Essays on Portfolio Optimization, Volatility Modelling and Risk Measurement

Liyuan Chen

PhD

University of York

Economics

September 2017

Abstract

This study comprises of three essays on the subject of financial risk management with applications in the fields of portfolio optimization, continuous and discrete time stochastic volatility (SV) modelling. We jointly consider two risk measures: Value-at-Risk (VaR) and conditional Value-at-Risk (CVaR) to measure the financial market risk. In order to model the distribution of financial asset returns which is characterized by skewness, heavy tails and leptokurtosis, we employ the Asymmetric Laplace distribution (ALD) in the first and third essay while constructing the risk model on the basis of the Heston stochastic volatility (SV) model in the second essay. Specifically, in the first essay, we provide a comprehensive empirical examination of the viability of the new proposed Mean-CVaR-Skewness optimization model under ALD by Zhao et al. (2015). In addition, we propose the Mean-VaR-Skewness model under ALD by employing VaR as risk measure. The closed-form solution of the two optimization models is shown to be consistent and is obtainable by using the Lagrange Multiplier approach. In the second essay, we construct the VaR and CVaR models for the financial dynamics that do not have a closed-form probability density function. The only input required in our approach is the knowledge of the characteristic function of the underlying asset. In the numerical analysis, we investigate the elements that could impact the VaR and CVaR approximations in the Heston model. The third essay contributes to the existing literature by extending the ALD (Kotz et al., 2001) to the return error term of a standard discrete time SV model. We give the closed-form VaR and CVaR formulas for oil supply and demand. As additional contribution, we propose a new scale mixture of uniform (SMU) representation for the AL density so that the model can be implemented efficiently within the Bayesian Markov Chain Monte Carlo framework.

Contents

List of Tables

List of Figures

Acknowledgements

The completion of my PhD has been a memorable moment in my life. It is no exaggeration to say that my years at the University of York have been the best of my life. In the Department of Economics and Related Studies (DERS), i was able to learn a huge amount from extremely knowledgeable colleagues. I am grateful to many individuals who have provided me with their sincere help in my educational and professional development, and who have given assistance in my daily life.

My deepest gratitude goes first and foremost to Dr. Paola Zerilli, my supervisor, for her constant support, warm encouragement and invaluable advice throughout my years at DERS. Without her consistent and illuminating instruction, this thesis could not have reached its present form. I also sincerely appreciate the help and guidance from my committee members including Professor Peter N. Smith and Professor Gulcin Ozkan during the past years of research.

I would like to thank the faculty and fellow students in my department for the opportunities to share and discuss innovative ideas with them in seminars and workshops. Thanks to Professor Peter Spencer (University of York), Professor Wolfgang Kürsten (University of Jena) and Professor Richard Gerlach (University of Sydney) for their constructive comments and suggestions on my work.

Throughout my studies, i have received wholehearted support and endless love from my parents and beloved sister Yangyang. My sincere gratitude and love goes to you. No words can express enough my gratitude towards my fiancée, Xinyu Xiao, thanks for your companionship and great confidence in me through all these years.

Declaration

I declare that the thesis I have presented for examination for the degree of PhD of the University of York is solely my own work other than where I have clearly indicated that it is a joint work with others.

This work has not previously been presented for an award at this, or any other, University. The copyright of this thesis rests with the author. All sources are acknowledged as references. This thesis may not be reproduced without my prior written consent. I warrant that this authorization does not, to the best of my belief, infringe the rights of any third party. I confirm that chapter 4 was jointly co-authored with Dr. Paola Zerilli. I completed chapter 4 independently under the guidance of Dr. Zerilli, who then later contributed to the work when it was submitted as a paper for publication to *Energy Economics*.

Chapter 1

Introduction

This thesis consists of three essays on financial risk measurement with applications in the fields of portfolio optimization, continuous and discrete time stochastic volatility models. In this thesis, we consider the measurement of a specific financial risk: the market risk. Market risk is due to unforeseen changes of asset prices, such as stock prices and commodity prices. Choosing an appropriate risk measure is of great practical and regulatory importance, as exemplified in the Basel II (2004) and Basel III (2010) accord, which recommends Value-at-Risk (VaR) as a risk measure to set capital requirements for banks' daily operations.

As a measure of market risk, VaR has been widely developed since its introduction in RiskMetrics by JP Morgan (1994). It is defined as the maximum potential loss of an underlying asset at a specific probability level over a certain horizon. Despite its popularity, an obvious and distinctive limitation of the VaR approach is that it only specifies the maximum one can lose at a given risk level, but provides no indication for how much more than VaR one can lose if extreme tail events happen. This may lead to an equivalent VaR estimate for two different positions, though they have completely different risk exposures. Artzner et al. (1999) proposed the concept of coherent risk measure, which has become the paradigm of risk measurement. To further illustrate this concept, let us consider W_1 and W_2, which represent the return of two portfolios. φ is a risk measure over a given horizon, so φ is said to be a coherent risk measure if the following four axioms are satisfied:

(i) *Monotonicity* : If $W_2 \geq W_1$, then $\varphi(W_2) \leq \varphi(W_1)$.

(ii) *Subadditivity* : $\varphi(W_1 + W_2) \leq \varphi(W_1) + \varphi(W_2)$.

(iii) *Positive homogeneity* : $\varphi(\alpha W_1) = a\varphi(W_1)$ for $a > 0$.

(iv) *Translation invariance* : $\varphi(W_1 + n) = \varphi(W_1) - n$.

The most important criterion a risk measure needs to satisfy in order to be considered "valid" is property (ii). Subadditivity, the idea behind portfolio diversification, reflects the expectation that the risk of a merged portfolio cannot be greater than the sum of the risks of its individual constituents. This spells trouble for VaR, because it does not satisfy the subadditivity property. A good alternative is conditional Value-at-Risk (CVaR), which is a coherent risk measure and retains the benefits of VaR in terms of the capability to define quantiles of the loss distribution. Hence, considering the fact that VaR is still employed by Basel II and III, as well as the additional desirable property of CVaR, our thesis contributes to the construction and modelling of these two risk measures in several fields.

In Chapter 2, we provide a comprehensive examination of the viability of a theoretical portfolio optimization model first proposed by Zhao et al. (2015), namely, the Mean-CVaR-Skewness model with incorporation of Asymmetric Laplace distribution (ALD) and an extension of this model by employing VaR as the risk measure. The superiority of Zhao et al.'s (2015) model is that it considers a series of stylized facts in realistic scenarios that relate to portfolio optimization analysis. For example, the importance of higher moments, the fat-tailed, skewed and leptokurtic features of financial asset returns, and the drawbacks of using Normal-based risk measures are all scenarios that Zhao et al.'s (2015) model considers. Our contribution in this Chapter is threefold. First, we propose a three-dimensional optimization model under ALD using VaR as a risk measure, called the Mean-VaR-Skewness model. We find that the solution of this model (or the analytical optimal weight expression) is exactly the same as the solution proposed by Zhao et al. (2015). Both solutions can be obtained by transforming the multi-objective optimization functions into an equivalent, single-objective, quadratic programming problem through the Lagrange Multiplier method. Second, as the main objective in this chapter, we empirically investigate the viability of Mean-CVaR-Skewness model under ALD in terms of the portfolio CVaR performance from in-sample, Monte-Carlo simulation and out-of-sample forecasting perspectives. This process is realized by conducting comparative analysis with other classical approaches, and the accuracy judgment is concluded according to a number of statistical measures. Notably, there have not been any attempts to explore the practical implications of this model empirically. Therefore, we study the empirical performance of this model, optimal portfolio allocations, and portfolio risk-adjusted returns in different economic periods, with a special focus on the variations of portfolio risk in the period of the 2007-2008 global financial crisis. The evaluation of economic significance represents the third contribution.

Chapter 3 addresses the problems of measuring parametric VaR and CVaR for heavy-tailed distributions (i.e. stable distribution) or the financial dynamics (i.e. Hull and White, 1987; Stein and Stein, 1991; Heston, 1993; Schöbel and Zhu, 1999) that do not have analytical representation for their probability density functions (*p.d.f.*). We build a general framework for the VaR and CVaR computations in a generalized Fourier transform scheme to apply to those financial dynamics that are fully characterized by their characteristic function, considering the fact that an asset distribution always has a characteristic function but not a *p.d.f.*. As the first paper to build the connection between risk measures and models defined by characteristic function, our paper differs from that of Bormetti et al. (2010) in several aspects, such as the method used to construct VaR and CVaR formula, in the adoption of characteristic function forms in risk measurement, and, most importantly, in the differences of concentrations on data analysis.[1] In our numerical experiment, we focus on the determinants of VaR and CVaR approximations by employing the Heston stochastic volatility model in a trapezoidal integration scheme. We investigate how the settings of grid size space could impact the accuracy of VaR and CVaR estimates. In addition, we provide evidence for the influence of the potential parameters in the Heston model on the movements of VaR and CVaR approximations given a series of risk levels.

[1]For details see the introduction of Chapter 3.

Considering the remarkably volatile nature of world crude oil markets and potential return uncertainties caused by the instability of oil prices, in Chapter 4, we propose a new parametric approach to estimate the market risk for crude oil prices using VaR and CVaR from the perspective of both oil supply and demand. To capture the potential heavy-tailed and leptokurtic features of oil return series, we employ the ALD to model the extreme tail risks. A standard discrete stochastic volatility (SV) model, which treats the latent volatilities as an unobservable stochastic process, is considered to characterize the behavior of return volatility with extension of adopting ALD as the distribution of return errors. Our theoretical contribution to the literature is reflected in the construction of the SV-ALD model, and the VaR and CVaR formula derivations for both oil supply and demand. Since the likelihood function of volatilities in the SV-ALD model is intractable, the Bayesian approach, which uses a simulation-based Markov Chain Monte Carlo (MCMC) algorithm, is employed in our paper for statistical inference.

However, implementing the SV-ALD model is often problematic. As in the framework of Bayesian MCMC, the full conditional posterior distributions are of non-closed forms. As such, our second contribution in terms of improving estimation methodology is to propose a new scale mixture of uniform (SMU) representation for the AL density. The use of SMU for scaled AL density is a data augmentation technique and its advantage is that some of the full conditional posterior distributions can be reduced to standard form, hence facilitating an efficient Gibbs sampling algorithm in the Bayesian MCMC framework. With this SMU representation, we can straightforwardly implement the SV-ALD model. Last, in the empirical analysis, a model comparison study from Bayesian statistical perspective is conducted between the SV-ALD model and the classical SV Normal model to test model fitting abilities for target oil return series. In addition, we investigate the market risk of two major oil markets using the SV-ALD model, along with an accuracy assessment by backtesting VaR and CVaR violations. Finally, the economic implications and applicability of the model are discussed.

Chapter 5 summarizes the thesis and discusses the avenues for future research.

Chapter 2

Portfolio Optimization in Higher Moments with Incorporation of Asymmetric Laplace Distribution

2.1 Introduction

The prevailing paradigm for optimal portfolio allocation is the seminal work by Markowitz (1952), which proposed the mean-variance model aimed at minimizing risk for a given level of expected return or, equivalently, maximizing expected return for a given level of risk. However, it remains controversial as to whether variance is appropriate for measuring risk and whether higher moments should be incorporated in the portfolio selection model.

The standard mean-variance model operates on the assumption that asset returns follow multivariate Normal distribution. Following Markowitz's seminal work, a series of optimization models were proposed such as, mean-lower partial moment model (Bawa and Lindenberg, 1977), mean-absolute deviation model (Konno and Yamazaki, 1991) and mean-semivariance model (Markowitz et al., 1993), etc. A distinct characteristic of those models is that only the first two moments of return distribution are considered. However, real financial asset returns may represent asymmetric leptokurtic features. Therefore, higher moments of return distributions cannot be neglected unless there is reason to believe that asset returns are normally distributed, investor's utility is a quadratic function of rate of return, or higher moments are irrelevant to investors' decision.[1] Samuelson (1970) introduces a critical work on the limitations of the mean-variance optimization model from the perspective of its non-practicability, showing the standpoint that higher moments are relevant to the portfolio selection problem. Arditti and Levy (1975) stress the importance of skewness in purchasing lottery tickets and pricing of stock. In addition to this, after replacing variance with absolute deviation as a risk measure, Konno et al. (1993) incorporate skewness in the optimization model and shed light on practical means to obtain a portfolio with large skewness. Konno and Suzuki (1995) extend the mean-variance model to the mean-variance-skewness model and show the importance of skewness in selecting an optimal portfolio. Therefore, in a multi-objective portfolio optimization problem, it is necessary to consider the role of portfolio skewness (see, e.g., Prakash et al., 2003; Jurczenko et al., 2005; Jondeau and Rockinger, 2006; Rubinstein et al., 2006.). Holding everything else constant, risk-averse investors should pursue portfolios that are right-skewed rather than left-skewed, i.e. they prefer fewer but high pay-offs rather than fewer but large losses.

Markowitz's model is being questioned with the use of variance as a risk measure (Artzner et al., 1999). Variance considering measures to positive and negative fluctuations of expected returns in the same way: this approach fails to capture the asymmetric features of real asset return distributions. On the other hand, Value at Risk (VaR) has been widely employed by regulators and financial institutions in real practice, because its representation of percentile losses is easy to understand (see, e.g., Wipplinger, 2007; Huang et al., 2009; Dimitrakopoulos et al., 2010). However, its limitations are obvious due to the undesirable properties such as lack of sub-additivity, which indicates that VaR is not a coherent risk measure (Artzner et al., 1999).[2] Besides, VaR as a risk control tool does not take into account the part of the distribution that falls beyond the confidence level, and thus minimizing

[1] On the other hand, the achievement of Markowitz's portfolio optimization model could not be denied as a number of classical models and works with less demanding in calculation are proposed successively based on the original ideal of that (i.e. CAMP and APT, etc.).

[2] That is, the VaR of a portfolio with two or more securities may be larger than the sum of the VaRs of the securities in the portfolio, which implies the potential failure of portfolio diversification.

VaR may increase extreme losses. Rockfeller and Uryasev (2000, 2002) thereafter proposed the Conditional Value-at-Risk (CVaR), also named the expected shortfall (ES), which is defined as the conditional expectation of losses exceeding VaR over a time period at a given confidence level. Its superiorities can be reflected in several aspects. CVaR is a coherent risk measure. Pflug (2000) demonstrates the convexity property of CVaR, which implies that CVaR is also a convex risk measure. In addition, CVaR implies less computational burden in modelling optimization than VaR as it can be expressed by a clear minimization formula (Rockfeller and Uryasev, 2000; Krokhmal et al., 2002). Such computational advantage of CVaR over VaR has provided a major stimulus for the development of CVaR as a key risk measure in portfolio optimization and other practical financial problems; see, for example, Zhu and Fukushima (2009), Stoyanov et al. (2013) and Dai and Wen (2014).

Despite growing interest in CVaR and higher moments of portfolio optimization, very few studies have considered higher moment (i.e. skewness) in the context of Mean-CVaR optimization model. Zhao et al. (2015) are a notable exception. They propose a theoretical Mean-CVaR-Skewness optimization model by adopting the Asymmetric Laplace distribution (ALD) to take into account the heavy-tailed and leptokurtic features of asset returns. The main advantage of using ALD for asset returns is that it allows for asymmetry and peakness observed in the data, and is thus more suitable and attractive for modelling fat-tailed and peaked return series, especially considering its capability for modelling multivariate asymmetric data from a portfolio management point of view (Kozubowski and Podgrski, 2001; Hürlimann, 2013). In addition, unlike other skewed distributions (i.e. stable Pareto distribution), ALD provides more flexibility and parsimony with fewer parameters to estimate.[3] Notably, there has not been any attempt to evaluate the model accuracy and performance in terms of modelling financial data. Therefore, this paper is expected to fill this gap.

By extending the framework of Zhao et al.'s (2015) model, we use VaR as a

[3]Studies enrich the application of Asymmetric Laplace law in financial fields can also refer to Ayebo and Kozubowski (2003), Kotz and Van Dorp (2005), Jayakumar and Kuttykrishnan (2007), Komunjer (2007) and Chen et al. (2012).

risk measure to study a new Mean-VaR-Skewness optimization model under ALD considering the fact that VaR is still widely adopted by regulators and financial intermediaries (see, e.g., Cuoco and Liu, 2006; Natarajan et al., 2008; Adrian and Shin, 2010). Alexander and Baptista (2004) study the portfolio selection implications by comparing the imposition of VaR and CVaR constraints on the mean-variance model. They find that for a given confidence level to control slightly risk-averse agents, the CVaR constraint is more effective than the VaR constraint. To help investors hedge their portfolios, Yamai and Yoshiba (2005) argue that CVaR is a better risk measure than VaR, but when asset returns are fat-tailed, the estimation error of CVaR is much greater than that of VaR. As a consequence, a larger sample size for CVaR estimation is a prerequisite to improve the accuracy level. These findings reveal that a single risk measure should not dominate the risk management decision, but rather that supplementing one with another represents an efficient way to monitor the tail risks. Moreover, extensive literatures have been developed to improve risk measures' estimation error and handle the unstable nature of asset return distributions in optimization problems, such as robust optimization in the mean-variance framework (DeMiguel and Nogales, 2009; Delage and Ye, 2010; Chen et al., 2011), robust VaR optimization (Ghaoui et al., 2003; Natarajan et al., 2008) and robust CVaR optimization (Quaranta and Zaffaroni, 2008; Huang et al., 2010). However, it is important to emphasize that unless the return distribution is known, the actual value of portfolio VaR/CVaR cannot be obtained, as argued by Natarajan et al. (2008). In this sense, the accuracy of the parametric portfolio VaR/CVaR relies heavily on the modelling ability of prescribed innovation of asset returns. In other words, the modelling ability of ALD in the parametric optimization model determines the model performance.

The goal of this paper is threefold. First, we are optimizing the reduced Mean-CVaR-Skewness model under ALD using the Lagrange multiplier method in order to find a closed-form solution rather than using the software, as Zhao et al. (2015) claimed. Following Zhao et al.'s (2015) work, we develop a Mean-VaR-Skewness optimization model under ALD, the solution of which is shown to be exactly the same as that of Zhao et al.'s (2015) model. We then calculate the VaR of a single asset in the portfolio based on Normal distribution in order to further explore the empirical performance of the single asset if using the conventional method. To evaluate the portfolio performance, we compare the new-constructed portfolio VaR model under ALD with conventional portfolio VaR model under Normal distribution and the benchmark historical portfolio VaR model.

Second, computational experiments are conducted to test the viability of the Mean-CVaR-Skewness model under ALD in terms of the portfolio CVaR performance by employing three different methods. The computed in-sample portfolio CVaR estimates under Mean-CVaR-Skewness model based on ALD are compared with historical CVaR estimates and estimates from classical Mean-CVaR model, where the portfolio return follows a jointly Normal distribution. In addition, a numerical experiment using Monte Carlo simulated data is introduced to evaluate the model performance. In order to test the accuracy of the above methods, three loss functions, root-mean-squared-error (RMSE), mean-absolute-error (MAE) and sum of squared relative errors (SSRE), are computed correspondingly. Moreover, the out-of-sample portfolio CVaR performance of Mean-CVaR-Skewness model and Mean-CVaR model are examined using monthly return data. The results of those experiments consistently confirm the practicability of the Mean-CVaR-Skewness optimization model under ALD in modelling extreme tail events, and it is thus suitable to measure tail risks when the distribution of portfolio asset returns demonstrate heavy-tailed and skewed features. In this regard, the Mean-CVaR-Skewness optimization model under ALD offers an efficient management tool in minimizing the portfolio CVaR, and at the same time maximizing the skewness for a given expected portfolio return. It should be noted that the analysis for both Mean-VaR-Skewness model and Mean-CVaR-Skewness models is implemented across three different economic periods in order to observe the consistency of model performance and to learn how people's investment behaviour is changing under different economic environments based on the target optimization models.[4]

[4] Saranya and Prasanna (2014) compare the performance of three models (mean-variance-skewness, mean-variance and mean-variance-skewness-kurtosis) using non-Normal stocks from the BSE 200 across bullish, bearish and crisis periods, showing that mean-variance-skewness and mean-variance-skewness-kurtosis model provide higher returns than that of mean-variance model.

Third, unlike conventional portfolio risk-adjusted return measures, i.e. Sharpe ratio, in order to evaluate the performance of a portfolio in the context of the Mean-CVaR-Skewness optimization model, we use CVaR deviation to replace standard deviation in the denominator and compare the calculated Sharpe-like ratios of the portfolio at the three economic periods for various confidence levels. Moreover, we discuss the optimal portfolio configurations and Mean-CVaR-Skewness efficient frontiers, and examine the economic implications under ALD assumption.

The structure of this paper is organized as follows. Section 2 studies the VaR, CVaR and skewness of a single asset under univariate ALD as well as the portfolio VaR, portfolio CVaR and portfolio skewness under multivariate ALD. Section 3 describes the construction of Mean-CVaR-Skewness model under ALD and places the proposed Mean-VaR-Skewness model under ALD in this section. Section 4 presents the algorithm to solve the optimization models. In section 5, we briefly explore the empirical performance of Mean-VaR-Skewness model under ALD. In addition, three experiments are implemented using in-sample, Monte Carlo simulation and out-of-sample methods to test the performance of Zhao et al.'s (2015) optimization model. Section 6 gives the optimal portfolio configuration and risk-adjusted performance. Section 7 concludes.

2.2 Preliminaries of VaR, CVaR and Skewness under ALD

This section first presents the assumptions for the discussed portfolio optimization models. Then, we introduce the single asset VaR, CVaR and skewness with incorporation of ALD, followed by the model specification of the portfolio VaR, CVaR and skewness using multivariate ALD.

2.2.1 Assumptions

A number of standard assumptions in the portfolio selection model are presented by Lai (1991). Similar hypotheses have been used by Kemalbay et al. (2011) and

Saranya et al. (2014). In this paper, the assumptions can be summarized as follows:

Assumption 2.1. *Investors are risk averse. Investors are more likely to invest their money on the assets with lower risk and they always pursue the maximization of their expected utility of personal wealth. Besides, investors are more concerned about the downside risk as it gives investors closer feelings for the risk of their investments.*

Assumption 2.2. *Portfolio consists of* n *(*$n > 1$*) risky assets and investors do not have access to a riskless asset, implying that the portfolio weights must sum to one.*

Assumption 2.3. *Single asset return in the portfolio jointly follows the multivariate Asymmetric Laplace Distribution, the correlation of them is a positive definite covariance matrix.*

Assumption 2.4. *There are no taxes and transaction costs in markets.*

Assumption 2.5. *Short selling is allowed, implying negative weights can occur.*

2.2.2 VaR, CVaR and Skewness of single asset under ALD

As developed by J.P. Morgan in 1994, VaR provides a measure for the largest potential loss that can be incurred by an investor over a given time period at a certain confidence level. It is a threshold value, such that probability of loss exceed the given amount of asset values.

Specifically, if the VaR of a portfolio return at a given confidence level $(1-\alpha) \in (0,1)$ is defined as $VaR_{1-\alpha}$. Then the mathematical formula of VaR can be shown as: $Prob(L \geq VaR_{1-\alpha}|\Omega_t) = \alpha$, where Ω_t denotes the information set up to time t, α is risk probability level and L are the possible losses of the portfolio for a holding time period. For CVaR, the mathematical expression can be formulated as: $CVaR_{1-\alpha} = E\{L|L \geq VaR_{1-\alpha}\}$. Given a random variable X denoting random portfolio returns, the equation $X = -L$ holds. Hence, CVaR can be equivalently written as: $CVaR_{1-\alpha} = -E\{X|X \leq -VaR_{1-\alpha}\}$. The skewness, which plays an important role in making investment choices, is a measure for the asymmetry of probability distribution of asset returns. The analytic formula is given by: $s = E[(X-\mu)^3/\sigma^3] = E[(X-\mu)^3]/E[(X-\mu)^2]^{(3/2)}$, where s denotes the skewness of

X, u is the mean, σ is the standard deviation and E is the expected value operator. A positive skewness indicates that the poor returns are more likely to occur but losses are small, while a negative skewness indicates that very high returns are less frequent and extreme losses are possible.

Asymmetric Laplace Distribution (ALD). As a member of the skewed distribution families, ALD has attracted many attentions in modelling financial asset returns as it provides more flexibilities, allows for asymmetry and symmetry (in which case ALD reduce to a symmetric Laplace distribution). In addition, its advantage in terms of application can be attributed to the simplicity with respect to mathematical computations. Explicit definition, distribution function and other characteristics of ALD allows for straightforward calculations, multivariate setting of ALD motivate further extensions like conventional Normal distribution, and the closed form expression of *p.d.f.* and *c.d.f.* ideally facilitate the implementation of model estimation, simulation and quantile calculations, etc. Before specifying the optimization models, we first introduce the properties of ALD.

Definition 2.6. *A random variable X is said to follow an Asymmetric Laplace Distribution if the characteristic function of X can be defined as:*

$$\psi(t) = E[e^{itX}] = \frac{1}{1 + \frac{\tau^2 t^2}{2} - i\mu t} \tag{2.1}$$

where i is the imaginary unit, $t \in R$ is the argument of the characteristic function, τ is the scale parameter with $\tau > 0$ and μ is the mean of X. Then, we have $X \sim AL(\mu, \tau)$.[5]

Proposition 2.7. *Let $f_{\kappa,\tau}(x)$ denote the p.d.f of an $AL^*(\kappa, \tau)$ distribution, then, the $f_{\kappa,\tau}(x)$ can be expressed as:*

$$f_{\kappa,\tau}(x) = \begin{cases} \frac{\sqrt{2}}{\tau}\frac{\kappa}{1+\kappa^2}\, exp(-\frac{\sqrt{2}\kappa}{\tau}x) & x \geq 0 \\ \frac{\sqrt{2}}{\tau}\frac{\kappa}{1+\kappa^2}\, exp(\frac{\sqrt{2}}{\tau\kappa}x) & x < 0 \end{cases} \tag{2.2}$$

[5]Note that this characteristic function is a standardized form with location parameter $\theta = 0$. A simplified notation for the distribution of X can be written as $AL(\mu, \tau)$ and $AL^*(\kappa, \tau)$ to replace of $AL(0, \mu, \tau)$ and $AL^*(0, \kappa, \tau)$, respectively. More details can refer to Kotz et al. (2001).

where κ *is the skewness parameter with* $\kappa = \sqrt{2}\tau/(\mu + \sqrt{\mu^2 + 2\tau^2}) > 0$, x *is a random variable representing return of assets. Then,* x *is said to be AL distributed if its p.d.f. is given by (2.2).*

Remark 2.8. For $\kappa = 1$, the asymmetric Laplace distribution degenerates to a symmetric Laplace distribution.

Remark 2.9. For $\kappa \neq 1$, the mode, median and mean of the ALD satisfies the following inequalities:

$$If\ \kappa \leq 1, then\ Mode \leq Median \leq Mean$$
$$If\ \kappa \geq 1, then\ Mode \geq Median \geq Mean$$

Parameter κ controls the probability assigned to each side of location parameter θ. $\kappa = 1$ means the probability of the two sides is equivalent and the distribution is symmetric with respect to θ. When $\kappa > 1$, the left tail of the distribution is thicker than the right, and the right tail is thicker than the left if $0 < \kappa < 1$. This is shown in Figure 2.1 which presents the impact of the variations of κ and τ on the density distribution while keeping the other parameters fixed.

Note that the three-parametric $p.d.f$ of ALD given by:

$$f_{\theta,\kappa,\tau}(x) = \begin{cases} \frac{\sqrt{2}}{\tau}\frac{\kappa}{1+\kappa^2}\ exp(-\frac{\sqrt{2}\kappa}{\tau}|x-\theta|) & x \geq 0 \\ \frac{\sqrt{2}}{\tau}\frac{\kappa}{1+\kappa^2}\ exp(-\frac{\sqrt{2}}{\tau\kappa}|x-\theta|) & x < 0 \end{cases} \tag{2.3}$$

takes into account the displacement characteristics of the distribution. Following the work of Kotz et al. (2001) and Zhao et al. (2014), the mean and variance of X are given as follows:

$$E(X) = \mu + \theta \tag{2.4}$$
$$Var(X) = \mu^2 + \tau^2 \tag{2.5}$$

Since the location parameter θ has no impact on the shape of density distribution, we can assume that $\theta = 0$ without loss of generality.

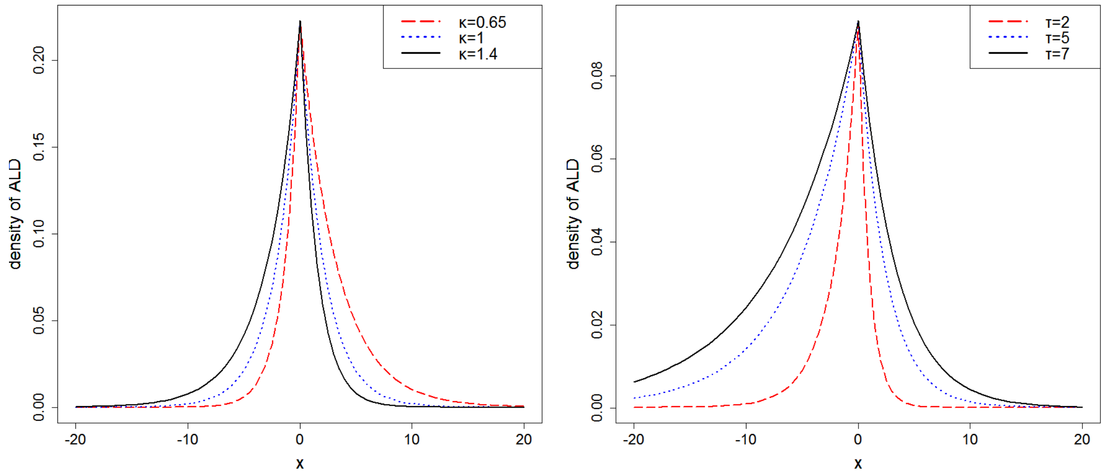

Figure 2.1: Asymmetric Laplace densities. Left: $\theta = 0$, $\tau = 3$, $\kappa = 0.65, 1, 1.4$; Right: $\theta = 0$, $\kappa = 1.5$, $\tau = 2, 5, 7$

Univariate VaR, CVaR and Skewness. Following the mathematical definition of VaR: $Prob(X \leq -VaR_{1-\alpha}|\Omega_t) = \alpha$ and CVaR: $CVaR_{1-\alpha} = -E[E|X \leq -VaR_{1-\alpha}]$, employing density function (2.2), the analytic expression of VaR and CVaR of a single asset under ALD can be obtained as follows:[6]

$$VaR = -\frac{\tau\kappa}{\sqrt{2}} ln \frac{\alpha(1+\kappa^2)}{\kappa^2} \tag{2.6}$$

$$CVaR = -\frac{\tau\kappa}{\sqrt{2}} ln \frac{\alpha(1+\kappa^2)}{\kappa^2} + \frac{\tau\kappa}{\sqrt{2}} \tag{2.7}$$

Here, the VaR and CVaR are positive values which indicate the inequality $0 < \alpha < \kappa^2/(1+\kappa^2)$ holds, considering the fact that κ estimate is close to 1 (see Appendix C) and α is generally selected as 0.01, 0.05 and 0.1.

The skewness of X is given by:[7]

$$s = \frac{2\mu^3 + 3\mu\tau^2}{\sigma^3} = \frac{2\mu^3 + 3\mu\tau^2}{(\mu^2 + \tau^3)^{3/2}} \tag{2.8}$$

[6]Derivation of VaR and CVaR see Appendix A.1 and A.2, respectively.
[7]See Appendix A.3 for derivation.

2.2.3 VaR, CVaR and Skewness of asset portfolio under multivariate ALD

Multivariate ALD. Since empirical data show that the return distribution of assets in the portfolio exhibits the characteristics of multivariate non-Normality, it is necessary to introduce the theory of multivariate ALD before constructing the portfolio. The multivariate ALD is employed to capture the multivariate non-Normal features, which as a combination, can be regarded as an extension of both the univariate ALD and multiple symmetric Laplace distribution. The univariate ALD has been successfully applied in financial markets because of the practicability of the explicit analytical density function and the finite second moments of random sums of *i.i.d.* random vectors.[8] As a result, the multivariate ALD, on the other hand, is believed to be effective, especially in describing skewed and leptokurtic multivariate datasets (see, e.g., Kotz et al., 2001; Lindsey and Lindsey, 2006). The multivariate ALD is defined via the form of characterization.

Definition 2.10. *A random vector Y in $\mathbb{R}^d$ is said to follow a multivariate asymmetric Laplace distribution if its characteristic function is in the form:*

$$\psi(t) = \frac{1}{1 + \frac{1}{2}t'\Sigma t - i\mu' t} \tag{2.9}$$

where $\mu \in \mathbb{R}^d$ and Σ is a $d \times d$ non-negative definite symmetric matrix.[9]

If quoting term $AL_d(\mu, \Sigma)$ as the distribution of the $d-$dimensional random vector Y and expressing as $Y \sim AL_d(\mu, \Sigma)$, then the relationship between mean vector $E(Y)$, covariance matrix $Cov(Y)$ and the parameters μ and Σ can be concluded as follows:[10]

$$E(Y) = \mu \tag{2.10}$$

$$Cov(Y) = V = \Sigma + \mu\mu' \tag{2.11}$$

[8]Random sums can be written as $X^1 + \cdots + X^{v_p}$, where v_p has a geometric distribution with the mean $1/p$. More details see Kotz et al. (2001).

[9]Note that Σ in multivariate ALD is no longer denotes the variance-covariance matrix of random vector Y unless mean vector μ is zero.

[10]Let $X \sim N_d(0, \Sigma)$, Z be a standard exponentially random variable, independently of X. Then, the following formula holds: $Y = \mu Z + \sqrt{Z}X$, correspondingly, $E(Y) = \mu$ and $Var(Y) = \Sigma + \mu\mu'$ exist. Further discussions and proofs see Kozubowski and Podgorski (2000).

Linear transformations are generally involved in portfolio optimization problems. Proposition 2.11 indicates that if $Y \sim AL_d(\mu, \Sigma)$, then all linear combinations of the components of Y are jointly AL distributed. In other words, if asset returns in the portfolio follow a multivariate AL distribution, their linear combination is subject to a one-dimensional ALD.

Proposition 2.11. *Let* $Y = (Y_1, Y_2 \cdots \cdots Y_n)' \sim AL_d(\mu, \Sigma)$ *and* W *is a real matrix whose size is* $n * 1$*, then the random variable* $W'Y \sim AL(\mu_W, \Sigma_W^2)$ *with* $\mu_W = W'\mu$*,* $\Sigma_W^2 = W'\Sigma W$*.*

Proof. The assertion follows the relation:

$$\psi_{W'Y}(t) = Ee^{i(W'Y)'t} = Ee^{iY'Wt} = \psi_Y(Wt)$$

Also, the characteristic function of $W'Y$ satisfies that:

$$\psi_Y(Wt) = \frac{1}{1 + \frac{1}{2}(Wt)'\Sigma(Wt) - i\mu'Wt} = \frac{1}{1 + \frac{1}{2}t'W'\Sigma Wt - i(W'\mu)'t}, \quad t \in R$$

Hence, we have $W'Y \sim AL_n(W'\mu, W'\Sigma W)$. □

Portfolio VaR, CVaR and Skewness. Suppose a portfolio consists of n risky assets with $W = (w_1, w_2, w_3 \cdots w_j)^T$, where w_j represents the weight of the jth asset with $j = 1, 2 \cdots n$. Setting $W^T\mathbf{1} = 1, \mathbf{1} = (1, 1, 1 \cdots 1)^T$ and $Y = (Y_1, Y_2, Y_3 \cdots Y_j)^T$, where Y_j is the rate of return of the jth asset. Then we can denote the return of the portfolio as $R(W, Y) = W^TY$ and the loss of the portfolio as $L(W, Y) = -W^TY$. If there is a feasibility condition such that $W = \{W \in R^n | W^T\mathbf{1} = 1\}$, then the mean and variance of the return of the portfolio can be shown as:

$$\mu_W = E[R(W, Y)] = W^T\mu \tag{2.12}$$

$$V_W = \sigma^2(W) = W^TVW \tag{2.13}$$

According to proposition 2.11, we have random variable $Y \sim AL_d(\mu, \Sigma)$ and the random vector $W'Y = AL_d(\mu_W, \Sigma_W^2)$, thus we can conclude that the return of the portfolio follows a one-dimensional AL distribution and therefore, an alternative expression for the variance of the portfolio return is obtained as:

$$V_W = \Sigma_W^2 + \mu_W^2 \tag{2.14}$$

Following proposition 2.11 and equation (2.11), we can obtain the analytic form of Σ_W^2 by transforming equation (2.14):

$$\Sigma_W^2 = W^T \Sigma W = W^T V W - W^T \mu (W^T \mu)^T \tag{2.15}$$

Substituting $\kappa = \sqrt{2}\tau/(\mu + \sqrt{\mu^2 + 2\tau^2})$ into CVaR formula of single asset, the portfolio CVaR can be formulated as:[11]

$$CVaR_W = (1 - ln\alpha) g(\mu_W, \Sigma_W) - g(\mu_W, \Sigma_W) ln(2 + \frac{\mu_W}{g(\mu_W, \Sigma_W)}) \tag{2.16}$$

where $CVaR_W$ denotes the CVaR of the portfolio and function $g(\mu_W, \Sigma_W)$ is defined as $g(\mu_W, \Sigma_W) = \Sigma_W^2/(\mu_W + \sqrt{\mu_W^2 + 2\Sigma_W^2})$ with parameter μ_W and Σ_W are the replacement of μ and τ, respectively.

Similarly, portfolio VaR can be shaped as:[12]

$$VaR_W = -ln(\alpha) g(\mu_W, \Sigma_W) - g(\mu_W, \Sigma_W) ln(2 + \frac{\mu_W}{g(\mu_W, \Sigma_W)}) \tag{2.17}$$

where VaR_W is the VaR of the portfolio, $g(\mu_W, \Sigma_W) = \Sigma_W^2/(\mu_W + \sqrt{\mu_W^2 + 2\Sigma_W^2})$. Taking S_W as the skewness of the portfolio and replace μ and τ in equation (2.8) by μ_W and Σ_W respectively, the portfolio skewness can be written as:

$$S_W = \frac{2\mu_W^3 + 3\mu_W \Sigma_W^2}{\sigma_W^3} = \frac{2\mu_W^3 + 3\mu_W \Sigma_W^2}{(\mu_W^2 + \Sigma_W^2)^{3/2}} \tag{2.18}$$

2.3 Portfolio optimization model

The objective of portfolio optimization is to reach an equilibrium status that can meet multiple objective functions at the same time, i.e. maximizing the expected

[11] See Appendix B.1 for proof.
[12] See Appendix B.2 for proof.

return of portfolio meanwhile minimizing the risk. In addition, the skewness of portfolio return is important in portfolio selection problems especially for investors who are completely risk averse.[13] Hence, for the Mean-CVaR-Skewness optimization model, the main aim is to minimize the portfolio CVaR and maximize the skewness simultaneously, at a given expected portfolio return. Using equation (2.16) and (2.18), the Mean-CVaR-Skewness model can be written as an optimization problem:

$$\begin{cases} \underset{W \in R^n}{Min} \quad CVaR_W = (1 - ln\alpha) g(\mu_W, \Sigma_W) - g(\mu_W, \Sigma_W) ln(2 + \frac{\mu_W}{g(\mu_W, \Sigma_W)}) \\ \underset{W \in R^n}{Max} \quad S_W = \frac{2\mu_W^3 + 3\mu_W \Sigma_W^2}{\sigma_W^3} = \frac{2\mu_W^3 + 3\mu_W \Sigma_W^2}{(\mu_W^2 + \Sigma_W^2)^{3/2}} \\ s.t. \quad \mu_W = W^T \mu = r, W^T \mathbf{1} = 1 \end{cases} \tag{2.19}$$

where r is the expected return of the portfolio, $g(\mu_W, \Sigma_W) = \Sigma_W^2 / (\mu_W + \sqrt{\mu_W^2 + 2\Sigma_W^2})$.

Obviously, this is a multi-target optimization problem with the first and second objective being nonlinear functions of W. Traditional mathematical approaches (i.e. simplex method) cannot perform well in solving the nonlinear programming problems. In addition, the process to solve the nonlinear programming problems tend to be difficult, existing techniques, including linear approximation and nonlinear goal programming, are overly complex and time consuming. A general hint behind optimizing multiple objectives is to reduce the multi-target functions to a single-target function, i.e. the Mean-variance-skewness optimization problem (Konno and Suzuki, 1995), which is suitable to apply here.

Monotonic analysis. Given the expected return of the portfolio r, replacing μ_W by r in equation (2.16), the CVaR of portfolio can be written as:

$$CVaR_W = (1 - ln\alpha) g(r, \Sigma_W) - g(r, \Sigma_W) ln(2 + \frac{r}{g(r, \Sigma_W)}) \tag{2.20}$$

[13]Empirical results from Lai (1991) indicate that the construction of an optimal portfolio is highly replying on the incorporation of skewness.

It is clear that $CVaR_W$ is a function of Σ_W and r. To simplify the analysis, equation (2.20) can be decomposed into two parts:

$$CVaR_W = f_1(\Sigma_W) + f_2(\Sigma_W)$$

where

$$f_1(\Sigma_W) = (1 - ln\alpha)g(r, \Sigma_W) \quad (2.21)$$

$$f_2(\Sigma_W) = -g(r, \Sigma_W)ln(2 + \frac{r}{g(r, \Sigma_W)}) \quad (2.22)$$

(1) Monotonicity of $f_1(\Sigma_W)$. It is obvious that $f_1(\Sigma_W)$ is a function of Σ_W, then we have:

$$\frac{dg(r, \Sigma_W)}{d\Sigma_W} = \frac{\Sigma_W}{r + \sqrt{r^2 + 2\Sigma_W^2}}(2 - \frac{2\Sigma_W^2}{(r\sqrt{r^2 + 2\Sigma_W^2})(r + \sqrt{r^2 + 2\Sigma_W^2})}) > 0 \quad (2.23)$$

The inequality $0 < \alpha < 1$ implies that $(1 - ln\alpha) > 0$, then we can get:

$$\frac{df_1(\Sigma_W)}{d\Sigma_W} = (1 - ln\alpha)\frac{dg(r, \Sigma_W)}{d\Sigma_W} > 0 \quad (2.24)$$

The equation (2.24) indicates that $f_1(\Sigma_W)$ is a monotonically increasing function with respect to Σ_W.

(2) Monotonicity of $f_2(\Sigma_W)$. It has been shown that Σ_W is positively correlated with $g(r, \Sigma_W)$, implying that $g(r, \Sigma_W)$ will become larger when Σ_W increases. Since r is positive, $2+\frac{r}{g(r, \Sigma_W)}$ tends to decrease when Σ_W increase, thus $-ln(2+\frac{r}{g(r, \Sigma_W)})$ increase. Applying the monotonicity rule of composite function, we conclude that $f_2(\Sigma_W)$ tends to rise as Σ_W increase. Therefore, $f_2(\Sigma_W)$ is a monotonically increasing function with respect to Σ_W. Overall, $CVaR_W = f_1(\Sigma_W) + f_2(\Sigma_W)$ is a monotonically increasing function with respect to Σ_W.

Taking VaR as a risk measure, the portfolio selection problem under Mean-VaR-Skewness model becomes:

$$\begin{cases} \underset{W\in R^n}{Min} & VaR_W = -ln(\alpha)g(\mu_W,\Sigma_W) - g(\mu_W,\Sigma_W)ln(2+\dfrac{\mu_W}{g(\mu_W,\Sigma_W)}) \\ \underset{W\in R^n}{Max} & S_W = \dfrac{2\mu_W^3+3\mu_W\Sigma_W^2}{\sigma_W^3} = \dfrac{2\mu_W^3+3\mu_W\Sigma_W^2}{(\mu_W^2+\Sigma_W^2)^{3/2}} \\ s.t. & \mu_W = W^T\mu = r, W^T\mathbf{1} = 1 \end{cases} \tag{2.25}$$

where $g(\mu_W,\Sigma_W) = \Sigma_W^2/(\mu_W+\sqrt{\mu_W^2+2\Sigma_W^2})$. It is not difficult to show that VaR_W is a monotonically increasing function with respect to Σ_W.

Replacing μ_W by r, the skewness of the portfolio becomes:

$$S_W = \frac{2\mu_W^3+3\mu_W\Sigma_W^2}{(\mu_W^2+\Sigma_W^2)^{3/2}} = \frac{2r^3+3r\Sigma_W^2}{(r^2+\Sigma_W^2)^{3/2}} \tag{2.26}$$

Since we have:

$$\frac{dS_W(r,\Sigma_W)}{d\Sigma_W} = -\frac{3r\Sigma_W^3}{(r^2+\Sigma_W^2)^{5/2}} < 0 \tag{2.27}$$

then we can conclude that S_W is a monotonically decreasing function of Σ_W. That is, if Σ_W^2 is minimized, VaR_W and $CVaR_W$ will be minimized and S_W will be maximized. This indicates that the original multi-objective optimization problems can be reduced to a one-objective problem by minimizing Σ_W^2. As a result, an equivalent reduced optimization model is introduced.

Proposition 2.12. *The Mean-CVaR-Skewness model has the same solution as that of Mean-VaR-Skewness model, both of which can be transferred to the following optimization problem:*

$$\begin{cases} \underset{W\in R^n}{Min} & \frac{1}{2}\Sigma_W^2 = \frac{1}{2}W^T\Sigma W = \frac{1}{2}W^T(V-\mu\mu^T)W \\ s.t. & \mu_W = W^T\mu = r, W^T\mathbf{1} = 1 \end{cases} \tag{2.28}$$

2.4 Algorithm of the optimization model

This section shows the computational scheme to solve the simplified optimization model using the Lagrange multiplier method.[14] This simplified form is a quadratic programming (QP) problem which involves minimizing a quadratic function Σ_W^2 subject to few linear constraints.[15] Following notations in section 2.2, the expected returns $\mu_j = E(Y_j)$ for $j = 1, 2, \cdots n$ is set to be in a one-column matrix $\mu = [\mu_1, \mu_2, \cdots \mu_n]^T$. The covariance of returns denoted by $V = Cov(Y_i, Y_j)$ for $i = 1, 2, \cdots n$ and $j = 1, 2, \cdots n$ is the entries of the $n * n$ variance-covariance matrix given by:

$$V = \begin{bmatrix} var(Y_1) & \ldots & cov(Y_1, Y_n) \\ \vdots & \ddots & \vdots \\ cov(Y_n, Y_1) & \ldots & var(Y_n) \end{bmatrix}$$

Then, we are able to show that the covariance matrix V in our datasets is a positive definite matrix and $\Sigma = V - \mu\mu^T$ is also a positive definite matrix.[16] This is consistent with the acknowledgement that covariance matrix of financial asset returns are generally positive-definite. A large number of mathematical techniques are designed to solve the QP problems (i.e. ellipsoid algorithm, interior point method, simplex type method and Lemke-Howson method, etc.). In addition, computer-based methods are studied and developed (see, e.g., Hasan, 2012; Hasan and Hasan, 2014). In this paper, we show that it is possible to reach a closed-form solution by using the Lagrange multiplier approach.

Suppose λ and β are two Lagrange multipliers, then the Lagrange function can

[14] This is different from Zhao et al. (2014) who claim to solve the model using a statistical software. Indeed, the Lagrange multiplier method has been shown to be a powerful tool in solving conventional mean-variance model, and the application here can be seen as a supplement to the existing studies.

[15] Quadratic programming problem is a subfield of nonlinear programming problem, and is defined based on its feature that the objective function is quadratic and constraints are linear. It has been widely applied in the fields of constrained regression and portfolio selection. Other subfield includes linear programming, geometric programming and convex programming, etc.

[16] This implies that $W^T \Sigma W$ is strictly convex on W, hence, this optimization model can also be classified as a quadratic convex programming problem (see Hasan, 2012). Besides, the positive-definite feature of Σ means the matrix is invertible which can be written in the form of Σ^{-1}.

be formulated as:

$$L(W, \lambda, \beta) = \frac{1}{2} W^T \Sigma W - \lambda W^T \mathbf{1} - \beta W^T \mu \tag{2.29}$$

Equating the partial derivative of $L(W, \lambda, \beta)$ with respect to the optimal weights (W) of portfolio to zero, we can obtain the optimal weight as:

$$W = \lambda \Sigma^{-1} \mathbf{1} + \beta \Sigma^{-1} \mu \tag{2.30}$$

Substituting equation (2.30) into the constraints of optimization model (2.28), a system of linear equations can be obtained as follows:

$$r = \mu^T \lambda \Sigma^{-1} \mathbf{1} + \mu^T \beta \Sigma^{-1} \mu \tag{2.31}$$

$$1 = \mathbf{1}^T \lambda \Sigma^{-1} \mathbf{1} + \mathbf{1}^T \beta \Sigma^{-1} \mu \tag{2.32}$$

which are equivalently to:

$$\lambda C + \beta A = 1 \tag{2.33}$$

$$\lambda A + \beta B = r \tag{2.34}$$

where $A = \mathbf{1}^T \Sigma^{-1} \mu = \mu^T \Sigma^{-1} \mathbf{1}$, $B = \mu^T \Sigma^{-1} \mu$ and $C = \mathbf{1} \Sigma^{-1} \mathbf{1}$. Then, we can obtain the solution of λ and β by solving the above binary linear equations:

$$\lambda = \frac{B - rA}{BC - A^2} \tag{2.35}$$

$$\beta = \frac{rC - A}{BC - A^2} \tag{2.36}$$

As a result, the analytical expression of portfolio weight vector can be obtained by substituting the solution (2.35) and (2.36) into (2.30):

$$\begin{aligned} W = \lambda \mathbf{1} \Sigma^{-1} + \beta \mu \Sigma^{-1} =& \frac{(B - rA) \mathbf{1} \Sigma^{-1}}{BC - A^2} + \frac{(rC - A) \mu \Sigma^{-1}}{BC - A^2} \\ =& \frac{B(\Sigma^{-1} \mathbf{1}) - A(\Sigma^{-1} \mu)}{BC - A^2} + \frac{C(\Sigma^{-1} \mu) - A(\Sigma^{-1} \mathbf{1})}{BC - A^2} r \end{aligned} \tag{2.37}$$

By setting a constant required return r, the optimal asset weights in the portfolio can be obtained using (2.37), and then the $CVaR_W$, VaR_W and S_W can be calculated following equations (2.16), (2.17) and (2.18), respectively.

2.5 Empirical analysis

2.5.1 Data and portfolio construction

The preliminary work for a portfolio selection problem is to choose the potential securities. There are two main possible approaches in the portfolio selection process. The first approach starts from the experience and observation of investors and ends with relevant beliefs about the future performance of potential securities, while the second approach is to initially consider the likely performance of potential securities and then the choice of portfolio (Markowitz, 1952). Our criterion for choosing assets is based on the second approach, that is, the top 20 constituents (with the exception of GOOG and PM) of Standard & Poor 500 by index weight in the year 2012 are considered.[17] The left 18 stocks are expressed with ticker as: AAPL, XOM, CVX, MSFT, IBM, GE, JNJ, PFE, PG, INTC, CSCO, KO, WMT, ORCL, ABT, MRK, T and COP. These stocks have large overall trading volume hence are highly representative in the stock market. In addition, we employ three datasets with cognitive standard as "pre-crisis" ($D1$), "crisis" ($D2$) and "all" ($D3$) to examine the consistency in model performance.

It is debatable to have a consistent criterion to define the time periods of global financial crisis. Frankel and Saravelos (2012) define financial crisis broadly from the view of both financial and real symptoms. They have considered the crisis period continued to 2009 rather than the end of 2008 because many asset prices and real output indicators still decreased after 2008. The starting point was marked from September 15, 2008 as the day of Lehman brother bankruptcy. Demirguc-Kunt et al. (2013) analyzed whether better capitalized banks experienced higher stock returns

[17] The reason to drop the two indexes is because the historical prices of them are dissatisfied our requirement for time periods. GOOG was first listed in NASDAQ on August 18, 2004 and PM started from March 16, 2008.

and they have defined the crisis time from Q3 2007 to Q1 2009. In this paper, we study a two-year cycle from September 2007 to September 2009 with dataset named as $D2$, considering both the time of US recession starting from December 2007 that claimed by NBER and progressively recovery of economy started from March of 2009 that indicated by MSCI world equity index. $D1$ was dated from January 2002 to the end of 2006 and there is an 8 months time gap between $D1$ and $D2$ due to the consideration of emerging of pre-recession (See Demirguc-Kunt et al., 2013). $D3$ was collected with time span from January 2002 to December 2012, covering both $D1$, $D2$ and three years lag. We obtain the raw data from Yahoo! Finance using historical daily adjusted close prices, which have been adjusted for dividends, stock splits and new stock offerings thus can be treated as an accurate technical signal. Daily log returns are calculated based on the natural logarithm of the price ratio for 2 consecutive days, which is defined by $R_t = [ln(p_t) - ln(p_{t-1})] * 100$ with p_t as the closing price on day t, and that would yield a total of 1259, 524 and 2769 observations in $D1$, $D2$ and $D3$, respectively.

The descriptive statistics of the portfolio stocks for the three datasets are summarized in Table 2.1. It is clear to see that stock returns during financial crisis time ($D2$) have exhibited considerable variance comparing to the mean than that of $D1$ and $D3$, and the mean value of returns in $D2$ are more likely to be negative. In addition, there is evidence that the stock returns in our portfolio exhibite asymmetric features, either positive or negative.

Since the purpose of incorporating higher moments is to illustrate the non-Normality feature of asset returns, then our empirical work starts from testing the Normality of stock returns in our portfolio. To check Normality, we employ the Jarque-Bera (J-B) test, which is regarded as an optimal tool with excellent asymptotic power under various alternative specifications of probability distribution (Jarque and Bera, 1987).[18] Results show that J-B statistics are significantly larger than 5.99 with

[18] The Jarque-Bera statistic has a chi-squared (χ^2) distribution with two degrees of freedom. Inverting the χ^2 at the 5% and 1% significance level lead to a critical value of 5.99 and 9.21, respectively. Hence, for instance, if a Jarque-Bera statistic is larger than 5.99, the null hypotheses of Normality will be rejected at 5% significance level, which means the distribution is not normally distributed.

