DO RANGE-BASED ESTIMATORS IMPROVE THE PERFORMANCE FORECASTING OF GARCH-*TYPE* MODELS? THE EVIDENCE FROM STOCK EXCHANGE OF THAILAND

BY

MR. KRITSANAPHUK LEEHANANTAKUL

AN INDEPENDENT STUDY SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF MASTER OF SCIENCE PROGRAM IN FINANCE (INTERNATIONAL PROGRAM)
FACULTY OF COMMERCE AND ACCOUNTANCY
THAMMASAT UNIVERSITY
ACADEMIC YEAR 2016

DO RANGE-BASED ESTIMATORS IMPROVE THE PERFORMANCE FORECASTING OF GARCH-*TYPE* MODELS? THE EVIDENCE FROM STOCK EXCHANGE OF THAILAND

BY

MR. KRITSANAPHUK LEEHANANTAKUL

AN INDEPENDENT STUDY SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF MASTER OF SCIENCE PROGRAM IN FINANCE (INTERNATIONAL PROGRAM)
FACULTY OF COMMERCE AND ACCOUNTANCY
THAMMASAT UNIVERSITY
ACADEMIC YEAR 2016

THAMMASAT UNIVERSITY
FACULTY OF COMMERCE AND ACCOUNTANCY

INDEPENDENT STUDY

BY

MR. KRITSANAPHUK LEEHANANTAKUL

ENTITLED

DO RANGE-BASED ESTIMATORS IMPROVE THE PERFORMANCE FORECASTING OF GARCH-*TYPE* MODELS? THE EVIDENCE FROM STOCK EXCHANGE OF THAILAND

was approved as partial fulfillment of the requirements for
the degree of Master of Science (Finance)

on 01 MAY 2017

Chairman P. Pav
(Associate Professor Pantisa Pavabutr, Ph.D.)

Member and Advisor Wasin S.
(Wasin Siwasarit, Ph.D.)

Dean P. Udorn
(Associate Professor Pipop Udorn, Ph.D.)

Ref. code: 25595802042225YAJ

Independent Study Title	DO RANGE-BASED ESTIMATORS IMPROVE THE PERFORMANCE FORECASTING OF GARCH-*TYPE* MODELS? THE EVIDENCE OF STOCK EXCHANGE OF THAILAND
Author	Mr. Kritsanaphuk Leehanantakul
Degree	Master of Science (Finance)
Major Field/Faculty/University	Master of Science Program in Finance (International Program) Faculty of Commerce and Accountancy Thammasat University
Independent Study Advisor	Wasin Siwasarit, Ph.D.
Academic Year	2016

ABSTRACT

As the actual volatility is inherently latent. We use the square return and three range-based estimators as the volatility proxy to evaluate the forecasting performance of various GARCH-*type* models from SET index return. The Root Mean Squared Error (RMSE) and Superior Predictive Ability (SPA) test of Hansen (2005) are applied to measure the goodness of fit. The empirical results indicate that the EGARCH model is superior in forecasting ability than other GARCH-*type* models for rolling out-of-sample forecasting. However, with the fixed-window, the findings are quite striking depend on the prediction horizon. In addition, the squared return as an exogenous variable can improve the forecasting performance of GARCH-*type* model while the range-based estimators cannot.

Keywords: GARCH-*type* models, Range-based estimators, Volatility forecast, Root Mean Squared Error (RMSE), Superior Predictive Ability (SPA) test

ACKNOWLEDGEMENTS

Primarily, I wish to express my sincerest appreciation to my advisor, Wasin Siwasarit, Ph.D. for his invaluable help, good guidance, suggestions, and criticism as well as encouragement throughout the processing of this research. This independent study would not have been successfully accomplished and my graduation would not be achieved without all his support and kindness. I would like to extend my special thank all the staff at MIF, Faculty of Commerce, and Accountancy at Thammasat University, for their generosity in providing all facilities and materials for my study.

Finally, I most gratefully acknowledge my family, especially my mother for her love, understanding, and encouragement during the past two years.

Mr. Kritsanaphuk Leehanantakul

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

Symbols/Abbreviations	Terms
ACF	Autocorrelation Function
ADF	Augmented Dickey–Fuller Test
AGARCH	Asymmetric Generalized Autoregressive Conditional Heteroskedasticity
AIC	Akaike Information Criterion
AR	Autoregressive Model
BIC	Bayesian Information Criterion
BM	Benchmark Model
C	Close Price
DM	Diabold and Mariano test
EWMA	Exponentially Weighted Moving Average
EGARCH	Exponential Generalized Autoregressive Conditional Heteroskedasticity
GARCH	Generalized Autoregressive Conditional Heteroskedasticity
GJR-GARCH	Glosten-Jagannathan-Runkle Generalized Autoregressive Conditional Heteroskedasticity
GK	Garman and Klass Estimation
H	High Price
JB	Jarque and Bera’s Test
KLCI	Kuala Lumpur Composite Index
KPSS	Kwiatkowski Phillips Schmidt Shin Test
L	Low Price
LB-Q	Ljung-Box Q test

Symbols/Abbreviations	**Terms**
LOG L	Log Likelihood Value
MAE	Mean Absolute Error
MAX	Maximum
MIN	Minimum
MSE	Mean Squared Error
NIID	Normal and Independent with Identical Distributions
O	Open Price
OMXC20	Offset Market Exchange Copenhagen20
OMXH25	Offset Market Exchange Helsinki 25
OMXS30	Offset Market Exchange Stockholm 30
PACF	Partial Autocorrelation Function
PK	Parkinson Estimator
PP	Phillips–Perron Test
RMSE	Root Mean Squared Error
RRV	Realized Range-based Volatility
RS	Rogers and Satchell Estimation
RV	Realized Volatility
SD	Standard Deviation
SET	Stock Exchange of Thailand
SMA	Simple Moving Average
SPA	Superior Predictive Ability Test
SPA_C	The consistent *p-value* of SPA test
SPA_L	The lower *p-value* of SPA test
SPA_U	The upper *p-value* of SPA test
SR	Squared Return
STI	Singapore Stock Market
TARCH	Threshold Autoregressive Conditional Heteroskedasticity

CHAPTER 1
INTRODUCTION

There are several type of news reports impacts the asset prices. The positive (negative) news will normally cause investors to buy (sell) the asset. Sometimes investors may notice that it difficult to interpret news that are issued at the same time. It depends on the individual who different background and experiences. It leads asset prices become volatility. A degree of variation, volatility, has been studied in several research fields, especially in empirical finance and financial econometrics. It is a crucial part of portfolio selection, risk management, asset allocation and capital asset pricing.

As the true volatility is inherently unobservable (latent). To assess and evaluate the volatility performance, we rely on ex post proxies for actual volatility. From the previous studies, Awartani and Corradi (2005) used squared returns (SR) as a proxy of actual volatility compared with different GARCH model. The results show that GARCH (1,1) is outperformed against the class of asymmetric GARCH. However, Andersen and Bollerslev (1998) pointed out squared return (SR) is unbiased estimator but extremely noisy because using only daily close price. When an open-high-low-close price are readily available, Parkinson (1980) introduced the estimation of the volatility using high-low range. In addition, Garman and Klass (1980) improved the Parkinson (1980) by adding open and close price with zero drift and no opening jumps. Moreover, Rogers and Satchell (1991) also estimated the variance from entire price process (high-low-open-close price) with non-zero drifts term and no opening jumps. We therefore apply Parkinson (1980), Garman and Klass (1980) and Rogers and Satchell (1991) as a volatility proxy.

To forecast the volatility, there are several models that try to mimic the characteristics of financial volatility. The first generation of parametric volatility models is an ARCH model of Engle (1982), which provided a way to model conditional heteroscedasticity. The GARCH, or Generalized ARCH model, proposed by Bollerslev (1986) which required fewer parameters to model the volatility process adequately. Both the ARCH and the GARCH models are designed to capture the volatility clustering effects, but cannot obtain the asymmetry effects. To overcome the limitation of those models, there are several class of GARCH-*type* models that includes leverage terms for modeling asymmetric volatility clustering. The well-known and frequency applied models are the Exponential GARCH (EGARCH) model of Nelson (1991), GJR-GARCH model of Glosten, Jagannathan, and Runkle (1993). In light of the non-normality distribution of asset returns, Kosapattarapim, C., Lin, Y. & McCrae, M. (2012) and Wennstrom (2014) suggested that assuming non-normal error distribution provides better out-of-sample forecast performance than a normal distribution. Moreover, Hung et al. (2013) found that the volatility forecasts can be enhanced by including range-based estimators.

To assess and evaluate the forecasting performance, we introduce two concepts of the Root Mean Squared Error (RMSE) and Superior Predictive Ability (SPA) test of Hansen (2005). First, Root Mean Squared Error (RMSE) with smallest RMSE is approved and accepted. In fact, it is difficult to decide whether the result is due to truly superior or merely lucky. Diabold and Mariano (1995) proposed DM test to compare the efficiency of two different forecasting model valid under general conditions including, for example, non-quadratic, asymmetric and non-Gaussian in forecast errors. However, the limitation of DM test is only compare two models in the same time. We then use Superior Predictive Ability (SPA) test followed Hansen (2005) to assessing the performance of range-based estimators. The main advantage of using the SPA test is that it can reduce potential data snooping bias.

This paper aims to investigate two objectives. The first objective is to find out the best volatility forecast of different GARCH-*type* models (ARCH, GARCH, EGARCH and GJR-GARCH) under normal and Student's t-distribution compared with various volatility proxies (SR, PK, GK and RS) for the Stock Exchange of Thailand index (SET). The second objective is to compare the performance between the normal GARCH-*type* model as the benchmark and other GARCH-*type* models incorporated with squared return and three range-based estimation (denoted as GARCH-*type*-SR, GARCH-*type*-PK, GARCH-*type*-GK and GARCH-*type*-SR) as an exogenous variable in improving the performance of volatility forecast.

CHAPTER 2
REVIEW OF LITERATURE

There are several studies investigating which GARCH-*type* models are outperformed in term of their ability to forecast compared with volatility proxy. The previous study of Phanathip (2011) test the forecasting performance in GARCH and EGARCH volatility models when adding range-based estimators' evidence in individual fourteen stocks in Thailand during the period 02 January 2002 to 30 December 2011. Moreover, Using Superior Predictive Ability (SPA) of Hansen (2005) to comparing the forecasting models and found GARCH model is more efficiency than EGARCH. The result appears to contradict from the previous study that EGARCH is more efficiency than GARCH model according to the properties of different data between in-of-sample and out-of-sample period.

Kosapattarapim, C., Lin, Y. & McCrae, M. (2012) evaluate the performance of volatility forecast of GARCH models with Six different types of error distributions (Normal distribution, Skewed Normal distribution, Student's t distribution, Skewed Student's t-distribution, Generalized Error distribution and Skewed Generalized Error distribution) using the dataset of three Asian stock markets (SET, KLCI and STI). The results show that a GARCH (p,q) model with non-normal error distributions tends to provide better out-of-sample forecast performance than a normal distributions. Moreover, the MSE and MAE given by the best fitted model is insignificantly different from that given by the best forecast performance model

Hung et al. (2013) investigate the volatility forecast of six GARCH-*type* model under normal distribution compared with the proxy of latent volatility. The proxy that Hung et al. (2013) use as volatility proxy is squared return, three range-based estimators and realized range-based volatility. The squared return is the simplest and commonly used to estimate the variance. However, this model is extremely noisy and less efficient. Parkinson (1980) introduces the estimation of the variance of the rate of return using daily high-low values and claims that using scaled high-low range value is more accurate and efficient about 2.5 - 5 times better than the squared return. Garman and Klass (1980) extended the Parkinson (1980) model by adding opening and closing values. The result finds that this simulation is better than the simplest one about eight times. Moreover, Rogers and Satchell (1991) estimates the variance from daily high-low opening and closing price but nonzero drifts term. Christensen and Podolskij (2007) , Martens and van Dijk (2007) further replace the intraday high-low range instead of point of data from Parkinson (1980). To evaluate and assess the performance of forecasting, Hung et al. (2013) use four loss function and the Superior Predictive Ability (SPA) test of Hansen (2005). The result shows that IGARCH models is better performance that other models when using range-based volatility as a proxy. The study of Hung et al. (2013) also examine the volatility proxy as an exogenous variable of GARCH-*t* model to improve the forecasting performance. The finding indicated that range-based volatility can improve the out-of-sample forecast. Furthermore, Hansen (2005) also suggests if intraday price is unavailable, Garman and Klass (1980) estimator can use to improve out-of-sample forecast of the GARCH-*t* model.

Wennstrom (2014) investigates the performance of six volatility forecasting models (SMA, EWMA, ARCH, GARCH, EGARCH and GJR-GARCH) from different Nordic equity indices (OMXS30, OMXC20 and OMXH25). There were three main themes in this study. First, the impact of assuming a student’s t-distribution provides a better in-sample fit than assuming a normal distribution. However, loss function suggested that assuming a normal error distribution provides a better out-of-sample fit. Second, normal error distribution. Second, EGARCH (1,1) is the best in-sample forecasting performance. Moreover, the higher order GARCH does not necessarily provide a better fit. Third, in term of the out-of-sample forecasting performance, the results are very inconclusive. The best and worst model respectively depends heavily on which loss function is used.

CHAPTER 3
RESEARCH METHODOLOGY

This study uses the nonparametric volatility models (Squared return and three of Range-based estimators) as a proxy of actual volatility ($\hat{\sigma}_t^2$) against parametric (GARCH family models) as the forecasting volatility ($\hat{h}_{k,t}$) in term of symmetric loss function (Root Mean Squared Error). After that, we use the superior predictive ability (SPA test) by testing the null hypothesis that "any loss function of alternative models inferior to the loss function benchmark."

The methodology as followed Section 3.1 reviews some of the non-parametric estimators (Volatility proxies). Section 3.2 explains the parametric estimators which included primitive four type of GARCH family with Gaussian and Student's t-distribution. We later added the exogenous variables into the general GARCH to test whether it improve the performance of volatility forecasting or not. We discussed the loss function (Root Mean Squared Error) in Section 3.3 and Superior Predictive Ability test in Section 3.4

3.1 Nonparametric Estimators of Volatility or Volatility Proxies ($\hat{\sigma}^2$)

Black-Scholes option pricing model assumes stock price movement is generated by the stochastic process follow a geometric Brownian motion and characterized by

$$dS_t = \mu S_t dt + \sigma S_t dZ. \qquad (1)$$

Here, S_t is the stock price, t is time, μ and σ are constant called the percentage drift and volatility which assumed constant for the moment. The variable dZ is an innovation term follows a stochastic process called Wiener process,

$$dZ = \varepsilon\sqrt{dt}, \qquad (2)$$

where ε is normal distribution zero mean and variance equal to one. Let's define the following variables for our method of estimating volatility:

O_t = daily opening price on the t^{th} trading day,

H_t = daily highest price on the t^{th} trading day,

L_t = daily lowest price on the t^{th} trading day,

C_t = daily closing price on the t^{th} trading day.

We now assume different methods of estimating the volatility such as Squared return, Parkinson (1980), Garman and Klass (1980) and Rogers and Satchell (1991)

3.1.1 Close to Close or Squared Return (SR)

The simplest and common type of calculation that use only the close price to be the volatility measure. The formula as followed

$$\widehat{\sigma^2}(SR) = \sqrt{\sum_{i=1}^{n} \ln(\frac{C_t}{C_t - 1})}. \quad (3)$$

3.1.2 Range-Base or Extreme Value

The high-low, also known as Range based estimation or extreme-value, is formed from the entire price process using daily opening, highest, lowest, and closing prices. It contains the essential information concerning the stock price volatility. A significant practical advantage of the price range provides more highly efficiency in contrast to the squared return which only using closed prices.

The classical range based estimator is developed by Parkinson (1980) based on the assumption that the asset price follows a driftless geometric Brownian motion and gives more efficiency 5.2 times than the classical estimator. His volatility estimator is given below

$$\hat{\sigma}^2_{PK} = \frac{1}{(4\, ln\, 2)} \left[\ln(\frac{H_t}{L_t})\right]^2. \quad (4)$$

Beckers (1983) proved theoretically Parkinson (1980) with Squared return estimators by using 208 stocks from 1 January 1973 through 31 March 1980. The result gives an accurately better indication of volatility because the inclusion of dispersion of prices observed over the entire day. However, instead of using closing price, Garman and Klass (1980) and Rogers and Satchell (1991) further extend Parkinson (1980) the range estimator by incorporate information about the opening and closing prices observed during the day. They assume the same diffusion process like Parkinson (1980). Their estimator provides more efficiency gain compared to the square return approximately 7.4 times. The volatility can be estimated by

$$\hat{\sigma}^2_{GK} = 0.511\left[\ln\left(\frac{H_t}{L_t}\right)\right]^2 - 0.019\left\{\left[\ln\left(\frac{C_t}{O_t}\right) x \ln\left(\frac{H_t * L_t}{{O_t}^2}\right)\right] - 2\left[\ln(\frac{H_t}{O_t}) x \ln(\frac{L_t}{O_t})\right]\right\} - 0.383[\ln\left(\frac{C_t}{O_t}\right)]^2. \quad (5)$$

This estimation is less biased because Garman and Klass (1980) subtract the squared open-to-close return to adjust the drift. However, Wiggins (1991) finds that both estimators of Parkinson (1980) and (Garman and Klass) are more efficient than close to close estimators when using continuous observation. Nevertheless, observed highest and lowest prices would understate and overstate the actual value if they calculated from discretely observed samples which give a downward biased and less efficient than the square return.

