

**ANALYZING VALUE AT RISK AND EXPECTED SHORTFALL  
METHODS: THE USE OF PARAMETRIC, NON-PARAMETRIC,  
AND SEMI-PARAMETRIC MODELS**

**by**

**Xinxin Huang**

**A Thesis Submitted to the Faculty of Graduate Studies**

**The University of Manitoba**

**in partial fulfillment of the requirements for  
the degree of**

**MASTER OF SCIENCE**

**Department of Agribusiness and Agricultural Economics**

**University of Manitoba**

**Winnipeg, Manitoba**

**Copyright © 2014 by Xinxin Huang**

## **ABSTRACT**

Value at Risk (VaR) and Expected Shortfall (ES) are methods often used to measure market risk. This is the risk that the value of assets will be adversely affected by the movements in financial markets, such as equity markets, bond markets, and commodity markets. Inaccurate and unreliable Value at Risk and Expected Shortfall models can lead to underestimation of the market risk that a firm or financial institution is exposed to, and therefore may jeopardize the well-being or survival of the firm or financial institution during adverse markets. Crotty (2009) argued that using inaccurate Value at Risk models that underestimated risk was one of the causes of 2008 US financial crisis. For example, past Value at Risk models have often assumed the Normal Distribution, when in reality markets often have fatter tail distributions. As a result, Value at Risk models based on the Normal Distribution have often underestimated risk. The objective of this study is therefore to examine various Value at Risk and Expected Shortfall models, including fatter tail models, in order to analyze the accuracy and reliability of these models.

Three main approaches of Value at Risk and Expected shortfall models are used. They are (1) 11 parametric distribution based models including 10 widely used and most studied models, (2) a single non-parametric model (Historical Simulation), and (3) a single semi-parametric model (Extreme Value Theory Method, EVT, which uses the General Pareto Distribution). These models are then examined using out of sample analysis for daily returns of S&P 500, crude oil, gold and the Vanguard Long Term Bond

Fund (VBLTX). Further, in an attempt to improve the accuracy of Value at Risk (VaR) and Expected Shortfall (ES) models, this study focuses on a new parametric model that combines the ARMA process, an asymmetric volatility model (GJR-GARCH), and the Skewed General Error Distribution (SGED). This new model, ARMA(1,1)-GJR-GARCH(1,1)-SGED, represents an improved approach, as evidenced by more accurate risk measurement across all four markets examined in the study. This new model is innovative in the following aspects. Firstly, it captures the autocorrelation in returns using an ARMA(1,1) process. Secondly, it employs GJR-GARCH(1,1) to estimate one day forward volatility and capture the leverage effect (Black, 1976) in returns. Thirdly, it uses a skewed fat tail distribution, skewed General Error Distribution, to model the fat tails of daily returns of the selected markets.

The results of this study show that the Normal Distribution based methods and Historical Simulation method often underestimate Value at Risk and Expected Shortfall. On the other hand, parametric models that use fat tail distributions and asymmetric volatility models are more accurate for estimating Value at Risk and Expected Shortfall. Overall, the proposed model here (ARMA(1,1)-GJR-GARCH(1,1)-SGED) gives the most balanced Value at Risk results, as it is the only model for which the Value at Risk exceedances fell within the desired confidence interval for all four markets. However, the semi-parametric model (Extreme Value Theory, EVT) is the most accurate Value at Risk model in this study for S&P 500 Index, likely due to fat tail behavior (including the out of sample data). These Results should be of interest to researchers, risk managers, regulators and analysts in providing improved risk measurement models.

Keywords: Risk Management, Volatility Estimate, Value at Risk, GARCH, ARMA, General Error Distribution (GED), ARMA(1,1)-GJR-GARCH(1,1)-SGED, Extreme Value Theory (EVT), General Pareto Distribution (GPD), Expected Shortfall (ES), Conditional Tail Expectation (CTE), Conditional Value at Risk (CVaR)

## **ACKNOWLEDGMENTS**

I would like to thank my advisors Dr. Milton Boyd and Dr. Jeffrey Pai for the encouragement, support, advice and inspiration that they have generously given me throughout the entire program. I would also like to thank my wife, Jenny, for her support and encouragement. Last but not least, I am grateful for the support of Winnipeg Commodity Exchange Fellowship, and for the assistance and guidance from committee members Dr. Lysa Porth and Dr. Barry Coyle.

# TABLE OF CONTENTS

<b>ABSTRACT.....</b>	<b>ii</b>
<b>ACKNOWLEDGMENTS .....</b>	<b>iv</b>
<b>LIST OF FIGURES .....</b>	<b>viii</b>
<b>LIST OF TABLES .....</b>	<b>x</b>
<b>CHAPTER 1. BACKGROUND AND INTRODUCTION .....</b>	<b>1</b>
BACKGROUND OF VALUE AT RISK AND EXPECTED SHORTFALL.....	1
THREE APPROACHES FOR VALUE AT RISK AND EXPECTED SHORTFALL.....	2
INTRODUCTION .....	3
<b>CHAPTER 2. ANALYZING VALUE AT RISK AND EXPECTED SHORTFALL METHODS: THE USE OF PARAMETRIC, NON-PARAMETRIC, AND SEMI-PARAMETRIC MODELS.....</b>	<b>7</b>
THEORY AND LITERATURE.....	7
Definition of Value at Risk (VaR) and Expected Shortfall (ES) .....	7
A Review of the Three Main Approaches for Value at Risk (VaR) and Expected Shortfall (ES) .....	8
Literature Review for Value at Risk and Expected Shortfall Model Comparisons .....	16
DATA .....	17
METHODS .....	18
The Proposed Model: ARMA(1,1)-GJR-GARCH(1,1)-SGED .....	18
Examining Goodness of Fit of Distributions Assumed by Parametric Models.....	22
Examining Goodness of Fit of General Pareto Distribution in the Semi-Parametric Model (Extreme Value Theory) .....	24
Out of Sample Test Procedures for Value at Risk and Expected Shortfall.....	25
OUT OF SAMPLE TEST RESULTS.....	29
Value at Risk Out of Sample Test Results for All Models .....	29
Expected Shortfall Out of Sample Test Results for All Models .....	30

CHAPTER SUMMARY.....	31
END NOTES .....	33
<b>CHAPTER 3. SUMMARY.....</b>	<b>63</b>
PROBLEM AND OBJECTIVE.....	63
DATA AND METHODS .....	64
RESULTS .....	65
CONCLUSION .....	67
<b>REFERENCES.....</b>	<b>69</b>
<b>APPENDIX A: ABBREVIATIONS LIST .....</b>	<b>72</b>
<b>APPENDIX B: LIST OF VALUE AT RISK AND EXPECTED SHORTFALL     MODEL TYPES .....</b>	<b>73</b>
<b>APPENDIX C: GRAPH FOR DATA PROBABILITY DENSITY VERSUS     NORMAL DISTRIBUTION PROBABILITY DENSITY FOR S&amp;P 500,     OIL, GOLD, AND VANGUARD LONG TERM BOND FUND.....</b>	<b>75</b>

## LIST OF FIGURES

Figure 2.1	QQ Plot for S&P 500 (Daily Return Standardized Residuals, 1950-2012): ARMA(1,1)-GJR-GARCH(1,1)-Norm (Normal Distribution).....	46
Figure 2.2	QQ Plot for S&P 500 (Daily Return Standardized Residuals, 1950-2012): ARMA(1,1)-GJR-GARCH(1,1)-STD (Student T Distribution).....	47
Figure 2.3	QQ Plot for S&P 500 (Daily Return Standardized Residuals, 1950-2012): ARMA(1,1)-GJR-GARCH(1,1)-SSTD (Skewed Student T Distribution) ....	48
Figure 2.4	QQ Plot for S&P 500 (Daily Return Standardized Residuals, 1950-2012): ARMA(1,1)-GJR-GARCH(1,1)-GED (General Error Distribution).....	49
Figure 2.5	QQ Plot for S&P 500 (Daily Return Standardized Residuals, 1950-2012): ARMA(1,1)-GJR-GARCH(1,1)-SGED (Skewed General Error Distribution).....	50
Figure 2.6	Extreme Value Theory (EVT): Actual Versus Estimated Left Tail for S&P 500 (Standardized Residuals).....	51
Figure 2.7	Extreme Value Theory (EVT): Actual Versus Estimated Left Tail for Crude Oil (Standardized Residuals).....	52
Figure 2.8	Extreme Value Theory (EVT): Actual Versus Estimated Left Tail for Gold (Standardized Residuals).....	53
Figure 2.9	Extreme Value Theory (EVT): Actual Versus Estimated Left Tail for Vanguard Long Term Bond Fund (Standardized Residuals).....	54
Figure 2.10	Actual Daily Return Versus Estimated Value at Risk (Model: ARMA(1,1)-GJR-GARCH(1,1)-SGED, Data: Out of Sample S&P 500 Daily Return N=1000, Confidence Level = 99%) .....	55
Figure 2.11	Actual Daily Return Versus Estimated Value at Risk (Model: ARMA(1,1)-GJR-GARCH(1,1)-SGED, Data: Out of Sample S&P 500 Daily Return N=1000, Confidence Level = 95%) .....	56
Figure 2.12	Actual Daily Return Versus Estimated Value at Risk (Model: ARMA(1,1)-GJR-GARCH(1,1)-SGED, Data: Out of Sample Crude Oil Daily Return N=1000, Confidence Level = 99%) .....	57
Figure 2.13	Actual Daily Return Versus Estimated Value at Risk (Model: ARMA(1,1)-GJR-GARCH(1,1)-SGED, Data: Out of Sample Crude Oil Daily Return N=1000, Confidence Level = 95%) .....	58



Figure 2.14	Actual Daily Return Versus Estimated Value at Risk (Model: ARMA(1,1)-GJR-GARCH(1,1)-SGED, Data: Out of Sample Gold Daily Return N=1000, Confidence Level = 99%) .....	59
Figure 2.15	Actual Daily Return Versus Estimated Value at Risk (Model: ARMA(1,1)-GJR-GARCH(1,1)-SGED, Data: Out of Sample Gold Daily Return N=1000, Confidence Level = 95%) .....	60
Figure 2.16	Actual Daily Return Versus Estimated Value at Risk (Model: ARMA(1,1)-GJR-GARCH(1,1)-SGED, Data: Out of Sample Vanguard Long Term Bond Fund Daily Return N=1000, Confidence Level = 99%) ...	61
Figure 2.17	Actual Daily Return Versus Estimated Value at Risk (Model: ARMA(1,1)-GJR-GARCH(1,1)-SGED, Data: Out of Sample Vanguard Long Term Bond Fund Daily Return N=1000, Confidence Level = 95%) ...	62

## LIST OF TABLES

Table 2.1	Descriptive Statistics and Sources of Data (Daily Return).....	34
Table 2.2	Leverage Effect Test: Sign Bias Test Results for ARMA(1,1)-GARCH(1,1) (Daily Returns) .....	35
Table 2.3	Leverage Effect Test: Sign Bias Test Results for ARMA(1,1)-GJR- GARCH(1,1) (Daily Return) .....	36
Table 2.4	Parameter Estimates for ARMA(1, 1)-GARCH(1, 1) (Daily Return) .....	37
Table 2.5	Parameter Estimates for ARMA(1, 1)-GJR-GARCH(1, 1) (Daily Returns)...	38
Table 2.6	Pearson Goodness-of-Fit Statistic Test Results for Five Distributions, ARMA(1,1)-GARCH(1,1) (Daily Return) .....	39
Table 2.7	Pearson Goodness-of-Fit Statistic Test Results for Five Distributions, ARMA(1,1)-GJR-GARCH(1,1) (Daily Return) .....	41
Table 2.8	Parameter Estimates for Extreme Value Theory (EVT) (Daily return, In Sample Data) .....	43
Table 2.9	Value at Risk Out of Sample Results: Number of Exceedances (Daily Return).....	44
Table 2.10	Expected Shortfall Out of Sample Test Results: P Values (Daily Return)....	45

# **CHAPTER 1**

## **BACKGROUND AND INTRODUCTION**

### **BACKGROUND OF VALUE AT RISK AND EXPECTED SHORTFALL**

The origin of quantifying financial losses can be traced back to New York Stock Exchange's capital requirement for its members in the 1920s (Holton, 2002). In early 1950s, statistically quantifying financial losses was studied by portfolio theorists for portfolio optimization purpose (Holton, 2002). It was in the 1980s when Value at Risk began to be used as a financial market risk measure by financial institutions, such as JP Morgan, a US bank. JP Morgan also publicized its internal Value at Risk model, RiskMetrics, in 1990s, which is defined in its Technical Document (1996), as "a measure of the maximum potential change in value of a portfolio of financial instruments with a given probability over a pre-set horizon". In 1995, the Basel Committee on Banking Supervision (BCBS) allowed internally developed Value at Risk (VaR) models for monitoring daily market risk and calculating capital reserves. Prior to this, a fixed percentage approach was required. The problem with the fixed percentage approach is that it does not adjust for portfolio specific risk, which could lead to excess reserve and inefficient use of capital.

Artzner (1999) theoretically criticized Value at Risk as an incoherent risk measure for its lack of sub-additive ability and information about the size of loss when the true loss does exceed the Value at Risk. An alternative to Value at Risk is Expected Shortfall (ES) [also known as Conditional Tail Expectation (CTE) or Conditional Value at risk (CVaR)]. This is a modified version of Value at Risk that overcomes the above

mentioned theoretical deficiencies. Expected Shortfall is defined by McNeil (1999) as “the expected loss given that the true loss does exceed the Value at Risk”.

### THREE APPROACHES FOR VALUE AT RISK AND EXPECTED SHORTFALL

Since the inception of using Value at Risk in risk management, three main approaches, namely, parametric, non-parametric and semi-parametric, have gained popularity over the others for either the ease of use, or better accuracy. Below is a brief introduction to the three approaches.

First, the parametric approach (the most common approach) attempts to model financial asset returns using parametric distributions. The RiskMetrics Method (J.P. Morgan and Reuters, 1996) is one of the most influential models in this category. It assumes that returns of financial assets follow the Normal Distribution, and it employs Exponentially Weighted Moving Average (EWMA) method to estimate the volatility of returns. Although EWMA puts less weight on older historical data, the normality assumption tends to cause underestimation of risk because empirical return distributions often have fatter tails than Normal Distribution. Historical data shows that returns of financial markets, such as equity markets, often have high kurtosis (fat tails) and skewness (Duffie and Pan, 1997; Taylor, 2005). To capture the high kurtosis and skewness in returns, researchers and risk managers use distributions that exhibit fat tails and skewness, such as Student T (Jorion, 1996), Student T with skewness (Giot and Laurent, 2004), and General Error Distribution (GED) (Kuester, Mittnik and Paolella, 2006). Researchers and risk managers attempt to improve the accuracy of volatility

estimation, in order to improve the accuracy of Value at Risk and Expected Shortfall models. The other widely used volatility models include the GARCH family and stochastic volatility models (Pan and Duffie, 1997).

Second, the non-parametric approach, also known as Historical Simulation (HS) method (Best, 1999), uses empirical percentiles of the observed data to estimate future risk exposure. The advantage of this approach is its ease of use as it does not require a large amount of computation. However, the Value at Risk calculated by this approach put too much reliance on older historical data, and therefore has low accuracy.

Third, the semi-parametric approach is a combined approach of both parametric and non-parametric methods. The Extreme Value Theory (EVT) method is considered one of the most practical semi-parametric models for the efficient use of data (McNeil and Frey, 2000). It separates the estimation of extreme tails and center quantiles. EVT models do well in estimating extreme Value at Risk and Expected Shortfall models with return data that exhibits high kurtosis (McNeil, 1999; Embrechts, Kluppelberg and Mikosch, 1997). The disadvantage of EVT is that it heavily relies on empirical data for estimating the center quantiles of the distribution.

## INTRODUCTION

### Objective

Value at Risk (VaR) and Expected Shortfall (ES) are financial risk methods that are often used to measure market (price) risk. This is the risk that the value of a portfolio

will be adversely affected by the movements in financial markets, such as equity markets, bond markets and commodity markets. Inaccurate Value at Risk and Expected Shortfall models can lead to underestimation of the market risk that a firm or financial institution is exposed to, and therefore jeopardize the well-being or survival of the firm or financial institution during adverse market movements. Crotty (2009) argued that using inaccurate Value at Risk models that underestimate risk led to inadequate capital reserves in large banks and therefore was one of causes of the 2008 US financial crisis. For example, past Value at Risk models have often assumed the Normal Distribution, when in reality markets often have fatter tail distributions. As a result, Value at Risk models based on the Normal Distribution have often underestimated risk. Therefore, the objective of this study is to examine various Value at Risk and Expected Shortfall models, including fatter tail models, in order to analyze the accuracy and reliability of these models.

### Data and Methods

In this study, the principles of selecting data are (1) to include a variety of financial assets that are exposed to daily market risks, (2) to include a sufficiently long time period of data for each of the selected markets, and (3) to ensure all data have the same start and end date for out of sample data. Based on these principles, data selected include S&P 500 Index (price adjusted for dividends), crude oil, gold and the Vanguard Long Term Bond Fund (VBLTX, adjusted for dividends). S&P 500 Index from January, 1950 to April, 2012 and Vanguard Long Term Bond Fund prices (VBLTX) from March, 1994 to April, 2012 are obtained from Yahoo Finance website. Crude oil prices from

January, 1986 to April, 2012 are obtained from US Energy Information Administration (EIA) website. Gold prices from April, 1968 to April, 2012 are obtained from Federal Reserve Bank of St. Louis website. Daily log returns are computed from the obtained data. The same 1000 days (approximately four years from May, 2008 to April, 2012) are used as the out of sample data.

In this study, three approaches are used for estimating Value at Risk and Expected Shortfall. They are (1) parametric approach (11 parametric distribution based models), (2) non-parametric approach (a single model, Historical Simulation), and (3) semi-parametric approach (a single model, Extreme Value Theory). The parametric approach includes the widely used and most studied models (Normal, Student T and General Error Distribution based models). In addition, this study also proposes a new method, a modified parametric model, ARMA(1,1)-GJR-GARCH(1,1)-SGED (Skewed General Error Distribution), aimed to increase the accuracy of Value at Risk and Expected Shortfall models. This new model is innovative in the following aspects. Firstly, it captures the autocorrelation in returns using ARMA(1,1) process. Secondly, it employs GJR-GARCH(1,1) to estimate one day forward volatility and capture the leverage effect (Black, 1976) in returns. Thirdly, it uses a skewed fat tail distribution, skewed General Error Distribution (SGED), to model the fat tails of daily returns. In order to analyze the accuracy of these Value at Risk and Expected Shortfall models, statistical tests and out of sample tests are conducted to ensure test results are robust.

## Thesis Structure

Following this Chapter 1 of background and introduction, is Chapter 2. This is an analysis of Value at Risk and Expected Shortfall using parametric, non-parametric and semi-parametric models, which includes an introduction of the theory of Value at Risk and Expected Shortfall, literature review, data, methods, and results. This is followed by Chapter 3, a comprehensive summary of the study.



## CHAPTER 2

### ANALYZING VALUE AT RISK AND EXPECTED SHORTFALL METHODS: THE USE OF PARAMETRIC, NON-PARAMETRIC, AND SEMI-PARAMETRIC MODELS

#### THEORY AND LITERATURE

##### Definition of Value at Risk (VaR) and Expected Shortfall (ES)

Value at Risk (VaR) is defined as “the maximum amount of money that may be lost on a portfolio over a given period of time, with a given level of confidence” (Best, 1999). Since Value at Risk was popularized by J.P. Morgan in monitoring daily market risk, it is typically calculated over one-day horizon. Statistically speaking, if the one day log return of a portfolio on day  $t$  is denoted by  $R_t$ <sup>1</sup> (end note 1), then  $(1-\alpha)$  % Value at Risk on day  $t$ ,  $VaR_{t,1-\alpha}$ , is the amount such that

$$P(R_t < VaR_{t,1-\alpha}) = \alpha\% \quad (1)$$

Expected Shortfall (ES), also known as Conditional Tail Expectation (CTE) or Conditional Value at Risk (CVaR), an alternative risk measure for Value at Risk, is defined as “the expected size of loss that exceeds VaR” (McNeil, 1999). The Expected Shortfall at  $(1-\alpha)\%$  confidence level is the expected loss on day  $t$  given that the loss does exceed  $VaR_{t,1-\alpha}$ , or mathematically (McNeil and Frey, 2000):

$$ES_{t,1-\alpha} = E[R_t | R_t < VaR_{t,1-\alpha}] \quad (2)$$

where  $R_t$  is the log return on day  $t$ .

## A Review of the Three Main Approaches for Value at Risk (VaR) and Expected Shortfall (ES)

There are three main approaches, namely parametric, non-parametric and semi-parametric, for estimating Value at Risk (VaR) and Expected Shortfall (ES). This section reviews the three approaches.

### *The Parametric Approach*

Value at Risk (VaR) and Expected Shortfall (ES) models under the parametric approach assume that the returns of financial assets over a given period of time can be approximated by a certain parametric distribution (e.g. Normal, Student T, or General Error Distribution). Assuming returns are conditional on the previous returns, an example of parametric method is illustrated as follows:

Let  $R_t$  be the daily return of some financial asset on day  $t$  and follow a parametric distribution  $D$  with conditional mean and variance. That is,

$R_t | R_1, R_2, \dots, R_{t-1} \sim D(\mu_t, \sigma_t^2)$ , where  $D$  is a known distribution (Taylor, 2005), or

$$R_t = \mu_t + \sigma_t X_t \quad (3)$$

where  $\sigma_t | R_1, R_2, \dots, R_{t-1}$  is the standard deviation of  $R_t$  (also referred to as the volatility of returns);  $X_t \sim i.i.d D(0,1)$ , is the standardized residual of  $R_t$ .

Now, substitute  $R_t$  in Equation (1) by Equation (3),

$$P(\mu_t + \sigma_t X_t < VaR_{t,1-\alpha}) = \alpha\% \quad (4)$$

Since  $D(0,1)$  is known, the  $\alpha$  percentile of  $D(0,1)$  is also known.  $VaR_{t,1-\alpha}$  can then be obtained by the following equation.

$$VaR_{t,1-\alpha} = \mu_t + \sigma_t Z_\alpha \quad (5)$$

where  $Z_\alpha$  is the  $\alpha$  percentile of  $D(0,1)$ ,  $\mu_t$  and  $\sigma_t$  follow the definition in Equation (3).

For example, assume  $D$  follows standard Normal Distribution [ $D \sim N(0,1)$ ] and  $\mu_t = 0$ , portfolio size \$1,000,000, daily volatility ( $\sigma_t$ ) 2%, desired confidence level is 99% or  $\alpha = 0.01$  ( $Z \approx -2.33$ ), the absolute value of value at risk on day  $t$  is then  $(0 + 2\% \times 2.33) \times \$1,000,000 = \$46,600$ . This means that the maximum amount that can be lost on the portfolio in one day is \$46,600, given a 99% confidence level.

The corresponding  $(1 - \alpha\%)$  Expected Shortfall on day  $t$  can then be expressed as below,

$$ES_{t,1-\alpha} = \mu_t + \sigma_t E[X_t | X_t < Z_\alpha] \quad (6)$$

where  $Z_\alpha$ ,  $\mu_t$  and  $\sigma_t$  follow the definition in Equation (5).

Continuing from the example for Value at Risk (VaR) above, given that the loss on day  $t$  will exceed \$46,600, the expected amount of money will be lost is  $(0 + 2\% \times 2.65) \times \$1,000,000 = \$53,000$  ( $2.65 \approx -E[X_t | X_t < Z_{0.01}] = -\int_{-\infty}^{-2.33} \left(\frac{1}{0.01}\right) x f(x) dx$ ,  $f(x)$  is standard normal pdf ). This means that given the loss on day exceeds the Value at Risk (VaR) of \$46,600, the expected amount of money lost on day  $t$  is \$53,000 with a 99% confidence level.

