Lund University

Department of Economics

Master Thesis in Economics, 15ECTS

Spring 2014

Reality Check for the Value-at-Risk Estimates of the Energy Commodities

---

Author: Dingquan Miao*   Supervisor: Professor Birger Nilsson

October 2014

*E-mail: eco13dmi@student.lu.se

Abstract

The fluctuations of the price in the energy market affect the households, firms and the government intuitions. We can perceive the information of the energy market from daily economic news. The entire society is concerned for the events that affect the energy market and the changing prices of the energy resources. It is thus meaningful and interesting to study the risk of the energy market. This paper provides empirical study for three representative energy resources from the year 1997 to 2013 by using the value-at-risk (VaR) estimation. The representative energy resources are natural gas, (Brent and WTI) crude oil and propane. In order to generate a serious study and consider both calm periods and volatile periods, the sample period is divided into 12 subsamples by using "rolling window" method. The investigation is designed to select the most adequate VaR estimates by applying three types of non-parametric approaches, namely the standard historical simulation (HS), the historical simulation with ARMA forecasting (HSAF) and the volatility weighted historical simulation (VWHS). In light of my empirical study, value-at-risk estimates at the 95% confidence level ($VaR_{95\%}$) generally perform poorly in explaining the risk of the three representative energy resources, and value-at-risk estimates at the 99% confidence level ($VaR_{99\%}$) are generally capable to explain the risk of the three representative energy resources (except for the financial crisis year 2008). Meanwhile, the results show that the HSAF approach and the VWHS approach perform slightly better than the standard HS approach, and the $VaR_{99\%}$ estimates of VWHS approach can explain the risk occurred in natural gas and Brent crude oil for all the subsample periods. More importantly, it seems that $VaR_{99\%}$ estimates of student t-distributed asymmetric VWHS models are qualified for both calm and volatile periods for natural gas and Brent crude oil.

Keywords: Energy commodities, Value-at-Risk (VaR), Historical Simulation (HS), Historical simulation with ARMA forecasting (HSAF), Autoregressive moving average process (ARMA), Volatility weighted historical simulation (VWHS), (Generalized) Autoregressive Conditional Heteroskedasticity (ARCH/GARCH), Threshold GARCH (TARCH), Exponential GARCH (EGARCH).

Acknowledgement

I would like thank my parents for their financial support and great confidence in me all through these years.

I would like to express my deepest gratitude to my supervisor professor Birger Nilsson, for his endless patience, constant encouragement and invaluable help.

Dingquan Miao
October 2014, Lund, Sweden

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

ACF – Autocorrelation Function

AIC – Akaike Information Criterion

ARCH/GARCH – Autoregressive Conditional Heteroskedasticity/ Generalized ARCH

ARMA Process – Autoregressive Moving-Average Process

BIC – Bayesian Information Criterion or Schwarz Criterion

EGARCH – Exponential GARCH

HS - Historical Simulation

HSAF - Historical Simulation with ARMA Forecasting

HSAF-AR(1)/MA(1) – Historical Simulation with AR(1) or MA(1) Forecasting

PACF – Partial Autocorrelation Function

TARCH – Threshold GARCH

VaR_α – Value-at-Risk at confidence level α

VWHS – Volatility Weighted Historical Simulation

VWHS-EGARCH(1,1) – EGARCH(1,1) Based Volatility Weighted Historical Simulation

VWHS-fat tailed-EGARCH(1,1) – Student t-distributed EGARCH(1,1) Based Volatility Weighted Historical Simulation

VWHS-TARCH(1,1) – TARCH(1,1) Based Volatility Weighted Historical Simulation

VWHS-fat tailed-TARCH(1,1) – Student t-distributed TARCH(1,1) Based Volatility Weighted Historical Simulation

1. INTRODUCTION

"Economics is the study of how societies use scarce resources to produce valuable commodities and distribute them among different people" (Samuelson & Nordhaus, 1992)

This paper aims to investigate the risk of different commodities in the energy market. More specifically, whether the selected risk quantification methods can really describe the risks occur in the energy market. The three representative energy commodities investigated in this thesis are natural gas, crude oil and propane, which all belong to the set of scarce resources in the modern society.

What is an appropriate measure to the risks occur in the energy market? As we know the underlying risk in the energy market may lead to economic depression. Sadorsky (1999) stated that oil prices and the volatility of oil price show significant impact on the economy, but changes in economic activity show limited impact on oil prices. For instance, the 1970s energy crisis made the oil price increase dramatically, and individuals and households suffered from unreasonable high oil price[1]. As a result, there is an increasing demand for the market and stakeholders (investors, regulators and etc.) to seek for the appropriate "risk indicators" to manage the risk. Standard deviation is considered as a primary risk measurement in financial studies. However, the application of standard deviation is not sufficient in some perspectives for the modern risk management, and there exists an alternative risk measurement, namely value-at-risk (VaR). Unlike the standard deviation, VaR gives an intuitive explanation that describes the risk of a loss in units, e.g. how much we will lose at most at the certain statistic confidence level. In many cases, VaR is a more preferable risk measurement than the standard deviation, since VaR allows for risk aggregation and takes an overall perspective on risk and has better explanations in terms of risk quantification[2].

Does VaR risk measurement always provide reliable estimates in practice? According to the concept of the VaR, we know that there are some potential limitations with this measurement

[1] The oil price has increased from \$3 to nearly \$12 per barrel from October 1973 to March 1974, see section 7 REFERENCES.
[2] The VaR definition and its quantification method will be introduced in section 3.1 Value-at-Risk.

and its estimation approaches. The concept of VaR can be explained by a critical point of the amount that we can lose at most at the certain level. In other words, the loss of "tail-events" under a certain probability, but value-at-risk does not make any interpretations about the size (amount) of these "tail-events". For example, investors estimate a portfolio's VaR at 95% confidence level to be 10 units (can be any currency), and at the same time there is 4% probability that 1 billion units loss can happen and it is not detected. Even the realized possibility of this 1 billion units' loss is smaller than 5%, it does not mean that this extreme large loss will not happen for sure, and it will cause a serious crisis if it happens. Hence, the VaR estimate is not always reliable, but it does offer a logic way of interpreting risk.

Do VaR estimation approaches provide satisfactory performance in the energy market? This paper contributes an answer to this question by implementing the reality check of the performance of VaR estimates in the energy market. In fact, value-at-risk is the most commonly used risk measurement in the energy market[3], and the application of value-at-risk in the energy market becomes more and more interesting and has drawn a lot of attentions since mid-1990s. It is thus meaningful and interesting to check the results of the VaR estimates in the energy market during a certain study period[4], e.g. can VaR estimates quantify and explain the risk happened in the energy market.

What are the most adequate VaR estimates in the energy market? The VaR estimation method can be classified into a non-parametric approach and a parametric approach, and the main distinction between the two approaches is the assumption of the distribution[5]. The non-parametric approach relaxes the assumption of a pre-determined loss distribution, instead it is based on the empirical loss distribution directly. This paper applies three non-parametric approaches, which are the standard historical simulation approach (HS), the historical simulation with ARMA forecasting approach (HSAF) and the volatility weighted historical

---

[3] Based on the Energy Risk's 2009 Risk Management Survey, there are 85% of participants answered that they use VaR metrics, and those participants are mainly traders, risk managers and senior executives from energy producers and energy trading firms.

[4] The empirical test period is from January 1997 to December 2013, more information of data processing procedure will be introduced in section 4.1. DATA.

[5] The VaR estimates are computed by estimating the important moments of the distributions. For example, the first moment μ and the second moment σ^2, and the assumed distribution can be normal distribution, student t distribution, lognormal distribution and etc.

simulation approach (VWHS). There are several merits to use a non-parametric approach instead of the parametric approach in my empirical study. First, it is more persuasive and trustful to let the historical data speak for themselves, since no matter which instruments we want to investigate we can always let the VaR estimation reflect its own historical data series. Second, by using the historical loss series it is relatively easier to interpret the VaR estimation, e.g. in light of its definition VaR is basically the percentile that we can lose at most. Third, to consider the investigated market in this paper, it is not clear how the energy commodities distribute during my study period and which specific parametric distribution to assume for the data series. Hence, it seems more reasonable to use the non-parametric approach, where the distributions for the VaR estimates are the distributions of the historical loss series.

What are the limitations of the non-parametric VaR estimation in the energy market? As every coin has two sides, there are certainly some disadvantages with the non-parametric estimation. For instance, there is no clear rule of how to select the appropriate sample size, because there is a tradeoff between VaR estimation and relevant information, e.g. a sufficient sample size is needed to estimate VaR, but this sufficient sample size may contain too much irrelevant information. More importantly, even if we can decide the appropriate sample size, it is difficult to decide which sample periods to use, since sometimes it is hard to distinguish the calm period and volatile period, which will affect the estimation results of VaR. Meanwhile, when VaR estimates are based on non-parametric historical simulation process, it usually takes time to detect shifts[6] and reflect new major events, and the energy price is very sensitive to the change of the macro-conditions, such as government regulations, market failure, and politics and wars etc. Can we alleviate the impact of the drawbacks derived from the non-parametric VaR estimation? This paper abates the effect of the above limitations by using subsample analysis technique and the composition of the model selection and VaR backtesting procedure[7].

The main idea of this thesis is to check the performance of VaR estimates for the energy commodities by connecting different estimated models with the real world VaR estimation

[6] For example, change in the market rates.

[7] Detailed arguments are given in section 4. EMPIRICAL ANALYSIS.

results. According to my empirical study, I found that the VWHS and the HSAF approaches perform slightly better than the standard HS approach. More importantly, I found the $VaR_{99\%}$ estimates of the VWHS approach are capable to explain the risk of the natural gas and the Brent crude oil for both calm and volatile periods[8].

The structure of the thesis is set as follows: section 2 provides some previous research related to the VaR applications. Section 3 explains the concept and the quantification method of VaR and the methodology of three non-parametric estimation processes. Section 4 illustrates the data processing procedure and presents the empirical results with detailed analysis. Section 5 summarizes the essential results of the entire empirical study.

2. PREVIOUS RESEARCH

From the 1970s, in order to tackle the problem of increasing risk from the financial firms, regulators require these firms to maintain a certain level of capital reverses, e.g. capital ratios. The process of setting the appropriate capital requirements raises the need of evaluating the potential loss of the investments at the certain statistical level with the certain holding periods, and this is the prototype of the value-at-risk or capital-at-risk. The concept VaR has been widely used and applied from the mid-1990s. In 1994, the Bank for International Settlements Fisher report required financial intermediaries release their measures of VaR publicly. In 1995, J.P. Morgan developed the "RiskMetrics" system and used this system to explain the term VaR and quantify the risk.

This paper is inspired by some previous studies of VaR in the energy market. Cabedo and Moya (2003) gave an example of employing historical simulation approach to estimate VaR for the oil market, and their sample data are the daily prices of Brent Crude Oil from 1992 to 1998 (in-sample period). Cabedo and Moya also introduced and developed the historical simulation ARMA forecasting (HSAF) approach in the same paper as a comparison to the standard HS, and they found that the VaR estimates with autoregressive moving average

[8] The notation of $VaR_{99\%}$ indicates the VaR estimation at 99% confidence level, which is introduced in section 3.1.2. Definition of Value-at-Risk.

models are more sensitive to the change of the variance in the oil market than the standard HS, and the HSAF approach provides efficient risk quantification for the oil market. Afterwards, Sadeghi and Shavvalpour (2005) gave elaborate complement of the HSAF methodology and GARCH models by using weekly OPEC prices from January 1997 to December 2002[9], and from their research the HSAF approach provides the most efficient results that are consistent with the results from Cabedo and Moya. In fact, many researchers have started to improve the HS approach by applying different variance-covariance models for many years. For example, Boudoukh, Richardson and Whitelaw (1998) studied VaR by applying the hybrid approach[10]; after that, Hull and White (1998) proposed the HW approach, which suggested a GARCH or exponential weighted moving average (EWMA) model to stress the volatility property in historical data. Additionally, Andriosopoulos and Nomikos (2013) presented the Monte Carlo simulation and hybrid Monte Carlo with the HS approach to model VaR, and they found that these two methods can explain the VaR of energy prices efficiently.