As Parkinson and Garman & Klass estimators of volatility become biased with percentage drift (μ) not equal to zero, Rogers and Satchell (1991) relax this assumption by extension the Parkinson (1980). They add a drift term in the stochastic process that can incorporate into a volatility estimator by using only daily opening, highest, lowest, and closing prices. The equation can define by

$$\hat{\sigma}^2_{RS} = \left[\ln\left(\frac{H_t}{C_t}\right) * \ln\left(\frac{H_t}{O_t}\right)\right] + \left[\ln\left(\frac{L_t}{C_t}\right) * \ln\left(\frac{L_t}{O_t}\right)\right]. \quad (6)$$

Rogers and Satchell (1991) and Rogers, Satchell, and Yoon (1994) Claims that RS estimator is outperforms than PK and GK estimators. We summarize all volatility proxies we used in Table 3.1

Table 3.1: Summarize of advanced volatility proxies

Volatility Proxies	Short Form	Prices Taken	Drift Term?	Efficiency
Close to Close	SR	C	No	1
Parkinson	PK	HL	No	5.2
Garman-Klass	GK	OHLC	No	7.4
Roger-Satchell	RS	OHLC	Yes	8

Note: C = Closing Price; O = Opening Price; H = High Price and L = Low Price

3.2 Parametric Estimators of Volatility or GARCH family ($\widehat{h}_t$)

3.2.1 The GARCH-*type* Models

To best understand the development of GARCH models, we first need to know the certain characteristics approximatively exhibit the same statistical properties that are commonly associated with all price time series of financial stocks and indexes. We list the most popular and well-known stylized facts in volatility analysis.

- **Volatility Clustering:** The shock at time $t-1$ increases not only the variance at time $t-1$ but also the variance at time t. Similarly, the volatility tends to be high at time t if it was also high at time $t-1$. Mandelbrot (1963) noted that "large changes tend to be follow by large changes and small changes tend to be follow small changes."
- **Fat Tails:** In financial data, the probability of time series return generally exhibit fatter tails than standard normal or Gaussian distribution. The fat tails also are known as leptokurtosis or kurtosis greater than three which depicts the situations that extreme outcomes have occurred more than expected or most of the variance is due to infrequent extreme deviations than predicted by the normal distribution.
- **Asymmetry:** The empirically observed sample fact that negative innovation at time $t-1$ have a greater impact on the return volatility at time t than positive innovation. On the other word, the risk increases from a negative shock more than positive shock. This asymmetry used to be called leverage effect. This is a stylized fact that the ARCH and GARCH model is not able to capture. However, there are several augmented GARCH models can capture this asymmetry effect such as TARCH, AGARCH, EGARCH, GJR-GARCH etc.

In next section, we will first start with the simple ARCH model. The fundamental tool for analyzing the time-variation of conditional variance. We second use GARCH model to address the ARCH drawbacks. Furthermore, we exhibit EGARCH and GJR-GARCH which can capture the leverage effect.

3.2.1.1 ARCH Model

In this section, we focus on financial applications. We first start with the assets return equation. Let r_t denote the daily log return defined relative to a past information set which consists of two variables. One is the conditional mean (μ_t) and another is the error term or innovation process (ε_t) which error variance is time-varying (so that there is heteroskedastic all the time). The formula is represented by

$$r_t = \mu_t + \varepsilon_t. \qquad (7)$$

The distribution of error term is conditionally normal as below

$$\varepsilon_t | I_{t-1} = h_t z_t \text{ and } \{ z_t \} \sim IID\ (0,1), \qquad (8)$$

where I_{t-1} is the information available at time *t-1* and z_t is independent identically distributed (i.i.d.) sequence with $E[z_t] = 0$ and $E[{z_t}^2] = 1$.

In fact, before ARCH, the primary descriptive tool to capture time-varying conditional volatilities is rolling sample windows which incorporate actual data for estimation. The standard deviation calculated using a fixed number of the most recent observations. We would be able to estimate the rolling sample of volatilities based on the *n* most recent observations,

$$h_t^2 = \frac{\sum_{i=1}^{n}(y_{t-i} - \hat{\mu})^2}{n} \equiv \frac{\sum_{i=1}^{n}\hat{\varepsilon}_{t-i}^2}{n}. \qquad (9)$$

The volatility h_t^2 depends on the average of squared residual terms. The total observation *(n)* directly determines the variance-bias of the estimator. The larger of observation decrease the variance but increase the bias. For instance, a rolling sample of variance could be calculated every day for twelve months using the most recent day of data (365 days). Similarly, we probably think of this formulation assumes that the variance of tomorrow's return is an equally weighted average of the squared residuals of the last 365 days.

Nevertheless, there is some drawback of using rolling sample window to estimate the volatility. The variance changes slowly over the sample period, and it is therefore approximately constant on a short rolling-time window. However, instead of equally weight each of the most recent *n* observations, it is reasonable if we consider that more recent samples are more relevant or should have higher weights than the longer sample.

An ARCH (autoregressive conditionally heteroscedastic) model are commonly used in model financial time series with time-varying volatility proposed by Engle (1982). ARCH model assumes the variance of the current error term or innovation is related to the previous time periods' error terms and weights are parameters to be estimated. Engle's ARCH model determines the best weights to use in forecasting the volatility. The formulation of the ARCH model, the variance is forecasted as a moving average of past error terms. The conditional volatility (h_t^2) modeled by

$$h_t^2 = \omega_0 + \alpha_1 \varepsilon_{t-1}^2 + \cdots + \alpha_n \varepsilon_{t-p}^2 = \omega_0 + \sum_{i=1}^{p} \alpha_i \varepsilon_{t-i}^2, \qquad (10)$$

where $\omega_0 > 0$ and $\alpha_i \geq 0$, i=1, 2, …, *p* to assure that the conditional variance value is positive and $\alpha_i < 1$ for stationary. The order of *p* is important to specify before fitting the model which can be found by Sample Partial Autocorrelation function of the squared returns. The weakness of ARCH model is usually required a high order of *p* such as the suggestion of ACF or PACF is more than 10 which is infeasible or unpractical. However, in this study will be restricted to the order of 1.

The forecasts of the ARCH model are obtained recursively as the forecasts of an AR model. If we consider an ARCH(*p*) model at the forecast origin *t*, the one-step ahead forecast of h_{t+1}^2 is

$$h_{t+1|t}^2 \; = \; \omega_0 + \alpha_1 \varepsilon_t^2 + \cdots + \alpha_p \varepsilon_{t+1-p}^2. \qquad (11)$$

The 2-step ahead forecast equation given the information at time *t* defined by

$$h_{t+2|t}^2 = \omega_0 + \alpha_1 h_{t+2|t}^2 + \alpha_2 \varepsilon_t^2 + \cdots + \alpha_p \varepsilon_{t+2-p}^2. \qquad (12)$$

We repeat the procedure to j-step ahead forecast for h_{t+j}^2 given the information at time t. The equation is represented by

$$h_{t+j|t}^2 = \omega_0 + \sum_{i=1}^{p} \alpha_i h_{t+(j-i)|t}^2,$$
$$\text{where } h_{t+(j-i)|t}^2 = \varepsilon_{t+j-1}^2 \text{ if } j-1 \leq 0 \qquad (13)$$

However, there are some major drawbacks of ARCH model. Firstly, ARCH model requires many parameters to precisely describe the volatility process. We extent ARCH (∞) in equation (10) as

$$h_t^2 = \omega_0 + \alpha_1 \varepsilon_{t-1}^2 + \alpha_2 \varepsilon_{t-2}^2 + \alpha_3 \varepsilon_{t-3}^2 + \cdots + \alpha_\infty \varepsilon_{t-\infty}^2. \qquad (14)$$

The conditional variance depends on ∞ lagged squared residual in which $\alpha_1, \alpha_2 \ldots \alpha_\infty$ must be estimated under restrictions. To address long lagged length effects with fewer parameters, Bollerslev (1986) purposed the generalized ARCH or GARCH (*p*,*q*) model that we will explain in next section

Another limitation that ARCH model unable to model the asymmetric effects of positive and negative shocks. However, there are many extensions of ARCH model that included the leverage effects into the conditional variance like EGARCH, GJR-GARCH, QGARCH, TARCH etc.

3.2.1.2 GARCH Model

Bollerslev (1986) introduced generalized autoregressive conditional heteroskedasticity (GARCH) which requires less parameters to adequately model the volatility process. However, GARCH model lack of ability to model the asymmetric effects of positive and negative shocks. The GARCH model is quite like the ARCH model by adding q lags of past conditional volatility,$h_{t-1}^2, h_{t-2}^2, \ldots, h_{t-q}^2$, which acts as a smoothing term to the ARCH model. The conditional volatility of GARCH model defined by

$$h_t^2 = \omega_0 + \sum_{p=1}^{p} \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^{q} \beta_j h_{t-j}^2, \quad (15)$$

where ω_0, α_i and β_j is nonnegative ($\omega_0 > 0$, $\alpha_i \geq 0$ and $\beta_j \geq 0$). However, there are some limitation to assume that the summation of the parameter must less than one ($\alpha_I + \beta_j < 1$). This condition is necessary and sufficient for the process for covariance stationary. Nevertheless, the GARCH model can alternatively be expressed as an ARCH model. We begin transform GARCH model into ARCH model by continually substitute its past conditional variance. GARCH (1,1) can rewritten in term of ARCH as

$$\begin{aligned}
h_t^2 &= \omega_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 h_{t-1}^2 \\
h_t^2 &= \omega_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta_1(\omega_0 + \alpha_1 \varepsilon_{t-2}^2 + \beta_1 h_{t-2}^2) \\
h_t^2 &= \omega_0 + \beta_1 \omega_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 \alpha_1 \varepsilon_{t-2}^2 + \beta_1^2 h_{t-2}^2 \\
h_t^2 &= \omega_0 + \beta_1 \omega_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 \alpha_1 \varepsilon_{t-2}^2 + \beta_1^2(\omega_0 + \alpha_1 \varepsilon_{t-3}^2 + \beta_1 h_{t-3}^2) \\
h_t^2 &= (\omega_0 + \beta_1 \omega_0 + \beta_1^2 \omega_0) + (\alpha_1 \varepsilon_{t-1}^2 + \beta_1 \alpha_1 \varepsilon_{t-2}^2 + \beta_1^2 \alpha_1 \varepsilon_{t-3}^2) + \beta_1^3 h_{t-3}^2 \\
h_t^2 &= \sum_{i=1}^{\infty} \beta_1^i \omega_0 + \sum_{i=1}^{\infty} \beta_1^i \alpha_1 \varepsilon_{t-i-1}^2.
\end{aligned} \quad (16)$$

As we can see that the conditional variance at time *t* consists of a constant, $\beta_1^i \omega_0$, and the past square residual term weighted with $\beta_1^i \alpha_1$. The equation (16) obviously same as the equation (10) for $\sum_{i=1}^{\infty} \beta_1^i \omega_0 = \omega_0$ and $\beta_1^i \alpha_1 = \alpha_1$

To forecast the one-step ahead forecast of h_{t+1}^2 equation given the information at time *t*, the forecasts of the GARCH model are obtained recursively as the forecasts of an ARMA model. The equation can be written by

$$h_{t+1|t}^2 = \omega_0 + \alpha_1 \varepsilon_t^2 + \beta_1 h_t^2. \quad (17)$$

However, we can rewrite the one step ahead of forecast the volatility equation (17) by substituting $\varepsilon_t^2 = h_t^2 z_t^2$ and adding $\alpha_1 h_t^2$. The equation as followed

$$h_{t+1|t}^2 = \omega_0 + \alpha_1 h_t^2 z_t^2 + \beta_1 h_t^2 + \alpha_1 h_t^2 - \alpha_1 h_t^2. \quad (18)$$

We rearrange the equation (18) which give

$$h_{t+1|t}^2 = \omega_0 + (\alpha_1 + \beta_1)h_t^2 + \alpha_1 h_t^2(z_t^2 - 1). \quad (19)$$

The two-step ahead forecast equation given the information at time *t* defined by

$$h_{t+2|t}^2 = \omega_0 + (\alpha_1 + \beta_1)h_{t+1|t}^2 + \alpha_1 h_{t+1|t}^2(z_{t+1}^2 - 1). \quad (20)$$

With $E(z_{t+1}^2 - 1) = 0$, the two-step ahead forecast can rewritten by

$$h_{t+2|t}^2 = \omega_0 + (\alpha_1 + \beta_1)h_{t+1|t}^2 \quad (21)$$

We repeat the procedure to *j*-step ahead forecast for h_{t+j}^2 with the information at time *t*. The equation is represented by

$$h_{t+j|t+j-1}^2 = \omega_0 + (\alpha_1 + \beta_1)h_{t+j-1|t}^2. \quad (22)$$

GARCH model is not a perfect model because GARCH model treats "bad news" and "good news" symmetrically as same as ARCH model. However, in financial world, positive and negative shocks impact asymmetric effects on volatility. In general, negative shocks tend to highly impact on volatility more positive shocks.

3.2.1.3 EGARCH Model

According to the limitation of ARCH and GARCH models, Exponential GARCH (EGARCH) is proposed by Nelson (1991). First, ARCH and GARCH models assume the only magnitude of unanticipated excess returns determines volatility. However, from Nelson (1991) argued that not only the magnitude but also the direction of the returns affects volatility. The negative shocks tend to impact volatility more than positive shocks. Another limitation is the persistence of volatility shock. How long does a shock influence the volatility estimation? Some may persist for a finite period while other might persist indefinitely which they may move the whole term structure market volatility. Moreover, nonnegative constraint imposed to ensure that h_t is nonnegative for all period which leads the estimated volatility ruling out random oscillatory behavior in h_t. The limitation can create difficulties in estimating GARCH model of Engle and Yoo (1987)

The EGARCH model provides the first explanation for the h_t in function of time and lagged z_t. The equation is followed

$$\ln(h_t^2) = \omega_0 + \sum_{i=1}^{p} g(z_{t-i}) + \sum_{j=1}^{q} \beta_j \ln(h_{t-j}^2). \quad (23)$$

The logarithm of h_t ensures that the conditional variances are positive and a function of past z_t. Additionally, the EGARCH specification can solve several limitations, such as large negative shocks having a greater impact on conditional volatility than positive shocks, and small positive shocks having a greater impact on conditional volatility than small negative shocks. To illustrate the asymmetrical effects of positive and negative return, we consider the function g_i defined by

$$g(z_{t-i}) = \alpha_i z_{t-i} + \gamma_i[|z_{t-i}| - E(z_{t-i})]. \quad (24)$$

The parameter ω_0, β_j, α_i, θ and λ are not required the restriction to assure that the conditional variance is nonnegative and z_t ~ NID (0,1) is the standardized residual which equal to $\frac{\varepsilon_{t-i}}{\sqrt{h_{t-i}}}$ and $E(z_t)$ depends on what error distribution we assume. If the distribution is normal, we have

$$E(z_t) = (2/\pi)^{\frac{1}{2}}. \quad (25)$$

The term of $g_i(z_{t-i})$ is a function of both magnitude and sign of z_t. The first term $\gamma[|z_{t-i}| - E(z_t)]$ represent a magnitude effect or the symmetric effect of the model. The second term $\alpha_i \, z_{t-i}$ measures the asymmetry or the leverage effect. This enables h_t to respond asymmetrically to positive and negative value of ε_t. In summary, the different impact on the volatility between the positive and negative shocks is the main advantage of the EGARCH model compared to the ARCH and GARCH models.

We can forecast the one-step ahead of h_{t+1}^2 from the origin *t*. The equation can be

$$\ln(h_{t+1|t}^2) = \omega_0 + \alpha_1 z_t + \gamma_1[|z_t| - E(z_t)] + \beta_1 \ln(h_t^2). \quad (26)$$

However, the j^{th} periods ahead volatility forecast of the EGARCH (1,1) model is not presented in this paper due to the limitation of time. Moreover, we decide to use ready-to-use MATLAB code to forecast j^{th} periods ahead volatility as same as other models.

3.2.1.4 GJR-GARCH Model

Glosten, Jagannathan, and Runkle (1993) constructed a popular alternative type of asymmetry model called GJR-GARCH model also discusses the importance of another type of leverage. In finance, a negative return is important because it refers the future losses, while positive returns are to be suppressed as they bring profits and not part of risk. GJR-GARCH model introduces two leverage parameters to capture the importance of negative returns. The GJR-GARCH equation as followed by

$$h_t^2 = \omega_0 + \sum_{i=1}^{p} \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^{q} \beta_j h_{t-j}^2 + \sum_{i=1}^{p} \gamma_i I_{t-1} \varepsilon_{t-i}^2, \quad \text{where: } I_{t-1} = \begin{cases} 1 & if \;\; \varepsilon_{t-i} < 0 \\ 0 & \;\;\;\;\;\; \varepsilon_{t-i} > 0 \end{cases}. \tag{27}$$

The parameter ω_0>0, $\alpha_i \geq 0$, $\beta_j \geq 0$ and $\gamma_i \geq 0$ to ensure that the conditional volatility is nonnegative. The first and second term as same as GARCH model, but the third term $\sum_{i=1}^{n} \gamma_i I_{t-1} \varepsilon_{t-i}^2$ depicts the asymmetry effect which I_{t-1} give the weighted of 0 if the error is positive and if it is negative it will assign a weight equal to 1. This will capture the effect of negative return more than the positive return. The one-step ahead forecast of h_{t+1}^2 is

$$h_{t+1}^2 = \omega_0 + \alpha_1 \varepsilon_t^2 + \gamma_1 \varepsilon_t^2 \, I_t + \beta_1 h_t^2. \tag{28}$$

When we compute multistep ahead forecasts of the conditional variance, we take an expectation (E) and rewritten using $\varepsilon_t^2 | I_{t-1}$ = $h_t^2 z_t^2$ which gives

$$E(h_{t+1}^2) = \omega_0 + \alpha_1 E(h_t^2) E(h_t^2) + \gamma_1 E(h_t^2\,) E(z_t^2)\; E(I_t) + \beta_1 E(h_t^2). \tag{29}$$

From $I_t = \begin{cases} 1 & if \;\; \varepsilon_i < 0 \\ 0 & \;\;\;\;\;\; \varepsilon_t > 0 \end{cases}$ which give $E[I_t] = 1/2$ and $E[{z_t}^2] = 1$,then we substitute $E[I_{t-1}] = 1/2$ and $E[{z_t}^2] = 1$ into the equation (29) as followed

$$E(h_{t+1}^2) = \omega_0 + \alpha_1 E(h_t^2) + \frac{\gamma_1}{2} E(h_t^2) + \beta_1 E(h_t^2). \tag{30}$$

We can rearrange the equation (30) which defined by

$$h_{t+1}^2 = \omega_0 + (\alpha_1 + \frac{\gamma_1}{2} + \beta_1) h_t^2. \tag{31}$$

We repeat the procedure to *j*-step ahead forecast for h_{t+j}^2 and the equation is represented by

$$h_{t+j}^2 = \omega_0 + (\alpha_1 + \frac{\gamma_1}{2} + \beta_1)^{j-1} h_{t+1}^2. \tag{32}$$

Table 3.2: Summarize GARCH family in term of one lag

ARCH	$h_t = \omega_0 + \alpha_1 \varepsilon_{t-1}^2$
GARCH	$h_t = \omega_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 h_{t-1}$
EGARCH	$\log(h_t) = \omega_0 + \alpha_1 \frac{\varepsilon_{t-1}}{\sqrt{h_{t-1}}} + \gamma_1 [\left\lvert \frac{\varepsilon_{t-1}}{\sqrt{h_{t-1}}} \right\rvert - (2/\pi)^{\frac{1}{2}}] + \beta_1 \ln(h_{t-1})$
GJR-GARCH	$h_t = \omega_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 h_{t-1} + \gamma_1 1_{(\varepsilon_{t-1}<0)} \varepsilon_{t-1}^2$

Note: ARCH: autoregressive conditional heteroskedasticity; GARCH: generalized autoregressive conditional heteroskedasticity autocorrelation; EGARCH: Exponential GARCH and GJR-GARCH: Glosten, Jagannathan, and Runkle GARCH.