Equation (5) and (6) imply that the estimation of Value at Risk and Expected Shortfall under parametric approach depends on (1) the estimation of conditional mean,

(2) the estimation of conditional variance or volatility and (3) the distribution assumed for standardized residual ( $X_t$ ). Empirical evidence often suggests that the one day mean return in many financial markets is very close to zero (Taylor, 2005). In this study, the conditional mean for daily returns will be assumed to be a constant, or

$$\mu_t = \mu \quad (7)$$

One of the most influential conditional variance models in the literature is Autoregressive Conditional heteroskedasticity (ARCH) (Engle, 1982). “Autoregressive” here means that the variance of some asset returns that is conditional on the information of previous returns is dependent on this information. GARCH (Bollerslev, 1986) is a generalized model of ARCH by including a lagged variance term in the model. GARCH is one of the most widely used conditional variance models because it is relatively consistent with financial market behavior. The first order GARCH(1,1) conditional variance model is defined as follows,

$$\sigma_t^2 = \beta_0 + \beta_1(R_{t-1} - \mu)^2 + \beta_2\sigma_{t-1}^2 \quad (8)$$

where  $\beta_0, \beta_1, \beta_2$  are the model parameters ( $\beta_i \geq 0$  ( $i = 0, 1, 2$ );  $\beta_1 + \beta_2 < 1$ ).  $\sigma_t$  is the volatility of returns on day  $t$ ;  $R_{t-1}$  is the asset return on day  $t-1$ .  $\mu_{t-1}$  is the mean of return on day  $t - 1$  (note this study assumes to the mean of return is a constant over time, therefore  $\mu_t$  and  $\mu$  are interchangeable in this study).

Using maximum likelihood estimation method (assuming standardized residuals follow standard Normal Distribution), consistent estimators can be obtained for parameters

$\beta_0, \beta_1, \beta_2, \mu$  and  $\sigma_1$ . Let these consistent estimates be  $\widehat{\beta}_0, \widehat{\beta}_1, \widehat{\beta}_2, \widehat{\mu}$  and  $\widehat{\sigma}_1$ . From Equation (5), the equation below can be obtained,

$$\widehat{VaR}_{t,1-\alpha} = \widehat{\mu} + \widehat{\sigma}_t Z_\alpha \quad (9)$$

and by Equation (6)

$$\widehat{ES}_{t,1-\alpha} = \widehat{\mu} + \widehat{\sigma}_t E[X_t | X_t < Z_\alpha] \quad (10)$$

where  $\widehat{\sigma}_t$  is obtained by back-iterating Equation (6) (t-1) times. That is,

$$\begin{aligned} \widehat{\sigma}_t &= \sqrt{\widehat{\beta}_0 + \widehat{\beta}_1(R_{t-1} - \widehat{\mu})^2 + \widehat{\beta}_2\widehat{\sigma}_{t-1}^2}, \widehat{\sigma}_{t-1} = \sqrt{\widehat{\beta}_0 + \widehat{\beta}_1(R_{t-2} - \widehat{\mu})^2 + \widehat{\beta}_2\widehat{\sigma}_{t-2}^2}, \dots, \\ \widehat{\sigma}_2 &= \sqrt{\widehat{\beta}_0 + \widehat{\beta}_1(R_1 - \widehat{\mu})^2 + \widehat{\beta}_2\widehat{\sigma}_1^2}. \end{aligned}$$

Another widely used conditional variance model is the Exponentially Weighted Moving Average (EWMA). The well known Riskmetrics (the Value at Risk model by J.P. Morgan) uses EWMA to estimate the volatility. The estimated one day forward volatility is given by the expression below (J.P. Morgan and Reuters, 1996),

$$\sigma_t = \sqrt{\lambda \sigma_{t-1}^2 + (1 - \lambda) R_{t-1}^2} \quad (11)$$

where  $\sigma_t$  is the volatility estimate for day t;  $\sigma_{t-1}$  is the volatility on day t-1;  $R_{t-1}$  is the return on day  $t - 1$ ;  $\lambda$  is the weight given to the previous day's volatility.  $(1 - \lambda)$  is the weight given to the previous day's return. ( $\lambda$  is also referred to as the decay factor). The RiskMetrics model pre-sets the decay factor to be 0.94 based on empirical data on 480 different time series data.

An interesting phenomenon about volatility, discussed by Black (1976), is that stock price movements are often negatively correlated with volatility. This is referred to as the leverage effect by Black (1976). Black (1976) argued that falling stock prices often imply an increased leverage and therefore higher market perceived uncertainty which leads to higher volatility. After that, the term “leverage effect” was widely used to describe the asymmetric behavior of volatility that for the same magnitude, losses are accompanied with higher volatility, compared to gains. To capture this asymmetry in volatility, Glosten, Jagannathan and Runkle (1993) introduced a modified GARCH by separating the positive and negative error terms. Its first order expression is as follows,

$$\sigma_t^2 = \beta_0 + \beta_1 \varepsilon_{t-1}^2 + \beta_{1-} I_{\varepsilon < 0} \varepsilon_{t-1}^2 + \beta_2 \sigma_{t-1}^2 \quad (12)$$

where  $\beta_i \geq 0$  ( $i = 0, 1, 2$ );  $\beta_1 + \beta_{1-} > 0$ ;  $\varepsilon_{t-1}$  is the error term at time t-1 [ $\varepsilon_{t-1} = \sigma_{t-1}^{-1} X_{t-1}$ , where  $\sigma_{t-1}$  and  $X_{t-1}$  follow the definition in Equation (3)];  $I_{\varepsilon < 0}$  is an indicator function such that  $I_{\varepsilon < 0} = 0$  when  $\varepsilon_{t-1} \geq 0$  and  $I_{\varepsilon < 0} = 1$  when  $\varepsilon_{t-1} < 0$ . This modified GARCH model is referred to as GJR-GARCH(1,1). Other major models that consider leverage effect include EGARCH (Nelson, 1991), APARCH (Ding et al., 1993) and FGARCH (Hentschel, 1995). [Implied volatility is not considered in this study, since the focus here is on historical volatility.]

Another variation in parametric models is the distributional assumptions for returns. Studies suggest that the empirical distribution of financial asset returns often exhibit high kurtosis and occasionally skewness in comparison to Normal Distribution (Taylor, 2005). Therefore assuming normality often leads to underestimation of Value at Risk. In attempts to capture the high kurtosis and skewness, distributions that exhibit fat

tails and skewness are used in Value at Risk and Expected Shortfall models. These distributions include Student T (Jorion, 1997), Student T with skewness (Giot and Laurent, 2004), and General Error Distribution (GED) (Kuester, Mittnik and Paolella, 2006; Fan, Zhang, Tsai and Wei, 2008).

### *The Non-Parametric Approach*

The non-parametric approach is also referred to as the Historical Simulation Method. The Historical Simulation method does not assume any distributions for return, so is referred to as a non-parametric method. It uses empirical percentiles to estimate Value at Risk and Expected Shortfall. To illustrate its Value at Risk estimation process, let  $R_1, R_2, \dots, R_{100}$  be a time series of daily financial returns from day one to day one hundred. The series is then sorted from smallest to largest. The  $VaR_{101,95\%}$  is simply the 5<sup>th</sup> return from the smallest (Best, 1999). Accordingly, the  $ES_{101,95\%}$  is the average of the 4 returns from 1<sup>st</sup> to the 4<sup>th</sup> from the smallest. The advantage of this method is apparently its simplicity. The disadvantage is the over reliance on past returns.

### *The Semi-Parametric Approach*

Value at Risk (VaR) and Expected Shortfall (ES) models under semi-parametric approach use both parametric and non-parametric methods jointly to estimate the return distribution, and further estimate Value at Risk (VaR) and Expected Shortfall (ES). In the example above for Historical Simulation method, suppose the 5 smallest returns are used

to estimate the tail of the return distribution, and the 95% Value at Risk is obtained parametrically using an estimated tail distribution. Then, the method used above is then a combination of parametric and non-parametric approach, or the semi-parametric approach. Models under this approach include Extreme Value Theory (EVT, based on General Pareto Distribution), Block Maxima and Hill Estimator methods (McNeil, 1999). The Extreme Value Theory (EVT) model is one of most practical semi-parametric models for its efficient use of data (McNeil and Frey, 2000), and therefore the only semi-parametric model examined in this study. An illustration of the EVT model is provided below.

Assuming  $F(x)$  is the CDF of the random variable  $X$  from an unknown distribution, the excess variable  $Y$  is defined as the excess of  $X$  over a chosen threshold  $u$ , or  $Y = X - u$  for all  $X > u$ . Then the CDF of  $Y$  can be written as follows,

$$F_u(y) = P\{X - u \leq y | X > u\} = \frac{F(y+u) - F(u)}{1 - F(u)} \quad (13)$$

Balkema et al. (1974) and Pickands (1975) showed that for a large class of distributions (including Pareto, Student T, Loggamma, Gamma, Normal, Lognormal), a positive function  $\beta(u)$  can be found such that

$$\lim_{u \rightarrow x_0} \sup_{0 \leq y < x_0 - u} |F_u(y) - G_{\xi, \beta(u)}(y)| = 0 \quad (14)$$

where  $x_0$  is the upper end point of distribution  $X$ , or  $x_0 = F_x^{-1}(1)$ ;  $G_{\xi, \beta}(y)$  is the CDF of General Pareto distribution (GPD),



$$G_{\xi,\beta}(y) = \begin{cases} 1 - \left(1 + \frac{\xi y}{\beta}\right)^{-\frac{1}{\xi}} & \xi \neq 0 \\ 1 - \exp\left(-\frac{y}{\beta}\right) & \xi = 0 \end{cases}$$

where  $\beta > 0$ ;  $y \geq 0$  when  $\xi \geq 0$ , and  $0 \leq y \leq -\frac{\beta}{\xi}$  when  $\xi < 0$ . The theory of Balkema et al. (1974) and Pickands (1975) implies that if  $u$  is chosen close enough to  $x_0$ , the excess distribution of  $X$ , or the tail of the distribution of  $X$  converges to GPD. Loosely speaking, if  $u$  is chosen large enough,  $F_u(y) = G_{\xi,\beta(u)}(y)$  (McNeil and Frey, 2000).

Thus, given  $X > u$ , the following can be obtained,

$$F(x) = (1 - F(u))G_{\xi,\beta}(x - u) + F(u) \quad (15)$$

where  $F(u)$  is the cumulative density of the chosen threshold  $u$  (can be obtained empirically).

The reason for using empirical data to estimate  $F(u)$  but not  $F(x)$  (for  $x > u$ ), is that empirical data tends to be sparse approaching to the tails. Considering the variable  $X$  the standardized residuals in Equation (3), Value at Risk and Expected Shortfall can be estimated once  $Z_\alpha$  [as defined in Equation (5)] is obtained by Equation (15). Volatility can be assumed to follow different processes such as GARCH(1,1). The threshold  $u$  must be chosen large enough so as to be close to  $x_0$  (the end point of a finite sample). However, for a finite sample,  $u$  needs to be chosen in a way such that there is a large enough sample in excess of  $u$ . The “large enough sample” generally needs to be larger than 50 observations (McNeil and Frey, 2000).

## Literature Review for Value at Risk and Expected Shortfall Model Comparisons

Huang and Lin (2004) examined a number of Value at Risk models including EWMA-Normal Distribution (RiskMetrics), APARCH-Normal Distribution, APARCH-Student T Distribution, and concluded that the APARCH-Normal (Distribution) model generates most accurate Value at Risk for financial asset returns at the lower confidence level (95%). At the higher confidence level (99%), Huang and Lin (2004) concluded that the APARCH-Student T (Distribution) model outperformed the rest of the models, using Taiwan Stock Exchange Index. Ünal (2011) compared the performance of Historical Simulation (HS), EWMA-Normal Distribution model (RiskMetrics) and the Extreme Value Theory (EVT) model using a variety of stock indices. The conclusion was that EVT had the best performance based on the numbers of exceedances. Ouyang (2009) reached the same conclusion in a similar study using Chinese stock index. Mitnik, Kuester and Paolella (2006) examined GARCH-Student T (Distribution) model, GARCH-skewed Student T (Distribution) model, GARCH-EVT model, and Historical Simulation (HS) using NASDAQ, and concluded that GARCH-EVT is the best model based on the number of exceedances in the out of sample test, followed by GARCH-GED.

Ozun, Cifter and Yimazer (2010) examined the performance of a semi-parametric Value at Risk and Expected Shortfall model (EVT) and parametric models (GARCH and FIGARCH- Student T) and concluded semi-parametric models outperform parametric models using ISE100 Index. McNeil and Frey (2000) reached the same conclusion as Ozun, Cifter and Yimazer (2010), in comparing EVT model, GARCH-Normal (Distribution) model and GARCH-Student T (Distribution) model using a variety of

financial data including stock index, individual stock, exchange rate and commodity returns.

## DATA

Given the objective of this study is to analyze the accuracy and reliability of Value at Risk and Expected Shortfall models, the principles of choosing data in this study are (1) to include a variety of financial assets (equities, bonds, and commodities) that are exposed to daily market risks, (2) to include a sufficiently long time period of data for each of the selected markets, and (3) to ensure all data have the same start and end date for out of sample data. Based on these principles, the data used in this study include S&P 500 Index, crude oil price, gold price and the Vanguard Long Term Bond Fund (Vanguard Long-Term Bond Index Fund, VBLTX : 41.11% government bonds, 51.37% corporate bonds and 0.47% asset backed securities. Average duration is 14.2 years).

S&P 500 daily prices (adjusted close<sup>2</sup>) from January, 1950 to April, 2012 and Vanguard Long Term Bond Fund (VBLTX) daily prices (adjusted close) from March, 1994 to April, 2012 are obtained from Yahoo Finance website. WTI (West Texas Intermediate) Crude oil daily prices from January, 1986 to April, 2012 are obtained from Energy Information Administration site. Gold (London Bullion Market) daily price from April, 1968 to April, 2012 are obtained from Federal Reserve Bank of St. Louis website. The same 1000 days (approximately 4 years from May, 2008 to April, 2012) are used as the out of sample data.

The daily returns are calculated by taking log differences of observations on two adjacent trading days. The descriptive statistics in Table 2.1 suggest some common properties of the calculated daily log returns. All four data sets have close to zero mean and negative skewness. The high kurtosis in all four groups of data suggests fatter tails or extreme changes. All computations in this study are conducted using statistical software R<sup>3</sup> (2.15.1) (end note 3) with RUGARCH<sup>4</sup> (end note 4) package and Microsoft Excel.

## METHODS

In this study, three approaches are used for estimating Value at Risk (VaR) and Expected Shortfall (ES) (Appendix B). They are (1) parametric approach (11 models), (2) non-parametric approach (a single model, Historical Simulation), and (3) semi-parametric approach (a single Extreme Value Theory model). The parametric approach includes the widely used and most studied models (based on Normal, Student T and General Error Distribution). In addition, this study also uses a new parametric model, ARMA(1,1)-GJR-GARCH(1,1)-SGED (Skewed General Error Distribution), aiming to improve the accuracy of existing Value at Risk (VaR) and Expected Shortfall (ES) models. The next section discusses the proposed new parametric model in detail.

### The Proposed Model: ARMA(1,1)-GJR-GARCH(1,1)-SGED

In addition to the widely used and most studied models under the three approaches of Value at Risk (VaR) and Expected Shortfall (ES) models, this study uses a

new modified parametric model, ARMA(1,1)-GJR-GARCH(1,1)-SGED, based on the model (ARMA(1,1)-GARCH(1,1)-Normal) by Berkowitz and O'Brien (2002). This model is innovative in the following three aspects. First, it captures the autocorrelation in returns using ARMA(1,1) process. Second, it employs GJR-GARCH(1,1) to estimate one day forward volatility, and captures the leverage effect (Black, 1976) in returns. Third, it uses a skewed fat tail distribution, skewed General Error Distribution, to model the extreme tails of daily returns of the selected financial assets.

#### *Using ARMA(1,1) to Capture Autocorrelation in Returns*

The demeaned ARMA(1,1) process is used to capture possible autocorrelation in the return time series data. Berkowitz and O'Brien (2002) find that the ARMA(1,1)-GARCH(1,1) can achieve similar performance to the multi-variate models. Tang, Chiu and Xu (2003) also showed that ARMA-GARCH combination produces better results than GARCH alone for stock price prediction purpose. The demeaned ARMA(1,1) process for returns can be expressed as follows,

$$R_t = \mu + \Phi(R_{t-1} - \mu) + \varepsilon_t + \theta\varepsilon_{t-1} \quad (16)$$

where  $\Phi$  and  $\theta$  are respectively the first order AR and MA parameters to be estimated;  $\varepsilon_t = \sigma_t X_t$ , where  $X_t$  and  $\sigma_t$  follow the definition in Equation (3) (Note that the log return on day t,  $R_t$ , is no longer a liner transformation of  $X_t$ ).  $\mu$  is the constant mean of returns. The  $(1-\alpha)\%$  VaR on day t is defined as below,

$$VaR_{t,1-\alpha} = \mu + \Phi(R_{t-1} - \mu) + \sigma_t Z_\alpha + \theta\sigma_{t-1}x_{t-1} \quad (17)$$

where  $Z_\alpha$  is the  $\alpha$  percentile of  $D(0,1)$ ;  $x_{t-1} = \varepsilon_{t-1} / \sigma_{t-1}$ .

Accordingly, under ARMA(1,1)-GARCH(1,1), the Expected Shortfall is defined as below

$$ES_{t,1-\alpha} = \mu + \Phi(R_{t-1} - \mu) + \sigma_t E[X|X < Z_\alpha] + \theta \sigma_{t-1} x_{t-1} \quad (18)$$

#### *Using GJR-GARCH(1,1) to Estimate One Day Ahead Volatility*

The proposed model uses GJR-GARCH(1,1) [Equation (12)] to capture the leverage effect (volatility asymmetry) in daily returns. For selecting the appropriate volatility model, Sign Bias Test (Engle and NG, 1993) is conducted as below. The methodology of this test is to regress the squared standardized residuals jointly on lagged positive and negative standardized residuals, as shown in the equation below,

$$\hat{x}_t^2 = c_0 + c_1 I_{\hat{x}_{t-1} < 0} + c_2 I_{\hat{x}_{t-1} < 0} \hat{x}_{t-1} + c_3 I_{\hat{x}_{t-1} \geq 0} \hat{x}_{t-1} + \epsilon_t \quad (19)$$

where  $\hat{x}_t$  is the filtered standardized residual (the innovation variable) on day t;

$c_0, c_1, c_2$  and  $c_3$  are the coefficients to be estimated;  $\epsilon_t$  is the error term at time t; the indicator functions are defined as

$$I_{\hat{x}_{t-1} < 0} = \begin{cases} 1, & \hat{x}_{t-1} < 0 \\ 0, & \hat{x}_{t-1} \geq 0 \end{cases}$$

and

$$I_{\hat{x}_{t-1} \geq 0} = \begin{cases} 1, & \hat{x}_{t-1} \geq 0 \\ 0, & \hat{x}_{t-1} < 0 \end{cases}$$

The estimated coefficients are then tested separately [ $H_0: c_i = 0$  ( $i = 1,2,3$ )] and jointly

( $H_0: c_1 = c_2 = c_3 = 0$ ). Significant coefficients from the test would suggest that there is leverage effect in the residuals, and that an asymmetric volatility model should be used.

Table 2.2 reports Sign Bias Test results for GARCH(1,1), which suggest that there is strong evidence (small p value) that both negative and positive sign biases exist in the standardized residuals of S&P 500 and crude oil data. On the other hand, the test results for GJR-GARCH(1,1), as reported in Table 2.3, show that there is no evidence for the negative or positive sign biases in the standardized residuals for S&P 500. The full parameter estimates for ARMA(1,1)-GARCH(1,1) and ARMA(1,1)-GJR-GARCH(1,1) are reported in Table 2.4 and Table 2.5. Based on the Sign Bias Test results, GJR-GARCH(1,1) is selected for the proposed new parametric model to better suit the S&P 500 data. Based on a study by Hansen and Lunde (2005) that concluded that GJR-GARCH is superior to a wide range of asymmetric volatility models, including EGARCH (Nelson, 1991), APARCH (Ding et al., 1993) and FGARCH (Hentschel, 1995) in terms of accuracy, GJR-GARCH is the only asymmetric model included in this study.

#### *Using Skewed General Error Distribution to Model Standardized Residuals of Returns*

Kuester, Mittnik and Paolella (2006) used a number of distributions including Normal, Student T and General Error Distribution (GED<sup>5</sup> end note 5), to pair with ARMA-GARCH. The conclusion is that ARMA-GARCH-GED has the best out of sample results for Value at Risk. Inspired by a Kuester et al (2006) study, this study introduces a skewed General Error Distribution to pair with ARMA(1,1)-GJR-GARCH(1,1) to capture the negative skewness in the data (all four groups of data have

negative skeweness as shown in Table 2.1). The method of Fernandez and Steel (1998) is used to introduce a skewness parameter into the pdf of GED, as follows. Considering a random variable  $X$  (pdf  $f(x)$ ), which is unimodal and symmetric about zero, or formally,  $f(x) = f(|x|)$ . Then the pdf of the skewed distribution is then given by introducing an inverse scale factor to the original pdf,

$$p\left(\frac{x}{\gamma}\right) = \frac{2}{\gamma + \frac{1}{\gamma}} \left\{ f\left(\frac{x}{\gamma}\right) I_{x \geq 0} + f(x\gamma) I_{x < 0} \right\} \quad (20)$$

where  $\gamma$  is the skewness parameter [to be estimated; when  $\gamma = 1$ ,  $p\left(\frac{x}{\gamma}\right) = f(x)$ ];  $f(x)$  is the pdf of General Error Distribution (End note 5). The indicator function  $I$  is defined as follows,

$$I_{x \geq 0} = \begin{cases} 1, & x \geq 0 \\ 0, & x < 0 \end{cases}$$

and

$$I_{x < 0} = \begin{cases} 1, & x < 0 \\ 0, & x \geq 0 \end{cases}$$

### Examining Goodness of Fit of Distributions Assumed by Parametric Models

This section examines the goodness of fit of distributions assumed in parametric models using the Adjusted Pearson Goodness of Fit Test (Palm, 1996). The null hypothesis of the adjusted Pearson Goodness of Fit is that the observations in the sample being tested are from a specific distribution. The test statistic is



$$X^2 = \sum_{i=1}^m \frac{(O_i - E_i)^2}{E_i} \quad (21)$$

where  $O_i$  is the observed frequency of a range of observations in the sample;  $E_i$  is the expected frequency of the same range of the known distribution;  $m$  is the number of bins that the distribution has been divided into. If the null hypothesis is true, then  $X^2$  asymptotically approaches to a chi-square Distribution with degree freedom  $n$  (number of estimated parameters). The data used here is the standardized residuals [the variable  $X$  in Equation (3)]. The adjusted Pearson Goodness of Fit test is conducted twice using two different volatility models, GARCH(1,1) and GJR-GARCH(1,1).

The test statistics and  $p$  values are summarized in Table 2.6 for GARCH(1,1) and Table 2.7 for GJR-GARCH(1,1). The results ( $p$  values) suggest that the skewed Student T Distribution has the strongest goodness of fit for crude oil data. The skewed General Error Distribution (SGED) and skewed Student T Distribution (SSTD) have the strongest goodness of fit for S&P 500. None of the distributions has strong goodness of fit for gold and Vanguard Long Term Bond Fund based on the test results. To visually examine the goodness of fit of the five distributions (Normal, Student T, skewed Student T, General Error Distribution, Skewed General Error Distribution) assumed by parametric models, Figure 2.1-2.5 show QQ plots of these distributions to data (S&P 500 is used for illustration purpose). Generally on QQ plots, goodness of fit is measured by the distances from hollow dots (represent the data) to the straight line (represents the theoretical quantiles of underlying distribution), and being closer the straight line indicates a better fit. Skewed Student T Distribution (SSTD) and skewed General Error Distribution

(SGED) show stronger goodness of fit compared to the other three, based on visual examination.