Table 2.1: Descriptive Statistics of assets in the portfolio for the three datasets

Number	Ticker	Stock	Datasets	Mean	Std.dev.	Skewness	Kurtosis	Jarque-Bera	Prob.
1	AAPL	Apple Inc.	D1	0.1624	2.5726	0.0885	6.3953	606.3732	0
			D2	0.0557	3.1001	-0.4784	7.1436	394.8604	0
			D3	0.1405	2.4514	-0.1379	7.5558	2403.3690	0
2	XOM	Exxon Mobil Corp.	D1	0.0622	1.4048	-0.2667	6.9668	840.3928	0
			D2	-0.0348	2.5533	0.2168	11.6369	1632.7890	0
			D3	0.0375	1.6500	-0.0217	14.9260	16410.0200	0
3	CVX	Chevron Corp.	D1	0.0530	1.3429	-0.3884	4.7707	196.1374	0
			D2	-0.0295	2.7575	0.2829	11.4668	1572.1630	0
			D3	0.0451	1.7193	0.0178	15.8632	19090.4000	0
4	MSFT	Microsoft Corp.	D1	0.0034	1.6929	-0.1677	9.1682	2001.7750	0
			D2	-0.0135	2.7171	0.3515	8.0438	566.2219	0
			D3	0.0022	1.8316	0.1597	10.8550	7130.5770	0
5	IBM	International Business Machines Corp.	D1	-0.0138	1.5662	0.0700	11.0497	3400.1890	0
			D2	0.0116	2.0431	0.2561	5.6950	164.3073	0
			D3	0.0220	1.5493	0.0860	9.1629	4385.5490	0
6	GE	General Electric Co.	D1	0.0046	1.5518	0.2169	8.3320	1501.2600	0
			D2	-0.1433	3.4332	0.1076	6.7236	303.7403	0
			D3	-0.0102	2.0509	0.0452	11.8384	9013.7430	0
7	JNJ	Johnson & Johnson	D1	0.0162	1.2701	-1.7189	31.6439	43660.6700	0
			D2	0.0086	1.4995	0.6980	14.0837	2724.7370	0
			D3	0.0165	1.1856	-0.6761	26.2610	62637.4600	0
8	PFE	Pfizer Inc.	D1	-0.0246	1.6607	-0.6402	10.1402	2760.4270	0
			D2	-0.0549	2.1557	-0.0447	6.7309	304.0925	0
			D3	-0.0022	1.6348	-0.3448	9.3319	4680.5770	0
9	PG	The Procter & Gamble Co.	D1	0.0469	1.0391	-0.1104	6.8954	798.5580	0
			D2	-0.0130	1.7112	-0.1569	7.3958	424.0443	0
			D3	0.0295	1.1516	-0.1658	9.7009	5193.2380	0
10	INTC	Intel Corp.	D1	-0.0312	2.4569	-0.7503	11.5816	3981.3810	0
			D2	-0.0415	2.8598	-0.0751	5.2486	110.8862	0
			D3	-0.0071	2.2709	-0.4802	10.0486	5838.6070	0
11	CSCO	Cisco Systems Inc.	D1	0.0327	2.4270	0.4190	11.2905	3642.4420	0
			D2	-0.0581	2.6993	0.0294	5.6515	153.5687	0
			D3	0.0043	2.3031	-0.0119	11.7397	8812.7160	0
12	KO	The Coca-Cola Co.	D1	0.0108	1.1821	-0.8805	12.3390	4737.9620	0
			D2	0.0126	1.8467	0.7211	10.7715	1364.0680	0
			D3	0.0259	1.2616	0.0453	14.1497	14343.8600	0
13	WMT	Wal-Mart Stores Inc.	D1	-0.0136	1.3573	0.2176	5.4132	315.4331	0
			D2	0.0298	1.8099	0.2800	7.8884	528.5792	0
			D3	0.0127	1.3525	0.1704	7.9878	2883.7040	0
14	ORCL	Oracle Corp.	D1	0.0172	2.4615	0.0897	6.3385	586.3673	0
			D2	0.0062	2.6522	0.2041	5.5245	142.7830	0
			D3	0.0331	2.2567	0.0306	6.9269	1779.6210	0
15	ABT	Abbott Laboratories	D1	0.0097	1.6336	-0.6601	18.6821	12992.3700	0
			D2	0.0017	1.7614	-0.2477	7.0839	369.4943	0
			D3	0.0218	1.4744	-0.4873	16.0235	19678.5300	0
16	MRK	Merck & Co. Inc.	D1	-0.0049	1.8491	-3.9804	70.2453	240537.2000	0
			D2	-0.0688	2.6871	-0.4724	8.3863	652.9196	0
			D3	0.0048	1.8750	-2.0067	37.5149	139301.8000	0
17	T	AT & T,Inc.	D1	0.0118	1.7368	-0.0764	7.2160	933.6592	0
			D2	-0.0546	2.3425	0.6772	8.4649	692.1116	0
			D3	0.0148	1.6748	0.2618	10.0753	5807.2190	0
18	COP	ConocoPhillips	D1	0.0792	1.5588	-0.2399	3.7806	44.0428	0
			D2	-0.1016	3.0821	-0.2793	7.0485	364.6614	0
			D3	0.0456	1.9251	-0.4112	10.1907	6043.5690	0

corresponding p-values equal to zero for the 18 assets in all three periods. Hence, there are sufficient evidences to against the null hypotheses that the asset returns in the portfolio are normally distributed. This result provides good evidence for using heavy-tailed distribution (i.e. ALD) to fit the datasets.[19]

2.5.2 Portfolio VaR model analysis

Single asset Normal VaR investigation. Methodology for VaR calculation can be summarized into three categories: Parametric-Normal approach, Non-parametric approach and Simulation approach. Parametric-Normal method assumes that asset returns are subject to a specific Normal distribution: the VaR at confidence level $(1-\alpha)$ is defined as the $(1-\alpha)$ quantile of the Normal distribution. Conversely, the non-parametric method does not take into account any parameters and depends only on the historical return series, therefore it is also called historical method or parameter-free method. As a result, the VaR estimates under non-parametric approach can be calculated immediately from real dataset at the corresponding confidence levels.

Figure 2.2 shows the differences between parametric-Normal VaR estimates and non-parametric VaR estimates of each single stock in our portfolio among three datasets at different confidence levels. Let $\tilde{d}$ denotes the differences between the parametric VaR estimates and the non-parametric VaR estimates. Then, at 90% conference level, we have $\tilde{d} > 0$ for all stocks in the three datasets with $\tilde{d}$ ranging from 0.05 to 0.74. At 95% conference level, $\tilde{d}$ fluctuates between -0.12 and 0.69 which implies that some of the parametric-Normal VaR estimates are lower than the historical VaR estimates. If the confidence level reaches 99%, almost all parametric-Normal VaR estimates are less than their counterparts. Therefore, we conclude that the parametric-Normal approach could underestimate the true VaRs at high confidence level for the heavy-tailed return series. The underestimation of VaR becomes much more obvious when moving further towards to the tail and this larger deviation may result in larger errors of exposure to the market risks. As a result, the potential losses

[19]Fitting $ALD(\kappa, \tau)$ via the maximum likelihood estimation (MLE) approach, estimates of the three datasets are shown in Appendix C.

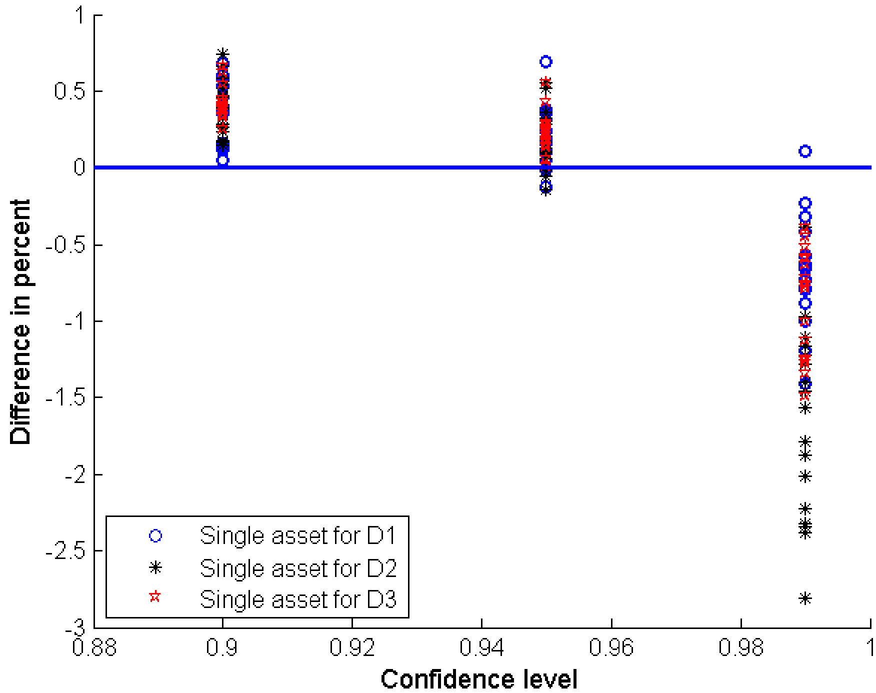

Figure 2.2: Differences between Parametric-Normal VaR estimates and Non-parametric VaR estimates for single asset in the portfolio of the three datasets at different conference levels

can be magnified if a risk manager is operating a large amount of asset positions.

Portfolio VaR comparisons. To test the viability of Mean-VaR-Skewness ALD optimization model in equation (2.25), we compare it with the portfolio optimization VaR model under Normal distribution and the historical portfolio VaR approach, the latter being regarded as a benchmark. To observe the consistency of model performance, the empirical analysis is implemented under three different time periods for a sequence of given confidence levels.

The algorithm of historical portfolio VaR approach firstly computes the equally weighted average return of assets in the portfolio, and then repeats over to the whole time periods thus eventually obtaining n sample portfolio returns. The portfolio VaR estimate is the threshold of this return distribution at a specific confidence level. For parametric-Normal portfolio VaR, we adopt the method proposed by Rockafellar and

Table 2.2: Comparison of portfolio VaR estimates under Historical, Normal and ALD approach for three datasets at different confidence levels

Confidence level	Dataset	Portfolio VaR		
		Historical	Normal ($r = 10\%$)	ALD ($r = 10\%$)
90%	D1	1.144	1.0918	1.0673
	D2	2.1187	1.4174	1.3950
	D3	1.3024	1.3665	1.3438
92.5%	D1	1.4256	1.1663	1.2667
	D2	2.4005	1.5122	1.6531
	D3	1.5617	1.4581	1.5927
95%	D1	1.6641	1.2407	1.5478
	D2	2.8763	1.6070	2.0168
	D3	1.8883	1.5498	1.9435
97.5%	D1	2.1636	1.3152	2.0282
	D2	4.0219	1.7018	2.6386
	D3	2.4099	1.6414	2.5432
99%	D1	2.7042	1.3599	2.6634
	D2	5.1383	1.7587	3.4605
	D3	3.5603	1.6964	3.3360

Note: The portfolio VaR optimization model under the two innovations are implemented at a given expected portfolio return $r = 10\%$.

Uryasev (2000) whose expression is:

$$VaR_{1-\alpha}(x) = c_1(1-\alpha) * \sigma(\alpha) - r \tag{2.38}$$

where $c_1(1-\alpha) = \sqrt{2}\, erf^{-1}(2*(1-\alpha)-1) > 0$, erf^{-1} represents the inverse of error function which is defined as $erf(z) = 2/\sqrt{2\pi}\int_0^z e^{-t^2}dt$ and $\sigma(x)$ is the minimized standard deviation of the portfolio for a given expected portfolio return r.

Table 2.2 compares the portfolio VaR estimates of various approaches for the three datasets over a range of confidence levels by setting the expected portfolio return to be 10%. Assuming there is a fund manager who has made an investment with \$100 million position in this portfolio, under the parametric-Normal approach, the portfolio VaR estimate is \$1.4174 million in the period of crisis ($D2$) at 90% confidence level, significantly lower than the maximum loss of the portfolio

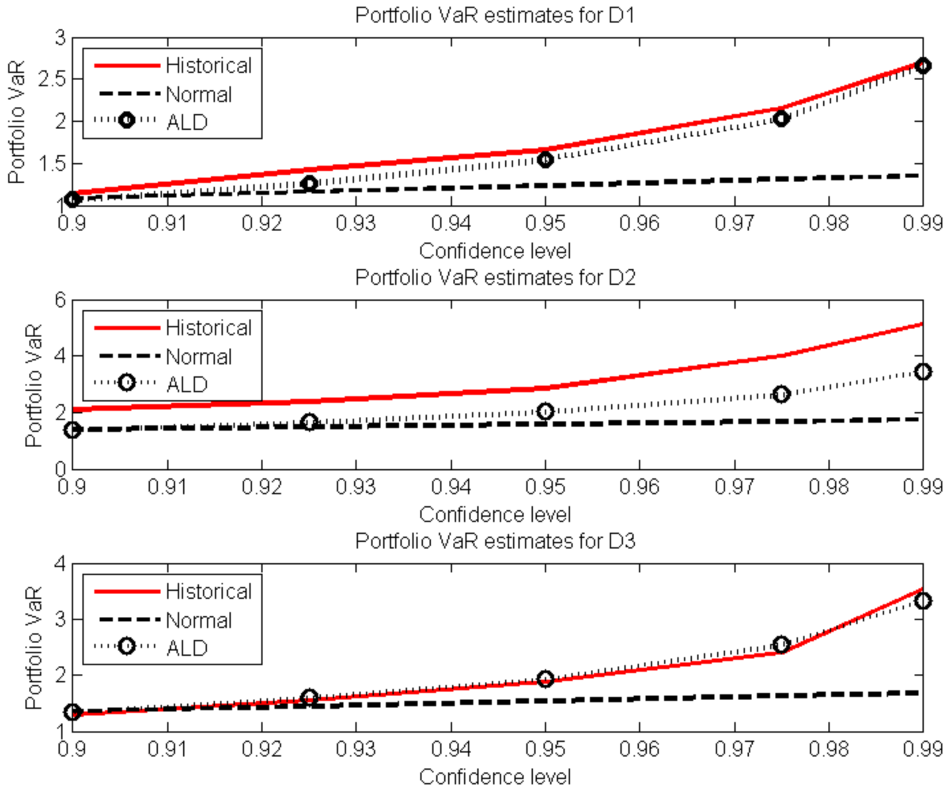

Figure 2.3: Portfolio VaR estimates of Historical, Normal and ALD approach for three datasets over a range of confidence levels

that calculated using historical distribution approach (\$2.1187 million).[20] The discrepancy of portfolio VaR between historical approach and parametric-Normal approach is rising from \$0.7013 million at 90% percentile to \$3.3796 million when the percentile reaches at 99%. Likewise, the portfolio loss of these two methods for the other two datasets ($D1$ and $D3$) have both shown an ascending tendency from \$0.0522 million and \$-0.0641 million at 90% percentile to \$1.3443 million and \$1.8639 million at 99% percentile, respectively. This indicates that the parametric-Normal portfolio VaR approach underestimates the true portfolio losses especially when focusing on the extreme tail events. A consistent finding is that portfolio VaR is obviously higher in the periods of crisis, no matter which methods we use.

The portfolio VaR estimates under ALD, along with estimates of the other two

[20]This is saying that the fund manager has 90% confidence to believe that the losses will not exceed \$1.4174 million if using parametric-Normal approach and \$2.1187 million using historical distribution approach.

approaches, are shown in Figure 2.3. We can conclude that in both the three datasets, portfolio VaR estimates obtained under ALD assumption are much closer to historical estimates than that calculated under Normal assumption. To further examine the accuracy of the two methods, we employ root-mean-square error (RMSE), mean absolute error (MAE) and sum of squared relative error (SSRE) to present model evaluation statistics.[21] The functional forms are defined as:

$$RMSE_{VaR} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(\widehat{VaR}_n(1-\alpha) - VaR_n(1-\alpha)\right)^2} \tag{2.39}$$

$$MAE_{VaR} = \frac{1}{N}\sum_{i=1}^{N}\left|\widehat{VaR}_n(1-\alpha) - VaR_n(1-\alpha)\right| \tag{2.40}$$

$$SSRE_{VaR} = \sqrt{\sum_{i=1}^{N}\left(\frac{\widehat{VaR}_n(1-\alpha) - VaR_n(1-\alpha)}{VaR_n(1-\alpha)}\right)^2} \tag{2.41}$$

where N represents the number of calculated VaRs, $\widehat{VaR}_n(1-\alpha)$ denotes the theoretical VaR values at (1-α) confidence level and $VaR_n(1-\alpha)$ is the actual VaR values at (1-α) confidence level.[22] According to the computational results in Table 2.3, both RMSE, MAE and SSRE values for the ALD approach are lower than the case where Normal distribution is assumed in the three datasets. This implies a better performance of modelling stock return series using ALD when calculating portfolio VaRs.

To conclude, we find that the portfolio VaR estimates under ALD outperform the portfolio VaR estimates under the assumption of Normal distribution in terms of tail risk modelling accuracy. Although VaR has gained great popularities and often be requested by regulators considering its computational simplicity and ease property

[21] Both RMSE, MAE and SSRE have been widely used as a standard metric to evaluate model performance. Although it is quite controversial about which one is better, there is no consensus on the most appropriate indicator for model errors (Willmott and Matsuura, 2005; Willmott et al., 2009; Yao et al., 2013). Since a single metric only provides a unique prediction of model errors, therefore, multiple metrics are often required for testing model performance (See Chai and Draxler, 2014).

[22] Theoretical VaR value means the minimized VaR value of the optimization model under the corresponding Normal distribution or ALD. For every given confidence level $(1-\alpha)$, the $\widehat{VaR}_n(1-\alpha)$ value can be obtained, here $N = 5$.

Table 2.3: Metrics for portfolio VaR errors under Normal and ALD approach in the three datasets

Dataset	RMSE_N	RMSE_{ALD}	MAE_N	MAE_{ALD}	SSRE_N	SSRE_{ALD}
D1	0.7452	0.1137	0.5855	0.1056	0.7077	0.1611
D2	1.9848	1.1445	1.7117	1.0783	1.0985	0.7265
D3	0.9159	0.1215	0.6277	0.0971	0.6440	0.0964

for calculation, the intrinsic limitations are still obvious. We find the VaR loss of a portfolio that consisting of all risky assets can be larger than the sum of VaR loss of each single asset in the portfolio. In addition, VaR could be a misleading indicator to investors as it does not give any indications for the risk that exceed a specific quantile, hence investors' optimizing behavior may lead to market positions that are subject to extreme losses and eventually resulting in market instability. In the following, we empirically investigate the viability of the Mean-CVaR-Skewness model under ALD that proposed by Zhao et al. (2014) considering three different sets of market data. The analysis is conducted by comparing with classical mean-CVaR optimization model under Normal distribution based on three methods: In-sample, Monte Carlo simulation and Out-of-sample.

2.5.3 Portfolio CVaR model performance: In-sample

We examine the model performance with regard to the CVaR losses for the two innovations. The original Mean-CVaR optimization model was first proposed by Rockafellar and Uryasev (2000) based on the Normality assumption of underlying risk factors. Analytic minimized CVaR expression can be expressed as:

$$CVaR_{1-\alpha}(x) = c_2(1-\alpha) * \sigma(x) - r \tag{2.42}$$

where $c_2(1-\alpha) = (\sqrt{2\pi} * (exp(erf^{-1}(2*(1-\alpha)-1))^2) * \alpha)^{-1}$, erf^{-1} denotes the inverse of error function which is defined as $erf(z) = (2/\sqrt{2\pi}) \int_0^z e^{-t^2} dt$ and $\sigma(x)$ is the minimized standard deviation of the portfolio for a given expected rate of return r. In the scenario of allowing short-selling, the analytic expression of $\sigma(x)$ can be

Table 2.4: In-sample portfolio CVaR estimates under various approaches for the three datasets at different confidence levels

Confidence level	Dataset	Portfolio CVaR		
		Historical	M-C(r=10%)	M-C-S(r=10%)
90%	D1	1.8964	2.1159	1.7605
	D2	3.449	2.7212	2.2921
	D3	2.2274	2.6267	2.2090
92.5%	D1	2.0945	2.6205	1.9599
	D2	3.847	3.3638	2.5501
	D3	2.4915	3.2477	2.4579
95%	D1	2.3631	3.6389	2.2409
	D2	4.4619	4.6604	2.9138
	D3	2.8766	4.5008	2.8087
97.5%	D1	2.8016	6.7171	2.7214
	D2	5.5472	8.5795	3.5356
	D3	3.8484	8.2885	3.4084
99%	D1	3.4318	15.9840	3.3565
	D2	7.0826	20.3782	4.3575
	D3	4.731	19.6916	4.2012

Note: *M-C* denotes the Mean-CVaR optimization model under Normal distribution and *M-C-S* represents Mean-CVaR-Skewness model under ALD. This two models are implemented at a given expected portfolio return $r = 10\%$.

derived in the Mean-Variance theoretical framework (See Yao et al., 2013) as:

$$\sigma(x) = \frac{1}{\sqrt{D}}\sqrt{Cx^2 - 2Ax + B}$$

where $A = \mathbf{1}^T\Sigma^{-1}\mu, B = \mu^T\Sigma^{-1}\mu, C = \mathbf{1}^T\Sigma^{-1}\mathbf{1}$ and $D = B * C - A * A$.

Table 2.4 contains the in-sample portfolio CVaR estimates for historical optimization model, Mean-CVaR model under Normal distribution and Mean-CVaR-Skewness model under ALD. The computed outcomes of CVaR represent the worst average portfolio loss (in percentage) per unit invested that can happen with probability α. Consider a portfolio manager who is managing an investment with a position of \$100 million, following the Mean-CVaR-Skewness ALD model at 10% expected portfolio return, for instance, the manager would expect to incur a loss of \$4.3573

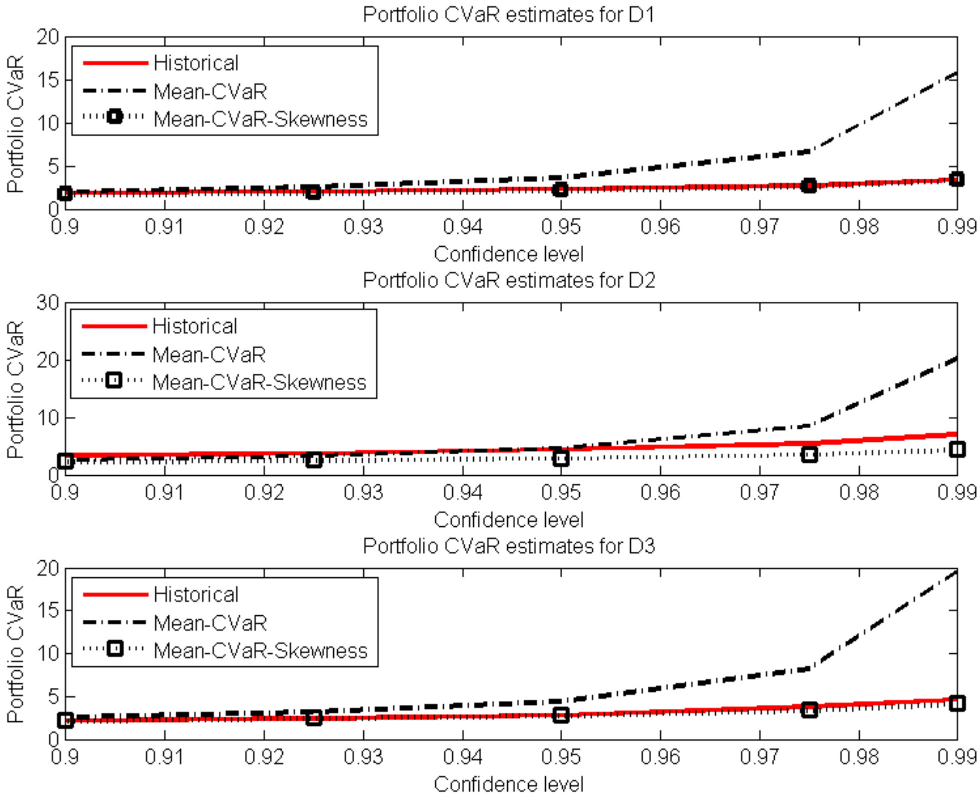

Figure 2.4: Portfolio CVaR estimates of Historical, Mean-CVaR model and Mean-CVaR-Skewness model for three datasets over a range of confidence levels

million in the period of crisis (D2) at 99% confidence level if things do get bad.

We find that the portfolio average losses are relatively high in the time of global financial crisis for all these three methods, no matter what confidence level we choose. In addition, the Mean-CVaR Normal model yields remarkably higher portfolio CVaR values than that of Mean-CVaR-Skewness ALD model at 99% quantile level.[23] To study the performance of these two models, we calculate the portfolio CVaR estimates under the historical approach and the comparisons of these three methods are represented in Figure 2.4. We can see that the Mean-CVaR-Skewness ALD model demonstrates the capability of modelling heavy-tailed data especially at high confidence levels. Compared to the performance of the Mean-CVaR Normal model,

[23]The reason of the distinct larger numbers under the Mean-CVaR model is because the value of coefficient c_2 in equation (2.42) varies significantly when confidence level is changing. For instance, the value of c_2 at quantile 97.5%, 98% and 98.5% is 6.4717, 7.9366 and 10.3799, respectively. However, when confidence level rises to 99%, c_2 becomes 15.2691 and at 99.5%, it even reaching at 29.9425.

Table 2.5: Metrics for portfolio CVaR errors under Mean-CVaR Normal and Mean-CVaR-Skewness ALD in the three datasets

Dataset	RMSE$_{MC}$	RMSE$_{MCS}$	MAE$_{MC}$	MAE$_{MCS}$	SSRE$_{MC}$	SSRE$_{MCS}$
D1	5.9134	0.1128	3.6978	0.1096	3.9622	0.1151
D2	6.1118	1.8379	3.5475	1.7477	1.9710	0.7913
D3	7.0271	0.3100	4.4361	0.2179	3.4313	0.1625

the Mean-CVaR-Skewness ALD model can accurately reflect the portfolio CVaR estimates for the whole range of confidence levels. To further test the estimation accuracy, we define the loss functions RMSE, MAE and SSRE as follows:

$$RMSE_{CVaR} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(\widehat{CVaR}_n(1-\alpha) - CVaR_n(1-\alpha)\right)^2} \tag{2.43}$$

$$MAE_{CVaR} = \frac{1}{N}\sum_{i=1}^{N}\left|\widehat{CVaR}_n(1-\alpha) - CVaR_n(1-\alpha)\right| \tag{2.44}$$

$$SSRE_{CVaR} = \sqrt{\sum_{i=1}^{N}\left(\frac{\widehat{CVaR}_n(1-\alpha) - CVaR_n(1-\alpha)}{CVaR_n(1-\alpha)}\right)^2} \tag{2.45}$$

where N is the number of calculated portfolio CVaRs, $\widehat{CVaR}_n(1-\alpha)$ denotes the theoretical portfolio CVaRs that are calculated based on the two parametric optimization models at (1-α) confidence level, $CVaR_n(1-\alpha)$ is the historical portfolio CVaRs at (1-α) confidence level.

Results from Table 2.5 show that both RMSE$_{MCS}$, MAE$_{MCS}$ and SSRE$_{MCS}$ are significantly lower than the corresponding RMSE$_{MC}$, MAE$_{MC}$ and SSRE$_{MC}$ in the three datasets, which provide evidence that portfolio CVaR estimates of Mean-CVaR-Skewness model under ALD outperforms Mean-CVaR model under Normal distribution in terms of the accuracy of modelling extreme tail risks in the portfolio framework.

2.5.4 Portfolio CVaR model performance: Monte Carlo simulations

This part aims to test the validity of the Mean-CVaR-Skewness model under ALD, in comparison with the Mean-CVaR model under Normal distribution using Monte Carlo simulations. Specifically, we compare the portfolio CVaR values produced by the two models with historical portfolio CVaR values. The data information for the two models stem from the generated random numbers.

For the Mean-CVaR Normal model, we assume that the vector of the underlying risk factors follows a multivariate Normal distribution $N(\mu, \Sigma)$. Then $r_i \sim N(\mu,\Sigma)$, where μ is the mean return of sample dataset ($D3$ daily returns) and Σ is the corresponding variance-covariance matrix. To calculate the minimized CVaR value for a given expected rate of return of the portfolio, we use the agreed algorithm following formula (2.42). The simulated sample size that is generated from the multivariate Normal distributions are 1,000, 5,000 and 10,000.[24]

The simulation algorithm for generating random numbers from multivariate Asymmetric Laplace distributions can be summarized into three steps (see Kotz et al., 2001):

- Generate a standard exponential variate M;
- Generate a multivariate Normal $N_d(0, \Sigma)$ variate N, independently of M;
- Let $y = \mu M + \sqrt{M}N$, then $y \sim AL_d(\mu, \Sigma)$

Table 2.6 shows the simulated experiments of Mean-CVaR Normal model and Mean-CVaR-Skewness ALD model. For a given required rate of return of the portfolio, simulated portfolio CVaRs based on the Mean-CVaR-Skewness ALD model exhibit smaller values than that of Mean-CVaR Normal model at an identical confidence level. This property holds even when the number of sample size varies. In addition, we can see that the simulated CVaR values for both the two models have demonstrated an ascending trend when the expected portfolio return becomes larger,

[24]Analogously to Proposition 2.11, if asset returns in the portfolio are jointly following a Normal distribution, then the linear combination of them is subject to a one-dimensional Normal distribution (See Rachev et al., 2003).

Table 2.6: Comparison of portfolio CVaR estimates for the two optimization models with different expected portfolio returns at different confidence levels: Monte Carlo simulations generated by random numbers.

Sample size	r	Portfolio CVaR					
		Mean-CVaR			Mean-CVaR-Skewness		
		90%	95%	99%	90%	95%	99%
1,000	0.02	1.8902	3.2032	13.8455	1.6352	2.0721	3.0868
	0.04	1.9030	3.2385	14.0634	1.6332	2.0723	3.0918
	0.06	1.9410	3.3164	14.4644	1.6542	2.1014	3.1397
	0.08	2.0021	3.4332	15.0330	1.6959	2.1567	3.2266
	0.1	2.0837	3.5847	15.7507	1.7559	2.2351	3.3479
	0.12	2.1832	3.7662	16.5977	1.8315	2.3334	3.4986
	0.14	2.2978	3.9734	17.5550	1.9205	2.4484	3.6742
5,000	0.02	1.9200	3.2534	14.0614	1.6516	2.0930	3.1177
	0.04	2.0009	3.4036	14.7736	1.6486	2.0917	3.1207
	0.06	2.1685	3.7002	16.1154	1.7301	2.1975	3.2827
	0.08	2.4032	4.1100	17.9446	1.8816	2.3918	3.5763
	0.1	2.6868	4.6022	20.1279	2.0857	2.6526	3.9689
	0.12	3.0049	5.1528	22.5624	2.3275	2.9612	4.4326
	0.14	3.3476	5.7448	25.1752	2.5961	3.3036	4.9466
10,000	0.02	1.9424	3.2913	14.2243	1.6897	2.1411	3.1894
	0.04	2.0595	3.5026	15.1994	1.7757	2.2527	3.3600
	0.06	2.2696	3.8709	16.8496	1.9449	2.4694	3.6871
	0.08	2.5484	4.3550	18.9983	2.1754	2.7636	4.1294
	0.1	2.8752	4.9203	21.4959	2.4484	3.1117	4.6519
	0.12	3.2353	5.5415	24.2345	2.7508	3.4970	5.2296
	0.14	3.6185	6.2018	27.1409	3.0736	3.9081	5.8457

Note: r denotes the expected rate of return of the portfolio. *Mean-CVaR* represents the Mean-CVaR optimization model under Normal distribution and *Mean-CVaR-Skewness* represents Mean-CVaR-Skewness model under ALD.

no matter what confidence level we consider. Then the question arises, how do those simulated portfolio CVaRs in the two models perform compared to historical CVaRs?

Taking 10,000 simulated samples as an example Figure 2.5 shows the corresponding comparison of portfolio CVaRs for the two simulated models with historical results under different expected portfolio returns at different confidence levels. At the 90th

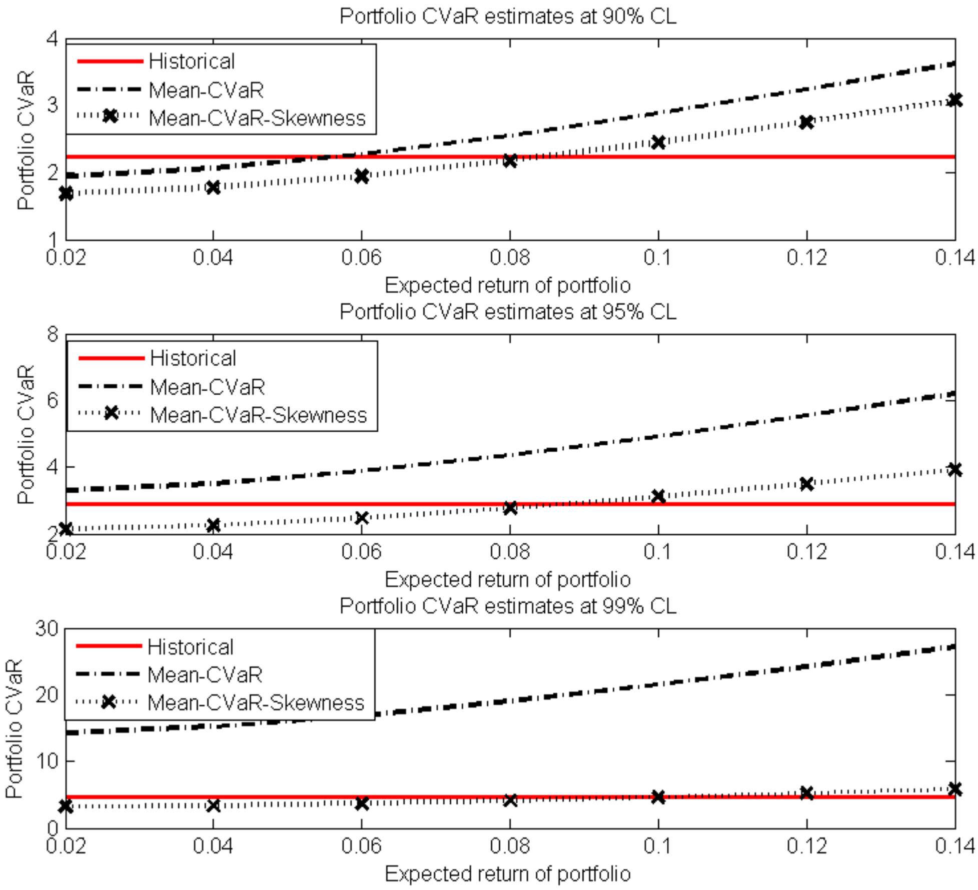

Figure 2.5: Comparison of portfolio CVaRs for the two Monte Carlo simulated models based on 10,000 sample size with historical results under different expected portfolio returns at different confidence levels

percentile, it is difficult to distinguish which line is closer to the historical one, while this uncertainty seems become clear at 95% percentile. When the confidence level increases up to 99%, we can see that the portfolio CVaRs generated by the Mean-CVaR-Skewness ALD model almost overlaps with historical losses. The line of portfolio CVaRs generated by the Mean-CVaR Normal model becomes further from the historical line than that at 90% and 95% percentiles, and the discrepancy is more obvious when expected portfolio return is high. We thus preliminarily conclude from the experiment that the Mean-CVaR-Skewness model under ALD is more capable to fit true portfolio CVaRs at high confidence level than the Mean-CVaR model which is relying on the Normality assumption.

Table 2.7: Metrics for portfolio CVaR errors under Mean-CVaR Normal and Mean-CVaR-Skewness ALD model using three simulated samples at different confidence levels

Sample size	Confidence level	RMSE_{MC}	RMSE_{MCS}	MAE_{MC}	MAE_{MCS}	SSRE_{MC}	SSRE_{MCS}
1,000	90%	0.2328	0.5054	0.2045	0.4951	0.2765	0.6003
	95%	0.6800	0.6869	0.6256	0.6738	0.6254	0.6318
	99%	10.6754	1.4508	10.5990	1.4359	5.9701	0.8114
5,000	90%	0.5670	0.4139	0.4466	0.3726	0.6735	0.4917
	95%	1.6477	0.5563	1.4044	0.4954	1.5155	0.5116
	99%	14.4672	1.1553	13.9491	1.0147	8.0906	0.6461
10,000	90%	0.7156	0.4802	0.5518	0.4163	0.8500	0.5704
	95%	1.9300	0.6110	1.6496	0.5381	1.7751	0.5619
	99%	15.6469	1.0145	15.0037	0.8929	8.7503	0.5674

Accuracy test based on RMSE, MAE and SSRE are defined as follows:

$$RMSE_{CVaR} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(\widehat{CVaR}_n(r) - CVaR_n\right)^2} \tag{2.46}$$

$$MAE_{CVaR} = \frac{1}{N}\sum_{i=1}^{N}\left|\widehat{CVaR}_n(r) - CVaR_n\right| \tag{2.47}$$

$$SSRE_{CVaR} = \sqrt{\sum_{i=1}^{N}\left(\frac{\widehat{CVaR}_n(r) - CVaR_n}{CVaR_n}\right)^2} \tag{2.48}$$

where N is the number of calculated portfolio CVaRs, $\widehat{CVaR}_n(r)$ is the theoretical portfolio CVaR values which are calculated based on the two optimization models at different expected portfolio returns for a specific confidence level, $CVaR_n$ denotes the historical portfolio CVaR values at the corresponding confidence levels.

Results are shown in Table 2.7. It is worth noting that when the number of simulated samples is small, i.e. 1,000, the model errors of the Mean-CVaR are lower than that of the Mean-CVaR-Skewness at 90% and 95%. Whereas, the model errors of the Mean-CVaR-Skewness ALD model are significantly lower than that of Mean-CVaR Normal model at all three confidence levels when the sample size is 5,000 and 10,000.[25] To conclude, the model evaluation statistics show a more accurate portfolio CVaR estimate, which indicate that ALD is more capable of modelling extreme tail events.

2.5.5 Portfolio CVaR model performance: Out of sample forecasting

The evaluation of the out-of-sample performance of the Mean-CVaR-Skewness ALD model is conducted using monthly data in comparison with the Mean-CVaR Normal model. Specifically, the first sample portfolio of the 18 risky assets is constructed using monthly returns from Jan 2002 to Dec 2006 and the second sample portfolio is constructed taking a rolling-window pattern from Feb 2002 to Jan 2007. Thus, 72 sample portfolios are constructed with a time frame up to Dec 2012. All the 72 sample portfolios are iteratively examined using both Mean-CVaR-Skewness ALD model and Mean-CVaR Normal model, which is altogether 144 times. To evaluate the out-of-sample performance of the two optimization models, we first calculate the optimal weights using the 60 in-sample returns for a given required portfolio return (i.e. 10%), then we multiply by the return vector of the first following month to ultimately get the out-of-sample portfolio return.[26] For instance, the first constructed portfolio taking data form Jan 2002 to Dec 2006 as in sample, following one month (Jan 2007) return as out-of-sample. The algorithm of portfolio CVaR under Mean-CVaR Normal model is the same as in section 2.5.3 (see Rockafellar and Uryasev, 2002).

Table 2.8 shows the average monthly out-of-sample portfolio returns and average monthly out-of-sample portfolio CVaR values for each year at different confidence levels. That is, each portfolio is constructed based on the 60 rolling monthly returns and for each year, the average value of 12 portfolios is calculated. We can observe that for a same out-of-sample portfolio return, the Mean-CVaR-Skewness model under ALD performs equivalently or better than the Mean-CVaR model at 95%

[25] It is believed that the reconstructed error distributions under the three metrics are more reliable when the sample size is large enough (see Chai and Draxler, 2014).

[26] Formulized as $\mu_{\text{p\_out}} = w^{'}x$, where w is the in-sample optimal weight, x is the one month out-of-sample return vector (Methodology see Ledoit and Wolf, 2003; Sarykalin et al., 2008).

Table 2.8: Out-of-sample portfolio performance of the two optimization models at different confidence levels

Year	Mean-CVaR (r=10%)				Mean-CVaR-Skewness (r=10%)			
	Return	CVaR(90%)	CVaR(95%)	CVaR(99%)	Return	CVaR(90%)	CVaR(95%)	CVaR(99%)
2007	0.5838	3.9993	7.1494	32.6827	0.5838	4.1560	5.2470	7.7800
2008	-1.0311	4.9989	7.7261	29.8313	-1.0311	8.9005	10.9403	15.6767
2009	-0.5611	5.7333	9.2884	38.1037	-0.5611	9.6566	11.9061	17.1291
2010	-0.3169	5.8797	9.7033	40.6949	-0.3169	7.4644	9.2753	13.4800
2011	2.1302	3.7807	7.8435	40.7740	2.1302	3.9949	5.1193	7.7302
2012	0.3674	5.3080	9.2090	40.8280	0.3674	5.6639	7.1180	10.4944
Average	0.1954	4.9500	8.4866	37.1524	0.1954	6.6394	8.2677	12.0484

Note: Results in the first 6 rows represent the average value of 12 months at that year and the “average” in the last row calculate the average value of the above 6 years. *Mean-CVaR* represents the Mean-CVaR optimization model under Normal distribution and *Mean-CVaR-Skewness* represents Mean-CVaR-Skewness model under ALD. This two models are implemented at a given expected portfolio return $r = 10\%$.

confidence level in the sense of a smaller risk, with exception of years 2008 and 2009.[27] In addition, we find that the Mean-CVaR-Skewness ALD model produces much smaller CVaR values than the Mean-CVaR Normal model at 99% confidence level which indicates that the former is superior to the latter in modelling extreme tail events.

To conclude, the in-sample, Monte Carlo simulation and Out-of-sample experiment have drawn consistent conclusions that the Mean-CVaR-Skewness optimization model with incorporation of ALD performs comparatively better than Mean-CVaR Normal optimization model when the stock returns in the portfolio exhibit heavy-tailed features, and the former is much more adoptable especially for investors who are more concerned about extreme tail risks.

2.6 Portfolio Configuration and risk-adjusted returns

Portfolio Configuration. We present the optimal portfolio weight, portfolio CVaR value and portfolio skewness value under different expected portfolio returns for the three datasets in Tables 2.9, 2.10 and 2.11 respectively, in order to observe the variation of stock allocation and portfolio risk level. From Table 2.9, we can see that asset 9 (“PG”) takes the largest weight (in a long position) of the portfolio in

[27]The year 2008 was the period of global financial crisis and the year 2009 was the recovery phase of this crisis (see Goh et al., 2012; Saranya and Prasanna, 2014).

Table 2.9: Optimal assets weights, portfolio CVaR and skewness values at different required return of the portfolio in $D1$

r	0.02	0.04	0.06	0.08	0.1	0.12	0.14
w1	-0.0205	0.0324	0.0853	0.1382	0.1911	0.2440	0.2969
w2	-0.1990	-0.1476	-0.0962	-0.0448	0.0066	0.0580	0.1095
w3	0.2133	0.1867	0.1602	0.1336	0.1070	0.0805	0.0539
w4	0.0109	0.0054	-0.0001	-0.0056	-0.0112	-0.0167	-0.0222
w5	0.1234	0.0831	0.0427	0.0024	-0.0380	-0.0783	-0.1187
w6	-0.0670	-0.0693	-0.0717	-0.0740	-0.0763	-0.0787	-0.0810
w7	0.1153	0.1147	0.1141	0.1134	0.1128	0.1122	0.1116
w8	-0.0085	-0.0506	-0.0927	-0.1348	-0.1769	-0.2190	-0.2611
w9	0.3068	0.3772	0.4476	0.5181	0.5885	0.6589	0.7293
w10	-0.0390	-0.0709	-0.1028	-0.1347	-0.1665	-0.1984	-0.2303
w11	-0.0070	0.0109	0.0289	0.0468	0.0647	0.0826	0.1006
w12	0.2034	0.1727	0.1419	0.1112	0.0804	0.0497	0.0189
w13	0.1538	0.1034	0.0530	0.0025	-0.0479	-0.0983	-0.1487
w14	-0.0042	0.0083	0.0208	0.0333	0.0458	0.0583	0.0708
w15	0.0623	0.0571	0.0519	0.0467	0.0415	0.0363	0.0311
w16	0.0178	0.0143	0.0107	0.0071	0.0035	0.0000	-0.0036
w17	0.0378	0.0387	0.0395	0.0403	0.0411	0.0420	0.0428
w18	0.1004	0.1337	0.1670	0.2003	0.2336	0.2669	0.3002
CVaR(0.99)	2.7062	2.6757	2.7903	3.0272	3.3565	3.7523	4.1949
CVaR(0.95)	1.8163	1.7925	1.8665	2.0226	2.2409	2.5039	2.7983
CVaR(0.90)	1.4330	1.4122	1.4686	1.5900	1.7605	1.9663	2.1969
Skewness	0.0753	0.1490	0.2101	0.2544	0.2839	0.3028	0.3147

$D1$ ("pre-crisis") for the given portfolio required return r, whereas, asset 7 ("JNJ") (in a long position) becomes more valuable in time of financial crisis ($D2$). We can notice the variations of stocks' weight with the changing of r in all the three datasets, either in a long position or in a short position. The long position or short position themselves are not a form of diversification in the Mean-CVaR-ALD model as the essence of diversification consists of reducing the risks by owning a variety of different securities. However, taking a short position can be more risky than taking a long since there are no limitations in terms of the losses. Therefore, shorting stocks in the portfolio may not reduce risks as much as buying them in this diversified portfolio.

For $D1$("pre-crisis") and $D3$ ("all"), an interesting finding is that when r rises,

Table 2.10: Optimal assets weights, portfolio CVaR and skewness values at different required return of the portfolio in $D2$

r	0.02	0.04	0.06	0.08	0.1	0.12	0.14
w1	0.0165	0.0534	0.0903	0.1272	0.1641	0.2010	0.2379
w2	-0.0325	-0.0198	-0.0072	0.0055	0.0181	0.0308	0.0434
w3	-0.1922	-0.1165	-0.0408	0.0348	0.1105	0.1862	0.2619
w4	-0.0173	-0.0161	-0.0149	-0.0138	-0.0126	-0.0115	-0.0103
w5	0.2154	0.2499	0.2845	0.3190	0.3536	0.3881	0.4226
w6	-0.0089	-0.0359	-0.0630	-0.0900	-0.1171	-0.1442	-0.1712
w7	0.4225	0.4678	0.5132	0.5585	0.6039	0.6493	0.6946
w8	0.0067	-0.0196	-0.0460	-0.0723	-0.0987	-0.1250	-0.1513
w9	0.2105	0.2010	0.1915	0.1820	0.1725	0.1629	0.1534
w10	-0.0211	-0.0219	-0.0227	-0.0235	-0.0243	-0.0251	-0.0258
w11	-0.0666	-0.0987	-0.1308	-0.1629	-0.1950	-0.2270	-0.2591
w12	0.1585	0.1637	0.1688	0.1740	0.1792	0.1844	0.1896
w13	0.1418	0.1622	0.1826	0.2030	0.2234	0.2438	0.2642
w14	-0.0505	-0.0398	-0.0291	-0.0185	-0.0078	0.0029	0.0135
w15	0.1950	0.1894	0.1838	0.1782	0.1726	0.1670	0.1613
w16	-0.0486	-0.0738	-0.0990	-0.1242	-0.1494	-0.1746	-0.1998
w17	0.0132	-0.0239	-0.0611	-0.0982	-0.1353	-0.1725	-0.2096
w18	0.0576	-0.0212	-0.1000	-0.1788	-0.2576	-0.3364	-0.4153
CVaR(0.99)	4.3752	4.2329	4.1838	4.2274	4.3575	4.5641	4.8357
CVaR(0.95)	2.9384	2.8395	2.8033	2.8295	2.9138	3.0496	3.2291
CVaR(0.90)	2.3196	2.2394	2.2088	2.2274	2.2921	2.3974	2.5372
Skewness	0.0470	0.0957	0.1432	0.1864	0.2233	0.2532	0.2766

the CVaR values at 99%, 95% and 90% confidence levels increase with the only exception of a slight decrease at 4% expected portfolio return. This phenomenon is basically consistent with the fact that higher returns are generally accompanied by higher risks. However, if we consider the financial crisis time period, that seems not to be the case. Results indicate that within a 10% amount of required returns, risks are quite volatile and could not convey certain information to investors. That is, higher returns are not accompanied by higher risks anymore, thus investors may able to cash higher returns with relatively low risks. By comparing portfolio CVaR values of the three different economic periods, another important observation is that the values of portfolio CVaR in $D1$ are always lower than the corresponding portfolio CVaR values in the periods that including financial crisis ($D2$ and $D3$). Interestingly, one identical finding is that the skewness of portfolio always exhibits

Table 2.11: Optimal assets weights, portfolio CVaR and skewness values at different required return of the portfolio in $D3$

r	0.02	0.04	0.06	0.08	0.1	0.12	0.14
w1	0.0071	0.1189	0.2306	0.3424	0.4542	0.5660	0.6778
w2	-0.0273	-0.0043	0.0188	0.0418	0.0648	0.0878	0.1108
w3	-0.1026	-0.0719	-0.0412	-0.0105	0.0201	0.0508	0.0815
w4	-0.0002	-0.0284	-0.0565	-0.0847	-0.1129	-0.1410	-0.1692
w5	0.2673	0.2612	0.2551	0.2490	0.2429	0.2368	0.2307
w6	-0.0308	-0.0602	-0.0896	-0.1190	-0.1485	-0.1779	-0.2073
w7	0.5025	0.4729	0.4433	0.4137	0.3841	0.3545	0.3249
w8	-0.0119	-0.0456	-0.0793	-0.1130	-0.1467	-0.1803	-0.2140
w9	0.1608	0.2271	0.2935	0.3599	0.4263	0.4927	0.5591
w10	0.0003	-0.0408	-0.0820	-0.1232	-0.1643	-0.2055	-0.2466
w11	-0.0885	-0.1309	-0.1732	-0.2155	-0.2579	-0.3002	-0.3425
w12	0.1587	0.1716	0.1845	0.1974	0.2103	0.2232	0.2361
w13	0.1812	0.1615	0.1419	0.1222	0.1025	0.0829	0.0632
w14	-0.0417	-0.0303	-0.0189	-0.0074	0.0040	0.0155	0.0269
w15	0.1807	0.1924	0.2042	0.2159	0.2276	0.2394	0.2511
w16	-0.0750	-0.0915	-0.1081	-0.1246	-0.1411	-0.1577	-0.1742
w17	-0.0278	-0.0482	-0.0686	-0.0890	-0.1094	-0.1298	-0.1502
w18	-0.0527	-0.0536	-0.0545	-0.0554	-0.0563	-0.0572	-0.0581
CVaR(0.99)	3.1414	3.0966	3.2909	3.6781	4.2012	4.8144	5.4870
CVaR(0.95)	2.1089	2.0755	2.2030	2.4602	2.8087	3.2178	3.6669
CVaR(0.90)	1.6642	1.6358	1.7345	1.9357	2.2090	2.5303	2.8831
Skewness	0.0651	0.1295	0.1799	0.2124	0.2310	0.2411	0.2463

a significant increasing trend if the required return of portfolio rises.

Given an expected portfolio return, we can calculate the portfolio CVaR and skewness, thus a three dimensional efficient frontier of the Mean-CVaR-Skewness model can be constructed. A portfolio, which is able to generate the minimum risk for a given expected return would be preferred by investors who are risk-averse, hence the most risk-efficient portfolios are those that exactly located at the efficient frontier. For a lower confidence level, the larger expected return seems to be a better choice because the portfolio CVaR values are more likely to be smaller.[28] This is what we commonly called "high return but low risk". The visualized efficient frontier

[28]Note that our result is consistent with the finding of Liu et al. (2005) who have theoretically investigated the evolution of Mean-CVaR efficient frontier based on the Normality assumption of risky assets in the portfolio.

at different time periods are shown in Figures 2.6, 2.7 and 2.8 respectively.

Sharpe ratio and Sharpe-like ratio. As Markowitz argued, risk/return profiles of assets should not be considered separately but in the context of a portfolio. A portfolio is considered efficient either when it has a minimal risk level for a given expected return or when it can maximize expected return for a given level of risk. The difference between these two concepts is reflected in algorithm distinctions. The former deals with a quadratic objective function with linear constraints while the latter is intended to solve a linear objective function with quadratic constraints. Therefore, in order to examine the performance of the investment by adjusting for its risks, we calculate the Sharpe ratio and Sharpe-like ratio of the portfolio at the three different time periods.[29] The top left part of Figure 2.9 plots the percentage Sharpe ratio values based on the minimal variance portfolio framework. It shows a clear changing tendency when the required rate of return of the portfolio increases. We can observe that the Sharpe ratio values of the portfolio in $D2$ are relatively lower than those in $D1$ especially at high expected returns, which indicate a less attractive risk-adjusted return by holding the portfolio in the period of global financial crisis.

However, it would be contentious to calculate a Sharpe ratio when asset returns are not normally distributed. The remaining three graphs in Figure 2.9, in which the portfolio CVaR based Sharpe ratio (or Sharpe-like ratio) at different confidence levels is presented, solve this problem. The Sharpe-like ratio, which is defined as the ratio of expected excess return to portfolio CVaR deviation, replaces the standard deviation with the coherent CVaR deviation measure.[30] We should distinguish between the CVaR risk measure and the CVaR deviation measure. The former defines losses versus zero while the latter defines mean values of the portfolio returns. The relationship between CVaR risk measure ($CVaR_\alpha(X)$) and CVaR deviation

[29] The calculation of Sharpe ratio is by: $(r - r_f)/V_W$, where r is the required portfolio return, r_f is the risk free rate and V_W is the portfolio standard deviation.

[30] Standard deviation is not a coherent risk measure and has received lots of criticisms for its Normality assumption for return distributions. Sarykalin et al. (2008) in their key observations argue that CVaR deviation is a strong "competitor" to the standard deviation and can be applied in financial concepts, such as sharpe ratio and portfolio beta, etc.