3.2.2 The GARCH-*type* Models under student's t-distribution

The normal ARCH model that we mentioned before are typically estimated by maximum likelihood under the assumption of the conditional error distribution is Gaussian (normal). However, in many empirical types of research, the standardized residuals appear to have fatter tail or leptokurtic than normality distributed. Bollerslev (1987) proposed Student's *t* GARCH model which assuming standardized Student's t-distribution instead of normality distributed. Furthermore, the evident in data description in chapter 4 display the return series is typically not normally distributed but significantly heavier tails which imply that in this paper two different types of error distributions are considered. The probability density of the normal distribution is

$$f(e) = \frac{1}{\sqrt{2\pi}\sigma} e^{\frac{(e-\mu)^2}{2\sigma^2}}, \qquad (33)$$

where $\mu, \sigma\ and\ \sigma^2$ denotes the expectation of the distribution, standard deviation and variance of sample error terms. The density function of the Student's t-distribution is defined by

$$f(e) = \frac{\Gamma(\frac{v+1}{2})}{\Gamma(\frac{v}{2})} \frac{1}{\sqrt{v\pi}} (1 + \frac{e^2}{v})^{-(\frac{v+1}{2})}, \qquad (34)$$

where v is the number of degrees of freedom and $\Gamma(.)$ is the gamma function.

3.2.3 Explanatory exogenous variable in the GARCH-*type* models

There are several studies added exogenous variables into the conditional mean equation. Not only the conditional mean equation, but exogenous explanatory variables also added into the conditional variance equation. To address the second objective of this paper that the squared return and three range-based estimation in case of Stock Exchange of Thailand index (SET) can be improve the forecasting performance of GARCH-*type* model under normal and Student’s t-distribution. We added the squared return and three range-based estimation (Parkinson (1980) , (Garman and Klass) and Rogers and Satchell (1991)) as the exogenous variables of GARCH-*type* models. We added exogenous parameter ($\delta proxy_{t-1}$) into the general GARCH-*type* models in equation (10), (15), (23) and (27) see Table 3.3.

Table 3.3: Summarize GARCH family with exogenous variable in term of one lag

ARCH	$h_t = \omega_0 + \alpha\varepsilon_{t-1}^2 + \delta proxy_{t-1}$
GARCH	$h_t = \omega_0 + \alpha\varepsilon_{t-1}^2 + \beta h_{t-1} + \delta proxy_{t-1}$
EGARCH	$\log(h_t) = \omega_0 + \alpha\frac{\varepsilon_{t-1}}{\sqrt{h_{t-1}}} + \gamma[\left\lvert\frac{\varepsilon_{t-1}}{\sqrt{h_{t-1}}}\right\rvert - (2/\pi)^{\frac{1}{2}}] +$ $\beta\ln(h_{t-1}) + \delta proxy_{t-1}$
GJR-GARCH	$h_t = \omega_0 + \alpha\varepsilon_{t-1}^2 + \beta h_{t-1} + \gamma 1_{(\varepsilon_{t-i}<0)}\varepsilon_{t-i}^2 + \delta proxy_{t-1}$

Note: ARCH: autoregressive conditional heteroskedasticity; GARCH: generalized autoregressive conditional heteroskedasticity autocorrelation; EGARCH: Exponential GARCH and GJR-GARCH: Glosten, Jagannathan, and Runkle GARCH.

3.3 Forecasting GARCH-*type* models

Suppose that we have T sample observations (1, 2, …, T). For the estimated conditional volatility ($\hat{h}_{k,t}$), we separate the entire period into two sub-periods. One is estimation period (in-of-sample period), another one is forecasting period (out-of-sample). To forecast the condition variance, we use two methods to forecast out-of-sample GARCH-*type* models. First, we estimate a fixed number of observation to predict the *h*-periods ahead. In other word, we estimate and forecast *h*-periods ahead from in-of-sample period. Second, we apply the rolling out-of-sample forecast to predict GARCH-*type* models by updated the new information and removed the old one. (see Figure 3.1).

Figure 3.1: Fixed window out-of-sample forecast

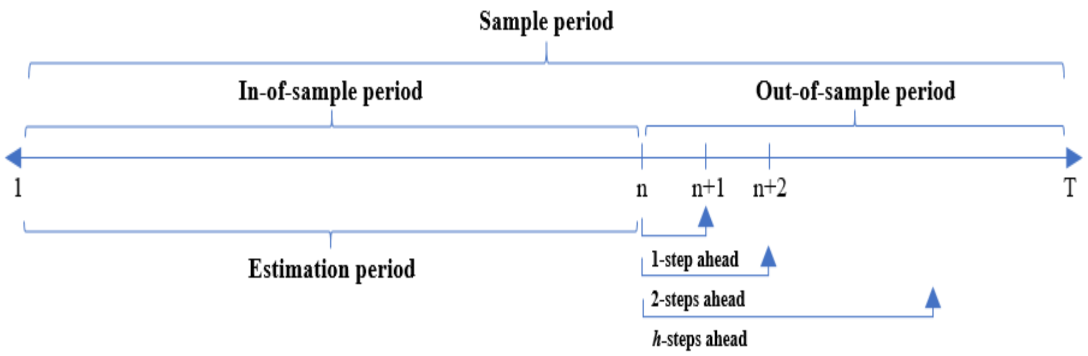
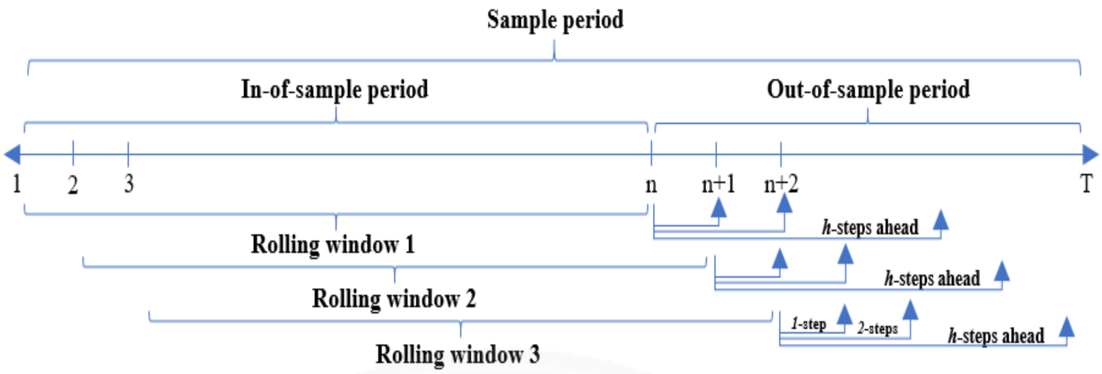


Figure 3.2: Rolling window out-of-sample forecast



However, for rolling window out-of-forecast (see Figure 3.2), We choose the observation per rolling window as the same size of in-of-sample period. The first rolling window sub-period contains observation from 1 through *n*. We delete one observation at the beginning and add new one observation. Then, second rolling window sub-period contains observation from 2 through *n+1* and so on. The rolling window sub-period is increase by one. We re-estimated and forecast *h*-periods ahead out-of-sample forecast compared with the volatility at the same time in terms of loss function and SPA test.

3.4 Loss Function

To evaluate the performance of volatility forecast models, we use symmetric loss functions which emphasize equally of the errors problems no matter what the errors resulting from over forecasting or under forecasting. To examine whether the forecasting performance of ranged-based estimators could be possible for true proxy in GARCH-*type* models, we use Root Mean Square Error (RMSE) define as follows,

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(\hat{\sigma}_t^2 - \hat{h}_{t,k}^2\right)^2}, \tag{35}$$

where n is the total number of daily observations.$\hat{\sigma}_t^2$ is the true proxy of volatility as we computed before and $\hat{h}_{k,t}$ is the estimated volatility produced by model k at time t

The limitations of loss function are assumed to be quadratic, Gaussian and serially uncorrelated in forecast errors. However, there are recent studies which test under more relaxed condition. Diebold and Mariano (1995) examines whether two forecasts have equal predictive ability. DM test also directly applicable to non-quadratic, non-Gaussian, non-zero-mean, serially correlated, and contemporaneously correlated in forecast errors. Moreover, the Reality Check of White (2000) address some drawback of DM test that examine a set of forecasting models instead of two competing models of Diebold and Mariano (1995).

3.5 Superior Predictive Ability (SPA test)

Hansen (2005) proposed Superior predictive ability test which quite similar to White (2000), but constructs in a different way. SPA test is a test that used to compare the forecasting performance of two or more models in term of loss function. The best performance of forecasting models is the model that give the smallest expected loss. The loss function we use is symmetric loss function as we mention above. SPA test helps to reduce the sensitivity of poor predicting models and improves the performance of the test.

We first set up a hypothesis test which consist of a null hypothesis and an alternative hypothesis test to determine the validity of a statistical claim. If the benchmark model is outperformed the other models, then the mean of each element of $\delta_{k,t}$ should be 0 or negative where $\delta_{k,t}$ is the vector of excess performance between the benchmark and alternative *m* models at time *t*. The null hypothesis in a test of SPA is

$$H_0: \max_{k=1,2,\dots,m}(E[\delta_{k,t}]) \leq 0, \qquad (36)$$

and alternative hypothesis,

$$H_1: \max_{k=1,2,\dots,m}(E[\delta_{k,t}]) > 0. \qquad (37)$$

The standard example of loss function which quantifies the amount of the prediction deviates from the actual values for comparing models is

$$L\left(\sigma^2_{t+j}, \hat{h}^2_{k,t+j|t}\right) = (\sigma^2_{t+j} - \hat{h}^2_{k,t+j|t})^2, \qquad (38)$$

where σ^2_{t+j} is the non-parametric models, Squared Return, Parkinson (1980), Garman and Klass (1980), and Rogers and Satchell (1991) as the volatility proxies or the actual value and $\hat{h}^2_{k,t+j|t}$ is *j*-step periods ahead given the information at time *t* of GARCH family models as the comparing volatility model. We can write the differential of loss function between benchmark (*BM*) and one alternative models (*k=1*) into *j* x *1* vector form,

$$\delta_{j,1} = \begin{bmatrix} \delta_{1,1} \\ \delta_{2,1} \\ \vdots \\ \delta_{n,1} \end{bmatrix} = \begin{bmatrix} L\left(\sigma^2_{t+1}, \hat{h}^2_{t+1|t,BM}\right) - L\left(\sigma^2_{t+1}, \hat{h}^2_{t+1|t,1}\right) \\ L\left(\sigma^2_{t+2}, \hat{h}^2_{t+2|t,BM}\right) - L\left(\sigma^2_{t+2}, \hat{h}^2_{t+2|t,1}\right) \\ \vdots \\ L\left(\sigma^2_{t+n}, \hat{h}^2_{t+n|t,BM}\right) - L\left(\sigma^2_{t+n}, \hat{h}^2_{t+n|t,1}\right) \end{bmatrix} \equiv \begin{bmatrix} (\sigma^2_{t+1} - \hat{h}^2_{t+1|t,BM})^2 - (\sigma^2_{t+1} - \hat{h}^2_{t+1|t,1})^2 \\ (\sigma^2_{t+2} - \hat{h}^2_{t+2|t,BM})^2 - (\sigma^2_{t+2} - \hat{h}^2_{t+2|t,1})^2 \\ \vdots \\ (\sigma^2_{t+n} - \hat{h}^2_{t+n|t,BM})^2 - (\sigma^2_{t+n} - \hat{h}^2_{t+n|t,1})^2 \end{bmatrix}, \tag{39}$$

where *j=1, 2, …, n.* We can rewrite equation (39) in to *j* x *k* matrix, then the loss differentials vector is

$$\delta_{j,k} = \begin{bmatrix} \delta_{1,1}, \delta_{1,2} & \cdots & \delta_{1,m} \\ \delta_{2,1}, \delta_{2,2} & & \delta_{2,m} \\ \vdots & \ddots & \vdots \\ \delta_{n,1}, \delta_{n,2} & \cdots & \delta_{n,m} \end{bmatrix} = \begin{bmatrix} L\left(\sigma^2_{t+1}, \hat{h}^2_{t+1|t,BM}\right) - L\left(\sigma^2_{t+1}, \hat{h}^2_{t+1|t,1}\right) & \cdots & L\left(\sigma^2_{t+1}, \hat{h}^2_{t+1|t,BM}\right) - L\left(\sigma^2_{t+1}, \hat{h}^2_{t+1|t,m}\right) \\ L\left(\sigma^2_{t+2}, \hat{h}^2_{t+2|t,BM}\right) - L\left(\sigma^2_{t+2}, \hat{h}^2_{t+2|t,1}\right) & & L\left(\sigma^2_{t+2}, \hat{h}^2_{t+2|t,BM}\right) - L\left(\sigma^2_{t+2}, \hat{h}^2_{t+2|t,m}\right) \\ \vdots & \ddots & \vdots \\ L\left(\sigma^2_{t+n}, \hat{h}^2_{t+n|t,BM}\right) - L\left(\sigma^2_{t+n}, \hat{h}^2_{t+n|t,1}\right) & \cdots & L\left(\sigma^2_{t+n}, \hat{h}^2_{t+n|t,BM}\right) - L\left(\sigma^2_{t+n}, \hat{h}^2_{t+n|t,m}\right) \end{bmatrix}, \tag{40}$$

where *j=1,2,…,n and k=1,2,…,m*. Suppose we first choose Parkinson (1980) as the volatility proxy and *j*-step periods ahead forecasting of ARCH as the benchmark model. Another comparing model is one-step period ahead of GARCH, EGARCH and GJR-GARCH models as the alternative models. We can write the simple form for comparing models using RMSE is

$$\delta_{j,k} = \begin{bmatrix} L\left(\sigma^2_{1,PK}, \hat{h}^2_{1|0,ARCH}\right) - L\left(\sigma^2_{1,PK}, \hat{h}^2_{1|0,GARCH}\right), & L\left(\sigma^2_{1,PK}, \hat{h}^2_{1|0,ARCH}\right) - L\left(\sigma^2_{1,PK}, \hat{h}^2_{1|0,EGARCH}\right), & L\left(\sigma^2_{1,PK}, \hat{h}^2_{1|0,ARCH}\right) - L\left(\sigma^2_{1,PK}, \hat{h}^2_{1|0,GJR}\right) \\ L\left(\sigma^2_{2,PK}, \hat{h}^2_{2|0,ARCH}\right) - L\left(\sigma^2_{2,PK}, \hat{h}^2_{2|0,GARCH}\right), & L\left(\sigma^2_{2,PK}, \hat{h}^2_{2|0,ARCH}\right) - L\left(\sigma^2_{2,PK}, \hat{h}^2_{2|0,EGARCH}\right), & L\left(\sigma^2_{2,PK}, \hat{h}^2_{2|0,ARCH}\right) - L\left(\sigma^2_{2,PK}, \hat{h}^2_{2|0,GJR}\right) \\ \vdots & \vdots & \vdots \\ L\left(\sigma^2_{n,PK}, \hat{h}^2_{n|0,ARCH}\right) - L\left(\sigma^2_{n,PK}, \hat{h}^2_{n|0,GARCH}\right), & L\left(\sigma^2_{n,PK}, \hat{h}^2_{n|0,ARCH}\right) - L\left(\sigma^2_{2,PK}, \hat{h}^2_{n|0,EGARCH}\right), & L\left(\sigma^2_{n,PK}, \hat{h}^2_{n|0,ARCH}\right) - L\left(\sigma^2_{n,PK}, \hat{h}^2_{n|0,GJR}\right) \end{bmatrix} \tag{41}$$