#### Examining Goodness of Fit of General Pareto Distribution in the Semi-Parametric Model (Extreme Value Theory)

This section discusses the goodness of fit of the General Pareto Distribution used by semi-parametric model (Extreme Value Theory). It has been mentioned in the previous section that the first step in applying the EVT model is choosing a threshold  $\mu$ . It has also been emphasized that  $\mu$  has to be large (or small if left tail is to be estimated) so it is close enough to the end point of empirical distribution yet small (or large) enough to leave a sufficient number of observations for estimation. In this study, the 5% empirical percentile (of the standardized residuals of returns) is chosen, which means the smallest (largest loss wise) 5% of observations are used for EVT estimation. Standardized residuals are obtained by assuming GARCH(1,1) process for returns. The parameter estimation is done by Maximum Likelihood Estimation Method. Using the S&P 500 data for an example, the reverse signed in sample standardized residuals ( $14,683 = 15,683$  [total data points] – 1,000 [out of sample data]) are sorted from smallest to the largest and  $\mu$  is chosen such that exact 5% from the top of the observations or 734 observations exceed  $\mu$ . The 734 observations are then used for GPD parameter estimation by Maximum Likelihood Estimation Method. The estimated parameters,  $\xi$  (the shape parameter) and  $\beta$  (the scaling parameter) are reported in Table 2.8. The estimated  $\xi$  is greater than zero for all four groups of data, indicating high kurtosis (generally speaking, a greater-than-zero  $\xi$  indicates a heavy tailed data.).

To visually examine the tail goodness of fit of the estimated GPD to data, the theoretical tail of estimated GPD is plotted against the data (S&P 500, crude oil, gold and Vanguard Long Term Bond Fund) in Figure 2.6-2.9. For the purpose of Value at Risk estimation, it is desirable that the dots fall on or under the solid curve. Figure 2.6-2.9 show that the tail of the estimated GPD fit the data tail well and the underestimation (estimated is less than data) only happens at the far left tail.

[Goodness of fit for the Historical Simulation method (non-parametric approach) is not included, as this method does not contain any distribution assumptions, therefore goodness of fit cannot be examined].

### Out of Sample Test Procedures for Value at Risk and Expected Shortfall

#### *Out of Sample Test Procedure for Value at Risk*

In the previous sections, statistical tests and various plots are used to examine the goodness of fit of the parametric and semi-parametric models. In this section, the out of sample test is conducted to examine model performance against past realized returns. For example, the model is constructed using earlier in sample data, then tested using more recent out of sample data, that the model has not seen. The general procedure for a dynamic out of sample test can be summarized as follows [For non-parametric model – Historical Simulation method, skip step 2, 3, 6, and 7. The semi-parametric model – Extreme Value Theory model does re-estimate model parameters as described in step 6.]

- 1) Given the log return data  $(R_1, R_2, R_3, \dots, R_n)$ , determine in sample  $(R_1, R_2, R_3, \dots, R_{t-1})$  and out of sample data  $(R_t, R_{t+1}, R_{t+2}, \dots, R_n)$ .
- 2) Estimate parameters of the chosen Value at Risk models using in sample data  $(R_1, R_2, R_3, \dots, R_{t-1})$ . [This step applies to parametric and semi-parametric models only.]
- 3) Estimate one day forward volatility  $\sigma_t$  using the selected conditional variance models [GARCH(1,1) or GJR-GARCH(1,1)]. [This step applies to parametric models only.]
- 4) Calculate one-day  $VaR_{t,1-\alpha}$
- 5) Measure  $R_t$  against  $VaR_{t,1-\alpha}$ . If  $R_t < VaR_{t,1-\alpha}$ , then count one exceedance.
- 6) Re-estimate parameters with the new in sample data  $(R_1, R_2, R_3, \dots, R_t)$ . [This step applies to parametric models only.]
- 7) Repeat step 3 – 6 for  $n - t + 1$  times.
- 8) The total number of exceedances is then compared to the expected number of exceedances at the given level of confidence.

In this study, the same number of out of sample data points, 1000, is chosen to create comparability across data for all three financial markets. Also, choosing a relatively large size of out of sample data will ensure more robust results.

A deviation to the general procedure above is that instead of re-estimating model parameters daily, re-estimation is done every 25 days. This is because the initial in sample size is large, and re-estimation with one addition of in sample data point is not important. For example, the S&P 500 data has 15683 observations and therefore 14683 initial in sample observations. It is unlikely that re-estimating with one additional

observation would produce significantly different parameters. For the consistent out of sample sizes (1000), exactly 40 parameter estimations will be conducted for each of the four data sets.

#### *Out of Sample Test Procedure for Expected Shortfall (ES)*

Expected Shortfall (ES), on the other hand, is a measure of expected values which are not directly comparable to empirical data. Therefore, the Value at Risk out of sample test cannot be used to examine Expected Shortfall models. Instead, the method introduced by McNeil and Frey (2000) is used to test Expected Shortfall. The procedure is introduced below.

In the case of a Value at Risk exceedance (i.e.  $R_t < VaR_{t,1-\alpha}$ ), define a new residual  $z_t$ , such that

$$z_t = x_t - E[X|X < Z_\alpha] \quad (22)$$

where  $x_t$  is the (GARCH or GJR-GARCH) filtered standardized residual on day  $t$ ;  $Z_\alpha$  is  $\alpha$  percentile of distribution  $X_t$  (since  $X_t$  is i.i.d. as defined in Equation (3),  $X$  is independent of  $t$ ).

Then by Equation (16) and (18),

$$x_t = \frac{R_t - \mu - \Phi(R_{t-1} - \mu) - \theta \sigma_{t-1} x_{t-1}}{\sigma_t} \quad (23)$$

$$E[X|X < Z_\alpha] = \frac{ES_{t,1-\alpha} - \mu - \Phi(R_{t-1} - \mu) - \theta \sigma_{t-1} x_{t-1}}{\sigma_t} \quad (24)$$

Now, if the estimated Expected Shortfall (by the underlying model) is accurate, then the new residual  $z_t$  should be an independently and identically distributed variable with zero mean. The one sided t test is conducted to test the null hypothesis,

$$H_0: \mu_z = 0 \text{ Versus } H_1: \mu_z < 0$$

where  $\mu_z$  is the mean of new residual  $z_t$ .

One sided t test is used because risk managers are mainly interested in detecting underestimation of Expected Shortfall (the absolute value of estimated Expected Shortfall is smaller than the absolute value of the true Expected Shortfall).

#### *Procedures for Measuring Exceedances and Confidence Interval for Value at Risk*

A confidence interval (based on Z test) is introduced here to facilitate the exceedance test. The confidence intervals are calculated assuming the total number of exceedances is random variable from a Binomial Distribution. That is,  $v \sim B(Np, Npq)$ , where  $p = (1-\alpha)\%$  and  $q = \alpha\%$ ;  $N$  is the total number of out of sample observations. Then by central limit theory,  $v$  is asymptotically normally distributed, or  $\frac{v-Np}{\sqrt{Npq}} \sim N(0, 1)$ . The  $(1-\pi)\%$  confidence interval for number of exceedances over  $(1-\alpha)\%$  level Value at Risk is then  $(Np - Z_{1-\frac{\pi}{2}}\sqrt{Npq}, Np + Z_{1-\frac{\pi}{2}}\sqrt{Npq})$ . The 95% confidence interval is used in the Value at Risk out of sample test because it is not only one of the most commonly used confidence level (Best, 1999), but also the standard that Basel Committee adopts in accessing a Value at Risk models (Basel II, 2006). Since the size of out of sample

observations is 1000, the expected number of exceedances is 10 for 99% confidence level and 50 for 95% confidence level.

## OUT OF SAMPLE TEST RESULTS

### Value at Risk Out of Sample Test Results for All Models

Table 2.9 reports the Value at Risk out of sample test results. The numbers of exceedances that fall outside the 95% confidence interval are shown with asterisk. Overall speaking, the proposed parametric model (ARMA(1,1)-GJR-GARCH(1,1)-SGED) is the most balanced model, as it is the only model for which the numbers of exceedances fall within 95% confidence interval (for the Z test) for all four markets in this study. Figure 2.10-2.17 show plots of the estimated Value at Risk versus actual out of sample returns using the proposed model (eight plots for four markets at two confidence levels). Therefore, it is considered an adequate VaR model for all four markets.

For S&P 500, the EVT model is superior to all other models in terms of accuracy, due to the smallest number of exceedances at both confidence levels. The only other model that passes the Z test at 95% level is the proposed parametric model (GJR-GARCH-SGED). The Normal (Distribution) based models (EWMA-Normal, GARCH-Normal and GJR-GARCH-Normal) and Historical Simulation (HS) have the largest numbers of exceedances and therefore underestimate Value at Risk.

For crude oil, all models except for EVT performed reasonably well with all numbers of exceedances falling within the confidence intervals. This could be because

the crude oil data has relatively low kurtosis (Table 2.1). For gold, the most accurate models are GARCH-GED and GARCH-SGED based (Appendix B) on the numbers of exceedances, followed closely by GJR-GED and GJR-SGED. Normal (Distribution) based parametric models and EVT all failed the Z test at 5% level ( $\alpha$ ) for Value at Risk (numbers of exceedances fall outside the confidence interval).

For Vanguard Long Term Bond Fund, most parametric models are statistically adequate for estimating Value at Risk [except for Student T and Skewed Student T based models (Appendix B) at 95% confidence level]. The Historical Simulation method (HS) underestimated risk evidenced by the large number of exceedances at both confidence levels. Extreme Value Theory (EVT, semi-parametric) overestimated Value at Risk due to the small number of exceedances.

#### Expected Shortfall Out of Sample Test Results for All Models

Table 2.10 reports out of sample test results for Expected Shortfall. For S&P 500 and 95% Expected Shortfall, the null hypothesis (the Expected Shortfall residual has a zero mean) is rejected at 5% level for five models, EWMA-Normal, ARMA(1,1)-GARCH(1,1)-Normal, ARMA(1,1)-GJR-GARCH(1,1)-Normal, ARMA(1,1)-GARCH(1,1)-GED and HS model. For 99% Expected Shortfall, the null hypothesis is rejected at 5% level for ARMA(1,1)-GARCH(1,1)-Normal and HS models.

For crude oil, the hypothesis is also rejected (at 5% level) for ARMA(1,1)-GJR-GARCH(1,1)-Normal and HS for Expected Shortfall confidence level of 99%. For gold, the null hypothesis cannot be rejected at 5% level using non-Normal (Distribution) based



parametric models. Normal (Distribution) based parametric models and HS failed the t test at 5% level. For Vanguard Long Term Bond Fund, EWMA-Normal (Distribution) failed the t test for 95% Expected Shortfall. It is worth to mention that the size of sample data for EVT is small for this test, especially for Expected Shortfall with confidence level of 99%. This is due to the overestimation of EVT for Value at Risk, and the returns that exceed Value at Risk are too few.

## CHAPTER SUMMARY

In this Chapter, three main Value at Risk (VaR) and Expected Shortfall (ES) approaches (parametric, semi-parametric and non-parametric), were analyzed. Then a modified parametric model was proposed, which uses ARMA(1,1) process, the GJR-GARCH(1,1) volatility model and skewed General Error Distribution (SGED). These models were then compared for accuracy using out of sample test. Based on the out of sample results, the proposed parametric model (ARMA(1,1)-GJR-GARCH(1,1)-SGED) is the most balanced model for both Value at Risk (VaR) and Expected Shortfall (ES) based on out of sample results. It passed all Z tests at 95% confidence level for Value at Risk and T tests for Expected Shortfall for all four markets. Figure 2.10-2.17 show plots for estimated Value at Risk versus actual out of sample returns using the proposed model (four markets and two confidence levels).

For S&P 500 data only, The EVT model had the smallest number of Value at Risk exceedances and therefore outperformed all other models. However, it overestimated Value at Risk for crude oil, gold and Vanguard Long Term Bond Fund, which also

resulted in the small sample sizes in the Expected Shortfall out of sample test. The Normal Distribution based parametric models (Appendix B) have the most unreliable performance in both Value at Risk (VaR) and Expected Shortfall (ES) out of sample tests. They mostly underestimated risk for all markets.

The results above should be of value to researchers, risk managers, regulators, and analysts for selection of Value at Risk (VaR) and Expect shortfall (ES) models. Based on the results of this study, the semi-parametric model, Extreme Value Theory (EVT), is recommended for estimating Value at Risk for S&P 500 Index, or similar indices, likely due to fat tail behavior for both in sample and out of sample data. As an overall model for Value at Risk (VaR) and Expect shortfall (ES) for equity, bond and commodity markets, the proposed parametric model (AMRA(1,1)-GJR-GARCH(1,1)-SGED) is recommended based on the results of this study. In all cases, the Normal Distribution based parametric models (Appendix B) and Historical Simulation models are inadequate for estimating Value at Risk (VaR) and Expect shortfall (ES).

## END NOTES

---

<sup>1</sup>  $R_t = \log P_t - \log P_{t-1}$ , where  $P_t$  is the price of some financial asset on day t and  $P_{t-1}$  is the price of same financial asset on day t-1.

<sup>2</sup> Adjusted close: the pre-dividend (coupon) data is adjusted to exclude the dividend (coupon) payout. <https://help.yahoo.com/kb/finance/historical-prices-sln2311.html?impressions=true>

<sup>3</sup> A statistical freeware <http://www.r-project.org/>

<sup>4</sup> A specialized package for R <http://cran.r-project.org/web/packages/rugarch/index.html>

<sup>5</sup> PDF of Generalized Error Distribution (Giller, 2005):

$$f(x) = \frac{\kappa e^{-0.5 \left| \frac{x-\alpha}{\beta} \right|^\kappa}}{2^{1+\kappa^{-1}} \beta \Gamma(\kappa^{-1})}$$

where  $\alpha, \beta, \kappa$  are respectively the location, scale and shape parameter.  
 $\Gamma(\cdot)$  is the gamma function.

**Table 2.1 Descriptive Statistics and Sources of Data (Daily Return)**

<b>Data</b>	<b># Obs</b>	<b>Mean</b>	<b>Std-dev</b>	<b>Min</b>	<b>Max</b>	<b>Skewness</b>	<b>Excess Kurtosis**</b>	<b>Source</b>
S&P 500	15683	0.0003	0.0098	-0.2290	0.1096*	-1.0384	27.83	Yahoo Finance
Crude Oil	6640	0.0002	0.0259	-0.4064	0.1915	-0.7668	14.52	EIA
Gold	11232	0.0003	0.0135	-0.3756	0.3589	-0.1154	104.22	FRED
Vanguard Long Term Bond Fund	4576	0.0003	0.0062	-0.0438	0.0324	-0.1004	1.82	Yahoo Finance

Notes: Obs = observations;

Yahoo Finance website <http://finance.yahoo.com/>

EIA = US Energy Information Administration, website <http://www.eia.gov/>

FRED = Federal Reserve Bank of St. Louis, website <http://research.stlouisfed.org/>

\*For example, the maximum daily return of S&P 500 is 10.96%.

\*\*Kurtosis values greater than zero indicate positive kurtosis.

Data: S&P 500 daily return (adj close) from 01/01/1950 to 04/30/2012; Crude oil price daily return from 01/02/1986 to 04/30/2012; Gold daily return from 04/02/1968 to 04/30/2012; Vanguard Long Term Bond Fund (Vanguard Long-Term Bond Index Fund, VBLTX), daily return (adj close) from 03/01/1994 to 04/30/2012 (41.11% government bonds, 51.37% corporate bonds and 0.47% asset backed securities. Average duration is 14.2 years).

**Table 2.2 Leverage Effect Test: Sign Bias Test Results for ARMA(1,1)-GARCH(1,1)  
(Daily Returns)**

Market	Test	t value	p value
S&P 500			
	Sign Bias ( $c_1$ )	0.6093	0.5420
	Negative Sign Bias ( $c_2$ )	3.7870	0.0002
	Positive Sign Bias ( $c_3$ )	2.6935	0.0071
	Joint Effect	46.0427	0.0000
Crude Oil			
	Sign Bias ( $c_1$ )	0.1537	0.8780
	Negative Sign Bias ( $c_2$ )	3.6892	0.0002
	Positive Sign Bias ( $c_3$ )	2.2421	0.0250
	Joint Effect	31.4795	0.0000
Gold			
	Sign Bias ( $c_1$ )	0.3191	0.7497
	Negative Sign Bias ( $c_2$ )	1.1457	0.2519
	Positive Sign Bias ( $c_3$ )	0.2310	0.8173
	Joint Effect	2.1441	0.5430
Vanguard Long Term Bond Fund			
	Sign Bias ( $c_1$ )	0.4668	0.6406
	Negative Sign Bias ( $c_2$ )	0.7504	0.4530
	Positive Sign Bias ( $c_3$ )	1.0716	0.2840
	Joint Effect	2.9891	0.3933

Notes: The purpose of this test is to detect leverage effect (asymmetric volatility) in the residuals (return mean residual divided by standardized deviation) assuming underlying volatility model and distribution. Small p values suggest existence of leverage effect, which was found for much of the S&P 500 and crude oil data.

Data: S&P 500 daily return from 01/01/1950 to 04/30/2012; Crude oil price daily return from 01/02/1986 to 04/30/2012; Gold daily return from 04/02/1968 to 04/30/2012; Vanguard Long Term Bond Fund (VBLTX), daily return from 03/01/1994 to 04/30/2012.

**Table 2.3 Leverage Effect Test: Sign Bias Test Results for ARMA(1,1)-GJR-GARCH(1,1) (Daily Return)**

Market	Test	t value	p value
S&P 500			
	Sign Bias ( $c_1$ )	0.9838	0.3252
	Negative Sign Bias ( $c_2$ )	1.4471	0.1479
	Positive Sign Bias ( $c_3$ )	1.5222	0.1280
	Joint Effect	14.8191	0.0020
Crude Oil			
	Sign Bias ( $c_1$ )	3.8670	0.0001
	Negative Sign Bias ( $c_2$ )	1.1040	0.2698
	Positive Sign Bias ( $c_3$ )	2.8150	0.0049
	Joint Effect	16.1790	0.0010
Gold			
	Sign Bias ( $c_1$ )	0.1860	0.8525
	Negative Sign Bias ( $c_2$ )	2.0545	0.0400
	Positive Sign Bias ( $c_3$ )	0.0061	0.9952
	Joint Effect	5.9358	0.1148
Vanguard Long Term Bond Fund			
	Sign Bias ( $c_1$ )	0.4233	0.6721
	Negative Sign Bias ( $c_2$ )	0.7694	0.4417
	Positive Sign Bias ( $c_3$ )	0.9986	0.3180
	Joint Effect	2.5649	0.4637

Notes: The purpose of this test is to detect leverage effect (asymmetric volatility) in the standardized residuals (return mean residual divided by standardized deviation) of returns, assuming underlying volatility model and distribution; Small p values suggest existence of leverage effect; Leverage effect is eliminated completely for S&P 500 and partially for crude oil, when replacing GARCH(1,1) with GJR-GARCH(1,1) for its asymmetric nature.

Data: S&P 500 daily return from 01/01/1950 to 04/30/2012; Crude oil price daily return from 01/02/1986 to 04/30/2012; Gold daily return from 04/02/1968 to 04/30/2012; Vanguard Long Term Bond Fund (VBLTX), daily return from 03/01/1994 to 04/30/2012.

**Table 2.4 Parameter Estimates for ARMA(1, 1)-GARCH(1, 1) (Daily Return)**

Market	Parameters	Estimate	Std. Error	t value	p value
S&P 500					
	$\mu$	0.0005	0.0000	6.6616	0.0000
	$\phi$	-0.1609	0.1146	-1.4044	0.1602
	$\theta$	0.2644	0.1101	2.4005	0.0164
	$\beta_0$	0.0000	0.0000	4.4824	0.0000
	$\beta_1$	0.0830	0.0126	6.5969	0.0000
	$\beta_2$	0.9114	0.0119	76.3090	0.0000
Crude Oil					
	$\mu$	0.0006	0.0001	4.5094	0.0000
	$\phi$	-0.1436	0.1188	-1.2089	0.2267
	$\theta$	0.1576	0.1229	1.2827	0.1996
	$\beta_0$	0.0000	0.0000	1.5845	0.1131
	$\beta_1$	0.0923	0.0417	2.2114	0.0270
	$\beta_2$	0.8996	0.0418	21.5059	0.0000
Gold					
	$\mu$	0.0001	0.0001	0.9010	0.3676
	$\phi$	-0.7109	0.0525	-13.5438	0.0000
	$\theta$	0.7406	0.0383	19.3445	0.0000
	$\beta_0$	0.0000	0.0000	1.0724	0.2835
	$\beta_1$	0.0946	0.0147	6.4568	0.0000
	$\beta_2$	0.9044	0.0213	42.5261	0.0000
Vanguard Long Term Bond Fund					
	$\mu$	0.0003	0.0001	4.7679	0.0000
	$\phi$	0.7927	0.0679	11.6682	0.0000
	$\theta$	-0.8245	0.0629	-13.1000	0.0000
	$\beta_0$	0.0000	0.0000	1.4401	0.1498
	$\beta_1$	0.0354	0.0026	13.4456	0.0000
	$\beta_2$	0.9548	0.0025	376.9083	0.0000

Notes: Table shows parameter estimates and their statistical significance. ARMA(1,1) parameters ( $\phi$  and  $\theta$ ) are shown to be most significant for Gold and Vanguard Long Term Bond Fund. GARCH parameters ( $\beta_0$ ,  $\beta_1$  and  $\beta_2$ ) are shown to be mostly significant in all four markets.

Data: S&P 500 daily return from 01/01/1950 to 04/30/2012; Crude oil price daily return from 01/02/1986 to 04/30/2012; Gold daily return from 04/02/1968 to 04/30/2012; Vanguard Long Term Bond Fund (VBLTX), daily return from 03/01/1994 to 04/30/2012.

**Table 2.5 Parameter Estimates for ARMA(1, 1)-GJR-GARCH(1, 1) (Daily Returns)**

Market	Parameters	Estimate	Std. Error	t value	p value
S&P 500					
	$\mu$	0.0003	0.0001	2.5879	0.0097
	$\phi$	0.7776	0.0461	16.8795	0.0000
	$\theta$	-0.7907	0.0455	-17.3595	0.0000
	$\beta_0$	0.0000	0.0000	2.8215	0.0048
	$\beta_1$	0.0103	0.0102	1.0079	0.3135
	$\beta_2$	0.9103	0.0226	40.3017	0.0000
	$\beta_{1-}$	0.1256	0.0345	3.6365	0.0003
Crude Oil					
	$\mu$	0.0004	0.0002	1.7641	0.0777
	$\phi$	0.8534	0.0597	14.2933	0.0000
	$\theta$	-0.8807	0.0551	-15.9761	0.0000
	$\beta_0$	0.0000	0.0000	2.9261	0.0034
	$\beta_1$	0.0963	0.0183	5.2671	0.0000
	$\beta_2$	0.9019	0.0163	55.3443	0.0000
	$\beta_{1-}$	-0.0063	0.0116	-0.5443	0.5862
Gold					
	$\mu$	0.0001	0.0001	1.3355	0.1817
	$\phi$	-0.7380	0.0493	-14.9756	0.0000
	$\theta$	0.7631	0.0386	19.7473	0.0000
	$\beta_0$	0.0000	0.0000	1.1339	0.2568
	$\beta_1$	0.1056	0.0161	6.5595	0.0000
	$\beta_2$	0.9137	0.0201	45.5324	0.0000
	$\beta_{1-}$	-0.0405	0.0123	-3.2844	0.0010
Vanguard Long-term Bond Fund					
	$\mu$	0.0003	0.0001	4.9467	0.0000
	$\phi$	0.7990	0.0666	12.0036	0.0000
	$\theta$	-0.8305	0.0613	-13.5482	0.0000
	$\beta_0$	0.0000	0.0000	1.4207	0.1554
	$\beta_1$	0.0365	0.0037	9.8203	0.0000
	$\beta_2$	0.9567	0.0019	510.5048	0.0000
	$\beta_{1-}$	-0.0040	0.0060	-0.6616	0.5082

Notes: Table shows parameter estimates and their statistical significance. ARMA(1,1) parameters ( $\phi$  and  $\theta$ ) are shown to be most significant for Gold and Vanguard Long Term Bond Fund. GARCH parameters ( $\beta_0$ ,  $\beta_1$  and  $\beta_2$ ) are shown to be mostly significant in all four markets. The GJR parameter,  $\beta_{1-}$ , is significant for S&P 500 and gold.