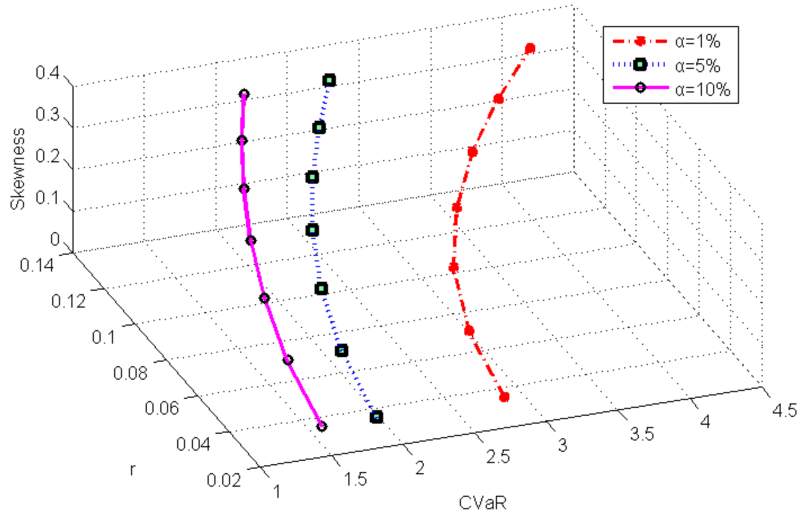

Figure 2.6: The mean-CVaR-skewness efficient frontier in 2002-2006

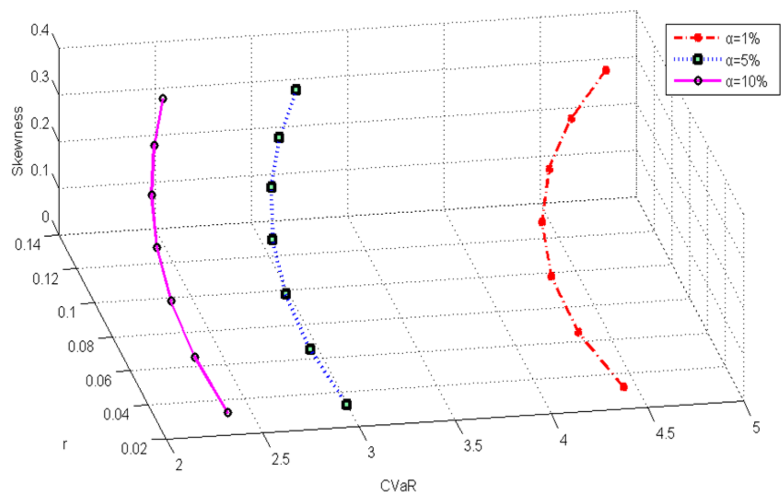

Figure 2.7: The mean-CVaR-skewness efficient frontier in 2007-2009

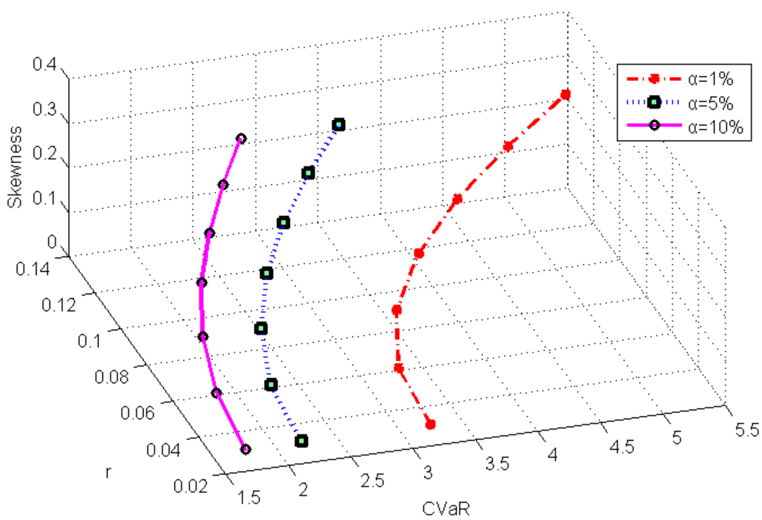

Figure 2.8: The mean-CVaR-skewness efficient frontier in 2002-2012

measure ($CVaR^{\Delta}_{\alpha}(X)$) can be formulated as:

$$CVaR_{\alpha}(X) = CVaR^{\Delta}_{\alpha}(X) + E(X) \tag{2.49}$$

From Figure 2.9, we can see that the Sharpe-like ratios depict similar trajectories as the Sharpe ratio at various confidence levels. However, compared to the values of the Sharpe ratio, the vertical Sharpe-like ratio values at various confidence levels are smaller.[31] A consistent finding is that, in terms of both Sharpe ratio and Sharpe-like ratio, the portfolio performance in $D2$ is more likely to be worse than the performance in $D1$ while portfolio performance in $D3$ falls across them. Hence, we conclude that in the framework of the Mean-CVaR-Skewness ALD optimization model, the inclusion of periods of global financial crisis can result in a relatively lower Sharpe-like ratio for the holding portfolio at any confidence levels.

Moreover, by using CVaR as a risk measure, investors can be provided with a wider range of choices according to their personal risk preferences. We find that conservative investors who may prefer a larger confidence level, and thus estimate a larger risk, could have a relatively lower Sharpe-like ratio than aggressive investors who may estimate a smaller risk with a smaller confidence level.

2.7 Conclusion

This paper presents a comprehensive empirical analysis of the Mean-CVaR-Skewness portfolio optimization model with incorporation of ALD. The theoretical analysis of the model is provided in Zhao et al. (2015), but there are as yet no relevant empirical investigations of the model performance. Therefore, one of the main objectives of this paper is to examine the model performance by constructing a risky portfolio. Extending Zhao et al.'s (2015) framework, we study a new optimization model using VaR as a risk measure and assuming asset returns are Asymmetric Laplace distributed, thus construct the Mean-VaR-Skewness model under ALD. We find that Mean-CVaR-Skewness and Mean-VaR-Skewness model have same

[31] Considering the fact that using factor variance as a risk measure in a non-Normal world may lead to underestimation of portfolio risks, one thus calculates a higher and improper Sharpe ratio.

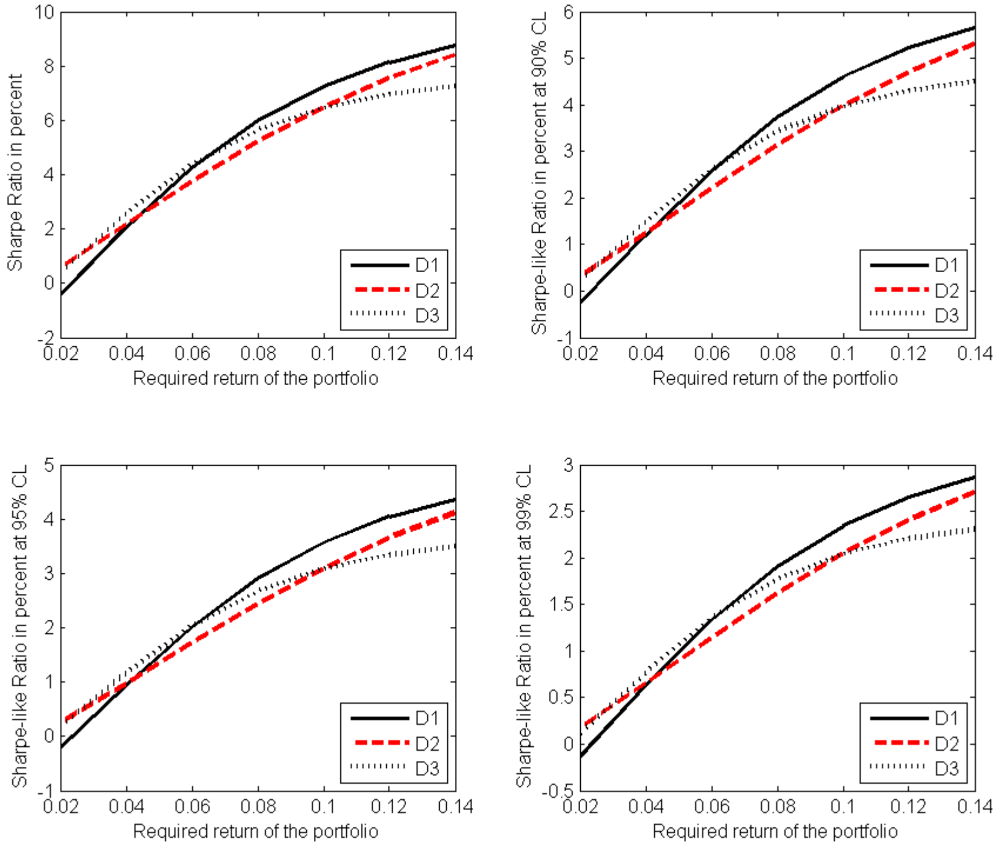

Figure 2.9: Sharpe ratio and Sharpe-like ratio of the portfolio in three time periods

solution after reducing the multi-objective portfolio selection problem into a single-target quadratic programming problem. The closed-form solution of the simplified optimization model can be straightforwardly obtained using the Lagrange multiplier method, the algorithm of which can be considered as an extension of the original Markowitz portfolio model (Mean-Variance model).

Both the Mean-CVaR-Skewness and Mean-VaR-Skewness models have been examined using data from three different time periods in order to observe the consistence of model performance. For the single asset VaR in the constructed portfolio, we find that the underestimation of single VaRs becomes more obvious when moving further to the tails if heavy-tailed asset returns are assumed to be normally distributed. Additionally, by comparing the portfolio VaR model under ALD with the parametric-Normal portfolio VaR and historical portfolio VaR model, we find that the ALD assumption can accurately replicate the tail characteristics of asset returns. The smaller RMSE, MAE and SSRE values further reinforce this finding.

We then empirically investigate the Mean-CVaR-Skewness model under ALD from in-sample, Monte Carlo simulations and out-of-sample forecasting perspectives, and simultaneously compare it with the existing benchmark Mean-CVaR model which assumes that asset returns are normally distributed. Results consistently show that the Mean-CVaR-Skewness optimization model with the incorporation of ALD performs comparatively better than the Mean-CVaR model under Normal distribution when the stock returns in the portfolio exhibit heavy-tailed features. In addition, we find that in $D1$ ("pre-crisis") and $D3$ ("all"), higher expected portfolio returns (r) are generally accompanied by higher risk levels at a specific confidence level. It should be emphasized that at lower confidence levels, we find evidence of "high return but low risk" in the Mean-CVaR-Skewness model under ALD. Furthermore, one identical finding is that the portfolio skewness responds positively to a higher r. Finally, we assess the risk-adjusted performance of the Mean-CVaR-Skewness model by calculating the Sharpe-like ratio at different economic periods. We find that the risk-adjusted portfolio returns are relatively low at a given confidence level when we consider the impact of the global financial crisis. Compared to aggressive investors at the same r level, conservative investors who are more sensitive to risk will inevitably have a lower Sharpe-like ratio.

Chapter 3

Risk Measuring under Heston Stochastic Volatility Model

3.1 Introduction

As an inherent component of the financial world, the market risk, in the sense of the possibility of suffering losses from unexpected movements of asset prices, is of utmost importance for individual investors, financial institutions and corporate entities. However, there is no unified agreement on the measure to quantify the market risk. A number of measures, such as standard deviation, lower partial moments and quantiles are generally used. A standard benchmark risk measure that has been widely adopted since 1994 is the Value-at-risk (VaR), which was first proposed by J.P. Morgan. VaR summarizes the potential largest loss over a target horizon for a given confidence level. To ensure that banks have adequate capital to expose themselves to the operational risks, the Basel Committee on Banking Supervision published the Basel II Accord (2004), which established the risk and capital management requirements and recommended VaR as the preferred approach to measure market risk. Despite its conceptual simplicity and applicability to most financial instruments, the drawbacks are obvious, and there is thus reason to be skeptical about its accuracy and reliability. VaR violates the subadditivity criterion of a coherent risk measure, which may lead to concentration rather than diversification of a portfolio. In addition, VaR fails to take account of the losses that exceed the threshold value for a given risk tolerance level. The Conditional

Value-at-risk (CVaR) that was proposed by Artzner et al. (1999) overcomes these disadvantages, as it is a coherent risk measure and considers the average losses that exceed VaR.

The existing approaches for VaR and CVaR estimation in practice can be classified into three categories: non-parametric historical simulation approach, Monte Carlo simulation approach and parametric-Normal approach. It is worth noting that the implementation of the parametric-Normal approach depends on the Normality assumption of the underlying asset distribution even if extensive studies have shown the existence of heavy-tailed and leptokurtic features in financial asset returns. Essentially, the VaR/CVaR calculation is to specify the quantiles of asset or portfolio returns. As such, the estimation results are very sensitive to the tail behavior of the distribution of the loss function. As a result, a broad range of heavy-tailed distributions are applied to modelling the extreme tail risk of the underlying return series. Dokov et al. (2007) construct the VaR and CVaR model using the skewed-t distribution, whilst Bali and Theodossiou (2007) build risk models based on the skewed generalized-t (SGT) distribution. Fan et al. (2008) embed the Generalized Error distribution (GED) in the Generalized Autoregressive Conditional Heteroscedasticity (GARCH) type models, finding that VaR is effective in in-sample and out-of-sample performance compared to other methods. Moreover, Gerlach et al. (2011) employ the Asymmetric Laplace distribution (ALD) as the error innovation of return equation in a quantile GARCH model, arguing that the proposed model provides more accurate VaR forecasting than other alternatives.

Unlike conventional methods that characterize the financial dynamics using probability density function ($p.d.f.$), this paper provides an alternative route to calculate VaR and CVaR that is fully characterized in terms of the characteristic function to describe the random variables. One of the important applications of characteristic function is the analysis of linear combinations of independent random variables in the context of portfolio construction (i.e. Chapter 2 in this thesis). In addition, it completely defines the probability distribution of a number of real-valued random variables. Considering the one-to-one relationship between the variables, the characteristic function of a distribution can be obtained by taking the Fourier transform of its *p.d.f.*.[1] This indicates that our methods are readily applicable to a wide range of models, not only those with an explicit closed-form *p.d.f.*, but are also especially suited to those distributions whose *p.d.f.* are not known analytically.[2] The construction of VaR and CVaR in our approach follows the framework of Lewis (2001) by using the generalized Fourier transform technique. This entails extending the argument of the characteristic function from a real number to a complex number (also see Bormetti et al., 2010).[3] This requires the target characteristic function to be reformulated as an extended form, and can thus be evaluated in our models.

One possible way to estimate VaR and CVaR is to apply the GARCH-type models using the Autoregressive Conditional Heteroscedasticity (ARCH) model introduced by Engle (1982). However, stochastic volatility (SV) model, as an innovative alternative for modelling the time-varying volatility process, is more appropriate mainly due to its natural capability to capture the volatility behavior of actual asset returns in financial markets (see, e.g., Pederzoli, 2006). The existing literature on risk management using SV models is generally based on the discrete time scenario (e.g., Li, 2006; Zhou and Liu, 2010; Zhou et al., 2012), while the SV model under continuous time scheme is intensively applied in option pricings (e.g., Nicolato and Venardos, 2003; Cont and Voltchkova, 2005; Cai and Kou, 2011; Tong, 2016). This paper contributes to the existing studies on VaR and CVaR calculation by employing the Heston (1993) continuous time stochastic volatility model, the characteristic function of which is analytically tractable, thus ensuring its applicability to our proposed risk models.

The development of Heston's (1993) stochastic volatility model was prompted by

[1] The Fourier transform method shows how to extract *p.d.f.* from the characteristic function, and the most extensive applications are in option pricings (see, e.g., Bakshi and Madan, 2000; Lewis, 2001; Wu, 2007; Hurd and Zhou, 2010).

[2] Note that the characteristic function of a distribution always exists, even if the corresponding *p.d.f.* is not analytically tractable.

[3] Lewis (2001) employ a generalized Fourier transform approach to evaluate option prices by allowing the argument of characteristic function $\phi \to z \in \mathbb{C}$, while Bormetti et al. (2010) is the first to apply this technique to risk management. More details will be discussed in the following section.

the flaw in the benchmark option pricing model: the Black-Scholes formula, which is based on the assumption of log-Normal return process with constant volatility. Two driving mechanisms for the diffusions are proposed in Heston dynamics, including the geometric Brownian motion process of return series and the mean-reversion Ornstein-Uhlenbeck process of volatility. The Heston model's prominence amongst financial practitioners can be attributed to its realistic-closed properties, such as the capturing ability of skewness and kurtosis for real financial data and the non-negative and mean-reverting feature of return volatility process. In addition, it allows for a correlation factor between return series and volatility process, which is generally shown to be negative in reality. Most importantly, the closed-form characteristic function is tractable by solving the systemic partial differential equations (PDE), and can thus be applied to a great deal of financial models. This computational efficiency is critical when evaluating the model using market prices and is the greatest advantage of Heston over other potentially realistic SV models, such as in the literature by Hull and White (1987), Johnson and Shanno (1987), Wiggins (1987) and Melino and Turnbull (1990) (also see Moon et al., 2009; Čížek et al., 2011). Therefore, in this paper, we study the evolution of the extended characteristic function (ECF) of the Heston model and its application to the proposed VaR and CVaR models.

The main advantages of our risk models are twofold. First, we provide a general framework for VaR and CVaR computation with the employment of function T_v and H_v.[4] These two functions are suited to a variety of financial models and especially work for those distributions whose *p.d.f.* are not known analytically, because the only input required in our approach is the closed-form expression of the characteristic function. When the characteristic function is obtainable, we can evaluate function T_v and H_v, and thus approximate the VaR and CVaR values. Second, our approach provides great flexibility for people who have different risk tolerance levels. The VaR and CVaR approximations can be produced using a numerical integration algorithm by setting a wide range of risk levels first. This overcomes the drawbacks of the

[4] Function T_v and H_v are two derived formulas for calculating VaR and CVaR, details see equation (3.17) and (3.18).

traditional VaR and CVaR models, in which the value has to be recalculated once the risk level is changed.

As an important precursor to our paper, the study of Bormetti et al. (2010) is the first to efficiently conjugate the risk measures to the models that are defined by the characteristic function. However, our study differs from that of Bormetti et al. (2010) in three important respects. First, we are deriving the VaR and CVaR model strictly starting from their mathematical definitions, and assuming that the asset returns are the natural logarithm of stock prices. While in Bormetti et al.'s (2010) paper, the centered logarithm returns are used. This leads to different CVaR expressions, though we have the same VaR formula.[5] Second, contrary to Bormetti et al. (2010), who adopted Heston's cumulant generating function for parameter estimation and then conducted empirical analysis, we derive an explicit ECF of the Heston model and investigate the impact of the grid size on the VaR and CVaR approximations in a numerical integration framework. This process is realized by setting the initial parameter values and selecting a range of risk levels in advance. Third, our study focuses on the determinants of VaR and CVaR approximations. Therefore, we seek to explain the VaR and CVaR variations via changing parameter values in the Heston model, while keeping everything else constant. Our aim is to answer the following questions: how the grid size spacing influences the accuracy of VaR and CVaR approximations in a trapezoidal integration scheme? Which factors in the Heston case can persistently affect the VaR and CVaR values? How the factors play a major role in determining the movements of risk estimates?

The paper unfolds as follows. In section 2, we present the construction of the VaR and CVaR model in a Fourier space using the generalized Fourier transform technique and give the general framework for risk measurement. Section 3 provides a discussion of Heston dynamics, determination of PDE and evolution of Heston ECF. Section 4 presents numerical analysis for the proposed VaR and CVaR model under the Heston framework and gives numerical findings. Section 5 concludes.

[5] Since asset returns are generally obtained using natural logarithm in financial empirical studies, hence this paper derives the risk measures in such a realistic and simpler manner.

3.2 Model setup

The first part in this section presents the application of the Fourier transform method and the construction of VaR in a generalized Fourier transform framework. Then, we introduce the CVaR measure in the Fourier space as well as an equivalent systemic risk measures following the algorithm of Euler translation.

3.2.1 VaR measure in a Fourier space

The Fourier transform and its inverse are generally well defined in most functions encountered in financial practical analysis, and have been successfully applied to determine option prices, i.e. Carr and Madan (1999), Duffie et al. (2000), Hurd and Zhou (2010), Escobar and Gschnaidtner (2016). As mentioned by Zhu (2009), almost all new developed option valuation models have extensively involved Fourier transforms and their inverse form in option pricing calculations, mainly due to their capability in modeling stochastic processes and loss distributions. There are several common conventions to define the Fourier inversion $\mathcal{F}_f(\phi)$ of an integrable function $f : \mathbb{R} \to \mathbb{C}$ which satisfies the integrability condition:

$$\int_{-\infty}^{+\infty} |f(x)|dx < \infty \tag{3.1}$$

The one we will use is in the following form:

$$\mathcal{F}_f(\phi) = \int_{-\infty}^{+\infty} e^{i\phi x} f(x)dx \tag{3.2}$$

where $i = \sqrt{-1} \in \mathbb{C}$ is the imaginary unit. As the key identity for the application of Fourier transform to risk management in this paper, formula (3.5) stems from the Parseval theorem which presents the fact that the scalar product of two integrable functions $m, n : \mathbb{R} \to \mathbb{C}$ on $(-\infty, +\infty)$ is preserved under Fourier transform.[6] Define the inner product of m and n by:

$$< m, n >= \int_{-\infty}^{+\infty} m(x)\overline{n}(x)dx \tag{3.3}$$

[6] Application of Parseval theorem in option pricing and insurance can refer to Dufresne et al. (2009).

where $\overline{n}(x)$ is the complex conjugate of $n(x)$, then we have:

$$< m, n >= \frac{1}{2\pi} < \mathcal{F}_m(\phi), \mathcal{F}_n(\phi) > \tag{3.4}$$

Hence, a direct consequence is that given $\mathcal{F}_f(\phi)$, the function f in equation (3.2) can be recovered using the inverse Fourier transform:

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{-i\phi x} \mathcal{F}_f(\phi) d\phi \tag{3.5}$$

We consider now a stochastic process X as the random variable of the rate of return of an asset and L (with $L > 0$) as the maximum potential loss of the asset over a fixed holding time horizon τ, then equation $X = -L$ holds.[7] Given a confidence level $(1 - \alpha)$, the mathematical formula of VaR can be formulated by $prob(L \geq VaR_{1-\alpha}) = \alpha$, or equivalently written as $prob(X \leq -VaR_{1-\alpha}) = \alpha$.[8] Then, we can explicitly express the integral equation of VaR as follows:

$$\alpha = \int_{-\infty}^{-VaR} f(x) dx \tag{3.6}$$

where x denotes the rate of return of an asset at time t and can be obtained using the natural logarithm (i.e. $x_t = ln(s_t/s_{t-1})$), α represents the risk level and $f(x)$ is the $p.d.f$ associated with X.

For a parametric VaR and CVaR computation, the prescribed density function is generally required. It has turned out that even though the $p.d.f.$ of asset returns cannot be explicitly computed, the characteristic function of the return is still tractable, for example, the Heston stochastic volatility model and Levy process, etc. Hence, the characteristic function can be an efficient alternative because of its one to one relationship with the $p.d.f.$. To be specific, the characteristic function of a given stochastic process X is a Fourier transform of its $p.d.f.$. Following the representation in equation (3.2) where $f(x)$ represents the density of X, then the

[7]Note that a stochastic process can be written, among other ways, as $X(t)$, $\{X_t\}_{t \in T}$ or X_t. We use notation X here to define a series of random variables.

[8]VaR and CVaR defined in this paper is a positive value which indicate that a negative VaR or CVaR represents a gain rather than a loss.

Fourier transform $\mathcal{F}_f(\phi)$ is the characteristic function of X:

$$\mathcal{F}_f(\phi) = \psi_X(\phi) = E[e^{i\phi X}] = \int_{-\infty}^{+\infty} e^{i\phi x} f(x)dx \tag{3.7}$$

A generalized Fourier transform. If we take argument ϕ to be a complex number z where $z = w + iv$ has $w, v \in \mathbb{R}$ with $v \neq 0$, then $\mathcal{F}_f(z)$ is defined as the generalized Fourier transform of function f.[9] Lewis (2001) employed the generalized Fourier transform of the derivative value f to valuate option prices:

$$\mathcal{F}_f(z) = \int_{-\infty}^{+\infty} e^{izx} f(x)dx \tag{3.8}$$

Extending Lewis' (2001) approach to the context of risk management, we are able to calculate the *p.d.f.* of a stochastic process X in terms of its characteristic function $\psi_X(z)$ by taking the inverse Fourier transform.[10]

Definition 3.1. (*p.d.f.* of X). *For a complex number z that has $a < \mathrm{Im}(z) < b$, and defining $\psi_X(z) = \mathbb{E}[\exp(izX)]$ as the characteristic function of the process X, then we can obtain the p.d.f. of X through:*

$$f(x) = \frac{1}{2\pi} \int_{iv-\infty}^{iv+\infty} \psi(z) e^{-izx} dz \tag{3.9}$$

where $z = w + iv$ is a complex number, w is the real part that can be any real numbers and v is the imaginary part which belongs to the proper strip of regularity of the extended characteristic function (ECF) $\psi(z)$.

Suppose the stochastic process X with ECF $\psi(z)$ is regulated in the strip $S_X = \{z = w + iv : v \in (a, b)\}$ for the real numbers a and b such that $a < b$ (See Lewis, 2001). Then, to derive the VaR in the generalized Fourier transform framework, we substitute the *p.d.f* of stochastic process X (equation 3.9) into the integral equation

[9] For a complex number with a formation function of $z = a + bi$, we define the real part of this complex number as: $Re(z) = a$ while the imaginary part is given as: $Im(z) = b$. In this paper, both a and b are real numbers. However, if $Re(z) = a = 0$, we say that the complex number z is pure imaginary and z is pure real when $Im(z) = b = 0$.

[10] The difference between regular Fourier transform and generalized Fourier transform is that the former assumes ϕ is real while the latter uses complex $z \in \mathbb{C}$. The generalized Fourier transform exist by integrating a straight line in a z plane which indicate that the characteristic function $\psi(z)$ is regulated in the z-plane strip.

of VaR (equation 3.6), a composite expression can be obtained:

$$\alpha = \frac{1}{2\pi} \int_{iv-\infty}^{iv+\infty} \left(-\frac{e^{izVaR}}{iz} \right) \psi(z) dz \tag{3.10}$$

or equivalently written as:[11]

$$\alpha = \frac{e^{-vVaR}}{2\pi} \int_{-\infty}^{+\infty} \left(\frac{e^{iwVaR}}{v - iw} \right) \psi(w + iv) dw \tag{3.11}$$

3.2.2 CVaR measure in a Fourier space

The construction of CVaR in a Fourier space starts strictly from the mathematical definition which calculates the conditional expected tail loss of an asset over a fixed period τ at a given confidence level $(1-\alpha)$. We define the mathematic expression of CVaR as $CVaR_{1-\alpha} = E[L|L \geq VaR_{1-\alpha}]$, or equivalently written as an integral equation:

$$\begin{aligned} CVaR_{1-\alpha} =& - E[X|X \leq -VaR_{1-\alpha}] = \frac{-\int_{-\infty}^{-VaR} xf(x)dx}{prob(x \leq -VaR)} \\ =& - \alpha^{-1} \int_{-\infty}^{-VaR} xf(x)dx \end{aligned} \tag{3.12}$$

where α denotes the risk level. We use consistent notations as those of VaR and plug the *p.d.f.* of stochastic process X (equation 3.9) in the above integral formula of CVaR, yield the following CVaR with respect to the generalized Fourier transform $\psi(z)$:

$$CVaR = -\frac{1}{2\pi\alpha} \int_{iv-\infty}^{iv+\infty} \frac{e^{izVaR}}{iz} \left(VaR - \frac{1}{iz} \right) \psi(z)\, dz \tag{3.13}$$

or eventually shown as:[12]

$$CVaR = \frac{e^{-vVaR}}{2\pi\alpha} \int_{-\infty}^{+\infty} \frac{e^{iwVaR}}{v - iw} \left(VaR + \frac{1}{v - iw} \right) \psi(w + iv)\, dw \tag{3.14}$$

[11]See Appendix D.1 for the derivation of VaR under the generalized Fourier transform.

[12]See Appendix D.2 for the derivation of CVaR under the generalized Fourier transform.

Thus, the two risk measures VaR and CVaR under the framework of generalized Fourier transform algorithm in this paper can be integrated as:

$$\begin{cases} \alpha = \dfrac{e^{-vVaR}}{2\pi} \int_{-\infty}^{+\infty} \psi(w+iv) \left(\dfrac{e^{iwVaR}}{v-iw} \right) dw \\ CVaR = \dfrac{e^{-vVaR}}{2\pi\alpha} \int_{-\infty}^{+\infty} \psi(w+iv) \left(\dfrac{e^{iwVaR}}{v-iw} \right) \left(VaR + \dfrac{1}{v-iw} \right) dw \end{cases} \tag{3.15}$$

The integrand of VaR equation in (3.15) is a complex number, whereas the risk level α is a real number. This implies that we can ignore the imaginary part of the integrand and only consider the real part, which is even-valued.[13] This is similar to achieving the *p.d.f* which is a real-valued function while the imaginary part must be offset when taking the integration over the real line.[14] In addition, if we consider the symmetric property of the real and imaginary part of $\psi(z)$ over the integration interval, we can eventually recover equation (3.15) as a simplified form:[15]

$$\begin{cases} \alpha = \dfrac{\mathbf{Re}\, T_v(VaR, v)}{\pi} \\ CVaR = VaR + \dfrac{\mathbf{Re}\, H_v(VaR, v)}{\mathbf{Re}\, T_v(VaR, v)} \end{cases} \tag{3.16}$$

where

$$T_v(VaR, v) = e^{-vVaR} \left[\int_0^{+\infty} \frac{\psi(w+iv)}{v-iw}\, e^{iwVaR}\, dw \right] \tag{3.17}$$

and

$$H_v(VaR, v) = e^{-vVaR} \left[\int_0^{+\infty} \frac{\psi(w+iv)}{(v-iw)^2}\, e^{iwVaR}\, dw \right] \tag{3.18}$$

An alternative form for systemic equation (3.16) is shown below by taking the real part of function T_v and H_v into further consideration. By applying Euler's translation in a complex plane, it is not difficult to get the following real part of

[13] The ECF $\psi(z)$ has the property $\psi(z) = \overline{\psi}(-z)$ with $\overline{\psi}(-z)$ to be the complex conjugate. Thus $\psi(z)$ is an even function in its real part over z while odd in its imaginary part. Similar study can refer to Rough (2013) who has studied a simplified call prices formula by considering the fact that the real part of the integrand is even and the imaginary part is odd.

[14] For example, one can recovers the *p.d.f.* using the inverse Fourier transform: $f(x) = \frac{1}{\pi} \int_0^{+\infty} Re[\psi(z)e^{-izx}]dz$.

[15] See Appendix D.3 for derivation.

function T_v and H_v, expressed as TT_v and HH_v, respectively.

Definition 3.2. (Euler's Formula). *Let φ be any real number or complex number, thus $e^{i\varphi} = cos\varphi + i\, sin\varphi$, where the argument φ is defined as radians, e is the base of natural logarithm, i is the imaginary unit, the trigonometric function "$cos\varphi$" represents real part of a complex number and the trigonometric function "$sin\varphi$" is the imaginary part.*

Remark 3.3. *For any two random complex numbers expressed in the form $a + bi$ and $c + di$ with $a, b, c, d \in \mathbb{R}$, we have:*

$$\frac{a+bi}{c+di} = \frac{(a+bi)(c-di)}{(c+di)(c-di)} = \frac{ac+bd}{c^2+d^2} + \frac{(bc-ad)i}{c^2+d^2} \tag{3.19}$$

By applying Euler's formula and Remark 3.3 to equation (3.16), we can explicitly write the real part of function T_v as a function of trigonometric function, which produces an alternative for the VaR computation:

$$\begin{aligned} & TT_v(VaR, v) = \mathbf{Re}\ T_v(VaR, v) \\ =& e^{-vVaR} \int_0^{+\infty} \frac{dw}{v^2+w^2} \{cos(wVaR)\, [v\, \mathbf{Re}\, \psi(w+iv) - w \mathbf{Im}\, \psi(w+iv)]\} \\ & - e^{-vVaR} \int_0^{+\infty} \frac{dw}{v^2+w^2} \{sin(wVaR)\, [w\, \mathbf{Re}\, \psi(w+iv) + v\, \mathbf{Im}\, \psi(w+iv)]\} \end{aligned} \tag{3.20}$$

Likewise, a new transform of the real part of function H_v in terms of CVaR computation can be formulated as:[16]

$$\begin{aligned} & HH_v(VaR, v) = \mathbf{Re}\ H_v(VaR, v) \\ =& e^{-vVaR} \int_0^{+\infty} \frac{dw}{(v^2+w^2)^2} \left\{cos(wVaR) \left[(v^2-w^2)\, \mathbf{Re}\, \psi(w+iv) - 2vw \mathbf{Im}\, \psi(w+iv)\right]\right\} \\ & - e^{-vVaR} \int_0^{+\infty} \frac{dw}{(v^2+w^2)^2} \left\{sin(wVaR) \left[2vw\, \mathbf{Re}\, \psi(w+iv) + (v^2-w^2)\, \mathbf{Im}\, \psi(w+iv)\right]\right\} \end{aligned} \tag{3.21}$$

[16]See Appendix D.4 and D.5 for derivation.

where $\mathbf{Re}\,\psi(w+iv)$ and $\mathbf{Im}\,\psi(w+iv)$ are the real and imaginary parts of $\psi(w+iv)$, respectively. As a result, we obtain an alternative expression of risk measures VaR and CVaR with respect to trigonometric function:

$$\begin{cases} \alpha = \dfrac{TT_v(VaR, v)}{\pi} \\ CVaR = VaR + \dfrac{HH_v(VaR, v)}{TT_v(VaR, v)} \end{cases} \tag{3.22}$$

The two equalities obtained in both equation (3.16) and (3.22) for calculating VaR and CVaR in a generalized Fourier transform framework represent the first contribution of this paper, which proposes the numerical calculation of the VaR and CVaR for an interest stochastic process X that is completely specified by its characteristic function.[17] As most studies have argued, the prerequisite to evaluate a parametric VaR and CVaR is undoubtedly to identify a closed-form $p.d.f.$ of the distribution. In this paper, equations (3.16) and (3.22) ease this restriction and are readily applicable to a number of interesting innovations whose $p.d.f.$ are not well-defined analytically. This is realized because the characteristic function of a stochastic process X always exists, and has a unique interdependent relationship with the $p.d.f.$. For example, Lewis (2001) listed a number of characteristic functions for Lévy Processes in Table 2.1 to evaluate option prices in a Fourier space.

In addition, an important factor for investors who are evaluating VaR and CVaR is the choice of the risk tolerance level α. In general, the VaR and CVaR are computed for a fixed α (e.g., Chapter 2 and Chapter 4 in this thesis). However, this seems a limitation for evaluating the risk of security prices using a continuous time stochastic process (e.g., geometric Brownian motion or Lévy Process). Equations (3.16) and (3.22) provide flexibility and are analytically tractable to overcome this drawback by varying a series of target α values. The only input required in our risk model is the extended characteristic function. As long as it is specified, a grid of w

[17] The first paper that conjugate the risk measuring tools with financial dynamic models that are fully characterized by its characteristic function is the one that studied by Bormetti et al. (2010), to the best of our knowledge. However, unlike Bormetti et al. (2010) who are using centered logarithmic returns to define VaR and CVaR. In this paper, we adopt the natural logarithm algorithm to define the asset returns like most empirical works did and the derivation starts strictly from the original mathematic definition of VaR and CVaR. The analytic difference reflects on the CVaR expression while VaR is consistent.

(real part of z) values can be set with an admission value v (imaginary part of z) so that the corresponding VaR and CVaR can be computed along with a number of prescribed α values.[18] The solution of equations (3.16) and (3.22) involves a numerical integration algorithm such as the Fast Fourier Transform (FFT) method or the trapezoidal integration rule, the latter of which that we adopted in this paper will be discussed in detail in a later section.

3.3 The Heston stochastic volatility model

As shown in equations (3.16) and (3.22), the VaR and CVaR computation requires the analytical expression of a characteristic function. The algorithm of our risk model is highly consistent with dynamic models used in the context of option valuation under stochastic volatility (i.e. exponential Ornstein-Uhlenbeck model, Heston model and Stein-Stein model, etc). Hence, it can be applied to thses models using equations (3.16) and (3.22) as long as the characteristic function is analytically tractable. Therefore, we explore the VaR and CVaR model on the basis of the extended characteristic function of the Heston model which has gained great popularity in option pricing but less in risk measuring.

3.3.1 Heston dynamics and parameters

The Heston model (1993) assumes that both the underlying asset price and its volatility follow diffusion processes. In contrast to Black and Scholes (1973) who used a constant volatility, the stochastic volatility over time can be more accurate for pricing options than the Black-Scholes model. The model is given by the bivariate system of stochastic differential equations (SDEs):

$$dS_t = \mu S_t d_t + \sqrt{V_t} S_t dW_t^S \tag{3.23}$$

$$dV_t = \kappa(\theta - V_t) d_t + \xi \sqrt{V_t} dW_t^V \tag{3.24}$$

[18] Essentially, equations (3.16) and (3.22) design a similar mechanism as, for example, the Heston option pricing model, in the sense that a number of call prices are obtainable when a series of strike prices are assigned.

where S_t and V_t are the underlying price and volatility, following a Black-Scholes type stochastic process and Cox-Ingersoll-Ross process (Cox et al., 1985), respectively. Other parameters of the model are:

- V_t is the variance for underlying asset price returns
- μ is the rate of return of the asset
- θ is long-term variance, if t tends to infinity, then $E(V_t)$ tends to θ
- κ is the speed of variance mean-reversion
- ξ is the non-stochastic volatility of variance process, or volatility of volatility
- dW_t^S and dW_t^V are two correlated Wiener processes: $E^{\mathbb{P}}[dW_t^S dW_t^V] = \rho dt$

The assumption of correlation between the two Brownian motions can be formulated as follows:

$$dW_t^V = \rho dW_t^S + \sqrt{1-\rho^2} dW_t \tag{3.25}$$

where W_t is a Wiener process independently of W_t^S and $\rho \in [-1, 1]$ is the correlation coefficient.

The Heston stochastic volatility model, among various volatility models, has desirable properties in modeling specific characteristics that generally are observed in the behavior of financial markets.[19] Moreover, the Heston can follow a number of different distributions and it allows for the statistical interdependence factor (ρ) between the underlying asset returns and volatility. Since empirical analysis has substantially demonstrated the heavy-tailed and leptokurtic features of asset returns in financial markets, the variation of ρ can flexibly reflect the asymmetry level of density distributions.

To illustrate, we plot Figure 3.1 which shows the simulated asymmetric variation

[19] Such as the assumption that V_t follows a mean-reverting Cox-Ingersoll-Ross process properly describing the behavior of volatility in real markets.

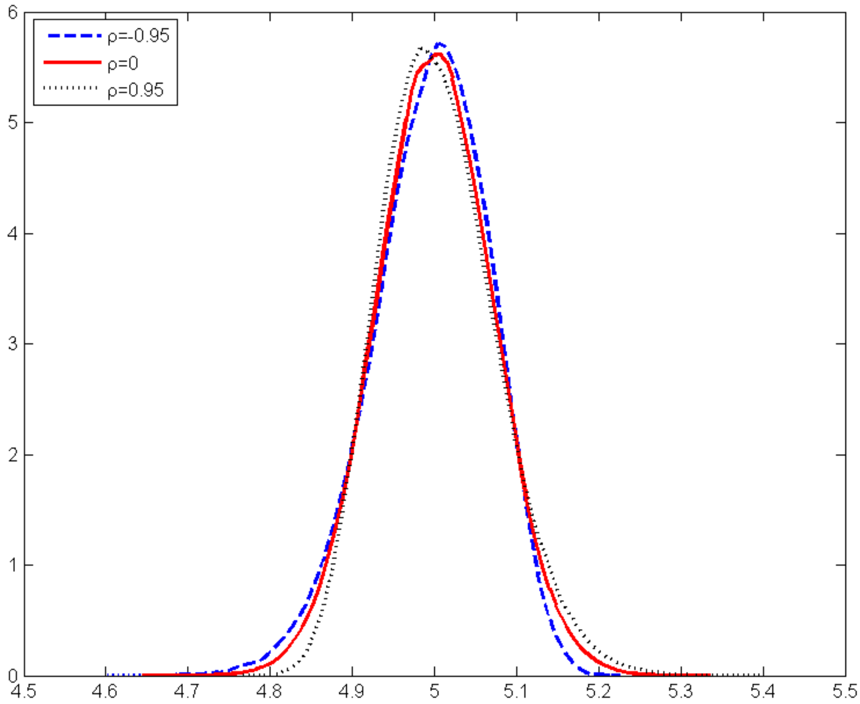

Figure 3.1: Asymmetry variations of the simulated density distributions under different ρ. The setting parameters: $S_0 = 5$, $V_0 = 0.01$, $\mu = 0$, $\kappa = 2$, $\Delta t = 0.02$, $\theta = 0.01$, $\xi = 0.1$

of the density distribution for different values of ρ.[20] The parameter setting is: $S_0 = 5$, $V_0 = 0.01$, $\mu = 0$, $\kappa = 2$, $\Delta t = 0.02$, $\theta = 0.01$, $\xi = 0.1$. Specifically, when $\rho < 0$, the volatility tends to decrease as asset returns increase, and vice versa. For example, if we set $\rho = -0.95$, the distribution is skewed to the left with heavier left tails comparing to the scenario that $\rho = 0$. Conversely, a positive ρ implies that volatility increases when asset returns rise. The graph with $\rho = 0.95$ suggests a right skewed distribution with fatter right tails. Empirically, ρ is usually negative for financial time series but the economic explanation regarding the leverage effect between asset price and volatility shocks is controversial.[21]

[20]The simulation work is conducted by discretizing the stochastic processes based on Euler-Maruyama method (Moodley, 2005). Specifically, Heston dynamics can be written as: $S_t = S_{t-1} + \mu S_{t-1}\Delta t + \sqrt{V_{t-1}}S_{t-1}\sqrt{\Delta t}Z_t^S$ and $V_t = V_{t-1} + \kappa(\theta - V_{t-1})\Delta t + \xi\sqrt{V_{t-1}}\sqrt{\Delta t}Z_t^W$, where $\{Z_t^S\}_{t\geq 0}$ and $\{Z_t^W\}_{t\geq 0}$ are two standard Normal random variables with correlations $Z_t^S = \varphi_t^S$ and $Z_t^W = \rho\varphi_t^S + \sqrt{1-\rho^2}\varphi_t^W$, in which φ_t^S and φ_t^W are two standard Normal random variables.

[21]The negative correlation reflects on the decline of asset price when volatility goes up, but for simplicity it is typically be assumed to be a constant in empirical scenarios (See Duffie et al., 2000).

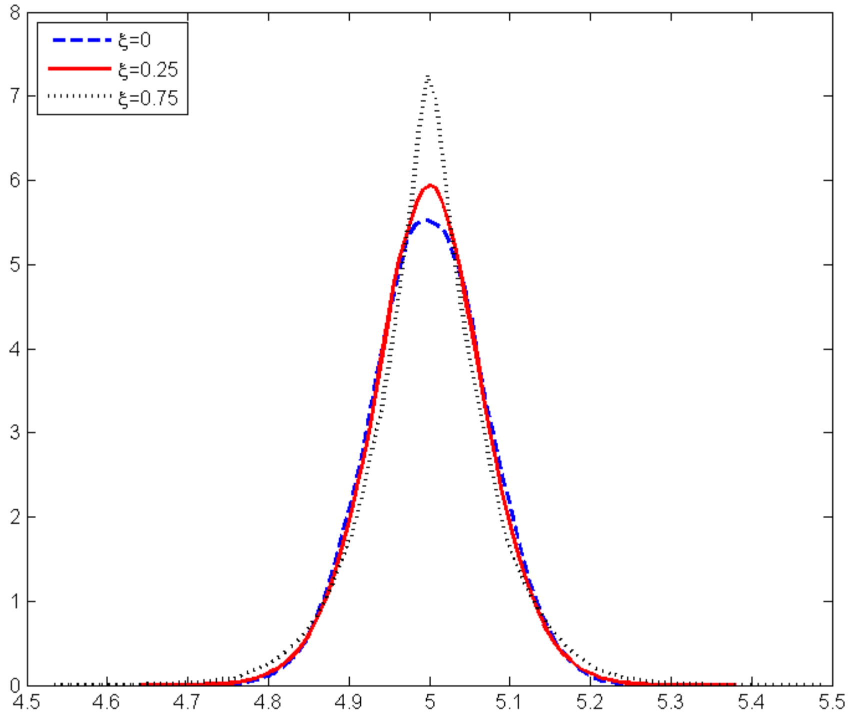

Figure 3.2: Kurtosis variations of the simulated density distribution under different ξ. The setting parameters: $S_0 = 5$, $V_0 = 0.01$, $\mu = 0$, $\kappa = 2$, $\Delta t = 0.02$, $\theta = 0.01$, $\rho = 0$

The volatility of variance parameter ξ controls the kurtosis of a distribution. This is shown in Figure 3.2, where the density distribution under different values of ξ is plotted. The parameter setting is: $S_0 = 5$, $V_0 = 0.01$, $\mu = 0$, $\kappa = 2$, $\Delta t = 0.02$, $\theta = 0.01$, $\rho = 0$. When ξ is 0, the volatility of the variance is zero and the log-return of assets follows a Normal distribution. While a value of ξ different from zero implies a more volatile behavior of the variance process, thus extreme movements of asset prices are more likely to appear, resulting in a higher kurtosis and fatter tails.

Now suppose that $f(t, S_t)$ is a twice-differentiable scalar function of two real variables t and S_t. Applying Itô's lemma with $f(t, S_t) = lnS_t$ to the SDE of asset price (equation 3.23) yield:[22]

$$ln(\frac{S_t}{S_0}) = (\mu - \frac{1}{2}V_t)t + \sqrt{V_t}W_t^S \tag{3.26}$$

We can rewrite the dynamics of log-price process lnS_t as $X_t = lnS_t$ in accordance with the literature (see Shreve, 2004), thus:

$$dX_t = (\mu - \frac{1}{2}V_t)dt + \sqrt{V_t}dW_t^S \tag{3.27}$$

The application of Itô's formula implies that $\{X_t, t \geq 0\}$ satisfies the new derived SDE (3.27). As a result, we are able to reformulate the bivariate system of Heston SDEs as follows:

$$dX_t = (\mu - \frac{1}{2}V_t)dt + \sqrt{V_t}dW_t^S \tag{3.28}$$

$$dV_t = \kappa(\theta - V_t)d_t + \xi\sqrt{V_t}dW_t^V \tag{3.29}$$

The SDE (3.28) is often used instead of (3.23) because the Heston characteristic function is derived using a log-price process X_t rather than the real price process S_t. In general, the characteristic function of a random variable X_t that is defined by a SDE can be obtained without actually solving the SDE. When the PDE is solvable, we are able to calculate the corresponding characteristic function using the Feynman-Kac representation theorem.

3.3.2 Partial differential equation

Given the SDEs (3.28) and (3.29), we introduce the following theorem to determine a pair of PDE via the multi-dimensional version of Itô's formula.

Theorem 3.4. (Partial differential equation). *Let* $\{X_t, t \geq 0\}$ *and* $\{V_t, t \geq 0\}$ *are*

[22] Transform equation (3.26) by taking the exponential e on both sides, then the formula becomes: $S_t = S_0\, exp\left((\mu - \frac{1}{2}V_t)t + \sqrt{V_t}W_t^S\right)$, implying that S_t is log-Normal distributed. See Appendix E.1 for the derivation of Itô's lemma.

two diffusion processes that satisfy the stochastic differential equations given by:

$$dX_t = a_1(t, X_t, V_t)dt + b_1(t, X_t, V_t)dW_t^1$$

$$dV_t = a_2(t, X_t, V_t)dt + b_2(t, X_t, V_t)dW_t^2$$

where $\{W_t^1, t \geq 0\}$ *and* $\{W_t^2, t \geq 0\}$ *are two standard one-dimensional Wiener processes. Let* f *be a twice-differentiable scalar function of its arguments and* $f \in C^{1,2}([0, \infty) \times \mathbb{R})$,[23] *then we have the partial differential equation as:*[24]

$$\begin{aligned} df(t, X_t, V_t) =& \dot{f}(t, X_t, V_t)d_t + f_1(t, X_t, V_t)dX_t + f_2(t, X_t, V_t)dV_t \\ &+ \frac{1}{2}f_{11}(t, X_t, V_t)d\langle X\rangle_t + \frac{1}{2}f_{22}(t, X_t, V_t)d\langle V\rangle_t + f_{12}(t, X_t, V_t)d\langle X, V\rangle_t \end{aligned}$$

where the partial derivatives are mathematically defined as:

$$\dot{f}(t, x, v) = \frac{\partial f(t, x, v)}{\partial t}, \quad f_1(t, x, v) = \frac{\partial f(t, x, v)}{\partial x}, \quad f_2(t, x, v) = \frac{\partial f(t, x, v)}{\partial v}$$

$$f_{11}(t, x, v) = \frac{\partial^2 f(t, x, v)}{\partial x^2}, \quad f_{22}(t, x, v) = \frac{\partial^2 f(t, x, v)}{\partial v^2}, \quad f_{12}(t, x, v) = \frac{\partial^2 f(t, x, v)}{\partial x \partial v}$$

and the computation result of $d\langle X, V\rangle_t$ *is expressed as:*

$$d\langle X, V\rangle_t = dX_t \times dV_t = b_1(t, X_t, V_t)\, b_2(t, X_t, V_t)\, dW_t^1 dW_t^2$$

Remark 3.5. *For the multi-dimensional Itô Processes, the correlation coefficient* ρ_{ij} *between the quadratic variation Wiener processes* W_t^i *and* W_t^j *is expressed as* $E^{\mathbb{P}}[d(W_t^i, W_t^j)] = \rho_{ij}dt$. *Two specific cases should be noticed that when* $\rho_{ij} = 1$, *then* $E^{\mathbb{P}}[dW_t^i dW_t^j] = dt$, *and* $E^{\mathbb{P}}[dW_t^i dW_t^j] = 0$ *if* $\rho_{ij} = 0$.

Given the Theorem 3.4 and Remark 3.5, quoting same notations and keeping

[23] $C^{1,2}([0, \infty) \times \mathbb{R})$ is defined to be the class of functions that are continuously differentiable in the first argument and are twice continuously differentiable in the second argument (See Chang, 2004).

[24] Here employing Newton's notation for differentiation where the independent variable is time t. For example, $\dot{f}()$ represents the first-order derivative of dependent variable $f(t, X_t, V_t)$ with respect to time t.

the functional form consistent with the Heston (1993) model, we can formulate the PDE as follows:[25]

$$df(t,X_t,V_t) = \frac{\partial f(t,X_t,V_t)}{\partial X_t}\sqrt{V_t}dW_t^S + \frac{\partial f(t,X_t,V_t)}{\partial V_t}\xi\sqrt{V_t}dW_t^V + (\mathcal{A}f)(t,X_t,V_t)dt \tag{3.30}$$

where $\mathcal{A}$ is the differential generator defined as:[26]

$$\begin{aligned}(\mathcal{A}f)(t,X_t,V_t) =& \frac{\partial f(t,X_t,V_t)}{\partial t} + (\mu + pV_t)\frac{\partial f(t,X_t,V_t)}{\partial X_t} + (a - qV_t)\frac{\partial f(t,X_t,V_t)}{\partial V_t} \\ &+ \frac{V_t}{2}\frac{\partial^2 f(t,X_t,V_t)}{\partial X_t^2} + \frac{\xi^2 V_t}{2}\frac{\partial^2 f(t,X_t,V_t)}{\partial V_t^2} + \xi V_t\rho\frac{\partial^2 f(t,X_t,V_t)}{\partial V_t \partial X_t}\end{aligned} \tag{3.31}$$

with parameter p, a and q in our case is given by: $p = -\frac{1}{2}$, $a = \kappa\theta$ and $q = \kappa$.

3.3.3 Closed-form extended characteristic function of Heston

As we have mentioned, the modern approach to determine the characteristic function is not to solve a SDE, but is simply to obtain the solution of the PDE. This method has been successfully applied by Christoffersen et al. (2009) in the context of double Heston model via the multi-dimensional Feynman-Kac theorem. We present the univariate Feynman-Kac theorem as follows.

Theorem 3.6. (Feynman-Kac Representation Theorem). *Suppose that* $u \in C^{1,2}(\mathbb{R})$ *and that the diffusion* $\{X_t, t \geq 0\}$ *is defined by an SDE:*

$$dX_t = \mu(t,X_t)dt + \sigma(t,X_t)dW_t$$

then, the unique bounded function $f(t,X_t)$ *which satisfies the partial differential equation is given by:*

$$(\mathcal{A}f)(t,X_t) = \dot{f}(t,X_t) + \mu(t,X_t)f'(t,X_t) + \frac{1}{2}\sigma^2(t,X_t)f''(t,x) = 0,\ 0 \leq t \leq T,\ X_t \in \mathbb{R}$$

[25]See Appendix E.2 for derivation.

[26]Equation (3.30) establishes a connection between stochastic calculus and the partial differential equations. As Itô's integral is a martingale, then function $f(t,X_t,V_t)$ is a martingale if the coefficient of term dt is 0. Namely, the PDE $(\mathcal{A}f)(t,X_t,V_t) = 0$. This idea is indicated and developed by the Feynman-Kac representation theorem.

subject to the terminal condition:

$$f(T,x) = u(x) = e^{i\phi x}, \quad x \in \mathbb{R}$$

has a unique bounded solution:[27]

$$f(t,x) = \mathbb{E}\left[u(X_T)|X_t = x\right] = \mathbb{E}\left[e^{i\phi X_T}|X_t = x\right]$$

If we jointly consider the two diffusions $\{X_t, t \geq 0\}$ and $\{V_t, t \geq 0\}$ in Heston model (Equations 3.28 and 3.29), then we can extend Theorem 3.6 to a multi-dimensional form. The ECF of X_t based on the multi-dimensional Feynman-Kac representation theorem can be written as:

$$\psi_X(z) = f(t,x,v) = \mathbb{E}\left[e^{izX_T}|X_t = x, V_t = v\right] \tag{3.32}$$

Guided by the equation of terminal condition and studies from Heston (1993), the functional form of ECF is formulated as:

$$f(t,x,v) = E(e^{izX_T}) = f(t,X_t,V_t) = e^{C(\tau)+D(\tau)V_t+izX_t} \tag{3.33}$$

where $C(\tau)$ and $D(\tau)$ are coefficients in a form of function and satisfy the Riccati differential equation, $\tau = T - t$ represents the holding time of an asset or the time to maturity in option pricings. Note that at maturity $\tau = 0$, $X_T = lnS_T$ is known, hence we have $f(t,x,v) = e^{izX_T}$. This indicate that the two initial conditions $D(0) = 0$ and $C(0) = 0$ are satisfied at $\tau = 0$ (also see Rieck, 2007 and Rouah, 2013).