To test the hypothesis, we followed White (2000) who developed a framework for evaluating the performance of multiple forecasting models and propose a Reality Check (RC) to eliminate data snooping bias. The statistic bias that appears when the researcher decides to use the dataset more than once for model selection or conclusion. When the set of data reuse, there are possibility using of satisfactory results obtained from the previous test to uncover misleading relationships in data. The two methods implementation has been used in reality check of White (2000). One is Monte Carlo Reality Check, and another is Bootstrap Reality Check which usually used in practice. However, the Reality Check of White (2000) uses the loss differentials directly which lead to a loss of power of the test if there is a large amount of cross-sectional heteroskedasticity or variance term time-varying. Depending on the nature of the heteroskedasticity, significance tests can be too high or too low because it provides the variance with the over-estimation or under-estimation. To address this problem, Hansen (2005) use the Studentized Bootstrap Reality Check. The basic idea of bootstrapping is inference the population of interest from our sample data. A practical bootstrapping procedure, therefore, is as follows:

1. We transform each of loss different between benchmark and alternative models from equation (39) into $\delta_{j,k}$ where *j=1, 2, …, n* step ahead of forecasting and *k=1 ,2 , …, m* comparison models as followed,

$$\delta_{j,k} = \begin{bmatrix} \delta_{1,1}, & \delta_{1,2}, & \dots, & \delta_{1,k} \\ \delta_{2,1}, & \delta_{2,2}, & \dots, & \delta_{2,k} \\ \vdots & \vdots & \ddots & \vdots \\ \delta_{n,1}, & \delta_{n,2}, & \dots, & \delta_{n,k} \end{bmatrix}. \quad (42)$$

2. Compute T^{SPA} by choosing maximum each of average of origin sample observation ($\bar{\delta}^k$)

$$\delta_{j,k} = \begin{bmatrix} \delta_{1,1}, & \delta_{1,2}, & \dots, & \delta_{1,k} \\ \delta_{2,1}, & \delta_{2,2}, & \dots, & \delta_{2,k} \\ \vdots & \vdots & \ddots & \vdots \\ \delta_{n,1}, & \delta_{n,2}, & \dots, & \delta_{n,k} \end{bmatrix} \equiv \begin{bmatrix} \bar{\delta}^1 \\ \delta_{1,1} \\ \delta_{2,1} \\ \vdots \\ \delta_{n,1} \end{bmatrix}, \begin{bmatrix} \bar{\delta}^2 \\ \delta_{1,2} \\ \delta_{2,2} \\ \vdots \\ \delta_{n,2} \end{bmatrix}, \dots, \begin{bmatrix} \bar{\delta}^k \\ \delta_{1,k} \\ \delta_{2,k} \\ \vdots \\ \delta_{n,k} \end{bmatrix} \quad (43)$$

3. Divided maximum of k loss different by a long run variance of $\bar{\delta}^k$ ($\widehat{\omega}_k^2$),

$$\widehat{\omega}_k^2 = \hat{\gamma}_{k,0} + 2\sum_{j=1}^{n-1} k_i\, \hat{\gamma}_{k,j}. \tag{44}$$

where $\hat{\gamma}_{k,j}$ is the j^{th} sample autocovariance of sequence of $\bar{\delta}^k$,and

$$k_i = \frac{n-j}{n}(1-\frac{1}{w})^j + \frac{j}{n}(1-\frac{1}{w})^{n-j} \tag{45}$$

where w is the window length in stationary bootstrap which we followed Politis and White (2004) method to calculate the optimal length of window. Then the statistic test of SPA can rewrite as,

$$T^{SPA} = \max_{k=1,2,\dots,m} (\bar{\delta}^k / \sqrt{\frac{\widehat{\omega}_k^2}{n}}). \tag{46}$$

4. To compute the Studentized Reality Check p-value as the percentage of the bootstrapped maxima which are larger than the sample maximum,

$$p - value = B^{-1}\sum_{b=1}^{B} I[T_b^{*SPA} > T^{SPA}], \tag{47}$$

where $I=1$ when $T_b^{*SPA} > T^{SPA}$. To find T_b^{*SPA}, let b denote the number of bootstrap replications where $b=1, 2, \dots, B$. We randomly re-sample $t=1, 2, \dots, n$ observation of the vector of each loss differentials ($\bar{\delta}_k$) for $b=1, 2, \dots, B$ times to construct a bootstrap sample ($\bar{\delta}_k^{b*}$) by using stationary bootstrap which widely used in time series data and computed mean of each of bootstrap replications,

$$\left\{\begin{array}{c} \text{Own sample} \\ \bar{\delta}_1 = \frac{\sum_{t=1}^{n} \delta_{1,t}^{B*}}{n} \\ \hline \delta_{1,1} \\ \delta_{1,2} \\ \vdots \\ \delta_{1,n} \end{array}\right\} \equiv \left\{\begin{array}{c} b=1 \\ \bar{\delta}_1^{1*} = \frac{\sum_{t=1}^{n} \delta_{1,t}^{1*}}{n} \\ \hline \delta_{1,1}^{1*} \\ \delta_{1,2}^{1*} \\ \vdots \\ \delta_{1,n}^{1*} \end{array}\right\}, \left\{\begin{array}{c} b=2 \\ \bar{\delta}_1^{2*} = \frac{\sum_{t=1}^{n} \delta_{1,t}^{2*}}{n} \\ \hline \delta_{1,1}^{2*} \\ \delta_{1,2}^{2*} \\ \vdots \\ \delta_{1,n}^{2*} \end{array}\right\}, \dots, \left\{\begin{array}{c} b=B \\ \bar{\delta}_1^{B*} = \frac{\sum_{t=1}^{n} \delta_{1,t}^{B*}}{n} \\ \hline \delta_{1,1}^{B*} \\ \delta_{1,2}^{B*} \\ \vdots \\ \delta_{1,n}^{B*} \end{array}\right\}, \tag{48}$$

$$\left\{\begin{array}{c}\text{Own sample}\\ \bar{\delta}_2=\sum_{t=1}^{n}\delta_{2,t}\\ \hline \delta_{2,1}\\ \delta_{2,2}\\ \vdots\\ \delta_{2,n}\end{array}\right\}\equiv\left\{\begin{array}{c}b=1\\ \bar{\delta}_2^{1*}=\frac{\sum_{t=1}^{n}\delta_{2,t}^{1*}}{n}\\ \hline \delta_{2,1}^{1*}\\ \delta_{2,2}^{1*}\\ \vdots\\ \delta_{2,n}^{1*}\end{array}\right\},\left\{\begin{array}{c}b=2\\ \bar{\delta}_2^{2*}=\frac{\sum_{t=1}^{n}\delta_{2,t}^{2*}}{n}\\ \hline \delta_{1}^{2*}\\ \delta_{2}^{2*}\\ \vdots\\ \delta_{2,n}^{2*}\end{array}\right\},\dots,\left\{\begin{array}{c}b=B\\ \bar{\delta}_2^{B*}=\frac{\sum_{t=1}^{n}\delta_{2,t}^{B*}}{n}\\ \hline \delta_{2,1}^{B*}\\ \delta_{2,2}^{B*}\\ \vdots\\ \delta_{2,n}^{B*}\end{array}\right\},$$

$$\vdots$$

$$\left\{\begin{array}{c}\text{Own sample}\\ \bar{\delta}_k=\sum_{t=1}^{n}\delta_{k,t}\\ \hline \delta_{k,1}\\ \delta_{k,2}\\ \vdots\\ \delta_{k,n}\end{array}\right\}\equiv\left\{\begin{array}{c}b=1\\ \bar{\delta}_1^{1*}=\frac{\sum_{t=1}^{n}\delta_{k,t}^{1*}}{n}\\ \hline \delta_{k,1}^{1*}\\ \delta_{k,2}^{1*}\\ \vdots\\ \delta_{k,n}^{1*}\end{array}\right\},\left\{\begin{array}{c}b=2\\ \bar{\delta}_k^{2*}=\frac{\sum_{t=1}^{n}\delta_{k,t}^{2*}}{n}\\ \hline \delta_{k,1}^{2*}\\ \delta_{k,2}^{2*}\\ \vdots\\ \delta_{k,n}^{2*}\end{array}\right\},\dots,\left\{\begin{array}{c}b=B\\ \bar{\delta}_k^{B*}=\frac{\sum_{t=1}^{n}\delta_{k,t}^{B*}}{n}\\ \hline \delta_{k,1}^{B*}\\ \delta_{k,2}^{B*}\\ \vdots\\ \delta_{k,n}^{B*}\end{array}\right\}.$$

Compute the p-value of SPA test as the percentage of the bootstrapped maxima which are larger than the sample maximum,

$$T_b^{*SPA}=\max(\frac{n^{-1}\sum_{t=1}^{n}\overline{\delta}_{k,t}^{b*}-\overline{\delta}_k}{\sqrt{\widehat{\omega}_k^2/n}}), \tag{49}$$

However, the $p-value$ of the T_U^{SPA} is for upper bound which assume that all models could as good as the benchmark or close to benchmark. However, there are some of the models have a very low mean and a high standard deviation. We then assume that all loss function that is worse than benchmark model is unimportant for the asymptotic distribution test or lower bound (T_l^{SPA}) Another test is only re-centers if the loss-differential is positive (e.g. the benchmark is out-performed) or lower bound (T_c^{SPA}). We then add I_k^s into the equation (49),

$$T_b^{*SPA}=\max(\frac{n^{-1}\sum_{t=1}^{n}\overline{\delta}_{k,t}^{b*}-I_k^s\overline{\delta}_k}{\sqrt{\widehat{\omega}_k^2/n}}), \tag{50}$$

where $I_j^U=1$, $I_j^C=\bar{\delta}_j/\sqrt{\widehat{\omega}_j^2/P}>-\sqrt{2\ln\ln P}$ and $I_j^L=\bar{\delta}_j>0$ and the studentized Reality Check p-value,

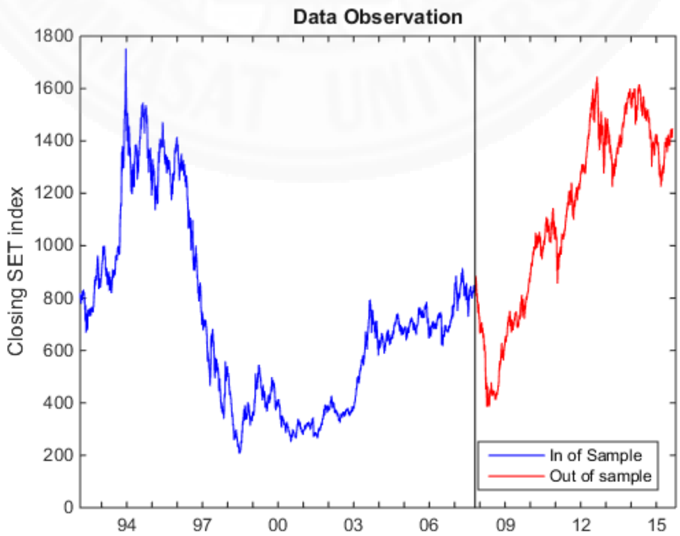
$$p-value=B^{-1}\sum_{b=1}^{B}I[T_{s,b}^{*SPA}>T^{SPA}]\text{ , s=u, c, l} \tag{51}$$

CHAPTER 4
RESULTS AND DISCUSSION

4.1 Data description

The observation data are daily high-low prices and open-close prices from Thailand Stock Index (SET) during the entire period from 17 February 1992 until 30 June 2016, for a total of 5,967 daily observations. We received the data from the database of DataStream. This study period divided data into two sub-periods. One is an estimation period (in-of-sample) in which the model parameters fitted, another sub-periods used to evaluate the forecasting performance of the models (out-of-sample). The reason for using out of sample evaluating is to ensure optimization methods from estimation period can give the best possible results if we use other different sets of data. If the entire set of data consists of T observation $X_1, X_2, \dots , X_T$. The estimation period is $X_1, X_2, \dots, X_n$ and $X_{n+1}, X_{n+2}, \dots , X_T$ for forecasting period. Following Tsay(2008), he demonstrates that $n = \frac{2T}{3}$ is the reasonable choice of n in divide total period into two sub-periods, which gives the root mean square error (RMSE) superior using $n = \frac{T}{2}$.

Figure 4.1: Daily closing SET index



Note: We separate our entire data set into in-of-sample (blue line) and out-of-sample (red line) based on the study of Tsay (2008) which cutting point is equal to n = 2T/3

Accordingly, the estimation data starts from 17 February 1992 to 7 May 2008 consisted of 3,978 observations and observation of forecasting period obtained 1,989 observations which starts from 8 May 2008 – 30 June 2016. The optimal cutting point of in-of-sample period and out-of-sample period will be explained later in this section.

In Figure 4.1, we plotted the entire period of a daily closing SET price. The vertical line (black) separate period of data set into two sub periods. The blue line represents the index level for the estimation period (in-of-sample) and forecasting period (out-of-sample) is represented by the red line.

In the real world, there are unpredictable shocks or unexpected changes that produce a significant impact on the whole system, especially in financial markets. The effect of shocks could die out gradually over the time if the systems are stationary or exist permanent if the systems is non-stationary. However, financial data sets are often non-stationary which mean, variance, and covariance change over the period. Non-stationary behaviors can be trends, cycles, and random walks, which cannot be modeled or forecasted. The results obtained by using non-stationary time series may be spurious that it may indicate a relationship between two variables where one does not exist. To test the stationarity of financial time series, there are many different unit-root tests and stationary test.

The Dickey–Fuller (DF) test is the simplest and popular model to test unit root, $\Delta y_t = \rho y_{t-1} + u_t$, where y_t is the variable of interest at time *t*, p is a coefficient and u_t is the error term that assume IID $(0,\sigma^2)$. The null hypothesis that a time series is integrated of order 1 or unit root ($p = 1$). The alternative hypothesis depends on which version of the test, stationary with drift term ($\Delta y_t = \propto + \rho y_{t-1} + u_t,$ where $p < 1$) and stationary with drift term and time trend ($\Delta y_t = \propto + \rho y_{t-1} + \gamma t + u_t,$ where $p < 1$). However, Augmented Dickey–Fuller (ADF) expands the DF regressions by adding lagged difference terms ($\sum \delta_i \, \Delta y_{t-i}$) if the error term ($u_t$) is serially correlated. While the Phillips–Perron (PP) estimates the equation by ordinary least squares (OLS) instead of incorporating lagged difference terms. The critical values for these tests are identical to those for the ADF equivalents. The ADF and PP unit root tests are for the null hypothesis that a time series is unit root. On the other hand, Kwiatkowski–Phillips–

Schmidt–Shin (KPSS) is stationary test which the null hypothesis that an observable time series is stationary around a deterministic trend.

In Figure 4.1, we perform unit root test included constant and trend term of SET index closing price (Panel A). Our finding shows that both of the PP and ADF tests fail to reject the null hypothesis of unit-root at 95.00% confidence level. On the other hand, KPSS rejects the null hypothesis of trend stationary. From all test, we conclude that daily SET index closing price is non-stationary.

However, one way to make a time series stationary is to compute the differences between observations. This is known as differencing. We then transform SET index closing price in term of logarithms return that can help to stabilize the variance of a time series. The results in Panel B also confirmed that the return of SET index is stationary. Therefore, we use SET return to model and forecast in the next section.

Table 4.1: Stationary test and Unit root test with constant term and time trend of SET index closing price and return

PP test	ADF test	KPSS test
Panel A: SET index closing price		
-1.0857	-1.0857	106.3702
(0.3181)	(0.3181)	(728.5627)
Panel B: SET index return		
-70.4131	-70.4131	0.1029
(20.6291)	(20.6291)	(0.7049)

Note: This table shows the test-statistic value from stationary test and unit root test with constant and trend term of SET index and SET return with standard deviation in parentheses from 17 February 1992 to 30 June 2016. The null hypothesis of Phillip-Perron (PP) and Augmented Dickey-Fuller (ADF) tests are that the variable contains a unit root. By the way, Kwiatkowski–Phillips–Schmidt–Shin (KPSS), the null hypothesis is trend stationary.

*denote significant at the 5% level.

We also present the basic statistical analysis of daily return of SET index in Table4.2. From panel A, the sample size, mean, max, min, standard deviation, skewness, and kurtosis is presented. The return of SET index is widely distribution from

-16.0633 to 11.3495. The average SET returns are slightly positive and are very small compared with standard deviation. Furthermore, the skewness defined as the standardized third central moment positive which indicated that the return series is skewed to the right side or right-tailed. Additionally, the probability distribution with kurtosis greater than three is said to be leptokurtic or fatter tail. Then, it necessary to correctly determine the distribution of error term.

Table 4.2: Preliminary analysis of SET return

Panel A: Summary statistics						
Mean	**Max**	**Min**	**SD**	**Skewness**	**Kurtosis**	**J-B**
0.0094	11.3495	-16.0633	1.5771	0.0178	10.1260	12,625.55
						(2110.343)
Panel B: ARCH test						
Ljung-Box		**Eagle's**				
500.5850		206.0492				
(130.3946)		(53.6377)				

Note: This table shows the characteristic test-statistic value for return of Stock Exchange of Thailand index with standard deviation in parentheses from 17 February 1992 to 30 June 2016. J-B represents the statistic of Jarque and Bera's (1987) normal distribution test. To check the serial correlation of residual terms, we use Ljung-Box test. Finally, the test statistic of Engle and Ng (1993) exhibit Autoregressive Conditional Heteroskedasticity effects in the residuals term.

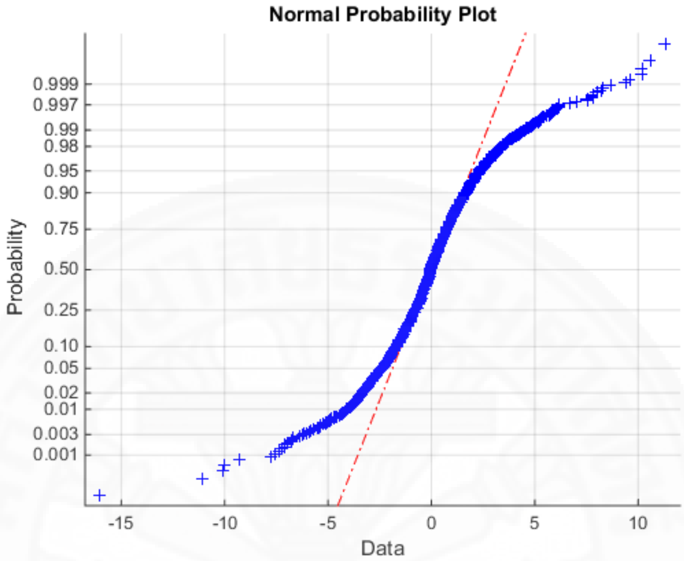
*denote significant at the 5% level.