Data: S&P 500 daily return from 01/01/1950 to 04/30/2012; Crude oil price daily return from 01/02/1986 to 04/30/2012; Gold daily return from 04/02/1968 to 04/30/2012; Vanguard Long Term Bond Fund (VBLTX), daily return from 03/01/1994 to 04/30/2012.



**Table 2.6 Pearson Goodness-of-Fit Statistic Test Results for Five Distributions, ARMA(1,1)-GARCH(1,1)  
(Daily Return)**

Market		S&P 500		Crude Oil		Gold		Vanguard Long Term Bond Fund	
Test #	bin	statistic	p value	statistic	p value	statistic	p value	statistic	p value
Normal									
1	20	282.8	6.64E-49	121.0	7.18E-17	584.2	1.10E-111	578.4	1.85E-110
2	30	323.7	1.58E-51	129.3	1.28E-14	647.8	7.54E-118	632.6	1.11E-114
3	40	350.2	1.14E-51	153.3	1.80E-15	726.0	1.91E-127	623.7	1.93E-106
4	50	373.5	1.69E-51	170.8	2.16E-15	756.1	2.02E-127	708	1.21E-117
Student T									
1	20	64.0	8.94E-07	27.4	0.09602	140.2	1.68E-20	91820	0
2	30	93.7	9.59E-09	32.9	0.28326	181.8	4.65E-24	142606	0
3	40	89.0	8.92E-06	38.9	0.47313	214.6	4.03E-26	193433	0
4	50	108.8	1.98E-06	56.6	0.21172	209.9	8.43E-22	244278	0
Skewed Student T									
1	20	48.3	0.0002305	15.8	0.6721	144.7	2.30E-21	91820	0
2	30	59.7	0.0006649	22.8	0.7869	180.5	8.00E-24	142606	0
3	40	70.9	0.0013501	29.9	0.8514	218.5	7.96E-27	193433	0
4	50	86.5	0.0007548	43.4	0.6973	216.1	7.47E-23	244278	0
GED									
1	20	67.5	2.37E-07	57.2	1.06E-05	72.6	3.36E-08	16679	0
2	30	80.2	1.09E-06	59.7	6.68E-04	95.4	5.18E-09	20939	0
3	40	88.0	1.21E-05	69.1	2.09E-03	120.5	2.99E-10	25431	0
4	50	104.1	7.50E-06	84.9	1.12E-03	128.7	4.49E-09	30640	0
Skewed GED									
1	20	43.3	0.001177	44.9	0.000705	70.2	8.54E-08	17724	0
2	30	43.6	0.039921	49.4	0.010549	92.9	1.30E-08	21129	0
3	40	57.0	0.031558	62.4	0.010184	123.4	1.09E-10	25631	0
4	50	77.0	0.0065	72.9	0.014959	132.0	1.53E-09	30938	0

**Table 2.6 (continued)**

Note: The purpose of this test is to examine the goodness of fit of the distributions assumed by parametric models. Test is conducted for five distribution assumptions for four markets, assuming daily returns follow ARMA(1,1) process, and the volatility follows GARCH(1,1) process. Data are grouped into 20, 30, 40 and 50 bins for each market, and the test is conducted independently for each of these bin numbers used. The results suggest that Skewed GED is the best fit for S&P 500 and skewed Student T Distribution is the best fit for crude oil. None of the distributions fits well with gold and Vanguard Long Term Bond Fund.

Data: S&P 500 daily return from 01/01/1950 to 04/30/2012; Crude oil price daily return from 01/02/1986 to 04/30/2012;  
Gold daily return from 04/02/1968 to 04/30/2012; Vanguard Long Term Bond Fund (VBLTX), daily return from 03/01/1994 to 04/30/2012.

**Table 2.7 Pearson Goodness-of-Fit Statistic Test Results for Five Distributions, ARMA(1,1)-GJR-GARCH(1,1) (Daily Return)**

Market		S&P 500		Crude Oil		Gold		Vanguard Long Term Bond Fund	
Test #	bin	statistic	p value	statistic	p value	statistic	p value	statistic	p value
Normal									
1	20	248.8	5.41E-42	125.7	9.50E-18	576.5	4.62E-110	579.6	1.01E-110
2	30	275.7	5.02E-42	136.0	8.64E-16	655.7	1.73E-119	623.4	9.06E-113
3	40	308.3	1.41E-43	163.5	3.47E-17	739.8	2.73E-130	628.8	1.72E-107
4	50	304.7	1.27E-38	177.3	1.99E-16	771.9	1.25E-130	722	1.80E-120
Student T									
1	20	52.0	6.70E-05	31.7	0.03422	143.5	3.93E-21	82597	0
2	30	72.5	1.35E-05	37.4	0.13628	178.8	1.69E-23	128281	0
3	40	79.3	1.46E-04	44.4	0.25593	208.3	5.53E-25	174002	0
4	50	94.3	1.07E-04	55.0	0.25878	211.3	4.75E-22	219740	0
Skewed Student T									
1	20	37.0	0.008061	14.9	0.7302	146.4	1.10E-21	82597	0
2	30	50.5	0.008021	16.3	0.9723	185.8	8.54E-25	128281	0
3	40	55.3	0.043349	30.9	0.8179	214.6	3.95E-26	174002	0
4	50	67.3	0.042597	36.9	0.8982	216.1	7.50E-23	219740	0
GED									
1	20	66.4	3.56E-07	57.1	1.11E-05	72.9	3.04E-08	13766	0
2	30	68.7	4.60E-05	63.1	2.50E-04	94.2	8.24E-09	18280	0
3	40	94.2	1.81E-06	68.9	2.20E-03	122.9	1.29E-10	23027	0
4	50	110.3	1.29E-06	86.9	6.87E-04	111.0	1.04E-06	26615	0
Skewed GED									
1	20	46.7	0.0003964	49.7	0.000146	70.2	8.67E-08	12963	0
2	30	55.0	0.002494	47.8	0.015554	89.5	4.36E-08	17781	0
3	40	59.7	0.0179324	68.0	0.002755	113.7	3.17E-09	22805	0
4	50	73.2	0.0142492	74.3	0.011254	106.7	3.68E-06	25743	0

**Table 2.7 (continued)**

Note: The purpose of this test is to examine the goodness of fit of the distributions assumed by parametric models. Test is conducted for five distribution assumptions for four markets, assuming daily returns follow ARMA(1,1) process, and the volatility follows GJR-GARCH(1,1) process. Data are grouped into 20, 30, 40 and 50 bins for each market, and the test is conducted independently for each of these bin numbers used. The result suggests that Skewed Student T is the best fit for S&P 500, closely followed by skewed GED. Student T, skewed Student T and Skewed GED (with 50 bins) are all acceptable for crude oil at 99% level. None of the distributions fit well with gold and Vanguard Long Term Bond Fund.

Data: S&P 500 daily return from 01/01/1950 to 04/30/2012; Crude oil price daily return from 01/02/1986 to 04/30/2012; Gold daily return from 04/02/1968 to 04/30/2012; Vanguard Long Term Bond Fund (VBLTX), daily return from 03/01/1994 to 04/30/2012.

**Table 2.8 Parameter Estimates for Extreme Value Theory (EVT) (Daily return, In Sample Data)**

	<b>S&amp;P 500</b>	<b>Crude Oil</b>	<b>Gold</b>	<b>Vanguard Long Term Bond Fund</b>
$\xi$	0.2993045*	0.1608582	0.257962	0.0360845
$\beta$	0.4662844	0.6627524	0.620441	0.5634607

Note:  $\xi$  and  $\beta$  are the shape and scaling parameters in the CDF of General Pareto Distribution respectively. A greater than zero  $\xi$  indicates fat tails. Threshold  $u$  is chosen to be the 95% percentile of empirical distribution.

\*For example, for S&P 500,  $\xi = 0.2993045$  indicates that the original distribution of S&P 500 daily returns contains fat tails, since it is greater than zero (the lower 5% of the daily return distribution is used for EVT estimation) .

Data: S&P 500 daily return from 01/01/1950 to 05/12/2008; Crude oil price daily return from 01/02/1986 to 05/12/2008; Gold daily return from 04/02/1968 to 05/15/2008; Vanguard Long Term Bond Fund (VBLTX), daily return from 03/01/1994 to 05/12/2008.

**Table 2.9 Value at Risk Out of Sample Results: Number of Exceedances (Daily Return)**

Market	VaR confd	EWMA- Norm	ARMA- GARCH- Norm	ARMA- GARCH- STD	ARMA- GARCH- SSTD	ARMA- GARCH- GED	ARMA- GARCH- SGED	ARMA- GJR- Norm	ARMA- GJR- STD	ARMA- GJR- SSTD	ARMA- GJR- GED	ARMA- GJR- SGED	EVT	HS
<b>S&amp;P 500</b>														
	99%	27*	26*	21*	18*	21*	16	28*	17*	15	17*	15	8	68*
	95%	64*	69*	72*	69*	68*	67*	67*	70*	68*	67*	62	43	140*
<b>Crude Oil</b>														
	99%	14	16	9	7	10	7	12	9	7	10	7	3*	19*
	95%	52	53	61	58	56	51	55	61	58	56	51	21*	80*
<b>Gold</b>														
	99%	23*	20*	13	13	12	12	20*	14	14	12	12	3*	16
	95%	54	52	65*	64*	56	56	52	63	63	57	57	28*	84*
<b>Vanguard LT Bond Fund</b>														
	99%	11	15	9	9	9	10	15	9	9	10	10	3*	35*
	95%	56	61	64*	64*	60	60	63	64*	64*	60	60	35*	125*

Note: The expected number of Value at Risk exceedances is 10 for 99% confidence level and 50 for 95% confidence level, given 1000 out of sample observations. More exceedances indicate the model underestimates Value at Risk, and fewer exceedances indicate the model overestimates Value at Risk. Z test confidence intervals are obtained by assuming each number of exceedances follows a Binomial Distribution. For 99% Value at Risk, the 95% confidence level is (3.83, 16.17). For 95% Value at Risk, the 95% confidence level is (36.49, 63.51).

\* indicates that the number of exceedances fall outside 95% confidence interval. ARMA-GJR-SGED model is the only model that passes the Z test at both confidence levels for all four markets.

ARMA = ARMA(1,1), GARCH = GARCH(1,1), GJR = GJR-GARCH(1,1), EWMA-Norm model has a fixed  $\lambda = 0.94$ .

Data: Out of sampled data, S&P 500 daily return from 05/13/2008 to 04/30/2012; Crude oil price daily return from 05/13/2008 to 04/30/2012; Gold daily return from 05/16/2008 to 04/30/2012; Vanguard Long Term Bond Fund (VBLTX), daily return from 05/12/2008 to 04/30/2012. N = 1000.

**Table 2.10 Expected Shortfall Out of Sample Test Results: P Values (Daily Return)**

Market	ES Confd	EWMA- Norm	ARMA- GARCH- Norm	ARMA- GARCH- STD	ARMA- GARCH- SSTD	ARMA- GARCH- GED	ARMA- GARCH- SGED	ARMA- GJR- Norm	ARMA- GJR- STD	ARMA- GJR- SSTD	ARMA- GJR- GED	ARMA- GJR- SGED	EVT	HS
<b>S&amp;P 500</b>														
	99%	0.0530	0.0182*	0.7479	0.7272	0.3067	0.1760	0.0855	0.3382	0.5614	0.1149	0.2676	0.9880	0.0004*
	95%	0.0009*	0.0001*	0.0580	0.1964	0.0448*	0.1258	0.0004*	0.0815	0.2797	0.0878	0.2243	1.0000	0.0000*
<b>Crude Oil</b>														
	99%	0.0578	0.0684	0.8451	0.7513	0.5841	0.3889	0.0492*	0.8469	0.7631	0.5817	0.3756	0.7278	0.0020*
	95%	0.1599	0.5773	0.9932	0.9998	0.9873	0.9918	0.4913	0.9932	0.9998	0.9864	0.9940	0.9611	0.4867
<b>Gold</b>														
	99%	0.0876	0.0235*	0.9931	0.9927	0.8559	0.8320	0.0254*	1.0000	1.0000	0.9702	0.9984	0.9999	0.0073*
	95%	0.0126*	0.0036*	0.7477	0.7416	0.6845	0.6456	0.0050*	0.7608	0.7694	0.7197	0.7382	0.9465	0.0916
<b>Vanguard LT Bond Fund</b>														
	99%	0.1323	0.1270	0.5448	0.6294	0.5089	0.5243	0.1321	0.4872	0.4409	0.3416	0.3894	0.3803	0.5000
	95%	0.0467*	0.0697	0.7883	0.8538	0.7507	0.7322	0.0595	0.7332	0.7407	0.6992	0.6878	0.9219	0.5000

Note: t-test with  $H_0: \mu_z = 0$  vs.  $H_1: \mu_z > 0$ , where z is the difference between out of sample returns and calculated Expected Shortfall

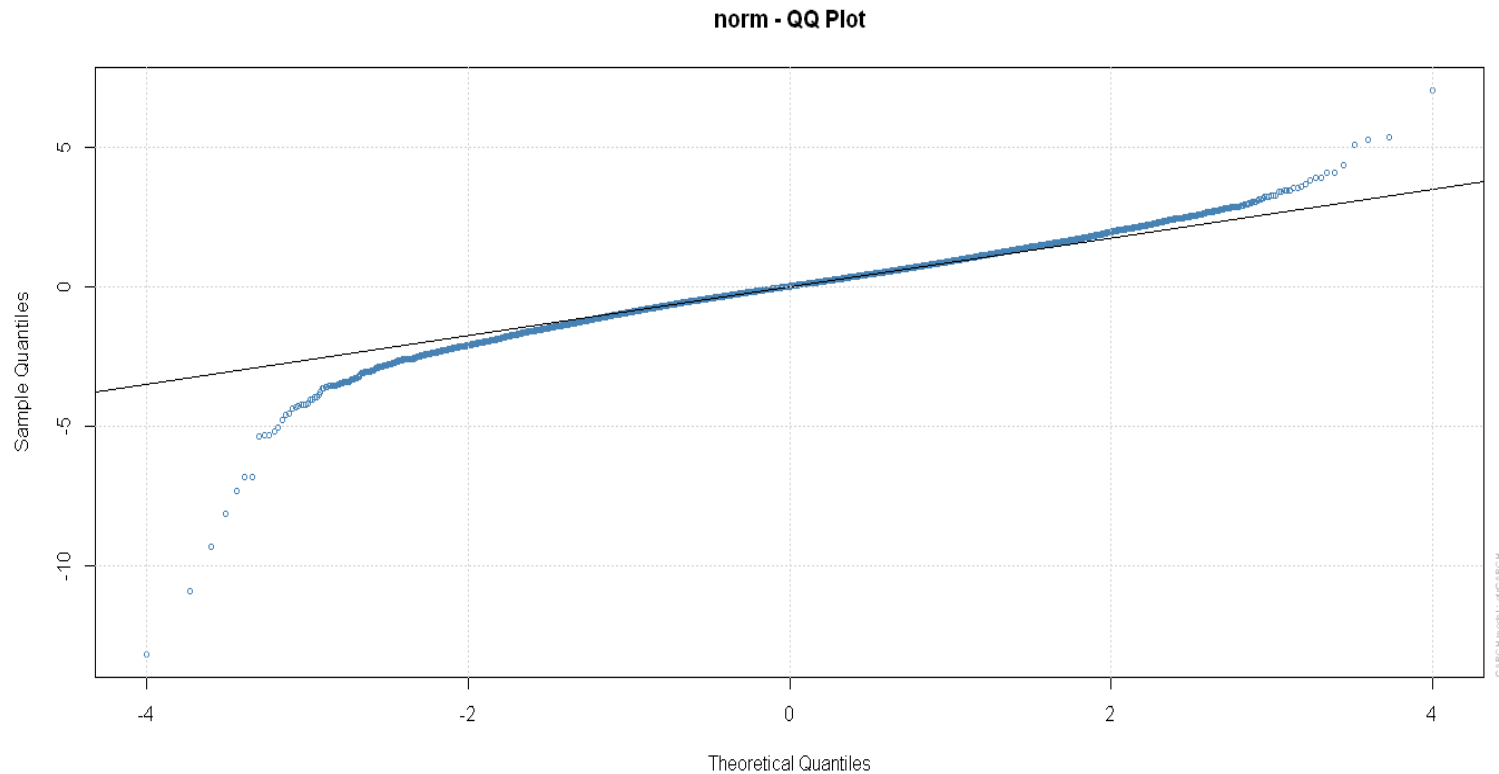
Small p values indicate less reliable Expected Shortfall.

\* indicates that model fails the 95% one sided t test. For example, the t test p value for EWMA-Norm and S&P 500 is 0.0009, based on which the null hypothesis (the estimated Expected Shortfall is reliable) is rejected at 5% level.

ARMA = ARMA(1,1), GARCH = GARCH(1,1), GJR = GJR-GARCH(1,1), EWMA-Norm model has a fixed  $\lambda = 0.94$

Data: Out of sample data, S&P 500 daily return from 05/13/2008 to 04/30/2012; Crude oil price daily return from 05/13/2008 to 04/30/2012; Gold daily return from 05/16/2008 to 04/30/2012; Vanguard Long Term Bond Fund (VBLTX), daily return from 05/12/2008 to 04/30/2012. N = 1000.

**Figure 2.1 QQ Plot for S&P 500 (Daily Return Standardized Residuals, 1950-2012): ARMA(1,1)-GJR-GARCH(1,1)-Norm (Normal Distribution)**

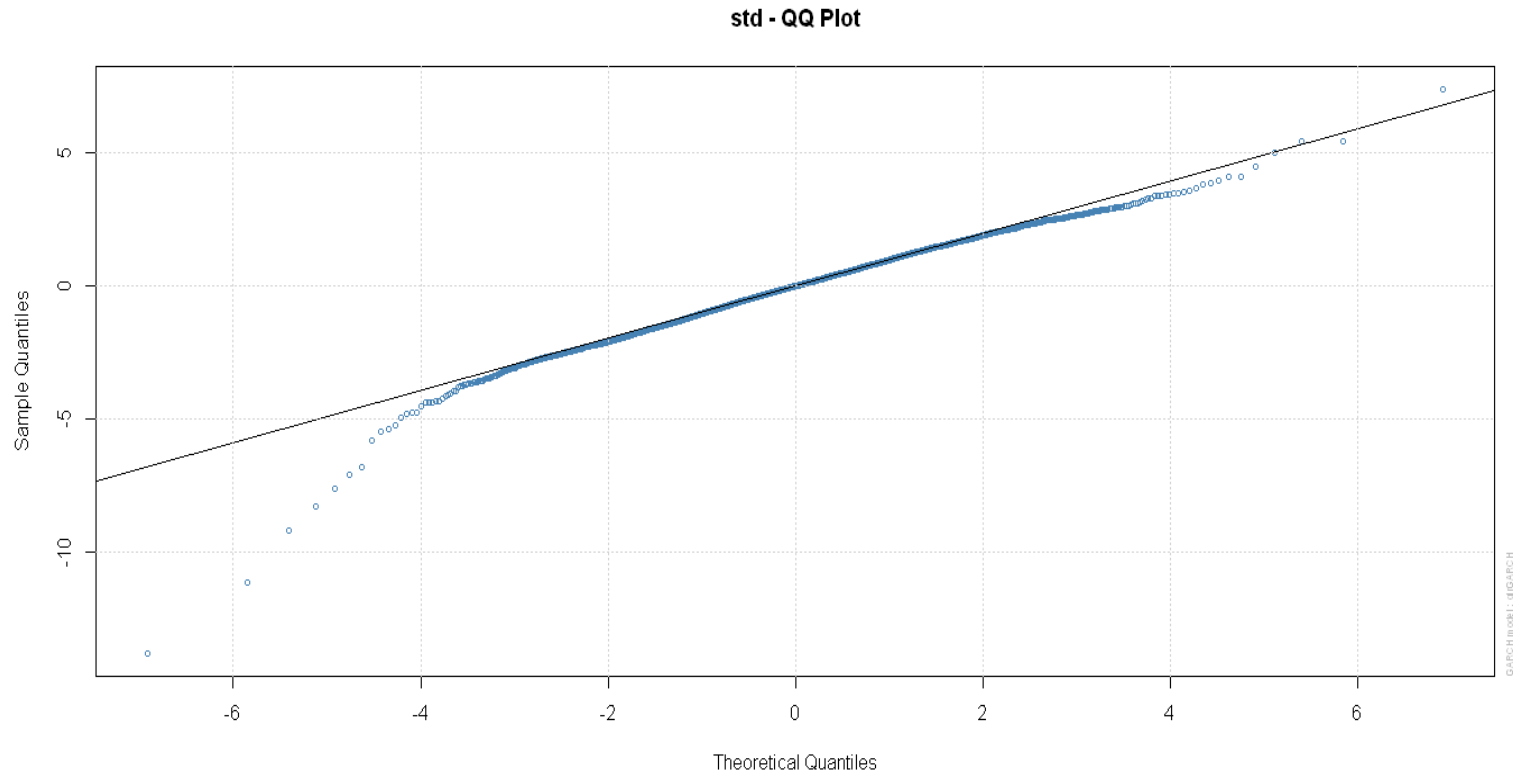


Notes: The solid straight line represents theoretical Normal quantiles, the hollow dots represent actual return standardized residuals. The hollow dots falling on or close to the solid straight line would indicate that the underlying distribution fits the data well. This figure suggests that Normal Distribution fits the data poorly.

Data: S&P 500 daily return from 01/01/1950 to 04/30/2012, entire data time period.  
Return standardized residual is defined as return mean residual divided by standardized deviation.