Differentiate the ECF (3.33) and substitute to $(\mathcal{A}f)(t,X_t,V_t) = 0$. Then, the corresponding results can lead us to obtain two reduced ordinary differential

[27] Note that the explicit solution equation indicated by the Feynman-Kac representation theorem conveys a information that the solution u(x) to the PDE $(\mathcal{A}f)(t,X_t)$ can be written as a form of conditional expectation, which provides the same way as defining the characteristic function.

equations (ODEs):[28]

$$\frac{\partial D(\tau)}{\partial \tau} + iz\xi\rho D(\tau) + \frac{1}{2}\xi^2 D^2(\tau) - \frac{1}{2}z^2 - qD(\tau) + izp = 0 \tag{3.34}$$

$$\frac{\partial C(\tau)}{\partial \tau} + iz\mu + aD(\tau) = 0 \tag{3.35}$$

where $p = -\frac{1}{2}$, $a = k\theta$ and $q = \kappa$. The first equation of the ODEs is a Riccati equation for coefficient $D(\tau)$, while the second one for coefficient $C(\tau)$ is a function of $D(\tau)$ which can be solved using integration methods as long as the solution of $D(\tau)$ is obtained first.[29]

Given the derived analytical solution of parameters $D(\tau)$ and $C(\tau)$ (equation E.15 and E.19), the ECF of Heston model can be eventually formulated as:

$$f(t, x, v) = E(e^{izX_T}) = f(t, X_t, V_t) = e^{C(\tau)+D(\tau)V_t+izX_t} \tag{3.36}$$

where

$$C(\tau) = iz\mu\tau + \frac{a}{\xi^2}\left[(q - \rho\xi iz + d)\tau - 2ln\left(\frac{1 - ge^{d\tau}}{1 - g}\right)\right] \tag{3.37}$$

$$D(\tau) = \frac{q - \rho\xi iz + d}{\xi^2}\left(\frac{1 - e^{d\tau}}{1 - ge^{d\tau}}\right) \tag{3.38}$$

and

$$g = \frac{q - \rho\xi iz + d}{q - \rho\xi iz - d} \tag{3.39}$$

$$d = \sqrt{(\rho\xi iz - q)^2 - \xi^2(2piz - z^2)} \tag{3.40}$$

$$p = -\frac{1}{2}, \quad a = \kappa\theta, \quad q = \kappa \tag{3.41}$$

[28]See Appendix E.3 for derivation.

[29]Appendix E.4 demonstrates the solution of $C(\tau)$ and $D(\tau)$.

3.4 Numerical illustration

In this section, we present and evaluate the experimental results of the proposed risk models based on the Heston ECF. The calculation of VaR and CVaR estimates requires the evaluation of an integral. Hence we first reformulate the risk models using the ECF of Heston and present the trapezoidal integration rule for numerical analysis. Subsequently, the convergence of the integrand of function T_v^* and H_v^* is discussed to ensure efficient approximations. In the end, we examine the impact of several key ingredients on the approximation of VaR and CVaR.

3.4.1 Trapezoidal integration rule for the reformulated VaR and CVaR model

The characteristic function is the prerequisite for evaluating the derived VaR and CVaR model (equatioin 3.16). Therefore, we substitute the extended form of Heston characteristic function (equation 3.36) into equation (3.16), so that the two risk measures can be reformulated as follows:

$$\begin{cases} \alpha = \dfrac{\mathbf{Re}\, T_v^*(VaR, v)}{\pi} \\ CVaR = VaR + \dfrac{\mathbf{Re}\, H_v^*(VaR, v)}{\mathbf{Re}\, T_v^*(VaR, v)} \end{cases} \tag{3.42}$$

where

$$T_v^*(VaR, v) = e^{-vVaR} \left[\int_0^{+\infty} \frac{e^{C(\tau)+D(\tau)V_t+(iw-v)X_t}}{v - iw} \, e^{iwVaR} \, dw \right] \tag{3.43}$$

$$H_v^*(VaR, v) = e^{-vVaR} \left[\int_0^{+\infty} \frac{e^{C(\tau)+D(\tau)V_t+(iw-v)X_t}}{(v - iw)^2} \, e^{iwVaR} \, dw \right] \tag{3.44}$$

and

$$C(\tau) = iz\mu\tau + \frac{a}{\xi^2} \left[(q - \rho\xi iz + d)\tau - 2ln \left(\frac{1 - ge^{d\tau}}{1 - g} \right) \right] \tag{3.45}$$

$$D(\tau) = \frac{q - \rho\xi iz + d}{\xi^2} \left(\frac{1 - e^{d\tau}}{1 - ge^{d\tau}} \right) \tag{3.46}$$

and

$$g = \frac{q - \rho\xi iz + d}{q - \rho\xi iz - d} \tag{3.47}$$

$$d = \sqrt{(q - iz\xi\rho)^2 - \xi^2(2izp - z^2)} \tag{3.48}$$

$$p = -\frac{1}{2}, \quad a = \kappa\theta, \quad q = \kappa \tag{3.49}$$

It is worth noting that the risk model consists of two functions T_v^* and H_v^*, which require the evaluation of an integral. In general, the antiderivative of the integrand can be found by applying the Fundamental Theorem of Calculus, according to which the integral values are obtainable by the difference between the antiderivative values evaluated at the endpoints of the interval. However, the antiderivative of the integrand (namely ECF) for the Heston model is analytically intractable, hence the integrals must be approximated using a numerical approach. This resembles most of the option pricing models that we have encountered by inverting the characteristic function to evaluate the option call prices (e.g. Stein and Stein, 1991; Rchöbel and Zhu, 1999; Lewis, 2000 and Attari, 2004), or the models that produce the probabilities for a given inverted characteristic function. The integral in functions T_v^* and H_v^* can be approximated by a quadrature and we will use the numerical approximation approach called the trapezoidal integration rule, to evaluate functions T_v^* and H_v^*.

Conventionally, the quadrature method approximates an integral as the weighted sum of functional values evaluated at discrete points over the integration domain $[\Phi_{min}, \Phi_{max}]$:

$$\int_{\Phi_{min}}^{\Phi_{max}} \Psi(x)dx \approx \sum_{j=1}^{N} \eta_i \Psi(x_j) \qquad j = 1, 2, ..., N \tag{3.50}$$

where $(x_1, ..., x_N)$ are called abscissas and $(\eta_1, ..., \eta_N)$ are the weights for each point. For the trapezoidal integration rule, however, the integral in (3.50) is approximated as the sum of a number of trapezoids and the integral interval $[\Phi_{min}, \Phi_{max}]$ is assumed to be split into equally spaced subintervals using the abscissas $(x_1, ..., x_N)$. This indicates that the size of abscissas needs to be large enough to ensure the accuracy of the approximations and this is particularly important if there are specific regions over $[\Phi_{min}, \Phi_{max}]$ where the function $\Psi(x)$ is substantially oscillatory.

The integration domain, as indicated by functions T_v^* and H_v^*, ranges from $(0, \infty)$ and also the characteristic function of Heston is involved in the integrand of T_v^* and H_v^*. This implies the possibility of the integrand being highly oscillatory over this integration range. Rouah (2013) pointed out that the Heston integrand has a fair amount of oscillation especially when τ is small. Indeed, in some scenarios, it could pose a numerical problem for the integration approximation if the integrand was not well-behaved. In other scenarios, the integrand might be not oscillating anywhere and could quickly decay to 0, not causing any numerical problems. We then investigate how large the upper limit (Φ_{max}) of the numerical integration in terms of function T_v^* and H_v^* need to be in order to ensure that the integrands are not oscillating.[30]

The choice of upper limit Φ_{max} means to truncate the interval $(0, \infty)$. Suppose we choose a point $\overline{\eta}_N$ which is large enough to ensure the convergence of the integrand to 0 for both T_v^* and H_v^* function so that the points after $\overline{\eta}_N$ are all negligible. Then, we divide the interval $[0, \overline{\eta}_N]$ into $N-1$ subintervals with equal width $\Delta w = w_{j+1} - w_j$, which yield the grid points $w = \{w_j = j\Delta w, \quad j = 0, 1, ..., N\}$.[31]

If we denote the real part of the integrand of functions T_v^* and H_v^* to be two functions $m(x, w)$ and $n(x, w)$, respectively:

$$m(x, w) = \mathbf{Re}\left[\frac{e^{C(\tau)+D(\tau)V_t+(iw-v)X_t}}{v - iw}\, e^{iwVaR}\right] \tag{3.51}$$

$$n(x, w) = \mathbf{Re}\left[\frac{e^{C(\tau)+D(\tau)V_t+(iw-v)X_t}}{(v - iw)^2}\, e^{iwVaR}\right] \tag{3.52}$$

[30] An alternative to choose the integration domain is by Kahl and Jäckel (2005), who used a closed interval [0, 1] instead of truncating domain $(0, \infty)$. However, their approach is especially appropriate for Gauss-Lobatto quadrature rather than the trapezoidal rule. More details can also refer to Rouah (2013).

[31] Trapezoidal rule use the weight $\eta_1 = \eta_N = \Delta w/2$ and $\eta_j = \Delta w$ with $j = 2, ..., N-1$.

then the height of each trapezoids can be evaluated by the functions of $m(x, w)$ and $n(x, w)$ at each endpoints. As a result, we are able to approximate the integrals of functions T_v^* and H_v^* using the trapezoidal rule as follows:

$$\int_0^{+\infty} m(x, w)dw \approx \frac{1}{2}m_0(x)\Delta w + \sum_{j=1}^{N-1} m_j(x)\Delta w + \frac{1}{2}m_N(x)\Delta w \tag{3.53}$$

$$\int_0^{+\infty} n(x, w)dw \approx \frac{1}{2}n_0(x)\Delta w + \sum_{j=1}^{N-1} n_j(x)\Delta w + \frac{1}{2}n_N(x)\Delta w \tag{3.54}$$

3.4.2 Convergence of the integrand and efficient computation

As prerequisites for the evaluation of VaR and CVaR models, the upper integration limit $\overline{\eta}_N$ and grid space Δw should be cautiously identified in order to receive an accurate approximation, and this is done through investigating the convergence of $m(x, w)$ and $n(x, w)$. The selection of $\overline{\eta}_N$ can be guided by the speed of decay of $m(x, w)$ and $n(x, w)$, while a fine grid of Δw is generally required in order to accurately calculate the density of the tail distribution.[32]

To illustrate this point, we investigate the impact of changing the values of three factors (ξ, VaR and τ) on the integrand $m(x, w)$ and $n(x, w)$, respectively, following the work of Moodley (2005). In addition, for setting the parameter values, we use the Heston estimates of Bormetti et al. (2010), which uses a series of 5000 observations of real daily returns of German DAX 30 Index ranging from 14th of November 1988 to 9th of September 2008. The parameter set Θ is given by: $\kappa = 0.86$, $\theta = 0.0471$, $\mu = 0.1102$, $\rho = -0.17$, $X_0 = 0.1102$, $V_0 = 0.0471$.

Figure 3.3 shows the effect of changing three factors on the integrand of T_v^*. In the graph, we have set the upper integration limit $\overline{\eta}_N = 30$ and use a relatively fine grid size $\Delta w = 0.1$, which generates 300 grid points. In the first scenario, we fix ξ and VaR whilst change τ. We can see that the integrand is sufficiently close to 0 before $\overline{\eta}_N = 10$ although the integrand for shorter τ may oscillate for much longer.

[32] Chourdakis (2008) empirically discussed the selection criterion about the size of grid space, arguing the necessities to employ a large amount of grid points if people are focusing on the tail of a distribution instead of the central part of a distribution.

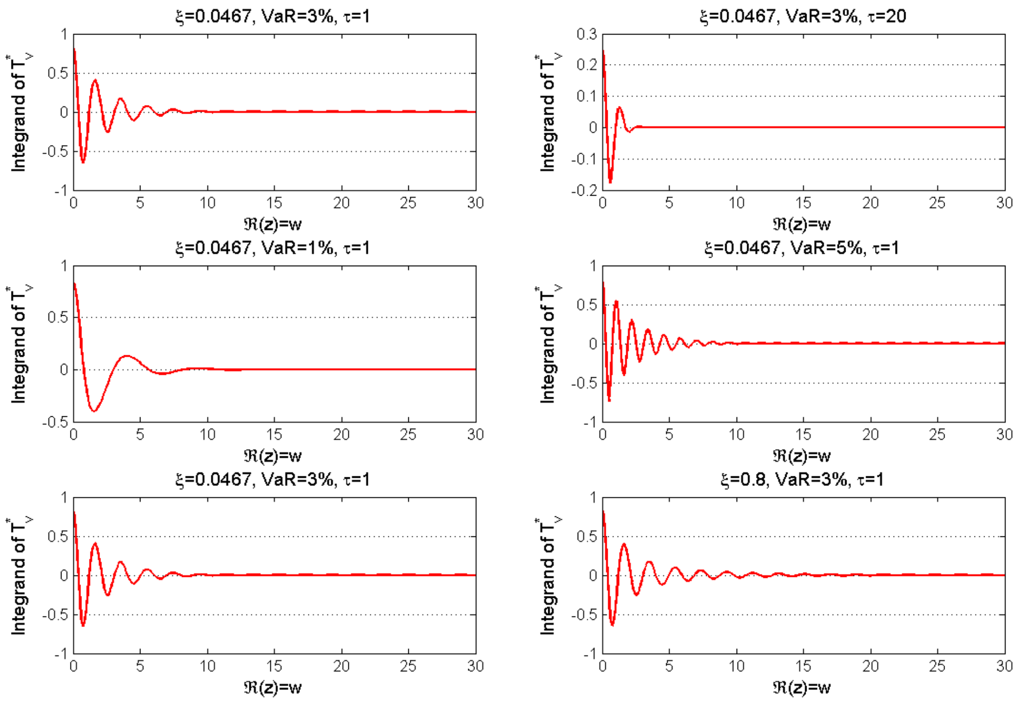

Figure 3.3: Convergence of the integrand of function T_v^* for different values of τ(upper), VaR(middle) and ξ(bottom) with a given parameter set Θ.

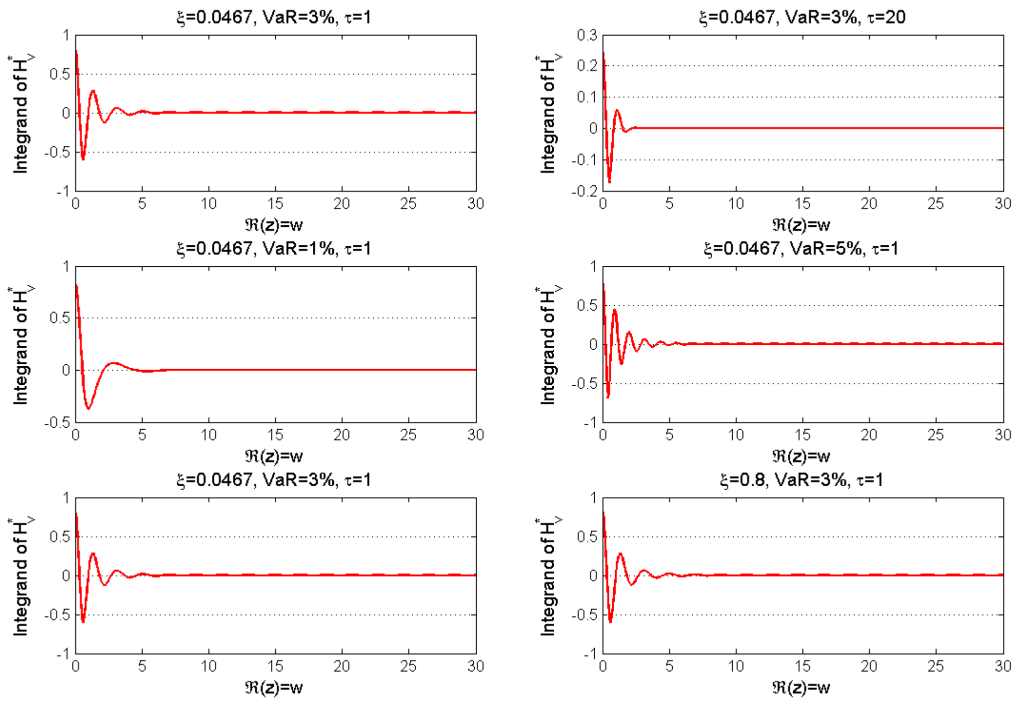

Figure 3.4: Convergence of the integrand of function H_v^* for different values of τ(upper), VaR(middle) and ξ(bottom) with a given parameter set Θ.

This indicates that high and long oscillations are more likely to be associated with shorter τ, which has also been claimed by Rouah (2013). In the second scenario, we fix ξ and τ and investigate the VaR variations. It is easy to see that larger VaR values can cause a steep oscillation of the integrand but they decay very quickly when approach to $\overline{\eta}_N = 10$. This effect is the same as using lower VaR values. In contrast to the quick convergence of integrand for lower ξ, in the last scenario, by fixing VaR and τ, a larger ξ apparently results in a slow convergence of the integrand until when approach to $\overline{\eta}_N = 20$. Note that our findings when varying ξ are consistent with the results of Chourdakis (2008) in terms of revealing the important relationship between kurtosis and the decay of the characteristic function.[33]

Since the calculation of CVaR relies on both T_v^* and H_v^*, the selected $\overline{\eta}_N$ and Δw must ensure the convergence of the two integrands simultaneously. We use the same parameter set Θ for the H_v^* integrand calculation and adopt consistent integral conditions with upper integration bound $\overline{\eta}_N = 30$ and grid size $\Delta w = 0.1$. Figure 3.4 shows the oscillation of the H_v^* integrand affected by ξ, VaR and τ, respectively. In all the three scenarios, we can see that the integrand of H_v^* is highly oscillatory at the beginning but rapidly goes close to 0 far before $\overline{\eta}_N = 10$. Therefore, we conclude that the prudent selection of $\overline{\eta}_N = 30$ and $\Delta w = 0.1$ should be able to ensure the convergence of the integrand of the functions T_v^* and H_v^*, therefore being suitable for the numerical approximations of the VaR and CVaR model.

3.4.3 Numerical results

To evaluate the reformulated VaR and CVaR in equation (3.42) under the trapezoidal integration rule, we use the parameter sets specified in Table 3.1 to produce VaR and CVaR approximations over a range of risk levels α. The target ingredients in Table 3.1 we are exploring are those that could have the most significant impact on the VaR and CVaR approximations. We aim to check how

[33] The results of Chourdakis (2008) indicate that a coarse grid may be sufficient if people are simply focusing on the central part of the Normal distribution. However, a fine grid size, or probably a larger interval is necessary if one is interested in the tail density. As we have argued in previous section that ξ determines the kurtosis of the simulated Heston *p.d.f.*, hence our finding confirms the opposite relationship between kurtosis and the speed of decay of Heston characteristic function.

Table 3.1: Grid size and Heston parameter specification for the VaR and CVaR integral approximations

Targets	w	κ	θ	ξ	ρ	τ	X_0	V_0	μ
w	**(0:0.1:30)** **(0 : 2 : 30)**	0.86	0.0471	0.0467	-0.17	20	0.1102	0.0471	0.1102
κ	(0:0.1:30)	**0.05→0.6**	0.0471	0.0467	-0.17	20	0.1102	0.0471	0.1102
θ	(0:0.1:31)	0.86	**0.1→0.8**	0.0467	-0.17	20	0.1102	0.0471	0.1102
ρ	(0:0.1:32)	0.86	0.0471	0.0467	**-0.95→0.95**	20	0.1102	0.0471	0.1102
τ	(0:0.1:33)	0.86	0.0471	0.0467	-0.17	**1→20**	0.1102	0.0471	0.1102

Note: This table lists the setting parameter values for the numerical integrations. "Targets" refers to the ingredients that could affect the VaR and CVaR approximations. The bold fonts represent the allowable variation range of the interested parameters that we explored. For instance, (0:0.1:30) represents an increment of 0.1 up to 30 and " → " means a successive increase.

these ingredients could impact on the approximations, how sensitive they are to the behavior of VaR and CVaR, and whether the impacts are sustainable.

Impact of the grid size. One main advantage of the VaR and CVaR risk models (equation 3.42) is the speed and efficiency of the algorithm, in the sense that they are able to produce a large number of approximations in one single calculation when setting a series of risk levels α. In the following VaR and CVaR calculation, we employ a grid of one hundred equally spaced α values ranging from 0.1% to 10% with an admission value of $v = 1$. Figure 3.5 first investigates the effect of changing grid size w on the VaR and CVaR approximations. The grid size Δw with $w = (0 : 0.1 : 30)$ has been proved to be large enough and the integrand can be sufficiently decaying to 0 without causing any loss of accuracy in the numerical approximation in equations (3.53) and (3.54). However, when we use a coarse grid size Δw with $w = (0 : 2 : 30)$, the VaR and CVaR approximations apparently become larger. This indicates that caution is called for when setting up the numerical integration procedures using formula (3.42), because a coarse grid size (or the shortage of trapezoids) will inevitably calculate extra trapezoidal areas, thus, reduce the integration accuracy. Conceivably, the VaR and CVaR estimates will gradually move downward with smaller Δw until they approximate to the theoretically "perfectly accurate" values.

Impact of changes in parameters κ, θ, ρ, τ. To investigate the effect of changing the Heston parameter values on the VaR and CVaR estimates for a given sets of α values, we consider four starting values of κ, θ, ρ and τ and then progressively raise their values.[34] Figure 3.6 presents the variation of VaR and CVaR estimates by varying the values of mean-reversion speed parameter κ. We begin with the initial value of κ as 0.05. It is clear to see that if we increase the value of κ, the VaR and CVaR estimates move downwards simultaneously until, for example, $\kappa = 0.6$. After $\kappa = 0.6$, the decline slows down and we find that the line of CVaR at $\kappa = 1$ completely coincides with the line of VaR for $\kappa = 0.05$. This result implies a negative relationship between the speed of mean-reversion and the approximation of the two risk measures.

Figure 3.7 shows the results of VaR and CVaR variations along with changing the long-term variance θ. By starting at $\theta = 0.1$ and progressively increasing the value of θ up to 0.8, we can see that the VaR and CVaR estimates demonstrate an upward trend. However, our experiment indicates that the magnitude of growth for VaR is much lower than the one for CVaR. In other words, a small amount of increment of θ can push the CVaR estimates remarkably higher. Overall, we conclude that the parameter θ can positively affect both the two risk measures, but that this effect is more obvious for the CVaR estimates.

In Figure 3.8, we examine the impact of changing values for the correlation coefficient ρ on the VaR and CVaR estimates over a sequence of α levels. In real financial markets, there is generally a negative correlation between asset returns and variance. We thus first set $\rho = -0.95$ as a benchmark and then change the value of ρ from -0.95 to 0 and from 0 to 0.95. The results show that the variations of VaR and CVaR are consistently towards one direction for $\rho \in [-0.95, 0.95]$, both the two lines move downward stably with the increase of ρ and there are no radical movements during the selected interval. As a result, we can conclude that there is a negative correlation between ρ and the two risk measures.

[34] In our experiment, we find that only these four parameters could have an obvious impact on the estimates of VaR and CVaR in a trapezoidal integration scheme.

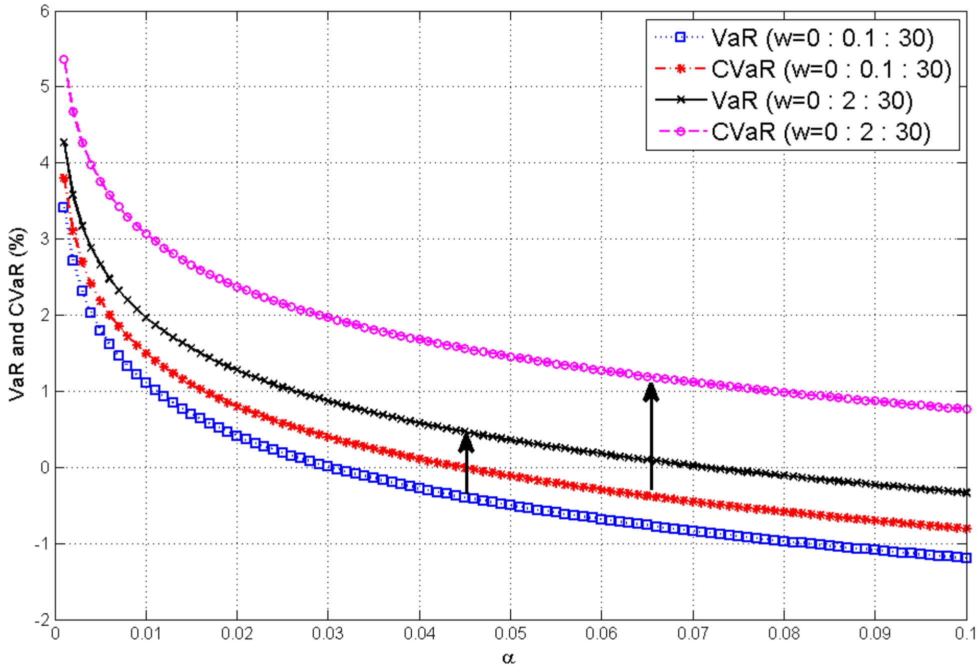

Figure 3.5: VaR and CVaR approximations using trapezoidal rule against a hundred equally spaced α level ranging from 0.1% to 10% at $\{\overline{\eta_N} = 30, \Delta w = 0.1\}$ and $\{\overline{\eta_N} = 30, \Delta w = 2\}$

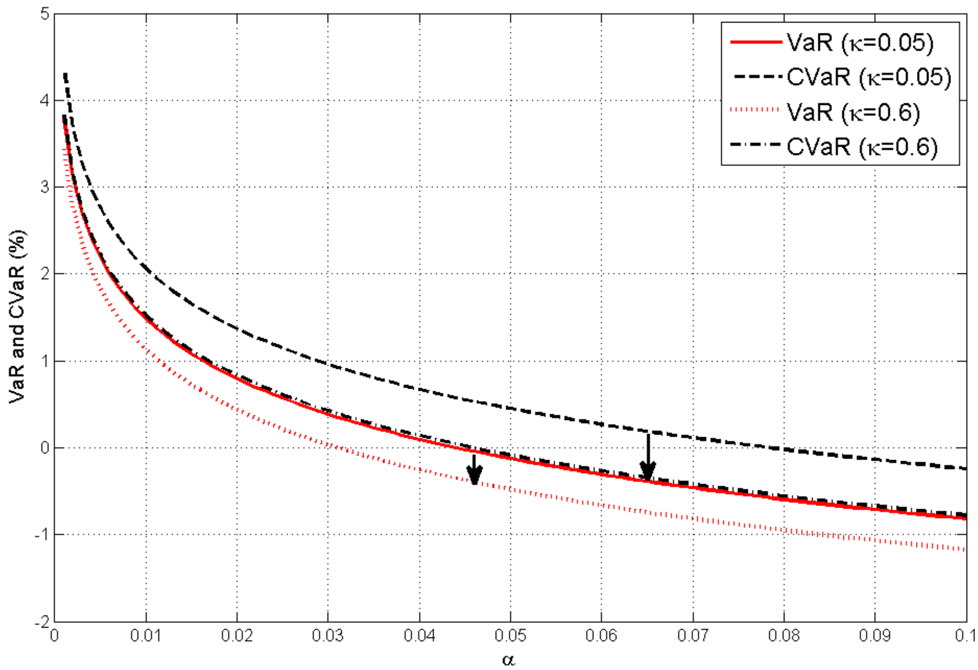

Figure 3.6: VaR and CVaR approximations using trapezoidal rule against a hundred equally spaced α level ranging from 0.1% to 10% at different values for the mean reversion speed $\kappa = 0.05, 0.6$

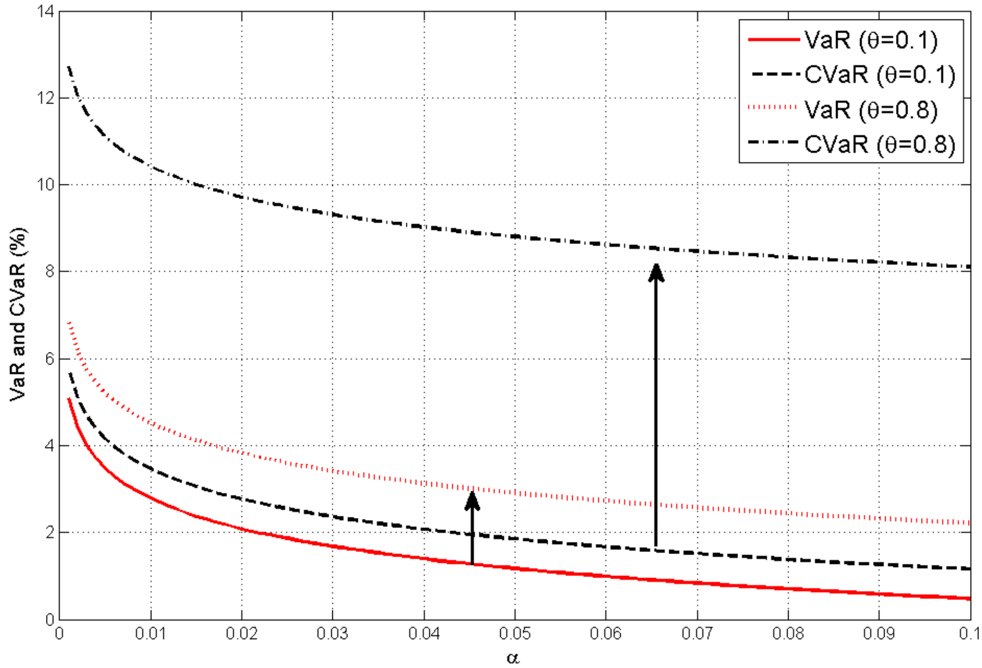

Figure 3.7: VaR and CVaR estimates using trapezoidal rule against a hundred equally spaced α level ranging from 0.1% to 10% at different values for the long-term variance $\theta = 0.1, 0.8$

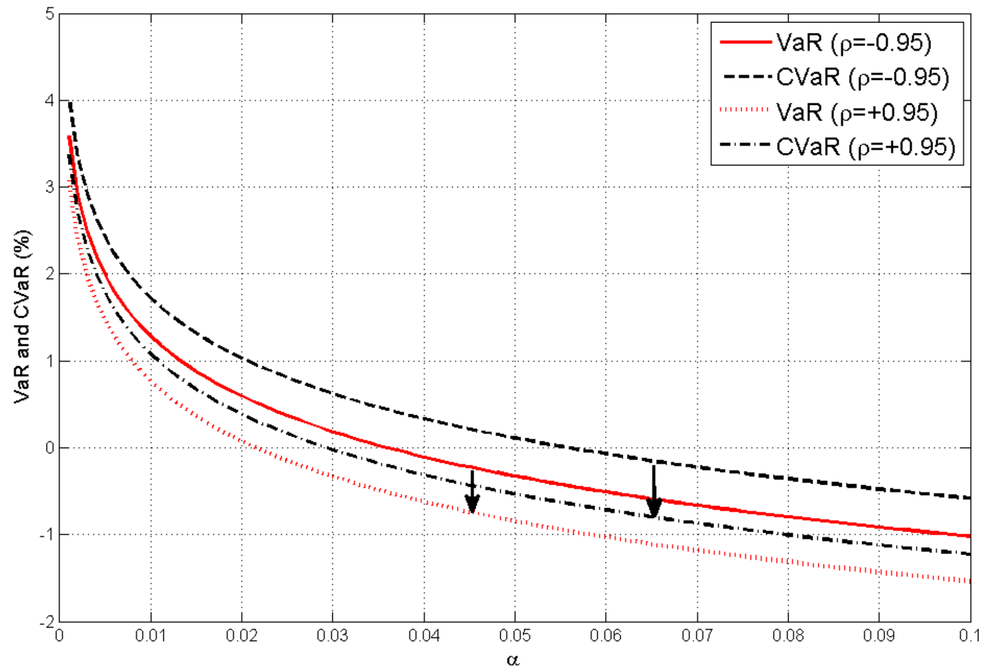

Figure 3.8: VaR and CVaR approximations using trapezoidal rule against a hundred equally spaced α level ranging from 0.1% to 10% at different values for the correlation $\rho = -0.95, +0.95$

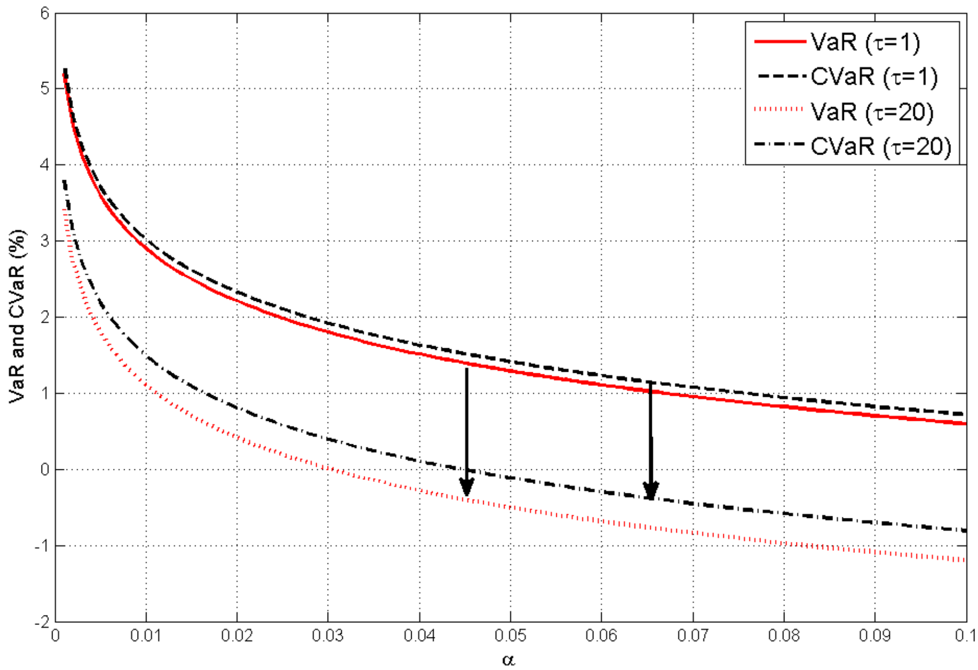

Figure 3.9: VaR and CVaR approximations using trapezoidal rule against a hundred equally spaced α level ranging from 0.1% to 10% at different values for the asset holding period $\tau = 1, 20$

Last, to study the impact of maturity τ, we adjust the maturity range from 1 to 20. Similarly to the findings of κ and ρ, the effect of τ on both the VaR and CVaR estimates are negative. As indicated by Figure 3.9, we can see that VaR and CVaR could estimate a relatively smaller risk value when the maturity becomes longer. This result is in line with our previous findings in Figures 3.3 and 3.4, where we observed that a short maturity is generally associated with higher and longer oscillations of the integrand of T_v^* and H_v^*. Combining the outcomes of all four cases, we can find that the CVaR estimates are comparatively larger than the estimates of VaR, which is reasonable as CVaR calculates the average losses that greater than VaR. To conclude, the mean-reversion speed κ, correlation factor ρ and maturity τ could negatively affect the estimates of VaR and CVaR in our model, while the long-term variance parameter θ is the only one that has a positive relationship with the VaR and CVaR estimates in the Heston model.

3.5 Conclusion

This paper develops a general framework by using the generalized Fourier transform approach in the context of risk management. We analytically build the VaR and CVaR formula with the introduction of two deterministic functions T_v and H_v by using only the closed-form solution for the characteristic function that describes the distribution of financial data. We also present alternative functions TT_v and HH_v when taking the real part of T_v and H_v into further consideration by applying Euler's translation. The proposed VaR and CVaR model extends risk analysis to those non-Gaussian models in a Fourier space, readily applicable to a number of distributions whose $p.d.f.$ is not analytically known.

In our application, we assume asset prices follow the Heston dynamics, which has become the most popular stochastic volatility model for option pricing but remains inadequate in the context of risk management. An important implication of the Heston model is that it allows a number of different non-Gaussian distributions, with correlation parameter ρ controlling the skewness of tail density and volatility of volatility parameter ξ affecting the kurtosis of the distribution. The inherent properties such as non-Gaussian distribution, mean-reversion and volatility persistence ensure the practicality for our VaR and CVaR analysis. The calculation of our VaR and CVaR involves the evaluation of integrals, which is analogous to most option pricing models we have encountered (see, e.g., Heston, 1993; Carr and Madan, 1999; Lewis, 2001). Since the anti-derivative of the integrand for Heston does not exist, the VaR and CVaR values must be approximated numerically. In our analysis a quadrature is adopted, namely the trapezoidal integration rule.

The integrand in our deterministic functions T_v and H_v involves the Heston characteristic function, which can sometimes be highly oscillatory (see Rouah, 2013). Thus, given a set of initial parameter values by employing estimates from Bormetti et al. (2010), we investigate the convergence of the integrands with a selection of appropriate upper integration bounds and grid space in order to ensure that the integrands can sufficiently decay to 0 so as to not cause a loss of accuracy for VaR and CVaR approximations. We find that the integral range of (0:0.1:30) is prudent and desirable in our case.

In addition, we investigate the impact of the grid size on VaR and CVaR approximations. Our results indicate that we will potentially have to carry out numerical integration with a fine grid because a coarse grid size would inevitably result in an overestimation of VaR and CVaR values. Finally, to provide answers to the key questions related to the determinants of VaR and CVaR approximations in the Heston scheme, we explore a series of potential parameters. Our results show that an increase of mean-reversion speed κ, correlation factor ρ and maturity τ can inversely affect VaR and CVaR approximations, for a sequence of risk levels α. The changing trajectory of the risk estimates is stable and does not create any significant jumps. In contrast, we find that the long-term variance θ is the only parameter that could positively impact the VaR and CVaR values and the influence on CVaR is remarkably higher than that on VaR.

Chapter 4

VaR and CVaR Estimation for Oil Prices via SV-ALD Model: A Bayesian Approach Using Scale Mixture of Uniform Distribution

4.1 Introduction

The world crude oil markets have been quite volatile and risky in the past few decades due to the large fluctuations of oil prices, which have become a principal concern for oil suppliers, oil consumers, relevant firms and governments. In addition, as a primary source of energy in the power industry, industrial production and transportation, volatile oil prices may lead to cost uncertainties for other markets, thus extensively affecting the development of the economy. A large number of studies have shown that oil price fluctuations could have considerable impact on economic activities. Papapetrou (2001) argue that the variability of oil prices plays a critical role in affecting real economic activity and employment. Lardic and Mignon (2008) explore the long-term relationship between oil prices and GDP, and find evidence that aggregate economic activity seems to be retarded particularly when oil prices increase. The asymmetry phenomenon is found in both the U.S. and European countries. Consequently, quantifying and managing the risks inherent to the volatility of oil prices has become critical for both researchers and energy market participants.

The Value at Risk (VaR) measure, which was first proposed by J.P. Morgan in the RiskMetrics model in 1994, has been developed as one of the most popular approaches in financial markets to manage market risk. VaR defines the maximum amount of loss of an investment over a given period of time at a specific confidence level. It answers the question as to how much an investor can lose for a given tolerance level over a certain time horizon. Although VaR is recommended by Basel II and III and has been widely adopted by financial institutions, it has been challenged by the Bank of International Settlements (BIS) Committee, who pointed out that VaR cannot measure market risk as it fails to consider the extreme tail events of a return distribution (see, Chen et al., 2012). In addition, Artzner et al. (1999) argue that VaR does not meet the requirements of sub-additivity and thus is not a coherent measure. As an alternative, they proposed a conservative, but more coherent measure, called Conditional VaR at risk (CVaR) or expected shortfall (ES), which considers the average loss that exceeding the VaR threshold. Given all these factors, in this paper, both measures are used to quantify oil return risks. Additionally, similarly to studies on financial markets, existing literatures using VaR and CVaR to measure oil risks generally focus on the scenario of declining oil price (i.e. downside risk). However, the oil market has its own traits which are quite different from those of financial assets. When oil prices fall due to sudden negative news, oil exporting countries or oil producers would inevitably incur losses while oil consumers would benefit from those negative extreme events. On the other hand, if oil prices rise suddenly, oil consumers might have to pay more to compensate this rising risk. Therefore, in comparison to risk measurement in financial markets, we quantify VaR and CVaR by considering both shocks to oil supply and oil demand.[1]

A parametric VaR/CVaR estimation relies heavily on the estimation of price volatility. In order to accurately measure volatility, Engle (1982) proposed the

[1]Note that we use the terminology "oil supply and oil demand" in this paper which refers to the long (or downside risk) and short (or upside risk) trading positions, respectively.

autoregressive conditional heteroscedasticity (ARCH) model based on the clustering and long-memory features of assets returns. Later, Bollerslev (1986) proposed the generalized ARCH (GARCH) model, which has been extensively used to model the time-varying oil price volatility (see, e.g., Sadorsky, 1999; Kang et al., 2009; Wei et al., 2010; Lux et al., 2016).[2] Alternatively, a growing number of empirical papers have adopted the stochastic volatility (SV) model, first proposed by Taylor (1994) to capture volatility features, where the unobserved volatility is treated as a latent variable that follows a stochastic process (see, e.g., Regnier, 2007; Vo, 2009; Vo, 2011; Brooks and Prokopczuk, 2013; Wichitaksorn et al., 2015). Debate subsequently arose as to which of these is more appropriate in modelling time-varying volatility. Hence, the recent literature has devoted extensive attention to studying the ability of the GARCH-type models and SV models to fit the volatility of the data. Wei (2012) explores the forecasting performance of volatility in the fuel oil futures market in China, and finds that the SV model performs better than many linear and non-linear GARCH-type models in capturing volatility features. Chan and Grant (2016) compare various GARCH models to a number of stochastic volatility models using nine series of energy prices, indicating that stochastic volatility models are almost always superior to their counterpart GARCH models when modelling time-varying volatility. Comparatively speaking, the SV model, which defines conditional variance as a latent stochastic process, is more capable of providing flexibility and practicality than GARCH-type models.[3]

However, a main challenge for the SV model is the more statistically and computationally demanding implementation. Unlike GARCH-type models where conditional volatility that can be easily estimated, the likelihood function of volatility in the SV model has no closed-form expression and thus parameters cannot be directly estimated using a classical approach, i.e. Maximum likelihood estimation. Hence, traditional model comparison criteria that are readily available in GARCH models, such as AIC and BIC, would be inapplicable in SV model

[2] GARCH model defines the conditional variance of assets returns as a deterministic function of model parameters and historical data. For instance, a GARCH(1,1) model is formulated as: return equation: $y_t = ax_t + e_t$; volatility equation: $Var(e_t) = h_t = b + b_1 e_{t-1}^2 + b_2 h_{t-1}$ where $e_t \sim N(0, \sigma_t^2)$.

[3] A comprehensive review of SV and GARCH model refers to Shephard (2005).

comparisons. Recently, a simulation-based technique, the Bayesian Markov Chain Monte Carlo (MCMC) approach, has been developed rapidly and is used extensively for estimating SV models. Therefore, considering the special advantage of Bayesian MCMC in dealing with latent variables, in this paper, the SV model is employed to model oil prices volatilities and interested parameters are estimated in the framework of the Bayesian MCMC algorithm.

It has long been recognized that the distribution of financial asset returns are rarely normally distributed. A Gaussian assumption for return errors in the SV model may fail to account for the skewness, heavy tails and leptokurtic features in financial returns. As a result, the simple SV model with normally distributed error has been widely extended to heavy-tailed errors in the literature. Student's t distribution is often used in return errors of the SV model (see, e.g., Chi et al., 2002; Omori et al., 2007; Asai, 2008; Nakajima and Omori, 2009). In addition, Tsiotas (2012) and Abanto-Valle et al. (2015) extend previous studies by employing Skew Student's t distribution in the SV model, Joanna et al. (2013) employ generalized-t distribution with SV to model financial time series, while Nakajima and Omori (2012) adopt generalized hyperbolic (GH) Skew Student's t error in the SV model.

In this paper, we extend the Bayesian analysis of the SV model by allowing the return error to follow an Asymmetric Laplace distribution (ALD), the density of which was proposed by Kotz et al. (2001). ALD has been successfully applied in many fields of finance and economics, i.e. modelling return skewness and leptokurtic features of financial data, incorporating Laplace noise in autoregressive moving average (ARMA) model, VaR estimation and portfolio optimizations, etc. However, to the best of our knowledge, the application of ALD within the SV framework has not been used, with one exception of Nuttanan et al. (2014) who have explored return asymmetry and quantiles using the SV model with incorporation of Asymmetric Laplace error in the return equation. Instead of employing the $p.d.f.$ proposed by Yu and Zhang (2005) as in Nuttanan et al. (2014), our paper focuses on extending the application of ALD proposed by Kotz et al. (2001) to the SV model.[4] In addition, we derive the closed form expression of VaR and CVaR based on SV-ALD model to measure oil price risks from the perspective of both supply and demand.

It is challenging to estimate volatility within the SV-ALD model because in the context of Bayesian inference via MCMC algorithm, the full conditional posterior distributions are of non-closed form. The data augmentation technique to express the *p.d.f.* of the heavy-tailed distribution using a scale mixture form has been extensively used. Meyer and Yu (2000), Abanto-Valle et al. (2010) and Wang et al. (2011) have studied Student t distribution in SV using a scale mixture of Normal (SMN) representation to replace its density. Choy and Chan (2008) propose the scale mixture of uniform (SMU) representation to substitute generalized t density function, so that Bayesian MCMC algorithm can be implemented and ease the computational burden. Nuttanan et al. (2014) incorporate Yu and Zhang's (2005) ALD to SV by transforming the scaled density to an SMU form. In this study, based on the constructed SV-ALD model, a new SMU representation is proposed for ALD error to solve the parameter estimation difficulties in the Bayesian MCMC framework. The main advantage of the proposed SMU is to ensure that some of the full conditional posterior distributions are obtainable in a standard form, hence facilitating an efficient Gibbs sampling algorithm for the parameter estimation.[5]

In summary, this paper studies a new parametric approach to estimate VaR and CVaR for crude oil prices from both perspective of oil supply and demand. We model the time-varying oil price volatilities using classical discrete SV model with the extension of adopting ALD (Kotz et al., 2001) for the return error distribution

[4]According to the *p.d.f.* of ALD studied by Zhang and Yu (2005), there are three parameters: skew p, scale σ and location u, where $0 < p < 1$, $\sigma > 0$ and $-\infty < u < +\infty$. When $p = \frac{1}{2}$, it becomes a symmetric distribution called Laplace double exponential distribution. ALD of Kotz et al. (2001) is another asymmetric extension of double exponential distribution by splitting it where two random variables are independent but no longer identically distributed. Further empirical explorations of ALD (Kotz et al., 2001) have not been conducted, especially in SV framework. More details can refer to Chapter 3 of Kotz et al. (2001).

[5]Choy et al. (2009) and Wang et al. (2013) study SMU for Student t distribution and generalized t distribution, respectively, demonstrating that Gibbs sampler for Bayesian MCMC computation can be substantially simplified when using SMU representation. In other words, most of the full conditional posterior distributions are of standard forms, thus satisfying the requirements for Gibbs sampling scheme.

in order to take account of the heavy-tailed and leptokurtic features. Hence we build the SV-ALD model and derive the corresponding analytic expressions of VaR and CVaR for oil supply and demand. To estimate the parameters, a Bayesian approach is employed to sample model parameters and unobserved volatilities from their posterior distributions via the simulation-based MCMC approach through Gibbs sampling algorithm. In order to overcome the difficulties to realize this procedure, we use a data augmentation technique by proposing a new SMU representation for the scaled AL density, which can simplify the algorithm of Bayesian MCMC. In practice, a model comparison study from Bayesian statistical perspectives is performed between the SV-ALD model and SV Normal model to test model fitting abilities using the oil return series. We then model the time-varying oil price volatilities in the WTI and Brent markets based on the SV-ALD model, and investigate the market risk of oil supply and demand. In addition, we conduct an accuracy test by backtesting VaR and CVaR violations. Considering the complexity of CVaR backtesting in comparison to VaR, an equal-quantile based method suggested by Kerkhol and Melenverg (2004) is used by confirming a specific quantile level that CVaR of ALD falls at. We prove that this quantile level depends only on the prescribed risk level α and is irrelevant to any other parameters. Thus, standard tests in the application of VaR can be applied to CVaR. The implementation of the SV-ALD model under SMU representation in this paper relies on the Bayesian specific software WinBugs.[6]

The rest of this paper is organized as follows. Section 2 gives a brief review of the market risk concept, discrete SV model and ALD structure, prior to providing details for the VaR and CVaR formula of oil supply and demand in the framework of SV-ALD model. Section 3 outlines the estimation methodology, including Maximum likelihood estimation (MLE) and Bayesian MCMC algorithm. Additionally, the proposed SMU representation of ALD is discussed. Section 4 is devoted to the empirical analysis of market risk in two major oil markets WTI and Brent. Finally, concluding remarks are given in Section 5.

[6]WinBugs is short for Windows version of Bayesian Analysis Using Gibbs Sampler, details about the application of WinBugs in SV can refer to Meyer and Yu (2000).

4.2 Model specification

In this section, the statistical concepts of VaR and CVaR for oil supply and demand are first introduced. Then, we discuss the specification of standard discrete stochastic volatility model.[7] Next, we introduce the mathematical definition and properties of ALD. Last, we propose the closed-form VaR and CVaR model for both oil supply and demand in the framework of SV-ALD.

4.2.1 VaR and CVaR risk measure

Although the VaR approach has been widely employed in measuring the market risk, existing studies have not extensively used yet the CVaR as risk measure in energy markets.[8] In this paper, we employ both the VaR and CVaR approach to measure the market risk with focus not only on the left tail risks for oil supply but on the right tail risks for oil demand. Specifically, the probability of left tail of the return distribution measures the potential risk when the oil price drops. The left α quantile is the point of VaR which indicates the extra expenses of oil supply as a result of a sharp fall of oil prices. On the other hand, the probability of the right tail represents the risk when the oil price raises, namely, the risk for oil supply.

Considering $VaR_{s,t}(l)$ and $VaR_{d,t}(l)$ as the VaR for oil supply and oil demand in l-period with confidence level $(1-\alpha) \in (0,1)$ respectively, then, we have:

$$Supply: \quad Prob\left(y_t(l) \leq -VaR_{s,t}(l)|\Omega_t\right) = \alpha \tag{4.1}$$

$$Demand: \quad Prob\left(y_t(l) \geq VaR_{d,t}(l)|\Omega_t\right) = \alpha \tag{4.2}$$

where $y_t(l)$ denotes the oil return series for l period of time (from t to $t+l$), Ω_t is the information set up to time t and α is the risk level. Notably, the value of $VaR_{s,t}$ and $VaR_{d,t}$ is defined to be positive. Likewise, $CVaR_{s,t}(l)$ and $CVaR_{d,t}(l)$ are defined as the CVaR for oil supply and demand respectively over l-period time

[7] Relevant studies of SV model can also refer to Jian et al. (2011) and Chan et al. (2016).

[8] One can be found is by Youssef et al. (2015) who have evaluated the VaR and CVaR for crude oil and gasoline markets by employing a series of GARCH-type models using extreme value theory to model the tail distributions. They advocate CVaR as an alternative risk measure to the quantile based VaR because of the desirable properties of CVaR over VaR.

at confidence level $(1-\alpha)$ and they can be mathematically expressed as:

$$Supply: \quad CVaR_{s,t}(l) = -E\{y_t(l)|y_t(l) \leq -VaR_{s,t}(l)\} \tag{4.3}$$

$$Demand: \quad CVaR_{d,t}(l) = E\{y_t(l)|y_t(l) \geq VaR_{d,t}(l)\} \tag{4.4}$$

4.2.2 Stochastic volatility model

A traditional approach to measure financial risk is to assume that the asset returns are normally distributed with constant variance. Undoubtedly, results obtained under this method are rough and inaccurate as a result of the time-varying volatility features in real markets. Therefore, a number of GARCH-type models are widely applied for modelling the dynamics of oil prices (see, e.g., Chkili et al., 2014; Youssef et al., 2015; Chan et al., 2016). Although EGARCH and EGARCH-GED models are generally have a good performance (see Jian et al., 2011), most GARCH-type models are not very suitable to model the asymmetry, kurtosis and leverage effect features of oil returns. On the other hand, extensive studies have shown that the SV models are more likely to outperform their GARCH counterparts. However, due to the difficulty to estimate the parameters, SV models were rarely used in practice until recently (e.g. Jian et al., 2011 and Chan et al., 2016, among others). Hence, it is of interest in this paper to explore the market risks by modelling the oil return series using the SV model.

The discrete SV model can be considered as the discretization of the Ornstein-Uhlenbeck stochastic differential equations (Scott, 1987; Taylor, 1994). In this paper, we adopt a standard discrete SV model to capture the volatility feature of oil markets which has been studied recently by Jian et al. (2011) and Chan et al. (2016).[9] To be specific, the return equation in the SV model is a linear function of conditional mean and latent volatilities, and the volatility equation has

[9] A number of empirical works via the extended SV models can be found from Breidt et al. (1998), Yu and Yang (2002), Koopman and Uspensky (2002) and Cappuccio et al. (2004), among others.

the first-order autoregressive component:

$$y_t = \mu + \sigma_t z_t \tag{4.5}$$

$$ln\sigma_t^2 = h_t = \alpha + \beta(h_{t-1} - \alpha) + \eta_t \qquad \eta_t \sim N(0, \sigma_\eta^2) \tag{4.6}$$

where $t = 1, 2, ..., T$ is the time horizon, y_t are the oil returns at time t which can be calculated through $y_t = ln(p_t/p_{t-1})$ with p_t being the daily spot price at time t, μ denotes conditional mean,[10] h_t is the log-volatility at time t,[11] $ln\sigma_t^2$ follows a stationary AR(1) process with persistence parameter β having $|\beta| < 1$,[12] z_t and η_t represents a series of independent identical distributed ($i.i.d.$) random errors in the return and volatility equations respectively and they are uncorrelated.