Previous researches provide two important evidences for the energy commodities, which will also be tested in my empirical study. Firstly, estimated models under fat-tailed distributions are usually more suitable to explain risk for the energy commodities; secondly, the risks of the energy commodities are affected by the asymmetric market information. For example, Hung, Lee and Liu (2008) studied five commodities from the energy market[11], and their results suggest that VaR estimates of heavy-tailed distributions are more suitable for the energy commodities. And in light of Giot and Laurent (2003) study results, the skewed Student APARCH model is the most preferable for all cases in their study[12], which are 5-year out-of-sample commodities research.

---

[9] For instance, Sadeghi and Shavvalpour use Augmented Dickey-Fuller unit roots test instead of the Ljung-Box test in the ARMA models testing procedure, and present more details implementation of the HSAF procedure and etc.

[10] It is the combination of historical simulation and RiskMetrics approach, and the improvement of hybrid approach is to use the decay factor to assign different weights for the realized returns.

[11] The five energy commodities are WTI crude oil, Brent crude oil, heating oil #2, propane and New York Harbor Conventional Gasoline Regular.

[12] The investigated commodities in their empirical study are aluminum, copper, nickel, Brent crude oil and WTI crude oil daily cash prices and cocoa nearby futures contracts.

The empirical study in this paper tries to combine some of methodologies from previous research, and aims to generate more accurate results by applying more subsample periods and the larger sample size. From the perspective of the non-parametric approaches, my empirical study models VaR estimates from the standard HS approach, the HSAF approach and the VWHS approach. From the distribution perspective, this paper estimates variance-covariance models by both normal distribution and student t-distribution. From the asymmetric effect modeling perspective, this empirical study applies conditional-variance model from standard GARCH to Threshold GARCH (TARCH) and Exponential GARCH (EARCH). From the view of the sample of my empirical study, this paper investigates the performance of 12 out-of-samples VaR estimates for each commodity, and each out-of-sample forecasting is based on 5 years in-sample period[13]. In a word, this paper aims to provide a relatively complete and careful empirical study for the VaR estimates of energy commodities.

3. METHODOLOGY

3.1. VALUE-AT-RISK

3.1.1. The rise of Value-at-Risk

In 1950s, Harry Markowitz developed the mean-variance model, which assumes rational investors prefer higher expected return and lower risk (risk averse). After that, the science of risk management has been widely studied and has become an important subfield in the finance discipline. Variance (or standard deviation) is an important component of the mean-variance model, and it describes the level of volatility, e.g. does one portfolio yield higher risk than another portfolio. Variance is thus considered as the standard measurements to measure the risk of assets.

However, because of the randomness and uncertainty of the stock market, variance tells limited information about future returns and losses. For example, a high forecasting volatility

[13] In total, there are 48 out-of-samples forecasting periods, and more details are introduced in section 4.1. DATA.

for tomorrow only infers that the future price maybe unusually high or low, i.e. we cannot know the accurate stock price and whether it is a return or loss. In addition, because of modern firms' complex structure, we cannot detect all the potential risks by using variance measurement. Hence, it's necessary to come up with a more efficient and intuitive measurement, and VaR is considered as an established risk measurement in recent decades. The VaR can quantify the risk in terms of units, and it offers a way to measure the aggregate risks, which takes all the risk factors into account. This means that we can estimate what we will lose at most simply based on the loss distribution.

3.1.2. Definition of Value-at-Risk

The VaR is a risk assessment that quantifies loss either in percentage term or in unit term, the value-at-risk equation can be expressed as:

$$VaR_\alpha = min\{l: Pr(L > l) \leq 1 - \alpha\} \quad (1)$$

Based on the above mathematical expression, we define VaR as the smallest loss l that the probability of future assets' (stocks, bonds, derivatives etc.) loss L larger than this smallest loss is equal to or smaller than $1 - \alpha$ within a holding period, and α is the confidence level that can be any value between 0 and 1. In a simpler way of explanation, VaR is the value you expect to lose at most at a certain percentage, e.g. the next year $VaR_{95\%} = 1000$ indicates that there is 95% probability that investment A is expected to lose at most 1000kr in next year. Because of its intuitive explanation, the concept of VaR is widely implied in the risk management, especially in the financial industry.

For a continuous distribution, α is the quantile of the loss distribution and it usually takes value as 90%, 95% or 99%. Note that some textbooks and research papers use profit and return distributions when estimate VaR. This paper uses loss distribution directly, since it is more intuitive and easier to interpret. For example, if we create a probability distribution function for the daily loss of the stock markets (include mean), positive values will indicate positive losses, and negative values will indicate negative losses that equivalent to gains. And for a commonly used confidence interval $\alpha = 99\%$, if we assume the underlying loss distribution follows the normal distribution, and then the corresponding $VaR_{99\%}$ is 2.326. If

$\alpha = 1\%$, and then the $VaR_{1\%}$ equals to -2.326. This is a reasonable result and is consist with the definition of the VaR, e.g. the most we can lose at 99% probability must be higher or equal to the most we can lose at 1% probability.

3.1.3. Value-at-Risk quantification method

Assume under a continuous loss distribution D, and loss L follows loss distribution D, with mean equal to μ and variance equal to σ^2, and the loss can thus be expressed as $L \sim D(\mu, \sigma)$.

$$Pr\big(L > VaR_\alpha(L)\big) = Pr\left(\frac{L-\mu}{\sigma} > \frac{VaR_\alpha(L)-\mu}{\sigma}\right)$$

$$= Pr\left(l > \frac{VaR_\alpha(L)-\mu}{\sigma}\right) = 1 - Pr(l \leq \frac{VaR_\alpha(L)-\mu}{\sigma}) = 1-\alpha$$

$$l_\alpha = \frac{VaR_\alpha(L)-\mu}{\sigma}, VaR_\alpha(L) = \mu + \sigma l_\alpha$$

From the above mathematical expressions, l denotes the standardized stochastic variable$L$, and l_α denotes the α-quantile of the loss distribution. The above expression can be expressed as $VaR_\alpha(L) = \sigma l_\alpha$ directly if we exclude the mean[14].

3.2. THE HISTORICAL SIMULATION METHODOLOGY

3.2.1. The Standard Historical Simulation Methodology

Historical simulation, also known as the standard (basic) historical simulation is the cornerstone of other non-parametric and semi-parametric VaR estimation methods. By contrast with the parametric approaches of VaR estimation, historical simulation does not rely on strict assumptions of the distributions like many other financial models, instead historical simulation is based on the historical data (e.g. sample of observed losses) to calculate VaR. With this attractive merit, we let the data speak for themselves based on previous data, which

[14] The excluded form of definition will be applied in the historical simulation with ARMA forecasting approach.

mimic the losses of all the market events. More importantly, we do not need to know the relevant parameters of the distribution.

In light of the introduction above, historical simulation is a method that simply uses historical data to forecast VaR at the certain statistical level. One of the important issues for the standard historical simulation approach is to assign equal weights to the historical data, and loss distributions are assumed to be independently and identically distributed. By using this standard historical simulation approach, the VaR estimates are just the percentile of the empirical loss distribution. We could implement the VaR estimation and backtesting process of the standard HS approach by using software MS Excel[15]. For instance, we obtain the first out-of-sample VaR estimate by using the PERCENTILE() formula in Excel, and the array within the parentheses is the first in-sample period data and the percentile applied in this paper is 95% and 99%. After obtaining the first out-of-sample estimation, we can then apply the rolling window method to compute the rest of the out-of-sample VaR estimates. It is worth to briefly introduce the process of deciding the in-sample size given the total sample size and the confidence level. Assume we have the sample size N and the confidence level α, based on the definition of VaR from the equation (1), we expect to find $N(1-\alpha)$ number of losses that are larger than $VaR_\alpha(L)$, and therefore $VaR_\alpha(L)$ is the $N(1-\alpha)+1$ largest loss. And no matter $N(1-\alpha)$ is an integer or not, we can always use the definition $Pr(L > l) \leq 1-\alpha$ to find the correct VaR at the confidence level α. This paper divides the total sample into several subsamples to increase the precision of the empirical test, and the notation of subsample size is n (in contrast to the notation of total sample size N)[16].

3.2.2. The Historical Simulation with ARMA forecasting (HSAF) Methodology

Cabedo and Moya (2003) suggested that by embedding the time series model in the historical simulation we can generate more accurate estimations of VaR, and they introduced the historical simulation ARMA forecasting approach that is derived from the historical simulation. After Cabedo and Moya (2003) proposed the HSAF approach, Sadeghi and

---

[15] The backtesting process is introduced in section 3.4. Backtesting.
[16] More information about sample data will be introduced in section 4.1. DATA.

Shavvalpour (2005) made more detailed explanation of the HSAF model. The HSAF model in this paper has made minor adjustments in the method implementation procedure[17]. Basically, the HSAF approach in this paper consists of several steps and these steps can be classified into two parts, which are ARMA generating process and VaR estimation process.

As the HSAF approach is the combination of standard historical simulation method and ARMA process, it is meaningful to briefly introduce the autoregressive moving-average model (ARMA) before we introduce the HSAF approach. The general ARMA(p, q) process can be expressed as:

$y_t = \alpha_0 + \alpha_1 y_{t-1} + \alpha_2 y_{t-2} + \cdots + \alpha_p y_{t-p} + \varepsilon_t + \beta_1 \varepsilon_{t-1} + \beta_2 \varepsilon_{t-2} + \cdots + \beta_q \varepsilon_{t-q}$, or in a more compact way:

$$y_t = \alpha_0 + \sum_{i=1}^{p} \alpha_i y_{t-i} + \sum_{i=0}^{q} \beta_i \varepsilon_{t-i} \qquad (2)$$

The above equation includes three components, which are the intercept α_0, the sum of autoregressive part y_{t-i} and the sum of the moving average of error terms ε_{t-i}. The interpretation of the expression is simply the current value y_t depends on its own lagged value and the moving average of current and previous shocks. The notation p is the number of lags contains in the model, and q indicates the number of pervious periods' shocks that affects y_t, and β_0 can take the normalized value 1.

Again, HSAF method stands for the historical simulation with ARMA forecasting, thus our primary task is to find the most appropriate ARMA model by applying BOX-JENKINS model selection criteria. After that I use the selected model to rescale the loss series and then apply historical simulation technique to estimate VaR.

First, we need to test the stationarity of the data series, since the non-stationary process might lead to spurious regression[18], which make the estimated models meaningless. In order to check the stationarity, we need to apply the unit roots testing method. There are four loss series tested at this stage, which are natural gas, Brent crude oil and WTI crude oil and propane. The common used method of unit roots test is the Augmented Dickey Fuller test, and

---

[17] The adjustments are introduced in the later part of this section.

[18] For example, a spurious regression may yield unreliable high R^2 and t-statistics, even the model has no economic meaning.

in order to strengthen the power of the test, this paper selects more rigorous unit roots testing approach which was suggested by Sims, Stock and Watson (1990). This unit roots testing approach helps to detect the potential deterministic regressors, which enhance the power of the simple Augmented Dickey Fuller test. The unit roots testing in this paper can thus test both difference stationarity and trend stationarity. If the results indicate stationarity we can then continue to test the autocorrelation of the data series; and if the data series is non-stationary, we can take the difference of the series until they become stationary series I(0).

Next step is to select the most appropriate ARMA model, which includes the case of no autocorrelation. As mentioned above Box-Jenkins model selection criterion is the essence of this part, and the ARMA estimation process employs the three-stage model selection method (Box and Jenkins, 1976), which is identification stage, estimation stage and diagnostic checking stage.