We also use Jarque and Bera (1987) to test of whether daily returns have the skewness and kurtosis matching a normal distribution. The JB test is computed from skewness and kurtosis which computed as

$$JB = \frac{N}{6}\left(\widehat{skew}^2 + \frac{(\widehat{kurt}-3)^2}{4}\right), \quad (50)$$

where $\widehat{skew}$ denotes the sample skewness, and $\widehat{kurt}$ denotes the sample kurtosis. For large sample size (>2000), the test statistic has a chi-square distribution with two degrees of freedom. The *p-value* of JB test of SET daily returns in Table 4.1 (Panel A) rejects the null hypothesis of normality. The figure 4.2 also demonstrates returns (plus sign line) are significantly different from a normal population (dash-dot line). Our data clearly shows that it is reasonable to assume $z_t \sim t(v)$ than to assume $z_t \sim N(0,1)$.

Figure 4.2: Normal probability plot of SET return

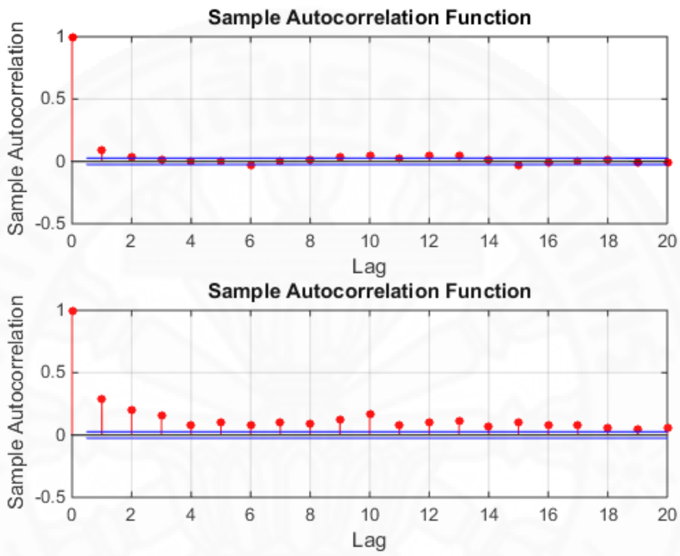


Note: This graph shows the residual of SET index is non-normal distribution which the plus sign line (return of SET index) do not array in straight line compare with dash-dot line (normal data)

From Panel B, we use regular hypothesis checks for correlation and ARCH effect, like Ljung-Box Q test, and Engle's ARCH test. First, the LB-Q function performs whether a series of residual over time are independent or not. If there are autocorrelation, it can decrease the accuracy of a time-based predictive model. Second, Eagle's ARCH test is the most popular which also tests the presence of significant conditional heteroskedasticity in financial time series. It tests the null hypothesis that a time series of sample residuals consists of independent identically distributed (I.I.D.). The test statistic of all methodologies rejects the null hypothesis of no ARCH effects exist at 95.00% of significant level.

Besides, in Figure 4.3 shows the Sample Autocorrelation Function and Sample Partial Autocorrelation Function for the daily return between lags 0 to 20. The Sample Autocorrelation function and Sample Partial Autocorrelation Function are commonly-used tools for checking the randomness in a data set. This randomness is verified by computing the autocorrelations for data values at varying time lags. If random, such autocorrelations should be near zero for any and all time-lag separations. If non-random, then one or more of the autocorrelations will be significantly non-zero. As we can see in Figure 4.3, the sample ACF illustrate no relationship in SET return. However, ACF of the squared return shows clearly significant autocorrelation. On the other word, SET return is serially uncorrelated but not independent

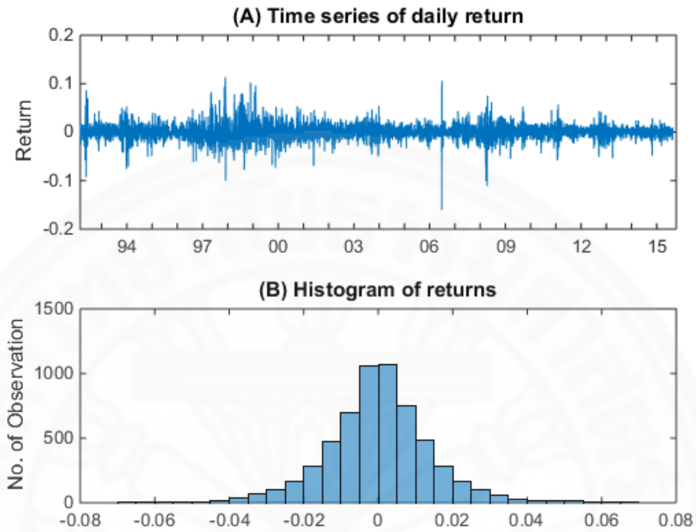
Figure 4.3: Sample autocorrelation function



Note: The sample ACF exhibits significant autocorrelation at lag 1. The sharp cutoff of the ACF combined with the more gradual decay suggests that a GARCH-*type* model at one lag might be appropriate for this data.

Let us look at some stylized facts about the behavior of financial variables, the logarithmic return of closing price graph. Figure 4.4 shows the time series of the daily returns (Panel A.) which change rapidly over the entire period. We can say that the time series are volatile. Furthermore, there is some volatility clustering, which indicates that there is an ARCH effect in observation period. Normally, the one of key characteristic of introduction of financial times series usually known as volatility clustering, in which large changes in price are often followed by larger changes, and small changes are more likely to follow small changes as shown as in daily return. We notice groups of volatilities from 1997 to 2000 and 2007-2009 which is an Asian crisis and Global Financial crisis. Panel (B) shows the histogram of the return which displays non-normal properties. Note that there are more observations around the mean and in the tails. Distribution of these properties is said to be “Leptokurtic.” Besides, probability distributions of financial data typically observed the fat-tail effect (or excess kurtosis.)

Figure 4.4: Daily logarithm return and histogram

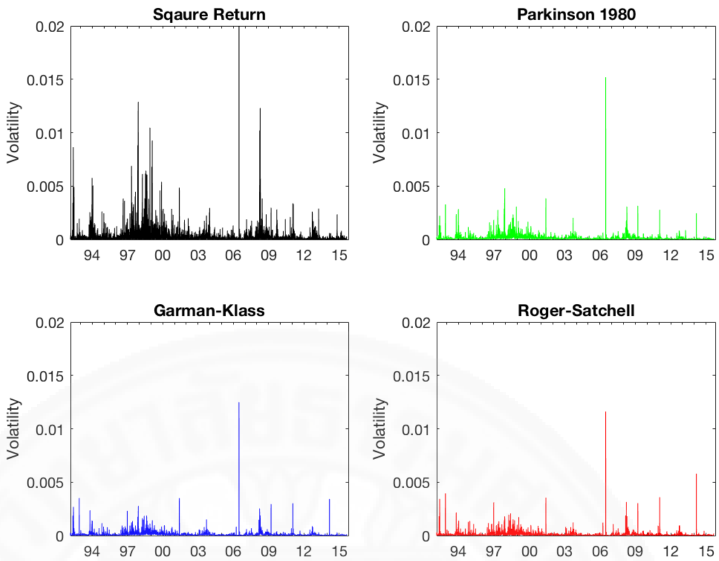


Note: In panel (a) shows SET index in term of logarithm from the beginning to end of data. We obviously noticed that there are some of the group of volatility clustering around the year 1997-2000 which is the Asian crisis and the period of Subprime crisis in 2007-2009. Moreover, the “leptokurtic” be notices in panel (b)

4.2 Volatility Proxies

As mentioned in the introduction, it is a well-known fact that volatility of asset returns is time-varying and unobservable. Volatility also has some commonly seen characteristics that volatility exhibits persistence, mean reverting and innovations have an asymmetric impact on volatility. In this paper, the focus will be examining the non-parametric models as the valid proxies of volatility; Square Return and Range-based Volatility. All the volatility proxies compared with the forecasting performance of commonly used GARCH-*type* models (Parametric model) under Gaussian and Student’s t-distribution to find the actual volatility.

Figure 4.5: Volatility Proxies



Note: This figure compares the level of volatility between square return and range-based estimators which are Parkinson (1980), Garman and Klass (1980) and Rogers and Satchell (1991). We rescale the y-axis to make this chart comparable

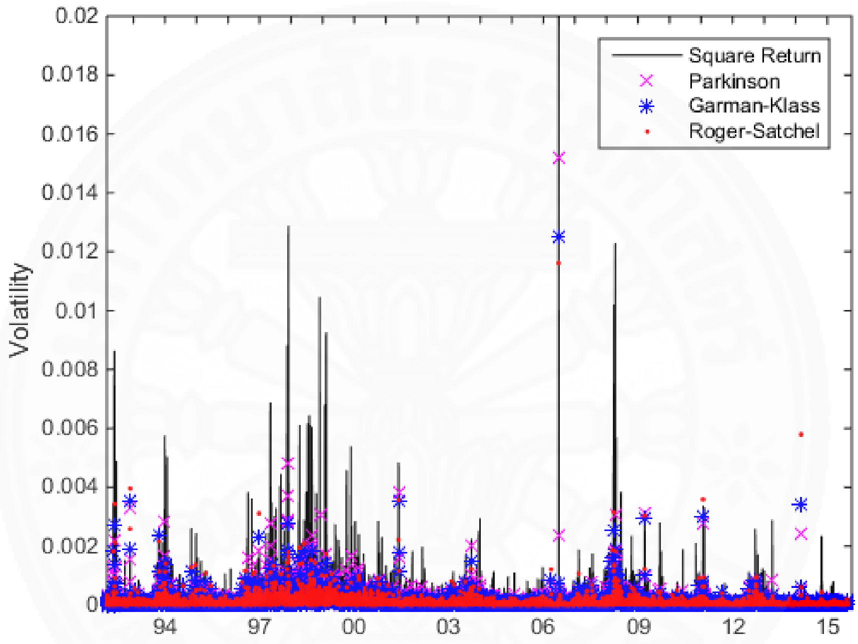
Square return is commonly used to determine instability of data set because it is simplest and easy to obtain closing price. However, when asset prices fluctuate during the day, to capture only closing price, we cannot capture the correct volatility. For example, if the closing price of the day equal to the previous close as price significantly swing over the day. It reflects low volatility instead of high volatility while the log range captures the price movement throughout the day, highest and lowest price, which seems to imply correct volatility and can be used for volatility assessment. Moreover, in the real world, opening, high, low and closing data are readily available. The previous study indicated the different in their efficiency see Parkinson (1980), Garman and Klass (1980) and Rogers and Satchell (1991).

Garman and Klass (1980) who defined the efficiency of a variance estimator to be the ratio[1] of the variance of the classical daily volatility estimator (Squared return),

[1] *Eff* = $\left(\frac{Var(\sigma^2_{benchmark})}{Var(\sigma^2_{Range-based})}\right)$, where *Eff* is variance efficiency ratio,
$Var(\sigma^2_{benchmark})$ is variance of the benchmark variance estimator (Square Return), and
$Var(\sigma^2_{Range-based})$ is variance of the range-based estimator (PK, GK and RS).

claimed that the higher of ratio give more efficient estimator. The most efficient would take us closer to the unbiased estimators of true volatility. We also computed the efficient of volatility proxies followed Garman and Klass (1980). The Parkinson (1980) has efficiency 5.82, Garman and Klass (1980) has the highest efficient 8.72 compared with three range-based and Rogers, Satchell has efficiency 8.18.

Figure 4.6: Comparing the volatility proxies



Note: This graph compares the level of volatility between square return and range-based estimators. We rescale the y-axis to make this chart comparable. The squared return is volatile than three range-based estimators means over-under prediction or unbiased estimators of true volatility.

We mutually comparing proxies of volatility is presented in Figure 4.5 and Figure 4.6. When comparing range-based measures of volatility (Figure 4.5), we rescale of y-axis to compared the level of volatility. We observe squared return estimators sometimes report a high volatility or perform poorly relative to range-based estimators. Similarly, Andersen and Bollerslev (1998) concluded that squared return is a very noisy volatility estimator because it computed by only closing price. By the way, the similarity of Range-based variance construction by formed from the entire price process using open, high, low, and close price. It is not surprising to find nearly identical level of volatility. Most notably, Parkinson's estimators sometimes report a high volatility compared with other range-based. However, Roger and Satchell has more efficiency than another range-based estimator.

Next, we firstly investigate the in-of-sample coefficient of GARCH with different error distribution. Secondly, we forecast j^{th} periods ahead out-of-sample of conditional variance using fixed and rolling windows approach. Thirdly, we evaluate the performance of volatility forecast with various volatility proxies in terms of loss function and SPA test. Last, we empirically improve the performance of GARCH-*type* models under normal and Student's t-distribution by adding various volatility proxies as exogenous variable.

4.3 The estimation and forecasting of GARCH family models

In this section, we computed the parametric estimators or GARCH family using 5,967 daily observations of SET index from 17 February 1992 to 30 June 2016. However, this study divided the observations into two sub-periods as we mentioned in data description part. For in-of-sample period, we estimated the parameters of GARCH-*type* models by using the data from 17 February 1992 to 7 May 2008 consisted of 3,978 observations. Another period obtained 1,989 observations which starts from 8 May 2008 to 30 June 2016 which used to test the out-of-sample forecasting performance.

4.3.1 In-of-sample volatility estimation

In Table 4.3 shows the estimation of different GARCH-*type* models under normal and Student's t-distribution from 17 February 1992 to 30 June 2016 (3,978 observations) of SET index return. The first row of each estimation shows the coefficient and the second row in the parenthesis form is robust standard error. Panel A presents GARCH-*type* model with normal (Gaussian) distribution which often used in general. To capture the stylized fact of SET return, we assume GARCH-*type* models with Student's t-distribution that presented in Panel B.

Based on the assumption of 5% significant, the parameters of all GARCH-*type* models are highly significant. The statistical significance of α, β and γ indicate that past innovations, past conditional variance, and news about volatility from the previous periods have an impact on the conditional variance.

The estimates of constant term (ω) are all positive and extremely smaller than the sample variances except EGARCH model due to the changing conditional variances over time.

The parameters α and β are significantly positive for all models. Recall that α is the coefficient in front of the innovation term at time *t-1*. Higher value of α means that shock yesterday more impacts the volatility of today. For the coefficient β is in front of the conditional variance at time *t-1*, so a high beta means that the yesterday's volatility dramatically impacts today's volatility.

Table 4.3: The parameters of GARCH-*type* with normal distribution and GARCH-*type* with Student's t-distribution

Model	**Constant (ω)**	**ARCH (α)**	**GARCH (β)**	**Leverage (γ)**
Panel A: GARCH-*type* with normal distribution				
ARCH	2.10E-04	0.2637	-	-
	(1.33E-05)	(0.0357)	-	-
GARCH	1.53E-05	0.1319	0.8157	-
	(9.09E-06)	(0.0194)	(0.0408)	-
EGARCH	-0.5227	0.2624	0.9357	-0.0484
	(0.2746)	(0.0325)	(0.0340)	(0.0198)
GJR-GARCH	1.53E-05	0.1002	0.8089	0.0831
	(8.77E-06)	(0.0168)	(0.0406)	(0.0278)
Panel B: GARCH-*type* with Student's t-distribution				
ARCH	2.03E-04	0.3240	-	-
	(1.08E-05)	(0.0369)	-	-
GARCH	7.17E-06	0.1261	0.8528	-
	(1.71E-06)	(0.0164)	(0.0185)	-
EGARCH	-0.2793	0.2269	0.9657	-0.0379
	(0.0623)	(0.0232)	(0.0075)	(0.0116)
GJR-GARCH	7.99E-06	0.0953	0.8427	0.0792
	(1.99E-06)	(0.0142)	(0.0201)	(0.0207)

Note: In panel A shows the parameter estimates of ARCH, GARCH, EGARCH, and GJR-GARCH with normal distribution and Student's t-distribution in panel B. The robust standard error depicts in parentheses form. The models are estimated on SET index return from 17 February 1992 to 30 June 2016 consisted of 3,978 observations (in-of-sample period)

Also, the coefficients β is significantly positive, and γ is significantly negative for EGARCH model. For $\beta/\gamma < 0$, then negative innovations have higher impact on volatility than positive innovations or asymmetric effect in SET index. If $0 < \beta/\gamma < 1$, then the positive innovations increase the volatility, but negative innovation decrease volatility. Moreover, the coefficients of GJR-GARCH is significantly positive indicated that the negative innovations have an impact on volatility.

Our results also suggest that the persistence in volatility in GARCH and GJR-GARCH models, as measured by the sum of two estimated ARCH and GARCH coefficients $\alpha + \beta$ (persistence coefficients), is less than one which is required to have a mean reverting variance process. In contrast, the sum of parameters of EGARCH model is larger than one, suggesting that shocks to the conditional variance are highly persistent, i.e. the conditional variance process is exploded. This implies that the tendency of large changes in returns of financial assets to cluster together. Therefore, it confirms that volatility clustering is observed in SET index returns series.

Further general evaluation method for comparison of each fitting models from in-of-sample periods are presented in Table 4.4. In this paper, we use two of the most well-known and popular criteria for model selection, Akaike information criterion (AIC) and Bayesian information criterion (BIC).