In the current setting, the error term z_t in the return equation is assumed to be Asymmetric Laplace distributed (Kotz et al., 2001), while η_t follows a Normal distribution with mean 0 and variance σ_η^2. Given the value of h_{t-1}, α and β, h_t follows a Normal distribution with conditional mean $E[h_t|h_{t-1}]$ and variance $Var(h_t|h_{t-1})$ expressed as:

$$E[h_t|h_{t-1}] = E[\alpha + \beta(h_{t-1} - \alpha)] + E[\eta_t] = \alpha + \beta(h_{t-1} - \alpha) \tag{4.7}$$

$$\begin{aligned} Var(h_t|h_{t-1}) &= E[(h_t|h_{t-1} - E[h_t|h_{t-1}])^2] = E[\eta^2] \\ &= Var(\eta_t) - (E[\eta_t])^2 = \sigma_\eta^2 \end{aligned} \tag{4.8}$$

or equivalently written as:

$$h_t|h_{t-1}, \alpha, \beta, \sigma_\eta^2 \sim N(\alpha + \beta(h_{t-1} - \alpha), \sigma_\eta^2) \qquad t = 2, 3, ..., T \tag{4.9}$$

[10] As indicated in Table 4.1, sample mean of return series is quite small comparing to its standard deviation, hence it is assumed to be 0 in this paper, follow the work of Koopman et al. (2005) and Wei et al. (2010). It should be noticed that the criteria are different regarding to the definition of return equation in discrete SV model, i.e. μ is simply abandoned in some studies (See Omori et al., 2007; Takahashi et al., 2009; Nakajima and Omori, 2012).

[11] For $h_t = ln\sigma_t^2$, equation (4.5) can also be written in another equivalent form: $y_t = \mu + z_t e^{h_t/2}$.

[12] This condition is necessary to satisfy the requirement for stationarity, it is not difficult from equation (4.6) to observe that persistence parameter β reflects the impact of current volatilities on future volatilities.

Similarly, the unconditional distribution of h_t follows a Normal distribution with unconditional mean and unconditional variance expressed as $E[h_t] = \alpha$ and $Var(h_t) = \sigma_\eta^2/(1-\beta^2)$, respectively. Since each of η_t is *i.i.d.*, then the unconditional expectation and variance of h_t can be formally specified as:[13]

$$h_t \sim N(\alpha, \frac{\sigma_\eta^2}{1-\beta^2}) \qquad t = 1, 2, ..., T \tag{4.10}$$

4.2.3 Asymmetric Laplace distribution

The Asymmetric Laplace Distribution, as one with good-property of skewed distribution families, is often adopted to describe the distribution of the losses of financial assets because it can feature skewness which is observed in the distribution of asset returns. An obvious advantage of the application of ALD can be attributed to the fact that it is simple to compute. A closed form expression for the *p.d.f.* would ideally make the simulation of the model simpler. Therefore, to measure the market risk of oil prices in a SV setting, we assume that the disturbance term of oil returns is AL distributed rather than Normal. Before conducting empirical analysis, it is necessary to review the statistical properties of ALD that proposed by Kotz et al. (2001).[14]

We consider the error term z_t of the return equation (formula 4.5) in SV model follows the ALD, thus $z_t \sim AL(\theta, \kappa, \tau)$, where θ is the location parameter, κ is the skewness parameter and τ is the scale parameter.[15] A simplified form of the *p.d.f.* for ALD is introduced via the following proposition:

Proposition 4.1. *Let $f(z|\kappa, \theta, \tau)$ denote the p.d.f of a $AL(\kappa, \theta, \tau)$ distribution,*

[13] See Appendix F.1 for derivation.

[14] Notably, as mentioned in the introduction, there are different analytic expressions of the *p.d.f.* of ALD. Guermat and Harris (2002) have explored short horizon assets returns by employing GARCH volatility model with symmetric Laplace distribution, this work was then being extended via AL distribution that proposed by Lu et al. (2010). Chen et al. (2012) slightly simplify and modify their model to estimate and forecast VaR and CVaR using GJR-GARCH model by assuming error term follows AL distribution. Another known formation of ALD is proposed by Yu and Zhang (2005). Based on this, Wichitaksorn et al. (2015) study asymmetries and quantiles of stock index returns using stochastic volatility model with ALD error.

[15] Following the suggestion of Kotz et al. (2001), location parameter θ in ALD is set to be 0 in our following empirical practice but still left in writings to keep the completeness of notions.

then, the $f(z|\kappa,\theta,\tau)$ *can be expressed as:*

$$f(z|\kappa,\theta,\tau) = \begin{cases} \frac{\sqrt{2}}{\tau}\frac{\kappa}{1+\kappa^2}\ exp(-\frac{\sqrt{2}\kappa}{\tau}(z-\theta)) & z \geq \theta \\ \frac{\sqrt{2}}{\tau}\frac{\kappa}{1+\kappa^2}\ exp(\frac{\sqrt{2}}{\tau\kappa}(z-\theta)) & z < \theta \end{cases} \tag{4.11}$$

The error term of return equation in SV model is said to behave as ALD if its *p.d.f.* follows equation (4.11). The parameter κ controls the probability assigned to each side of location parameter θ, $\kappa = 1$ implies the probability are equivalent and distribution is symmetric of θ. When $\kappa > 1$, the left tail of the distribution is thicker than the right. When $0 < \kappa < 1$, the right rail is thicker than the left.

4.2.4 VaR and CVaR in SV-ALD setting

Risk of oil Supply

VaR: Follow the definition, the VaR for oil supply can be written as:

$$P(y_t \leq -VaR_{\mathrm{s,t}}|\Omega_t) = \int_{-\infty}^{-VaR_{s,t}} f(y_t|\Omega_t)\, dy_t = \alpha \tag{4.12}$$

where $f(y_t|\Omega_t)$ is the conditional *p.d.f.* of y_t that follows ALD.

Substituting the return equation of SV model (formula 4.5) into the above definition and rewrite it, we thus have a standard form:

$$\begin{aligned} P(y_t \leq -VaR_{\mathrm{s,t}}|\Omega_t) &= P\left(\frac{y_t-\mu}{\sigma_t} \leq -\frac{VaR_{s,t}+\mu}{\sigma_t}\middle|\Omega_t\right) \\ &= P\left(z_t \leq -m_{s,q} = -\frac{VaR_{s,t}+\mu}{\sigma_t}\right) = \int_{-\infty}^{-m_{s,q}} f^{-}(z_t)\, dz_t = \alpha \end{aligned} \tag{4.13}$$

where $m_{s,q}$ is designed as the left α-quantile of the ALD, which is used for the residual series of SV model, $m_{s,q}$ is set to be equal to $(VaR_{s,t}+\mu)/\sigma_t$ and $f^{-}(z_t)$ represents the negative part of the *p.d.f.* of ALD. The analytic expression of $m_{s,q}$

can be obtained by solving the integral equation (4.13):[16]

$$m_{s,q} = -\frac{\kappa\tau}{\sqrt{2}} ln \frac{\alpha(1+\kappa^2)}{\kappa^2} \tag{4.14}$$

Hence, the analytic form of VaR for oil supply under SV-ALD model can be expressed as follows:

$$VaR_{s,t} = -\mu + m_{s,q}\sigma_t = -\mu - \frac{\kappa\tau\sigma_t}{\sqrt{2}} ln \frac{\alpha(1+\kappa^2)}{\kappa^2} \tag{4.15}$$

CVaR: Similarly, using the same notation as in the VaR framework, a standard form of CVaR for oil supply can be written as[17]:

$$CVaR_{s,t} = -\, E\left[y_t | y_t \leq -VaR_{s,t}\right] = -(\mu + \sigma_t E\left[z_t | z_t \leq -m_{s,q}\right]) \tag{4.16}$$

As a result, the ultimate closed-form expression of CVaR for oil supply under SV-ALD model can be shaped as:[18]

$$CVaR_{s,t} = -\mu - \frac{\kappa\tau\sigma_t}{\sqrt{2}} \left(1 - ln \frac{\alpha(1+\kappa^2)}{\kappa^2}\right) \tag{4.17}$$

or equivalently written as:

$$CVaR_{s,t} = VaR_{s,t} + \frac{\kappa\tau\sigma_t}{\sqrt{2}} \tag{4.18}$$

Risk of oil Demand

VaR: The initial VaR of oil demand can be formulated as follows:

$$P(y_t > VaR_{\mathrm{d,t}} | \Omega_t) = \int_{VaR_{d,t}}^{+\infty} f(y_t | \Omega_t)\, dy_t = \alpha \tag{4.19}$$

where $f(y_t|\Omega_t)$ is the conditional $p.d.f.$ of y_t that follows ALD.

[16]See Appendix F.2 for derivation of $m_{s,q}$.

[17]A similar way for defining CVaR under the dynamic mean equation can refer to Youssef et al. (2015), which has studied the estimation of extreme risks in crude oil and gasoline market via GARCH-EVT model.

[18]See Appendix F.3 for derivation of $CVaR_{s,t}$.

By incorporating SV model, a standard form can be written as:

$$\begin{aligned} P(y_t > VaR_{\mathrm{d,t}}|\Omega_t) &= P\left(\frac{y_t - \mu}{\sigma_t} > \frac{VaR_{d,t} - \mu}{\sigma_t}\middle|\Omega_t\right) \\ &= P\left(z_t > m_{d,q} = \frac{VaR_{d,t} - \mu}{\sigma_t}\right) = \int_{m_{d,q}}^{+\infty} f^+(z_t)\, dz_t = \alpha \end{aligned} \tag{4.20}$$

where $m_{d,q} = (VaR_{d,t} - \mu)/\sigma_t$ is the right α-quantile of the AL distribution and $f^+(z_t)$ denotes the positive part of the $p.d.f.$ of ALD. Solving the above integral, an analytic expression of $m_{d,q}$ can be obtained:[19]

$$m_{d,q} = -\frac{\tau}{\sqrt{2}\kappa}\, ln(\alpha(1+\kappa^2)) \tag{4.21}$$

Hence, VaR for oil demand under SV-ALD model can be formulated as:

$$VaR_{d,t} = \mu + m_{d,q}\sigma_t = \mu - \frac{\tau\sigma_t}{\sqrt{2}\kappa}\, ln(\alpha(1+\kappa^2)) \tag{4.22}$$

CVaR: Analogously, the standard formation of CVaR for oil demand is given by:

$$CVaR_{d,t} = E\left[y_t | y_t > VaR_{d,t}\right] = \mu + \sigma_t E\left[z_t | z_t > m_{d,t}\right] \tag{4.23}$$

As a result, the analytic expression of CVaR for oil demand under SV-ALD model can be written as follows:[20]

$$CVaR_{d,t} = \mu + \frac{\tau\sigma_t}{\sqrt{2}\kappa}\left(1 - ln\left(\alpha(1+\kappa^2)\right)\right) \tag{4.24}$$

or equivalently:

$$CVaR_{d,t} = VaR_{d,t} + \frac{\tau\sigma_t}{\sqrt{2}\kappa} \tag{4.25}$$

4.3 Estimation methodology

Given the explicit analytical expressions of VaR and CVaR, related parameters (i.e. κ, τ and σ_t) are necessary for risk measuring. We employ the Maximum Likelihood

[19] See Appendix F.4 for derivation of $m_{d,q}$.

[20] See Appendix F.5 for derivation of $CVaR_{d,t}$.

Estimation method to estimate density parameter κ and τ in ALD and Bayesian Markov Chain Monte Carlo approach for model parameters and latent volatilities in SV scheme. In addition, a new scale mixture of uniform representation is proposed for the implementation of the Bayesian SV-ALD model.

4.3.1 Maximum likelihood estimation

Since the *p.d.f* of ALD is analytically tractable, hence the MLE approach is applicable to the parameter estimations of ALD. Suppose that $Y_1, \ldots Y_n$ are *i.i.d.* random samples from $Y \sim AL(\theta, \kappa, \tau)$, and let $y_1, \ldots y_n$ be the nth order statistics of the random sample. Then the maximum likelihood function is defined as:

$$L(y; \theta; \kappa; \tau) = \frac{2^{\frac{n}{2}}}{\tau^n} \frac{k^n}{(1+k^2)^n} exp\left[-\frac{\sqrt{2}\kappa}{\tau}\sum_{i=1}^{n}(y_i-\theta)^+ - \frac{\sqrt{2}}{\kappa\tau}\sum_{i=1}^{n}(y_i-\theta)^-\right] \quad (4.26)$$

where

$$(y_i-\theta)^+ = \begin{cases} y_i - \theta & if\ y_i \geq \theta \\ 0 & if\ y_i < \theta \end{cases}$$

and

$$(y_i-\theta)^- = \begin{cases} \theta - y_i & if\ y_i < \theta \\ 0 & if\ y_i \geq \theta \end{cases}$$

Taking the natural logarithm of equation (4.26) yields the log-likelihood function:

$$lnL(\theta; \kappa; \tau) = \frac{n}{2}ln(2) - nln(\tau) + nln(\frac{k}{1+\kappa^2}) - \frac{\sqrt{2}}{\tau}M \quad (4.27)$$

where

$$M(y; \kappa; \theta) = k\sum_{i=1}^{n}(y_i-\theta)^+ + \frac{1}{\kappa}\sum_{i=1}^{n}(y_i-\theta)^-$$

For a given θ (i.e. 0), then our aim becomes to maximize the following function:[21]

$$P(\kappa;\tau) = ln(\kappa) - ln(1+\kappa^2) - ln(\tau) - \frac{\sqrt{2}}{\tau}\{\kappa, \frac{1}{\kappa}\}\bar{Z}^{(n)} \tag{4.28}$$

where the vector $\bar{Z}^{(n)}$ is given by:

$$\bar{Z}^{(n)} = \frac{1}{n}\sum_{i=1}^{n} Z^{(i)} = \left[\frac{1}{n}\sum_{i=1}^{n} Z_1^{(i)}, \frac{1}{n}\sum_{i=1}^{n} Z_2^{(i)}\right]'$$

and the two components of the vector $\bar{Z}^{(n)}$ are:

$$\bar{Z}_1^{(n)} = \frac{1}{n}\sum_{i=1}^{n} Z_1^{(i)} = \frac{1}{n}\sum_{i=1}^{n}(y_i-\theta)^+, \qquad \bar{Z}_2^{(n)} = \frac{1}{n}\sum_{i=1}^{n} Z_2^{(i)} = \frac{1}{n}\sum_{i=1}^{n}(y_i-\theta)^-$$

4.3.2 Scale mixture of uniform representation of ALD

Expressing ALD as an SMU representation can alleviate the computational burden when using Gibbs sampling algorithm in MCMC and thus simplifying the estimation method in the Bayesian analysis. In this paper, a new scale mixture of uniform (SMU) representation for the AL density of Kotz et al. (2001) is proposed to facilitate the implementation of SV-ALD model. According to Choy and Chan (2008), the definition of SMU can be presented as:

Definition 4.2. *For a continuous random variable z^* with location parameter θ^* and scale parameter τ^*, the probability density function of z^* is said to have an SMU if it can be expressed as:*

$$f(z^*|\theta^*,\tau^*) = \int_0^\infty f_U(z^*|\theta^* - \psi(\lambda)\tau^*, \theta + \psi(\lambda)\tau^*)\, f_\psi(\lambda)\, d\lambda \tag{4.29}$$

where $f_U(a,b)$ denotes the uniform density function with interval [a,b], λ is the scale mixing parameter that has been commonly used as a global diagnostic check for outliers (Choy and Smith, 1997), $\psi(\cdot)$ is a positive function and $f_\psi(\cdot)$ is a density function defined on $\mathbb{R}^+ = (0,\infty)$.

[21] MLE methodology for ALD has been explicitly described by Kotz et al. (2001) and more discussions of the case when $\theta \neq 0$ can be found in Kotz et al. (2001) ch.3.

Combining the method proposed by Wichitaksorn et al. (2015) and definition 4.2, the AL density of Kotz et al. (2001) can be expressed equivalently by an SMU.[22]

Proposition 4.3. *If* $\lambda \sim Ga(2,1)$ *and* $z \sim U(\theta - \frac{\kappa\tau\lambda}{\sqrt{2}}, \theta + \frac{\tau\lambda}{\sqrt{2}\kappa})$*, then the SMU density:*

$$f(z|\kappa,\tau,\theta,\lambda) = \int_0^\infty f_U(z|\theta - \frac{\kappa\tau\lambda}{\sqrt{2}}, \theta + \frac{\tau\lambda}{\sqrt{2}\kappa}) \times f_{Ga}(\lambda|2,1)\, d\lambda \tag{4.30}$$

has the same form as that of the AL density function given by:

$$f(z|\kappa,\tau,\theta) = \begin{cases} \frac{\sqrt{2}}{\tau}\frac{\kappa}{1+\kappa^2}\, exp(-\frac{\sqrt{2}\kappa}{\tau}(z-\theta)) & z \geq \theta \\ \frac{\sqrt{2}}{\tau}\frac{\kappa}{1+\kappa^2}\, exp(\frac{\sqrt{2}}{\tau\kappa}(z-\theta)) & z < \theta \end{cases} \tag{4.31}$$

where z *is a random variable follows ALD,* κ*,* τ *and* θ *denote skewness, scale and location parameters in the AL density respectively and* $f_{Ga}(c,d)$ *is the gamma density function of the form:*

$$Ga(\lambda|c,d) = \frac{1}{\Gamma(c)d^c}\lambda^{c-1}\, exp(-\frac{\lambda}{d}) \qquad \lambda, c, d > 0 \tag{4.32}$$

with shape parameter c *and scale parameter* d*, and* $\Gamma(c)$ *is the gamma function evaluated at* c*.*

To estimate the parameters in the context of SV model, we use the scaled ALD (SALD) which means that the ALD random variable is scaled by its standard deviation.[23] The *p.d.f.* of SALD is characterized as follows.

Proposition 4.4. *Let* z *be an ALD random variable with* $z \sim ALD(\kappa,\tau,\theta)$*, then the random variable* $\varepsilon_t = \frac{z}{S.D.[z]}$ *has SALD with p.d.f. given by:*

$$f(\varepsilon_t|\kappa,\theta,\sigma_t) = \begin{cases} \frac{\sqrt{1+\kappa^4}}{1+\kappa^2}\frac{1}{\sigma_t}\, exp(\frac{-\sqrt{1+\kappa^4}}{\sigma_t}(\varepsilon_t-\theta)) & \varepsilon_t \geq \theta \\ \frac{\sqrt{1+\kappa^4}}{1+\kappa^2}\frac{1}{\sigma_t}\, exp(\frac{\sqrt{1+\kappa^4}}{\kappa^2\sigma_t}(\varepsilon_t-\theta)) & \varepsilon_t < \theta \end{cases} \tag{4.33}$$

where κ *is skewness parameter and* σ_t *is the standard deviation (or the time-varying*

[22] See appendix G.1 for the derivation.

[23] More details of this algorithm can refer to Chen et al.(2009) and Wichitaksorn et al. (2015).

volatility) of z_t.[24]

Hence, the corresponding SMU of SALD can be obtained as follows:

Proposition 4.5. *If* $\lambda \sim Ga(2,1)$ *and* $\varepsilon_t \sim U(\varepsilon_t|\theta - \frac{\lambda\kappa^2\sigma_t}{\sqrt{1+\kappa^4}}, \theta + \frac{\lambda\sigma_t}{\sqrt{1+\kappa^4}})$, *then the SMU density:*

$$f(\varepsilon_t|\kappa,\theta,\lambda,\sigma_t) = \int_0^\infty f_U(\varepsilon_t|\theta - \frac{\lambda\kappa^2\sigma_t}{\sqrt{1+\kappa^4}}, \theta + \frac{\lambda\sigma_t}{\sqrt{1+\kappa^4}}) \times f_{Ga}(\lambda|2,1)\, d\lambda \qquad (4.34)$$

has the same form as the SALD density function given in equation (4.33).[25]

Using the SMU representation of SALD, an efficient simulation algorithm is developed to overcome the parameter estimation difficulties. It facilitates the Gibbs sampling for Bayesian computation in SV model where return errors are modelled by ALD while volatility errors are modelled by Normal distribution. As a result, the SV model discussed in section 4.2.2 can be written hierarchically as:

Return equation:

$$y_t|\kappa,\theta,\lambda,h_t \sim U(\theta - \frac{\lambda\kappa^2 e^{h_t/2}}{\sqrt{1+\kappa^4}}, \theta + \frac{\lambda e^{h_t/2}}{\sqrt{1+\kappa^4}}) \qquad (4.35)$$

$$\lambda \sim Ga(2,1) \qquad (4.36)$$

Volatility equation:

$$h_t|\alpha,\beta,\sigma_\eta^2,h_{t-1} \sim N(\alpha + \beta(h_{t-1}-\alpha), \sigma_\eta^2) \qquad t = 1,2,...,T \qquad (4.37)$$

$$h_1|\alpha,\beta,\sigma_\eta^2 \sim N(\alpha, \frac{\sigma_\eta^2}{1-\beta^2}) \qquad (4.38)$$

4.3.3 Bayesian Markov Chain Monte Carlo

Parameter estimation of the SV model is not straightforward due to the intractable form of likelihood function, hence, traditional method, such as MLE, cannot be used.[26] Several estimation methods have been proposed in literature in order to

[24] Note that scale parameter τ has been canceled out in this derivation, location parameter θ is set to be 0 in real practice. Relevant derivations see appendix G.2.

[25] See appendix G.3 for the derivation.

[26] In the SV-ALD model, the likelihood function is a t-dimensional integration with unknown latent volatilities, which means the closed-form could not been identified.

solve this problem, including the Generalized Method of Moments (Melino and Turnbull, 1990), the Quasi-Maximum Likelihood (Nelson, 1988 and Harvey et al., 1994), the Efficient Method of Moments (Gallant, 1997) and the Simulated Maximum Likelihood (Danielsson, 1994 and Sandmann and Koopman, 1998). Unlike conventional statistical inference methods based on the maximum-likelihood approach, Bayesian inference simply relies on the joint posterior distributions, which according to the Bayes' rule, can be factored as:

$$f(h,\dot{\theta}|y) \propto f(y|h,\dot{\theta})\, f(h|\dot{\theta})\, f(\dot{\theta}) \tag{4.39}$$

where sign $\propto$ stands for proportion, $y = (y_1, ..., y_T)$ is a series of observed returns, $h = (h_1, ..., h_T)$ denotes the unobserved latent volatilities, $\dot{\theta}$ represents the estimated parameters in SV model (i.e. $\dot{\theta} = (\alpha, \beta, \sigma_\eta^2)$), $f(y|h,\dot{\theta})$ is the full-information likelihood function, $f(h|\dot{\theta})$ is conditional distribution of the state variables and $f(\dot{\theta})$ represents the prior distribution that summarizes all initial beliefs in parameters.[27] However, using the posterior distribution generally involves computing integrals and the target posterior distribution $f(h,\dot{\theta}|y)$ in equation (4.39) is analytically intractable due to the non-linearity in equation (4.5), hence it is not easy to sample immediately. Yu and Meyer (2006) point out that the difficulties of multidimensional numerical integration involved in the posterior computations have been overcome by the development of MCMC techniques following and extending the seminal work of Jacquier et al. (1994). As a simulation based approach, MCMC is particularly well suited to the estimation of non-linear state space models (Karali et al., 2011) and has been extensively studied in empirical applications. It works through generating a number of random samples and then constructing a Markov chain whose stationary distribution converges to the joint posterior distribution.

The Bayesian MCMC algorithm via Gibbs sampling algorithm is employed in this

[27] A mathematical expression according to Bayes' theorem is shaped as: $f(h,\dot{\theta}|y) = \frac{f(y|h,\dot{\theta})\, f(h,\dot{\theta})}{f(y)}$. In the application of Bayesian inference, however, the posterior density distribution $f(h,\dot{\theta}|y)$ does not relying on density $f(y) = \int f(y|h,\dot{\theta}) f(h,\dot{\theta}) dh d\dot{\theta}$. The reason is that for a fixed return series, the density is irrelevant to h and $\dot{\theta}$, thus $f(y)$ is regarded as a constant. Bayes' theorem then shows that the posterior is proportional to the prior times likelihood function.

paper to make posterior inferences. To implement the MCMC, we set priors as:

$$\alpha \sim N(\mu_\alpha, \sigma_\alpha^2) \tag{4.40}$$

$$\frac{1+\beta}{2} = \beta^* \sim Be(a_\beta, b_\beta) \tag{4.41}$$

$$\sigma_\eta^2 \sim IG(a_\sigma, b_\sigma) \tag{4.42}$$

where $Be(\cdot,\cdot)$ denotes a beta distribution and $IG(\cdot,\cdot)$ is an inverse-gamma distribution. Let $f(\alpha)$, $f(\beta)$ and $f(\sigma_\eta^2)$ denote the prior probability densities of α, β and σ_η^2 respectively, then, the conditional posterior distributions of model parameters and latent variables can be formulated as follows:

$$f(\alpha|\beta, \sigma_\eta^2, h, y) \propto f(y|\alpha, \beta, \sigma_\eta^2, h) f(h|\alpha, \beta, \sigma_\eta^2) f(\alpha) \tag{4.43}$$

$$f(\beta|\alpha, \sigma_\eta^2, h, y) \propto f(y|\alpha, \beta, \sigma_\eta^2, h) f(h|\alpha, \beta, \sigma_\eta^2) f(\beta) \tag{4.44}$$

$$f(\sigma_\eta^2|\alpha, \beta, h, y) \propto f(y|\alpha, \beta, \sigma_\eta^2, h) f(h|\alpha, \beta, \sigma_\eta^2) f(\sigma_\eta^2) \tag{4.45}$$

$$f(h_t|h_{-t}, \alpha, \beta, \sigma_\eta^2, y) \propto f(y|h_t, \alpha, \beta, \sigma_\eta^2) f(h_t|h_{-t}, \alpha, \beta, \sigma_\eta^2) \tag{4.46}$$

where h_{-t} represents all the elements of $h = (h_1, ..., h_T)$ except h_t. The system of full conditional distributions via the SMU of ALD are:[28]

- **Full conditional distribution for** α:

$$\alpha|\beta, \sigma_\eta^2, h, y \sim N(\hat{\mu}_\alpha, \hat{\sigma}_\alpha^2)$$

where $\hat{\mu}_\alpha = \hat{\sigma}_\alpha^2[\frac{h_1(1-\beta^2)+(1-\beta)\sum_{t=2}^T(h_t-\beta h_{t-1})}{\sigma_\eta^2} + \frac{\mu_\alpha}{\sigma_\alpha^2}]$ and $\hat{\sigma}_\alpha^2 = [\frac{1-\beta^2+(T-1)(1-\beta)^2}{\sigma_\eta^2} + \frac{1}{\sigma_\alpha^2}]^{-1}$.

- **Full conditional distribution for** β^*:

$$f(\beta^*|\alpha, \sigma_\eta^2, h, y) = f(h_1|\alpha, \frac{\sigma_\eta^2}{1-\beta^2}) \times \sum_{t=2}^T f(h_t|\alpha + \beta(h_{t-1} - \alpha), \sigma_\eta^2) \times f(\beta^*|a_\beta, b_\beta)$$

[28]Derivations can be found in Appendix G.4.

- **Full conditional distribution for σ_η^2:**

$$\sigma_\eta^2|\alpha,\beta,h,y \sim IG(\hat{a}_\sigma,\hat{b}_\sigma)$$

where $\hat{a}_\sigma = a_\sigma + \frac{T}{2}$ and $\hat{b}_\sigma = b_\sigma + \frac{1}{2}(h_1-\alpha)^2(1-\beta^2) + \frac{1}{2}\sum_{t=2}^{T}(h_t-\alpha-\beta(h_{t-1}-\alpha))^2$.

- **Full conditional distribution for h_t:**

$$h_t|h_{-t},\alpha,\beta,\sigma_\eta^2,y \sim N(\hat{\mu}_h,\hat{\sigma}_h^2)$$

where

$$\hat{\mu}_h = \begin{cases} \alpha+\beta(h_{t+1}-\alpha)-\dfrac{\sigma_\eta^2}{2} & if \quad t=1, \\ \alpha+\dfrac{\beta[(h_{t-1}-\alpha)+(h_{t+1}-\alpha)]}{1+\beta^2} & if \quad t=2,3,...,T-1 \; and \\ \alpha+\beta(h_{T-1}-\alpha)-\dfrac{\sigma_\eta^2}{2} & if \quad t=T, \end{cases}$$

$$\hat{\sigma}_h^2 = \begin{cases} \sigma_\eta^2 & if \quad t=1,T \;\; and \\ \dfrac{\sigma_\eta^2}{1+\beta^2} & if \quad t=2,...,T-1 \end{cases}$$

Thus, it is possible to directly sample from the full conditional posterior distributions by sweeping each variable while keeping the remaining variables fixed. The resulting Gibbs sampling algorithm is summized as Algorithm 1.

Algorithm 1 Gibbs sampler

1: Initialize $\alpha^{(0)}$, $\beta^{(0)}$, $\sigma_\eta^{2(0)}$ and $h^{(0)}$.
2: Sample α from $f(\alpha|\beta,\sigma_\eta^2,h,y)$.
3: Sample β from $f(\beta|\alpha,\sigma_\eta^2,h,y)$.
4: Sample σ_η^2 from $f(\sigma_\eta^2|\alpha,\beta,h,y)$.
5: Sample h_t from $f(h_t|h_{-t},\alpha,\beta,\sigma_\eta^2,y)$.
6: Return to step 2 and repeat this procedure untile convergence is achieved.

This procedure is implemented in the Bayesian specific software WinBugs which gives the flexibility of using various density functions (i.e. Normal, t, Gamma, Beta, etc.) to specify the prior distributions and likelihood functions for Bayesian inference. The proposed SMU representation of ALD enables some of the full conditional distributions reduced to standard forms to simplify the Gibbs sampling algorithm, hence making the statistical inference of the SV-ALD model easy to implement within the WinBugs environment.[29]

4.4 Empirical analysis

4.4.1 Data description and preliminary results

The objective of this study is to estimate the crude oil price risk by modelling its returns and volatilities using the proposed SV-ALD model. With this aim, two major crude oil markets are considered: West Texas intermediate crude oil (WTI) and Europe Brent oil (Brent). Daily closing spot prices, which are quoted in US dollars per barrel, are obtained from the U.S. Energy Information Administration (EIA) covering the periods from May 22, 2006 to May 20, 2016 with 2520 observations in WTI and 2522 observations in Brent. Let p_t denotes the oil price on day t, then the sample price can be converted into daily price return y_t by: $y_t = ln(p_t/p_{t-1})$.

The time variations of daily prices and returns for WTI and Brent are plotted in Figure 4.1. We can see that the trajectory of historical prices for WTI and Brent have exhibited high similarities during May, 2006 to May, 2016. The graphs of daily returns show the volatility clustering effect in both the two oil markets which reveal the presence of heteroscedasticity. In addition, we can observe that both WTI and Brent oil prices are quite volatile in the 2007-2009 global financial crisis characterized by a succession of large fluctuations of oil returns within a short time period, and this in turn reflect the necessities of strengthening oil risk control. The largest changes of oil returns for WTI and Brent occurred on September 22, 2008 and January 2, 2009 with a record of 16.41% and 18.13% surge respectively. Figure 4.2 demonstrates a scatterplot of the observed oil returns versus quantiles calculated from a standard Normal distribution. This subjective check indicates the plausibility of using heavy-tailed distribution for oil returns as there are more extreme values

[29] This paper gives a major contribution from an implementation point of view as develops the statistical modelling of the SV-ALD model using WinBugs.

Table 4.1: Descriptive statistics of crude oil price returns

	WTI	Brent
Panel A: Descriptive statistics		
Mean	-0.000144	-0.000127
Std.dev.	0.024863	0.021998
Maximum	0.164137	0.181297
Minimum	-0.128267	-0.168320
Skewness	0.1567	0.1443
Kurtosis	7.6122	8.8043
J-B test	2243.0570***	3547.5790***
Q(10)	30.6030***	16.9600*
Q(20)	60.8980***	54.2270***
ARCH(10)	475.9680***	215.7230***
ARCH(20)	575.8620***	409.0370***
Panel B: Unit roots and stationarity tests		
ADF	-51.4930***	-48.9570***
PP	-51.5220***	-48.9660***
Zivot-Andrews	-13.5590***	-12.2245***
KPSS	0.0507	0.0690

Note: Q(l) is the Ljung-Box statistic with order up to l. Test statistic of ARCH(m) is obtained using chi-squared distribution with lag up to m while ADF, PP and Zivot-Andrews statistics are based on t distribution. The largest value from the first 8 lags of KPSS test is listed. $*$, $**$ and $***$ denote rejection of null hypothesis at 10%, 5% and 1% significance level respectively.

that are curve off in the extremities of the Normal line.

It is evident that large fluctuations of crude oil prices can be caused by aggregate demand and supply shocks. Looking back to history, a typical example of an oil demand shock is the Asian financial crisis in 1998, which led to large crashes in WTI and Brent crude oil prices. On the other hand, oil production decreased in 2003 due to the second Gulf War, thus correspondingly stimulating the rise of oil prices in a short period of time. After that, oil prices experienced a persistent increase from 2003 to 2008 which was driven by the prosperity of emerging economies. With time span shortened to the past ten years, the U.S. subprime mortgage crisis has aggravated the turbulence of global economic markets, leading to a quick drop in

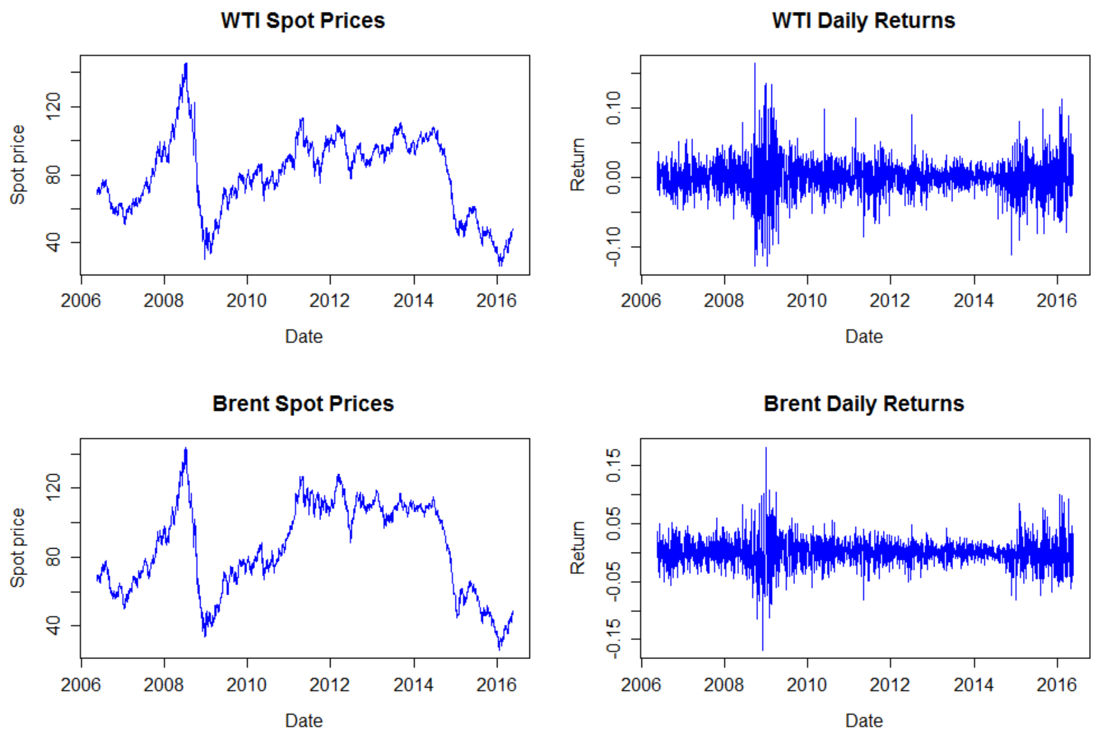

Figure 4.1: Daily spot prices and returns for WIT and Brent from May 1987 to May 2016

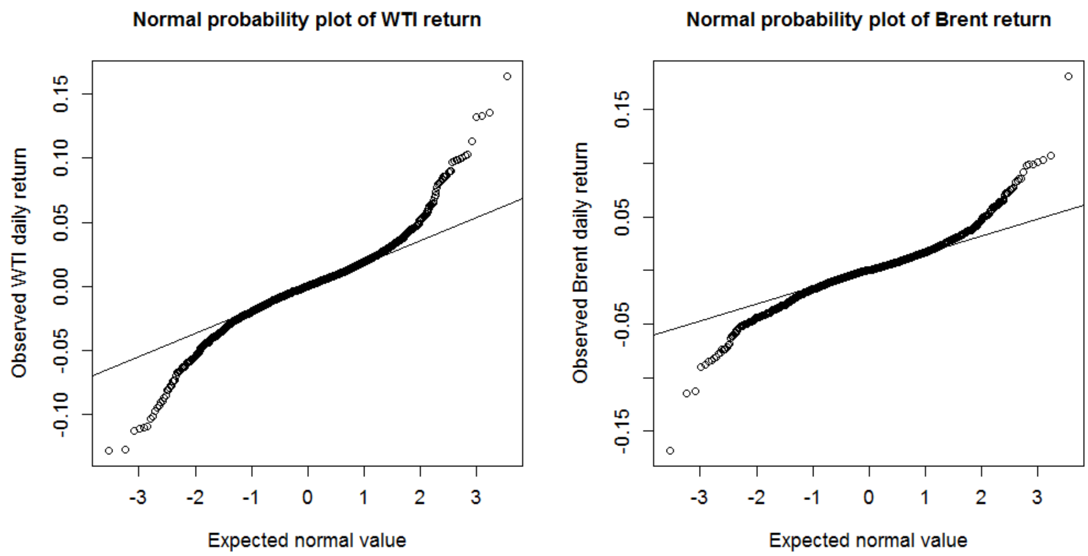

Figure 4.2: Normal probability plots of oil returns

crude oil prices from the third quarter of 2008. In addition, it is clear to see from Figure 4.1 that there is a sharp drop in oil prices for the two markets from the second half of year 2014 because of the worldwide imbalance in aggregate demand and supply. Overall, we can summarize from the historical oil prices that there exist high uncertainty over time for both the main oil markets. Although many other factors, such as geopolitical, weather conditions and cash flow fluctuations etc., could affect the movements of oil prices, the aggregate demand and supply level are undoubtedly playing a more important role, especially in the long run.

Descriptive statistics for the WTI and Brent returns as well as their stochastic characteristics are provided in Table 4.1. From *Panel A*, the Jarque-Bera statistics suggest the rejection of the null hypothesis of Gaussian distribution for the two returns at 1% significance level.[30] This is confirmed by the mild skewness and kurtosis. Results from Ljung-Box Q statistics of order up to 20 indicate that the null hypothesis of no serial correlation is rejected at 1% significance level, implying the existence of autocorrelation in the WTI and Brent return series. In addition, Engel's test shows the presence of autoregressive conditional heteroscedasticity (ARCH) effects in WTI and Brent returns.[31] These results can be observed from the pattern of return series in Figure 4.1 where large price movements are followed by large movements while small changes are followed by small changes.

Before fitting return series, we employ some tests to check the presence of unit roots and to examine stationarity. In *Panel B*, both the Augmented Dicky-Fuller (ADF) test and Phillips-Perron (PP) test significantly reject the null hypothesis of unit root in the studied WTI and Brent series at 1% significant level, and results from the Zivot-Andrews unit root test are robust to the presence of a potential structural break in the two markets.[32] Statistics from Kwiatkowski-Phillips-Schmidt-Shin

[30] This corresponds to Figure 4.2 which demonstrates significant fat tail features of the observed returns comparing to Normal probability values.

[31] The return series demonstrate strong ARCH effects from lag(1), this feature may explains the reason why GARCH-type models are widely applied by researchers and practitioners for the analysis of financial time series and appreciated for measuring risks.

[32] The most significant structural break point selected by the Zivot-Andrews test in the WTI and Brent markets are in the period of Global financial crisis, at 22nd September 2008 and 2nd January 2009, respectively.

(KPSS) test indicate do not reject the null hypothesis of stationarity. These results imply that the return series in the WTI and Brent markets are stationary and appropriate for further analysis.

4.4.2 Estimation results of SV-ALD model

The computation of VaR and CVaR under SV-ALD model can be summarized in the following steps:

-Step 1: *SV-ALD model estimation*: We fit the proposed SV-ALD model to oil return datasets using the Bayesian MCMC approach, that is, we implement Bayesian inference by sampling from the joint posterior distributions via the SMU representation. Then we estimate the time-varying volatilities (σ_t) from the fitted model and then extract the standardized residual series z_t.

-Step 2: *Density parameter estimation*: We employ the standardized residuals obtained in Step 1, which can be seen as the realization of a white noise process, to estimate parameter κ and τ in AL density function using the Maximum likelihood estimation method. Then, we calculate the corresponding α-quantile of the ALD according to the derived formulas ($m_{s,q}$ for oil supply and $m_{d,q}$ for oil demand).

-Step 3: *VaR and CVaR quantification*: We compute the dynamics of VaR and CVaR for both oil supply and demand at a given confidence interval using the estimated parameters coming from the above two steps.[33]

Convergence diagnostics. Prior to estimating the parameters from the joint posterior distribution, convergence diagnostic of the constructed Markov chain in the MCMC algorithm is necessary in order to ensure the estimation accuracy from the posterior distributions. If there are any parameters that are not converging in the chain, the corresponding estimation for other parameters would be deviating from its original path, leading to large estimation errors. Generally, preliminary convergence tests can be monitored using different methods: (i) observing the produced trace plots for parameters; (ii) using autocorrelation plots; (iii) monitoring

[33] Related work can refer to Bali et al. (2008) and Marimoutou et al. (2009).

the evolution of selected quantiles; (iv) checking the MC errors by comparing them to the corresponding estimated posterior standard deviations. Nevertheless, other formal diagnostic tests have been developed in the literature and can be directly implemented through Convergence Diagnostic and Output Analysis (CODA) or Bayesian output analysis (BOA) packages in R. This paper employs the Brooks-Gelman-Rubin (BGR) diagnostics, proposed by Gelman and Rubin (1992) and developed by Brooks and Gelman (1998), to assess the convergence of the Markov chain in the MCMC scheme.[34]

To complete a Bayesian paradigm, the prior distribution of the estimated SV model parameters are set as: $\alpha \sim N(-10, 0.001)$, $\tau_\eta \sim Ga(2.5, 0, 025)$ with $\tau_\eta = \frac{1}{\sigma_\eta^2}$, $\beta^* \sim Be(20, 1.5)$ with $\beta^* = \frac{\beta+1}{2}$ and the prior distribution of skewness parameter is chosen as: $\kappa \sim U(0, 2)$, where τ_η is the precision parameter, $Ga(\cdot,\cdot)$, $Be(\cdot,\cdot)$ and $U(\cdot,\cdot)$ represents the Gamma distribution, the Beta distribution and the Uniform distribution, respectively.[35] To implement the BGR diagnostic test, two independent Markov chains are generated in parallel by setting different initial values. Four parameters are assigned to initial values in each chain and for other parameters, i.e. the latent parameters, are generated randomly from the prior distributions. The initial values of the two Markov chains are set as: (1) $\alpha = 1.5, \tau_\eta = 100, \kappa = 0.95$ and $\beta_0 = 0.95$; (2) $\alpha = 2.5, \tau_\eta = 100, \kappa = 1.05$ and $\beta_0 = 0.95$. The BGR diagnostic relies on a statistic R, known as the shrinking factor or the potential scale reduction factor, which is estimated by calculating and comparing the between-chain variance and the within-chain variance of the two chains, defined by:

$$\hat{R} = \frac{\hat{V}_p}{\hat{V}_w} = 1 - \frac{1}{N} + \frac{\hat{V}_b}{\hat{V}_w N}\frac{M+1}{M} \tag{4.47}$$

where M is the number of generated Markov chains which in this case is equal to 2, N is the number of iterations after discarding the burn-in period samples in each chain, $\hat{V}_b$ and $\hat{V}_w$ represent the mean of variances between each chain and within

[34] BGR is one approach that recommended by CODA and BOA in R.

[35] The selection of prior distribution follows Kim et al. (1998) and recently see Takahashi et al. (2009), Abanto-Valle et al. (2010), Wang (2012), Nakajima and Omori (2012) and Abanto-Valle et al. (2015).

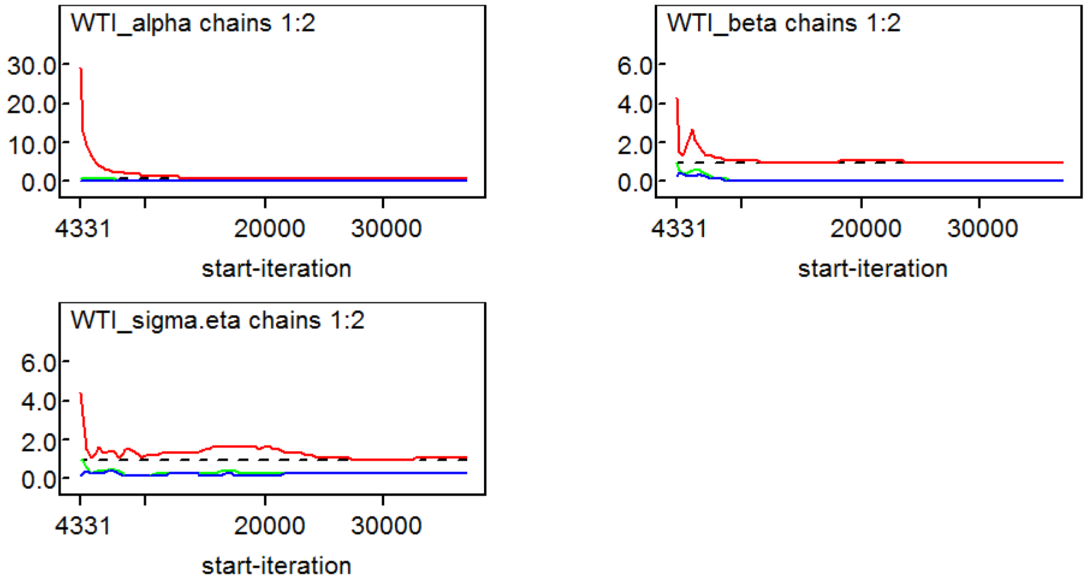

Figure 4.3: Gelman-Rubin convergence diagnostic for the parameters in SV-ALD model of WTI market

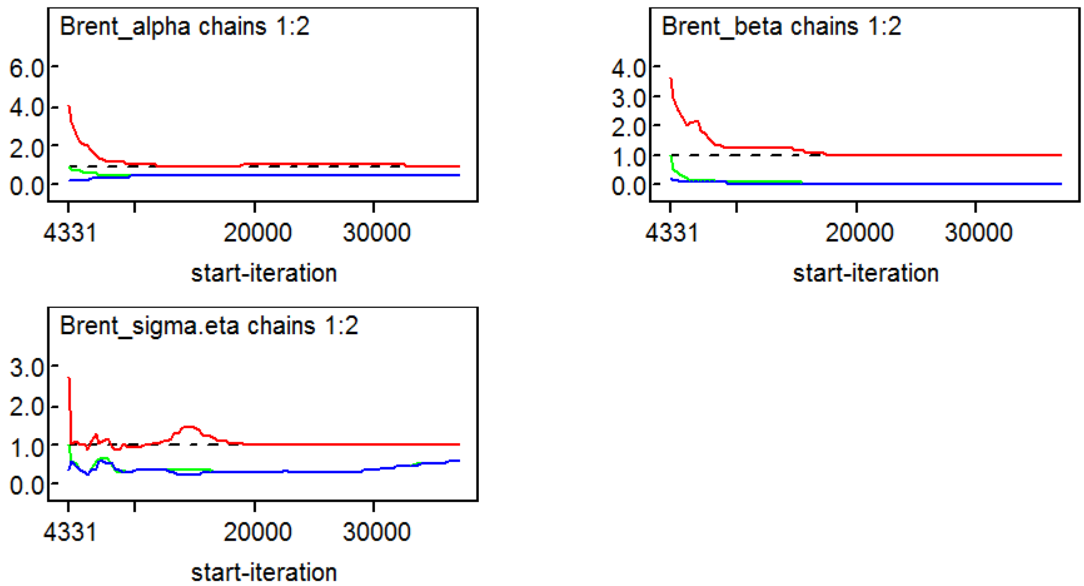

Figure 4.4: Gelman-Rubin convergence diagnostic for the parameters in SV-ALD model of Brent market

each chain respectively, and $\hat{V}_p$ is defined as:

$$\hat{V}_p = (1 - \frac{1}{N})\hat{V}_w + \frac{\hat{V}_b(1+M)}{MN} \tag{4.48}$$

representing the pooled estimate of posterior variance.[36]

The criteria for checking convergence of the Markov chain using BGR approach is to observe the variations of the estimated shrinking factor $\hat{R}$ as the number of iterations increases. When convergence is reached, values of $\hat{V}_p$ and $\hat{V}_w$ should coincide and $\hat{R}$ should be close to 1. Results of BGR diagnostic test based on WIT and Brent return series are shown in Figure 4.3 and Figure 4.4 respectively. Under the Gibbs sampling principle, two Markov chains are running and corresponding parameters (α, β and σ_η) in the SV-ALD model are simulated simultaneously with each of them drawing 60,000 times in one chain. In the WTI market, the estimated pooled posterior variance estimate $\hat{V}_p$ (green line) and within-chain variance $\hat{V}_w$ (blue line in) stabilize and the estimated shrinking factor $\hat{R}$ (red line) converges to 1 quickly for parameters α and β after drawing 15,000 times. However, the estimated $\hat{R}$ for parameter σ_η appears to be unstable between 15,000 to 25,000 draws which implies that more draws may be appropriate. On the contrary, the convergence for parameters in Brent market is faster at approximately 20000 draws.[37] Therefore, for consistency in the empirical applications, 60,000 MCMC draws are performed in one chain and the first 30,000 draws are discarded as the burn-in period in both of the two oil markets, thus remaining 2 $\times$ 30,000=60,000 simulated samples for posterior inference.

Posterior estimates and model comparison. Posterior summaries of parameters in SV-ALD model are shown in the first half of Table 4.2. To compare the performance of the SV-ALD model with classical SV model using Normal distribution for return errors, datasets for WTI and Brent are also fitted by the SV-N model with estimates listed in the second half of Table 4.2. In order to assess the performance of

[36]More details for the mathematical algorithm of BGR refer to Ntzoufras (2011).

[37]Concrete numerical estimates of $\hat{V}_p$, $\hat{V}_w$ and $\hat{R}$ for parameters are not listed here but is available upon request.

the generated posterior distribution under SV-N model, BGR convergence diagnostic are employed by running two Markov chains with 60,000 simulations each. Results suggest that the estimated shrinking factor $\hat{R}$ is able to closely converge to 1 very fast within 5,000 iterations for all relevant parameters in both the WTI and Brent markets.[38] To be consistent with SV-ALD model in terms of the iteration times, the first 30,000 draws are discarded as burn-in period and the remaining 60,000 samples are used for posterior inference.

The model comparison criterion between SV-ALD and SV-N is based on Deviance Information Criterion (DIC) (Spiegelhalter et al., 2002) which is computationally applicable and has been successfully applied to the family of SV models under Bayesian MCMC scheme (see, e.g., Berg et al., 2004; Choy and Chan, 2008; Abanto-Valle et al., 2011), the function of DIC can be defined as:

$$DIC = E[D(\ddot{\theta})] + D(E[\ddot{\theta}]) \tag{4.49}$$

where $\ddot{\theta}$ is the vector of model parameters, $D(\cdot)$ denotes the posterior deviance function, $E[D(\ddot{\theta})]$ is the posterior mean of deviance which measures the goodness-of-fit of the Bayesian model and $D(E[\ddot{\theta}])$ represents the deviance function with posterior expectation of the model parameters. The model with smallest DIC value is regarded to be superior to the other one. Although there are other model selection criteria, such as traditional AIC and BIC criteria, they are inappropriate in this case as the number of parameters to be estimated is not well-defined due to the large amount of latent variables (see, Wang, 2012). Besides, using Bayes factor as an alternative is unrealistic because it is difficult to implement especially for models with numerous unknown parameters, i.e. latent variables in SV-ALD model and models using vague/non-informative priors (Wichitaksorn et al., 2015).[39]

Results for posterior mean, standard derivation (SD), MC errors and estimated values within their 95% confidence interval (95% CI) for the SV-ALD and SV-N

[38]Plots of BGR test for parameters in SV-N model are not listed here but is available upon request.

[39]More details about computations and discussions of Bayes factor can refer to Carlin and Chib (1995) and Dellaportas et al. (2002).

Table 4.2: Posterior summary statistics of the parameters in SV-ALD and SV-N model in WTI and Brent markets

Market	Parameter	Mean	SD	MC error	95% CI	DIC	Sample
				SV-ALD			
WTI	α	-7.58700	0.48730	0.00408	(-8.42200,-6.69800)	-15898.3	60,000
	β	0.99470	0.00224	0.00005	(0.98990,0.99870)		
	σ_η	0.08891	0.00992	0.00051	(0.07065,0.10810)		
Brent	α	-7.75000	0.56200	0.00587	(-8.66000,-6.69600)	-16595.5	60,000
	β	0.99590	0.00184	0.00004	(0.99200,0.99910)		
	σ_η	0.07351	0.00812	0.00042	(0.06106,0.09410)		
				SV-N			
WTI	α	-7.87200	0.32060	0.00218	(-8.48000,-7.26900)	-12566.3	60,000
	β	0.98990	0.00378	0.00012	(0.98150,0.99640)		
	σ_η	0.12960	0.01677	0.00080	(0.10020,0.16560)		
Brent	α	-7.95400	0.49910	0.00334	(-8.76800,-7.03900)	-13028.7	60,000
	β	0.99450	0.00247	0.00007	(0.98900,0.99870)		
	σ_η	0.09329	0.01252	0.00060	(0.07317,0.1207)		

models are shown in Table 4.2. The posterior means of β in WTI and Brent markets under SV-ALD model are highly close to 1. This is consistent with our general beliefs that there exist a strong persistence of volatility in oil returns. It is worth noting that the estimated posterior mean of β in WTI and Brent markets under the SV-N model are slightly lower than their counterparts in the SV-ALD model and the estimates of σ_η for the SV-ALD model in both of the two oil markets are lower than these in the SV-N model. These results are consistent with the findings in Chib et al. (2002) and Abanto-Valle et al. (2010), which indicate that introducing an heavy-tailed error distribution (ALD) in mean equation appears to explain the excess of returns, thus decreasing the variance of volatility process. The lower posterior standard deviations of parameters β and σ_η under SV-ALD model in the two markets provide further evidence for this. In addition, MC errors, which measure the variations of parameter mean in the simulation, are relatively low compared to the corresponding estimated posterior SD for all parameters in the two competing models, thus confirming the high precision of the estimated posterior mean. Most importantly, the evidence in favor of the SV-ALD model is supported by DIC values, which for the WTI and Brent markets are -15898.3 and -16595.5 respectively, significantly lower than the values generated in the SV-N model.