**Figure 2.2 QQ Plot for S&P 500 (Daily Return Standardized Residuals, 1950-2012): ARMA(1,1)-GJR-GARCH(1,1)-STD (Student T Distribution)**

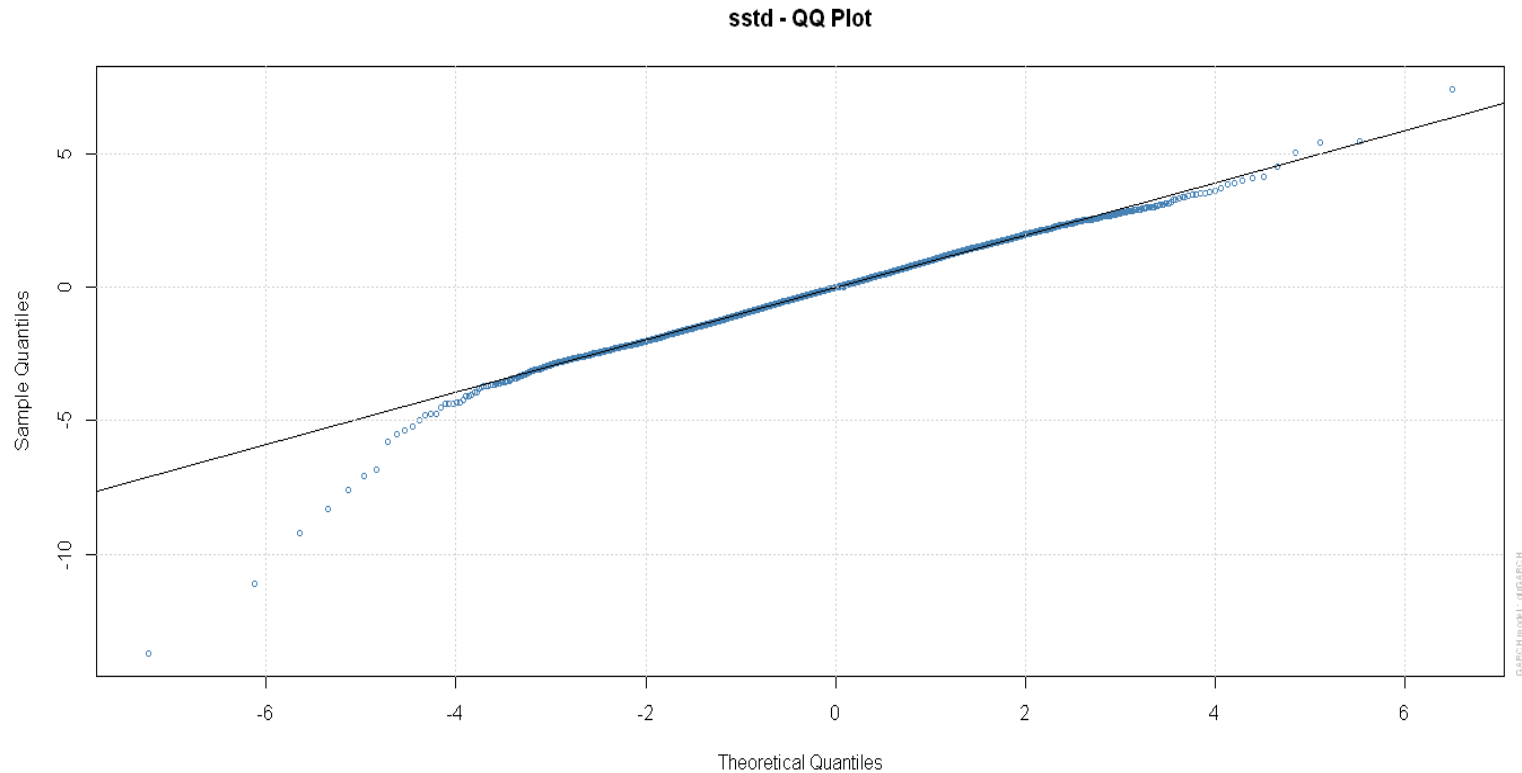


Notes: The solid straight line represents theoretical quantiles of Student T with  $\text{DOF} = 7.367265$ . The hollow dots represent actual return standardized residuals. The hollow dots falling on or close to the solid straight line would indicate that the underlying distribution fits data well.

Comparing to Figure 2.1 (Normal Distribution), Student T Distribution better fits the left tail and right tail of return standardized residuals distribution. Standardized residual is defined as return mean residual divided by standardized deviation

Data: S&P 500 daily Index return from 01/01/1950 to 04/30/2012, entire data time period

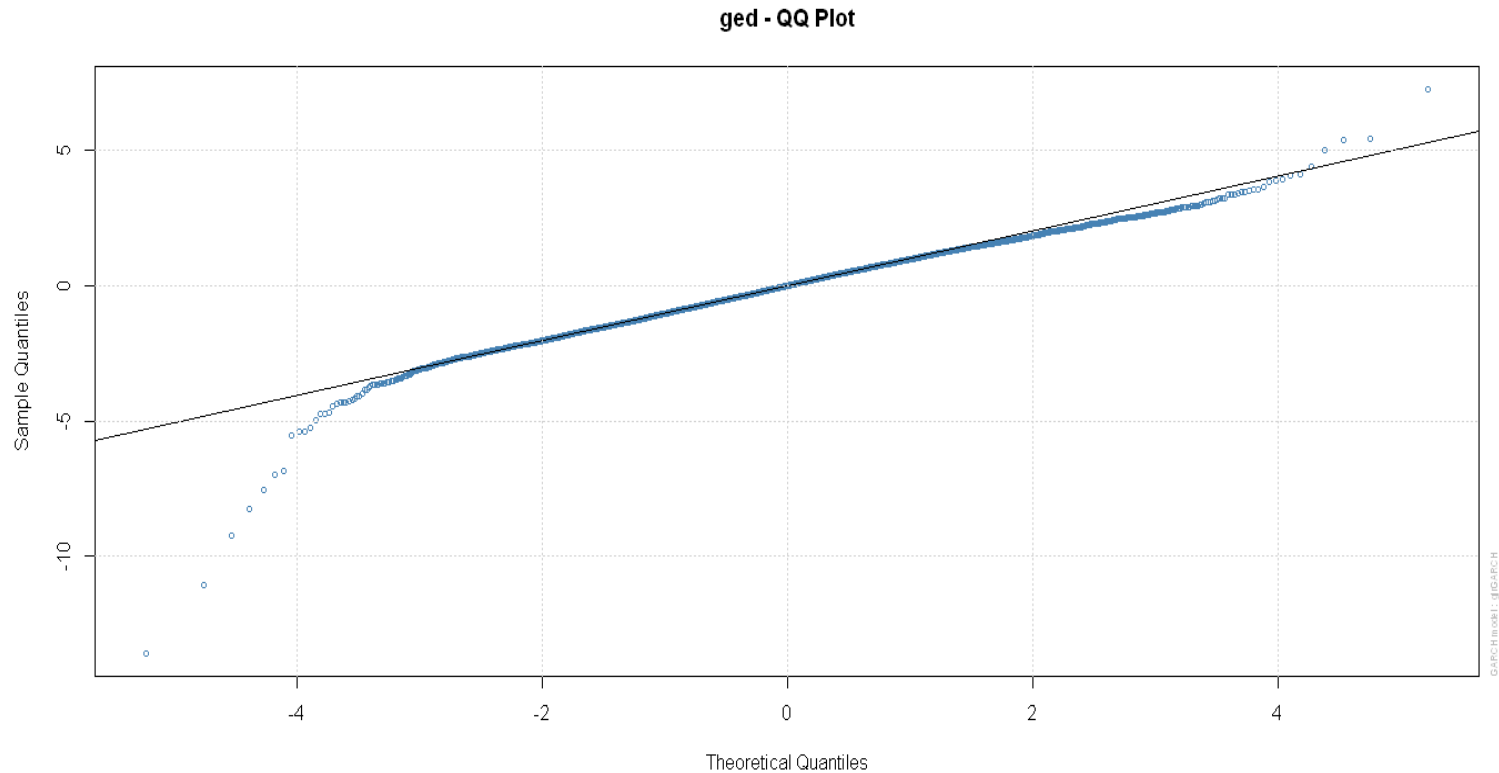
**Figure 2.3 QQ Plot for S&P 500 (Daily Return Standardized Residuals, 1950-2012): ARMA(1,1)-GJR-GARCH(1,1)-SSTD (Skewed Student T Distribution)**



Notes: The solid straight line represents theoretical quantiles of Student T with  $\text{DOF} = 7.435878$ ; Skewness = 0.943460; The hollow dots represent actual return standardized residuals. The hollow dots falling on or close to the solid straight line indicate that the underlying distribution fits data well. Skewed Student T Distribution fits the data better than both Student T (Figure 2.2) and Normal Distribution (Figure 2.1). Standardized residual is defined as return mean residual divided by standardized deviation.

Data: S&P 500 daily Index return from 01/01/1950 to 04/30/2012, entire data time period.

**Figure 2.4 QQ Plot for S&P 500 (Daily Return Standardized Residuals, 1950-2012): ARMA(1,1)-GJR-GARCH(1,1)-GED (General Error Distribution)**



Notes: The solid straight line represents theoretical quantiles of GED with  $K$  (shape) = 1.374324.

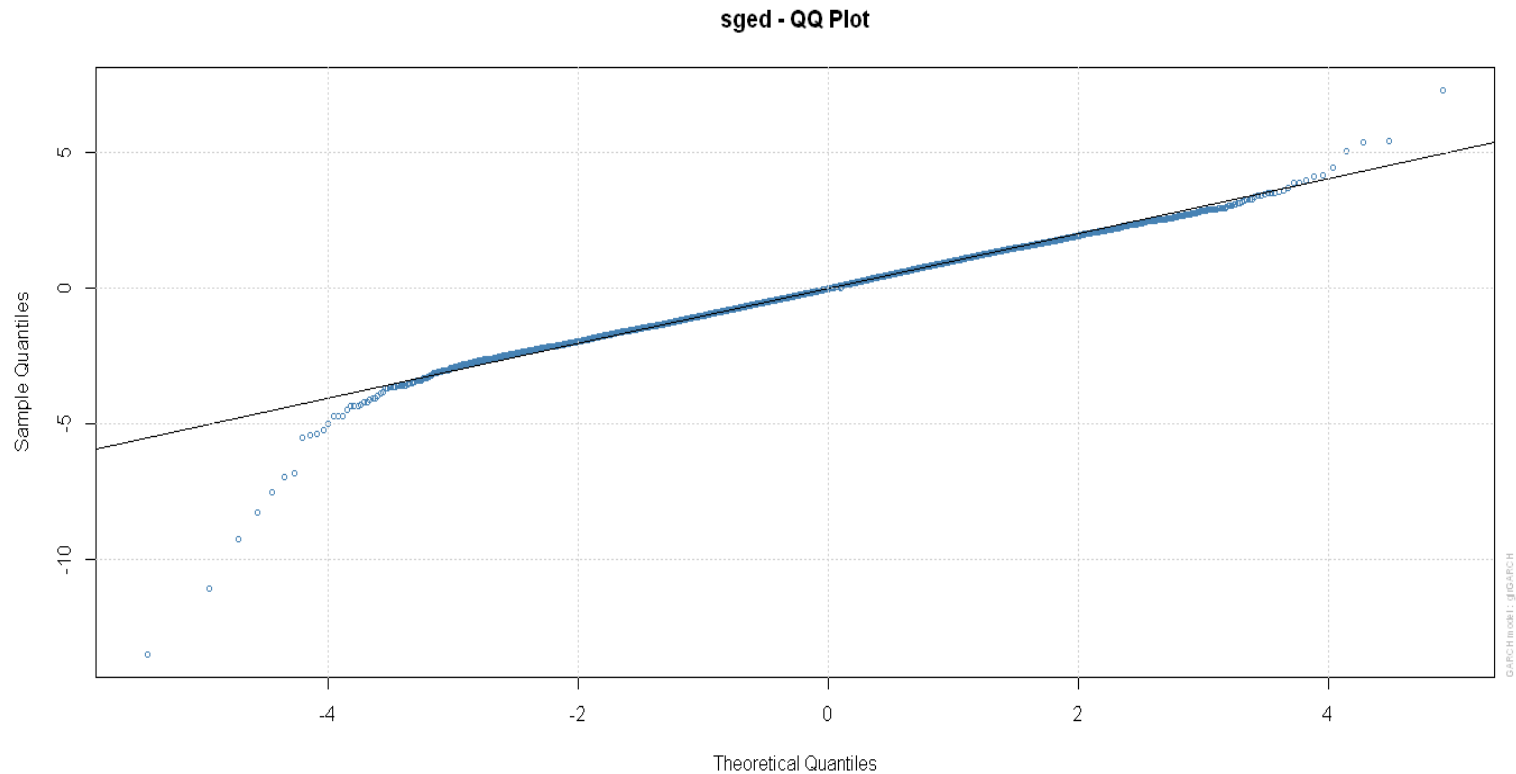
The hollow dots represent actual return standardized residuals. The hollow dots falling on or close to the solid straight line indicate that the underlying distribution fits data well.

This figure shows GED is superior to Normal Distribution (Figure 2.1) in approximating standardized residuals. However, it is inclusive that GED is superior to Student T (Figure 2.2) or skewed Student T Distribution (Figure 2.3).

Standardized residual is defined as return mean residual divided by standardized deviation.

Data: S&P 500 daily Index return from 01/01/1950 to 04/30/2012, entire data time period.

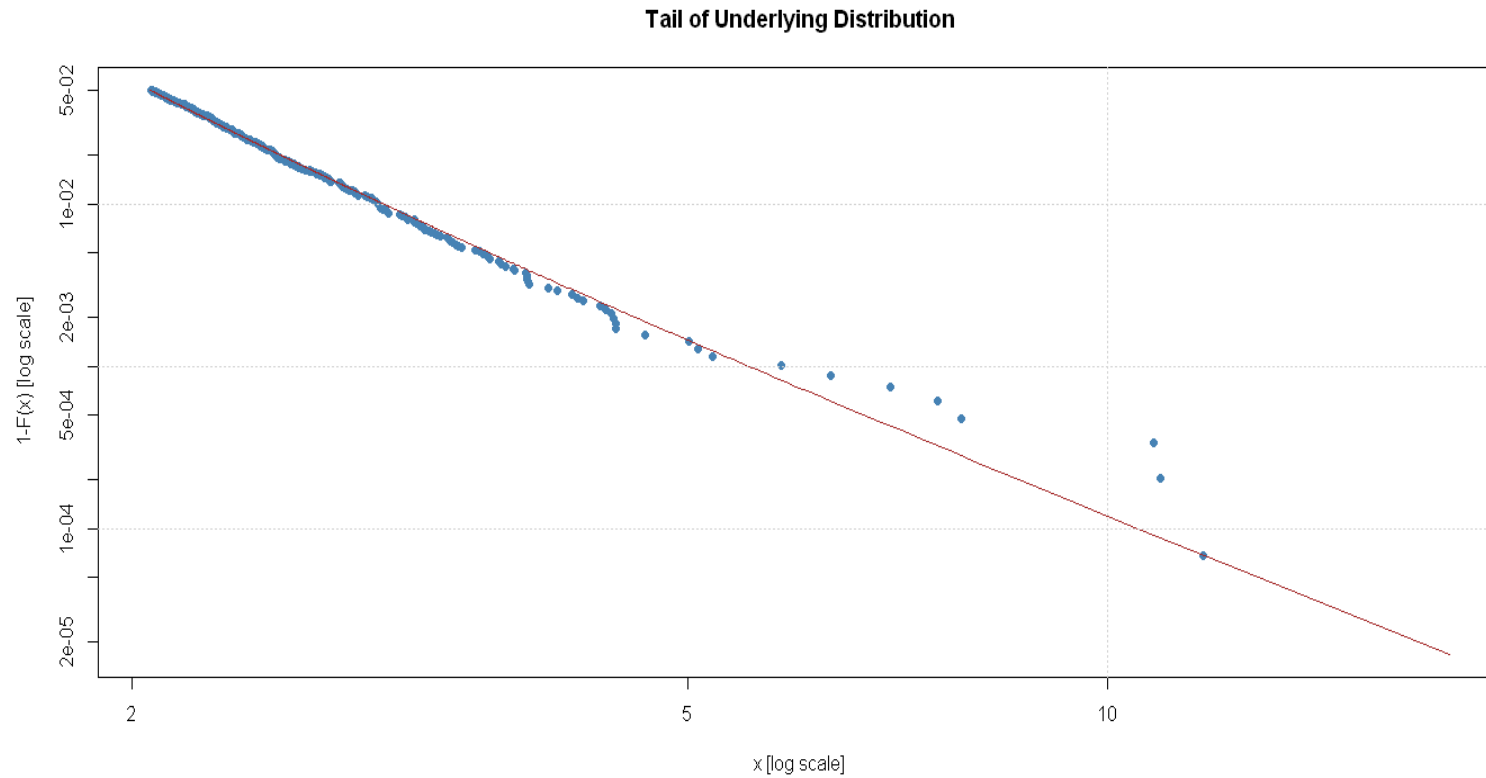
**Figure 2.5 QQ Plot for S&P 500 (Daily Return Standardized Residuals, 1950-2012): ARMA(1,1)-GJR-GARCH(1,1)-SGED (Skewed General Error Distribution)**



Note: The solid straight line represents theoretical quantiles of SGED with  $K(\text{shape}) = 1.387967$ ; Skewness = 0.937046  
The Hollow dots represent actual return standardized residuals. The hollow dots falling on or close to the solid straight line indicate that the underlying distribution fits the data well.  
This figure shows that SGED fits the data better than Normal Distribution (Figure 2.1), Student T Distribution (Figure 2.2), and GED Distribution (Figure 2.4). Standardized residual is defined as return mean residual divided by standardized deviation.

Data: S&P 500 daily return from 01/01/1950 to 04/30/2012, entire data time period.

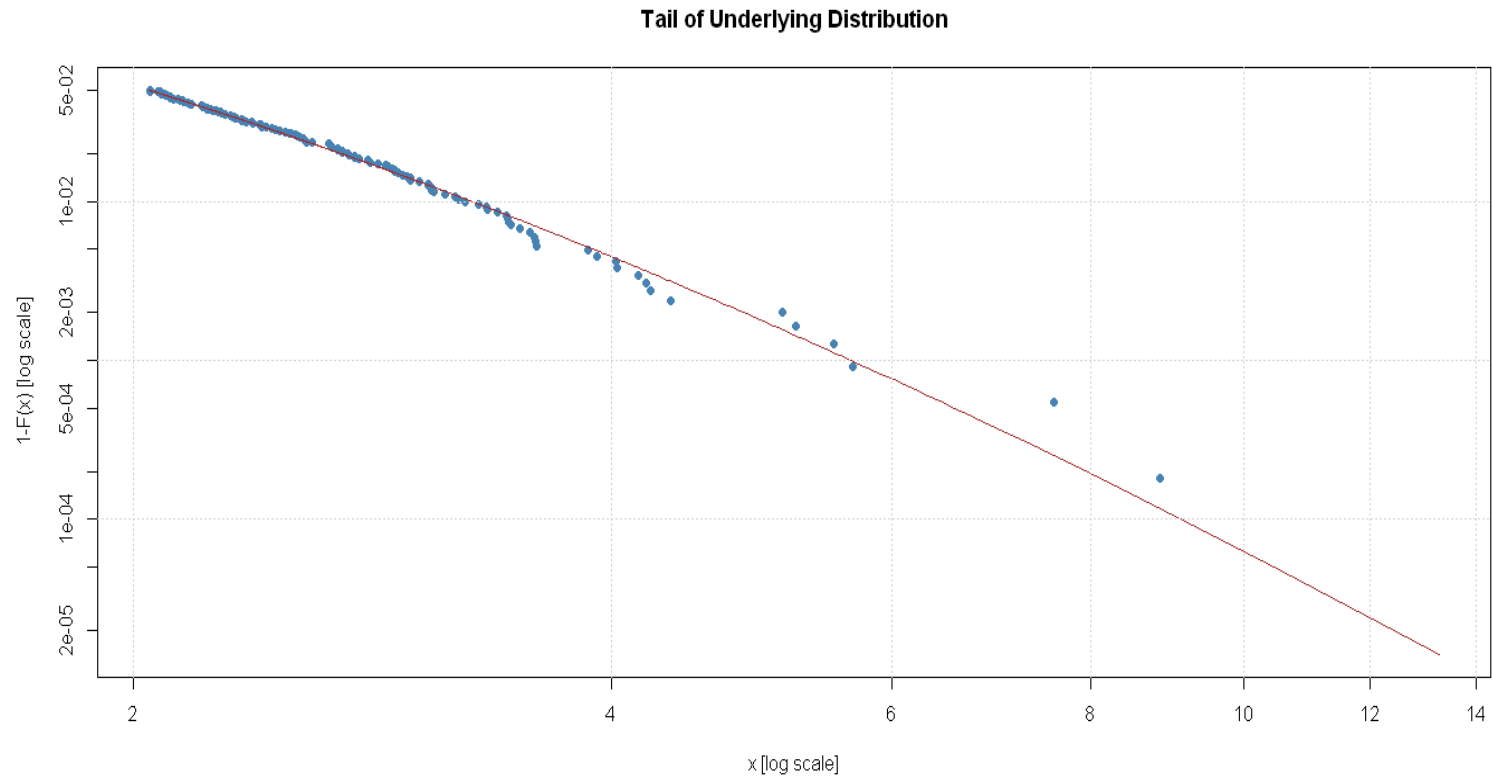
**Figure 2.6 Extreme Value Theory (EVT): Actual Versus Estimated Left Tail for S&P 500 (Standardized Residuals)**



Notes: Solid line represents the estimated left tail by EVT. Dots represent actual data. Dots sitting above the solid line suggest underestimation of risk by EVT at the corresponding percentile, and vice versa. Figure suggests that EVT captures the tail risk well. The threshold  $u$  is chosen to be the lower 5 percentile of the in sample returns. Residuals are obtained by assuming GARCH(1,1).

Data: S&P 500 daily Index daily return, the lower 5% of the in sample data (from 01/01/1950 to 04/30/2012).

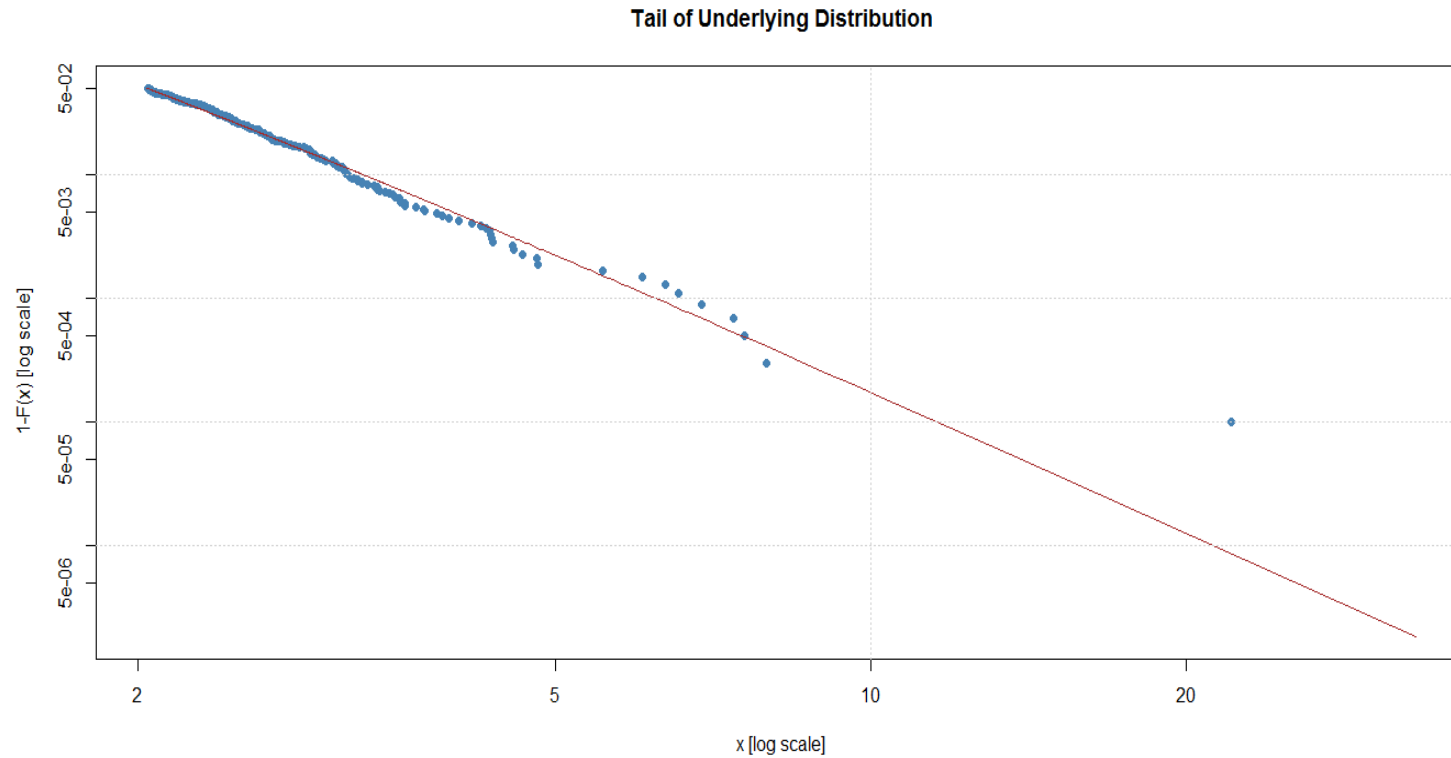
**Figure 2.7 Extreme Value Theory (EVT): Actual Versus Estimated Left Tail for Crude Oil (Standardized Residuals)**



Note: Solid line represents the estimated left tail by EVT. Dots represent actual left tail. Dots sitting above the solid line suggest underestimation of risk by EVT at the corresponding percentile, and vice versa. Figure suggests that EVT overestimates risk except in the far left tail. The threshold  $u$  is chosen to the lower 5 percentile of the in sample returns. Residuals are obtained by assuming GARCH(1,1).