Table 4.4: Diagnostics from GARCH-*type* with normal distribution and GARCH-*type* with Student's t-distribution

Model	ARCH	GARCH	EGARCH	GJR-GARCH
Panel A: GARCH-*type* with normal distribution				
AIC	-21,525.16	-22,019.35	-22,023.14	-22,044.75
BIC	-21,518.88	-22,006.78	-22,004.28	-22,025.89
Log L	10,763.58	11,011.68	11,014.57	11,025.38
Panel B: GARCH-*type* with Student's t-distribution				
AIC	-22,044.97	-22,393.32	-22,410.03	-22,412.04
BIC	-22,032.40	-22,374.46	-22,384.87	-22,386.89
Log L	11,024.49	11,199.66	11,209.01	11,210.02

Note: AIC (Akaike Information Criterion) and BIC (Bayesian Information Criterion) considers goodness-of-fit and parsimony. The formula of AIC is -2(*Log L*) + 2*k* and BIC is -2(*Log L*) + ln*(n)k* where *Log L*, *k* and *n* denoted log likelihood value, number of estimated parameters and sample size, respectively. To select models, we choose the models give minimize AIC and BIC value.

We compare the maximum value of likelihood functions if the number of parameters of each model is the same. However, if the models have a different number of parameters. The Akaike information criterion (AIC) adjusts the likelihood function to account for the number of parameters. The AIC is given by -2(*Log L*) + 2*k* where *k* denoted the parameter. However, AIC gives a penalty of 2 for an extra parameter (2*k*) which leads us to choose models with too many parameters. Another model selection is Bayesian information criterion (BIC) which gives a penalty of *ln(n)k* for an additional parameter. The penalty term of BIC is larger than in AIC when adding more parameters. The BIC is given by -2(*Log L*) + *ln(n)k* where *k* denoted the parameter and *n* denoted the parameter sample size. Based on model selection criterion, the GJR-GARCH model with normal and Student's t-distribution is better fit than other models for both AIC and BIC which give the lowest values.

4.3.2 Out-of-sample volatility forecasting

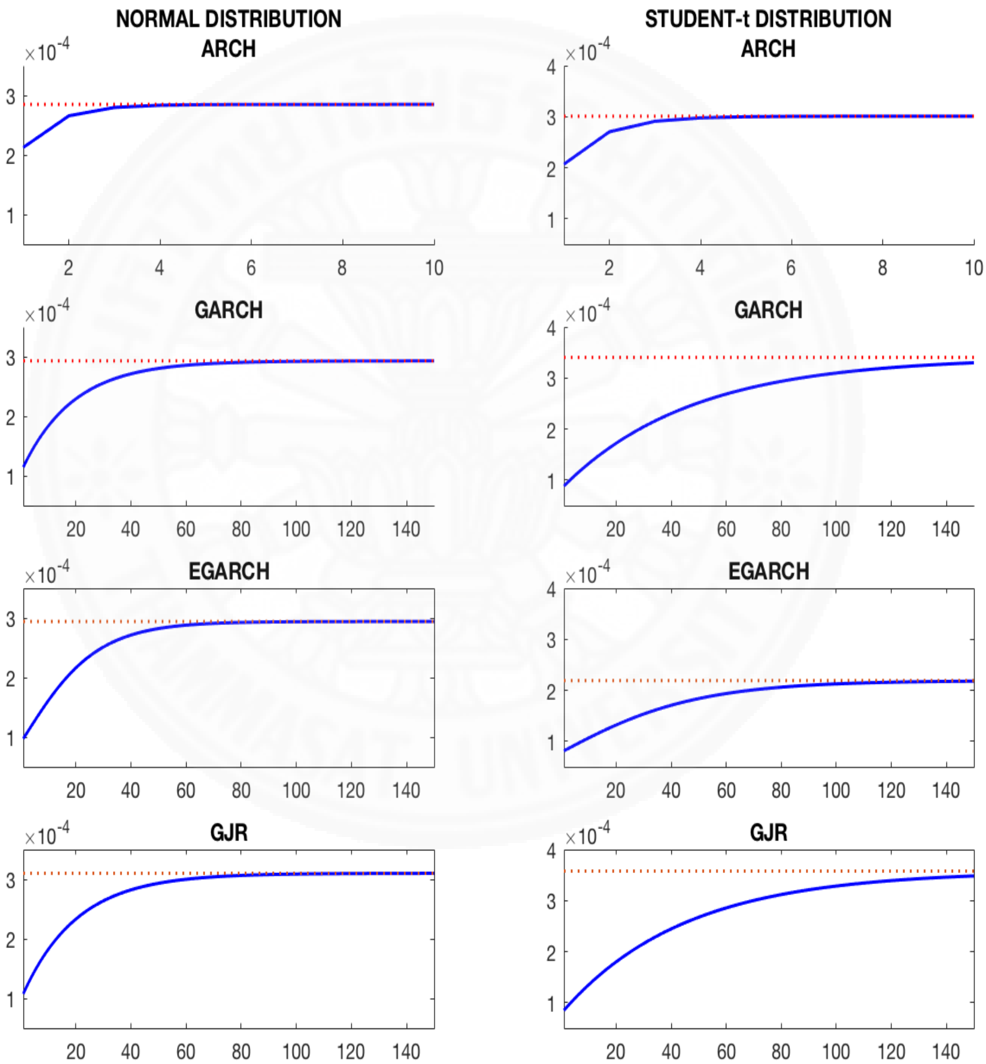
When the four GARCH models have been fully specified with their respective order and a probability distribution for the error term. We estimate the in-of-sample coefficient, and the out-of-sample evaluation is the next step in the study performed in this paper. Since we do not know the optimal sampling frequency, we then forecast using various popular used sampling frequencies: monthly, quarterly, half-yearly and yearly periods ahead forecast of the conditional variance (j^{th} = 22, 55, 110, and 220).

Figure 4.7, We plot out-of-sample forecasts of conditional variance using fitted four GARCH models with both error distributions from in-of-sample period (3,978 observations). The rescaled horizontal line is the number of estimation periods of volatility and the vertical line presents the estimated conditional and unconditional variance from the fitted models. The red dot horizontal lines exhibit in figure 4.7 is the unconditional volatility level or long-term average volatility which the prior information does not give any information to forecast the volatility at infinite horizon. We measured the unconditional volatility[2] of ARCH, GARCH, EGARCH and GJR-GARCH model is calculated as 0.000284, 0.000292, 0.000294 and 0.00031,

[2] The unconditional variance of ARCH, GARCH, EGARCH and GJR-GARCH is computed as $\sigma^2_{ARCH} = \frac{\omega}{1-\alpha}$, $\sigma^2_{GARCH} = \frac{\omega}{1-\alpha-\beta}$, $\sigma^2_{EGARCH} = EXP(\frac{\omega}{1-\beta})$, and $\sigma^2_{GJR} = \frac{\omega}{1-\alpha-\beta-\frac{\gamma}{2}}$, respectively

respectively. For Student's-t error distribution, long-term average volatility of ARCH, GARCH, EGARCH and GJR-GARCH model is 0.0003, 0.000339, 0.000290 and 0.000356, respectively. The unconditional variance of four GARCH models under normal and Student's t-distribution is quite the same except EGARCH model.

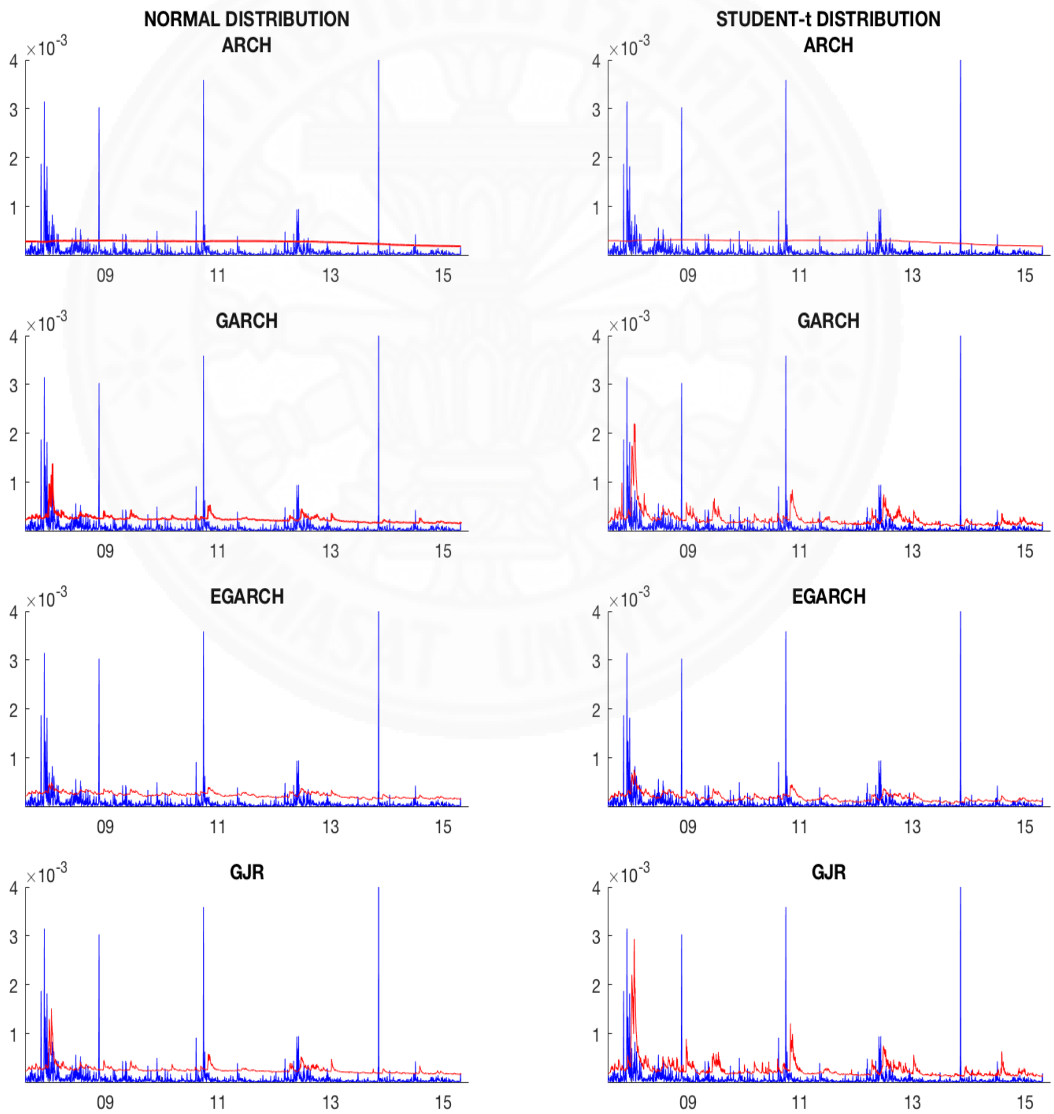
Figure 4.7: Out-of-sample conditional variance forecasts from in-sample estimation period compared with unconditional variance



Note: The blue line presents the conditional variance forecast which converges asymptotically to the unconditional innovation variance (the red dot).

From the beginning of the forecast period, the conditional variance forecast is less than the unconditional variance for all models. The forecasts revert upward toward the unconditional volatility level. For ARCH model, the speed of volatility rapidly reverts to the unconditional level about three days. However, the conditional variance of GARCH, EGARCH and GJR models under normal distribution are increasing and mean-reverting to the long-term average about quarter. However, for four GARCH under Student's t-distribution, the conditional forecast obviously take longer period back to the unconditional variance level.

Figure 4.8: 22 periods ahead forecast of EGARCH-*type* model using rolling out-of-sample compared with RS estimators



Note: This figure shows the different between 22 periods ahead rolling out-of-sample forecast of EGARCH model under normal and Student’s-t distribution (red line). The blue line is RS volatility proxy

The estimation and forecasting using rolling window approach provide the simplest way to incorporate actual data into the estimation of time-varying volatilities. We set the number of consecutive observation per rolling window equal to in-of-sample observations (*m*=3,978 observations). The first rolling window subsample contains 1 through 3,978 observations. For next rolling window subsample, we include the new observation and remove the first one. Then, the second rolling window contains observations for period 2 through *m + 1*, and so on for 1,989 rolling window subsample

After that, we re-estimate the parameters of each model with one more observation added to rolling window sub period. That give estimated 1,989 set of coefficients for each model. These values are used to generate j^{th} periods ahead forecast of the conditional volatility for 1,989 periods. We keep out-of-sample forecasting using rolling window method compared with volatility proxies (SR, PK, GK and RS) in the same period to measure the predictive performance in next section.

Figure 4.8 illustrates the example of the 22 periods ahead forecast of EGARCH model under normal and Student's t-distribution using 1,989 rolling out-of-sample compared with RS estimator. We noticed that 22 periods ahead forecast of ARCH is similar to the unconditional variance presented in Figure 4.7 which has fast mean-reverting. Figure 4.8 also shows the condition variance of GARCH and GJR under Student's t-distribution is more volatile than EGARCH because it takes shorter speed of mean reversion back to long-term variance than GARCH and GJR models

However, we cannot conclude which models have better performance against alternative volatility proxies from plotting the graph. To measure the discrepancy between proxy of actual volatility and different forecasting models. We use loss function and superior predictive ability test which explain the results in section 4.4 and section 4.5

4.4 Evaluating the forecasting performance of GARCH family models

To measure the forecasting performance of the different models, it is necessary to specify an actual volatility. Of course, the volatility is unobservable and an ex-ante proxy must be utilized. We choose a standard squared return and three type of Range-based estimators as volatility proxies which are Parkinson (1980), Garman and Klass (1980) and Rogers and Satchell (1991) estimators. With the volatility proxies ($\hat{\sigma}_{t+j}^{2}$)

and j^{th} periods ahead volatility forecasts ($\hat{h}^2_{t+j|t}$), computed for each of the days in the out-of-sample period using fixed window or rolling window. A well-known way to evaluate the out-of-sample fit is through loss function. The Root Mean Squared Error[3] (RMSE) is widely common metrics used and easy tools to measure accuracy for continuous variables which computed the square root of the average of the squared deviations between volatility proxies and predicted conditional volatility. The criterion for selecting the "Best" model is which model produces the smallest values of RMSE

In Table 4.5 and 4.6 contain the loss function value (RMSE) between alternative volatility proxies in the rows and out-of-sample forecasts of various GARCH-*type* models under normal distribution in four columns on the left and GARCH-*type* models under student's t-distribution in four columns on the right. The 22, 55, 110 and 220 periods out-of-sample forecast presented in Panel A, B, C, and D.

Table 4.5 represents RMSE values of different periods ahead out-of-sample forecast using fixed window in-of-sample estimation periods start from 17 February 1992 to 7 May 2008 (3,978 observation). After the results in Table 4.4 which GJR that is assuming both normal and Student's t-distribution provides the best in-of-sample fit which gives the smallest AIC and BIC. However, it is quite surprising that RMSE in Table 4.5, For normally distributed error terms, EGARCH model gives a better performance among others when using PK, GK, and RS estimators as volatility proxy for all periods forecast. For Student's t distribution of error term, the result is quite mixed. EGARCH model still give lowest RMSE using range-based volatility proxies. However, GJR is better model if we used SR as volatility proxy.

Moreover, we noticed RMSE of short period forecasting is less than long period forecasting ($RMSE_{22}$ < $RMSE_{55}$ < $RMSE_{110}$ < $RMSE_{220}$). If the forecast period is too long, there are more factors can affect the forecast. It leads long-term forecast is less accurate than a short-term forecast. Furthermore, using RS estimator as a volatility proxy compared with normal and Student's t-distribution of different GARCH models give us the lowest RMSE value according to the efficiency of range-based estimators which RS provides the best efficiency.

---

[3] $RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(\hat{\sigma}_t^2 - \hat{h}_{t,k}^2\right)^2}$

For evaluation of the performance of out-of-sample using rolling window presented in Table 4.6, we also use loss function to assess the best models. It confirmed that EGARCH model with Student's t-distribution significantly provides the lowest RMSE values for all period forecasts. Moreover, for normal distribution of error term, EGARCH model can beat other models when forecasting for a short-term period. However, ARCH better performance for longer period forecast.