To have an overall inspection about the data series, we need to go through the identification stage. And in order to identify the basic condition of the series, we can plot the data series and draw the autocorrelation function (ACF) and partial autocorrelation function (PACF). These will give us a general idea of whether it includes a deterministic trend and outliers etc.; and by comparing the ACF and PACF from Ljung-Boxtest statistic, we can obtain the rough information of the possible ARMA process. If there is no autocorrelation then the HSAF approach is equivalent to the standard HS approach.

In the estimation stage, we can derive the estimated ARMA process by estimating the coefficients of each model. In light of the Box-Jenkins model selection criteria, we need to select the most parsimonious model. For instance, the most parsimonious model should satisfy some desirable conditions simultaneously: high value of goodness-of-fit and low Akaike's information Criterion (AIC) (Akaike, 1973) value or Schwarz Bayesian Information Criterion (BIC) (Schwarz, 1978) value, and the model match the trend found in ACF and PACF. Note that, the sample size in my empirical test is considered to be large, and even for subsamples [19]. In this case, the Schwarz Bayesian Information Criterion (BIC) punish

[19] As mentioned in section 4.1. Data, each subsample contains approximately 260 observations.

parameters more seriously than the Akaike's information Criterion (AIC), and BIC is a more preferable reference in the empirical test as it is more likely to select the correct model.

After generating the ARMA model, we can test the residuals of the selected model, and the residuals should follow the white noise process. The detection of the serial correlation can be applied by using the Ljung-Box portmanteau test statistic (Q-statistic), and $Q_k = T(T+2)\sum_{k=1}^{K}(\frac{1}{T-k}r_k^2)$ follows the chi square distribution. If there exists no residual autocorrelation the selected ARMA model passes the diagnostic test, otherwise the suggested model is rejected.

The second part is to use the historical simulation approach to estimate VaR based on the selected ARMA model. After estimating the appropriate ARMA process for different subsample periods for all energy commodities, I need to rescale the loss distributions of the energy resources, and then estimate VaR based on rescaled loss distributions. It is worth to notice that the rescaling process of the HSAF approach is not the same with the rescaling process with VWHS introduced in the section 3.3.4. For example, for AR(1) process I can first obtain the estimated losses by using the estimated ARMA models, and then compute residuals by using realized loss subtract the estimated loss. After that I compute the percentile of the residual distribution at the 95% and 99% level, and then I add the estimated loss with the 95% and 99% residual percentile respectively. The summations are thus the estimated VaR at the 95% and 99% confidence level. For MA(1) process, I need to generate the estimated losses to calculate the corresponding residuals that is the same with the AR(1) process, and the difference here is to calculate the unexpected losses for the last period instead of using last period loss directly, and it can be computed by $\varepsilon_{t-1} = l_{t-1} - \bar{l}$, $\bar{l}$ is the average loss of the subsample. After we get the residuals of MA(1), the VaR estimation method is the same with AR(1) process. For ARMA(1,1) process, the rescaling method is merely the combination of AR(1) and MA(1).

3.3. THE WEIGHTED HISTORICAL SIMULATION METHODOLOGY

3.3.1. The Volatility Weighted Historical Simulation-ARCH/GARCH Methodology

As Pritsker (2003) mentioned, age weighed historical simulation based VaR estimates are not sufficient to reflect the changeable underlying risk in the financial market, which makes it necessary to derive time related models to estimate VaR, such as ARCH and its relevant variance models. In the stock market, the variance of data series (such as stock returns) is not constant over time, and it is frequently seen that the current period is more likely to show high/low volatilities, if its previous periods show high/low volatilities. In 1982, Engle first proposed the concept of autoregressive conditional heteroskedasticity (ARCH) model, which explains this volatility clustering phenomenon.

As the energy market is affected by market information and shocks, it is natural to suspect that the price in the energy market also exhibits volatility clustering phenomenon. To elaborate the ARCH model, the essence of this model is to relax the constant variance assumption of the white noise process and allows the error term to be conditional on the previous error term, which means the volatility of the current shock depends on the volatility of previous realized shocks. The ARCH(1) model can be shown as $\varepsilon_t = v_t\sqrt{\alpha_0 + \alpha_1\varepsilon_{t-1}^2}$, where v_t should satisfy for certain assumptions and limitations. For example, v_t follows a white noise process with variance $\sigma_{v_t} = 1$, and v_t and ε_{t-1} are independent, and the restriction for lagged parameter $0 \leq \alpha_1 < 1$ to be stationary, α_0 should be positive. The expression of ARCH(1) model is shown as[20]:

$$h_t = E[\varepsilon_t^2|\varepsilon_{t-1}, \varepsilon_{t-2}, \varepsilon_{t-3} \dots] = \alpha_0 + \alpha_1\varepsilon_{t-1}^2 \quad (4a)$$

Where we denote h_t as the conditional variance at time t, and the more general ARCH(p) model can be simply expressed as two following ways:

$$\varepsilon_t = v_t\sqrt{\alpha_0 + \alpha_1\varepsilon_{t-1}^2 + \alpha_2\varepsilon_{t-2}^2 + \alpha_3\varepsilon_{t-3}^2 + \cdots + \alpha_p\varepsilon_{t-p}^2} \quad (4b)$$

$$h_t = \alpha_0 + \alpha_1\varepsilon_{t-1}^2 + \alpha_2\varepsilon_{t-2}^2 + \alpha_3\varepsilon_{t-3}^2 + \cdots + \alpha_p\varepsilon_{t-p}^2 \quad (4c)$$

[20] The brief derivation of ARCH(1) model is given in APPENDIX 6.5.

In light of the ARCH process, Bollerslev (1986) proposed the generalized autoregressive conditional heteroskedasticity (GARCH) process to capture the volatility dynamics. The GARCH model suggests that the variance of a shock in the current period not only depends on the volatility of previous realized shocks but also depends on the previous conditional variance of shocks. The standard GARCH(1,1) model can be expressed as $h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 h_{t-1}$ with $\varepsilon_t = v_t \sqrt{h_t}$, which is similar to the ARMA process, the conditional variance h_t depends on both autoregressive process and moving average process. For a more general GARCH (p, q) model can then be expressed as:

$$h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \cdots + \alpha_p \varepsilon_{t-p}^2 + \beta_1 h_{t-1} + \cdots + \beta_q h_{t-q} \ \textit{with} \ \varepsilon_t = v_t \sqrt{h_t} \tag{5}$$

Notice that there is an alternative approach that can take the volatility clustering into account, which is the exponentially weighted moving average (EWMA). And EWMA can be expressed as a special case of GARCH(1,1) when there is a reasonable large observing periods in the sample[21].

3.3.2. The Volatility Weighted Historical Simulation – TARCH Methodology

It is not hard to understand nearly all the markets are affected by the information released from their market, the question is how the markets react to this information, and do they react similarly to positive and negative shocks? Zakoian (1994) provided the evidence that volatility of stock returns react differently to positive and negative shocks by using the TARCH model, which proved an asymmetry property in the stock market. Does the energy market also react asymmetrically when facing positive and negative shocks? In order to detect the asymmetry in the energy market, we need to introduce the TARCH model.

The TARCH stands for Threshold-ARCH model, this model is based on the essence that shocks greater than the threshold have different effects than shocks below the threshold. The word "threshold" is derived from the leverage effect, as the market is affected by both good and bad news, the threshold point is the point where the market is neither affected by good

---

[21] In that case, EWMA can be expressed as $h_t = (1-\lambda)\varepsilon_{t-1}^2 + \lambda h_{t-1}$, which is equivalent to a GARCH(1,1) model with $\alpha_0 = 0, \alpha_1 = 1-\lambda$ and $\beta_1 = \lambda$.

nor bad news. Leverage effect states a phenomenon that volatilities after bad information increase more than volatilities following good information. And the TARCH(1,1) model can be expressed as

$$h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \lambda_1 d_{t-1} \varepsilon_{t-1}^2 + \beta_1 h_{t-1} \tag{6}$$

Here, a dummy variable $d_{t-1} = \begin{cases} 1 \; if \; \varepsilon_{t-1} < 0 \\ 0 \; if \; \varepsilon_{t-1} \geq 0 \end{cases}$ is included. From the equation (6) we can see $\varepsilon_{t-1} = 0$ is a threshold that partition two different types of shocks. The effect is $\alpha_1 + \lambda_1$ when ε_{t-1} is negative, and the effect is α_1 when ε_{t-1} is positive. It is worth to notice that given $\lambda_1 > 0$, negative shocks will have larger effects on h_t than positive ones.

3.3.3. The Volatility Weighted Historical Simulation – EGARCH Methodology

Another model that can explain the asymmetric effect in the market is the EGARCH model, and EGARCH stands for Exponential-GARCH model, and the EGARCH(1,1) is expressed as:

$$\ln(h_t) = \alpha_0 + \alpha_1 \left(\frac{\varepsilon_{t-1}}{\sqrt{h_{t-1}}}\right) + \lambda_1 \left|\frac{\varepsilon_{t-1}}{\sqrt{h_{t-1}}}\right| + \beta_1 \ln(h_{t-1}) \tag{7}$$

There are several merits reveal from the equation (7). Firstly, the model is capable to measure the asymmetric effect or leverage effect. As we can see from the equation, if the effect $\frac{\varepsilon_{t-1}}{\sqrt{h_{t-1}}}$ is negative, the effect of the shock will be $-\alpha_1 + \lambda_1$, and if the past shock $\frac{\varepsilon_{t-1}}{\sqrt{h_{t-1}}}$ is positive, the effect of the shock will be $\alpha_1 + \lambda_1$, which also mimic the effect of the threshold. However, unlike the TARCH model, coefficients in EGARCH model are not restricted to be nonnegative. Since the equation (7) takes the logarithm of h_t , it can never be negative. And also the EGARCH model uses the standardized residuals $s_{t-1} = \frac{\varepsilon_{t-1}}{\sqrt{h_{t-1}}}$, which gives a more natural interpretation of the shocks.

3.3.4. The Rescaling Method for Volatility Weighted Historical Simulation

Same as the HSAF approach, in order to estimate VaR under the volatility weighted historical simulation, we should first derive the above ARCH/GARCH family models from the data series, and then rescale the losses by the appropriate volatility models, and the rescaled loss series can be shown as:

$$l_T^* = \frac{\sqrt{h_{T+1}}}{\sqrt{h_T}} l_T \tag{8}$$

Where l_T is the loss from at time T from the original loss series and l_T^* is rescaled loss at time T from the rescaled loss series; $\sqrt{h_{T+1}}$ is the forecast volatility for the next holding period T+1, $\sqrt{h_T}$ is the conditional variance at time T. The scaling process is completed in the software MS Excel, in order to compute the conditional variance, we usually set the initial values $\varepsilon_0 = 0$ and $h_0 = sample\ variance\ of\ l_1, l_2, \ldots, l_T$.

This rescaled loss can simply be interpreted as the standardized loss multiplied by the forecast of tomorrow's volatility. The core concept in this volatility weighted historical simulation approach is to derive the forecast volatilities from different variance-covariance models. Finally, we can, based on this rescaled loss series, estimate VaR by finding the percentiles of the distribution.

3.4. The Backtesting Approach

This backtesting procedure uses the standard Kupiec frequency test to detect the stability of the VaR estimation models. The Kupiec test was presented by Kupiec in 1995, and this test aims to test whether the actual frequency of VaR violations deviates too much from the predicted frequency of violations[22]. The essence of the Kupiec test is a binomial test[23], which we take event as either violation or non-violation. In empirical testing, we code the violation as 1, and non-violation as 0. The cumulative probability of a binomial distribution is expressed as

$$\Pr(X \leq x) = \sum_{i=0}^{x} \binom{n}{i} p^i (1-p)^{n-i} \tag{9}$$

In the equation (9), p stands for the expected frequency of VaR violations that can be expressed as $p = 1 - \alpha$, and α is the certain confidence level is given by the equation (1); x

22 A VaR violation is when an actual loss exceeds corresponding VaR estimate.
23 See APPENDIX 6.3.1. Binomial Distribution.

indicates the actual number of VaR violations; n is the number of observations of subsamples in this paper.