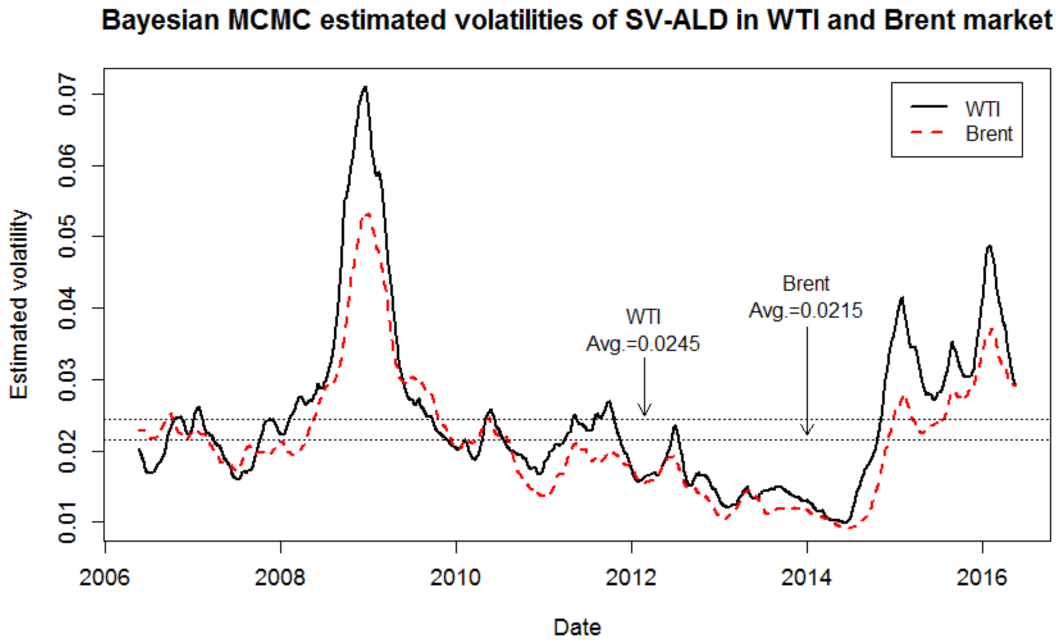

Figure 4.5: Bayesian MCMC estimation of the latent volatilities for WTI and Brent oil returns from May 2006 to May 2016

The time-varying posterior means of volatilities for the WTI and Brent oil returns based on the SV-ALD model are depicted in Figure 4.5. We can notice that the volatility estimates in the WTI and Brent oil markets produce similar trajectories in the ten-year time. Nevertheless, the ten-year average value of posterior mean of volatilities (0.0215) in the Brent market, represented by the dotted horizontal lines in Graph 4.5, is slightly lower than the average value (0.0245) in the WTI market. This is in line with the findings reported in Table 4.2 where the estimated σ_η (0.07351) for the Brent series is lower than the estimated σ_η (0.08891) in the WTI market. Noticeably, there are two large fluctuations of the estimated volatilities in both of the two major oil markets during this time: one during the global financial crisis and the other one from the second half of year 2014 characterized by the worldwide imbalance of aggregate demand and supply for oil. Those large vibrations of volatility again are a symptom for the existence of extreme risks in international crude oil markets, thus is necessary to control them.

ML estimation for κ and τ. To estimate the VaR and CVaR for oil supply and demand, it is necessary to estimate the skewness (κ) and scale (τ) statistic

Table 4.3: Maximum likelihood estimates and standard errors ($s.e.$) for the skewness and scale parameters in $AL(\kappa,\tau)$ fit to the standardized residuals

Market	$\tilde{\kappa}$	($s.e.$)	$\tilde{\tau}$	($s.e.$)
WIT	0.99562***	(0.01403)	0.97557***	(0.01944)
Brent	0.99926***	(0.01407)	0.98778***	(0.01967)

Note: $***$ denotes statistically significant of estimates with p-value less than 0.01.

of the standardized residuals. We fit the $AL(\kappa, \tau)$ distribution via the Maximum Likelihood approach for each of the standardized residual series. Results for the estimated $\tilde{\kappa}$ and $\tilde{\tau}$ in the two markets are shown in Table 4.3.

4.4.3 VaR and CVaR estimations under SV-ALD model

This part focuses on the application of the Bayesian SV-ALD model on the measurement of market risks. Figure 4.6 and 4.7 show oil returns in the WTI and Brent markets, along with the VaR estimates for oil supply and demand at different risk levels produced by the Bayesian MCMC SV-ALD model. The graphs reveal that the VaR estimates under SV-ALD model are very flexible in both the two oil markets and adjust efficiently to the fluctuations of oil returns. One quick response to the high volatility is evident during the 2008-2009 global financial crisis. Notice that some violations can be found for both oil supply and demand in the two markets, even at 1% risk level.

Figure 4.8 and 4.9 show oil returns in the two markets, superimposed by the CVaR estimates for both supply and demand at different confidence intervals. Clearly, the trends of those CVaR estimates have high similarities compared to the VaR estimates, but appear to be more dispersed in netting actual losses. Although there are some exceptions for CVaR at different tail risk levels, these are fewer than for VaR. To identify the viability of the dynamic VaR and CVaR estimates before risk managers proceed to any actual financial risk management, it is necessary to backtest the model in order to help them making appropriate decisions on when to utilize these risk tools. This process is put in place by further studying the behavior of VaR and CVaR violations over time. Results from the backtesting exercise are

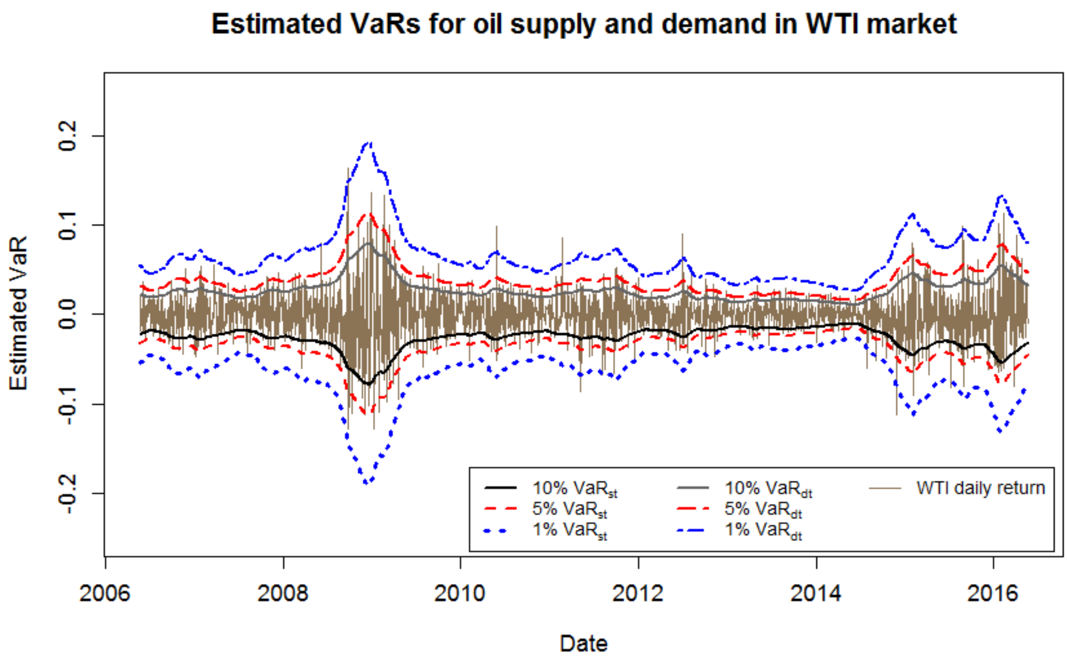

Figure 4.6: Dynamic VaR estimates for oil supply and demand in WTI market using SV-ALD model at different risk levels α

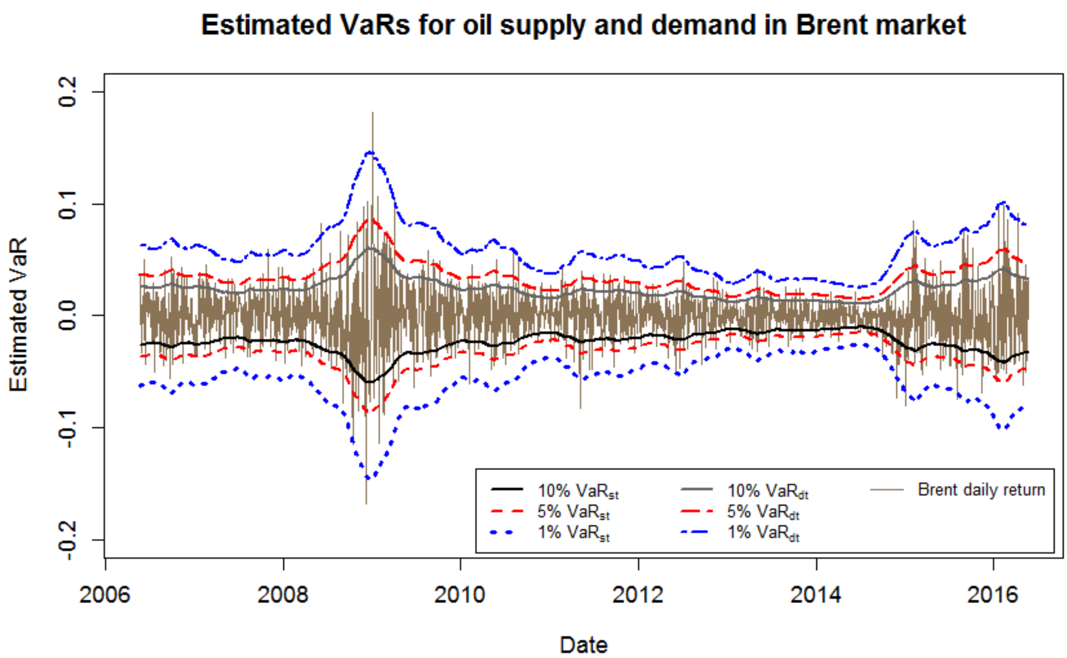

Figure 4.7: Dynamic VaR estimates for oil supply and demand in Brent market using SV-ALD model at different risk levels α

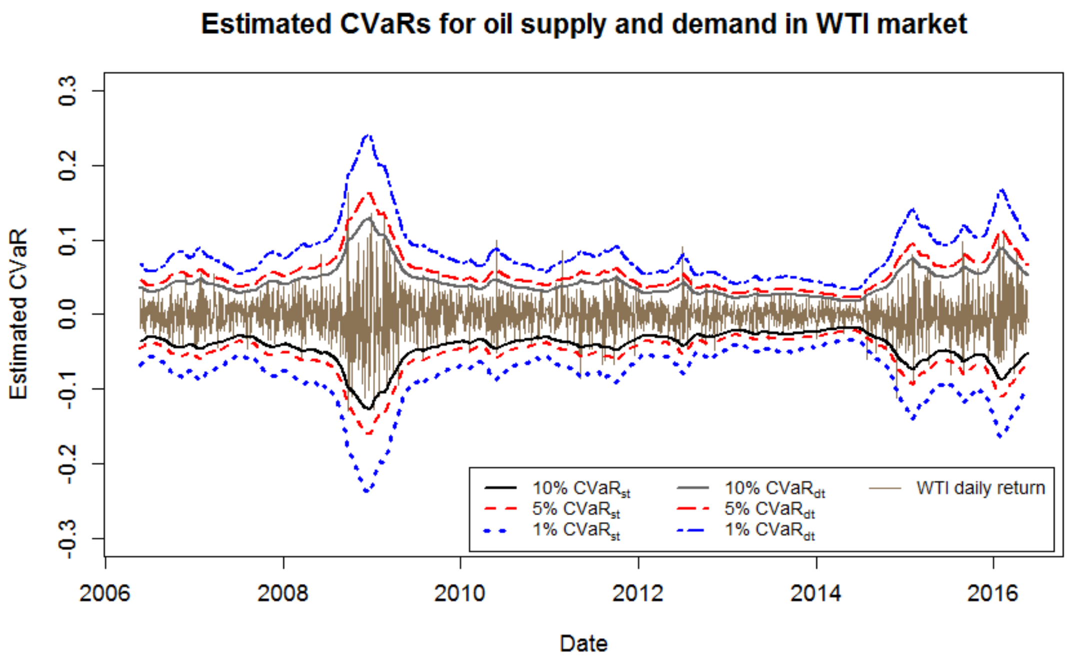

Figure 4.8: Dynamic CVaR estimates for oil supply and demand in WTI market using SV-ALD model at different risk levels α

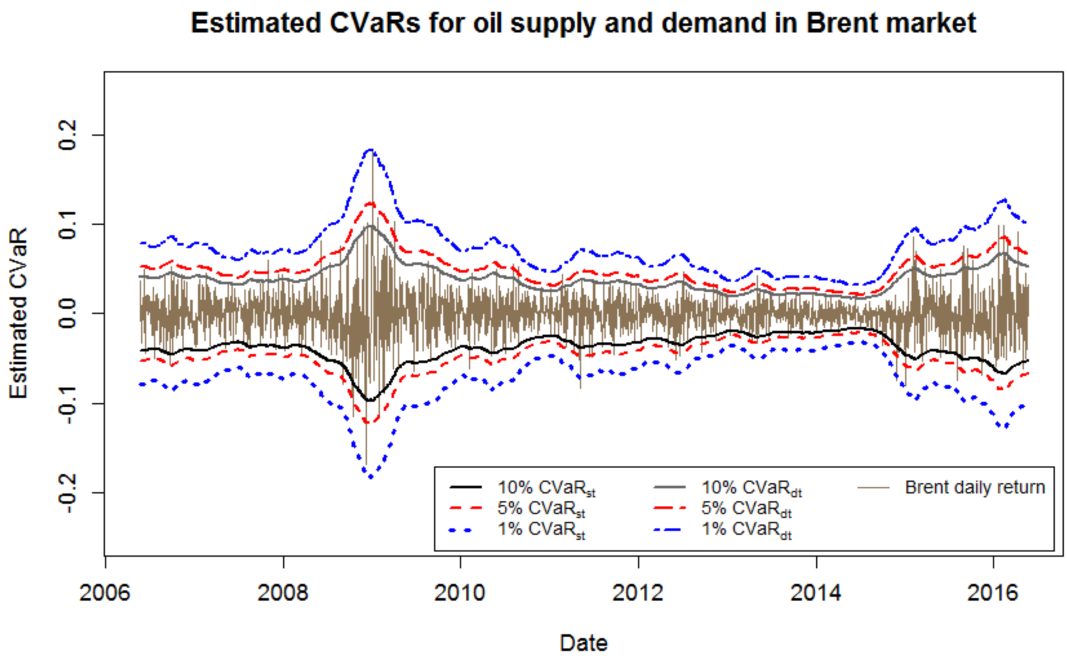

Figure 4.9: Dynamic CVaR estimates for oil supply and demand in Brent market using SV-ALD model at different risk levels α

shown in the following section.

4.4.4 Backtesting VaR model

In order to back-examine the accuracy of the estimated VaRs, we calculate the empirical failure rates for both oil supply and demand. The failure rate (FR) or violation rate, computes the ratio of the number of times oil returns exceed the estimated VaRs over the total number of observations. The model is said to be correctly specified if the calculated ratio is equal to the pre-specified VaR level α (i.e. 10%, 5% and 1%). If the ratio is greater than α, we can conclude that the model underestimates the risks, and vice versa. In our work, the failure rate $FRVaR_s$ for oil supply (or in long trading position) is calculated as the percentage of negative returns that are smaller than the left quantile VaRs, while the failure rate $FRVaR_d$ for oil demand (or in short trading positions) is the ratio of positive returns larger than the right quantile VaRs. We define $FRVaR_s$ and $FRVaR_d$ as follows:

$$FRVaR_s = \frac{1}{T}\sum_{t=1}^{T} I_t(y_t < -VaR_{s,t}) \tag{4.50}$$

$$FRVaR_d = \frac{1}{T}\sum_{t=1}^{T} I_t(y_t > VaR_{d,t}) \tag{4.51}$$

where $VaR_{s,t}$ and $VaR_{d,t}$ are the estimated VaRs for supply and demand at time t for a given confidence interval, T is the number of observations and $I_t(\cdot)$ is the indicator function which is defined as:

$$Supply : I_t = \begin{cases} 1, if y_t < -VaR_{s,t} \\ 0, if y_t \geq -VaR_{s,t} \end{cases} \tag{4.52}$$

$$Demand : I_t = \begin{cases} 1, if y_t > VaR_{d,t} \\ 0, if y_t \leq VaR_{d,t} \end{cases} \tag{4.53}$$

Furthermore, we consider three formal tests based on the above criteria to backtest the VaR estimates. The unconditional coverage test (LR_{uc}), proposed by Kupiec (1995), is to examine whether the null hypothesis $H_0 : FR = \alpha$ can be satisfied. A good performance of VaR model should be accompanied by an accurate unconditional coverage, that is, the failure rate is statistically expected to be equal to the prescribed VaR level α. The likelihood ratio statistic is given by:

$$LR_{uc} = -2\,log\left\{\alpha^N (1-\alpha)^{T-N}\right\} + 2\,log\left\{\left(\frac{N}{T}\right)^N \left(1-\frac{N}{T}\right)^{T-N}\right\} \tag{4.54}$$

where $N = \sum_{t=1}^{T} I_t$ is the number of failures in the sample size T. Under the null hypothesis, the LR_{uc} statistic is asymptotically distributed as $\chi^2(1)$.

The method proposed by Kupiec (1995) is capable to test the overestimates or underestimates of a VaR model. It does not, however, consider whether the exceptions are scattered or if they appear in clusters.[40] In order to examine whether the VaR violations are serially uncorrelated over time, Christoffersen (1998) propose the following likelihood ratio test with null hypothesis of serial independence:

$$LR_{ind} = 2\,log\left\{(1-\pi_{01})^{n_{00}}\,\pi_{01}^{n_{01}}\,(1-\pi_{11})^{n_{10}}\,\pi_{11}^{n_{11}}\right\} - 2\,log\left\{(1-\pi)^{n_{00}+n_{10}}\,\pi^{n_{01}+n_{11}}\right\} \tag{4.55}$$

where n_{ij} is the times of the transform from state i to state j for $i, j = 0, 1$, $\pi_{ij} = \frac{n_{ij}}{n_{i0}+n_{i1}}$ is the probability that state i is followed by state j, $\pi = \frac{n_{01}+n_{11}}{n_{01}+n_{11}+n_{00}+n_{10}}$ is the probability of transform to state 1. Under the null hypothesis, the LR_{ind} statistic is asymptotically distributed as $\chi^2(1)$. In addition, a more strict and elaborate conditional coverage test (LR_{cc}) to jointly examine the unconditional coverage and independence of violations has been developed by Christoffersen (1998). This test investigates if the failure rate is equal to the expected prescribed risk level and if the exceptions are independently distributed over time. Under the null hypothesis that the exceptions are independent and that the expected failure rate is equal to

[40] Kupiec's (1995) approach is an unconditional test. On the other hand, we need to conditionally examine the VaR performance under the time-varying volatility framework. A good VaR model should be able to reflect this dynamic behavior, which implies the losses exceed VaR should be independent and unpredictable.

Table 4.4: VaR backtesting results under SV-ALD and SV-N model for WTI and Brent markets at different confidence intervals

α	Risk	Failure times		Failure rate		LR_{uc}		LR_{ind}		LR_{cc}	
		WTI	Brent	WTI	Brent	WTI	Brent	WTI	Brent	WTI	Brent
						SV-ALD					
10%	$VaR_{s,t}$	250	267	9.925%	10.591%	0.8995	0.3267	0.3618	0.3317	0.5896	0.3450
	$VaR_{d,t}$	252	257	10.004%	10.194%	0.9947	0.7457	0.8635	0.9635	0.8867	0.8511
5%	$VaR_{s,t}$	81	104	3.216%	4.125%	0.0000*	0.0380*	0.4058	0.8819	0.0000*	0.1101
	$VaR_{d,t}$	87	103	3.454%	4.086%	0.0002*	0.0298*	0.2738	0.1922	0.0004*	0.0387*
1%	$VaR_{s,t}$	7	8	0.279%	0.317%	0.0000*	0.0000*	0.8434	0.8214	0.0000*	0.0003*
	$VaR_{d,t}$	5	7	0.199%	0.278%	0.0000*	0.0000*	0.8878	0.8435	0.0000*	0.0000*
						SV-N					
10%	$VaR_{N,s,t}$	235	243	9.329%	9.639%	0.2568	0.5436	0.6304	0.1466	0.4246	0.2620
	$VaR_{N,d,t}$	245	244	9.726%	9.679%	0.6454	0.5890	0.0615*	0.7086	0.1413	0.7279
5%	$VaR_{N,s,t}$	99	122	3.93%	4.839%	0.0106*	0.7099	0.9549	0.2110	0.0366*	0.4061
	$VaR_{N,d,t}$	104	119	4.129%	4.720%	0.0388*	0.5156	0.7291	0.4516	0.1068	0.5810
1%	$VaR_{N,s,t}$	19	24	0.754%	0.952%	0.1951	0.8071	0.5909	0.4969	0.3710	0.7633
	$VaR_{N,d,t}$	15	21	0.595%	0.833%	0.0273	0.3856	0.6716	0.5525	0.0796	0.5705

Note: α of 10%, 5% and 1% represent prescribed VaR level corresponding to 90%, 95% and 99% CI respectively, LR_{uc} columns show p-values of Kupiec's (1995) unconditional coverage test, LR_{ind} columns are p-values of Christofferson's (1998) independent test and LR_{cc} columns are p-values of Christofferson's (1998) conditional coverage test, * denotes significance at its corresponding risk level.

prescribed risk level, the likelihood ratio statistic is defined as:

$$LR_{cc} = -2\,log\left\{\alpha^N(1-\alpha)^{T-N}\right\} + 2\,log\left\{(1-\pi_{01})^{n_{00}}\,\pi_{01}^{n_{01}}\,(1-\pi_{11})^{n_{10}}\,\pi_{11}^{n_{11}}\right\} \tag{4.56}$$

where the LR_{cc} statistic is asymptotically $\chi^2(2)$ distributed.

Table 4.4 shows the VaR backtesting summary results under the SV-ALD and SV-N model for WTI and Brent markets considering both supply and demand risks.[41] In the SV-ALD model, it is clear from the LR_{ind} test that the null hypothesis, that the exceptions are independent, cannot be rejected at all three risk levels in the two markets, for either oil supply or demand suggesting that there are few consecutive violations, or none at all. At 10% risk level, the failure rate for oil suppler and demand in both the WTI and Brent markets are approximately equal to the pre-specified risk level, indicating the capability of the model to accurately specify the market risks. This is confirmed by the LR_{uc} test which has statistically insignificant

[41] See Appendix H for the VaR setting under SV-N model.

p-values that do not reject the null hypothesis of "FR=α". The LR_{cc} results indicate that the events of losses exceeding VaR are independent and that the expected failure rate is equal to the prescribed risk level. Moving further to 5% and 1% tail risk, both markets are estimated to have a higher level of risks for both oil supply and demand, though the failure rate in the Brent market is closer to the pre-specified risk level than that in the WTI market. The deduced p-values from the LR_{uc} and LR_{cc} tests at 5% and 1% risk level remain consistent with those findings, except for the oil supply in the Brent market at 95% confidence level. Moreover, the Bayesian SV-ALD model shows a much more rigorous performance when moving further to the extreme tails (or equivalent increases in confidence level). This is evidenced by the fact that the overestimation of risk is smaller in magnitude at the 5% risk level than at the 1% risk level. On the other hand, as a competitor, we find that the SV model with Normally distributed errors is more capable of estimating the tail risk for both supply and demand in the Brent market while it tends to overestimate risk in WTI market at 5% risk level over the study period.

It should be noticed that the obtained backtesting results of VaR under SV-ALD are similar to the findings of Chen et al. (2012), who have constructed a VaR and CVaR model by employing the Asymmetric Laplace form as error distribution for a well-known GJR-GARCH model.[42] Their outcomes suggest that the models with AL errors are the only ones that consistently estimate a higher level of risk when focusing on the extreme tail risk, i.e. at 1%. Moreover, by studying the periods of pre-financial-crisis and post-financial-crisis, the same conclusions are reached that GJR-AL models are the only consistently conservative models in risk forecasting.

4.4.5 Backtesting CVaR model

Although CVaR approach has been widely used for risk measuring, the implementation of backtesting for CVaR models is much harder than VaR models. Nevertheless, formal backtesting methods can be found in literature, such as the most commonly

[42]The employed AL density has different forms in this two papers. This paper extended the application of AL density form of Kotz et al. (2001) to a SV model while they are working on the basis of the one proposed by Yu and Zhang (2005). More discussions of ALD can be found in the introduction of this paper.

used approach zero-mean residual test by McNeil and Frey (2000) which rely on bootstrapping or one sample t principle, censored Gaussian method by Berkowitz (2001) and the functional delta approach by Kerkhol and Melenverg (2004).[43] However, applying these methods tend to be difficult and overly complex. The application of these methods is based upon the realization of specific conditions, hence only available to backtest CVaR in appropriate circumstances. Kerkhol and Melenverg (2004) suggest a viable and simpler alternative to backtest CVaR on the basis of equal quantiles, after finding a nominal risk level $\tilde{\alpha}$ that CVaR is located at. This method relies on a parametric model for calculating specific quantiles of CVaR. This is straightforward in this paper as ALD is employed and analytic CVaR expression for oil supply and demand can be easily derived. Specifically, we use the cumulative distribution function ($c.d.f$) of the ALD which, according to Kotz et al. (2001), is given by:

$$F(z|\kappa,\theta,\tau) = \begin{cases} 1 - \dfrac{1}{1+\kappa^2}\ exp\Big(-\dfrac{\sqrt{2}\kappa}{\tau}(z-\theta)\Big) & z \geq \theta \\ \dfrac{\kappa^2}{1+\kappa^2}\ exp\Big(\dfrac{\sqrt{2}}{\tau\kappa}(z-\theta)\Big) & z < \theta \end{cases} \tag{4.57}$$

Then, this $c.d.f$ is evaluated at the CVaR level. As a consequence, the probability ($\tilde{\alpha}$) that CVaR occurs under ALD for oil supply and demand can be mathematically expressed as:[44]

$$Supply: \quad \tilde{\alpha} = F(CVaR_s|\alpha) = \frac{\alpha}{e} \tag{4.58}$$

$$Demand: \quad \tilde{\alpha} = 1 - F(CVaR_d|\alpha) = \frac{\alpha}{e} \tag{4.59}$$

where e is the natural exponent. For both oil supply and demand, the quantile level of CVaR under ALD is simply a function of α and e and is irrelevant with respect to any other parameters in the AL density. This finding is consistent with the results by Chen et al. (2012). Hence, according to formula (4.58) and (4.59), the nominal risk level $\tilde{\alpha}$ for CVaR under ALD at 10%, 5% and 1% can be correspondingly obtained as 3.68%, 1.84% and 0.37%, respectively. Using $\tilde{\alpha}$ as the prescribed risk level for

[43] A comprehensive discussion of various CVaR backtesting methodologies as well as their implementations at different circumstances can refer to Wimmerstedt (2015).

[44] See Appendix I for derivation.

Table 4.5: CVaR backtesting results under SV-ALD and SV-N model for WTI and Brent markets at different confidence intervals

$\tilde{\alpha}$	Risk	Failure times		Failure rate		LR_{uc}		LR_{ind}		LR_{cc}	
		WTI	Brent	WTI	Brent	WTI	Brent	WTI	Brent	WTI	Brent
SV-ALD											
3.68%	$CVaR_{s,t}$	54	63	2.144%	2.499%	0.0000*	0.0008*	0.1239	0.6154	0.0000*	0.0033*
	$CVaR_{d,t}$	52	61	2.064%	2.420%	0.0000*	0.0003*	0.4132	0.6747	0.0000*	0.0015*
1.84%	$CVaR_{s,t}$	23	17	0.913%	0.674%	0.0001*	0.0000*	0.5149	0.6308	0.0005*	0.0000*
	$CVaR_{d,t}$	14	19	0.556%	0.754%	0.0000*	0.0000*	0.6924	0.5911	0.0000*	0.0000*
0.37%	$CVaR_{s,t}$	3	2	0.119%	0.079%	0.0155	0.0035*	0.9326	0.9551	0.0533	0.0141
	$CVaR_{d,t}$	3	0	0.119%	0.000%	0.0155	0.0000*	0.9326	1.0000	0.0533	0.0000*
SV-N											
3.96%	$CVaR_{N,s,t}$	80	100	3.176%	3.967%	0.0368*	0.9863	0.7160	0.9868	0.1025	0.9601
	$CVaR_{N,d,t}$	84	94	3.335%	3.729%	0.0983	0.5477	0.0160*	0.3658	0.0136*	0.5338
1.96%	$CVaR_{N,s,t}$	40	46	1.588%	1.825%	0.1637	0.6200	0.2558	0.1909	0.1957	0.3691
	$CVaR_{N,d,t}$	33	45	1.310%	1.785%	0.0123*	0.5199	0.3492	0.2008	0.0278	0.3523
0.38%	$CVaR_{N,s,t}$	8	9	0.318%	0.357%	0.6002	0.8496	0.8213	0.7995	0.8470	0.9476
	$CVaR_{N,d,t}$	7	9	0.278%	0.357%	0.3816	0.8496	0.8434	0.7995	0.6669	0.9476

Note: $\tilde{\alpha}$ of 3.68% (3.96%), 1.84% (1.96%) and 0.37% (0.38%) represent nominal CVaR level corresponding to 90%, 95% and 99% CI respectively in the context of SV-ALD model (SV-N model), LR_{uc} columns show p-values of Kupiec's (1995) unconditional coverage test, LR_{ind} columns are p-values of Christofferson's (1998) independent test and LR_{cc} columns are p-values of Christofferson's (1998) conditional coverage test, * denotes significance at its corresponding risk level.

CVaR backtesting, the failure rate of CVaR for oil supply and demand, denoted as $FRCVaR_s$ and $FRCVaR_d$, can be defined as follows:

$$FRCVaR_s = \frac{1}{T}\sum_{t=1}^{T} I_t(y_t < -CVaR_{s,t}) \tag{4.60}$$

$$FRCVaR_d = \frac{1}{T}\sum_{t=1}^{T} I_t(y_t > CVaR_{d,t}) \tag{4.61}$$

where the indicator function $I_t(\cdot)$ is equal to one if the condition is satisfied and zero otherwise. As a consequence, three formal statistic test LR_{uc}, LR_{ind} and LR_{cc} can be run to examine the CVaR model accuracy based on the nominal risk level $\tilde{\alpha}$.

Table 4.5 presents the CVaR backtesting results under the SV-ALD and SV-N model for oil supply and demand in the WTI and Brent markets.[45] The performance of CVaR is very similar to the performance of VaR. Looking at the failure rate in the SV-ALD framework, for example, the value for both oil supply and demand in the WTI and Brent markets is lower than the corresponding nominal risk level.

[45]See Appendix H for the CVaR setting under SV-N model.

This means they overestimate the risk and is consistent with the fact that CVaR gives conservative estimates at the other two nominal risk levels for both supply and demand in the two markets. Moreover, LR_{uc} and LR_{cc} statistic results indicate that there are no differences with the above findings at 3.68% and 1.84% risk levels. At the 0.37% risk level, however, there is a slight difference between these two markets. LR_{uc} and LR_{cc} agree that the SV-ALD model cannot be rejected to capture the dynamic risks in the WTI market for both oil supply and demand, while it is controversial in terms of oil supply in the Brent market. Results of CVaR backtesting in this paper are again consistent with the findings by Chen et al. (2012), who forecast a higher level of CVaR risks at both 5% and 1% levels when employing the AL error to the GJR-GARCH model. In comparison, we find that the CVaR model under SV-N can statistically accurately estimate the tail risks in most cases for both types of investors.

To summarize, the two risk measures are consistent in estimating WTI and Brent market risks for oil supply and demand over the ten-year study period and they both reach conservative positions in the framework of the SV-ALD model when focusing on extreme tail risks. According to LR_{uc} and LR_{cc}, VaR under SV-ALD is accurately estimated at 10% risk level, but tends to be more conservative at 5% and 1% levels in both markets. The exception is the LR_{cc} test for oil supply in the Brent market at 95% CI. In addition, the SV-ALD model overestimates the CVaR risk level at 10% and 5% in the WTI and Brent markets and at 1% in the Brent markets for oil demand, while it is statistically significant to capture the 1% tail risks in WTI market.

4.5 Conclusion

As the volatility of crude oil markets fluctuate dramatically, especially in the period of global financial crisis, it is important to implement appropriate risk measure tools that are suitable to different market participants. In this context, VaR and CVaR are employed to quantify the financial market risk embedded in oil prices in WTI and Brent markets for both oil supply and demand. We extend the AL density function (Kotz et al., 2001) to a standard discrete stochastic volatility model, to take account of the potential heavy-tailed and leptokurtic features of oil return series. This new model is called SV-ALD. Because VaR is based on the Normal distribution assumption, for assets with fat-tailed distributions, the minimization of VaR may result in a significant increase in the frequency of high losses exceeding VaR. Our model solves this problem, as we assume that the distribution of the asset returns is SV-ALD and therefore capable of modeling moments of market stress.

Considering the difficulty to evaluate the likelihood function of the SV-ALD model analytically, the Bayesian MCMC approach via the Gibbs sampling algorithm is employed for posterior inference of model parameters and latent volatilities. To facilitate an efficient Gibbs sampling, we propose a new SMU form to replace the heavy-tailed error distribution ALD, which enables the SV model to be straightforwardly implemented within the Bayesian MCMC scheme. The produced estimates of latent volatilities using this SMU representation not only promote our computation for VaR and CVaR in the SV-ALD model but are also applicable to other fields of financial modelling that related to time-varying volatilities. Regarding the empirical results, we find that the introduction of ALD in the return errors of the SV model explains the excess returns better than the SV-N model. The VaR and CVaR under SV-ALD can adjust flexibly to the fluctuations of oil returns in the two markets. From a Bayesian statistical viewpoint, the SV-ALD model is more capable to fit oil returns than the SV-N model, this is evidenced by lower DIC value in both of WTI and Brent markets.

On the basis of equal quantiles as suggested by Kerkhof and Melenberg (2004), a viable backtesting procedure of CVaR is conducted when obtaining a nominal risk level $\tilde{\alpha}$. We find that $\tilde{\alpha}$ depends only on the risk level α and is independent with respect to other parameters in the AL density, in either of the two sided tails. Consequently, three formal backtesting methods, LR_{uc}, LR_{ind} and LR_{cc} are applicable to examine model accuracy. Results from the empirical investigation of the proposed VaR and CVaR model for both oil supply and demand in the two key markets indicate that the two risk measures are consistently conservative in modelling tail risks at high confidence intervals, i.e. 95% and 99%, this indicates that the models generate higher risk measurement at a given risk exposure. Comparatively, traditional SV model with Normal distributed errors provides better tail risk measures throughout the period studied.

To conclude, we use the SV-ALD model and Bayesian simulation-based approach to estimate the dynamics of VaR and CVaR for oil prices. Backtesting results show that the SV-ALD model is able to model the tail risk for oil supply and demand, and that both VaR and CVaR overestimate the extreme tail risk leading to more conservative investment choices. As a result, we are proposing two new types of risk management tools which can be regarded as a valuable addition to the toolbox of risk management especially in the framework of stochastic volatility models and the techniques applied in this study demonstrate the feasibility to utilize more sophisticated SV-type models.

Chapter 5

Conclusion and Future Work

In this thesis, we have studied the measurement of market risk using VaR and CVaR as two risk tools by focusing on the construction and modelling in the fields of portfolio optimizations, continuous and discrete time stochastic volatility models. Our main contribution can be summarized in three aspects: the extension of existing theoretical models, improvements in the estimation methodology in Bayesian MCMC inference and practical implications.

In Chapter 2, we prove that the analytical solution of our proposed Mean-VaR-Skewness optimization model under ALD is consistent with that of Zhao et al.'s (2015) model, which can be obtained using the Lagrange Multiplier method. Our preliminary study, in the context of the Mean-VaR-Skewness model, shows the ability of ALD in describing the heavy-tailed features of stock returns. We then focus on examining the performance of Zhao et al.'s (2015) model from in-sample, Monte Carlo simulation and out-of-sample perspectives. Our results support the benefits of using ALD to model the asymmetry of asset returns. In terms of the empirical analysis under the Mean-CVaR-Skewness model, our findings indicate that in the two economic periods D1 (pre-crisis-period) and D3 (all-periods), higher expected portfolio returns are accompanied by higher risks, while the findings are ambiguous for the D2 (crisis-period). In addition, we find evidence that for a given risk level, the risk-adjusted return (measured by Sharpe-like ratio) of our portfolio is relatively low in the periods that include the financial crisis, and the outcomes are irrelevant to the chosen risk levels. Considering the applicability of CVaR for different risk levels, our results imply that aggressive investors, who may estimate a smaller risk level with their investments, have a relatively higher Sharpe-like ratio than conservative investors.

Chapter 3 has introduced a general framework for VaR and CVaR measurements using a generalized Fourier transform approach. We provide two deterministic functions: T_v and H_v, which are applicable to various financial dynamic models, especially to those distributions whose $p.d.f.$ are not known analytically and fully characterized by its characteristic function. Our numerical explorations in the context of the Heston stochastic volatility model provide answers to the key questions that arose at the beginning. We conclude that it is important to identify the number of integration points along an integral interval for the VaR and CVaR calculations in order to ensure the accuracy of risk approximations. In addition, the results from our numerical analysis indicate the presence of three key parameters: κ, ρ and τ, which could uniformly affect the movements of VaR and CVaR estimates in opposite direction; while parameter θ is the only one that could positively impact the VaR and CVaR values.

In Chapter 4, we employ VaR and CVaR to quantify the market risk embedded in oil prices in WTI and Brent markets for both oil supply and demand. We extend the AL density function (Kotz et al., 2001) to a standard discrete SV model to take into account the potential heavy-tailed and leptokurtic features in oil return series and thus construct the SV-ALD model. To overcome estimation difficulties, we propose a new SMU representation for the error distribution (ALD) to facilitate an efficient Gibbs sampling in the Bayesian MCMC framework. Regarding the empirical analysis, our results indicate that the introduction of ALD in return errors of the SV model explains the excess returns better than the SV-N model. The VaR and CVaR under SV-ALD can adjust flexibly to the fluctuations of oil returns in the two markets. From a Bayesian statistical viewpoint, the SV-ALD model is more capable to fit oil returns than the SV-N model, as evidenced by lower DIC value in both the WTI and Brent markets. In addition, the backtesting results of the estimated dynamics of VaR and CVaR show that the SV-ALD model is able to model the tail risk for both oil supply and demand, and that both VaR and CVaR overestimate the risk, leading to more conservative investment choices. As a result, we are proposing two new types of risk management tools, which could be very useful for managers who are risk averse.

There are a number of avenues for future research. Developing on Chapter 2, a future research project could look at the forecasting performance of the Mean-VaR-Skewness model under ALD, and comparative analysis with conventional VaR optimization models. Although we have preliminarily studied the portfolio VaR in an in-sample scenario, in-depth work on this model will improve the understanding of its application prospective. In addition, given the closed-form formula for single asset and multiple asset portfolios in the VaR and CVaR optimization models, it would be of interest to further investigate portfolio diversifications and offer references for practical applications in various markets. Moreover, undertaking a more extensive series of analyses by handling a large-scale portfolio consisting of a larger number of assets would be valuable for practical applications.

Developing from Chapter 3, a possible extension to this study could be to construct the VaR and CVaR models by employing other financial dynamic models, such as the Stein-Stein model (1991) and Schöbel-Zhu model (1999). Extending the present study to a valid empirical investigation could also be a constructive step in financial risk management. Finally, developing from Chapter 4, further research can be devoted to the assessment of VaR and CVaR forecasting abilities under SV-ALD scheme. In addition, it would be of interest to quantify the market risk using the SV-ALD model in comparison to those heavy-tailed GARCH-type models using oil data. Notably, the risk models that we propose under SV-ALD provide large flexibilities for a variety of financial models whose densities are non-Gaussian, hence would be applicable to various markets, including: exchange rate market, stock market and commodity market. Last, in terms of volatility forecasting accuracy, more research can aim to use high-frequency intra-day data to estimate the daily "realized volatility".

Appendix A

VaR, CVaR and Skewness of single asset under ALD

A.1 Derivation of single asset VaR

Following the mathematical definition of VaR and incorporating the negative part of $p.d.f.$ of ALD to the integration equation, VaR under ALD can be formulated as:

$$
\begin{aligned}
prob(x \leq -VaR) = & \int_{-\infty}^{-VaR} \frac{\sqrt{2}}{\tau} \frac{\kappa}{1+\kappa^2} exp(\frac{\sqrt{2}x}{\tau\kappa}) dx \\
= & \frac{\sqrt{2}}{\tau} \frac{\kappa}{1+\kappa^2} \int_{-\infty}^{-VaR} \frac{\tau\kappa}{\sqrt{2}} d(exp(\frac{\sqrt{2}x}{\tau\kappa})) \\
= & \frac{\kappa^2}{1+\kappa^2} (exp(\frac{\sqrt{2}(-VaR)}{\tau\kappa}) - exp(\frac{\sqrt{2}(-\infty)}{\kappa\tau})) \\
= & \frac{\kappa^2}{1+\kappa^2} exp(\frac{\sqrt{2}(-VaR)}{\kappa\tau}) = \alpha
\end{aligned}
$$

Rearranging the equation, we have:

$$
\begin{aligned}
& exp(\frac{\sqrt{2}(-VaR)}{\kappa\tau}) = \frac{\alpha(1+\kappa^2)}{\kappa^2} \\
\Longrightarrow & -\frac{\sqrt{2}VaR}{\tau\kappa} = ln\frac{\alpha(1+\kappa^2)}{\kappa^2} \\
\Longrightarrow & VaR = -\frac{\kappa\tau}{\sqrt{2}} ln\frac{\alpha(1+\kappa^2)}{\kappa^2}
\end{aligned}
\tag{A.1}
$$

A.2 Derivation of single asset CVaR

Suppose that we only consider the downside risk of asset return (i.e. $x < 0$), then the evolution of CVaR under ALD is shown as:

$$\begin{aligned}
CVaR_{(1-\alpha)} =& -E[X|X \leq -VaR_{(1-\alpha)}] = \frac{-\int_{-\infty}^{-VaR} x f_X(x)dx}{prob(x \leq -VaR)} \\
=& -\alpha^{-1}\int_{-\infty}^{-VaR} x f_X(x)dx = -\alpha^{-1}\int_{-\infty}^{-VaR} x\frac{\sqrt{2}}{\tau}\frac{\kappa}{1+\kappa^2}exp(\frac{\sqrt{2}x}{\kappa\tau})dx \\
=& -\frac{1}{\alpha}\frac{\sqrt{2}}{\tau}\frac{\kappa}{1+\kappa^2}\int_{-\infty}^{-VaR} x\ exp(\frac{\sqrt{2}x}{\kappa\tau})dx \\
=& -\frac{1}{\alpha}\frac{\sqrt{2}}{\tau}\frac{\kappa}{1+\kappa^2}\frac{\tau\kappa}{\sqrt{2}}\int_{-\infty}^{-VaR} x\ d(exp(\frac{\sqrt{2}x}{\kappa\tau})) \\
=& -\frac{\kappa^2}{\alpha(1+\kappa^2)}\left[exp(\frac{\sqrt{2}x}{\kappa\tau})\,x\Big|_{-\infty}^{-VaR} - \int_{-\infty}^{-VaR} exp(\frac{\sqrt{2}x}{\kappa\tau})dx\right] \\
=& -\frac{\kappa^2}{\alpha(1+\kappa^2)}\left[(-VaR)exp(\frac{\sqrt{2}(-VaR)}{\kappa\tau}) - \int_{-\infty}^{-VaR} exp(\frac{\sqrt{2}x}{\kappa\tau})dx\right] \\
=& -\frac{\kappa^2}{\alpha(1+\kappa^2)}\Bigg[(-VaR)\ exp(\frac{\sqrt{2}(-VaR)}{\kappa\tau}) \\
&-(\frac{\kappa\tau}{\sqrt{2}}exp(\frac{\sqrt{2}(-VaR)}{\kappa\tau}) - \frac{\tau\kappa}{\sqrt{2}}exp(\frac{\sqrt{2}(-\infty)}{\kappa\tau}))\Bigg] \\
=& -\frac{\kappa^2}{\alpha(1+\kappa^2)}\left[(-VaR - \frac{\tau\kappa}{\sqrt{2}})\frac{\alpha(1+\kappa^2)}{\kappa^2}\right] = VaR + \frac{\kappa\tau}{\sqrt{2}} \\
=& -\frac{\tau\kappa}{\sqrt{2}}ln(\frac{\alpha(1+\kappa^2)}{\kappa^2}) + \frac{\tau\kappa}{\sqrt{2}}
\end{aligned} \tag{A.2}$$

A.3 Derivation of single asset skewness

Expressing skewness of X as a form of non-central moment EX^3, then we have:

$$\begin{aligned} s = & E\left[\left(\frac{X-\mu}{\sigma}\right)^3\right] = \frac{EX^3 - 3\mu EX^2 + 3\mu^2 EX - \mu^3}{\sigma^3} \\ = & \frac{EX^3 - 3\mu(EX^2 - \mu EX) - \mu^3}{\sigma^3} = \frac{EX^3 - 3\mu\sigma^2 - \mu^3}{\sigma^3} \\ = & \frac{EX^3 - 3\mu(\mu^2+\tau^2) - \mu^3}{\sigma^3} = \frac{EX^3 - 4\mu^3 - 3\mu\tau^2}{\sigma^3} \end{aligned} \tag{A.3}$$

Since we have:

$$EX^3 = \int_{-\infty}^{+\infty} x^3 f_X(x)dx = \int_{-\infty}^{0} x^3 f_X(x) + \int_{0}^{+\infty} x^3 f_X(x) \tag{A.4}$$

and the $p.d.f.$ of ALD, thus the first term of equation (A.4) can be written as:

$$\begin{aligned} \int_{-\infty}^{0} x^3 f_X(x) = & \int_{-\infty}^{0} x^3 \frac{\sqrt{2}}{\tau}\frac{k}{1+k^2}\, exp(\frac{\sqrt{2}}{\tau\kappa}x)\, dx = \frac{\sqrt{2}}{\tau}\frac{k}{1+k^2}\int_{-\infty}^{0} x^3\, exp(\frac{\sqrt{2}}{\tau\kappa}x)\, dx \\ = & \frac{\sqrt{2}}{\tau}\frac{k}{1+k^2}\left[\frac{\tau\kappa}{\sqrt{2}}\int_{-\infty}^{0} x^3\, d(exp(\frac{\sqrt{2}}{\tau\kappa}x))\right] \\ = & \frac{\sqrt{2}}{\tau}\frac{k}{1+k^2}\left[\frac{\tau\kappa}{\sqrt{2}}\left(x^3\, exp(\frac{\sqrt{2}}{\tau\kappa}x)|_{-\infty}^{0} - \int_{-\infty}^{0} exp(\frac{\sqrt{2}}{\tau\kappa}x)\, 3x^2 dx\right)\right] \\ = & \frac{\sqrt{2}}{\tau}\frac{k}{1+k^2}\frac{\tau\kappa}{\sqrt{2}}\left(-3\int_{-\infty}^{0} exp(\frac{\sqrt{2}}{\tau\kappa}x)x^2 dx\right) \\ = & \frac{\kappa^2}{1+\kappa^2}\left[\frac{-3\tau\kappa}{\sqrt{2}}\left(x^2\, exp(\frac{\sqrt{2}}{\tau\kappa}x)|_{-\infty}^{0} - \int_{-\infty}^{0} exp(\frac{\sqrt{2}}{\tau\kappa}x)\, 2x\, dx\right)\right] \\ = & \frac{\kappa^2}{1+\kappa^2}\frac{6\tau\kappa}{\sqrt{2}}\int_{-\infty}^{0} exp(\frac{\sqrt{2}}{\tau\kappa}x)\, x\, dx = \frac{\kappa^2}{1+\kappa^2}\frac{6\tau\kappa}{\sqrt{2}}\frac{\tau\kappa}{\sqrt{2}}\int_{-\infty}^{0} x\, d(exp(\frac{\sqrt{2}}{\tau\kappa}x)) \\ = & \frac{\kappa^2}{1+\kappa^2}\, 3\tau^2\kappa^2\left[x\, exp(\frac{\sqrt{2}}{\tau\kappa}x)|_{-\infty}^{0} - \int_{-\infty}^{0} exp(\frac{\sqrt{2}}{\tau\kappa}x)\, dx\right] \\ = & \frac{\kappa^2}{1+\kappa^2}\, 3\tau^2\kappa^2\left[-\frac{\tau\kappa}{\sqrt{2}}\int_{-\infty}^{0} d(exp(\frac{\sqrt{2}}{\tau\kappa}x))\right] = \frac{\kappa^2}{1+\kappa^2}\, 3\tau^2\kappa^2\left[-\frac{\tau\kappa}{\sqrt{2}}(1-0)\right] \\ = & -\frac{\kappa^2}{1+\kappa^2}\, 3\tau^3\kappa^3\,\frac{1}{\sqrt{2}} \end{aligned}$$

Likewise, the second term of equation (A.4) can be reformulated as:

$$\int_0^{+\infty} x^3 f_X(x) = \int_{-\infty}^0 x^3 \frac{\sqrt{2}}{\tau} \frac{k}{1+k^2}\, exp(-\frac{\sqrt{2}\kappa}{\tau}x)\, dx = \frac{\sqrt{2}}{\tau}\frac{k}{1+k^2}\int_0^{+\infty} x^3\, exp(-\frac{\sqrt{2}\kappa}{\tau}x)\, dx$$

$$=\frac{\sqrt{2}}{\tau}\frac{k}{1+k^2}(-\frac{\tau}{\sqrt{2}\kappa})\int_0^{+\infty} x^3\, d(exp(-\frac{\sqrt{2}\kappa}{\tau}x))$$

$$=-\frac{1}{1+\kappa^2}\left[x^3\, exp(-\frac{\sqrt{2}\kappa}{\tau}x)|_0^{+\infty} - \int_0^{+\infty} exp(-\frac{\sqrt{2}\kappa}{\tau}x)\, 3x^2\, dx\right]$$

$$=\frac{3}{1+\kappa^2}\int_0^{+\infty} exp(-\frac{\sqrt{2}\kappa}{\tau}x)\, x^2\, dx$$

$$=\frac{3}{1+\kappa^2}\left(-\frac{\tau}{\sqrt{2}\kappa}\right)\int_0^{+\infty} x^2\, d(exp(-\frac{\sqrt{2}\kappa}{\tau}x))$$

$$=\frac{3}{1+\kappa^2}\left(-\frac{\tau}{\sqrt{2}\kappa}\right)\left[-\int_0^{+\infty} exp(-\frac{\sqrt{2}\kappa}{\tau}x)\, 2x\, dx\right]$$

$$=\frac{6}{1+\kappa^2}\frac{\tau}{\sqrt{2}\kappa}\int_0^{+\infty} exp(-\frac{\sqrt{2}\kappa}{\tau}x)\, x\, dx$$

$$=\frac{6}{1+\kappa^2}\frac{\tau}{\sqrt{2}\kappa}\left(-\frac{\tau}{\sqrt{2}\kappa}\right)\int_0^{+\infty} x\, d(exp(-\frac{\sqrt{2}\kappa}{\tau}x))$$

$$=\frac{6}{1+\kappa^2}\frac{\tau}{\sqrt{2}\kappa}\left(-\frac{\tau}{\sqrt{2}\kappa}\right)\left[-\int_0^{+\infty} exp(-\frac{\sqrt{2}\kappa}{\tau}x)\, dx\right]$$

$$=\frac{6}{1+\kappa^2}\frac{\tau^2}{2\kappa^2}\left(-\frac{\tau}{\sqrt{2}\kappa}\right)\left[\int_0^{+\infty} d(exp(-\frac{\sqrt{2}\kappa}{\tau}x))\right]$$

$$=\frac{3\tau^3}{\sqrt{2}(1+\kappa^2)\kappa^3}$$

Given condition $\kappa = \sqrt{2}\tau/(\mu + \sqrt{\mu^2 + 2\tau^2})$, then we have equation (A.4) as:

$$EX^3 = \frac{3\tau^3}{\sqrt{2}}\frac{(1+\kappa^4)(1-\kappa^2)}{\kappa^3} = \frac{3\tau^3}{\sqrt{2}\kappa^3} - \frac{3\tau^3}{\sqrt{2}\kappa} + \frac{3\tau^3}{\sqrt{2}}\kappa - \frac{3\tau^3}{\sqrt{2}}\kappa^3 = 6\mu^3 + 6\mu\tau^2$$

Consequently, we can express the skewness of ALD as follows:

$$s = \frac{EX^3 - 4\mu^3 - 3\mu\tau^2}{\sigma^3} = \frac{2\mu^3 + 3\mu\tau^2}{\sigma^3} = \frac{2\mu^3 + 3\mu\tau^2}{(\mu^2+\tau^2)^{3/2}}$$

Appendix B

VaR and CVaR of portfolio under multivariate ALD

B.1 Derivation of portfolio CVaR

As indicated by proposition 2.11 that portfolio return follows a one-dimensional ALD, then we are able to obtain portfolio CVaR by substituting the portfolio parameters for single asset CVaR parameters. Note that expression of CVaR for single asset involves parameter κ. Then for simplicity of the monotonicity analysis, we can first transform the single CVaR to the following form:

$$
\begin{aligned}
CVaR =& -\frac{\tau\kappa}{\sqrt{2}} ln(\frac{\alpha(1+\kappa^2)}{\kappa^2}) + \frac{\tau\kappa}{\sqrt{2}} \\
=& \frac{\tau\kappa}{\sqrt{2}} ln(\frac{\kappa^2}{\alpha(1+\kappa^2)}) + \frac{\tau\kappa}{\sqrt{2}} \\
=& \frac{\tau\kappa}{\sqrt{2}} ln(\frac{\kappa^2}{1+\kappa^2}) - \frac{\tau\kappa}{\sqrt{2}} ln\alpha + \frac{\tau\kappa}{\sqrt{2}} \\
=& \frac{\tau\kappa}{\sqrt{2}} ln(\frac{\kappa^2}{1+\kappa^2}) + \frac{\tau\kappa}{\sqrt{2}}(1 - ln\alpha) \\
=& \frac{\tau\kappa}{\sqrt{2}}(1 - ln\alpha) - \frac{\tau\kappa}{\sqrt{2}} ln(1 + \frac{1}{\kappa^2})
\end{aligned}
$$

Given $\kappa = \sqrt{2}\tau/(\mu + \sqrt{\mu^2 + 2\tau^2})$, then we have:

$$CVaR = \frac{1 - ln\alpha}{\sqrt{2}}\tau\frac{\sqrt{2}\tau}{\mu + \sqrt{\mu^2 + 2\tau^2}} - \frac{\tau}{\sqrt{2}}\frac{\sqrt{2}\tau}{\mu + \sqrt{\mu^2 + 2\tau^2}}ln(1 + \frac{1}{(\frac{\sqrt{2}\tau}{\mu + \sqrt{\mu^2 + 2\tau^2}})^2})$$

or in a simpler form:

$$CVaR = (1 - ln\alpha)g(\mu, \tau) - g(\mu, \tau)ln(2 + \frac{\mu}{g(\mu, \tau)}) \tag{B.1}$$

where the new function $g(\mu, \tau)$ is defined as $g(\mu, \tau) = \tau^2/(\mu + \sqrt{\mu^2 + 2\tau^2})$. Expressing $CVaR_W$ as the CVaR of the portfolio, and changing μ, τ in equation (B.1) by μ_W and Σ_W, respectively. The portfolio CVaR can be expressed as:

$$CVaR_W = (1 - ln\alpha)g(\mu_W, \Sigma_W) - g(\mu_W, \Sigma_W)ln(2 + \frac{\mu_W}{g(\mu_W, \Sigma_W)}) \tag{B.2}$$

where $g(\mu_W, \Sigma_W) = \Sigma_W^2/(\mu_W + \sqrt{\mu_W^2 + 2\Sigma_W^2})$.