Data: crude oil daily return, the lower 5% of the in sample data (from 01/02/1986 to 05/12/2008).

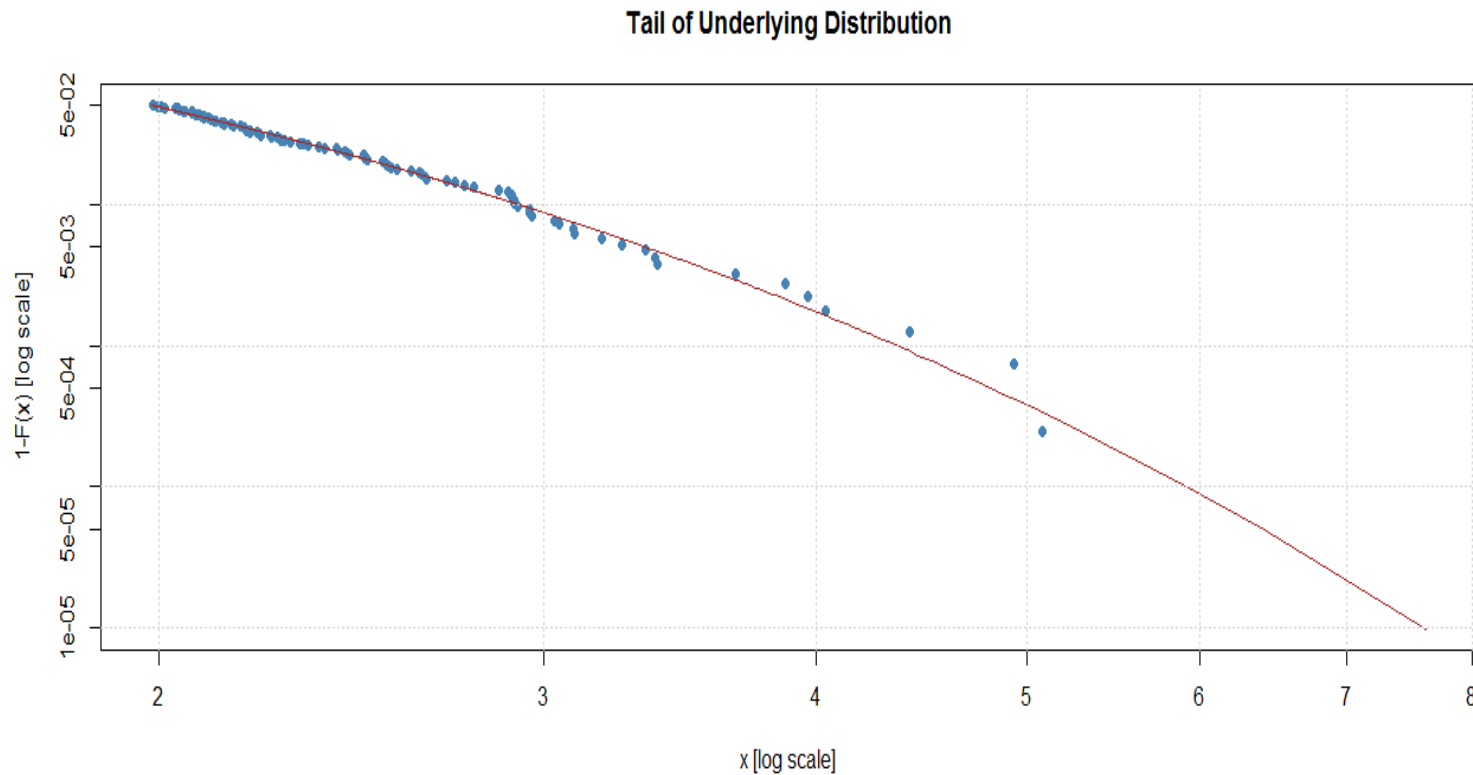
**Figure 2.8 Extreme Value Theory (EVT): Actual Versus Estimated Left Tail for Gold (Standardized Residuals)**



Note: Solid line represents the estimated left tail by EVT. Dots represent actual left tail. Dots sitting above the solid line suggest underestimation of risk by EVT at the corresponding percentile, and vice versa. Figure suggests that EVT overestimates risk except in the far left tail. The threshold  $u$  is chosen to be the lower 5 percentile of the in-sample returns. Residuals are obtained by assuming GARCH(1,1).

Data: gold daily returns, the lower 5% of the in sample data (from 04/02/1968 to 05/15/2008)

**Figure 2.9 Extreme Value Theory (EVT): Actual Versus Estimated Left Tail for Vanguard Long Term Bond Fund (Standardized Residuals)**

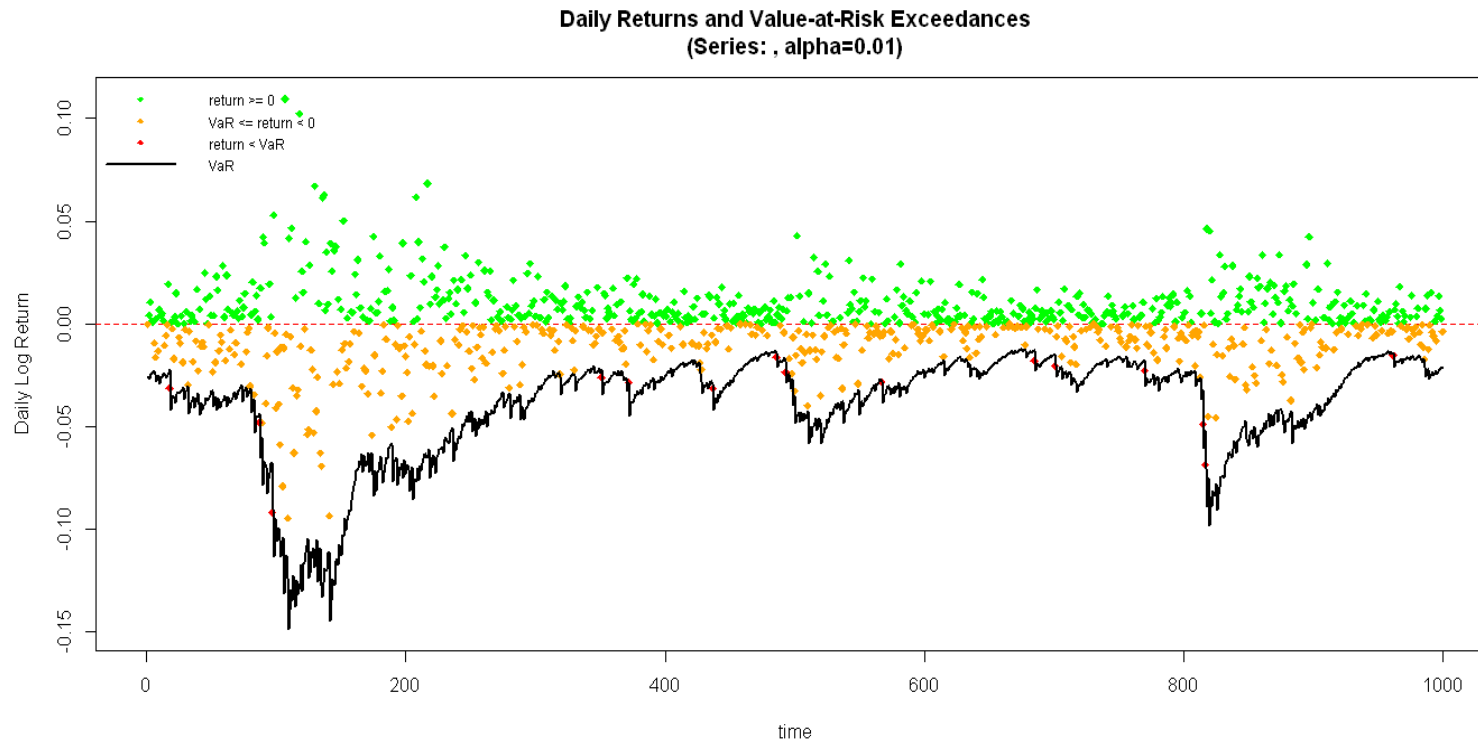


Note: Solid line represents the estimated left tail by EVT. Dots represent actual left tail. Dots sitting above the solid line suggest underestimation of risk by EVT at the corresponding percentile, and vice versa. Figure suggests that EVT overestimates risk except in the far left tail. The threshold  $u$  is chosen to be lower 5 percentile of the in-sample returns. Residuals are obtained by assuming GARCH(1,1).

Data: Vanguard Long Term Bond Fund (VBLTX) daily return, the lower 5% of the in sample data (from 03/01/1994 to 05/12/2008).



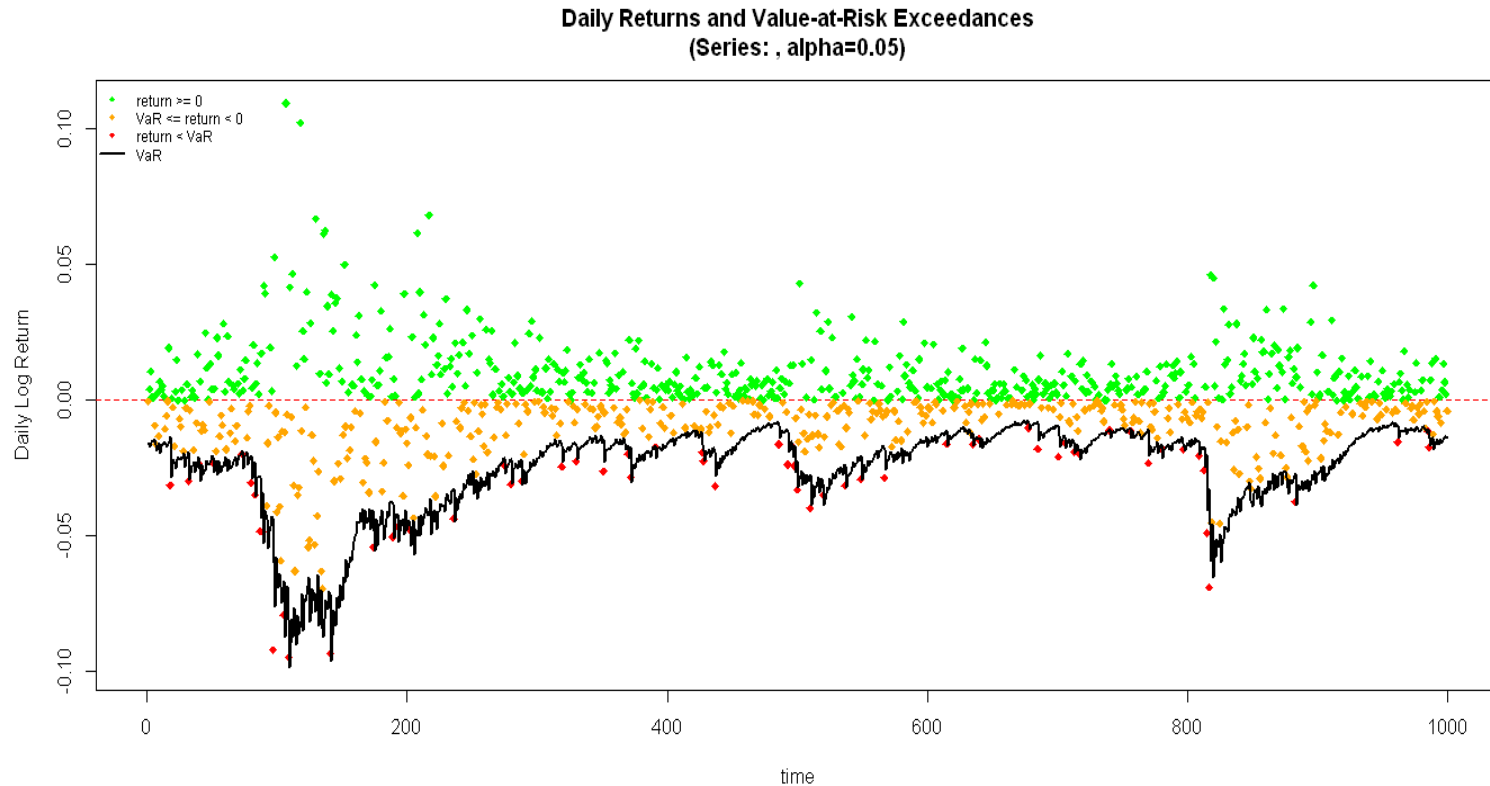
**Figure 2.10 Actual Daily Return Versus Estimated Value at Risk (Model: ARMA(1,1)-GJR-GARCH(1,1)-SGED, Data: Out of Sample S&P 500 Daily Return N=1000, Confidence Level = 99%)**



Note: Figure compares Value at Risk estimated by the proposed parametric model, ARMA(1,1)-GJR-GARCH(1,1)-SGED and actual S&P 500 return on each of 1000 out of sample days. Dots represent log daily returns of S&P 500. Solid line represents estimated Value at Risk at 99% confidence level by the proposed parametric model, GJR-GARCH(1,1)-SGED.

Data: S&P 500 log daily return from 05/13/2008 to 04/30/2012. N = 1000.

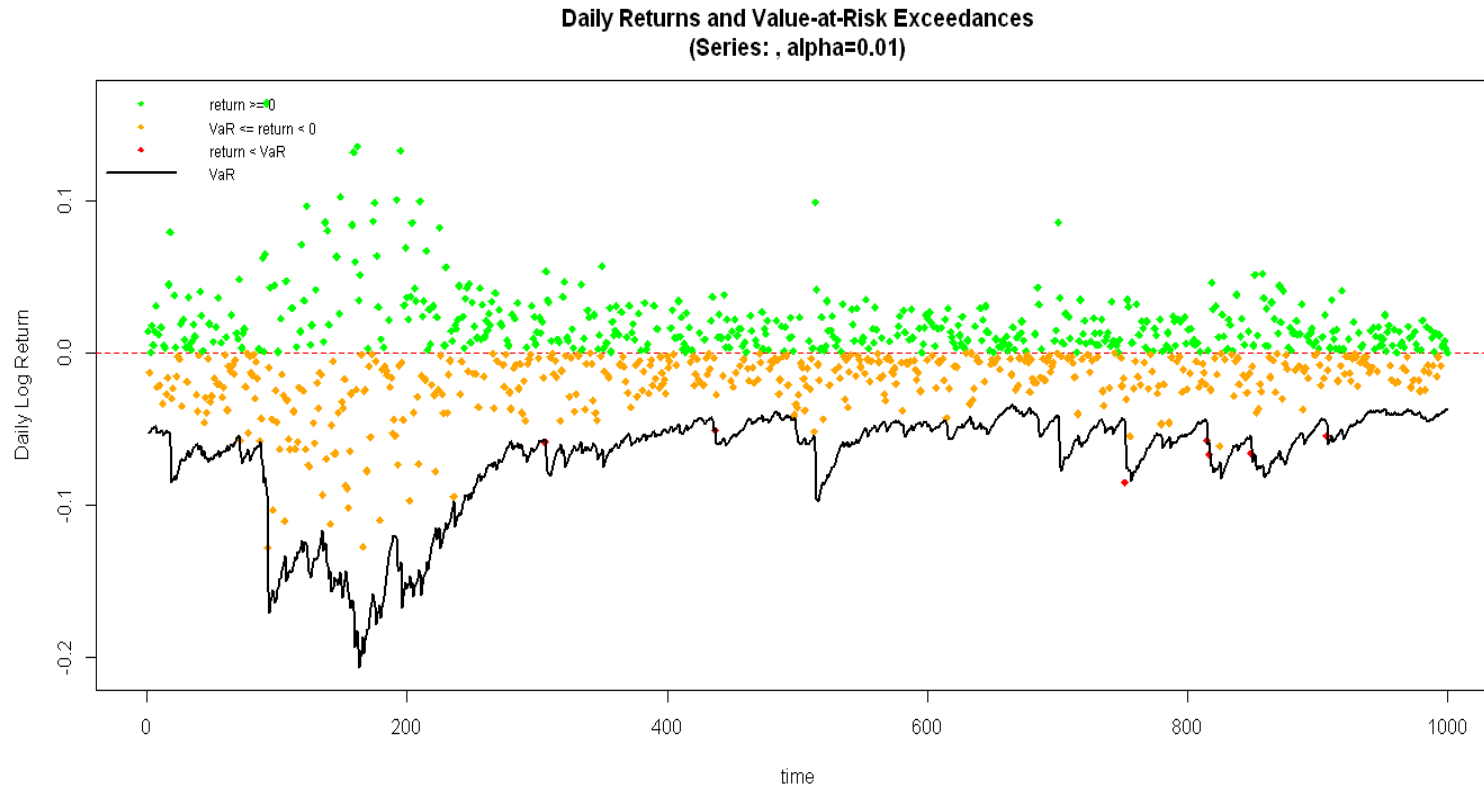
**Figure 2.11 Actual Daily Return Versus Estimated Value at Risk (Model: ARMA(1,1)-GJR-GARCH(1,1)-SGED, Data: Out of Sample S&P 500 Daily Return N=1000, Confidence Level = 95%)**



Note: Figure compares Value at Risk estimated by the proposed parametric model, ARMA(1,1)- GJR-GARCH(1,1)-SGED and actual S&P 500 return on each of 1000 out of sample days. Dots represent log daily returns of S&P 500. Solid line represents estimated Value at Risk at 95% confidence level by the proposed parametric model, GJR-GARCH(1,1)-SGED.

Data: S&P 500 log daily return from 05/13/2008 to 04/30/2012. N = 1000.

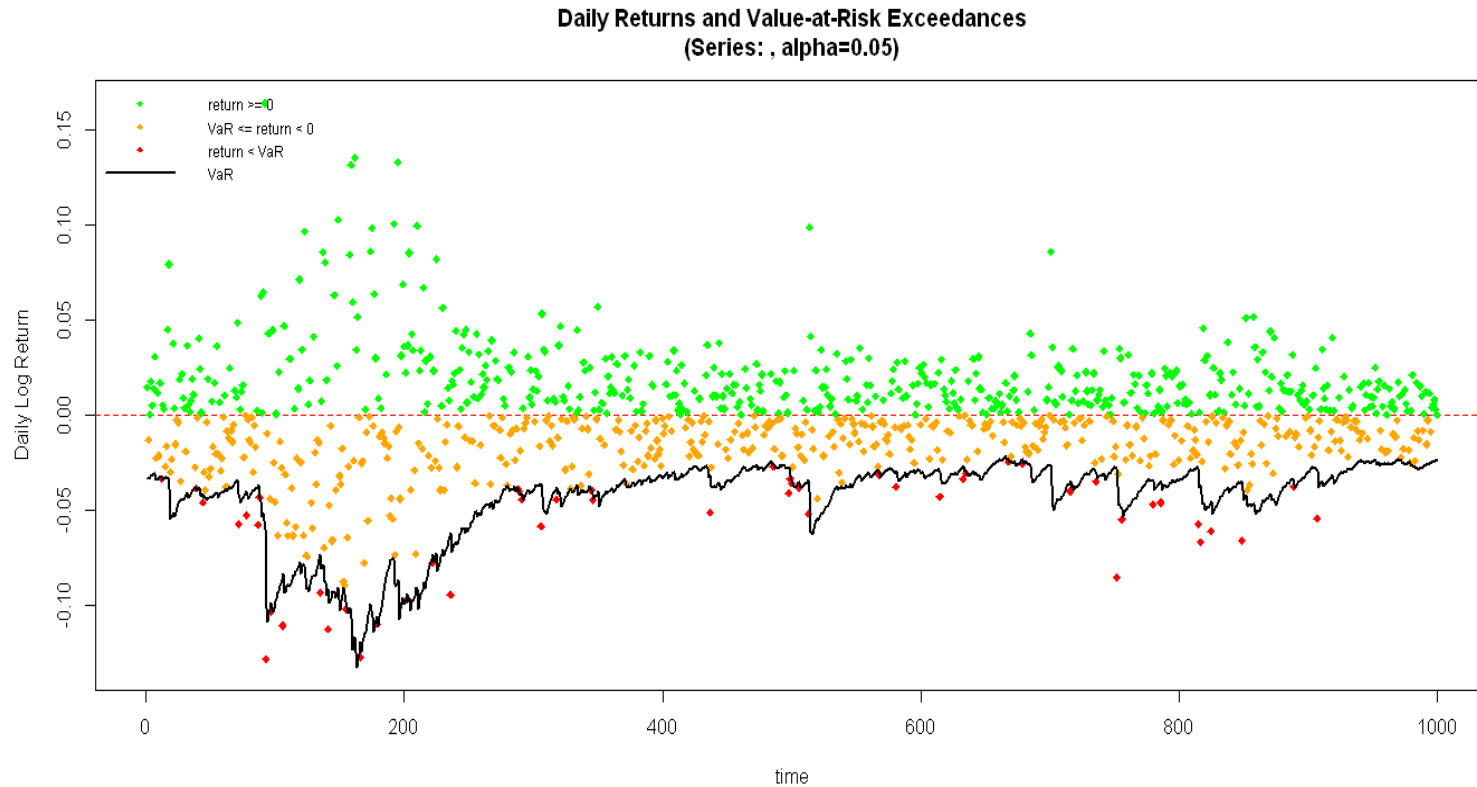
**Figure 2.12 Actual Daily Return Versus Estimated Value at Risk (Model: ARMA(1,1)-GJR-GARCH(1,1)-SGED, Data: Out of Sample Crude Oil Daily Return N=1000, Confidence Level = 99%)**



Note: Figure compares Value at Risk estimated by the proposed parametric model, ARMA(1,1)-GJR-GARCH(1,1)-SGED and actual crude oil return on each of 1000 out of sample days. Dots represent log daily returns of crude oil. Solid line represents estimated Value at Risk at 99% confidence level by the proposed parametric model, GJR-GARCH(1,1)-SGED.

Data: Crude oil daily return from 05/13/2008 to 04/30/2012. N = 1000.