The backtesting procedure in this paper applies the two-sided test at 95% confidence level, and the general idea is to compare if the actual number of VaR violations is too few or too many relative to the expected number of VaR violations at a certain confidence level, e.g. the standard confidence level is 95%. This approach consists of four steps:
First, we can compute the number of expected violations by $(1-\alpha)n$ as a baseline; Second, we should count the actual number of VaR violations x, and the actual number of violations x can either $x \geq (1-\alpha)n$ or $x \leq (1-\alpha)n$, since there exist both upper risk bound and lower risk bound in the two-sided backtesting process; Third, this step is to calculate $\Pr(X \geq x)$ or $\Pr(X \leq x)$, which depends on whether the actual number of VaR violations exceed the predicted number of VaR violations or not. Finally, we can compare the above calculated probability with the 95% confidence level.

There are mainly two ways to compare the actual frequency of violations with a certain confidence level in the two-sided test.

1. Calculate the probability of actual violation directly by applying the cumulative binomial distribution functions, and in this two-sided test we can calculate the upper and lower bound of the actual violations within 95% confidence level, and compare if the expected number violations falls within this confidence interval. In Excel, we can use the formula "BINOM.DIST()" to find the confidence interval or simply use "BINOM.INV()" to find upper and lower risk bound at 95% confidence level.
2. The alternative way is more preferable in this paper. Instead of calculating the confidence interval for each actual violation, we could choose to compute the confidence interval for expected violations and check if the actual number violations fall within this confidence interval. It can simply obtain from Excel formula "BINOM.INV()".

As we know from the next section, the data in this paper are divided into small subsamples, and each subsample contains approximately 260 observations. As mentioned above the upper and lower risk bound is computed in Excel by

$$x_{low} = BINOM.INV(n, p, 0.025)$$

$$x_{high} = BINOM.INV(n, p, 0.975)$$

Where $n = 260$ that is the size of the subsample; and $p = 1 - \alpha$ could be either 5% or 1% depends on different VaR's confidence level α; and since this is a two-sided test at 95% confidence level each side of the test will be 2.5%, then for the lower bound and upper bound we use 0.025 and 0.975 respectively.

The result of the backtesting is shown in Table 6.4, from the table we can see that the "acceptance interval" for $VaR_{95\%}$ with two-sided test at 95% confidence level and 260 observations sample size is $[7,20]$, and the boundary interval for $VaR_{99\%}$ with the same condition is [0,6], which is very natural that as $\alpha = 99\%$ is "wider" than $\alpha = 95\%$ and will generate less violations in general. Again, $(1 - \alpha)n$ is the expected number of violations, and there are 13 expected violations for $VaR_{95\%}$, and for $\alpha = 99\%$ it is 2.6 which can be rounded to 3. Based on the result from Table 6, we can compare the performance of different models by comparing their violation of the VaR estimates. Table 6.4 also illustrates the upper and lower bound for the entire sample at 95% and 99% confidence level, and this can give an overall inspection of the performance of VaR estimates.

4. EMPIRICAL ANALYSIS

4. 1. DATA

The empirical data applied in this paper are time series data and consists of three types of energy resources, which are natural gas, crude oil and propane. All of the three different types of energy data are from the data source THOMSON REUTERS. These resources are the main energy for daily production of the households and firms, which plays an important role in modern economics.

This paper uses weekly data of four data series from January 10, 1997 to December 27, 2013[24], and the sample contains 886 price observations for crude oil and petroleum products, and 885 price observations for natural gas[25]. In order to be consistent, the numbers of

[24] The four data series are natural gas, Brent crude oil, WTI crude oil and propane.
[25] Natural gas missed one observation on the day 30 September, 2005.

observations from four data series are set to be equal. The loss observation is generated by using the spot price of the previous period subtracts the spot price of the current period, and the size of the final loss observations of one loss distribution is 884 for three energy resources. The empirical testing uses the rolling window approach, and I divide the sample of loss into small subsamples. Each subsample is the composition of five years of in sample loss data and one year out of sample loss forecasting, and in total there are 12 in sample periods and 12 out-of-sample periods. As there are four data series, and each data series contains 12 subsamples, and in total there are 48 (12*4) subsamples to analyze, and each subsample contains approximately 260 observations to provide more accurate results.

As shown from the table 1, the first out of sample forecasting period is the year 2002, and the corresponding in sample period is from January 1997 to December 2001, and then the second out of sample forecasting period is the year 2003, and the matched in sample period is from January 1998 to December 2002 and so on. And for the last out of sample test, we use the data from January 2008 to December 2012 to forecast the loss occurred in 2013. The reason to use this forecasting approach is obvious. On the one hand, for this long span of sample period, it is more precise to divide the sample into small subsamples, since subsample approach allows the estimates respond quicker to recent losses. On the other hand, as this approach views specific sample periods for both calm period and volatile periods, e.g. we can generate a comparative analysis between the global financial crisis years and other calm years.

After dividing the sample into different subsamples, we can then estimate the value-at-risk at 95% and 99% for each out of sample time period, and finally compare the performance of different models by applying the Kupiec backtesting procedure.

4.2. HISTORICAL SIMULATION

4.2.1. Standard Historical Simulation Approach

As introduced in section 3.2.1, the basic historical simulation lets the data speak for themselves, and based on the definition of VaR we can estimate VaR by computing the percentile of the subsamples. Table 6 is a summary table that indicates the number of VaR

violations at 95% and 99% significance level, and VaR violation can be defined as the case when the actual loss exceed the estimated loss. In other words, Table 6.1 is a performance indicator for standard historical simulation estimations. From Table 6.1, it seems that the VaR estimates show higher violations in crude oil resources, while natural gas has least number of violations under historical simulation. To be more accurate, we need to compare the results with the boundary interval derived from the backtesting approach. By comparison with Table 6.4, we can see that the $VaR_{95\%}$ under historical simulation generally perform poorly, and we observe too few observations for all the energy resources. For example, almost all the estimates are smaller than lower boundary of the Kupiec frequency test. However, it seems like historical simulation model obtains a satisfactory $VaR_{99\%}$ estimation, e.g. only "reject" once in the forecast year 2008 for crude oil (both WTI and Brent) and propane, and it performs quite well after that. The year 2008 is within the global financial crisis period, and it seems like an acceptable result that we get too many violations for only one year. We should thus conclude that the standard historical simulation based $VaR_{95\%}$ is not capable to explain risk occurred in the energy market, but $VaR_{99\%}$estimates provide reliable results for energy resources I studied in this paper.

4.2.2. Historical Simulation with ARMA Forecasting Approach

As mentioned above (section 3.2.2), the HSAF approach can be divided into two parts. The first step of this unit root tests is to use ADF Dickey Fuller test with the trend and intercept, and under the null hypothesis the data series contains unit roots and follows Dickey Fuller distribution. The result of unit roots testing is shown in table 2, and from the table we can see that the probability of the test statistic is statistically significant at 5% significance level of all the energy commodities, and we can thus reject the null hypothesis and conclude that there is no unit root and all the loss series are stationary. In this case, there is no need to continue further tests. It is an advantage to use loss series rather than the price series, price might have a tendency to go up or fluctuate depending on the time. Since the energy prices are affected by many factors such as the consumption from the market, inflation, wars and etc., which means the price series are not stationary (tested).

After applying general unit roots testing of the data series, we can then use Box-Jenkins model selection approach to determine the most appropriate model. Firstly, we can plot the data of all tree energy resources, which are shown from Figure 1 to Figure 3. The graphs indicate that all the data series are suspected to have conditional heteroskedasticity, especially during the period of financial crisis, and the crude oil shows unusual high volatility. Meanwhile, we can see that there is no obvious evidence of the deterministic trend, which is consistent with the above section of Dickey Fuller tests. This can only give rough information of the data series property, and in order to select the most appropriate ARMA model, we need to continue further tests by employing Box-Jenkins model selection criterion.

The Box-Jenkins model selection method is introduced in section 3.2.3, and it can be summarized and adapted to five main criteria in this empirical study. The five criteria are namely: the correlogram of the data series; the significance of F-statistic of the joint regression coefficients, the p-value of coefficients; the AIC/BIC of the estimated models and residual autocorrelation of the selected model. Based on these five criteria, I can then select adequate models for four data series. The empirical study for HSAF approach is aimed to select one or two most preferred models from a set of eight candidate models[26], namely AR(1), MA(1), AR(2), MA(2), ARMA(1,1), ARMA(1,2), ARMA(2,1), ARMA(2,2).

As mentioned above, the first step is to detect the independence of the series, and correlogram is graphed in this step. The correlogram depicts the ARMA(p,q) process by plotting the autocorrelation function (ACF) and partial autocorrelation function (PACF), and we can get a rough indication of possible process from the ACF and PACF. For instance, the PACF of AR(p) process is close to 0 after the pth lag, and the ACF of MA(q) process is close to 0 after the qth lag[27]. And the second step is to check the overall significance of the model, and the value of the F-statistic is compared with its critical value at 5% significance level. After that, we can then inspect the p-value of coefficients, e.g. whether they are significant at 5% significance level. Akaike information Criterion and Schwarz Bayesian Information Criterion (BIC) are two commonly used criteria in various model selection processes, and since the sample size of each loss series is 884, the BIC criterion is considered to be more reliable in

[26] It has been illustrated in Table 3 that two ARMA process models are selected to rescale the loss series of the energy resources, which are AR(1) and MA(1).

[27] More information related ACF and PACF are introduced in APPENDIX 6.4.

this case, and the smaller the BIC value, the more parsimonious the estimated models. Finally, it is important to test whether the residual follows the white noise process.

After implementing all the above steps, we should be able to choose the most appropriate models. The results of ARMA model selection process for four loss series are shown from Table 3.1 to Table 3.4. As we can see these tables give ordinal rankings of the estimated models, and these rankings mainly depend on five model selection criterion mentioned above. According to the results of the tables, we observe that AR(1) and MA(1) are the two most preferred models for the four data series, and they are almost equivalently preferable than other models. We can thus employ these two models in the VaR estimation process, and the historical simulation with AR(1)/MA(1) forecasting process can be denoted as HSAF-AR(1)/MA(1). Additionally, both of the unit root test and Box-Jenkins model selection process are implemented by using the software EViews.

Based on the selected ARMA models from Table 3, I can then rescale the loss distribution by applying these models and derive the VaR estimates. And the results of violations of VaR estimates for three energy commodities are shown from Table 6.2.1 to Table 6.2.4. These four tables depict the violations of VaR estimates by using the HSAF-AR(1)/MA(1) approach, the number of violations of $VaR_{99\%}$ estimates for almost all resources are within the boundary interval, which means the $VaR_{99\%}$ of HSAF-AR(1)/MA(1) approach is an adequate estimation to quantify the risk in the energy market. The tables show that the violations only exceed the upper boundary for crude oil (Brent and Crude) and propane in the forecasting year 2008. However, the $VaR_{95\%}$ estimates are not satisfactory, and we should thus reject the HSAF model for $VaR_{95\%}$ estimates. Interestingly, HS approach has one more qualified $VaR_{95\%}$ estimate than HSAF-AR(1)/MA(1) approach in Brent crude oil, and has one less qualified $VaR_{95\%}$ estimated than HSAF approach in the WTI crude oil, which make no big difference in comparing two approaches. From an overall perspective, the VaR estimates of HSAF-AR(1)/MA(1) approach are slightly better than VaR estimates of HS approach for the loss series of natural gas, which indicates the HSAF approach is more preferable even it is only slightly better.