B.2 Derivation of Portfolio VaR

To derive portfolio VaR, we first transfer the single VaR as follows:

$$VaR = -\frac{\tau\kappa}{\sqrt{2}}ln(\frac{\alpha(1 + \kappa^2)}{\kappa^2}) = -\frac{\tau\kappa}{\sqrt{2}}ln\alpha - \frac{\tau\kappa}{\sqrt{2}}ln(1 + \frac{1}{\kappa^2})$$

Substituting $\kappa = \sqrt{2}\tau/(\mu + \sqrt{\mu^2 + 2\tau^2})$ into this equation, then we have:

$$VaR = -ln(\alpha)g(\mu, \tau) - g(\mu, \tau)ln(2 + \frac{\mu}{g(\mu, \tau)}) \tag{B.3}$$

where $g(\mu, \tau) = \tau^2/(\mu + \sqrt{\mu^2 + 2\tau^2})$. Expressing VaR_W as the VaR of the portfolio, and changing μ, τ in equation (B.3) by μ_W and Σ_W, respectively. Then, the portfolio VaR can be written as:

$$VaR_W = -ln(\alpha)g(\mu_W, \Sigma_W) - g(\mu_W, \Sigma_W)ln(2 + \frac{\mu_W}{g(\mu_W, \Sigma_W)}) \tag{B.4}$$

where $g(\mu_W, \Sigma_W) = \Sigma_W^2/(\mu_W + \sqrt{\mu_W^2 + 2\Sigma_W^2})$.

Appendix C

Maximum likelihood estimates for ALD parameters

Fitting parameter κ and τ in ALD to the datasets, we can obtain the parameter estimates and corresponding standard errors ($s.e.$) of each stock via the maximum likelihood estimation approach. The maximum log-likelihood function of ALD is given by:[1]

$$lnL(\theta;\kappa;\tau) = \frac{n}{2}ln(2) - nln(\tau) + nln(\frac{k}{1+\kappa^2}) - \frac{\sqrt{2}}{\tau}M \tag{C.1}$$

where

$$M(x;\kappa;\tau) = k\sum_{i=1}^{n}(x_i-\theta)^{+} + \frac{1}{\kappa}\sum_{i=1}^{n}(x_i-\theta)^{-} \tag{C.2}$$

Table C.1 reports the estimates of stock returns among the three datasets. In D1, we can observe that there are five assets (IBM, PFE, INTC, WMT and MRK) are left skewed. This situation is more obvious in the time of financial crisis with 11 stocks shown negative skewness. The number of left-skewed stocks is smaller under D3, only 3 stocks left. Stock PFE is the only one that has exhibited negative skewness in all three datasets.

[1]More details about MLE methodology refer to Kotz et al. (2001).

Table C.1: Maximum likelihood estimates and standard errors ($s.e.$) for the skewness and scale parameters in $AL(\kappa,\tau)$ fit to the returns of the 18 stocks under three datasets

Number	Ticker	D1				D2				D3			
		κ	$(s.e)$	τ	$(s.e)$	κ	$(s.e)$	τ	$(s.e)$	κ	$(s.e)$	τ	$(s.e)$
1	AAPL	0.9579	0.0191	2.6673	0.0752	0.9878	0.0305	3.2013	0.1399	0.9610	0.0129	2.4955	0.0474
2	XOM	0.9709	0.0194	1.4876	0.0419	1.0105	0.0312	2.3651	0.1033	0.9836	0.0132	1.6013	0.0304
3	CVX	0.9747	0.0194	1.4603	0.0412	1.0080	0.0311	2.6031	0.1137	0.9813	0.0132	1.6871	0.0321
4	MSFT	0.9985	0.0199	1.6229	0.0457	1.0036	0.0310	2.6559	0.1160	0.9991	0.0134	1.7545	0.0333
5	IBM	1.0066	0.0201	1.4842	0.0418	0.9961	0.0308	2.1026	0.0919	0.9897	0.0133	1.5033	0.0286
6	GE	0.9978	0.0199	1.5077	0.0425	1.0315	0.0319	3.2616	0.1425	1.0039	0.0135	1.8782	0.0357
7	JNJ	0.9905	0.0197	1.1975	0.0338	0.9955	0.0308	1.3484	0.0589	0.9894	0.0133	1.1011	0.0209
8	PFE	1.0107	0.0201	1.6383	0.0462	1.0181	0.0315	2.1604	0.0944	1.0010	0.0135	1.6139	0.0307
9	PG	0.9697	0.0193	1.0781	0.0304	1.0056	0.0311	1.6642	0.0727	0.9816	0.0132	1.1246	0.0214
10	INTC	1.0091	0.0201	2.4423	0.0688	1.0101	0.0312	2.9243	0.1278	1.0022	0.0135	2.2638	0.0430
11	CSCO	0.9903	0.0197	2.3769	0.0670	1.0151	0.0314	2.7525	0.1202	0.9986	0.0134	2.2166	0.0421
12	KO	0.9934	0.0198	1.1406	0.0321	0.9951	0.0307	1.8014	0.0787	0.9850	0.0132	1.2094	0.0230
13	WMT	1.0068	0.0201	1.4231	0.0401	0.9885	0.0305	1.8173	0.0794	0.9935	0.0133	1.3622	0.0259
14	ORCL	0.9951	0.0198	2.4568	0.0692	0.9984	0.0308	2.7134	0.1185	0.9897	0.0133	2.2545	0.0428
15	ABT	0.9956	0.0198	1.5478	0.0436	0.9993	0.0309	1.8124	0.0792	0.9893	0.0133	1.4254	0.0271
16	MRK	1.0021	0.0200	1.6269	0.0459	1.0188	0.0315	2.6092	0.1140	0.9980	0.0134	1.7081	0.0325
17	T	0.9950	0.0198	1.6870	0.0475	1.0167	0.0314	2.3335	0.1019	0.9935	0.0134	1.6153	0.0307
18	COP	0.9676	0.0193	1.6997	0.0479	1.0242	0.0316	3.0068	0.1314	0.9833	0.0132	1.9181	0.0365

Appendix D

VaR and CVaR derivations in a generalized Fourier transform framework

D.1 VaR derivation in a Fourier space

Given the generalized Fourier transformation of stochastic process X, we substitute the corresponding *p.d.f.* (equation 3.9) into the original integral formula of VaR (equation 3.6), then we have:

$$
\begin{aligned}
\alpha =& \frac{1}{2\pi}\int_{iv-\infty}^{iv+\infty}\psi(z)dz\int_{-\infty}^{-VaR}e^{-izx}dx \\
=& \frac{1}{2\pi}\int_{iv-\infty}^{iv+\infty}\psi(z)dz(-\frac{1}{iz})\int_{-\infty}^{-VaR}d(e^{-izx}) \\
=& \frac{1}{2\pi}\int_{iv-\infty}^{iv+\infty}\psi(z)dz\left[(-\frac{1}{iz})e^{izVaR}-0\right] \\
=& \frac{1}{2\pi}\int_{iv-\infty}^{iv+\infty}\psi(z)dz\left(-\frac{e^{izVaR}}{iz}\right)
\end{aligned}
\tag{D.1}
$$

In order to ensure the convergence of the second integral, we follow Bormetti et al. (2010) to define the imaginary part v to be greater than 0. Given conditions $z = w + iv$ and $iv - \infty < z = w + iv < iv + \infty$, we have $-iz = -i(w + iv) = -iw - i^2v = v - iw$ with $-\infty < w < +\infty$.

As a result, equation (D.1) becomes:

$$
\begin{aligned}
\alpha =& \frac{1}{2\pi}\int_{-\infty}^{+\infty}\psi(w+iv)dw\left(\frac{e^{iwVaR}\,e^{-vVaR}}{v-iw}\right)\\
=& \frac{e^{-vVaR}}{2\pi}\int_{-\infty}^{+\infty}\psi(w+iv)dw\left(\frac{e^{iwVaR}}{v-iw}\right)
\end{aligned}
\tag{D.2}
$$

D.2 CVaR derivation in a Fourier space

To derive the CVaR in a Fourier space, we plug the *p.d.f.* (equation 3.9) in terms of the generalized Fourier transform $\psi(z)$ in equation (3.12), then the CVaR is established as follows:

$$
\begin{aligned}
CVaR =& \frac{1}{2\pi}\int_{iv-\infty}^{iv+\infty}\psi(z)dz\ (-\frac{1}{\alpha})\int_{-\infty}^{-VaR}xe^{-izx}dx\\
=& -\frac{1}{2\pi\alpha}\int_{iv-\infty}^{iv+\infty}\psi(z)dz\ (-\frac{1}{iz})\int_{-\infty}^{-VaR}xd(e^{-izx})\\
=& -\frac{1}{2\pi\alpha}\int_{iv-\infty}^{iv+\infty}\psi(z)dz\ (-\frac{1}{iz})\left[xe^{-izx}\big|_{-\infty}^{-VaR}-\int_{-\infty}^{-VaR}e^{-izx}dx\right]\\
=& -\frac{1}{2\pi\alpha}\int_{iv-\infty}^{iv+\infty}\psi(z)dz\ (-\frac{1}{iz})\left[-VaR\,e^{izVaR}-0-\int_{-\infty}^{-VaR}e^{-izx}dx\right]\\
=& -\frac{1}{2\pi\alpha}\int_{iv-\infty}^{iv+\infty}\psi(z)dz\ (\frac{1}{iz})\left[VaR\,e^{izVaR}+(-\frac{1}{iz})\int_{-\infty}^{-VaR}d(e^{-izx})\right]\\
=& -\frac{1}{2\pi\alpha}\int_{iv-\infty}^{iv+\infty}\psi(z)dz\ (\frac{1}{iz})\left[VaR\,e^{izVaR}-\frac{1}{iz}(e^{izVaR}-0)\right]\\
=& -\frac{1}{2\pi\alpha}\int_{iv-\infty}^{iv+\infty}\psi(z)dz\ \frac{e^{izVaR}}{iz}\left(VaR-\frac{1}{iz}\right)
\end{aligned}
\tag{D.3}
$$

Given conditions $z = w + iv$ and $iv - \infty < z = w + iv < iv + \infty$, then we have $-iz = -i(w+iv) = -iw - i^2v = v - iw$ with $-\infty < w < +\infty$. This provides a new transformation of equation (D.3) as follows:

$$
\begin{aligned}
CVaR =& -\frac{1}{2\pi\alpha}\int_{-\infty}^{+\infty}\psi(w+iv)dw\ \frac{e^{iwVaR}\,e^{-vVaR}}{iw-v}\left(VaR-\frac{1}{iw-v}\right)\\
=& \frac{e^{-vVaR}}{2\pi\alpha}\int_{-\infty}^{+\infty}\psi(w+iv)dw\ \frac{e^{iwVaR}}{v-iw}\left(VaR+\frac{1}{v-iw}\right)
\end{aligned}
\tag{D.4}
$$

D.3 Simplified transform of VaR and CVaR

Considering the symmetric property of the real and imaginary parts of the ECF $\psi(z)$, then we are able to rewrite the first equation in (3.15) as follows:

$$\alpha = \frac{e^{-vVaR}}{\pi} Re\left[\int_0^{+\infty} \psi(w+iv)dw \left(\frac{e^{iwVaR}}{v-iw}\right)\right] \tag{D.5}$$

Thus, we have:

$$\alpha = \frac{Re\, T_v(VaR, v)}{\pi} \tag{D.6}$$

where

$$T_v(VaR, v) = e^{-vVaR}\left[\int_0^{+\infty} \frac{\psi(w+iv)}{v-iw}\, e^{iwVaR}\, dw\right]$$

Accordingly, the expression of CVaR in equation (3.15) can be reformulated as:

$$\begin{aligned}
CVaR =& \frac{e^{-vVaR}}{2\pi\alpha}\int_{-\infty}^{+\infty} \frac{\psi(w+iv)}{v-iw}\, e^{iwVaR}\, VaR\, dw \\
&+ \frac{e^{-vVaR}}{2\pi\alpha}\int_{-\infty}^{+\infty} \frac{\psi(w+iv)}{v-iw}\, e^{iwVaR}\, \frac{1}{v-iw} dw \\
=& \frac{e^{-vVaR}}{2Re\, T_v(VaR, v)}\int_{-\infty}^{+\infty} \frac{\psi(w+iv)}{v-iw}\, e^{iwVaR}\, VaR\, dw \\
&+ \frac{e^{-vVaR}}{2\pi\alpha}\int_{-\infty}^{+\infty} \frac{\psi(w+iv)}{v-iw}\, e^{iwVaR}\, \frac{1}{v-iw} dw \\
=& \frac{VaR\, Re\, T_v(VaR, v)}{Re\, T_v(VaR, v)} + \frac{e^{-vVaR}}{Re\, T_v(VaR, v)} Re\left[\int_0^{+\infty} \frac{\psi(w+iv)}{(v-iw)^2}\, e^{iwVaR}\, dw\right] \\
=& VaR + \frac{Re\, H_v(VaR, v)}{Re\, T_v(VaR, v)}
\end{aligned} \tag{D.7}$$

where

$$H_v(VaR, v) = e^{-vVaR}\left[\int_0^{+\infty} \frac{\psi(w+iv)}{(v-iw)^2}\, e^{iwVaR}\, dw\right]$$

D.4 Real part of function T_V

We define a new function TT_v representing the real part of function T_v. Following Definition 3.2, we have:

$$TT_v(VaR, v) = \mathbf{Re}\ T_v(VaR, v) = \mathbf{Re}\left\{e^{-vVaR}\left[\int_0^{+\infty} \frac{\psi(w+iv)}{v-iw}\ e^{iwVaR}\ dw\right]\right\}$$
$$= \mathbf{Re}\left\{e^{-vVaR}\left[\int_0^{+\infty} \frac{\psi(w+iv)\,cos(wVaR) + \psi(w+iv)\,i\,sin(wVaR)}{v-iw} dw\right]\right\} \tag{D.8}$$

Since e^{-vVaR} is simply a real number for an admission value v, thus by applying the Remark 3.19, we can rewrite the integrand in $(D.8)$ as follows:[1]

$$\mathbf{Re}\,\frac{\psi(w+iv)\,cos(wVaR) + \psi(w+iv)\,i\,sin(wVaR)}{v-iw}$$
$$= \mathbf{Re}\,\frac{v\,\psi(w+iv)\,cos(wVaR) + \psi(w+iv)\,sin(wVaR)\,(-w)}{v^2+w^2}$$
$$+ \mathbf{Re}\,\frac{[v\,\psi(w+iv)\,sin(wVaR) + w\,\psi(w+iv)\,cos(wVaR)]\ i}{v^2+w^2}$$
$$= \frac{v\,cos(wVaR)\,\mathbf{Re}\,\psi(w+iv) - w\,sin(wVaR)\,\mathbf{Re}\,\psi(w+iv)}{v^2+w^2}$$
$$+ \frac{-v\,sin(wVaR)\,\mathbf{Im}\,\psi(w+iv)\ - w\,cos(wVaR)\,\mathbf{Im}\,\psi(w+iv)}{v^2+w^2}$$
$$= \frac{cos(wVaR)\,[v\,\mathbf{Re}\,\psi(w+iv) - w\mathbf{Im}\,\psi(w+iv)]}{v^2+w^2}$$
$$- \frac{sin(wVaR)\,[w\,\mathbf{Re}\,\psi(w+iv) + v\,\mathbf{Im}\,\psi(w+iv)]}{v^2+w^2}$$

Substituting this expression into equation (D.8) yield:

$$TT_v(VaR, v) = \mathbf{Re}\ T_v(VaR, v)$$
$$= e^{-vVaR}\int_0^{+\infty} \frac{dw}{v^2+w^2}\left\{cos(wVaR)\,[v\,\mathbf{Re}\,\psi(w+iv) - w\mathbf{Im}\,\psi(w+iv)]\right\} \tag{D.9}$$
$$- e^{-vVaR}\int_0^{+\infty} \frac{dw}{v^2+w^2}\left\{sin(wVaR)\,[w\,\mathbf{Re}\,\psi(w+iv) + v\,\mathbf{Im}\,\psi(w+iv)]\right\}$$

[1] The application of $Im(\cdot)$ in the derivation process is based on the hint that if $f = a+bi$, then $fi = ai - b$, since $Im(f) = b$, thus we have $Re(fi) = -b = -Im(f)$.

D.5 Real part of function H_v

Analogously to TT_v, we are able to define another new function HH_v, which represents the real part of function H_v. Given Definition 3.2, we have:

$$HH_v(VaR, v) = \mathbf{Re}\ H_v(VaR, v) = \mathbf{Re} \left\{ e^{-vVaR} \left[\int_0^{+\infty} \frac{\psi(w+iv)}{(v-iw)^2}\ e^{iwVaR}\ dw \right] \right\}$$
$$= \mathbf{Re} \left\{ e^{-vVaR} \left[\int_0^{+\infty} \frac{\psi(w+iv)\, cos(wVaR) + \psi(w+iv)\, i\, sin(wVaR)}{v^2 - w^2 - 2vwi} dw \right] \right\} \tag{D.10}$$

By applying the Remark 3.19 to the integrand of $(D.10)$, we have:

$$\mathbf{Re}\, \frac{\psi(w+iv)\, cos(wVaR) + \psi(w+iv)\, i\, sin(wVaR)}{v^2 - w^2 - 2vwi}$$
$$= \mathbf{Re}\, \frac{(v^2-w^2)\, \psi(w+iv)\, cos(wVaR) + \psi(w+iv)\, sin(wVaR)\, (-2vw)}{(v^2-w^2)^2 + 4v^2w^2}$$
$$+ \mathbf{Re}\, \frac{[(v^2-w^2)\, \psi(w+iv)\, sin(wVaR) + 2vw\, \psi(w+iv)\, cos(wVaR)]\ i}{(v^2-w^2)^2 + 4v^2w^2}$$
$$= \frac{(v^2-w^2)\, cos(wVaR)\, \mathbf{Re}\, \psi(w+iv) - 2vw\, sin(wVaR)\, \mathbf{Re}\, \psi(w+iv)}{(v^2+w^2)^2}$$
$$+ \frac{-(v^2-w^2)\, sin(wVaR)\, \mathbf{Im}\, \psi(w+iv)\ - 2vw\, cos(wVaR)\, \mathbf{Im}\, \psi(w+iv)}{(v^2+w^2)^2}$$
$$= \frac{cos(wVaR)\, [(v^2-w^2)\, \mathbf{Re}\, \psi(w+iv) - 2vw\, \mathbf{Im}\, \psi(w+iv)]}{(v^2+w^2)^2}$$
$$- \frac{sin(wVaR)\, [2vw\, \mathbf{Re}\, \psi(w+iv) + (v^2-w^2)\, \mathbf{Im}\, \psi(w+iv)]}{(v^2+w^2)^2}$$

As a consequence, it is not difficult to obtain the following function if we substitute the above expression into equation (D.10):

$$HH_v(VaR, v) = \mathbf{Re}\ H_v(VaR, v)$$
$$= e^{-vVaR} \int_0^{+\infty} \frac{dw}{(v^2+w^2)^2} \left\{ cos(wVaR) \left[(v^2-w^2)\, \mathbf{Re}\, \psi(w+iv) - 2vw\mathbf{Im}\, \psi(w+iv) \right] \right\}$$
$$- e^{-vVaR} \int_0^{+\infty} \frac{dw}{(v^2+w^2)^2} \left\{ sin(wVaR) \left[2vw\, \mathbf{Re}\, \psi(w+iv) + (v^2-w^2)\, \mathbf{Im}\, \psi(w+iv) \right] \right\} \tag{D.11}$$

Appendix E

Relevant proofs in Heston model

E.1 Application of Itô's lemma

The Itô's lemma is an important foundmental tool in stochastic calculus. It acts like the chain rule in solving the PDEs and the application of Itô's lemma can simplify the SDEs. Suppose the asset prices process $\{S_t, t \geq 0\}$ satisfies the SDE given in equation (3.23). Applying Itô's lemma to $f(t, S_t) = lnS_t$, where $f(\cdot)$ depends on the particulars of t and S_t, we can specify the dynamics of this process as follows:

$$\begin{aligned}
f(t,S_t) =& f(0,S_0) + \int_0^t \frac{\partial f(s,S_s)}{\partial s} ds + \int_0^t \frac{\partial f(s,S_s)}{\partial S_s} dS_s + \frac{1}{2}\int_0^t \frac{\partial^2 f(s,S_s)}{\partial S_s^2} V_s S_s^2 ds \\
ln(S_t) =& ln(S_0) + 0 + (\int_0^t \frac{1}{S_s}\mu S_s ds + \int_0^t \frac{1}{S_s}\sqrt{V_s} S_s dW_s^S) - \frac{1}{2}\int_0^t \frac{1}{S_s^2} V_s S_s^2 ds \\
ln(S_t) =& ln(S_0) + \mu t + \sqrt{V_t} W_t^S - \frac{1}{2} V_t t \\
ln(\frac{S_t}{S_0}) =& (\mu - \frac{1}{2}V_t)t + \sqrt{V_t} W_t^S
\end{aligned} \tag{E.1}$$

where $f(t, S_t)$ is a twice-differentiable scalar function of two real variables t and S.

E.2 Derivation of Heston PDE

Given the algorithm of PDE defined in Theorem 3.31 and Remark 3.5, and the SDEs of Heston model (formula 3.28 and 3.29), we can derive the PDE in terms of log-return $X_t = lnS_t$ as follows:[1]

$$\begin{aligned}
df(t,X_t,V_t) =& \dot{f}(t,X_t,V_t)d_t + f_1(t,X_t,V_t)dX_t + f_2(t,X_t,V_t)dV_t \\
&+ \frac{1}{2}f_{11}(t,X_t,V_t)d\langle X\rangle_t + \frac{1}{2}f_{22}(t,X_t,V_t)d\langle V\rangle_t + f_{12}(t,X_t,V_t)d\langle X,V\rangle_t \\
=& \frac{\partial f(t,X_t,V_t)}{\partial t}dt + \frac{\partial f(t,X_t,V_t)}{\partial X_t}dX_t + \frac{\partial f(t,X_t,V_t)}{\partial V_t}dV_t \\
&+ \frac{1}{2}\frac{\partial^2 f(t,X_t,V_t)}{\partial X_t^2}V_t dt + \frac{1}{2}\frac{\partial^2 f(t,X_t,V_t)}{\partial V_t^2}\xi^2 V_t dt + \xi V_t\rho\frac{\partial^2 f(t,X_t,V_t)}{\partial V_t \partial X_t}dt \\
=& \frac{\partial f(t,X_t,V_t)}{\partial t}d_t + \frac{\partial f(t,X_t,V_t)}{\partial X_t}\left((\mu - \frac{1}{2}V_t)dt + \sqrt{V_t}dW_t^S\right) \\
&+ \frac{\partial f(t,X_t,V_t)}{\partial V_t}\left(\kappa(\theta - V_t)d_t + \xi\sqrt{V_t}dW_t^V\right) + \frac{1}{2}\frac{\partial^2 f(t,X_t,V_t)}{\partial X_t^2}V_t dt \\
&+ \frac{1}{2}\frac{\partial^2 f(t,X_t,V_t)}{\partial V_t^2}\xi^2 V_t dt + \xi V_t\rho\frac{\partial^2 f(t,X_t,V_t)}{\partial V_t \partial X_t}dt \\
=& \frac{\partial f(t,X_t,V_t)}{\partial X_t}\sqrt{V_t}dW_t^S + \frac{\partial f(t,X_t,V_t)}{\partial V_t}\xi\sqrt{V_t}dW_t^V + \mathcal{A}f(t,X_t,V_t)dt
\end{aligned} \tag{E.2}$$

where $\mathcal{A}$ represents differential generator defined as:

$$\begin{aligned}
(\mathcal{A}f)(t,X_t,V_t) =& \frac{\partial f(t,X_t,V_t)}{\partial t} + (\mu - \frac{1}{2}V_t)\frac{\partial f(t,X_t,V_t)}{\partial X_t} + \kappa(\theta - V_t)\frac{\partial f(t,X_t,V_t)}{\partial V_t} \\
&+ \frac{V_t}{2}\frac{\partial^2 f(t,X_t,V_t)}{\partial X_t^2} + \frac{\xi^2 V_t}{2}\frac{\partial^2 f(t,X_t,V_t)}{\partial V_t^2} + \xi V_t\rho\frac{\partial^2 f(t,X_t,V_t)}{\partial V_t \partial X_t}
\end{aligned} \tag{E.3}$$

Time evolution function $f(t,X_t,V_t)$ must be a martingale by iterated expectations.[2] Given $E[df] = 0$, and substitute it to equation (E.2) yields a new PDE:

$$\begin{aligned}
&\frac{\partial f(t,X_t,V_t)}{\partial t} + (\mu - \frac{1}{2}V_t)\frac{\partial f(t,X_t,V_t)}{\partial X_t} + \kappa(\theta - V_t)\frac{\partial f(t,X_t,V_t)}{\partial V_t} \\
&+ \frac{V_t}{2}\frac{\partial^2 f(t,X_t,V_t)}{\partial X_t^2} + \frac{\xi^2 V_t}{2}\frac{\partial^2 f(t,X_t,V_t)}{\partial V_t^2} + \xi V_t\rho\frac{\partial^2 f(t,X_t,V_t)}{\partial V_t \partial X_t} = 0
\end{aligned} \tag{E.4}$$

[1]Note that in the derivation, a series of simple rules in the statement of Itô's lemma are applied. All terms with $(dt)^2$, $dt\,dW_t^S$ and $dt\,dW_t^V$ are deleted as a consequence of the limit $dt \to 0$, and terms with $(dW_t^S)^2$ and $(dW_t^V)^2$ are substituted by dt.

[2]$f(t,X_t,V_t)$ is a martingale if and only if the differential generator $\mathcal{A}$ satisfies a condition that the PDE $\mathcal{A}f(t,X_t,V_t) = 0$.

E.3 Reduced form of ODEs

As indicated by Feynman-Kac representation theorem, the ECF (equation 3.33) of Heston dynamics follows the PDE in equation (E.4). That is, if a function $f(t, X_t, V_t)$ of the Heston SEDs satisfies the PDE defined in equation (E.4), then the solution of function $f(t, X_t, V_t)$ is the conditional expectation of $u(X_t) = e^{izX_t}$.

To evaluate the PDE (equation E.4) of the ECF of Heston (equation 3.33), following derivatives are required:

$$\frac{\partial f(t,X_t,V_t)}{\partial t} = \left(\frac{\partial C(\tau)}{\partial \tau} + \frac{\partial D(\tau)}{\partial \tau}V_t\right) f(t,X_t,V_t), \qquad \frac{\partial f(t,X_t,V_t)}{\partial X_t} = izf(t,X_t,V_t)$$

$$\frac{\partial f(t,X_t,V_t)}{\partial V_t} = D(\tau)f(t,X_t,V_t), \qquad \frac{\partial^2 f(t,X_t,V_t)}{\partial X_t^2} = -z^2 f(t,X_t,V_t)$$

$$\frac{\partial^2 f(t,X_t,V_t)}{\partial V_t^2} = D^2(\tau)f(t,X_t,V_t), \qquad \frac{\partial^2 f(t,X_t,V_t)}{\partial V_t \partial X_t} = izD(\tau)f(t,X_t,V_t)$$

Substituting those derivatives to the PDE (equation E.4) and cancel out the common term $f(t, X_t, V_t)$ yields:

$$\left(\frac{\partial D(\tau)}{\partial \tau} + iz\xi\rho D(\tau) + \frac{1}{2}\xi^2 D^2(\tau) - \frac{1}{2}z^2 - qD(\tau) + izp\right) V_t + \frac{\partial C(\tau)}{\partial \tau} + iz\mu + aD(\tau) = 0$$

where $p = -\frac{1}{2}$, $a = k\theta$ and $q = \kappa$.

This equality can be statisfied if and only if the coefficient part of V_t equals to 0 and the constant term to be 0, then it can be hierarchically written as two ordinary differential equations (ODEs):

$$\frac{\partial D(\tau)}{\partial \tau} + iz\xi\rho D(\tau) + \frac{1}{2}\xi^2 D^2(\tau) - \frac{1}{2}z^2 - qD(\tau) + izp = 0 \tag{E.5}$$

and

$$\frac{\partial C(\tau)}{\partial \tau} + iz\mu + aD(\tau) = 0 \tag{E.6}$$

E.4 Closed-form solution of parameters in ODEs

Let us write the ODE (equation 3.34) in a form of non-linear Riccati equation:

$$\frac{\partial D(\tau)}{\partial \tau} = RD^2(\tau) - QD(\tau) + P \tag{E.7}$$

where $R = \frac{1}{2}\xi^2$, $Q = q - iz\xi\rho$ and $P = izp - \frac{1}{2}z^2$.

This equaton can be solved by considering the following second-order ODE:

$$w'' + Qw' + PRw = 0 \tag{E.8}$$

which can also be written as $w'' + bw' + c = 0$. One solution of equation (E.8) implies a solution of equation (E.7) which can be expressed as:

$$D(\tau) = -\frac{w'}{w}\frac{1}{R} \tag{E.9}$$

According to the auxiliary equation, two solutions (α and β) of equation (E.8) are shown as:

$$\alpha = \frac{-b + \sqrt{b^2 - 4c}}{2} = \frac{-Q + \sqrt{Q^2 - 4PR}}{2} = \frac{-Q + d}{2} \tag{E.10}$$

$$\beta = \frac{-b - \sqrt{b^2 - 4c}}{2} = \frac{-Q - \sqrt{Q^2 - 4PR}}{2} = \frac{-Q - d}{2} \tag{E.11}$$

where

$$d = \alpha - \beta = \sqrt{Q^2 - 4PR} = \sqrt{(q - iz\xi\rho)^2 - \xi^2(2izp - z^2)} \tag{E.12}$$

The solution to the second-order ODE in equation (E.8) is formulated as:

$$w = Me^{\alpha\tau} + Ne^{\beta\tau} \tag{E.13}$$

where M and N are two constant, $\tau = T - t$ denotes the time to maturity. Therefore, the solution of the non-linear Riccati equation (E.7) is given by:

$$D(\tau) = -\frac{M\alpha e^{\alpha\tau} + N\beta e^{\beta\tau}}{Me^{\alpha\tau} + Ne^{\beta\tau}}\frac{1}{R} = -\frac{1}{R}\frac{K\alpha e^{\alpha\tau} + \beta e^{\beta\tau}}{Ke^{\alpha\tau} + e^{\beta\tau}} \tag{E.14}$$

where $K = \frac{M}{N}$. Given the initial condition $D(0) = 0$ at maturity T, the numerator becomes $K\alpha + \beta = 0$, thus $K = -\frac{\beta}{\alpha}$, then the solution can be rearranged as:

$$\begin{aligned} D(\tau) = &-\frac{\beta}{R}\left(\frac{-e^{\alpha\tau} + e^{\beta\tau}}{-ge^{\alpha\tau} + e^{\beta\tau}}\right) = -\frac{\beta}{R}\frac{1 - e^{d\tau}}{1 - ge^{d\tau}} \\ =&\frac{Q + d}{2R}\frac{1 - e^{d\tau}}{1 - ge^{d\tau}} = \frac{q - iz\xi\rho + d}{\xi^2}\frac{1 - e^{d\tau}}{1 - ge^{d\tau}} \end{aligned} \tag{E.15}$$

where

$$g = -K = \frac{\beta}{\alpha} = \frac{Q + d}{Q - d} = \frac{q - iz\xi\rho + d}{q - iz\xi\rho - d} \tag{E.16}$$

The solution of $C(\tau)$ in the ODE (equation 3.35) can be derived by taking integration:[3]

$$C(\tau) = \int_0^\tau iz\mu(dy) + a\frac{Q + d}{2R}\int_0^\tau \frac{1 - e^{dy}}{1 - ge^{dy}}(dy) + F \tag{E.17}$$

where F is arbitrarily a constant.

Substituting:

$$x = e^{dy}, \qquad (dx) = de^{dy}(dy), \qquad (dy) = \frac{(dx)}{de^{dy}} = \frac{(dx)}{dx}$$

into equation (E.17), then we have:

$$C(\tau) = iz\mu\tau + \frac{a}{d}\frac{Q + d}{\xi^2}\int_1^{e^{d\tau}} \frac{1 - x}{1 - gx}\frac{1}{x}(dx) + F \tag{E.18}$$

[3]For simplicity and clarification, here we difine the notations as following which are only applied in this appendix. “(dy)” or “(dx)” represents the differential symbol while independent “d” is simply a parameter which is equal to $\alpha - \beta$, i.e. dx means parameter d times x.

This equation can be further rearranged as:[4]

$$\begin{aligned}
C(\tau) =& iz\mu\tau + \frac{a}{d}\frac{Q+d}{\xi^2}\int_1^{e^{d\tau}}\left(\frac{1}{x} - \frac{1-g}{1-gx}\right)(dx) \\
=& iz\mu\tau + \frac{a}{d}\frac{Q+d}{\xi^2}\left[lnx + \frac{1-g}{g}ln(1-gx)\right]_{x=1}^{x=e^{d\tau}} \\
=& iz\mu\tau + \frac{a}{d}\frac{Q+d}{\xi^2}\left[d\tau + \frac{1-g}{g}ln\left(\frac{1-ge^{d\tau}}{1-g}\right)\right] \\
=& iz\mu\tau + \frac{a}{\xi^2}\frac{Q+d}{d}\left[d\tau - \frac{2d}{Q+d}ln\left(\frac{1-ge^{d\tau}}{1-g}\right)\right] \\
=& iz\mu\tau + \frac{a}{\xi^2}\left[(q - iz\xi\rho + d)\tau - 2ln\left(\frac{1-ge^{d\tau}}{1-g}\right)\right]
\end{aligned} \tag{E.19}$$

where

$$a = \kappa\theta, \quad q = \kappa \tag{E.20}$$

[4] The constant F becomes 0 in the calculation due to the initial condition $C(0) = 0$, then we have $C(\tau) = 0 + \frac{a}{d}\frac{Q+d}{\xi^2}\int_1^1 \frac{1-x}{1-gx}\frac{1}{x}(dx) + F = 0$, thus $F = 0$.

Appendix F

Derivations in the section of model specification

F.1 Derivation of unconditional expectation and variance of function h_t

Expectation. Given the volatility function of SV model (formula 4.6), the unconditional expectation of h_t can be derived in the following procedure:[1]

$$\begin{aligned} & E[h_t] = \alpha + \beta E[h_{t-1}] - \alpha\beta \\ \Longrightarrow & E[h_t] - \beta E[h_{t-1}] = \alpha(1-\beta) \\ \Longrightarrow & E[h_t](1-\beta) = \alpha(1-\beta) \Longrightarrow E[h_t] = \alpha \end{aligned} \tag{F.1}$$

Variance. According to the law of total variance, the unconditional variance of h_t is formulated as follows:

$$\begin{aligned} Var(h_t) &= E[Var(h_t|h_{t-1})] + Var(E[h_t|h_{t-1}]) \\ &= \sigma_\eta^2 + Var[\alpha + \beta(h_{t-1} - \alpha)] \\ &= \sigma_\eta^2 + Var(\beta h_{t-1}) = \sigma_\eta^2 + \beta^2 Var(h_{t-1}) \end{aligned} \tag{F.2}$$

[1]Here by assuming $h_0 \sim N(\alpha, \sigma_t^2)$, then we can obtain $E[h_1] = \alpha$, $E[h_2] = \alpha$..., hence there exist an equality where $E[h_t] = E[h_{t-1}] = \alpha$.

and since h_t is a stationary process, then we have:

$$\begin{aligned} Var(h_t)(1-\beta^2) &= \sigma_\eta^2 \\ \Longrightarrow Var(h_t) &= \frac{\sigma_\eta^2}{1-\beta^2} \end{aligned} \tag{F.3}$$

Consequently, it can be formally expressed as:

$$h_t \sim N(\alpha, \frac{\sigma_\eta^2}{1-\beta^2}) \qquad t = 1, 2, ..., T \tag{F.4}$$

F.2 Process to derive function $m_{s,q}$

According to the mathematical definition of VaR, it is not difficult to find a correlation between $f(z_t)$ and α, since oil suppliers are more considered about the downside risk of oil price returns, thus we substitute the negative part of *p.d.f.* of ALD into this correlation function. Corresponding analytic quantile expression ($m_{s,q}$) for the ALD can be obtained with the calculation process shown as follows:

$$\begin{aligned} P(y_t \leq -VaR_{s,t}|\Omega_t) &= P\left(\frac{y_t-\mu}{\sigma_t} \leq -\frac{VaR_{s,t}+\mu}{\sigma_t}\middle|\Omega_t\right) \\ &= P\left(z_t \leq -m_{s,q} = -\frac{VaR_{s,t}+\mu}{\sigma_t}\right) = \int_{-\infty}^{-m_{s,q}} f^-(z_t)\,dz_t \\ &= \int_{-\infty}^{-m_{s,q}} \frac{\sqrt{2}}{\tau}\frac{\kappa}{1+\kappa^2} exp(\frac{\sqrt{2}z_t}{\tau\kappa})\,dz_t \\ &= \frac{\sqrt{2}}{\tau}\frac{\kappa}{1+\kappa^2}\int_{-\infty}^{-m_{s,q}} \frac{\tau\kappa}{\sqrt{2}} d(exp(\frac{\sqrt{2}z_t}{\tau\kappa})) \\ &= \frac{\kappa^2}{1+\kappa^2}\left(exp(\frac{\sqrt{2}(-m_{s,q})}{\tau\kappa}) - exp(\frac{\sqrt{2}(-\infty)}{\kappa\tau})\right) \\ &= \frac{\kappa^2}{1+\kappa^2} exp(\frac{\sqrt{2}(-m_{s,q})}{\kappa\tau}) = \alpha \end{aligned}$$

where $f^-(z_t)$ represents the negative part of the *p.d.f.* of ALD. Transforming this equation, it is easy to obtain:

$$m_{s,q} = -\frac{\kappa\tau}{\sqrt{2}} ln \frac{\alpha(1+\kappa^2)}{\kappa^2} \tag{F.5}$$

F.3 Process to derive $CVaR_{s,t}$ under SV-ALD

The CVaR for supply (in the case $y_t \leq 0$) under SV-ALD model is shown as:

$$CVaR_{s,t} = -E\left[y_t|y_t \leq -VaR_{s,t}\right] = -(\mu + \sigma_t E\left[z_t|z_t \leq -m_{s,q}\right]) \tag{F.6}$$

Since we have:

$$\begin{aligned}
E\left[z_t|z_t \leq -m_{s,q}\right] =& \frac{\int_{-\infty}^{-m_{s,q}} z_t\, f^-(z_t)\, dz_t}{prob(z_t \leq -m_{s,q})} \\
=& \alpha^{-1}\int_{-\infty}^{-m_{s,t}} z_t f^-(z_t) dz_t = \alpha^{-1}\int_{-\infty}^{-m_{s,q}} z_t \frac{\sqrt{2}}{\tau}\frac{\kappa}{1+\kappa^2} exp(\frac{\sqrt{2}z_t}{\kappa\tau})dz_t \\
=& \frac{1}{\alpha}\frac{\sqrt{2}}{\tau}\frac{\kappa}{1+\kappa^2}\int_{-\infty}^{-z_{s,t}} z_t\ exp(\frac{\sqrt{2}z_t}{\kappa\tau})dz_t \\
=& \frac{1}{\alpha}\frac{\sqrt{2}}{\tau}\frac{\kappa}{1+\kappa^2}\frac{\tau\kappa}{\sqrt{2}}\int_{-\infty}^{-z_{s,t}} z_t\ d(exp(\frac{\sqrt{2}z_t}{\kappa\tau})) \\
=& \frac{\kappa^2}{\alpha(1+\kappa^2)}\left[exp(\frac{\sqrt{2}z_t}{\kappa\tau})\ z_t\Big|_{-\infty}^{-m_{s,q}} - \int_{-\infty}^{-m_{s,q}} exp(\frac{\sqrt{2}z_t}{\kappa\tau})\ dz_t \right] \\
=& \frac{\kappa^2}{\alpha(1+\kappa^2)}\Bigg[(-m_{s,q})\ exp(\frac{\sqrt{2}(-m_{s,q})}{\kappa\tau}) \\
&- (\frac{\kappa\tau}{\sqrt{2}} exp(\frac{\sqrt{2}(-m_{s,q})}{\kappa\tau}) - \frac{\tau\kappa}{\sqrt{2}} exp(\frac{\sqrt{2}(-\infty)}{\kappa\tau}))\Bigg] \\
=& \frac{\kappa^2}{\alpha(1+\kappa^2)}\left[(-m_{s,q} - \frac{\tau\kappa}{\sqrt{2}}) exp(\frac{\sqrt{2}(-m_{s,q})}{\kappa\tau}) \right] \\
=& -m_{s,q} - \frac{\kappa\tau}{\sqrt{2}} = \frac{\kappa\tau}{\sqrt{2}} ln\frac{\alpha(1+\kappa^2)}{\kappa^2} - \frac{\kappa\tau}{\sqrt{2}}
\end{aligned}$$

Substituting the above equation to (F.6), then $CVaR_{s,t}$ can be formulated as:

$$CVaR_{s,t} = -\mu - \sigma_t \frac{\kappa\tau}{\sqrt{2}}\left(1 - ln\frac{\alpha(1+\kappa^2)}{\kappa^2}\right) \tag{F.7}$$

or equivalently expressed as:

$$CVaR_{s,t} = VaR_{s,t} + \frac{\kappa\tau\sigma_t}{\sqrt{2}} \tag{F.8}$$

F.4 Process to derive function $m_{d,q}$

Regarding the risk of oil demand, $f^{+}(z_t)$ should be substituted by the positive part of the *p.d.f.* of ALD. We have the following derivation process:[2]

$$
\begin{aligned}
P(y_t > VaR_{d,t}|\Omega_t) &= P\left(\frac{y_t-\mu}{\sigma_t} > \frac{VaR_{d,t}-\mu}{\sigma_t}\middle|\Omega_t\right) \\
&= P\left(z_t > m_{d,q} = \frac{VaR_{d,t}-\mu}{\sigma_t}\right) = \int_{m_{d,q}}^{+\infty} f^{+}(z_t)\,dz_t \\
&= \int_{m_{d,q}}^{+\infty} \frac{\sqrt{2}}{\tau}\frac{\kappa}{1+\kappa^2}\, exp(-\frac{\sqrt{2}\kappa z_t}{\tau})\,dz_t \\
&= \frac{\sqrt{2}}{\tau}\frac{\kappa}{1+\kappa^2}\int_{m_{d,q}}^{+\infty} \frac{-\tau}{\sqrt{2}\kappa} d(exp(-\frac{\sqrt{2}\kappa z_t}{\tau})) \\
&= -\frac{1}{1+\kappa^2}\left(exp(-\frac{\sqrt{2}\kappa(+\infty)}{\tau}) - exp(-\frac{\sqrt{2}\kappa\, m_{d,q}}{\tau})\right) \\
&= \frac{1}{1+\kappa^2} exp(-\frac{\sqrt{2}\kappa\, m_{d,q}}{\tau}) = \alpha
\end{aligned}
$$

Then, we have:

$$
\begin{aligned}
& exp(-\frac{\sqrt{2}\kappa\, m_{d,q}}{\tau}) = \alpha(1+\kappa^2) \\
\Longrightarrow & -\frac{\sqrt{2}\kappa\, m_{d,q}}{\tau} = ln\left(\alpha(1+\kappa^2)\right) \\
\Longrightarrow & m_{d,q} = -\frac{\tau}{\sqrt{2}\kappa} ln\left(\alpha(1+\kappa^2)\right)
\end{aligned} \tag{F.9}
$$

F.5 Process to derive $CVaR_{d,t}$ under SV-ALD

For oil demand, risk measure CVaR under SV-ALD model (in the case $y_t > 0$) can be mathematically defined as:

$$
CVaR_{d,t} = E\left[y_t|y_t > VaR_{d,t}\right] = \mu + \sigma_t E\left[z_t|z_t > m_{d,q}\right] \tag{F.10}
$$

[2] Another way to derive function $m_{d,t}$ is to consider the other part of the integration interval and transform the equation to the form as: $P(y_t \leq VaR_{d,t}|\Omega_t) = \int_{-\infty}^{m_{d,t}} f(z_t)dz_t = \int_{-\infty}^{0} f^{-}(z_t)dz_t + \int_{0}^{m_{d,t}} f^{+}(z_t)dz_t = 1-\alpha$, where $f^{-}(z_t)$ and $f^{+}(z_t)$ represent negative and positive part of the *p.d.f.* of ALD respectively. Consistent result can be obtained in this logic.

For the conditional expectation of z_t, we have:

$$
\begin{aligned}
E\left[z_t|z_t > m_{d,q}\right] =& \frac{\int_{m_{d,q}}^{+\infty} z_t\, f^{+}(z_t)\, dz_t}{prob(z_t > m_{d,q})} \\
=& \alpha^{-1}\int_{m_{d,q}}^{+\infty} z_t f^{+}(z_t) dz_t = \alpha^{-1}\int_{m_{d,q}}^{+\infty} z_t \frac{\sqrt{2}}{\tau}\frac{\kappa}{1+\kappa^2} exp(\frac{-\sqrt{2}\kappa z_t}{\tau}) dz_t \\
=& \frac{1}{\alpha}\frac{\sqrt{2}}{\tau}\frac{\kappa}{1+\kappa^2}\int_{m_{d,q}}^{+\infty} z_t\ exp(\frac{-\sqrt{2}\kappa z_t}{\tau}) dz_t \\
=& \frac{1}{\alpha}\frac{\sqrt{2}}{\tau}\frac{\kappa}{1+\kappa^2}(-\frac{\tau}{\sqrt{2}\kappa})\int_{m_{d,q}}^{+\infty} z_t\ d(exp(\frac{-\sqrt{2}\kappa z_t}{\tau})) \\
=& -\frac{1}{\alpha(1+\kappa^2)}\left[exp(\frac{-\sqrt{2}\kappa z_t}{\tau})\ z_t\Big|_{m_{d,q}}^{+\infty} - \int_{m_{d,q}}^{+\infty} exp(\frac{-\sqrt{2}\kappa z_t}{\tau})\ dz_t \right] \\
=& -\frac{1}{\alpha(1+\kappa^2)}\left[(-m_{d,q}) exp(\frac{-\sqrt{2}\kappa m_{d,q}}{\tau}) - \int_{m_{d,q}}^{+\infty} exp(\frac{-\sqrt{2}\kappa z_t}{\tau}) dz_t \right] \\
=& -\frac{1}{\alpha(1+\kappa^2)}\Bigg[(-m_{d,q})\ exp(\frac{-\sqrt{2}\kappa m_{d,q}}{\tau}) \\
& - \Bigg(-\frac{\tau}{\sqrt{2}\kappa} exp(\frac{-\sqrt{2}\kappa(+\infty)}{\tau}) + \frac{\tau}{\sqrt{2}\kappa} exp(\frac{-\sqrt{2}\kappa m_{d,q}}{\tau}) \Bigg)\Bigg] \\
=& -\frac{1}{\alpha(1+\kappa^2)}\left[(-m_{d,q} - \frac{\tau}{\sqrt{2}\kappa}) exp(\frac{-\sqrt{2}\kappa m_{d,q}}{\tau}) \right] \\
=& \frac{1}{\alpha(1+\kappa^2)}(m_{d,q} + \frac{\tau}{\sqrt{2}\kappa})(\alpha(1+\kappa^2)) \\
=& m_{d,q} + \frac{\tau}{\sqrt{2}\kappa} = -\frac{\tau}{\sqrt{2}\kappa} ln\left(\alpha(1+\kappa^2)\right) + \frac{\tau}{\sqrt{2}\kappa}
\end{aligned}
\tag{F.11}
$$

Substituting the above equation to formula (F.10), then $CVaR_{d,t}$ can be obtained:

$$
CVaR_{d,t} = \mu + \sigma_t \frac{\tau}{\sqrt{2}\kappa}\left(1 - ln\left(\alpha(1+\kappa^2)\right)\right) \tag{F.12}
$$

or equivalently written as:

$$
CVaR_{d,t} = VaR_{d,t} + \frac{\tau\sigma_t}{\sqrt{2}\kappa} \tag{F.13}
$$

Appendix G

Derivations in the section of methodology

G.1 Derivation of ALD as an SMU

Consider a random variable z follows the ALD (equation 4.11), then the ALD can be expressed as a scale mixture of $f_U(z|\theta - \frac{\kappa\tau\lambda}{\sqrt{2}}, \theta + \frac{\tau\lambda}{\sqrt{2}\kappa})$ and $f_{Ga}(\lambda|2,1)$:[1]

$$\begin{aligned} f(z|\kappa,\tau,\theta,\lambda) &= \int_0^\infty f_U(z|\theta - \frac{\kappa\tau\lambda}{\sqrt{2}}, \theta + \frac{\tau\lambda}{\sqrt{2}\kappa}) \times f_{Ga}(\lambda|2,1)\, d\lambda \\ &= \int_0^\infty \frac{\sqrt{2}\kappa}{\tau(1+\kappa^2)\lambda}\, I(\theta - \frac{\kappa\tau\lambda}{\sqrt{2}} < z < \theta + \frac{\tau\lambda}{\sqrt{2}\kappa})\, \lambda\, exp(-\lambda)\, d\lambda \end{aligned} \tag{G.1}$$

In equation (G.1), there are two cases for z where (1): $z > \theta - \frac{\kappa\tau\lambda}{\sqrt{2}}$ or equivalently $\lambda > \frac{-\sqrt{2}(z-\theta)}{\tau\kappa}$ and (2): $z < \theta + \frac{\tau\lambda}{\sqrt{2}\kappa}$ or equivalently $\lambda > \frac{\sqrt{2}\kappa(z-\theta)}{\tau}$.

Case (1):

$$\begin{aligned} &\int_0^\infty \frac{\sqrt{2}\kappa}{\tau(1+\kappa^2)}\, I(\lambda > \frac{-\sqrt{2}(z-\theta)}{\tau\kappa})\, exp(-\lambda)\, d\lambda \\ =&\frac{-\sqrt{2}\kappa}{\tau(1+\kappa^2)} \int_{\frac{-\sqrt{2}(z-\theta)}{\tau\kappa}}^\infty exp(-\lambda)\, d(-\lambda) \\ =&\frac{\sqrt{2}\kappa}{\tau(1+\kappa^2)}\, exp(\frac{\sqrt{2}(z-\theta)}{\tau\kappa}) \end{aligned} \tag{G.2}$$

[1]Here, location parameter θ is included, though it is assumed to be 0 in the empirical part.

