticity Test: ARCH

F-statistic	339.9062	Prob. F(1,1987)	0.0000
Obs*R-squared	290.5461	Prob. Chi-Square(1)	0.0000

Test Equation:
Dependent Variable: RESID^2
Method: Least Squares
Date: 01/14/23 Time: 14:18
Sample (adjusted): 1/08/2015 11/30/2022
Included observations: 1989 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.844157	0.127230	6.634871	0.0000
RESID^2(-1)	0.382227	0.020732	18.43655	0.0000

Heteroskedasticity Test: ARCH

F-statistic	228.5931	Prob. F(1,1987)	0.0000
Obs*R-squared	205.2144	Prob. Chi-Square(1)	0.0000

Test Equation:
Dependent Variable: RESID^2
Method: Least Squares
Date: 01/14/23 Time: 14:25
Sample (adjusted): 1/08/2015 11/30/2022
Included observations: 1989 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.068047	0.005547	12.26704	0.0000
RESID^2(-1)	0.321850	0.021287	15.11929	0.0000

Heteroskedasticity Test: ARCH

F-statistic	346.1834	Prob. F(1,1987)	0.0000
Obs*R-squared	295.1156	Prob. Chi-Square(1)	0.0000

Test Equation:
Dependent Variable: RESID^2
Method: Least Squares
Date: 01/14/23 Time: 14:24
Sample (adjusted): 1/08/2015 11/30/2022
Included observations: 1989 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.821220	0.106825	7.687501	0.0000
RESID^2(-1)	0.385416	0.020715	18.60600	0.0000

Figure 12: ARCH test for Dow Jones returns, the internal components and external trend respectively