4.3. VOLITILITY WEIGHTED HISTORICAL SIMULATION APPROACH

As mentioned in the HSAF section above, we can see the plotted graphs of the three energy commodities exhibit volatility clustering phenomenon from Figure 1 to Figure 3. Additionally, there is an interesting fact that crude oil (Bren and WTI) and propane show volatility clustering around the 2008 global financial crisis period, but natural gas shows volatility clustering before the global financial crisis period. However, graphical observations cannot be regarded as a strong evidence of the existence of the ARCH effect. In order to make more careful testing, I need to take the heteroskedasticity test, and the heteroskedasticity test is based on estimated mean equations. The reason is that ARCH/GARCH family models consist of two parts, which are the mean equation and the variance equation. And to detect the ARCH effect, I can estimate the mean models first, and then take the heteroskedasticity test of the mean models. The mean equations of all the four loss series are AR(1)[28]. If the results support the existence of the ARCH effect, I will then continue to estimate ARCH/GARCH family models. Otherwise, it makes no sense to continue further estimations. The results of the ARCH effect test can be seen from Table 4, all the p-value of chi-square statistics are statistically significant at the 5% significance level, which means we can reject the null hypothesis of no ARCH effect, and conclude that all the loss series of energy commodities exhibit ARCH effect. Hence, it seems necessary to estimate the ARCH/GARCH family models for our loss series on energy commodities.

The empirical study in this section test 16 models simultaneously, which are ARCH(1), ARCH(2), GARCH(1,1), GARCH(1,2), GARCH(2,1), GARCH(2,2), TARCH(1,1) and EGARCH(2,2), and all of them are estimated by both normal distribution and t-distribution. These models are introduced in the section 3.3, which ordered from the standard conditional heteroskedasticity models to asymmetric effect capturing models. Since t-distribution shows strong explanatory power in the stock returns of financial market, and it is meaningful to test whether the t-distribution can generate better models of the loss series of the energy commodities[29]. To employ the ARCH/GARCH family models into the historical simulation

---

[28] AR(1) model is estimated and selected from the historical simulation with ARMA process from the section 4.2.2. Historical Simulation with ARMA Forecasting Approach.

[29] More information of t-distribution is shown in section 6.3.3. Student t-distribution.

process, I need to estimate these variance-covariance models first, and then I select the most adequate models and plug into VWHS process.

Same with the ARMA model selection process in the previous section, the ARCH/GARCH model selection procedure consists of several steps. There are basically five criteria to choose the most appropriate ARCH/GARCH family models, which are the p-value of estimated coefficients, the maximum likelihood value of the estimated models, the AIC/BIC model selection criterion, the diagnostic test of the residual autocorrelation and the remaining ARCH/GARCH effect. The heteroskedasticity test and model selection stage are processed in software EViews 8.1. The results of model selection are shown in Table 5, from the results we can see the t-distributed TARCH(1,1) model has the best explanatory power for natural gas; and t-distributed EGARCH(1,1) model is the most adequate model for crude oil (Brent and WTI) and propane series. I can then denote these two selected VWHS approaches as VWHS-TARCH(1,1) and VWHS-EGARCH(1,1) in the later analysis. It seems that the risks of energy commodities are also affected by market information asymmetrically, since the threshold GARCH and exponential GARCH perform better based on the model selection results. And according to the test results, the t-distributed estimation models are generally more parsimonious and generate better fitness.

In order to check the reliability of the above results and make detailed analysis, I will apply the estimated models to the real world case. The empirical study can be done by employing the estimated models to rescale losses and estimate VaR at the certain percentile, and then use the Kupiec backtesting approach to test whether the selected model really explains the risk of energy resources. The results of violations of VWHS approach based VaR estimates are shown from Table 6.3.1 to Table 6.3.4, and each table represents the performance of VaR estimates for one energy commodity. To analyze the performance of each variance-covariance model's VaR estimates, we also need Table 6.4 as a "reference table" that is introduced in section 3.4.

As we can see from the Table 6.3.1, the $VaR_{0.99}$ estimates of VWHS-TARCH(1,1) provide reliable estimates for natural gas, and the table shows the number of violations of $VaR_{0.99}$ estimates are all within the backtesting confidence interval at the 95% significance level, which generates the same results with the standard HS and HSAF approach. However, the

TARCH based historical simulation exhibit weak $VaR_{0.95}$ estimations, only the number of violations of the forecasting year 2008 is qualified, which may indicate $VaR_{0.95}$ estimates of VWHS-TARCH(1,1) are generally not able to explain the risk occurred in the natural gas market, only can explain the risk during the financial crisis period in a small probability. Whereas the overall performance of the sample is not as good as the HSAF or the standard HS approach.

The number of violations of VAR estimates of VWHS-EGARCH(1,1) for Brent crude oil is depicted in the Table 6.3.2. The table shows that $VaR_{0.95}$ estimates of VWHS-EGARCH(1,1) method underestimate the number of violations in most forecasting years except the year of 2008, and standard HS in fact performs slightly better in the forecasting year 2006. It is less useful to apply $VaR_{0.95}$ estimates of VWHS approach for the Brent crude oil. It is worth to notice that $VaR_{0.99}$ estimates of VWHS-EGARCH(1,1) approach exhibit the best performance among all the three non-parametric approaches, which qualified for all the calm and volatile periods.

The result of VWHS based VaR estimation for WTI crude oil is shown in Table 6.3.3, from the table we can see that there is no difference of the performance of VaR estimation between the VWHS approach and the HS approach in WTI crude oil. The $VaR_{0.95}$ estimates of HSAF approach perform slightly better than other approaches. The violations of $VaR_{0.99}$ estimates of VWHS-EGARCH(1,1) generally do not deviate too much from the expected violations, but it cannot capture the risk for the financial crisis year 2008. Table 6.3.4 illustrates the performance of the EGARCH(1,1) weighted VaR estimates for propane. According to the Table 6.3.4, the VWHS based VaR estimation for propane is not as good as the standard HS or the HSAF approach, since from an overall perspective $VaR_{99\%}$ estimation violates too much from the expected violations and $VaR_{95\%}$ estimation gives the same result as the HS and the HSAF approach.

Based on the model selection process and VaR estimation results, it seems that all the student t-distributed asymmetric variance-covariance models are more suitable to explain the risk for different energy commodities. For instance, the t-distributed VWHS-TARCH(1,1) is the most suitable model for the risk estimation of natural gas, and the t-distributed VWHS-EGARCH(1,1) gives the best explanation of risk occurred in the Brent crude oil market.

5. CONCLUSION

What does this paper study? This paper studies three representative energy resources in the energy market by employing the VaR estimation technique, i.e. whether VaR estimates can explain the risk occurred in the energy market. In order to study the performance of VaR estimates, this paper applies different non-parametric approaches, which are the standard historical simulation (HS), the historical simulation with ARMA forecasting (HSAF) and the volatility weighted historical simulation (VWHS).

How does the study implement? The main idea of this paper is to combine data selected models with the VaR estimation technique, and then use the backtesting approach to analyze the performance of the VaR estimates. To be more specific, I select the most adequate models from all the candidate models by employing the model selection criteria for the HSAF approach and VWHS approach. The selected models are thus based on the data of the loss series, and in order to test whether the selected models work in the real word, I then apply the selected model to VaR estimation procedure. The reason is that data estimated models may not be strictly correct, and to implement the VaR estimation can in turn strengthen the explanatory power of the estimated model. And the performance results are tested by the backtesting process.

What are the results of the study? The results of study are shown in Table 6, we can obviously detect that $VaR_{95\%}$ estimates are generally not sufficient to estimate the risk in the energy market. On the contrary, $VaR_{99\%}$ estimates can explain the risk for three energy commodities in almost every out-of-sample forecasting years; except for the financial crisis year 2008 of the WTI crude oil and Propane. Hence, it is interesting and meaningful to study the confidence level in between for the future empirical study in the energy market, e.g. what is the performance of $VaR_{97.5\%}$ estimates in the energy market.

Result 1.1: The $\boldsymbol{VaR_{95\%}}$ estimates are generally not capable to explain risks of the four energy resources, for any non-parametric historical simulation approaches applied in this paper.

Result 1.2: The $VaR_{99\%}$ ***estimates are generally capable to explain risks of the four energy resources but not for the financial crisis year 2008, for any non-parametric historical simulation approaches applied in this paper.***

In light of the empirical study of this paper, I found that the $VaR_{99\%}$ estimates of all three estimation approaches can explain the risk happened in the energy market, no matter for clam periods or the volatile periods such as financial crisis periods. Moreover, the number of violations of $VaR_{99\%}$ estimates of VWHS-EGARCH(1,1) are all within the "acceptance interval" for all the subsample periods[30]. It is worth to mention that all the variance-covariance models of VWHS approach are proved to be more suitable under the student t-distribution. We can then denote the two preferred $VaR_{99\%}$ estimates of VWHS models as VWHS-fat tailed-TARCH(1,1) and VWHS-fat tailed-EGARCH(1,1). Meanwhile, $VaR_{99\%}$ estimates of HSAF-AR(1)/MA(1) give the best performance in estimating the risk of the WTI Crude Oil and Propane, and the violation of $VaR_{99\%}$ only exceed once in the forecasting year 2008. As 2008 is the peak year of the global financial, it is hard to capture the risk happened in the market. We can get the *Result 2* based on the above statements:

Result 2.1: The $VaR_{99\%}$ ***estimates of VWHS-fat tailed-TARCH(1,1) are capable to explain risks of the natural gas for all the subsample periods, e.g. both calm periods and volatile periods.***

Result 2.2: The $VaR_{99\%}$ ***estimates of VWHS-fat tailed-EGARCH(1,1) are capable to explain risks of the Brent crude oil for all the subsample periods, e.g. both calm periods and volatile periods.***

By comparing the violation results of VaR estimates across different non-parametric approaches, we know the VWHS approach and the HSAF approach perform better than the standard HS, but only at an insignificant level. Again, the only distinct point is the *Result 2* that $VaR_{99\%}$ estimates of VWHS are considered as a reliable risk measurement for natural and Brent crude oil.

[30] "Acceptance interval" is discussed in the section 3.4. Backtesting Approach.

Result 3: The HSAF approach and VWHS approach do generate better results of VaR estimation, but only in an insignificant way.

There are some areas left out in this empirical research, and it might be interesting if these problems can be studied in future research. Firstly, based on *Result 1* it is meaningful to study the performance of VaR estimates at different confidence level α; For instance, to find the critical confidence level of α^* that makes $VaR_{1-\alpha^*}$ give acceptable estimation result where $\alpha^* \in (0.01,0.05)$. Secondly, *Result 2* indicates that $VaR_{99\%}$ estimates provide a reliable risk measurement for both calm periods and volatile periods, it is interesting to investigate whether this result holds by using other backtesting approaches or other time span for the same energy commodities I studied. Finally, it is interesting to investigate which estimation methods can give significantly better $VaR_{95\%}$ estimation, the parametric approaches or the non-parametric approaches. To evaluate the performance of $VaR_{95\%}$ estimates from different approaches, we can then set the $VaR_{95\%}$ estimates of standard HS as a benchmark.

6. APPENDIX

6.1. TABLES

Table 1: List of Testing Samples for Empirical Studies

	In-sample period				Out-of-sample period
1997	1998	1999	2000	2001	2002
1998	1999	2000	2001	2002	2003
1999	2000	2001	2002	2003	2004
2000	2001	2002	2003	2004	2005
2001	2002	2003	2004	2005	2006
2002	2003	2004	2005	2006	2007
2003	2004	2005	2006	2007	2008
2004	2005	2006	2007	2008	2009
2005	2006	2007	2008	2009	2010
2006	2007	2008	2009	2010	2011
2007	2008	2009	2010	2011	2012
2008	2009	2010	2011	2012	2013

Each row stands for one testing period that consists of 5 years in-sample and 1 year out-of-sample period, and there are in total 12 testing periods.