Since $\frac{-\sqrt{2}(z-\theta)}{\tau\kappa} > 0$, thus we have $z < \theta$ which implies:

$$f^{-}(z|\kappa,\tau,\theta) = \frac{\sqrt{2}\kappa}{\tau(1+\kappa^2)}\, exp(\frac{\sqrt{2}(z-\theta)}{\tau\kappa}) \qquad z < \theta$$

Case (2):

$$\begin{aligned}&\int_0^\infty \frac{\sqrt{2}\kappa}{\tau(1+\kappa^2)}\, I(\lambda > \frac{\sqrt{2}\kappa(z-\theta)}{\tau})\, exp(-\lambda)\, d\lambda \\ =&\frac{-\sqrt{2}\kappa}{\tau(1+\kappa^2)} \int_{\frac{\sqrt{2}\kappa(z-\theta)}{\tau}}^\infty exp(-\lambda)\, d(-\lambda) \\ =&\frac{\sqrt{2}\kappa}{\tau(1+\kappa^2)}\, exp(\frac{-\sqrt{2}\kappa(z-\theta)}{\tau})\end{aligned} \tag{G.3}$$

Since $\frac{\sqrt{2}\kappa(z-\theta)}{\tau} \geq 0$, then we have $z \geq \theta$ which implies:

$$f^{+}(z|\kappa,\tau,\theta) = \frac{\sqrt{2}\kappa}{\tau(1+\kappa^2)}\, exp(\frac{-\sqrt{2}\kappa(z-\theta)}{\tau}) \qquad z \geq \theta$$

Hence, it is shown that the Asymmetric Laplace density function of the random variable z:

$$f(z|\kappa,\tau,\theta) = \begin{cases} \frac{\sqrt{2}}{\tau}\frac{\kappa}{1+\kappa^2}\;\; exp(-\frac{\sqrt{2}\kappa}{\tau}(z-\theta)) & z \geq \theta \\ \frac{\sqrt{2}}{\tau}\frac{\kappa}{1+\kappa^2}\;\; exp(\frac{\sqrt{2}}{\tau\kappa}(z-\theta)) & z < \theta \end{cases} \tag{G.4}$$

has been equivalently transformed into the SMU distribution with density representation expressed as:

$$f(z|\kappa,\tau,\theta,\lambda) = \int_0^\infty f_U(z|\theta - \frac{\kappa\tau\lambda}{\sqrt{2}}, \theta + \frac{\tau\lambda}{\sqrt{2}\kappa}) \times f_{Ga}(\lambda|2,1)\, d\lambda \tag{G.5}$$

G.2 Derivation of the pdf of scaled ALD

Consider a random variable z follows the Asymmetric Laplace density function in equation (4.11) with mean and variance given by:[2]

$$E(z) = \theta + \frac{\tau}{\sqrt{2}}(\frac{1}{\kappa} - \kappa) \qquad\qquad Var(z) = \frac{\tau^2}{2}(\frac{1}{\kappa^2} + \kappa^2)$$

A random variable z can be transformed to another random variable ε_t by taking:

$$\varepsilon_t = \frac{z}{\sqrt{Var(z)}} \tag{G.6}$$

Taking partial derivatives of ε_t with respect to z, then we have:

$$dz = \sqrt{Var(z)}\, d\varepsilon_t = \frac{\tau}{\sqrt{2}} \frac{\sqrt{1+\kappa^4}}{\kappa}\, d\varepsilon_t \tag{G.7}$$

In the case $z \geq 0$ or $\varepsilon_t \geq 0$, by substituting (G.6) and (G.7) into density function (G.4), we are able to obtain:

$$\begin{aligned} Pr^{+}(\varepsilon_t) &= \int_0^{+\infty} \frac{\sqrt{2}}{\tau} \frac{\kappa}{1+\kappa^2} \frac{\tau}{\sqrt{2}} \frac{\sqrt{1+\kappa^4}}{\kappa}\; exp(\frac{-\sqrt{2}\kappa}{\tau} \frac{\tau}{\sqrt{2}} \frac{\sqrt{1+\kappa^4}}{\kappa} (\varepsilon_t - \theta))\, d\varepsilon_t \\ &= \int_0^{+\infty} \frac{\sqrt{1+\kappa^4}}{1+\kappa^2}\; exp(-\sqrt{1+\kappa^4}\,(\varepsilon_t - \theta))\, d\varepsilon_t \end{aligned} \tag{G.8}$$

Similarly, in the case $z < 0$ or $\varepsilon_t < 0$, it has:

$$\begin{aligned} Pr^{-}(\varepsilon_t) &= \int_{-\infty}^{0} \frac{\sqrt{2}}{\tau} \frac{\kappa}{1+\kappa^2} \frac{\tau}{\sqrt{2}} \frac{\sqrt{1+\kappa^4}}{\kappa}\; exp(\frac{\sqrt{2}}{\tau\kappa} \frac{\tau}{\sqrt{2}} \frac{\sqrt{1+\kappa^4}}{\kappa} (\varepsilon_t - \theta))\, d\varepsilon_t \\ &= \int_{-\infty}^{0} \frac{\sqrt{1+\kappa^4}}{1+\kappa^2}\; exp(\frac{\sqrt{1+\kappa^4}}{\kappa^2}\,(\varepsilon_t - \theta))\, d\varepsilon_t \end{aligned} \tag{G.9}$$

As a result, the *pdf* of SALD of random variable ε_t given σ_t can be written as:[3]

$$f(\varepsilon_t|\kappa,\theta,\sigma_t) = \begin{cases} \frac{\sqrt{1+\kappa^4}}{1+\kappa^2} \frac{1}{\sigma_t}\; exp(\frac{-\sqrt{1+\kappa^4}}{\sigma_t}\,(\varepsilon_t - \theta)) & \varepsilon_t \geq \theta \\ \frac{\sqrt{1+\kappa^4}}{1+\kappa^2} \frac{1}{\sigma_t}\; exp(\frac{\sqrt{1+\kappa^4}}{\kappa^2\sigma_t}\,(\varepsilon_t - \theta)) & \varepsilon_t < \theta \end{cases} \tag{G.10}$$

[2]See Kotz et al. (2001) for more details.

[3]Note that scale parameter τ has been canceled out in this derivation.

where κ is skewness parameter and σ_t is the time-varying volatility of return series.

G.3 Derivation of scaled ALD as an SMU

This part demonstrates the derivation of SALD as a scale mixture of $f_U(\varepsilon_t|\theta - \frac{\lambda\kappa^2\sigma_t}{\sqrt{1+\kappa^4}}, \theta + \frac{\lambda\sigma_t}{\sqrt{1+\kappa^4}})$ and $f_{Ga}(\lambda|2,1)$:

$$\begin{aligned} f(\varepsilon_t|\kappa,\theta,\lambda,\sigma_t) &= \int_0^\infty f_U(\varepsilon_t|\theta - \frac{\lambda\kappa^2\sigma_t}{\sqrt{1+\kappa^4}}, \theta + \frac{\lambda\sigma_t}{\sqrt{1+\kappa^4}}) \times f_{Ga}(\lambda|2,1)\, d\lambda \\ &= \int_0^\infty \frac{\sqrt{1+\kappa^4}}{1+\kappa^2}\frac{1}{\sigma_t}\frac{1}{\lambda}\, I(\theta - \frac{\lambda\kappa^2\sigma_t}{\sqrt{1+\kappa^4}} < \varepsilon_t < \theta + \frac{\lambda\sigma_t}{\sqrt{1+\kappa^4}})\, \lambda\, exp(-\lambda)\, d\lambda \end{aligned} \tag{G.11}$$

Consider two cases for random variable ε_t where (1): $\varepsilon_t > \theta - \frac{\lambda\kappa^2\sigma_t}{\sqrt{1+\kappa^4}}$ or equivalently $\lambda > \frac{-\sqrt{1+\kappa^4}(\varepsilon_t-\theta)}{\kappa^2\sigma_t}$ and (2): $\varepsilon_t < \theta + \frac{\lambda\sigma_t}{\sqrt{1+\kappa^4}}$ or equivalently $\lambda > \frac{\sqrt{1+\kappa^4}(\varepsilon_t-\theta)}{\sigma_t}$.

Case (1):

$$\begin{aligned} &\int_0^\infty \frac{\sqrt{1+\kappa^4}}{1+\kappa^2}\frac{1}{\sigma_t}\, I(\lambda > \frac{-\sqrt{1+\kappa^4}(\varepsilon_t-\theta)}{\kappa^2\sigma_t})\, exp(-\lambda)\, d\lambda \\ =& \frac{-\sqrt{1+\kappa^4}}{1+\kappa^2}\frac{1}{\sigma_t}\int_{\frac{-\sqrt{1+\kappa^4}(\varepsilon_t-\theta)}{\kappa^2\sigma_t}}^\infty exp(-\lambda)\, d(-\lambda) = \frac{\sqrt{1+\kappa^4}}{1+\kappa^2}\frac{1}{\sigma_t}\, exp(\frac{\sqrt{1+\kappa^4}(\varepsilon_t-\theta)}{\kappa^2\sigma_t}) \end{aligned} \tag{G.12}$$

Since $\frac{-\sqrt{1+\kappa^4}(\varepsilon_t-\theta)}{\kappa^2\sigma_t} > 0$, thus we have $\varepsilon_t < \theta$, which follows:

$$f^-(\varepsilon_t|\kappa,\theta,\sigma_t) = \frac{\sqrt{1+\kappa^4}}{1+\kappa^2}\frac{1}{\sigma_t}\, exp(\frac{\sqrt{1+\kappa^4}(\varepsilon_t-\theta)}{\kappa^2\sigma_t}) \qquad \varepsilon_t < \theta$$

Case (2):

$$\begin{aligned} &\int_0^\infty \frac{\sqrt{1+\kappa^4}}{1+\kappa^2}\frac{1}{\sigma_t}\, I(\lambda > \frac{\sqrt{1+\kappa^4}(\varepsilon_t-\theta)}{\sigma_t})\, exp(-\lambda)\, d\lambda \\ =& \frac{-\sqrt{1+\kappa^4}}{1+\kappa^2}\frac{1}{\sigma_t}\int_{\frac{\sqrt{1+\kappa^4}(\varepsilon_t-\theta)}{\sigma_t}}^\infty exp(-\lambda)\, d(-\lambda) = \frac{\sqrt{1+\kappa^4}}{1+\kappa^2}\frac{1}{\sigma_t}\, exp(\frac{-\sqrt{1+\kappa^4}(\varepsilon_t-\theta)}{\sigma_t}) \end{aligned} \tag{G.13}$$

Since $\frac{\sqrt{1+\kappa^4}(\varepsilon_t-\theta)}{\sigma_t} \geq 0$, thus we have $\varepsilon_t \geq 0$, which follows:

$$f^+(\varepsilon_t|\kappa,\theta,\sigma_t) = \frac{\sqrt{1+\kappa^4}}{1+\kappa^2}\frac{1}{\sigma_t}\,exp(\frac{-\sqrt{1+\kappa^4}(\varepsilon_t-\theta)}{\sigma_t}) \qquad \varepsilon_t \geq \theta$$

As a result, it is demonstrated that the scaled Asymmetric Laplace density function of random variable ε_t:

$$f(\varepsilon_t|\kappa,\theta,\sigma_t) = \begin{cases} \frac{\sqrt{1+\kappa^4}}{1+\kappa^2}\frac{1}{\sigma_t}\,exp(\frac{-\sqrt{1+\kappa^4}}{\sigma_t}(\varepsilon_t-\theta)) & \varepsilon_t \geq \theta \\ \frac{\sqrt{1+\kappa^4}}{1+\kappa^2}\frac{1}{\sigma_t}\,exp(\frac{\sqrt{1+\kappa^4}}{\kappa^2\sigma_t}(\varepsilon_t-\theta)) & \varepsilon_t < \theta \end{cases} \tag{G.14}$$

can be replaced by an SMU distribution given by:

$$f(\varepsilon_t|\kappa,\theta,\lambda,\sigma_t) = \int_0^\infty f_U(\varepsilon_t|\theta - \frac{\lambda\kappa^2\sigma_t}{\sqrt{1+\kappa^4}}, \theta + \frac{\lambda\sigma_t}{\sqrt{1+\kappa^4}}) \times f_{Ga}(\lambda|2,1)\,d\lambda \tag{G.15}$$

G.4 Derivation of full conditional distributions

The full conditional distributions of model parameters and latent volatilities are:

- For parameter α, we have:

$$\begin{aligned} f(\alpha|\beta,\sigma_\eta^2,h,y) &\propto f(h_1|\alpha,\beta,\sigma_\eta^2)\prod_{t=2}^{T} f(h_t|h_{t-1},\alpha,\beta,\sigma_\eta^2) f_N(\mu_\alpha,\sigma_\alpha^2) \\ &= exp\left[-\frac{(h_1-\alpha)^2(1-\beta^2)}{2\sigma_\eta^2} - \frac{\sum_{t=2}^{T}(h_t-\alpha-\beta(h_{t-1}-\alpha))^2}{2\sigma_\eta^2}\right] \\ &\left(\frac{1}{\sigma_\eta^2}\right)^{\frac{T}{2}} \frac{1}{\sqrt{2\pi}\sigma_\alpha} exp\left[-\frac{(\alpha-\mu_\alpha)^2}{2\sigma_\alpha^2}\right] \\ &\propto exp\Bigg\{-\frac{1}{2}\Bigg\{\underbrace{\left[\frac{1-\beta^2+(T-1)(1-\beta^2)}{\sigma_\eta^2} + \frac{1}{\sigma_\alpha^2}\right]}_{\text{A}}\alpha^2 \\ &- 2\alpha\underbrace{\left[\frac{h_1(1-\beta^2)+(1-\beta)\sum_{t=2}^{T}(h_t-\beta h_{t-1})}{\sigma_\eta^2} + \frac{\mu_\alpha}{\sigma_\alpha^2}\right]}_{\text{B}}\Bigg\}\Bigg\} \end{aligned} \tag{G.16}$$

Hence, we can obtain:

$$\alpha|\beta, \sigma_\eta^2, h, y \sim N(\frac{B}{A}, \frac{1}{A})$$

• The beta prior distribution is assigned to $\beta^* = \frac{\beta+1}{2} \sim Be(a_\beta, b_\beta)$, then we have:

$$\begin{aligned}
f(\beta^*|\alpha, \sigma_\eta^2, h, y) &\propto f(h_1|\alpha, \beta, \sigma_\eta^2) \prod_{t=2}^{T} f(h_t|h_{t-1}, \alpha, \beta, \sigma_\eta^2) f_{Be}(a_\beta, b_\beta) \\
&= exp\left[-\frac{(h_1-\alpha)^2(1-\beta^2)}{2\sigma_\eta^2} - \frac{\sum_{t=2}^{T}(h_t-\alpha-\beta(h_{t-1}-\alpha))^2}{2\sigma_\eta^2}\right] \\
&\left(\frac{1}{\sigma_\eta^2}\right)^{\frac{T}{2}} \frac{\beta^{*(a_\beta-1)}(1-\beta^*)^{(b_\beta-1)}}{B(a_\beta, b_\beta)} \\
&\propto exp\left[\frac{\beta\sum_{t=2}^{T}(h_t-1)(h_{t-\alpha}-\alpha)}{\sigma_\eta^2} + \frac{\beta^2[(h_1-\alpha)^2 - \sum_{t=2}^{T}(h_{t-1}-\alpha)^2]}{2\sigma_\eta^2}\right] \\
&(1+\beta)^{(a_\beta-1)}(1-\beta)^{(b_\beta-1)}
\end{aligned} \tag{G.17}$$

where $B(\cdot, \cdot)$ is beta function with $B(a_\beta, b_\beta) = \frac{\Gamma(a_\beta)\Gamma(b_\beta)}{\Gamma(a_\beta+b_\beta)}$, and $\Gamma(\cdot)$ is gamma function.

• For parameter σ_η^2, we have:

$$\begin{aligned}
f(\sigma_\eta^2|\alpha, \beta, h, y) &\propto f(h_1|\alpha, \beta, \sigma_\eta^2) \prod_{t=2}^{T} f(h_t|h_{t-1}, \alpha, \beta, \sigma_\eta^2) f_{IG}(a_\sigma, b_\sigma) \\
&= \frac{1}{\sqrt{2\frac{\sigma_\eta^2}{1-\beta^2}\pi}} exp\left[\frac{-(h_1-\alpha)^2}{2\frac{\sigma_\eta^2}{1-\beta^2}}\right] \frac{1}{\sqrt{2\sigma_\eta^2\pi}} \\
&exp\left[\frac{-\sum_{t=2}^{T}(h_t-\alpha-\beta(h_{t-1}-\alpha))^2}{2\sigma_\eta^2}\right] \frac{b_\sigma^{a_\sigma}}{\Gamma(a_\sigma)} \sigma_\eta^{2(-a_\sigma-1)} exp\left(-\frac{b_\sigma}{\sigma_\eta^2}\right) \\
&\propto exp\left[-\frac{b_\sigma + \frac{1}{2}(h_1-\alpha)^2(1-\beta^2) + \frac{1}{2}\sum_{t=2}^{T}(h_t-\alpha-\beta(h_{t-1}-\alpha))^2}{\sigma_\eta^2}\right] \\
&\left(\frac{1}{\sigma_\eta^2}\right)^{(a_\sigma+\frac{T}{2})+1}
\end{aligned} \tag{G.18}$$

Therefore, we can obtain:

$$\sigma_\eta^2|\alpha, \beta, h, y \sim IG(\hat{a}_\sigma, \hat{b}_\sigma)$$

where $\hat{a}_\sigma = a_\sigma + \frac{T}{2}$ and $\hat{b}_\sigma = b_\sigma + \frac{1}{2}(h_1 - \alpha)^2(1-\beta^2) + \frac{1}{2}\sum_{t=2}^{T}(h_t - \alpha - \beta(h_{t-1} - \alpha))^2$.

- For latent variables h_t, we have:

$$\begin{aligned} f(h_t|h_{-t}, \alpha, \beta, \sigma_\eta^2, y) &\propto f(y|h_t, \alpha, \beta, \sigma_\eta^2) f(h_t|h_{-t}, \alpha, \beta, \sigma_\eta^2) \\ &= \frac{1}{\frac{\lambda e^{h_t/2}}{\sqrt{1+\kappa^4}} + \frac{\lambda\kappa^2 e^{h_t/2}}{\sqrt{1+\kappa^4}}} \frac{1}{\sqrt{2\pi B^2}} \, exp\left[-\frac{(h_t - A)^2}{2B^2}\right] \\ &\propto e^{-\frac{h_t}{2}} \, exp\left[-\frac{1}{2}\left(\frac{h_t^2}{B^2} - 2h_t \frac{A}{B^2}\right)\right] \\ &= exp\Bigg\{ -\frac{1}{2}\Bigg[\underbrace{\frac{1}{B^2}}_{\text{C}} h_t^2 - 2h_t \underbrace{\left(\frac{A}{B^2} - \frac{1}{2}\right)}_{\text{D}}\Bigg]\Bigg\} \end{aligned} \tag{G.19}$$

where

$$A = \alpha + \frac{\beta[(h_{t-1} - \alpha) + (h_{t+1} - \alpha)]}{1+\beta^2}, \qquad B^2 = \frac{\sigma_\eta^2}{1+\beta^2}$$

is the mean and variance of Normal density function $f_N(h_t|A, B^2)$, which has equality:

$$f(h_t|h_{-t}, \alpha, \beta, \sigma_\eta^2) = f(h_t|h_{t-1}, h_{t+1}, \alpha, \beta, \sigma_\eta^2) = f_N(h_t|A, B^2)$$

Hence, it can be shown that:

$$h_t|h_{-t}, \alpha, \beta, \sigma_\eta^2, y \sim N(\frac{D}{C}, B^2) \quad or \quad N(A - \frac{B^2}{2}, B^2) \tag{G.20}$$

Appendix H

VaR and CVaR setting under SV-N model

Risk for oil Supply

(1) VaR: $VaR_{N,s,t} = -\mu - \sigma_t \Phi^{-1}(\alpha)$

where $\Phi^{-1}(\alpha)$ is the inverse cumulative distribution function of a N(0,1).

(2) CVaR: $CVaR_{N,s,t} = -E\left[y_t | \, y_t \leq -VaR_{N,s,t}\right] = -\mu - \frac{\sigma_t}{\alpha}\phi(\Phi^{-1}(\alpha))$

where $\phi(\cdot)$ denotes the probability density function of a N (0,1).[1]

Risk for oil demand

(1) VaR: $VaR_{N,d,t} = \mu + \sigma_t \Phi^{-1}(\alpha)$

where $\Phi^{-1}(\alpha)$ is the inverse cumulative distribution function of a N(0,1).

(2) CVaR: $CVaR_{N,d,t} = E\left[y_t | \, y_t \geq VaR_{N,d,t}\right] = \mu + \frac{\sigma_t}{\alpha}\phi(\Phi^{-1}(\alpha))$

where $\phi(\cdot)$ is the probability density function of a N (0,1).

[1]The nominal risl level $\tilde{\alpha}$ for CVaR backtesting under SV-N model is found depends only on the risk level α. That is, $\tilde{\alpha}$ of 3.96%, 1.96% and 0.38% corresponds to 10%, 5% and 1%, respectively (details see Chen et al., 2012).

Appendix I

Derivation of nominal risk level $\tilde{\alpha}$

To calculate the nominal risk level $\tilde{\alpha}$ (or the probability that CVaR falls at under ALD), the *c.d.f.* of ALD is employed given by:

$$F(z|\kappa,\theta,\tau) = \begin{cases} 1 - \dfrac{1}{1+\kappa^2}\ exp\Big(-\dfrac{\sqrt{2}\kappa}{\tau}(z-\theta)\Big) & z \geq \theta \\ \dfrac{\kappa^2}{1+\kappa^2}\ exp\Big(\dfrac{\sqrt{2}}{\tau\kappa}(z-\theta)\Big) & z < \theta \end{cases} \tag{I.1}$$

Then, we need to evaluate the *c.d.f.* at the point that equate to the CVaR level. A closed form solution of CVaR under ALD is firstly required for oil supply and demand, which can be obtained as:[1]

$$CVaR_s = \frac{\kappa\tau}{\sqrt{2}}\left(1 - ln\frac{\alpha(1+\kappa^2)}{\kappa^2}\right) \tag{I.2}$$

$$CVaR_d = \frac{\tau}{\sqrt{2}\kappa}\left(1 - ln\left(\alpha(1+\kappa^2)\right)\right) \tag{I.3}$$

Substituting the negative (I.2) and positive (I.3) into the second and first formula of (I.1) respectively, then we have:

[1]Derivation of $CVaR_s$ and $CVaR_d$ is straightforward and not to be shown here, it is similar to the derivation of $CVaR_{s,t}$ and $CVaR_{d,t}$ but in a more simpler manner.

For supply:

$$\begin{aligned}
F(CVaR_s|\kappa,\tau,\alpha) =& \frac{\kappa^2}{1+\kappa^2}\ exp\left[\frac{\sqrt{2}}{\tau\kappa}\frac{\kappa\tau}{\sqrt{2}}\left(ln\frac{\alpha(1+\kappa^2)}{\kappa^2}-1\right)\right] \\
=& \frac{\kappa^2}{1+\kappa^2}\ exp\left[ln\frac{\alpha(1+\kappa^2)}{\kappa^2}-1\right] \\
=& \frac{\kappa^2}{1+\kappa^2}\frac{exp\left[ln\frac{\alpha(1+\kappa^2)}{\kappa^2}\right]}{e} \\
=& \frac{\kappa^2}{1+\kappa^2}\frac{\frac{\alpha(1+\kappa^2)}{\kappa^2}}{e} \\
=& \frac{\alpha}{e} \\
=& \tilde{\alpha}
\end{aligned} \tag{I.4}$$

For demand:

$$\begin{aligned}
F(CVaR_d|\kappa,\tau,\alpha) =& 1-\frac{1}{1+\kappa^2}\ exp\left[-\frac{\sqrt{2}\kappa}{\tau}\frac{\tau}{\sqrt{2}\kappa}\left(1-ln\left[\alpha(1+\kappa^2)\right]\right)\right] \\
=& 1-\frac{1}{1+\kappa^2}\ exp\left(ln\left[\alpha(1+\kappa^2)\right]-1\right) \\
=& 1-\frac{1}{1+\kappa^2}\frac{exp(ln\left[\alpha(1+\kappa^2)\right])}{e} \\
=& 1-\frac{1}{1+\kappa^2}\frac{\alpha(1+\kappa^2)}{e} \\
=& 1-\frac{\alpha}{e}
\end{aligned} \tag{I.5}$$

Since for oil demand, they are focusing on the right tail of return distribution. Thus, the probability that CVaR of oil demand falls at is calculated as:

$$\tilde{\alpha} = 1-(1-\frac{\alpha}{e}) = \frac{\alpha}{e} \tag{I.6}$$

In summary, we conclude that the nominal risk level $\tilde{\alpha}$ of CVaR under ALD for oil supply and demand are identical, relying only on α and irrelevant to the skewness and scale parameter in ALD.

Bibliography

[1] Abanto-Valle, C. A., Lachos, V. and Dey, D. K. (2015). Bayesian estimation of a skew-student-t stochastic volatility model. *Methodology and Computing in Applied Probability*, 17(3), 721.

[2] Abanto-Valle, C. A., Bandyopadhyay, D., Lachos, V.H. and Enriquez, I. (2010). Robust Bayesian analysis of heavy-tailed stochastic volatility models using scale mixtures of normal distributions. *Computational Statistics & Data Analysis*, 54(12), 2883-2898.

[3] Abanto-Valle, C. A., Migon, H. and Lachos, V. (2011). Stochastic volatility in mean models with scale mixtures of normal distributions and correlated errors: A Bayesian approach. *Journal of Statistical Planning and Inference*, 141(5), 1875-1887.

[4] Adrian, T. and Shin, H. S. (2010). *Financial Intermediaries and Monetary Economics*. [Online]. Available at https://ssrn.com/abstract=1491603 [Accessed 25 September 2014].

[5] Alexander, G. J. and Baptista, A. M. (2004). A comparison of VaR and CVaR constraints on portfolio selection with the mean-variance model. *Management Science,* 50(9), 1261-1273.

[6] Arditti, F. D. and Levy, H. (1975). Portfolio efficiency analysis in three moments: the multiperiod case. *The Journal of Finance,* 30(3), 797-809.

[7] Artzner, P., Delbaen, F., Eber, J.M. and Heath, D. (1999). Coherent measures of risk. *Mathematical Finance,* 9(3), 203-228.

[8] Asai, M. (2008). Autoregressive stochastic volatility models with heavy-tailed distributions: A comparison with multifactor volatility models. *Journal of Empirical Finance*, 15(2), 332-341.

[9] Attari, M. (2004). *Option Pricing Using Fourier Transforms: A Numerically Efficient Simplification.* [Online]. Available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=520042 [Accessed 12 June 2016].

[10] Ayebo, A. and Kozubowski, T. J. (2003). An asymmetric generalization of Gaussian and Laplace laws. *Journal of Probability and Statistical Science,* 1(2), 187-210.

[11] Bakshi, G. and Madan, D. (2000). Spanning and derivative-security valuation. *Journal of Financial Economics,* 55(2), 205-238.

[12] Bali, T. G., Mo, H. and Tang, Y. (2008). The role of autoregressive conditional skewness and kurtosis in the estimation of conditional VaR. *Journal of Banking & Finance*, 32(2), 269-282.

[13] Bali, T. G. and Theodossiou, P. (2007). A conditional-SGT-VaR approach with alternative GARCH models. *Annals of Operations Research,* 151(1), 241-267.

[14] Bawa, V. S. and Lindenberg, E. B. (1977). Capital market equilibrium in a mean-lower partial moment framework. *Journal of Financial Economics,* 5(2), 189-200.

[15] Berg, A., Meyer, R. and Yu, J. (2004). Deviance information criterion for comparing stochastic volatility models. *Journal of Business & Economic Statistics*, 22(1), 107-120.

[16] Berkowitz, J. (2001). Testing density forecasts, with applications to risk management. *Journal of Business& Economic Statistics*, 19(4), 465-474.

[17] Black, F. and Scholes, M. (1973). The pricing of options and corporate liabilities. *Journal of Political Economy,* 81(3), 637-654.

[18] Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 31(3), 307-327.

[19] Bormetti, G., Cazzola, V., Livan, G., Montagna, G. and Nicrosini, O. (2010). A generalized Fourier transform approach to risk measures. *Journal of Statistical Mechanics: Theory and Experiment, 2010(01),* P01005.

[20] Breidt, F. J., Crato, N. and De Lima, P. (1998). The detection and estimation of long memory in stochastic volatility. *Journal of Econometrics*, 83(1), 325-348.

[21] Brooks, C. and Prokopczuk, M. (2013). The dynamics of commodity prices. *Quantitative Finance*, 13(4), 527-542.

[22] Brooks, S. P. and Gelman, A. (1998). General methods for monitoring convergence of iterative simulations. *Journal of Computational and Graphical Statistics*, 7(4), 434-455.

[23] Cai, N. and Kou, S. G. (2011). Option pricing under a mixed-exponential jump diffusion model. *Management Science,* 57(11), 2067-2081.

[24] Cappuccio, N., Lubian, D. and Raggi, D. (2004). MCMC Bayesian estimation of a skew-GED stochastic volatility model. *Studies in Nonlinear Dynamics* & *Econometrics*, 8(2).

[25] Carlin, B. P. and Chib, S. (1995). Bayesian model choice via Markov chain Monte Carlo methods. *Journal of the Royal Statistical Society. Series B (Methodological)*, 473-484.

[26] Carr, P. and Madan, D. (1999). Option valuation using the fast Fourier transform. *Journal of Computational Finance,* 2(4), 61-73.

[27] Chai, T. and Draxler, R. R. (2014). Root mean square error (RMSE) or mean absolute error (MAE)? Arguments against avoiding RMSE in the literature. *Geoscientific Model Development,* 7(3), 1247-1250.

[28] Chan, J. C. and Grant, A. L. (2016). Modeling energy price dynamics: GARCH versus stochastic volatility. *Energy Economics*, 54, 182-189.

[29] Chang, F.-R. (2004). *Stochastic Optimization in Continuous Time.* New York: Cambridge University Press.

[30] Chen, C. W., Gerlach, R. and Wei, D. (2009). Bayesian causal effects in quantiles: Accounting for heteroscedasticity. *Computational Statistics* & *Data Analysis*, 53(6), 1993-2007.

[31] Chen, L., He, S. and Zhang, S. (2011). Tight bounds for some risk measures, with applications to robust portfolio selection. *Operations Research,* 59(4), 847-865.

[32] Chen, Q., Gerlach, R. and Lu, Z. (2012). Bayesian Value-at-Risk and expected shortfall forecasting via the asymmetric Laplace distribution. *Computational Statistics* & *Data Analysis, 56(11),* 3498-3516.

[33] Chib, S. T. B., Nardari, F. and Shephard, N. (2002). Markov chain Monte Carlo methods for stochastic volatility models. *Journal of Econometrics*, 108(2), 281-316.

[34] Chourdakis, K. (2008). *Financial Engineering: A Brief Introduction Using the Matlab System.* [Online]. Available at http://cosweb1.fau.edu/ jmirelesjames/MatLabCode/Lecture_notes_2008d.pdf [Accessed 22 June 2016].

[35] Choy, S. T. B. and Chan, J. S. (2008). Scale mixtures distributions in statistical modelling. *Australian* & *New Zealand Journal of Statistics*, 50(2), 135-146.

[36] Choy, S. T. B., Wan, W. Y. and Chan, C. M. (2009). *Bayesian Student-t Stochastic Volatility Models via Scale Mixtures.* [Online]. Available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1456822 [Accessed 12 December 2016].

[37] Choy, S. T. B. and Smith, A. (1997). Hierarchical models with scale mixtures of normal distributions. *Test*, 6(1), 205-221.

[38] Christoffersen, P. F. (1998). Evaluating interval forecasts. *International Economic Review*, 841-862.

[39] Christoffersen, P., Heston, S. and Jacobs, K. (2009). The shape and term structure of the index option smirk: Why multifactor stochastic volatility models work so well. *Management Science,* 55(12), 1914-1932.

[40] Čížek, P., Härdle, W. K. and Weron, R. (2011). *Statistical Tools for Finance and Insurance*. 2^{nd} edn. Heidelberg: Springer.

[41] Cont, R. and Voltchkova, E. (2005). A finite difference scheme for option pricing in jump diffusion and exponential Lvy models. *SIAM Journal on Numerical Analysis,* 43(4), 1596-1626.

[42] Cox, J. C., Ingersoll Jr, J. E. and Ross, S. A. (1985). A theory of the term structure of interest rates. *Econometrica: Journal of the Econometric Society,* 385-407.

[43] Cuoco, D. and Liu, H. (2006). An analysis of VaR-based capital requirements. *Journal of Financial Intermediation,* 15(3), 362-394.

[44] Dai, Z. and Wen, F. (2014). Robust CVaR-based portfolio optimization under a genal affine data perturbation uncertainty set. *Journal of Computational Analysis & Applications,* 16(1), 93-103.

[45] Danielsson, J. (1994). Stochastic volatility in asset prices estimation with simulated maximum likelihood. *Journal of Econometrics*, 64(1), 375-400.

[46] Delage, E. and Ye, Y. (2010). Distributionally robust optimization under moment uncertainty with application to data-driven problems. *Operations Research,* 58(3), 595-612.

[47] Dellaportas, P., Forster, J. J. and Ntzoufras, I. (2002). On Bayesian model and variable selection using MCMC. *Statistics and Computing*, 12(1), 27-36.

[48] DeMiguel, V. and Nogales, F. J. (2009). Portfolio selection with robust estimation. *Operations Research,* 57(3), 560-577.

[49] Demirguc-Kunt, A., Detragiache, E. and Merrouche, O. (2013). Bank capital: Lessons from the financial crisis. *Journal of Money, Credit and Banking,* 45(6), 1147-1164.

[50] Dimitrakopoulos, D. N., Kavussanos, M. G. and Spyrou, S. I. (2010). Value at risk models for volatile emerging markets equity portfolios. *The Quarterly Review of Economics and Finance,* 50(4), 515-526.

[51] Dokov, S., Stoyanov, S. V. and Rachev, S. (2007). *Computing VaR and AVaR of Skewed-t Distribution.* [Online]. Available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1730263 [Accessed 5 March 2017].

[52] Duffie, D., Pan, J. and Singleton, K. (2000). Transform analysis and asset pricing for affine jumpdiffusions. *Econometrica,* 68(6), 1343-1376.

[53] Dufresne, D., Garrido, J. and Morales, M. (2009). Fourier inversion formulas in option pricing and insurance. *Methodology and Computing in Applied Probability,* 11(3), 359-383.

[54] Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica: Journal of the Econometric Society,* 987-1007.

[55] Escobar, M. and Gschnaidtner, C. (2016). Parameters recovery via calibration in the Heston model: A comprehensive review. *Wilmott,* 2016(86), 60-81.

[56] Fan, Y., Zhang, Y. -J., Tsai, H. -T. and Wei, Y. -M. (2008). Estimating Value at Riskof crude oil price and its spillover effect using the GED-GARCH approach. *Energy Economics,* 30(6), 3156-3171.

[57] Frankel, J. and Saravelos, G. (2012). Can leading indicators assess country vulnerability? Evidence from the 200809 global financial crisis. *Journal of International Economics,* 87(2), 216-231.

[58] Gallant, A. R., Hsieh, D. and Tauchen, G. (1997). Estimation of stochastic volatility models with diagnostics. *Journal of Econometrics*, 81(1), 159-192.

[59] Gelman, A. and Rubin, D. B. (1992). Inference from iterative simulation using multiple sequences. *Statistical Science*, 457-472.

[60] Gerlach, R. H., Chen, C. W. and Chan, N. Y. (2011). Bayesian time-varying quantile forecasting for value-at-risk in financial markets. *Journal of Business & Economic Statistics*, 29(4), 481-492.

[61] Ghaoui, L. E., Oks, M. and Oustry, F. (2003). Worst-case value-at-risk and robust portfolio optimization: A conic programming approach. *Operations Research,* 51(4), 543-556.

[62] Goh, J.W., Lim, K.G., Sim, M. and Zhang W. (2012). Portfolio value-at-risk optimization for asymmetrically distributed asset returns. *European Journal of Operational Research,* 221, 397-406.

[63] Guermat, C. and Harris, R. D. (2002). Robust conditional variance estimation and value-at-risk. *Journal of Risk*, 4, 25-42.

[64] Harvey, A., Ruiz, E. and Shephard, N. (1994). Multivariate stochastic variance models. *The Review of Economic Studies*, 61(2), 247-264.

[65] Hasan, M. B. (2012). A technique for solving special type quadratic programming problems. *Dhaka University Journal of Science,* 60(2), 209-215.

[66] Hasan, M. R. and Hasan, M. B. (2014). An alternative method for solving quadratic fractional programming problems with homogenous constraints. *Journal of Emerging Trends in Engineering and Applied Sciences,* 5(1), 11-19.

[67] Heston, S. L. (1993). A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Review of Financial Studies,* 6(2), 327-343.

[68] Huang, X., Zhou, H. and Zhu, H. (2009). A framework for assessing the systemic risk of major financial institutions. *Journal of Banking & Finance,* 33(11), 2036-2049.

[69] Huang, D., Zhu, S., Fabozzi, F.J. and Fukushima, M. (2010). Portfolio selection under distributional uncertainty: A relative robust CVaR approach. *European Journal of Operational Research,* 203(1), 185-194.

[70] Hurd, T. R. and Zhou, Z. (2010). A Fourier transform method for spread option pricing. *SIAM Journal on Financial Mathematics,* 1(1), 142-157.

[71] Hürlimann, W. (2013). A moment method for the multivariate asymmetric Laplace distribution. *Statistics & Probability Letters,* 83(4), 1247-1253.

[72] Jayakumar, K. and Kuttykrishnan, A. (2007). A time-series model using asymmetric Laplace distribution. *Statistics & Probability Letters,* 77(16), 1636-1640.

[73] Johnson, H. and Shanno, D. (1987). Option pricing when the variance is changing. *Journal of Financial and Quantitative Analysis*, 143-151.

[74] Jondeau, E. and Rockinger, M. (2006). Optimal portfolio allocation under higher moments. *European Financial Management,* 12(1), 29-55.

[75] Jurczenko, E., Maillet, B. B. and Merlin, P. (2005). *Hedge Funds Portfolio Selection with Higher-order Moments: A Non-parametric Mean-Variance-Skewness-Kurtosis Efficient Frontier.* [Online]. Available at https://ssrn.com/abstract=676904 [Accessed 12 March 2014].

[76] Kahl, C. and Jäckel, P. (2005). Not-so-complex logarithms in the Heston model. *Wilmott Magazine,* 19(9), 94-103.

[77] Kang, S. H., Kang, S.-M. and Yoon, S.-M. (2009). Forecasting volatility of crude oil markets. *Energy Economics*, 31(1), 119-125.

[78] Karali, B., Power, G. J. and Ishdorj, A. (2011). Bayesian state-space estimation of stochastic volatility for storable commodities. *American Journal of Agricultural Economics*, 93(2), 434-440.

[79] Kerkhof, J. and Melenberg, B. (2004). Backtesting for risk-based regulatory capital. *Journal of Banking & Finance*, 28(8), 1845-1865.

[80] Kim, S., Shephard, N. and Chib, S. (1998). Stochastic volatility: likelihood inference and comparison with ARCH models. *The Review of Economic Studies*, 65(3), 361-393.

[81] Komunjer, I. (2007). Asymmetric power distribution: theory and applications to risk measurement. *Journal of Applied Econometrics,* 22(5), 891-921.

[82] Konno, H. and Yamazaki, H. (1991). Mean-absolute deviation portfolio optimization model and its applications to Tokyo stock market. *Management Science,* 37(5), 519-531.

[83] Konno, H., Shirakawa, H. and Yamazaki, H. (1993). A mean-absolute deviation-skewness portfolio optimization model. *Annals of Operations Research,* 45(1), 205-220.

[84] Konno, H. and Suzuki, K.-I. (1995). A mean-variance-skewness portfolio optimization model. *Journal of the Operations Research Society of Japan,* 38(2), 173-187.

[85] Koopman, S. J. and Hol Uspensky, E. (2002). The stochastic volatility in mean model: empirical evidence from international stock markets. *Journal of Applied Econometrics*, 17(6), 667-689.

[86] Koopman, S. J., Jungbacker, B. and Hol, E. (2005). Forecasting daily variability of the S&P 100 stock index using historical, realised and implied volatility measurements. *Journal of Empirical Finance*, 12(3), 445-475.

[87] Kotz, S., Kozubowski, T. and Podgorski, K. (2001). *The Laplace Distribution and Generalizations: A Revisit with Applications to Communications, Economics, Engineering, and Finance.* New York: Springer Science & Business Media.

[88] Kotz, S. and Van Dorp, J. R. (2005). A link between two-sided power and asymmetric Laplace distributions: with applications to mean and variance approximations. *Statistics & Probability Letters,* 71(4), 383-394.

[89] Kozubowski, T. J. and Podgorski, K. (2000). A multivariate and asymmetric generalization of Laplace distribution. *Computational Statistics,* 15(4), 531-540.

[90] Kozubowski, T. J. and Podgrski, K. (2001). Asymmetric Laplace laws and modeling financial data. *Mathematical and Computer Modelling,* 34(9-11), 1003-1021.

[91] Krokhmal, P., Palmquist, J. and Uryasev, S. (2002). Portfolio optimization with conditional value-at-risk objective and constraints. *Journal of Risk,* 4, 43-68.

[92] Kupiec, P. H. (1995). Techniques for verifying the accuracy of risk measurement models. *The Journal of Derivatives*, 3(2), 73-84.

[93] Lai, T.-Y. (1991). Portfolio selection with skewness: a multiple-objective approach. *Review of Quantitative Finance and Accounting,* 1(3), 293-305.

[94] Lardic, S. and Mignon, V. (2008). Oil prices and economic activity: An asymmetric cointegration approach. *Energy Economics*, 30(3), 847-855.

[95] Ledoit, O. and Wolf, M. (2003). Honey, I shrunk the sample covariance matrix. *UPF Economics and Business Working Paper,* (691).

[96] Lewis, A. L. (2000). *Option Valuation under Stochastic Volatility with Mathematica Code.* California: Finance Press.

[97] Lewis, A. L. (2001). *A simple Option Formula for General Jump-diffusion and Other Exponential Lvy Processes.* [Online]. Available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=282110 [Accessed 15 January 2015].

[98] Lindsey, J. and Lindsey, P. (2006). Multivariate distributions with correlation matrices for nonlinear repeated measurements. *Computational Statistics & Data Analysis,* 50(3), 720-732.

[99] Li, F.-j. (2006). Measuring VaR and ES of Stock Market Based on SV-GED Model [J]. *Journal of Systems & Management,* 1, 007.

[100] Liu, X., Li, C. and Wang, J. (2005). MeanCVaR Efficient Frontier and Its Economic Implications (II)[J]. *Journal of Industrial Engineering and Engineering Management,* 1, 001.

[101] Lu, Z., Huang, H. and Gerlach, R. (2010). Estimating Value at Risk: from JP Morgans standard-EWMA to skewed-EWMA forecasting. *University of Sydney Working Paper.*

[102] Lux, T., Segnon, M. and Gupta, R. (2016). Forecasting crude oil price volatility and value-at-risk: Evidence from historical and recent data. *Energy Economics*, 56, 117-133.

[103] Marimoutou, V., Raggad, B. and Trabelsi, A. (2009). Extreme value theory and value at risk: application to oil market. *Energy Economics*, 31(4), 519-530.

[104] Markowitz, H. (1952). Portfolio selection. *The Journal of Finance,* 7(1), 77-91.

[105] Markowitz, H., Todd, P., Xu, G. and Yamane Y. (1993). Computation of mean-semivariance efficient sets by the critical line algorithm. *Annals of Operations Research,* 45(1), 307-317.

[106] 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.

[107] Melino, A. and Turnbull, S. M. (1990). Pricing foreign currency options with stochastic volatility. *Journal of Econometrics*, 45(1), 239-265.

[108] Meyer, R. and Yu, J. (2000). BUGS for a Bayesian analysis of stochastic volatility models. *The Econometrics Journal*, 3(2), 198-215.

[109] Moodley, N. (2005). The Heston model: A practical approach with Matlab code. *University of the Witwa-tersrand, Johannesburg.*

[110] Moon, K.-S., Seon, J.-Y., Wee, I.-S. and Yoon, C. (2009). Comparison of stochastic volatility models: Empirical study on Kospi 200 index options. *Bulletin of the Korean Mathematical Society*, 46(2), 209-227.

[111] Nakajima, J. and Omori, Y. (2009). Leverage, heavy-tails and correlated jumps in stochastic volatility models. *Computational Statistics & Data Analysis*, 53(6), 2335-2353.

[112] Nakajima, J. and Omori, Y. (2012). Stochastic volatility model with leverage and asymmetrically heavy-tailed error using GH skew Students t-distribution. *Computational Statistics & Data Analysis*, 56(11), 3690-3704.

[113] Natarajan, K. and Pachamanova, D.and Sim, M. (2008). Incorporating Asymmetric Distributional Information in Robust Value-at-Risk Optimization, *Management Science,* vol.54, No.3, 573-585.

[114] Nelson, D. B. (1988). *The time Series Behavior of Stock Market Volatility and Returns.* [Online]. Available at https://dspace.mit.edu/handle/1721.1/14363#files-area [Accessed 4 September 2016].

[115] Nicolato, E. and Venardos, E. (2003). Option Pricing in Stochastic Volatility Models of the OrnsteinUhlenbeck type. *Mathematical Finance,* 13(4), 445-466.

[116] Ntzoufras, I. (2011). *Bayesian Modeling Using WinBUGS.* Hoboken: John Wiley & Sons.

[117] Omori, Y., Chib, S., Shephard, N. and Nakajima, J. (2007). Stochastic volatility with leverage: Fast and efficient likelihood inference. *Journal of Econometrics*, 140(2), 425-449.

[118] Papapetrou, E. (2001). Oil price shocks, stock market, economic activity and employment in Greece. *Energy Economics*, 23(5), 511-532.

[119] Prakash, A. J., Chang, C.-H. and Pactwa, T. E. (2003). Selecting a portfolio with skewness: Recent evidence from US, European, and Latin American equity markets. *Journal of Banking & Finance,* 27(7), 1375-1390.

[120] Pederzoli, C. (2006). Stochastic volatility and GARCH: A comparison based on UK stock data. *European Journal of Finance,* 12(1), 41-59.

[121] Pflug, G. C. (2000). Some remarks on the value-at-risk and the conditional value-at-risk. *Probabilistic Constrained Optimization.* Springer.

[122] Quaranta, A. G. and Zaffaroni, A. (2008). Robust optimization of conditional value at risk and portfolio selection. *Journal of Banking & Finance,* 32(10), 2046-2056.

[123] Rachev, S., Schwartz, E. and Khindanova, I. (2003). Stable modeling of market and credit value at risk. *Handbook of Heavy Tailed Distributions in Finance,* 249-328.

[124] Regnier, E. (2007). Oil and energy price volatility. *Energy Economics*, 29(3), 405-427.

[125] Rieck, A. (2007). *Characteristic Function Evaluation in Heston's Model on Stochastic Volatility.* [Online]. Available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=958895 [Accessed 24 March 2015].

[126] Rockafellar, R. T. and Uryasev, S. (2000). Optimization of conditional value-at-risk. *Journal of Risk,* 2, 21-42.

[127] Rockafellar, R. T. and Uryasev, S. (2002). Conditional value-at-risk for general loss distribution. *Journal of Banking & Finance,* 26, 1443-1471.

[128] Rouah, F. D. (2013). *The Heston Model and Its Extensions in Matlab and C.* Hoboken: John Wiley & Sons.

[129] Rubinstein, M., Jurczenko, E. and Maillet, B. (2006). *Multi-moment Asset Allocation and Pricing Models.* Chippenham: John Wiley & Sons.

[130] Sadorsky, P. (1999). Oil price shocks and stock market activity. *Energy Economics*, 21(5), 449-469.

[131] Samuelson, P. A. (1970). The fundamental approximation theorem of portfolio analysis in terms of means, variances and higher moments. *The Review of Economic Studies,* 537-542.

[132] Sandmann, G. and Koopman, S. J. (1998). Estimation of stochastic volatility models via Monte Carlo maximum likelihood. *Journal of Econometrics*, 87(2), 271-301.

[133] Saranya, K. and Prasanna, P. K. (2014). Portfolio Selection and Optimization with Higher Moments: Evidence from the Indian Stock Market. *Asia-Pacific Financial Markets,* 21(2), 133-149.

[134] Sarykalin, S., Serraino, G. and Uryasev, S. (2008). Value-at-risk vs. conditional value-at-risk in risk management and optimization. *Tutorials in Operations Research. INFORMS, Hanover, MD,* 270-294.

[135] Schöbel, R. and Zhu, J. (1999). Stochastic volatility with an OrnsteinUhlenbeck process: an extension. *Review of Finance,* 3(1), 23-46.

[136] Scott, L. O. (1987). Option pricing when the variance changes randomly: Theory, estimation, and an application. *Journal of Financial and Quantitative Analysis*, 22(04), 419-438.

[137] Shephard, N. (2005). *Stochastic Volatility: Selected Readings.* New York: Oxford University Press.

[138] Shreve, S. E. (2004). *Stochastic Calculus for Finance II: Continuous-time Models.* New York: Springer Science & Business Media.

[139] Spiegelhalter, D. J., et al. (2002). Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 64(4), 583-639.

[140] Stein, E. M. and Stein, J. C. (1991). Stock price distributions with stochastic volatility: an analytic approach. *The Review of Financial Studies,* 4(4), 727-752.

[141] Stoyanov, S. V., Rachev, S. T. and Fabozzi, F. J. (2013). Sensitivity of portfolio VaR and CVaR to portfolio return characteristics. *Annals of Operations Research,* 205(1), 169-187.

[142] Takahashi, M., Omori, Y. and Watanabe, T. (2009). Estimating stochastic volatility models using daily returns and realized volatility simultaneously. *Computational Statistics* & *Data Analysis*, 53(6), 2404-2426.

[143] Taylor, S. J. (1994). Modeling stochastic volatility: A review and comparative study. *Mathematical Finance*, 4(2), 183-204.

[144] Tong, Z. (2016). Option pricing in stochastic volatility models driven by fractional Lvy processes. *International Journal of Financial Markets and Derivatives,* 5(1), 56-75.

[145] Tsiotas, G. (2012). On generalised asymmetric stochastic volatility models. *Computational Statistics* & *Data Analysis*, 56(1), 151-172.

[146] Vo, M. (2011). Oil and stock market volatility: A multivariate stochastic volatility perspective. *Energy Economics*, 33(5), 956-965.

[147] Vo, M. T. (2009). Regime-switching stochastic volatility: evidence from the crude oil market. *Energy Economics*, 31(5), 779-788.

[148] Wang, J. J. (2012). On asymmetric generalised t stochastic volatility models. *Mathematics and Computers in Simulation*, 82(11), 2079-2095.

[149] Wang, J. J., Chan, J. S. and Choy, S. T. B. (2011). Stochastic volatility models with leverage and heavy-tailed distributions: A Bayesian approach using scale mixtures. *Computational Statistics* & *Data Analysis*, 55(1), 852-862.

[150] Wang, J. J., Choy, S. T. B. and Chan, J. S. (2013). Modelling stochastic volatility using generalized t distribution. *Journal of Statistical Computation and Simulation*, 83(2), 340-354.

[151] Wei, Y. (2012). Forecasting volatility of fuel oil futures in China: GARCH-type, SV or realized volatility models? *Physica A: Statistical Mechanics and its Applications*, 391(22), 5546-5556.

[152] Wei, Y., Wang, Y. and Huang, D. (2010). Forecasting crude oil market volatility: Further evidence using GARCH-class models. *Energy Economics*, 32(6), 1477-1484.

[153] Wichitaksorn, N., Wang, J. J., Choy, S. T. B. and Gerlach, R. (2015). Analyzing return asymmetry and quantiles through stochastic volatility models using asymmetric Laplace error via uniform scale mixtures. *Applied Stochastic Models in Business and Industry*, 31(5), 584-608.

[154] Wiggins, J. B. (1987). Option values under stochastic volatility: Theory and empirical estimates. *Journal of Financial Economics*, 19(2), 351-372.

[155] Willmott, C. J. and Matsuura, K. (2005). Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. *Climate Research,* 30(1), 79.

[156] Willmott, C. J., Matsuura, K. and Robeson, S. M. (2009). Ambiguities inherent in sums-of-squares-based error statistics. *Atmospheric Environment,* 43(3), 749-752.

[157] Wimmerstedt, L. (2015). *Backtesting Expected Shortfall: The Design and Implementation of Different Backtests.* [Online]. Available at http://www.diva-portal.org/smash/record.jsf?pid=diva2%3A848996&dswid=1158 [Accessed 10 January 2017].

[158] Wipplinger, E. (2007). Philippe Jorion: Value at RiskThe New Benchmark for Managing Financial Risk. *Financial Markets and Portfolio Management,* 21(3), 397-398.

[159] Wu, L. (2007). Modeling financial security returns using Lvy processes. *Handbooks in Operations Research and Management Science,* 15, 117-162.

[160] Yamai, Y., Yoshiba, T. (2005). Value-at-risk versus expected shortfall: a practical perspective. *Journal of Banking and Finance,* 29, 997-1015.

[161] Yao, H., Li, Z. and Lai, Y. (2013). Mean-CVaR portfolio selection: A nonparametric estimation framework. *Computers & Operations Research,* 40(4), 1014-1022.

[162] Youssef, M., Belkacem, L. and Mokni, K. (2015). Value-at-Risk estimation of energy commodities: A long-memory GARCHEVT approach. *Energy Economics*, 51, 99-110.

[163] Yu, J. and Meyer, R. (2006). Multivariate stochastic volatility models: Bayesian estimation and model comparison. *Econometric Reviews*, 25(2-3), 361-384.

[164] Yu, J. and Yang, Z. (2002). *A Class of Nonlinear Stochastic Volatility Models.* [Online]. Available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=307731 [Accessed 20 November 2016].

[165] Yu, K. and Zhang, J. (2005). A three-parameter asymmetric Laplace distribution and its extension. *Communications in StatisticsTheory and Methods*, 34(9-10), 1867-1879.

[166] Zhao, S., Lu, Q., Han, L., Liu, Y. and Hu, F. (2015). A mean-CVaR-skewness portfolio optimization model based on asymmetric Laplace distribution. *Annals of Operations Research,* 226(1), 727-739.

[167] Zhou, X. and Liu, Q.-s. (2010). Portfolio Risk Analysis Based on Copula-SV-GED Model [J]. *Science Technology and Engineering,* 14, 058.

[168] Zhou, X.-h., Dong, Y.-w. and Jiang, T. (2012). Extreme risk measurement based on EVT-POT-SV-GED model [J]. *Journal of Systems Engineering,* 2, 004.

[169] Zhu, J. (2009). *Applications of Fourier Transform to Smile Modeling: Theory and Implementation.* 2^{nd} edn. Heidelberg: Springer Science & Business Media.

[170] Zhu, S. and Fukushima, M. (2009). Worst-case conditional value-at-risk with application to robust portfolio management. *Operations Research,* 57(5), 1155-1168.