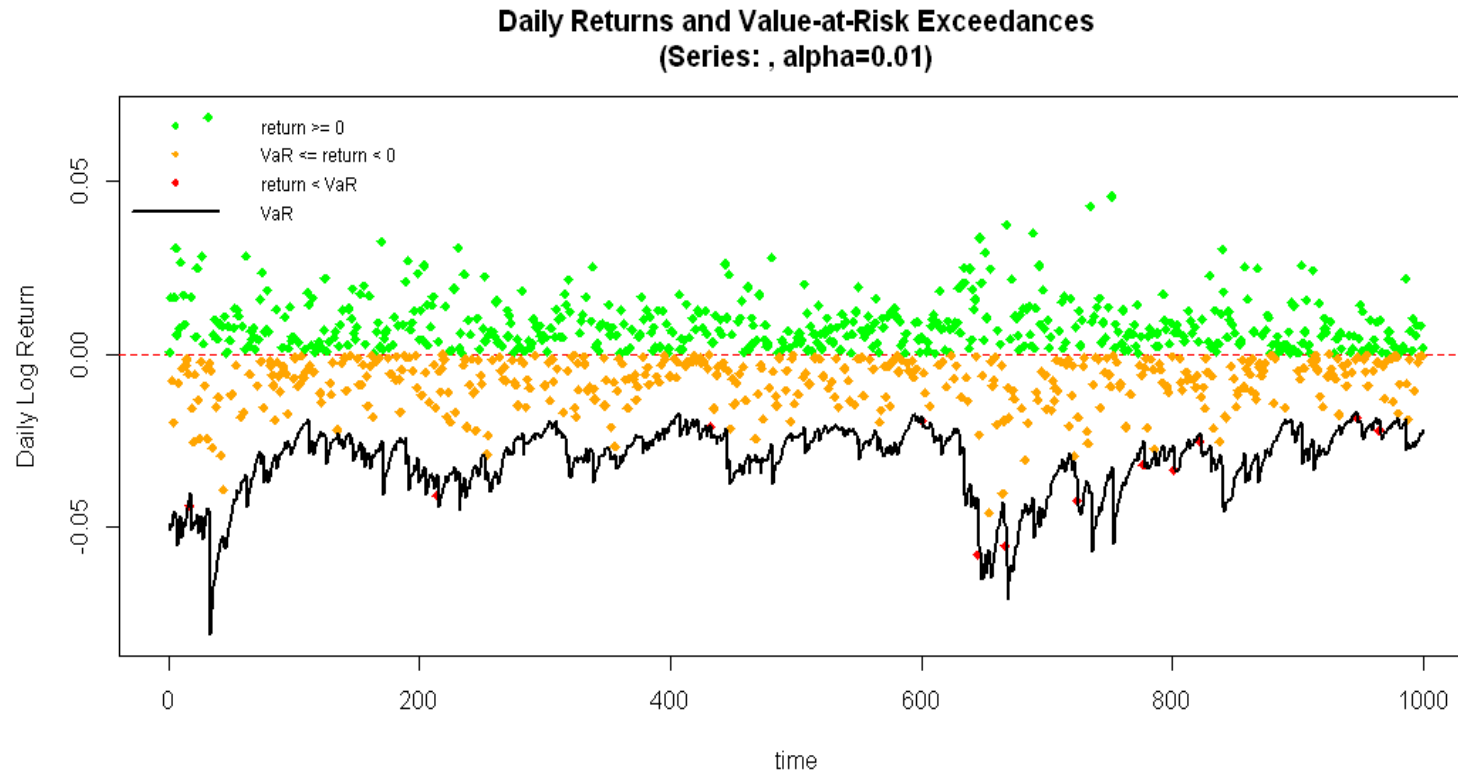
**Figure 2.13 Actual Daily Return Versus Estimated Value at Risk (Model: ARMA(1,1)-GJR-GARCH(1,1)-SGED, Data: Out of Sample Crude Oil Daily Return N=1000, Confidence Level = 95%)**



Note: Figure compares Value at Risk estimated by the proposed parametric model, ARMA(1,1)-GJR-GARCH(1,1)-SGED and actual crude oil return on each of 1000 out of sample days. Dots represent log daily returns of crude oil. Solid line represents estimated Value at Risk at 95% confidence level by the proposed parametric model, GJR-GARCH(1,1)-SGED.

Data: Crude oil daily return from 05/13/2008 to 04/30/2012. N = 1000.

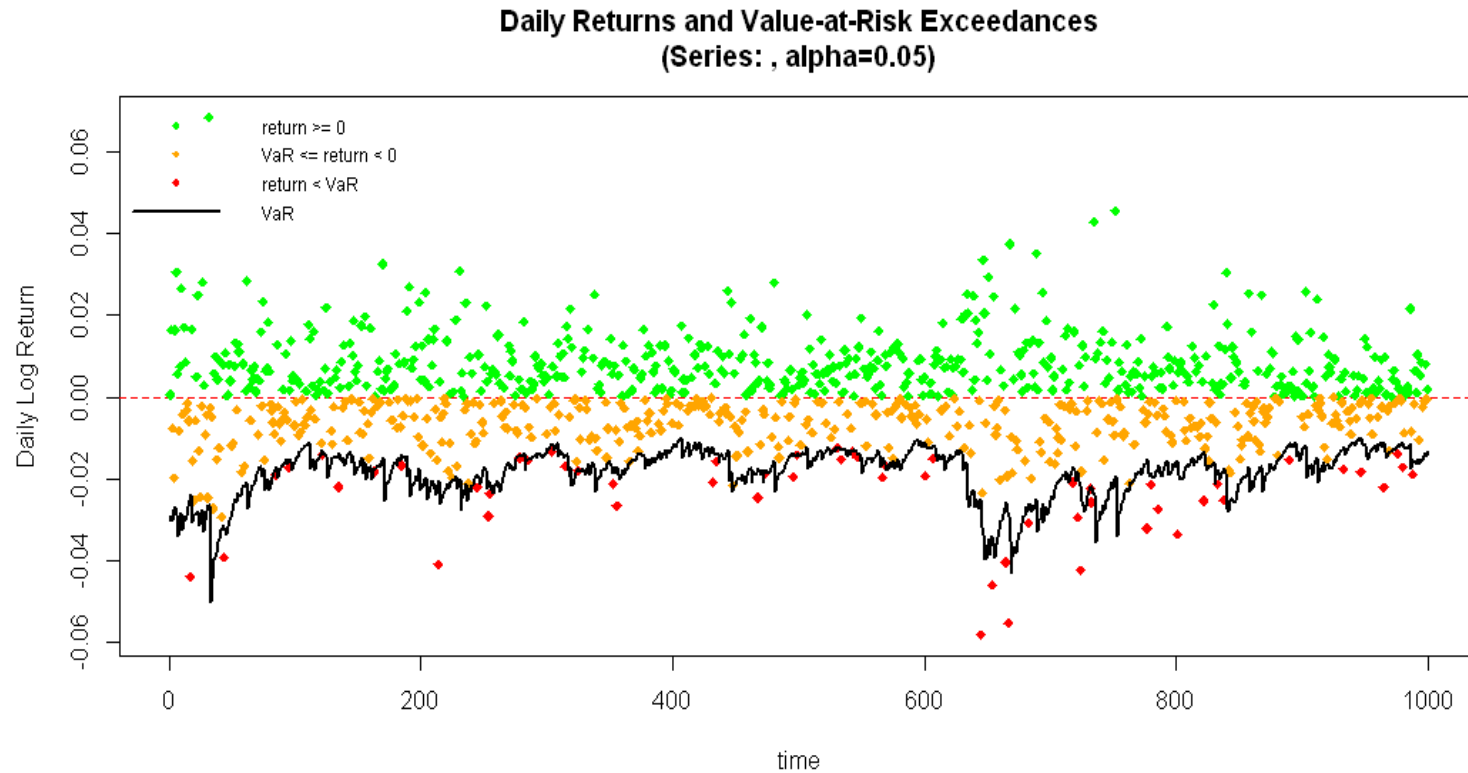
**Figure 2.14 Actual Daily Return Versus Estimated Value at Risk (Model: ARMA(1,1)-GJR-GARCH(1,1)-SGED, Data: Out of Sample Gold Daily Return N=1000, Confidence Level = 99%)**



Note: Figure compares Value at Risk estimated by the proposed parametric model, ARMA(1,1)-GJR-GARCH(1,1)-SGED and actual gold return on each of 1000 out of sample days. Dots represent log daily returns of gold. Solid line represents estimated Value at Risk at 99% confidence level by the proposed parametric model, GJR-GARCH(1,1)-SGED.

Data: Gold daily return from 05/16/2008 to 04/30/2012. N = 1000.

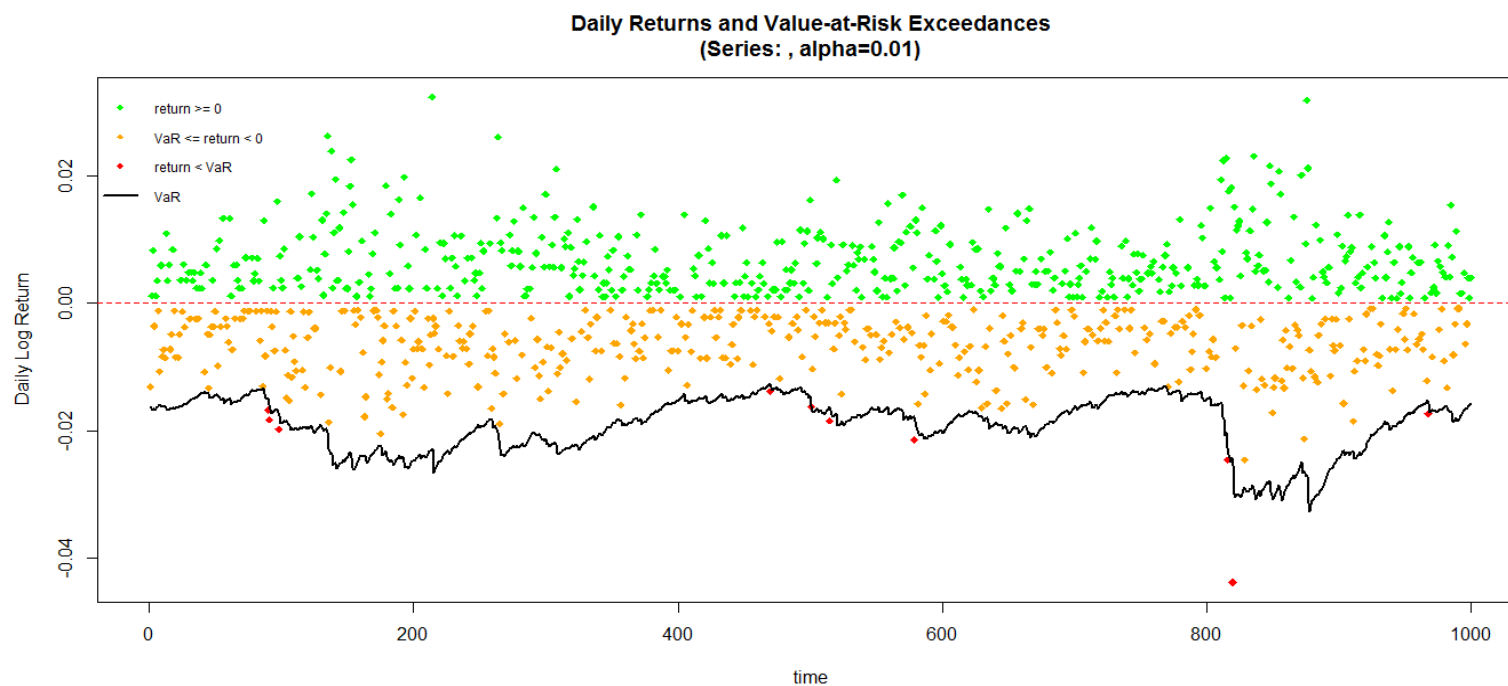
**Figure 2.15 Actual Daily Return Versus Estimated Value at Risk (Model: ARMA(1,1)-GJR-GARCH(1,1)-SGED, Data: Out of Sample Gold Daily Return N=1000, Confidence Level = 95%)**



Note: Figure compares Value at Risk estimated by the proposed parametric model, ARMA(1,1)-GJR-GARCH(1,1)-SGED and actual gold return on each of 1000 out of sample days. Dots represent log daily returns of gold. Solid line represents estimated Value at Risk at 95% confidence level by the proposed parametric model, GJR-GARCH(1,1)-SGED.

Data: Gold daily return from 05/16/2008 to 04/30/2012. N = 1000.

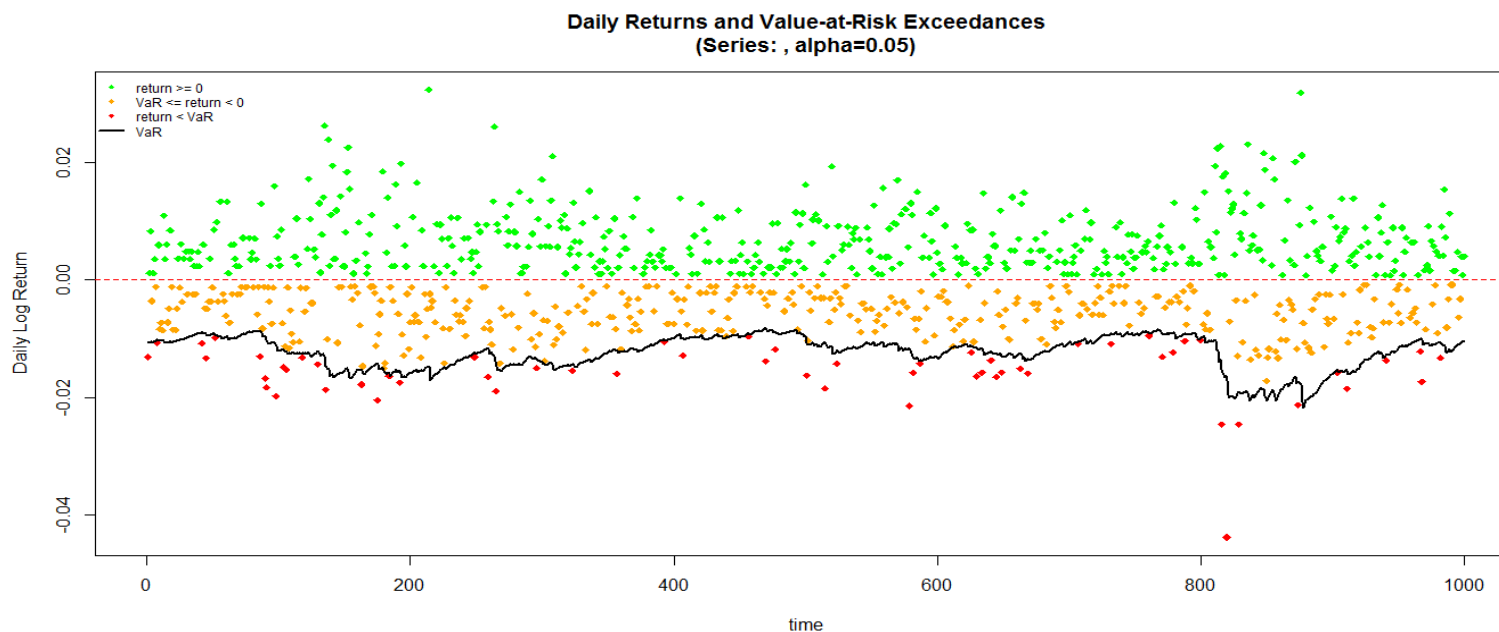
**Figure 2.16 Actual Daily Return Versus Estimated Value at Risk (Model: ARMA(1,1)-GJR-GARCH(1,1)-SGED, Data: Out of Sample Vanguard Long Term Bond Fund Daily Return N=1000, Confidence Level = 99%)**



Note: Figure compares Value at Risk estimated by the proposed parametric model, ARMA(1,1)-GJR-GARCH(1,1)-SGED and actual Vanguard Long Term Bond Fund return on each of 1000 out of sample days. Dots represent log daily returns of Vanguard Long Term Bond Fund. Solid line represents estimated Value at Risk at 99% confidence level by the proposed parametric model, GJR-GARCH(1,1)-SGED.

Data: Vanguard Long Term Bond Fund (VBLTX), daily return from 03/01/1994 to 04/30/2012. N = 1000.

**Figure 2.17 Actual Daily Return Versus Estimated Value at Risk (Model: ARMA(1,1)-GJR-GARCH(1,1)-SGED, Data: Out of Sample Vanguard Long Term Bond Fund Daily Return N=1000, Confidence Level = 95%)**



Note: Figure compares Value at Risk estimated by the proposed parametric model, ARMA(1,1)-GJR-GARCH(1,1)-SGED and actual Vanguard Long Term Bond Fund return on each of 1000 out of sample days. Dots represent log daily returns of Vanguard Long Term Bond Fund. Solid line represents estimated Value at Risk at 95% confidence level by the proposed parametric model, GJR-GARCH(1,1)-SGED.

Data: Vanguard Long Term Bond Fund (VBLTX), daily return from 03/01/1994 to 04/30/2012. N = 1000.



## **CHAPTER 3**

### **SUMMARY**

#### **PROBLEM AND OBJECTIVE**

Value at Risk (VaR) and Expected Shortfall (ES) are methods that are often used to measure market risk. This is the risk that the value of an asset will be adversely affected by the movements in financial markets, such as equity markets, bond markets, and commodity markets. Inaccurate and unreliable Value at Risk and Expected Shortfall models can lead to underestimation of market risk that a firm or financial institution is exposed to, and therefore may jeopardize the well-being or survival of the firm or financial institution during adverse market movements. Crotty (2009) argued that using (inaccurate) Value at Risk models led to inadequate capital reserves in large banks, and therefore this was one of the causes of the 2008 US financial crisis. For example, past Value at Risk models have often assumed the Normal Distribution, when in reality markets often have fatter tail distributions. As a result, Value at Risk models based on the Normal Distribution have often underestimated risk. The objective of this study is therefore to examine various Value at Risk and Expected Shortfall models, including fatter tail models, in order to analyze the accuracy and reliability of these models.

## DATA AND METHODS

The principles for selecting data in this study are (1) to include a variety of financial assets (equities, bonds, and commodities) that are exposed to daily market risk, (2) to include a sufficiently long time period of data for each of the selected markets, and (3) to ensure all data have the same start and end date for out of sample data. Based on these principles, data used in this study include daily returns of S&P 500 Index, crude oil, gold and the Vanguard Long Term Bond Fund (VBLTX). S&P 500 daily prices from January, 1950 to April, 2012 and Vanguard Long Term Bond Fund daily prices (VBLTX) from March, 1994 to April, 2012 are obtained from Yahoo Finance website. Crude oil daily prices from January, 1986 to April, 2012 are obtained from US Energy Information Administration (EIA) website. Gold daily prices from April, 1968 to April, 2012 are obtained from Federal Reserve Bank of St. Louis website. The same 1000 days (approximately four years from May, 2008- April, 2012) are used as the out of sample data in the out of sample test, to create comparability across data sets.

Three approaches are used for estimating Value at Risk and Expected Shortfall (Appendix B). They are (1) parametric approach (including 11 models), (2) a single model non-parametric approach [Historical Simulation (HS)], and (3) a single model semi-parametric approach [Extreme Value Theory (EVT), with the use of General Pareto Distribution]. The parametric models used in this study involve three volatility models including Exponential Weighted Moving Average (EWMA), GARCH(1,1), and GJR-GARCH(1,1), and three distribution assumptions including Normal, Student T, General Error Distribution (GED). The ARMA(1,1)-GJR-GARCH(1,1)-SGED (Skewed General Error Distribution) model is the new and proposed model of this study. It is innovative in

the following aspects. Firstly, it captures the autocorrelation in returns using ARMA(1,1) process. Secondly, it employs GJR-GARCH(1,1) to estimate one day forward volatility and captures the leverage effect (Black, 1976) in returns. Thirdly, it uses a skewed fat tail distribution, skewed General Error Distribution, to model the extreme tails of daily returns of the selected financial markets.

Statistical tests and out of sample test are used to examine the accuracy and reliability of Value at Risk and Expected Shortfall models. Statistical tests include sign bias test for leverage effect and Pearson test for goodness of fit for each distribution assumption. The Value at Risk out of sample test directly examines Value at Risk by measuring the total number of exceedances of each model using 1000 out of sample data. The Expected Shortfall out of sample test uses a one-tailed t test to examine the residuals and therefore indirectly examines the accuracy of the Expected Shortfall models.

## RESULTS

### Value at Risk Out of Sample Test Results

First, regarding the parametric models, the proposed model, ARMA(1,1)-GJR-GARCH(1,1)-SGED, is the most balanced model for estimating Value at Risk and Expected Shortfall based on the out of sample test. This model is statistically adequate for estimating Value at Risk for all four markets, based on the exceedance results for the out of sample test. Figure 2.10-2.17 show plots of estimated Value at Risk versus out of sample returns using the proposed model (eight plots for four markets and two confidence levels). Second, the non-parametric model (Historical Simulation method)

and Normal Distribution based parametric models underestimate risk under most circumstances based on the results of this study. Third, the semi-parametric model (Extreme Value Theory) is the most accurate Value at Risk model for S&P 500, based on the results of out of sample test. However, for the other three markets (crude oil, gold, and Vanguard Long Term Bond Fund), the the semi-parametric model (Extreme Value Theory) overestimates risk as evidenced by the overly small numbers of exceedances in the out of sample test.

#### Expected Shortfall Out of Sample Test

First, all parametric models except for ARMA(1,1)-GARCH(1,1)-GED and the Normal Distribution based models (Appendix B) are statistically adequate for estimating Expected Shortfall at both 95% and 99% level, based on the out of sample test results. Second, the non-parametric model (Historical Simulation) and Normal Distribution based models are unreliable for estimating Expected Shortfall compared to other models, based on the number of failed t test. Third, the semi-parametric approach (Extreme Value Theory) is adequate for Expected Shortfall estimation for all four markets based on the test p values. However, due to the small sample size for the t test, the reliability of these p values is questionable for Expected Shortfall with 99% confidence level.

## CONCLUSION

### Main Conclusions

In summary, four conclusions are drawn based on the results of this study. First, Historical Simulation (HS) method and Normal Distribution based models tend to underestimate financial market risk (i.e. Value at Risk and Expected Shortfall), based on the test results in this study. Second, the parametric models using skewed and fat tailed distributions (such as Student T and General Error Distribution) are superior to the Normal Distribution based models for accurately estimating Value at Risk and Expected Shortfall. The proposed modified parametric model, ARMA(1,1)-GJR-GARCH(1,1)-SGED, is the most balanced Value at Risk model in this class, as it is the only model that is considered statistically adequate for estimating risk for all four markets. This is advantageous as a Value at Risk model because it opens up the possibility of measuring the risk of a variety of assets within a portfolio without having to change models. Third, modeling asymmetric volatility increases the accuracy of Value at Risk and Expected Shortfall models, as demonstrated by the proposed model, ARMA(1,1)-GJR-GARCH(1,1)-SGED, in estimating risk in S&P 500. Finally, the semi-parametric model (EVT) is most accurate Value at Risk model for S&P 500 Index, likely due to fat tail behavior for both in sample and out of sample data.

### Implications, Limitations and Future Research

This study should be of value to researchers, risk managers and regulators, and analysts for selection of Value at Risk (VaR) and Expect shortfall (ES) models. To

accurately estimate Value at Risk for S&P 500 Index or equity indices that have similar distributional properties to S&P 500, the semi-parametric model, Extreme Value Theory (EVT) model is recommended based on this study. As an overall Value at Risk and Expected Shortfall model for risk in equity, commodity and bond markets, the proposed parametric model, AMRA(1,1)-GJR-GARCH(1,1)-SGED, is recommended based on the results of this study. This model is advantageous because it opens up the possibility of measuring the risk of a variety of financial assets within a portfolio without having to change models. In all cases, the Normal Distribution based parametric models (Appendix B) and Historical Simulation method are not recommended for estimating Value at Risk and Expected Shortfall due to the general underestimation of risk.

This study has a few limitations. First, it focuses on examining the accuracy and reliability of Value at Risk and Expected Shortfall models for only four individual markets (S&P 500 equity, crude oil, gold and Vanguard Long Term Bond Fund) as opposed to portfolios of assets from many countries and markets. Second, in relation to the first limitation, correlations of different markets are not considered in this study when computing Value at Risk and Expected Shortfall, though it would be expected that these correlations may considerably reduce risk for a portfolio of assets if the portfolio is well diversified, and could be examined in future research.