Table 2: Augmented Dickey-Fuller Test for the Loss Series

Table 2.1: ADF Test for the Loss Series of Natural Gas

Natural Gas		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-27.13678	0.0000
Test critical values:	1% level	-3.437533	
	5% level	-2.864600	
	10% level	-2.568453	
***MacKinnon (1996) one-sided p-values.**			

Table 2.2: ADF Test for the Loss Series of Brent Crude Oil

Brent Crude Oil		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-23.67996	0.0000
Test critical values:	1% level	-3.437533	
	5% level	-2.864600	
	10% level	-2.568453	

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

Table 2.3: ADF Test for the Loss Series of WTI Crude Oil

WTI Crude Oil		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-24.76207	0.0000
Test critical values:	1% level	-3.437533	
	5% level	-2.864600	
	10% level	-2.568453	
***MacKinnon (1996) one-sided p-values.**			

Table 2.4: ADF Test for the Loss Series of Propane

Propane		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-11.99743	0.0000
Test critical values:	1% level	-3.437558	
	5% level	-2.864611	
	10% level	-2.568459	
***MacKinnon (1996) one-sided p-values.**			

Table 3: Historical Simulation ARMA Model Selection Process

Table 3.1: Historical Simulation ARMA Model Selection Process for Natural Gas

Natural Gas	Correlogram	F-statistics	P-value of coefficients	AIC/BIC	Residual Autocorrelation	Rank
AR(1)	significant	7.153481	significant	1.300808	no serial correlation exist up to 4th lag	1
MA(1)	significant	7.394404	significant	1.300246	no serial correlation exist up to 4th lag	1
AR(2)	insignificant	3.996748	insignificant	1.300869	no serial correlation exist up to 4th lag	5
MA(2)	insignificant	3.746721	insignificant	1.307800	no serial correlation exist up to 4th lag	7
ARMA(1,1)	significant	5.506311	significant	1.304140	no serial correlation exist up to 4th lag	3
ARMA(1,2)	insignificant	4.218131	insignificant	1.309965	no serial correlation exist up to 4th lag	8

ARMA(2,1)	insignificant	5.971897	significant	1.297411	no serial correlation exist up to 4th lag	4
ARMA(2,2)	insignificant	4.475278	insignificant	1.305094	no serial correlation exist up to 4th lag	6

Table 3.2: Historical Simulation ARMA Model Selection Process for Brent Crude Oil

Brent Crude Oil	Correlogram	F-statistics	P-value of coefficients	AIC/BIC	Residual Autocorrelation	Rank
AR(1)	significant	45.76670	significant	4.527739	no serial correlation exist up to 7th lag	1
MA(1)	significant	42.61502	significant	4.530529	no serial correlation exist up to 7th lag	2
AR(2)	insignificant	22.95379	insignificant	4.536275	no serial correlation exist up to 12th lag	6
MA(2)	insignificant	22.69543	insignificant	4.535151	no serial correlation exist up to 7th lag	6
ARMA(1,1)	significant	23.01905	insignificant	4.535072	no serial correlation exist up to 7th lag	5
ARMA(1,2)	insignificant	16.47831	significant	4.539032	no serial correlation exist up to 12th lag	3
ARMA(2,1)	insignificant	16.35784	significant	4.540487	no serial correlation exist up to 12th lag	4
ARMA(2,2)	insignificant	12.34243	insignificant	4.547796	no serial correlation exist up to 12th lag	8

Table 3.3: Historical Simulation ARMA Model Selection Process for WTI Crude Oil

WTI Crude Oil	Correlogram	F-statistics	P-value of coefficients	AIC/BIC	Residual Autocorrelation	Rank
AR(1)	significant	29.22777	significant	4.578697	no serial correlation exist up to 7th lag	1
MA(1)	significant	28.73481	significant	4.578309	no serial correlation exist up to 7th lag	1
AR(2)	insignificant	14.55226	insignificant	4.587360	no serial correlation exist up to 7th lag	6
MA(2)	insignificant	14.49446	insignificant	4.585669	no serial correlation exist up to 7th lag	6
ARMA(1,1)	significant	14.59913	insignificant	4.586376	no serial correlation exist up to 7th lag	5
ARMA(1,2)	insignificant	11.63137	significant	4.587769	no serial correlation	3

					exist up to 7th lag	
ARMA(2,1)	insignificant	11.20316	significant	4.590059	no serial correlation exist up to 7th lag	4
ARMA(2,2)	insignificant	8.886975	insignificant	4.595580	no serial correlation exist up to 7th lag	8

Table 3.4: Historical Simulation ARMA Model Selection Process for Propane

Propane	Correlogram	F-statistics	P-value of coefficients	AIC/BIC	Residual Autocorrelation	Rank
AR(1)	significant	10.19579	significant	-3.540380	no serial correlation exist up to 3rd lag	1
MA(1)	significant	11.08280	significant	-3.540845	no serial correlation exist up to 3rd lag	1
AR(2)	insignificant	5.595307	insignificant	-3.534037	no serial correlation exist up to 3rd lag	4
MA(2)	insignificant	5.994067	insignificant	-3.534199	no serial correlation exist up to 3rd lag	4
ARMA(1,1)	significant	5.737377	insignificant	-3.534147	no serial correlation exist up to 3rd lag	3
ARMA(1,2)	insignificant	3.919534	insignificant	-3.526798	no serial correlation exist up to 3rd lag	4
ARMA(2,1)	insignificant	3.874665	insignificant	-3.526849	no serial correlation exist up to 3rd lag	4
ARMA(2,2)	insignificant	3.865497	insignificant	-3.523484	no serial correlation exist up to 3rd lag	4

The indication of correlalogram is based on spikes of ACF and PACF;

F-statistics is compared with 5% significance level;

P-value of coefficients is tested at 5% significance level;

The columns under AIC/BIC are mainly use BIC as an indicator;

Residual correlations are based on the Q-statistic at the 5% significance level;

Same rank of the models indicates two models are equivalent.

Table 4: Heteroskedasticity Test for the Loss Series

Table 4.1: Heteroskedasticity Test for Natural Gas

Natural Gas	Heteroskedasticity Test: ARCH		
F-statistic	211.9842	Prob. F(1,880)	0.0000
Obs*R-squared	171.2205	Prob. Chi-Square(1)	0.0000

Table 4.2: Heteroskedasticity Test for Brent Crude Oil

Brent Crude Oil	Heteroskedasticity Test: ARCH		
F-statistic	165.7363	Prob. F(1,880)	0.0000
Obs*R-squared	139.7861	Prob. Chi-Square(1)	0.0000

Table 4.3: Heteroskedasticity Test for WTI Crude Oil

WTI Crude Oil	Heteroskedasticity Test: ARCH		
F-statistic	205.2722	Prob. F(1,880)	0.0000
Obs*R-squared	166.8246	Prob. Chi-Square(1)	0.0000

Table 4.4: Heteroskedasticity Test for Propane

Propane	Heteroskedasticity Test: ARCH		
F-statistic	245.5218	Prob. F(1,880)	0.0000
Obs*R-squared	192.3999	Prob. Chi-Square(1)	0.0000

Table 5: ARCH/GARCH Family Model Selection for VWHS Approach

Table 5.1: VWHS-ARCH/GARCH Family Model Selection for Natural Gas

Natural Gas	p-value of coefficients	ML	AIC/BIC	Q1	Q2	selected model
Normal distribution	All are significant	TARCH(1,1)	TARCH(1,1)	passed	passed	TARCH(1,1) (N-dist)
Student t-distribution	All are significant except GARCH(2,1) and GARCH(2,2)	TARCH(1,1)	TARCH(1,1)	passed	passed	TARCH(1,1) (t-dist)
						TARCH(1,1) (t-dist)

Table 5.2: VWHS-ARCH/GARCH Family Model Selection for Brent Crude Oil

Brent Crude Oil	p-value of coefficients	ML	AIC/BIC	Q1	Q2	Selected model
Normal distribution	ARCH(1); ARCH(2); GARCH(1,1); EGARCH(1,1)	EGARCH(1,1)	EGARCH(1,1)	passed	passed	EGARCH(1,1) (N-dist)
Student t-distribution	ARCH(1); ARCH(2); GARCH(1,1); EGARCH(1,1)	EGARCH(1,1)	EGARCH(1,1)	passed	passed	EGARCH(1,1) (t-dist)
						EGARCH(1,1) (t-dist)

Table 5.3: VWHS-ARCH/GARCH Family Model Selection for WTI Crude Oil

WTI Crude Oil	p-value of coefficients	ML	AIC/BIC	Q1	Q2	selected model
Normal	All are significant except	EGARCH(1,1)	EGARCH(1,1)	passed	passed	EGARCH(1,1) (N-dist)

distribution	GARCH(2,1), GARCH(2,2) and TARCH(1,1)					
Student t-distribution	All are significant except GARCH(2,1), GARCH(2,2) and TARCH(1,1)	EGARCH(1,1)	EGARCH(1,1)	passed	passed	EGARCH(1,1) (t-dist)
						EGARCH(1,1) (t-dist)

Table 5.4: VWHS-ARCH/GARCH Family Model Selection for Propane

Propane	**p-value of coefficients**	**ML**	**AIC/BIC**	**Q1**	**Q2**	**selected model**
Normal distribution	All are significant	EGARCH(1,1)	EGARCH(1,1)	passed	passed	EGARCH(1,1) (N-dist)
Student t-distribution	All are significant except GARCH(1,2) and GARCH(2,1)	EGARCH(1,1)	EGARCH(1,1)	passed	passed	EGARCH(1,1) (t-dist)
						EGARCH(1,1) (t-dist)

Columns under "p-value of coefficients" indicate the models with significant p-value of coefficients;
Columns under "ML" are the models with the highest maximum likelihood value;
Columns under "AIC/BIC" show models with the lowest BIC/AIC (BIC) value;
"Passed" under column Q1 and Q2 indicate that the selected models (based on previous model selection criteria) passed the serial correlation and remaining ARCH effect diagnostic test.

Table 6: Summary Table of the Violations of VaR Estimates

Table 6.1: Violations of VaR Estimates Based on Standard HS Approach

		Henry hub natural gas		**Crude Oil-WTI**		**Crude Oil-Brent**		**Propane**	
In-sample Period	Out-of-sample	95%	99%	95%	99%	95%	99%	95%	99%
1997/1/10-2001/12/28	Year 2002	1	0	4	0	1	0	0	0
1998/1/2-2002/12/27	Year 2003	6	2	4	1	4	1	4	1
1999/1/1-2003/12/26	Year 2004	5	0	6	0	5	1	5	1
2000/1/7-2004/12/31	Year 2005	4	2	5	0	5	1	3	0
2001/1/5-2005/12/30	Year 2006	4	0	4	0	8	1	3	0
2002/1/4-2006/12/29	Year 2007	2	0	6	2	4	1	1	0
2003/1/3-2007/12/28	Year 2008	4	0	18	8	17	8	13	8
2004/1/2-2008/12/26	Year 2009	1	0	2	0	1	0	2	0
2005/1/7-2009/12/25	Year 2010	1	0	1	0	1	0	2	1
2006/1/6-2010/12/31	Year 2011	0	0	4	0	3	0	1	0
2007/1/5-2011/12/30	Year 2012	0	0	1	0	0	0	3	0
2008/1/4-2012/12/28	Year 2013	0	0	0	0	0	0	0	0

	28	4	55	11	49	13	37	11

HSAF-AR(1)/MA(1) indicates historical simulation with AR(1)/MA(1) forecasting process.