Table 4.5: The RMSE of alternative volatility proxies and different *j* periods forecast of GARCH-*type* with normal distribution and GARCH-*type* with Student's t-Distribution models from 3,978 fixed windows in-of-sample estimation period (17 February 1992 to 7 May 2008)

Proxy	**GARCH-*type* with normal distribution**				**GARCH-*type* with Student's t-Distribution**			
	ARCH	**GARCH**	**EGARCH**	**GJR**	**ARCH**	**GARCH**	**EGARCH**	**GJR**
Panel A: 22 periods ahead forecast (22 observations)								
SR	2.6903	2.2769	**2.2451**	2.2732	2.7645	2.2350	2.2639	**2.2315**
PK	2.1205	1.2237	**1.0614**	1.2231	2.2589	0.8016	**0.6157**	0.8311
GK	2.1244	1.2165	**1.0466**	1.2166	2.2655	0.7657	**0.5484**	0.7992
RS	2.1172	1.2259	**1.0619**	1.2254	2.2567	0.8025	**0.6182**	0.8322
Panel B: 55 periods ahead forecast (55 observations)								
SR	3.4043	3.2717	**3.2606**	3.2670	3.4313	3.2797	3.3821	**3.2654**
PK	2.0375	1.6225	**1.5669**	1.6690	2.1635	1.3322	**1.1443**	1.3915
GK	2.0295	1.5834	**1.5230**	1.6464	2.1748	1.1992	**0.8274**	1.2910
RS	2.1218	1.6946	**1.6386**	1.7601	2.2684	1.3121	**0.9197**	1.4068
Panel C: 110 periods ahead forecast (110 observation)								
SR	12.0964	12.0650	12.0613	**12.0375**	12.0746	12.0269	12.2187	**11.9953**
PK	3.3359	3.2163	**3.2045**	3.2344	3.3822	**3.1111**	3.1203	3.1341
GK	3.5270	3.4090	**3.3974**	3.4338	3.5803	3.2969	**3.2595**	3.3284
RS	3.9843	3.8790	**3.8690**	3.9024	4.0337	3.7757	**3.7339**	3.8054
Panel D: 220 periods ahead forecast (220 observations)								
SR	13.5947	13.5721	13.5690	**13.5421**	13.5677	13.5289	13.7335	**13.4976**
PK	3.6885	3.6358	**3.6305**	3.6481	3.7113	**3.6237**	3.6272	3.6467
GK	3.2082	3.1517	**3.1465**	3.1848	3.2533	3.1594	**3.0428**	3.2068
RS	3.4783	3.4282	**3.4238**	3.4648	3.5261	3.4425	**3.2955**	3.4930

Note: The RMSE value which presents in this table multiply by 10^{-4}. The bold letter provides the smallest RMSE values. The formula of $RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(\hat{\sigma}_t^2 - \hat{h}_{t,k}^2\right)^2}$, where $\hat{\sigma}_t^2$ denoted alternative volatility proxies in the row and $\hat{h}_{t,k}^2$ denoted estimated condition variance of different GARCH models

Table 4.6: The RMSE of alternative volatility proxies and different *j* periods forecast of GARCH-*type* with normal distribution and GARCH-*type* with Student's t-Distribution models from 1,989 Rolling windows out-of-sample forecast period (8 May 2008 – 30 June 2016)

Proxy	GARCH-*type* with normal distribution				GARCH-*type* with Student's t-Distribution			
	ARCH	**GARCH**	**EGARCH**	**GJR**	**ARCH**	**GARCH**	**EGARCH**	**GJR**
Panel A: 22 periods ahead forecast (1,967 observations)								
SR	5.4475	5.3906	**5.3616**	5.4241	5.4716	5.5140	**5.3007**	5.6017
PK	2.6118	2.4738	**2.4081**	2.5881	2.7055	2.7994	**2.0976**	3.0042
GK	2.6137	2.4833	**2.4131**	2.6038	2.7114	2.8276	**2.0932**	3.0374
RS	2.9119	2.7990	**2.7361**	2.9076	3.0007	3.1164	**2.4586**	3.3084
Panel B: 55 periods ahead forecast (1,934 observations)								
SR	5.4757	**5.4665**	5.4762	5.5049	5.4991	5.5414	**5.3671**	5.7060
PK	**2.6339**	2.6351	2.6546	2.7672	2.7280	3.0284	**2.1667**	3.3356
GK	**2.6358**	2.6399	2.6583	2.7759	2.7337	3.0567	**2.1493**	3.3628
RS	**2.9344**	2.9399	2.9553	3.0627	3.0233	3.3290	**2.5063**	3.6114
Panel C: 110 periods ahead forecast (1,879 observations)								
SR	**4.7989**	4.8165	4.8219	4.8856	4.8309	5.2891	**4.6841**	5.6203
PK	**2.6103**	2.6618	2.6763	2.8346	2.7104	3.5956	**2.1741**	4.1139
GK	**2.5939**	2.6476	2.6626	2.8262	2.6988	3.6005	**2.1327**	4.1251
RS	**2.8672**	2.9163	2.9297	3.0805	2.9633	3.8079	**2.4539**	4.3080
Panel D: 220 periods ahead forecast (1,769 observations)								
SR	**3.2316**	3.2705	3.2788	3.4239	3.3001	5.0115	**2.9882**	6.0066
PK	**2.5186**	2.5848	2.5987	2.8152	2.6389	4.7629	**2.0111**	5.8598
GK	**2.6010**	2.6661	2.6797	2.8931	2.7199	4.8214	**2.1006**	5.9147
RS	**2.8921**	2.9510	2.9630	3.1579	2.9996	4.9892	**2.4516**	6.0547

Note: The RMSE value which present in this table multiply by 10^{-4}. The bold letter provides the smallest RMSE values. The formula of $RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(\hat{\sigma}_t^2 - \hat{h}_{t,k}^2\right)^2}$, where $\hat{\sigma}_t^2$ denoted alternative volatility proxies in the row and $\hat{h}_{t,k}^2$ denoted estimated condition variance of different GARCH models

4.5 Superior Predictive Ability (SPA) results

As we evaluate the performance of in-of-sample conditional volatility and out-of-sample forecast using fixed and rolling in term of loss function in the previous section. The result indicated that EGARCH model with alternative volatility proxies gives us the smallest RMSE. In this section, we also perform the superior predictive ability test to determine the null hypothesis is that none of the alternative models outperform the benchmark model in term of loss function.

To calculate the *p-values* of SPA, we first generate the loss function of benchmark model and alternative models; $L\left(\sigma_{t+j}^2, \hat{h}_{k,t+j|t}^2\right) = (\sigma_{t+j}^2 - \hat{h}_{k,t+j|t}^2)^2$ where σ_{t+j}^2 denoted the volatility proxies at time $t+j$ and $\hat{h}_{k,t+j|t}^2$ denoted *j*-step ahead forecast of GARCH family given the information at time *t*. We replicated each loss function for 10,000 times and the optimal of block length for each out-of-sample period followed Politis and White (2004). The *p-values* for SPA tests show how all the tests fail to reject the null hypothesis when the benchmark model outperform alternative models.

The value reports in Table 4.7, Table 4.8, Table 4.9, and Table 4.10 is the *p-value* of Superior Predictive Ability test for each GARCH models against alternative volatility proxies. The table presents for three range-based volatilities (PK, GK, and RS) and SR as the volatility proxies in the rows to assess the performance of fixed and rolling window out-of-sample forecast of four GARCH models. The family of GARCH benchmark models is presented in the columns and each three *p-values* of SPA: SPA_U, SPA_C and SPA_L are the upper, consistent, and lower *p-values* of SPA test, respectively. The three p-values correspond to different re-center of the losses. The *p-value* of upper bound (SPA_U) is a White's (2000) Reality Check test of the conservative test which assumes that all the alternative models are as accurate as the benchmark model. However, SPA_U is sensitive to including weak and irrelevant models in the comparison. While the consistent and liberal tests are not. The consistent *p-value* (SPA_C) is produced by the test for SPA of Hansen (2005) determine which models are worse than the benchmark. The lower bound (SPA_L) is the *p-value* of a liberal test which the null hypothesis assumes that the alternative models with worse performance than the benchmark model are poor models.

Table 4.7: The *p-value* of SPA test for alternative volatility proxies and different *j* periods forecast of GARCH-*type* with normal distribution models from <u>3,978 fixed windows in-of-sample estimation period</u> (17 February 1992 to 7 May 2008)

	Benchmark: ARCH (1)			**Benchmark: GARCH (1,1)**			**Benchmark: EGARCH (1,1)**			**Benchmark: GJR (1,1)**		
Proxy	**SPA$_U$**	**SPA$_C$**	**SPA$_L$**	**SPA$_U$**	**SPA$_C$**	**SPA$_L$**	**SPA$_U$**	**SPA$_C$**	**SPA$_L$**	**SPA$_U$**	**SPA$_C$**	**SPA$_L$**
Panel A: 22 periods ahead forecast (22 observations)												
SR	0.0013	0.0013	0.0013	0.2664	0.1668	0.1668	**1.0000**	**1.0000**	**1.0000**	0.3066	0.2159	0.1421
PK	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	**1.0000**	**1.0000**	**1.0000**	0.0000	0.0000	0.0000
GK	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	**1.0000**	**1.0000**	**1.0000**	0.0000	0.0000	0.0000
RS	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	**1.0000**	**1.0000**	**1.0000**	0.0000	0.0000	0.0000
Panel B: 55 periods ahead forecast (55 observations)												
SR	0.0189	0.0189	0.0189	0.0882	0.0673	0.0673	**1.0000**	**1.0000**	**1.0000**	0.4962	0.3325	0.2727
PK	0.0007	0.0007	0.0007	0.0045	0.0027	0.0027	**1.0000**	**1.0000**	**1.0000**	0.0000	0.0000	0.0000
GK	0.0004	0.0004	0.0004	0.0046	0.0029	0.0029	**1.0000**	**1.0000**	**1.0000**	0.0000	0.0000	0.0000
RS	0.0002	0.0002	0.0002	0.0047	0.0024	0.0024	**1.0000**	**1.0000**	**1.0000**	0.0000	0.0000	0.0000
Panel C: 110 periods ahead forecast (110 observations)												
SR	0.0533	0.0533	0.0533	0.1649	0.1649	0.0907	0.3128	0.3128	0.1526	**1.0000**	**1.0000**	**1.0000**
PK	0.0464	0.0464	0.0464	0.3357	0.1924	0.1391	**1.0000**	**1.0000**	**1.0000**	0.1375	0.1375	0.1101
GK	0.0441	0.0441	0.0441	0.3564	0.2116	0.1452	**1.0000**	**1.0000**	**1.0000**	0.0373	0.0373	0.022
RS	0.0417	0.0417	0.0417	0.3527	0.2111	0.1448	**1.0000**	**1.0000**	**1.0000**	0.03	0.03	0.0132
Panel D: 220 periods ahead forecast (220 observations)												
SR	0.0159	0.0159	0.0159	0.0682	0.0492	0.0492	0.2613	0.2437	0.1236	**1.0000**	**1.0000**	**1.0000**
PK	0.0832	0.0832	0.0832	0.3197	0.3197	0.157	**1.0000**	**1.0000**	**1.0000**	0.4089	0.4089	0.312
GK	0.1155	0.1155	0.1155	0.4567	0.4567	0.1814	**1.0000**	**1.0000**	**1.0000**	0.1586	0.1586	0.1201
RS	0.1243	0.1243	0.1243	0.5112	0.5112	0.2041	**1.0000**	**1.0000**	**1.0000**	0.0626	0.0626	0.0437

Note: SPA$_U$, SPA$_C$, and SPA$_L$ denoted *p-value* of White's (2000) Reality Check test, *p-value* of Hansen's consistent and liberal test. The bold letter shows *p-value* of alternative volatility proxies in the rows and estimated GARCH-*type* models with normal distribution in the columns fail to reject the null hypothesis which none of competing models outperforms the benchmark model. To calculated the *p-value* of SPA, we use 10,000 times of bootstrap replications and optimal of block length followed Politis and White (2004)

Table 4.8: The *p-value* of SPA test for alternative volatility proxies and different *j* periods forecast of GARCH-*type* with Student's t-distribution from 3,978 fixed windows in-of-sample estimation period (17 February 1992 to 7 May 2008)

Proxy	Benchmark: ARCH (1)			Benchmark: GARCH (1,1)			Benchmark: EGARCH (1,1)			Benchmark: GJR (1,1)		
	SPA_U	**SPA_C**	**SPA_L**	**SPA_U**	**SPA_C**	**SPA_L**	**SPA_U**	**SPA_C**	**SPA_L**	**SPA_U**	**SPA_C**	**SPA_L**
Panel A: 22 periods ahead forecast (22 observations)												
SR	0.0032	0.0032	0.0032	0.6715	0.5878	0.4177	0.2945	0.2609	0.2609	**1.0000**	**1.0000**	**1.0000**
PK	0.0000	0.0000	0.0000	0.0001	0.0001	0.0001	**1.0000**	**1.0000**	**1.0000**	0.0003	0.0003	0.0003
GK	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	**1.0000**	**1.0000**	**1.0000**	0.0000	0.0000	0.0000
RS	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	**1.0000**	**1.0000**	**1.0000**	0.0000	0.0000	0.0000
Panel B: 55 periods ahead forecast (55 observations)												
SR	0.0335	0.0335	0.0335	0.2751	0.1703	0.1109	0.0516	0.0516	0.0432	**1.0000**	**1.0000**	**1.0000**
PK	0.0000	0.0000	0.0000	0.0002	0.0002	0.0002	**1.0000**	**1.0000**	**1.0000**	0.0002	0.0002	0.0002
GK	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	**1.0000**	**1.0000**	**1.0000**	0.0000	0.0000	0.0000
RS	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	**1.0000**	**1.0000**	**1.0000**	0.0000	0.0000	0.0000
Panel C: 110 periods ahead forecast (110 observations)												
SR	0.0582	0.0582	0.0508	0.2953	0.2953	0.1156	0.059	0.059	0.059	**1.0000**	**1.0000**	**1.0000**
PK	0.0016	0.0016	0.0016	**1.0000**	**1.0000**	**1.0000**	0.5618	0.4409	0.4376	0.3361	0.2515	0.2515
GK	0.0002	0.0002	0.0002	0.8103	0.7052	0.4497	**1.0000**	**1.0000**	**1.0000**	0.2383	0.1764	0.1764
RS	0.0002	0.0002	0.0002	0.7653	0.6472	0.3994	**1.0000**	**1.0000**	**1.0000**	0.2129	0.1507	0.1507
Panel D: 220 periods ahead forecast (220 observations)												
SR	0.0490	0.0490	0.0467	0.3039	0.3039	0.1284	0.0819	0.0819	0.0819	**1.0000**	**1.0000**	**1.0000**
PK	0.2079	0.2079	0.2079	**1.0000**	**1.0000**	**1.0000**	0.5421	0.5421	0.4822	0.3678	0.3678	0.2838
GK	0.2762	0.2762	0.2762	0.6629	0.6629	0.3238	**1.0000**	**1.0000**	**1.0000**	0.0929	0.0929	0.0665
RS	0.1619	0.1619	0.1619	0.4917	0.3314	0.2138	**1.0000**	**1.0000**	**1.0000**	0.0358	0.0358	0.0221

Note: SPA_U, SPA_C, and SPA_L denoted *p-value* of White's (2000) Reality Check test, *p-value* of Hansen's consistent and liberal test. The bold letter shows *p-value* of alternative volatility proxies in the rows and estimated GARCH-*type* models with Student's t-distribution in the columns fail to reject the null hypothesis which none of competing models outperforms the benchmark model. To calculated the *p-value* of SPA, we use 10,000 times of bootstrap replications and optimal of block length followed Politis and White (2004)

The table 4.7 contains the *p-value* of SPA test for 22, 55, 110 and 220 periods ahead forecast conditional variance from estimation period of four GARCH-*type* models compared with different volatility proxies. For each j^{th} period ahead forecast, the nominal *p-value* of SPA_C for EGARCH model are all well above 0.10 and give the highest *p-value* which fails to reject the null hypothesis of no alternative models outperform EGARCH when using range-based estimators as actual volatility. However, for SR as volatility proxy, using GJR is the benchmark model is the best one.

The *p-value* of SPA test for all periods forecast of conditional variance of various GARCH models with Student's t-distribution presented in Table 4.8. Using range-based estimators as a proxy (PK, GK, and RS), EGARCH model tends to be preferred one because it has the largest *p-value* for all period ahead forecast except long term forecast (110 and 220 days forecast) which GARCH is better model. For SR as a proxy and GJR as a benchmark model, the *p-value* gives us the largest SPA_C value.

Besides, we also use SPA test with out-of-sample forecast of condition variance using rolling window which shows in Table 4.9 and Table 4.10. The result in Table 4.9 indicated the *p-value* of each GARCH under normal distribution. The EGARCH model can be the benchmark model when using various volatility proxies for short period forecast. However, for long-term prediction, using ARCH model as benchmark outperform than other types of GARCH models. According to the Table 4.2 and Figure 4.2 show that SET return has excess kurtosis or fat tail. We then assume the Student's t-distribution stead of normal distribution. The result in Table 4.10 presented EGARCH model beats three alternative GARCH models when using different volatility proxy for all period forecasts.