## REFERENCES

- Artzner, P. A. (1999). Coherent Measures of Risk. *Mathematical Finance*, 9(3), 203-228.
- Balkema, A. A., and De Haan, L. (1974). Residual Life Time at Great Age. *The Annals of Probability*, 792-804.
- Basel Committee. (2006). International Convergence of Capital Measurement and Capital Standards: A Revised Framework, Comprehensive Version. *Switzerland: Bank for International Settlements*.
- Berkowitz, J., and O'Brien, J. (2002). How Accurate are Value-at-Risk Models at Commercial Banks? *The Journal of Finance*, 57(3), 1093-1111.
- Berthelsen C. and Dicolo J. (2013). Commodities Squeeze Banks. *Wall Street Journal*, Retrieved March 3, 2014, from <http://online.wsj.com/news/articles/SB10001424127887324096404578352492477050724>
- Best, P. (1999). *Implementing Value at Risk*. New York: John Wiley & Sons.
- Black, F. (1976). Studies of Stock Market Volatility Changes. *Proceedings of the American Statistical Association, Business and Economic Statistics Section*, 177-181.
- Bollerslev, T. (1986). Generalized Autoregressive Conditional Heteroskedasticity. *Journal of Econometrics*, 31(3), 307-327.
- Crotty, J. (2009). Structural Causes of the Global Financial Crisis: A Critical Assessment of the 'New Financial Architecture'. *Cambridge Journal of Economics*, 33(4), 563-580.
- Ding, Z., Granger, C. W., and Engle, R. F. (1993). A Long Memory Property of Stock Market Returns and A New Model. *Journal of Empirical Finance*, 1(1), 83-106.
- Duffie, D., and Pan, J. (1997). An Overview of Value at Risk. *The Journal of Derivatives*, 4(3), 7-49.
- Embrechts, P., Klüppelberg, C., and Mikosch, T. (1997). Modelling External Events. *Applications of Mathematics: Stochastic Modelling and Applied Probability*, 33.
- Engle, R. F. (1982). Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation. *Econometrica: Journal of the Econometric Society*, 987-1007.
- Engle, R. F., and Ng, V. K. (1993). Measuring and Testing the Impact of News on Volatility. *The Journal of Finance*, 48(5), 1749-1778.
- Fan, Y., Zhang, Y., Tsai, H., and Wei, Y. (2008). Estimating 'Value at Risk' of Crude Oil

Price and Its Spillover Effect Using the GED-GARCH Approach. *Energy Economics*, 30(6), 3156-3171.

Fernández, C., and Steel, M. F. (1998). On Bayesian Modeling of Fat Tails and Skewness. *Journal of the American Statistical Association*, 93(441), 359-371.

Giller, G. (2005). A generalized error distribution. *Giller Investments Research Note*, 20031222(1).

Giot, P., and Laurent, S. (2004). Modelling Daily Value-at-Risk Using Realized Volatility and ARCH Type Models. *Journal of Empirical Finance*, 11(3), 379-398.

Glosten, L. R., Jagannathan, R., and Runkle, D. E. (1993). On the Relation between the Expected Value and the Volatility of the Nominal Excess Return on Stocks. *Journal of Finance*, 48(5), 1779-1801.

Gouriéroux, C. (1997). *ARCH Models and Financial Applications*. Vol.1. New York: Springer.

Hansen, P. R., and Lunde, A. (2005). A Forecast Comparison of Volatility Models: Does Anything Beat A GARCH (1, 1)? *Journal of Applied Econometrics*, 20(7), 873-889.

Hentschel, L. (1995). All in the Family Nesting Symmetric and Asymmetric Garch Models. *Journal of Financial Economics*, 39(1), 71-104.

Holton, G.A. (2002). History of Value-at-Risk: 1922-1998. Working paper. *Contingency Analysis, Boston*. Retrieved February 5, 2014, from <http://128.118.178.162/eps/mhet/papers/0207/0207001.pdf>

Huang, Y. C., and Lin, B. (2004). Value-at-Risk Analysis for Taiwan Stock Index Futures: Fat Tails and Conditional Asymmetries in Return Innovations. *Review of Quantitative Finance and Accounting*, 22(2), 79-95.

Jorion, P. (1996). Risk2: Measuring the Risk in Value at Risk. *Financial Analysts Journal*, 47-56.

J.P. Morgan and Reuters (1996). *RiskMetrics Technical Document*. Retrieved October 5, 2013, from [www.jpmorgan.com](http://www.jpmorgan.com)

Kuester, K., Mittnik, S., and Paolella, M. S. (2006). Value-at-Risk Prediction: A Comparison of Alternative Strategies. *Journal of Financial Econometrics*, 4(1), 53-89.

McNeil, A. J. (1999). Extreme Value Theory for Risk Managers. *Departement Mathematik ETH Zentrum*. Retrieved April 5, 2012, from <http://www.macs.hw.ac.uk/~mcneil/ftp/cad.pdf>



- McNeil, A. J., and Frey, R. (2000). Estimation of Tail-Related Risk Measures for Heteroscedastic Financial Time Series: An Extreme Value Approach. *Journal of Empirical finance*, 7(3), 271-300.
- Nelson, D. B. (1991). Conditional Heteroskedasticity in Asset Returns: A New Approach. *Econometrica: Journal of the Econometric Society*, 347-370.
- Ouyang, Z. (2009). Model Choice and Value-at-Risk Estimation. *Quality and Quantity*, 43(6), 983-991.
- Ozun, A., Cifter, A., and Yilmazer, S. (2010). Filtered Extreme-Value Theory for Value at Risk Estimation: Evidence from Turkey. *Journal of Risk Finance*, 11(2), 164-179.
- Palm, F. C. (1996). 7 GARCH Models of Volatility. *Handbook of Statistics*, 14, 209-240.
- Pickands III, J. (1975). Statistical Inference Using Extreme Order Statistics. *The Annals of Statistics*, 119-131.
- Tang, H., Chun, K. C., and Xu, L. (2003). Finite Mixture of ARMA-GARCH Model for Stock Price Prediction. *Proc. of 3<sup>rd</sup> International Workshop on Computational Intelligence in Economics and Finance (CIEF2003), North Carolina, USA*
- Taylor, S. J. (2005). *Asset Price Dynamics, Volatility, and Prediction*. Princeton University Press.
- Ünal, G. (2011). Value-at-Risk Forecasts: A Comparison Analysis of Extreme-Value Versus Classical Approaches. *The Journal of Risk Model Validation*, 5(3), 59-76.

## APPENDIX A: ABBREVIATION LIST

ARCH	Autoregressive Conditional Heteroscedasticity
ARMA	Autoregressive Moving Average
BCBS	Basel Committee on Banking Supervision
CDF	Cumulative probability Density Function
CVAR	Conditional Value at Risk
CTE	Conditional Tail Expectation
ES	Expected Shortfall
EVT	Extreme Value Theory
EWMA	Exponentially Weighted Moving Average
GARCH	Generalized Autoregressive Conditional Heteroscedasticity
GED	General Error Distribution
GJR-GARCH	Glosten Jagannathan Runkle (GARCH)
GPD	General Pareto Distribution
HS	Historical Simulation
PDF	Probability Density Function
SGED	Skewed General Error Distribution
STD	Student T Distribution
SSTD	Skewed Student T Distribution
VaR	Value at Risk

## APPENDIX B: LIST OF VALUE AT RISK AND EXPECTED SHORTFALL MODEL TYPES

### *Description of 13 Value at Risk and Expected Shortfall Models*

**Non-Parametric Model (1):** Historical Simulation (HS)

**Semi-Parametric Model (1):** Extreme Value Theory (EVT)

**Parametric Models (11):**

EWMA-Norm (Normal Distribution)

ARMA(1,1)-GARCH(1,1)-Norm (Normal Distribution)

ARMA(1,1)-GARCH(1,1)-STD (Student T Distribution)

ARMA(1,1)-GARCH(1,1)-SSTD (Skewed Student T Distribution)

ARMA(1,1)-GARCH(1,1)-GED (General Error Distribution)

ARMA(1,1)-GARCH(1,1)-SGED (Skewed General Error Distribution)

ARMA(1,1)-GJR-GARCH(1,1)-Norm (Normal Distribution)

ARMA(1,1)-GJR-GARCH(1,1)-STD (Student T Distribution)

ARMA(1,1)-GJR-GARCH(1,1)-SSTD (Skewed Student T Distribution)

ARMA(1,1)-GJR-GARCH(1,1)-GED (General Error Distribution)

ARMA(1,1)-GJR-GARCH(1,1)-SGED (Skewed General Error Distribution)

### *Subgroups of the 13 Value at Risk and Expected Shortfall Models*

**Parametric Models with Asymmetric Volatility Model:**

ARMA(1,1)-GJR-GARCH(1,1)-Norm (Normal Distribution)

ARMA(1,1)-GJR-GARCH(1,1)-STD (Student T Distribution)

ARMA(1,1)-GJR-GARCH(1,1)-SSTD (Skewed Student T Distribution)

ARMA(1,1)-GJR-GARCH(1,1)-GED (General Error Distribution)

ARMA(1,1)-GJR-GARCH(1,1)-SGED (Skewed General Error Distribution)

**Normal Distribution Based Parametric Models:**

EWMA-Norm (Normal Distribution)

ARMA(1,1)-GARCH(1,1)-Norm (Normal Distribution)

ARMA(1,1)-GJR-GARCH(1,1)-Norm (Normal Distribution)

**Student T Distribution Based Parametric Models:**

ARMA(1,1)-GARCH(1,1)-STD (Student T Distribution)

ARMA(1,1)-GARCH(1,1)-SSTD (Skewed Student T Distribution)

ARMA(1,1)-GJR-GARCH(1,1)-STD (Student T Distribution)

ARMA(1,1)-GJR-GARCH(1,1)-SSTD (Skewed Student T Distribution)

**General Error Distribution (GED) Based Parametric Models:**

ARMA(1,1)-GARCH(1,1)-GED (General Error Distribution)

ARMA(1,1)-GARCH(1,1)-SGED (Skewed General Error Distribution)

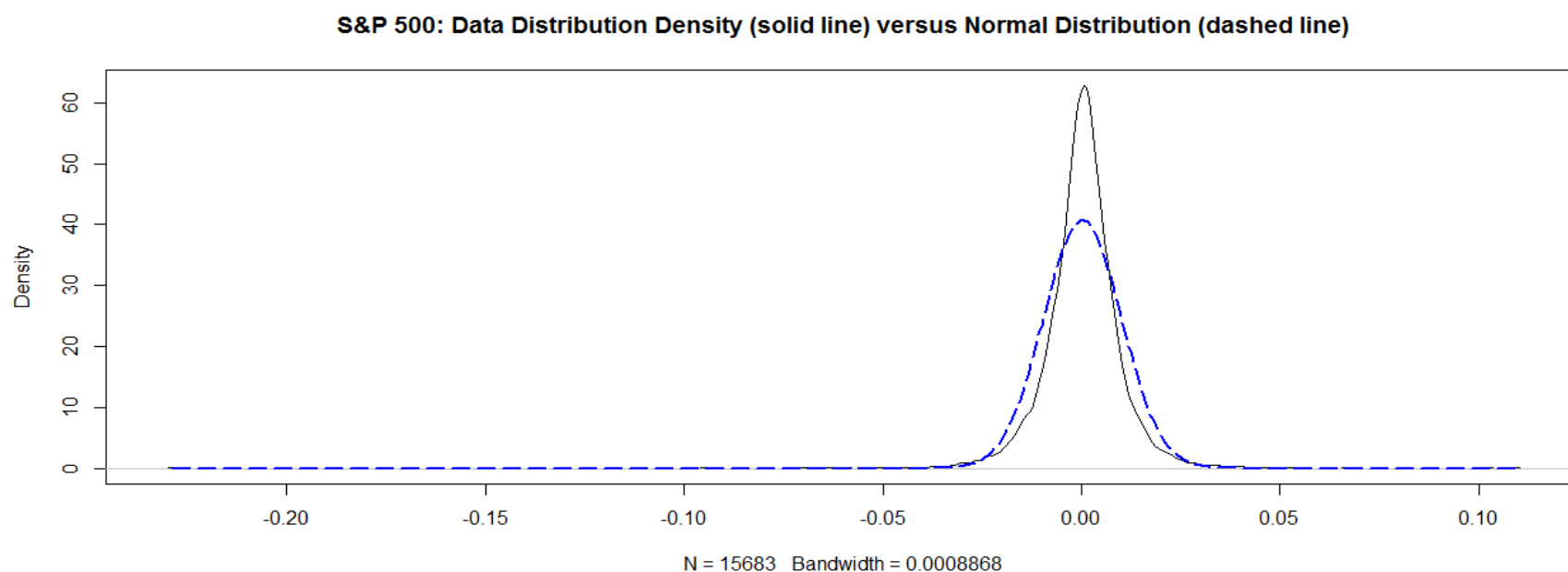
ARMA(1,1)-GJR-GARCH(1,1)-GED (General Error Distribution)

ARMA(1,1)-GJR-GARCH(1,1)-SGED (Skewed General Error Distribution)

**The New Proposed Parametric Model:**

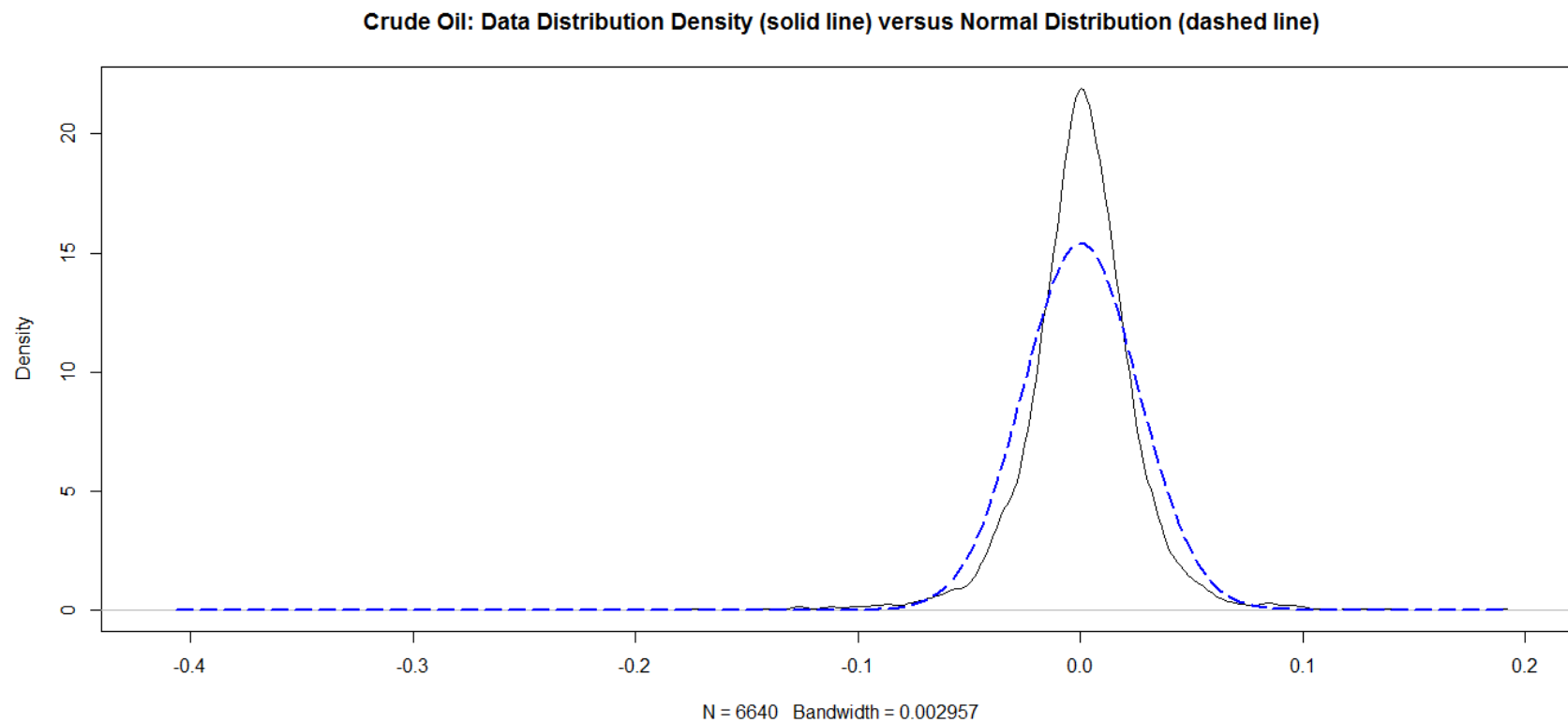
ARMA(1,1)-GJR-GARCH(1,1)-SGED (Skewed General Error Distribution)

## APPENDIX C: GRAPH OF DAILY RETURN PROBABILITY DENSITY VERSUS NORMAL DISTRIBUTION PROBABILITY DENSITY FOR S&P 500, OIL, GOLD, AND VANGUARD LONG TERM BOND FUND



Notes: Solid curve represents the kernel probability density of data. Dashed curve represents the kernel probability density of Normal distribution with the same mean and standard deviation.

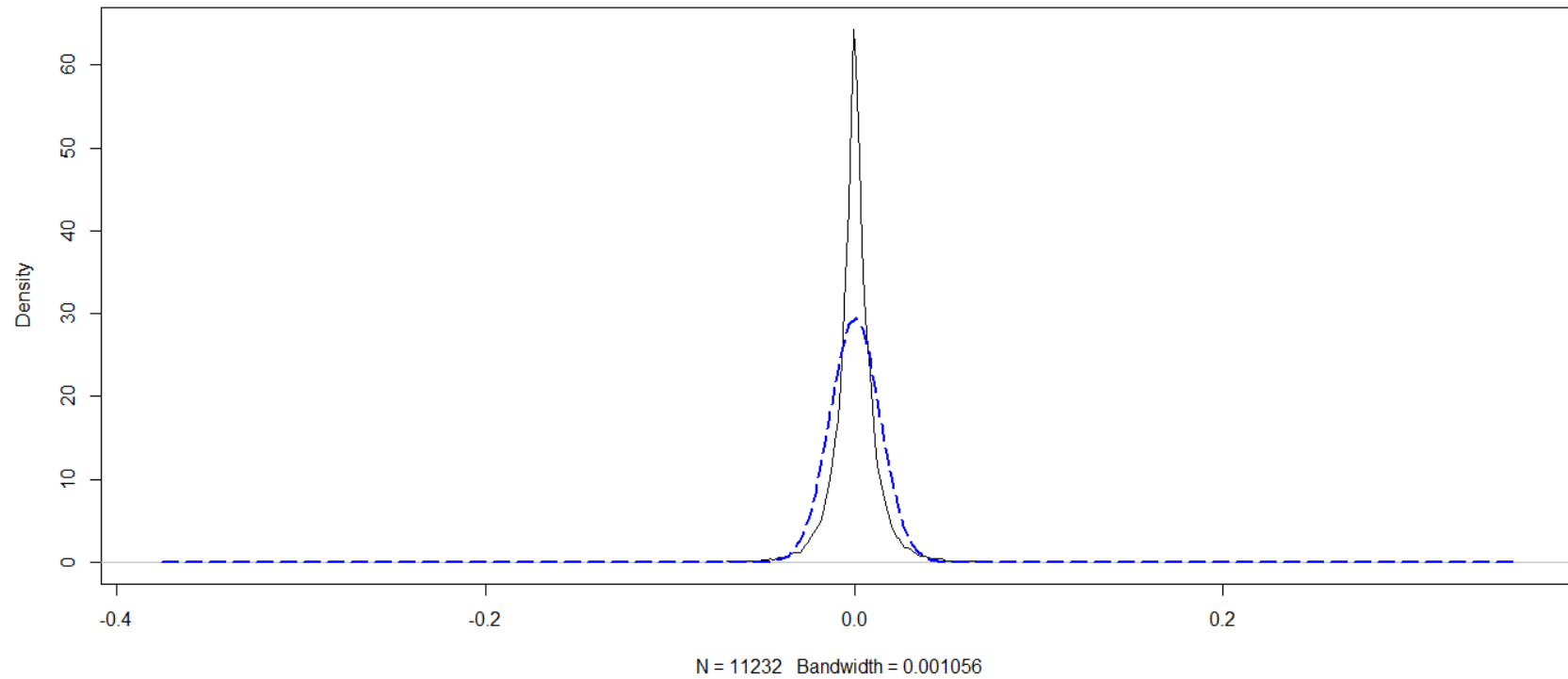
Data: S&P 500 daily return from 01/01/1950 to 04/30/2012.



Notes: Solid curve represents the kernel probability density of data. Dashed curve represents the kernel probability density of Normal distribution with the same mean and standard deviation.

Data: Crude oil price daily return from 01/02/1986 to 04/30/2012.

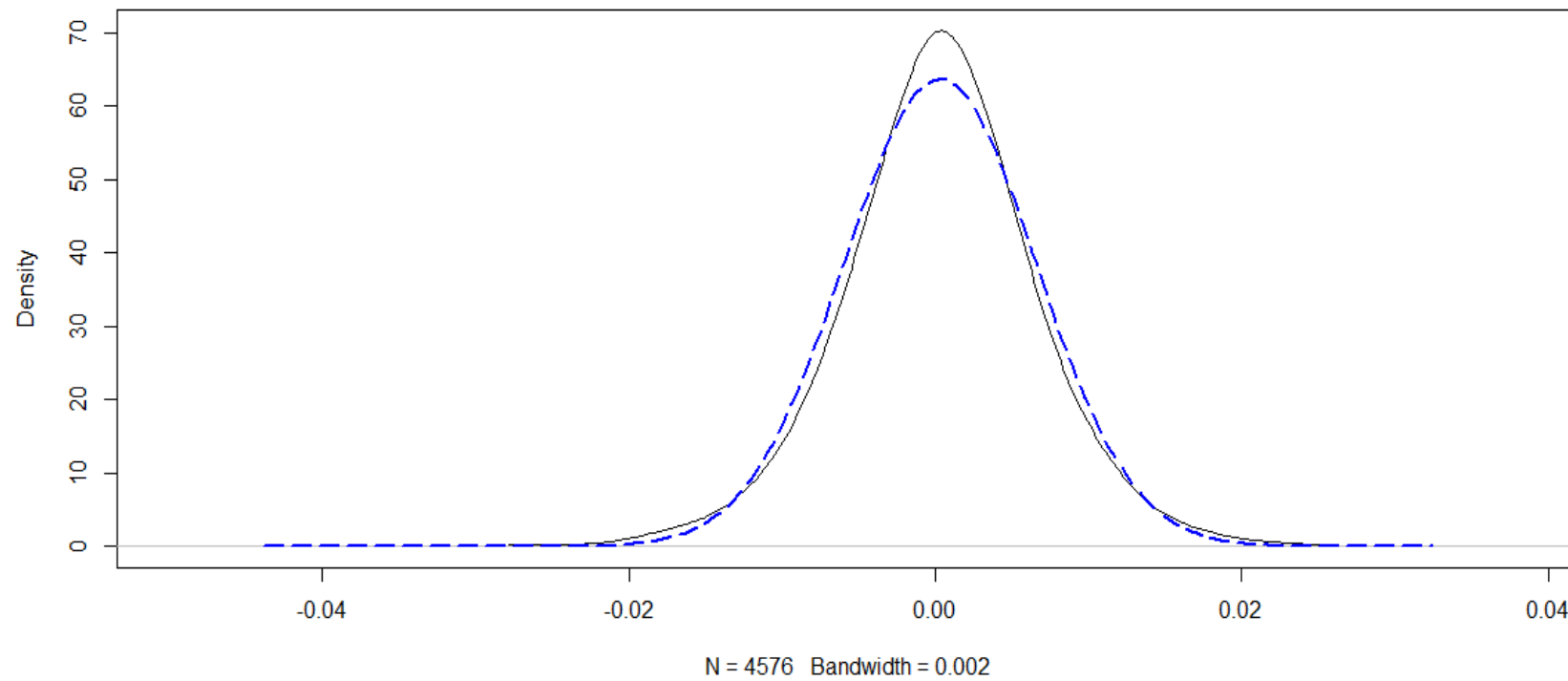
**Gold: Data Distribution Density (solid line) versus Normal Distribution (dashed line)**



Notes: Solid curve represents the kernel probability density of data. Dashed curve represents the kernel probability density of Normal distribution with the same mean and standard deviation.

Data: Gold daily return from 04/02/1968 to 04/30/2012.

**Vanguard Long Term Bond Fund: Data Distribution Density (solid line) versus Normal Distribution (dashed line)**



Notes: Solid curve represents the kernel probability density of data. Dashed curve represents the kernel probability density of Normal distribution with the same mean and standard deviation.

Data: Vanguard Long Term Bond Fund (VBLTX), daily return from 03/01/1994 to 04/30/2012.