Table 6.2.1: VaR ***Violations from HSAF Approach - Natural Gas***

Henry hub natural gas		**Standard HS**		**HSAF-AR(1)**		**HSAF-MA(1)**	
In-sample Period	Out-of-sample	95%	99%	95%	99%	95%	99%
1997/1/10-2001/12/28	year 2002	1	0	3	0	3	0
1998/1/2-2002/12/27	year 2003	6	2	5	2	5	2
1999/1/1-2003/12/26	year 2004	5	0	6	0	6	0
2000/1/7-2004/12/31	year 2005	4	2	5	2	5	2
2001/1/5-2005/12/30	year 2006	4	0	5	0	5	0
2002/1/4-2006/12/29	year 2007	2	0	3	0	3	0
2003/1/3-2007/12/28	year 2008	4	0	3	0	3	0
2004/1/2-2008/12/26	year 2009	1	0	1	0	1	0
2005/1/7-2009/12/25	year 2010	1	0	1	0	1	0
2006/1/6-2010/12/31	year 2011	0	0	0	0	0	0
2007/1/5-2011/12/30	year 2012	0	0	0	0	0	0
2008/1/4-2012/12/28	year 2013	0	0	0	0	0	0
		28	4	32	4	32	4

HSAF-AR(1)/MA(1) indicates historical simulation with AR(1)/MA(1) forecasting process.

Table 6.2.2: VaR ***Violations from HSAF Approach – Brent Crude Oil***

Brent Crude Oil		**Standard HS**		**HSAF-AR(1)**		**HSAF-MA(1)**	
In-sample Period	Out-of-sample	95%	99%	95%	99%	95%	99%
1997/1/10-2001/12/28	year 2002	1	0	3	0	3	0
1998/1/2-2002/12/27	year 2003	4	1	4	1	4	1
1999/1/1-2003/12/26	year 2004	5	1	5	0	5	0
2000/1/7-2004/12/31	year 2005	5	1	6	0	6	0
2001/1/5-2005/12/30	year 2006	8	1	4	0	4	0
2002/1/4-2006/12/29	year 2007	4	1	3	2	3	2
2003/1/3-2007/12/28	year 2008	17	8	18	7	19	7
2004/1/2-2008/12/26	year 2009	1	0	3	1	3	1
2005/1/7-2009/12/25	year 2010	1	0	0	0	0	0
2006/1/6-2010/12/31	year 2011	3	0	3	0	3	0
2007/1/5-2011/12/30	year 2012	0	0	1	0	1	0
2008/1/4-2012/12/28	year 2013	0	0	0	0	0	0
		49	13	50	11	51	11

HSAF-AR(1)/MA(1) indicates historical simulation with AR(1)/MA(1) forecasting process.

Table 6.2.3: VaR ***Violations from HSAF Approach – WTI Crude Oil***

WTI Crude Oil		Standard HS		HSAF-AR(1)		HSAF-MA(1)	
In-sample Period	Out-of-sample	95%	99%	95%	99%	95%	99%
1997/1/10-2001/12/28	year 2002	4	0	4	0	4	0
1998/1/2-2002/12/27	year 2003	4	1	4	1	4	1
1999/1/1-2003/12/26	year 2004	6	0	4	1	4	1
2000/1/7-2004/12/31	year 2005	5	0	5	1	5	1
2001/1/5-2005/12/30	year 2006	4	0	7	1	7	1
2002/1/4-2006/12/29	year 2007	6	2	5	1	5	1
2003/1/3-2007/12/28	year 2008	18	8	17	8	17	8
2004/1/2-2008/12/26	year 2009	2	0	2	0	2	0
2005/1/7-2009/12/25	year 2010	1	0	2	0	2	0
2006/1/6-2010/12/31	year 2011	4	0	3	0	3	0
2007/1/5-2011/12/30	year 2012	1	0	0	0	0	0
2008/1/4-2012/12/28	year 2013	0	0	0	0	0	0
		55	11	53	13	53	13

HSAF-AR(1)/MA(1) indicates historical simulation with AR(1)/MA(1) forecasting process.

Table 6.2.4: VaR ***Violations from HSAF Approach – Propane***

Propane		Standard HS		HSAF-AR(1)		HSAF-MA(1)	
In-sample Period	Out-of-sample	95%	99%	95%	99%	95%	99%
1997/1/10-2001/12/28	year 2002	0	0	1	0	2	0
1998/1/2-2002/12/27	year 2003	4	1	4	1	4	1
1999/1/1-2003/12/26	year 2004	5	1	5	1	5	1
2000/1/7-2004/12/31	year 2005	3	0	4	0	4	0
2001/1/5-2005/12/30	year 2006	3	0	3	0	3	0
2002/1/4-2006/12/29	year 2007	1	0	1	0	2	0
2003/1/3-2007/12/28	year 2008	13	8	13	8	13	8
2004/1/2-2008/12/26	year 2009	2	0	2	0	2	0
2005/1/7-2009/12/25	year 2010	2	1	1	1	1	1
2006/1/6-2010/12/31	year 2011	1	0	2	0	2	0
2007/1/5-2011/12/30	year 2012	3	0	2	0	2	0
2008/1/4-2012/12/28	year 2013	0	0	0	0	0	0
		37	11	38	11	40	11

HSAF-AR(1)/MA(1) indicates historical simulation with AR(1)/MA(1) forecasting process.

Table 6.3.1: VaR ***Violations from VWHS Approach - Natural Gas***

Henry hub natural gas		Standard HS		VWHS-TARCH(1,1) t-distribution	
In sample Period	Out of sample	95%	99%	95%	99%
1997/1/10-2001/12/28	year 2002	1	0	3	0
1998/1/2-2002/12/27	year 2003	6	2	4	3
1999/1/1-2003/12/26	year 2004	5	0	0	0
2000/1/7-2004/12/31	year 2005	4	2	4	2
2001/1/5-2005/12/30	year 2006	4	0	0	0
2002/1/4-2006/12/29	year 2007	2	0	2	1
2003/1/3-2007/12/28	year 2008	4	0	10	5
2004/1/2-2008/12/26	year 2009	1	0	3	1
2005/1/7-2009/12/25	year 2010	1	0	0	0
2006/1/6-2010/12/31	year 2011	0	0	0	0
2007/1/5-2011/12/30	year 2012	0	0	1	0
2008/1/4-2012/12/28	year 2013	0	0	2	1
		28	4	29	13

VWHS-TARCH(1,1) t-distribution indicates VaR estimates are estimated based on the TARCH(1,1) weight loss distribution, and TARCH(1,1) is estimated under t-distribution.

Table 6.3.2: VaR ***Violations from VWHS Approach – Brent Crude Oil***

Brent Crude Oil		Standard HS		VWHS-EGARCH(1,1) t-distribution	
In-sample Period	Out-of-sample	95%	99%	95%	99%
1997/1/10-2001/12/28	year 2002	1	0	4	0
1998/1/2-2002/12/27	year 2003	4	1	4	1
1999/1/1-2003/12/26	year 2004	5	1	4	1
2000/1/7-2004/12/31	year 2005	5	1	3	0
2001/1/5-2005/12/30	year 2006	8	1	4	0
2002/1/4-2006/12/29	year 2007	4	1	4	2
2003/1/3-2007/12/28	year 2008	17	8	13	6
2004/1/2-2008/12/26	year 2009	1	0	2	0
2005/1/7-2009/12/25	year 2010	1	0	1	0
2006/1/6-2010/12/31	year 2011	3	0	5	4
2007/1/5-2011/12/30	year 2012	0	0	1	0
2008/1/4-2012/12/28	year 2013	0	0	2	1
		49	13	47	15

VWHS-EGARCH(1,1) t-distribution indicates VaR estimates are estimated based on the EGARCH(1,1) weight loss distribution, and EGARCH(1,1) is estimated under t-distribution.

Table 6.3.3: VaR ***Violations from VWHS Approach – WTI Crude Oil***

WTI Crude Oil		Standard HS		VWHS-EGARCH(1,1) t-distribution	
In-sample Period	Out-of-sample	95%	99%	95%	99%
1997/1/10-2001/12/28	year 2002	4	0	4	0
1998/1/2-2002/12/27	year 2003	4	1	1	1
1999/1/1-2003/12/26	year 2004	6	0	5	2
2000/1/7-2004/12/31	year 2005	5	0	2	0
2001/1/5-2005/12/30	year 2006	4	0	3	1
2002/1/4-2006/12/29	year 2007	6	2	4	2
2003/1/3-2007/12/28	year 2008	18	8	12	7
2004/1/2-2008/12/26	year 2009	2	0	0	0
2005/1/7-2009/12/25	year 2010	1	0	1	0
2006/1/6-2010/12/31	year 2011	4	0	5	2
2007/1/5-2011/12/30	year 2012	1	0	0	0
2008/1/4-2012/12/28	year 2013	0	0	1	0
		55	11	38	15

VWHS-EGARCH(1,1) t-distribution indicates VaR estimates are estimated based on the EGARCH(1,1) weight loss distribution, and EGARCH(1,1) is estimated under t-distribution.

Table 6.3.4: VaR ***Violations from VWHS Approach – Propane***

Propane		Standard HS		VWHS-EGARCH(1,1) t-distribution	
In-sample Period	Out-of-sample	95%	99%	95%	99%
1997/1/10-2001/12/28	year 2002	0	0	2	0
1998/1/2-2002/12/27	year 2003	4	1	4	4
1999/1/1-2003/12/26	year 2004	5	1	6	3
2000/1/7-2004/12/31	year 2005	3	0	2	0
2001/1/5-2005/12/30	year 2006	3	0	0	0
2002/1/4-2006/12/29	year 2007	1	0	3	1
2003/1/3-2007/12/28	year 2008	13	8	13	7
2004/1/2-2008/12/26	year 2009	2	0	1	0
2005/1/7-2009/12/25	year 2010	2	1	1	1
2006/1/6-2010/12/31	year 2011	1	0	6	0
2007/1/5-2011/12/30	year 2012	3	0	4	0
2008/1/4-2012/12/28	year 2013	0	0	0	0
		37	11	42	16

VWHS-EGARCH(1,1) t-distribution indicates VaR estimates are estimated based on the EGARCH(1,1) weight loss distribution, and EGARCH(1,1) is estimated under t-distribution.

Table 6.4: Kupiec Frequency Backtesting Results of VaR Estimations

Subsample	VaR95%	VaR99%	Sample	VaR95%	VaR99%
x_{low}	7	0	x_{low}	32	4
x_{high}	20	6	x_{high}	57	15
Expected violations	13	2,6	**Expected violations**	44,2	8,84

Expected violations are calculated by $(1-\alpha)n$, and in MS Excel we can compute the lower and upper boundary of the two-sided confidence interval at 95% by using formula $x_{low} = BINOM.INV(n, p, 0.025)$, $x_{high} = BINOM.INV(n, p, 0.975)$. In empirical study, we simply compare the violations of VaR estimates from different approaches with the interval $[x_{low}, x_{high}]$.

6.2. FIGURES

Figure 1: Loss Series for Natural Gas

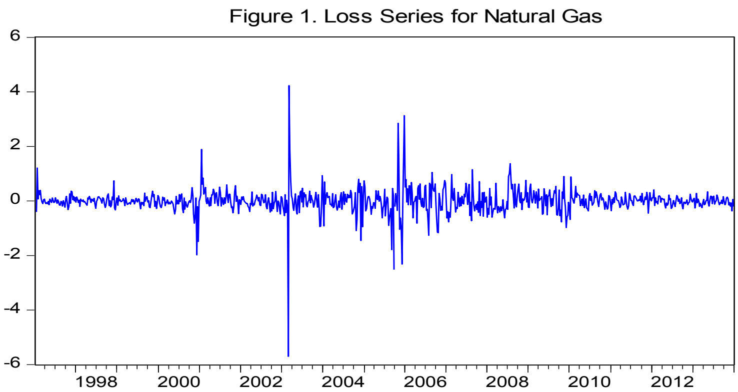

Figure 1 illustrates the loss series for natural gas, and the series exhibits volatility phenomena after the 1997 Asian financial crisis and before the 2008 financial crisis. Hence, it is reasonable to first take the heteroskedasticity test, and if we reject the null hypothesis of no ARCH/GARCH effect, we should then apply the VWHS approach to estimate VaR for natural gas.