To summarize, result from SPA tests provides evidence that the range-based volatility estimators could be possibly used as the proxy of actual volatility while evaluating several periods forecasting of various GARCH-*type* models under normal and Student's t-distribution. Moreover, the results from SPA test are consistent with the loss function RMSE in the previous section

Table 4.9: The *p-value* of SPA test for alternative volatility proxies and different *j* periods forecast of GARCH-*type* with normal distribution benchmark models from <u>1,989 rolling windows out-of-sample forecast period</u> (8 May 2008 – 30 June 2016)

	Benchmark: ARCH (1)			**Benchmark: GARCH (1,1)**			**Benchmark: EGARCH (1,1)**			**Benchmark: GJR (1,1)**		
Proxy	**SPA_U**	**SPA_C**	**SPA_L**	**SPA_U**	**SPA_C**	**SPA_L**	**SPA_U**	**SPA_C**	**SPA_L**	**SPA_U**	**SPA_C**	**SPA_L**
Panel A: 22 periods ahead forecast (1,967 observations)												
SR	0.0000	0.0000	0.0000	0.2235	0.1058	0.1058	**1.0000**	**1.0000**	**1.0000**	0.0022	0.0022	0.0021
PK	0.0001	0.0001	0.0001	0.2995	0.1782	0.1384	**1.0000**	**1.0000**	**1.0000**	0.0007	0.0007	0.0007
GK	0.0000	0.0000	0.0000	0.3051	0.1799	0.1355	**1.0000**	**1.0000**	**1.0000**	0.0013	0.0013	0.0013
RS	0.0001	0.0001	0.0001	0.3067	0.1779	0.137	**1.0000**	**1.0000**	**1.0000**	0.0015	0.0015	0.0015
Panel B: 55 periods ahead forecast (1,934 observations)												
SR	0.1292	0.1292	0.1059	**1.0000**	**1.0000**	**1.0000**	0.0107	0.0038	0.0038	0.0012	0.0012	0.0012
PK	**1.0000**	**1.0000**	**1.0000**	0.7198	0.5292	0.444	0.0032	0.0022	0.0022	0.0000	0.0000	0.0000
GK	**1.0000**	**1.0000**	**1.0000**	0.6418	0.4388	0.37	0.0011	0.001	0.001	0.0000	0.0000	0.0000
RS	**1.0000**	**1.0000**	**1.0000**	0.6008	0.4027	0.3447	0.0006	0.0005	0.0005	0.0000	0.0000	0.0000
Panel C: 110 periods ahead forecast (1,879 observations)												
SR	**1.0000**	**1.0000**	**1.0000**	0.0013	0.001	0.0003	0.002	0.0017	0.0017	0.0008	0.0008	0.0008
PK	**1.0000**	**1.0000**	**1.0000**	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
GK	**1.0000**	**1.0000**	**1.0000**	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
RS	**1.0000**	**1.0000**	**1.0000**	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Panel D: 220 periods ahead forecast (1,769 observations)												
SR	**1.0000**	**1.0000**	**1.0000**	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0008	0.0008	0.0008
PK	**1.0000**	**1.0000**	**1.0000**	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0001	0.0001
GK	**1.0000**	**1.0000**	**1.0000**	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
RS	**1.0000**	**1.0000**	**1.0000**	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0001	0.0001

Note: SPA_U, SPA_C, and SPA_L denoted *p-value* of White's (2000) Reality Check test, *p-value* of Hansen's consistent and liberal test. The bold letter shows *p-value* of alternative volatility proxies in the rows and estimated GARCH-*type* models with normal distribution in the columns fail to reject the null hypothesis which none of competing models outperforms the benchmark model. To calculated the *p-value* of SPA, we use 10,000 times of bootstrap replications and optimal of block length followed Politis and White (2004)

Table 4.10: The *p-value* of SPA test for alternative volatility proxies and different *j* periods forecast of GARCH-*type* with Student's t-distribution benchmark models from 1,989 rolling windows out-of-sample forecast period (8 May 2008 – 30 June 2016)

	Benchmark: ARCH (1)			**Benchmark: GARCH (1,1)**			**Benchmark: EGARCH (1,1)**			**Benchmark: GJR (1,1)**		
Proxy	**SPA$_U$**	**SPA$_C$**	**SPA$_L$**	**SPA$_U$**	**SPA$_C$**	**SPA$_L$**	**SPA$_U$**	**SPA$_C$**	**SPA$_L$**	**SPA$_U$**	**SPA$_C$**	**SPA$_L$**
Panel A: 22 periods ahead forecast (1,967 observations)												
SR	0.0000	0.0000	0.0000	0.0679	0.0679	0.0678	**1.0000**	**1.0000**	**1.0000**	0.069	0.069	0.069
PK	0.0000	0.0000	0.0000	0.0532	0.0531	0.0531	**1.0000**	**1.0000**	**1.0000**	0.0396	0.0396	0.0396
GK	0.0000	0.0000	0.0000	0.049	0.049	0.049	**1.0000**	**1.0000**	**1.0000**	0.0393	0.0393	0.0393
RS	0.0000	0.0000	0.0000	0.0462	0.0461	0.0461	**1.0000**	**1.0000**	**1.0000**	0.0421	0.0421	0.0421
Panel B: 55 periods ahead forecast (1,934 observations)												
SR	0.0148	0.0137	0.0055	0.0689	0.0489	0.0489	**1.0000**	**1.0000**	**1.0000**	0.0002	0.0002	0.0002
PK	0.0000	0.0000	0.0000	0.0005	0.0005	0.0005	**1.0000**	**1.0000**	**1.0000**	0.0000	0.0000	0.0000
GK	0.0000	0.0000	0.0000	0.0011	0.0011	0.0011	**1.0000**	**1.0000**	**1.0000**	0.0000	0.0000	0.0000
RS	0.0000	0.0000	0.0000	0.0007	0.0007	0.0007	**1.0000**	**1.0000**	**1.0000**	0.0001	0.0001	0.0001
Panel C: 110 periods ahead forecast (1,879 observations)												
SR	0.0034	0.0002	0.0002	0.0000	0.0000	0.0000	**1.0000**	**1.0000**	**1.0000**	0.0000	0.0000	0.0000
PK	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	**1.0000**	**1.0000**	**1.0000**	0.0000	0.0000	0.0000
GK	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	**1.0000**	**1.0000**	**1.0000**	0.0000	0.0000	0.0000
RS	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	**1.0000**	**1.0000**	**1.0000**	0.0000	0.0000	0.0000
Panel D: 220 periods ahead forecast (1,769 observations)												
SR	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	**1.0000**	**1.0000**	**1.0000**	0.0000	0.0000	0.0000
PK	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	**1.0000**	**1.0000**	**1.0000**	0.0000	0.0000	0.0000
GK	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	**1.0000**	**1.0000**	**1.0000**	0.0000	0.0000	0.0000
RS	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	**1.0000**	**1.0000**	**1.0000**	0.0000	0.0000	0.0000

Note: SPA$_U$, SPA$_C$, and SPA$_L$ denoted *p-value* of White's (2000) Reality Check test, *p-value* of Hansen's consistent and liberal test. The bold letter shows *p-value* of alternative volatility proxies in the rows and estimated GARCH-*type* models with Student's t-distribution in the columns fail to reject the null hypothesis which none of competing models is outperform the benchmark model. To calculated the *p-value* of SPA, we use 10,000 times of bootstrap replications and optimal of block length followed Politis and White (2004)

4.6 The result of GARCH family models with exogenous variables

To address the second objective, we incorporated lag of volatility proxies as the exogenous variables to improve the out-of-sample performance of forecast as we mentioned in section 3.2.3. We first estimate the in-of-sample coefficient and in-of-sample condition variance. After that, we compare the loss function between in-of-sample conditional variance of normal GARCH and GARCH with exogenous variables. The table 4.11 and table 4.12 shows the loss function of in-of-sample conditional variance of normal GARCH and GARCH incorporated with SR and range-based estimators compared with volatility proxies. Mostly, RMSE of normal GARCH provides the smallest value of losses when using three range-based estimators as proxy.

Table 4.11: The RMSE of in-of-sample conditional variance of GARCH-*type* under normal distribution and GARCH-*type* under normal distribution with exogenous variables, using different volatility proxies

	In-of-Sample Conditional Variance	**In-of-Sample Conditional Variance of GARCH with different exogenous variables (SR, PK, GK and RS)**			
Panel A: ARCH model					
Proxy	**ARCH**	**ARCH-SR**	**ARCH-PK**	**ARCH-GK**	**ARCH-RS**
SR	**5.6520**	5.6527	-	-	-
PK	2.5064	-	**2.5060**	-	-
GK	2.8463	-	-	**2.8457**	-
RS	**3.2762**	-	-	-	3.2770
Panel B: GARCH model					
Proxy	**GARCH**	**GARCH-SR**	**GARCH-PK**	**GARCH-GK**	**GARCH-RS**
SR	7.0411	**7.0369**	-	-	-
PK	**3.4866**	-	3.4878	-	-
GK	**3.4307**	-	-	3.4336	-
RS	**3.6479**	-	-	-	3.6518
Panel C: EGARCH model					
Proxy	**EGARCH**	**EGARCH-SR**	**EGARCH-PK**	**EGARCH-GK**	**EGARCH-RS**
SR	6.9765	**6.7506**	-	-	-
PK	**3.1583**	-	3.1624	-	-
GK	**3.0972**	-	-	5.5675	-
RS	**3.3293**	-	-	-	5.4284
Panel D: GJR model					
Proxy	**EGARCH**	**EGARCH-SR**	**EGARCH-PK**	**EGARCH-GK**	**EGARCH-RS**
SR	6.9844	**6.9821**	-	-	-
PK	**3.3974**	-	3.3984	-	-
GK	**3.3491**	-	-	3.3513	-
RS	**3.5697**	-	-	-	3.5723

Note: The RMSE value which present in this table multiply by 10^{-4}. The bold letter provides the smallest RMSE values. The formula of $RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(\hat{\sigma}_t^2 - \hat{h}_{t,k}^2\right)^2}$, where $\hat{\sigma}_t^2$ denoted alternative volatility proxies in the row and $\hat{h}_{t,k}^2$ denoted estimated condition variance of different GARCH-*type* models

Table 4.12: The RMSE of in-of-sample conditional variance of GARCH-*type* under Student's t-distribution and GARCH-*type* under Student's t-distribution with exogenous variables, using different volatility proxies.

	In-of-Sample Conditional Variance	In-of-Sample Conditional Variance of GARCH with different exogenous variables (SR, PK, GK and RS)			
Panel A: ARCH model					
Proxy	**ARCH**	**ARCH-SR**	**ARCH-PK**	**ARCH-GK**	**ARCH-RS**
SR	6.1555	**0.6043**	-	-	-
PK	**2.5329**	-	5.4457	-	-
GK	**2.6692**	-	-	6.4773	-
RS	**3.0468**	-	-	-	7.0078
Panel B: GARCH model					
Proxy	**GARCH**	**GARCH-SR**	**GARCH-PK**	**GARCH-GK**	**GARCH-RS**
SR	6.9553	**6.9511**	-	-	-
PK	3.2582	-	**3.2578**	-	-
GK	**3.1865**	-	-	3.1878	-
RS	**3.4163**	-	-	-	3.4182
Panel C: EGARCH model					
Proxy	**EGARCH**	**EGARCH-SR**	**EGARCH-PK**	**EGARCH-GK**	**EGARCH-RS**
SR	**6.6314**	6.8797	-	-	-
PK	**2.9925**	-	6.8973	-	-
GK	**3.0649**	-	-	5.6328	-
RS	**3.3445**	-	-	-	5.5414
Panel D: GJR model					
Proxy	**EGARCH**	**EGARCH-SR**	**EGARCH-PK**	**EGARCH-GK**	**EGARCH-RS**
SR	6.8967	**6.8932**	-	-	-
PK	3.2021	-	**3.2018**	-	-
GK	**3.1486**	-	-	3.1498	-
RS	**3.3832**	-	-	-	3.3849

Note: The RMSE value which present in this table multiply by 10^{-4}. The bold letter provides the smallest RMSE values. The formula of $RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(\hat{\sigma}_t^2 - \hat{h}_{t,k}^2\right)^2}$, where $\hat{\sigma}_t^2$ denoted alternative volatility proxies in the row and $\hat{h}_{t,k}^2$ denoted estimated condition variance of different GARCH-*type* models

However, using SR estimator as a volatility proxy, we notice GARCH-family incorporated with lag of SR as an exogenous variable give the better performance except ARCH model. For Student's distribution, the results indicated that adding exogenous variable in EGARCH model give worse performance. When using RS and PK as a volatility proxy, the results are quite mixed depend on type of GARCH models.

According to RMSE of in-of-sample conditional volatility, we decide not to forecast the conditional variance. We conclude that incorporating volatility proxies into the GARCH variance equation not improve the performance of four GARCH models in evidence of stock exchange of Thailand.

CHAPTER 5
CONCLUSIONS AND RECOMMENDATIONS

This paper uses the different parametric volatility models to analyze and forecast the conditional variance of SET index return. To exhibit more of the stylized facts and characteristics of asset price volatility, the Gaussian distribution and the Student's t-distribution is assumed to be the error term's distribution in this study. To compare the performance of GARCH models which are ARCH, GARCH, EGARCH and GJR-GARCH, the four nonparametric volatility models (Squared return, Parkinson (1980), Garman and Klass (1980) and Rogers and Satchell (1991) estimators) are introduced as a proxy of the actual volatility of the daily returns. The concept of Root Mean Squared Error (RMSE) and Superior Predictive Ability (SPA) of Hansen (2005) are applied in this study to evaluate the goodness of fit of the models. Thank to the SPA test of Hansen (2005), we can reduce the risk of biased statistical inference due to data snooping.

There are 5,979 daily entire price observations (Close, High, Low, and Close price) of SET index during the observation period from 17 February 1992 until 30 June 2016. The in-of-sample period consisted of 3,978 daily observations (17 February 1992 to 7 May 2008) and the out-of-sample period consisted of 1,989 daily observations (8 May 2008 – 30 June 2016). We take the first difference to transform SET index because financial time series data is non-stationary which unpredictable and cannot be modeled or forecasted. After the data has been stationarized by differencing, we then check the preliminary statistic test of SET return. The finding presents the return of SET index is well-modeled by Student's t-distribution according to stylized fact of financial time-series data. Moreover, the SET index return exhibits conditional heteroscedasticity or autocorrelation in the squared series.

To address the first objective of this study, we compare GARCH models under normal and Student's t-distribution with various volatility proxies.

We first estimate the GARCH models from 8 May 2008 – 30 June 2016 (1,989 daily observations). Second, we forecast 22, 55, 110 and 220 days out-of-sample forecast using the fixed-window. With this model specification setting, in light of the lowest RMSE and p-value of the SPA test, EGARCH model under normal distribution provides a better performance among other models (ARCH, EGARCH and GJR-GARCH) compared to PK, GK, and RS estimators as volatility proxy for all periods forecast. With the assumption of Student's t-distribution, the results were striking depend on volatility proxy. With squared return, GJR-GARCH model tends to be the most preferred one while EGARCH model clearly better than others when using range-based volatility as the volatility proxy. Third, we forecast 22, 55, 110 and 220 days ahead out-of-sample with rolling-window by adding new observation and removing the first one for 1,989 observations. For normally distributed error term, the evidence points out that EGARCH model outperforms when we compare it with range-based models as a proxy for short-term rolling out-of-sample forecast. However, with squared return as a proxy, ARCH model is superior for long-term rolling out-of-sample forecast. Furthermore, EGARCH model under Student's t-distribution beats three alternative GARCH models when using different volatility proxy for all period forecasts.

To address the second objective, we add the volatility proxy as the exogenous variable in the GARCH-*type* models under normal and Student's t-distribution. We first calculated the RMSE of the conditional variance of GARCH-*type* models with volatility proxies (SR, PK, GK and RS estimators). Next, we computed the RMSE of four volatility proxies with conditional variance of GARCH-*type* models that includes the lag of each volatility proxy (GARCH-*type*-RS, GARCH-*type* -PK, GARCH-*type* -GK and GARCH-*type*-RS). In general, the results show that adding the range-based estimators into the GARCH-*type* models cannot improve the performance of volatility forecast because the values RMSE of normal GARCH-*type* models are still smaller than GARCH-*type* models with exogenous variables. However, adding SR as an exogenous variable can improve the GARCH-*type* volatility forecast.

For further research, another proxy for the latent volatility could be used is realized volatility which provides better efficient than range-based volatility. To further increase the understanding of volatility forecasting could consider different forecast horizons, other asset classes than equity indices or investigate the use of other error distributions such as the skew-normal distribution.

REFERENCES

1. Andersen, Torben G. 2000. "Some Reflections on Analysis of High-Frequency Data." In, 146. American Statistical Association.

2. Andersen, Torben G., and Tim Bollerslev. 1997. 'Intraday periodicity and volatility persistence in financial markets', *Journal of Empirical Finance*, 4: 115-58.

3. Andersen, Torben G., and Tim Bollerslev. 1998. "Answering the Skeptics: Yes, Standard Volatility Models do Provide Accurate Forecasts." In, 885. The Economics Department of the University of Pennsylvania, and the Osaka University Institute of Social and Economic Research Association.

4. Awartani, Basel M. A., and Valentina Corradi. 2005. 'Predicting the volatility of the S&P-500 stock index via GARCH models: the role of asymmetries', *International Journal of Forecasting*, 21: 167-83.

5. Beckers, Stan. 1983. "Variances of Security Price Returns Based on High, Low, and Closing Prices." In, 97. University of Chicago Press.

6. Bollerslev, Tim. 1986. 'Generalized autoregressive conditional heteroskedasticity', *Journal of Econometrics*: 307.

7. Christensen, Kim, and Mark Podolskij. 2007. 'Realized range-based estimation of integrated variance', *Journal of Econometrics*, 141: 323-49.

8. Engle, Robert F. 1982. "Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation." In, 987. The Econometric Society.

9. Engle, Robert F., and Byung Sam Yoo. 1987. 'Forecasting and testing in co-integrated systems', *Journal of Econometrics*: 143.

10. Garman, Mark B., and Michael J. Klass. 1980. "On the Estimation of Security Price Volatilities from Historical Data." In, 67. University of Chicago Press.

11. Hansen, Peter Reinhard. 2005. "A Test for Superior Predictive Ability." In, 365. American Statistical Association.

12. Hung, Jui-Cheng, Tien-Wei Lou, Yi-Hsien Wang, and Jun-De Lee. 2013. 'Evaluating and improving GARCH-based volatility forecasts with range-based estimators', *Applied Economics*, 45: 4041-49.

13. Martens, Martin, and Dick van Dijk. 2007. 'Measuring volatility with the realized range', *Journal of Econometrics*, 138: 181-207.

14. Nelson, Daniel B. 1991. 'Conditional Heteroskedasticity in Asset Returns: A New Approach', *Econometrica*, 59: 347-70.

15. Parkinson, Michael. 1980. "The Extreme Value Method for Estimating the Variance of the Rate of Return." In, 61. University of Chicago Press.

16. Phanathip, Chitayapuntagool. 2011. *On Forecasting Daily Stock Volatility: The role of intraday information and Market Conditions in Thailand*.

17. Politis, Dimitris N., and Halbert White. 2004. 'Automatic Block-Length Selection for the Dependent Bootstrap', *Econometric Reviews*, 23: 53-70.

18. Rogers, L. C. G., and S. E. Satchell. 1991. "Estimating Variance From High, Low and Closing Prices." In, 504. Institute of Mathematical Statistics.

19. Vortelinos, Dimitrios I. 2014. 'Optimally sampled realized range-based volatility estimators', *Research in International Business and Finance*, 30: 34-50.

20. White, Halbert. 2000. "A Reality Check for Data Snooping." In, 1097. Econometric Society.

21. Wiggins, James B. 1991. "Empirical Tests of the Bias and Efficiency of the Extreme-Value Variance Estimator for Common Stocks." In, 417. University of Chicago Press.

BIOGRAPHY

Name	Mr. Kritsanaphuk Leehanantakul
Date of Birth	November 29, 1990
Educational Attainment	2017: Master in Finance, Thammasat University
	2012: Bachelor of Economics, Thammasat University
Work Position	Senior Account Officer
	Krung Thai Bank Public Company Limited
Work Experiences	2013-present: Senior Account Officer
	Krung Thai Bank Public Company Limited
	2012: Online Marketing
	S&P Miracle Import Company Limited