Figure 2: Loss Series for Crude Oil

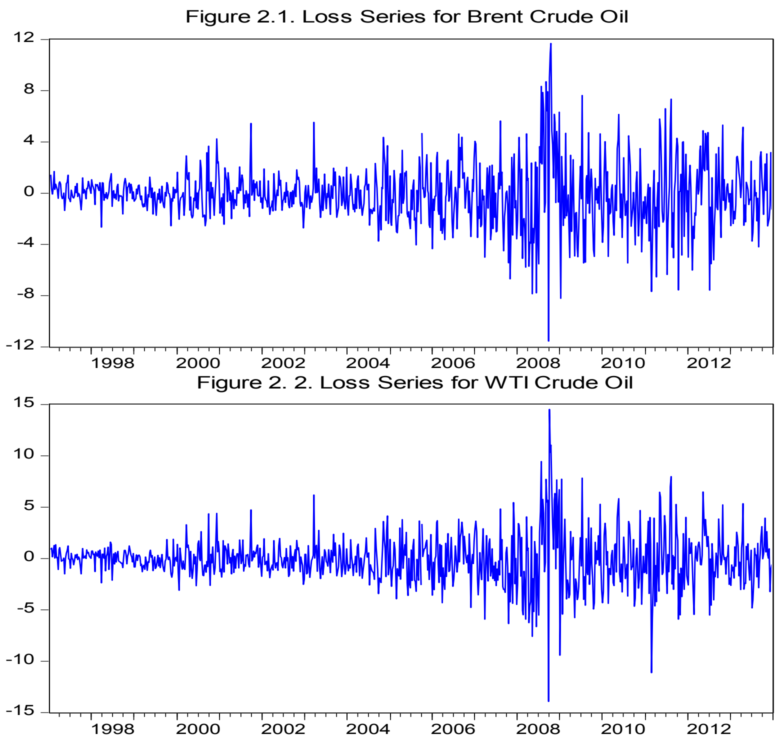

Figure 2.1 and 2.2 depict the loss series for Crude Oil, and we can detect the volatility clustering phenomena around the global financial crisis period. To compare with loss series of natural gas, it seems crude oil series are more volatile.

Figure 3: Loss Series for Propane

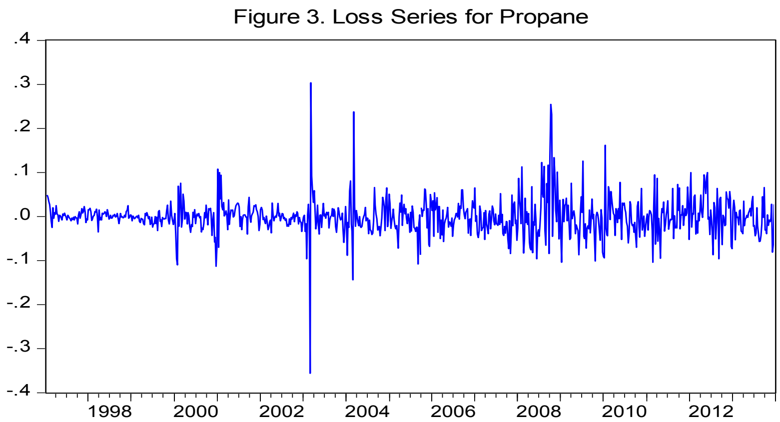

Figure 3 graphs the loss series for propane, the loss series seems less volatile during the financial crisis periods, but there are significant volatility clustering exhibit. It is thus necessary to take the heteroskedasticity test for the loss series of propane.

6.3. DISTRIBUTIONS

6.3.1. Binomial Distribution

The definition of Binomial Distribution is derived from Bernoulli distribution. If we have a random variable X and it can only take two values 0 and 1, and we set $f(1) = p$ and $f(0) = 1 - p$, and then X follows a Bernoulli distribution with parameter p. It is already obvious to see that Kupiec frequency test is basically a binomial test, the way of coding violations and non-violations are exactly 0 and 1, and binomial distribution is the set of all Bernoulli variables. There are a set of variables $X_1, X_2, \dots X_n$, and if these n variables are independent Bernoulli random variables with parameter p, then $B = \sum_{i=1}^{n} X_i$ has a binomial distribution with parameters p and n.

6.3.2. Normal distribution

Normal distribution is frequently employed in financial models, and it plays a central role in the science of financial econometrics. This paper studies few volatility weighted historical simulation models, and these models are derived from both the normal distribution and student t-distribution. Hence, it is necessary to give some brief information of these two models. The normal distribution function can be expressed as $\mathrm{f(z)} = \frac{1}{\sqrt{2\pi\sigma^2}} \mathrm{e}\text{^}(-\frac{(z-\mu)^2}{2\sigma^2})$, and we say that Z follows a normal distribution with mean equals μ and variance is σ^2, which can be written as $Z \sim N(\mu, \sigma^2)$.

6.3.3. Student t-distribution

As mentioned above, there is another distribution commonly used in financial models, namely student t-distribution. To compare with the case of normally distributed, the financial returns usually exhibit fat tails, or excess kurtosis, which are better explained by student t-distribution. It is thus meaningful to detect whether the loss series of energy resources is more likely to follow t-distribution. The student t-distribution function can be expressed as $\mathrm{T} =$

$\frac{Z}{\sqrt{\xi/J}}$. We say that T follows a t-distribution with J degrees of freedom ($T \sim t_J$), if ξ follows a chi-square distribution with J degrees of freedom ($\xi \sim \chi_J^2$) and Z follows a standard normal distribution ($Z \sim N(1,0)$).

6.4. AUTOCORRELATION FUNCTION AND PARTIAL AUTOCORRELATION FUNCATION

The autocorrelation function (ACF) can be expressed as $\rho_k = \frac{cov(Y_t, Y_{t-k})}{V(Y_t)} = \frac{\gamma_k}{\gamma_0}$, where $V(Y_t)$ is the variance of the Y_t, and γ_k is the kth-order autocovariance that is depend only on the distance in time between two observations. Hence, the mathematical expression of the kth-order autocovariance can be expressed as $\gamma_k = cov(Y_t, Y_{t-k}) = cov(Y_{t-k}, Y_t)$. And when t=0, $\gamma_k = cov(Y_t, Y_t) = V(Y_t)$. The kth-order partial autocorrelation function (PACF) for AR(K) model is simply the estimated coefficients θ_k.

6.5. THE BRIEF DEVIATION OF ARCH(1) MODEL

$\varepsilon_t = v_t\sqrt{\alpha_0 + \alpha_1 \varepsilon_{t-1}^2}$

$E[\varepsilon_t^2 | \varepsilon_{t-1}, \varepsilon_{t-2}, \varepsilon_{t-3} \dots] = E[v_t^2(\alpha_0 + \alpha_1 \varepsilon_{t-1}^2) | \varepsilon_{t-1}, \varepsilon_{t-2}, \varepsilon_{t-3} \dots]$ (*)

$E[\varepsilon_t^2 | \varepsilon_{t-1}, \varepsilon_{t-2}, \varepsilon_{t-3} \dots] = E[v_t^2] * E[\alpha_0 + \alpha_1 \varepsilon_{t-1}^2 | \varepsilon_{t-1}, \varepsilon_{t-2}, \varepsilon_{t-3} \dots]$ (**)

$E[\varepsilon_t^2 | \varepsilon_{t-1}, \varepsilon_{t-2}, \varepsilon_{t-3} \dots] = \alpha_0 + E[\alpha_1 \varepsilon_{t-1}^2 | \varepsilon_{t-1}, \varepsilon_{t-2}, \varepsilon_{t-3} \dots]$ (***)

$h_t^2 = E[\varepsilon_t^2 | \varepsilon_{t-1}, \varepsilon_{t-2}, \varepsilon_{t-3} \dots] = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2$ (****)

The step * is because of $\varepsilon_t = v_t\sqrt{\alpha_0 + \alpha_1 \varepsilon_{t-1}^2}$; and the next step ** is based on the assumption that v_t and ε_{t-1} are independent; the third step *** can be obtained by $E[v_t^2] = 1$; and the final step we denote h_t^2 as the conditional variance at time t for simplicity.

7. REFERENCES

Akaike, H., 1973, 'Information theory and an extension of the maximum likelihood principle', *Proc. 2nd Inter. Symposium on Information Theory*, 267-281, Budapest.

Box, George E. P. and Jenkins, Gwilym M., 1976, *Time Series Analysis: Forecasting and Control*, 2nd Edition, Holden Day, San Francisco.

Boudoukh, J., Richardson, M. and Whitelaw R., 1998, 'The Best of Both Worlds: A Hybrid Approach to Calculating Value at Risk', *Risk*, **Vol. 11**, No. 5, 64-67.

Bollerslev, T., 1986, 'Generalized autoregressive conditional heteroskedasticity', *Journal of Econometrics*, **31**, 307-327.

Cabedo, J. D. and Moya, I., 2003, 'Estimating oil price "Value at risk" using the historical simulation approach', *Energy Economics*, **25**, 239-253.

Campbell, John Y., Lo, Andrew W. and Mackinlay, A. Craig., 1998, *The Econometrics of Financial Markets*, Princeton University Press.

CBC News, 2006, 'The price of oil – in context'. [Online] https://web.archive.org/web/20070609145246/http://www.cbc.ca/news/background/oil/, [accessed 01 April 2014].

Dowd, K., 2005, *Measuring Market Risk*, 2nd Edition, John Wiley & Sons, Ltd, The Atrium, Southern Gate, Chichester, England.

Engle, Robert F., 1982, 'Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation', *Econometrica*, **Vol. 50**, No. 4, 987-1007.

Enders, W., 2010, *Applied Econometric Time Series*, 3rd Edition, John Wiley & Sons, Ltd, The Atrium, Southern Gate, Chichester, England.

Giot, P., Laurent, S., 2003, ‘Market risk in commodity markets: a VaR approach’, *Energy Economics*, **25,** 435-457.

Hung, Cheng J., Lee, Chih M., Liu, Chun H., 2008, ‘Estimation of value-at-risk for energy commodities via fat-tailed GARCH models’, *Energy Economics*, **30,** 1173-1191.

Hendricks, D., 1996, ‘Evaluation of value at risk modeling using historical data’. Federal Reserve Bank of New York. *Economic Policy Review*, 36-39.

Marimoutou, V., Raggad, B. and Trabelsi A., 2009, ‘Extremes Value Theory and Value at Risk: Application to oil market’, *Energy Economics*, **31**, 519-530.

McAleer, M. and Veiga, Da. B., 2008, ‘Forecasting value-at-risk with a parsimonious portfolio spillover GARCH (PS-GARCH) model’, *Journal of Forecasting*, **27**, 1-19.

McAleer, M. and Veiga, Da. B., 2008, ‘Single-index and portfolio models for forecasting value-at-risk thresholds’, *Journal of Forecasting*, **27**, 217-235.

Pritsker, M., 2006, ‘The hidden dangers of historical simulation’, *Journal of Banking and finance*, **30**, 561-582.

Sadeghi, M. and Shavvalpour, S., 2006, ‘Energy risk management and value at risk modeling’, *Energy Policy*, **34**, 3367-3373.

Sadorsky, P., 1999, ‘Oil price shocks and stock market activity’, *Energy Economics*, **21**, 449-469.

Samuelson, Paul A. and Nordhaus, William D., 1992, *Economics*, 14th Edition, McGraw-Hill Inc., US.

Schwarz G., 1978, ‘Estimating the dimension of a model’, *The Annals of Statistics*, **Vol. 6**, No. 2, 461-464.

Sims, Christopher A., Stock, James H. and Watson, Mark W., 1990, 'Inference in linear time series models with some unit roots', *Econometirca*, **Vol. 58**, No.1, 113-144.

Verbeek, M., 2012, *A Guide to Modern Econometrics,* 4th Edition, John Wiley & Sons, Ltd, The Atrium, Southern Gate, Chichester, England.

Zakoian, Jean-Michel, 1994, 'Threshold heteroskedastic models', *Journal of Economic Dynamics and Control*, **18**, 931-955.