

STOCHASTIC CORRELATION AND PORTFOLIO OPTIMIZATION BY  
MULTIVARIATE GARCH

by

Cuicui Luo

A thesis submitted in conformity with the requirements  
for the degree of Doctor of Philosophy  
Graduate Department of Mathematics  
University of Toronto

© Copyright 2016 by Cuicui Luo

# Abstract

Stochastic Correlation and Portfolio Optimization by Multivariate Garch

Cuicui Luo

Doctor of Philosophy

Graduate Department of Mathematics

University of Toronto

2016

Modeling time varying volatility and correlation in financial time series is an important element in derivative pricing, risk management and portfolio management. The main goal of this thesis is to investigate the performance of multivariate GARCH model in stochastic correlation forecast and apply these techniques to develop a new model to enhance the dynamic portfolio performance in several context, including hedge fund portfolio construction.

First, we examine the performance of various univariate GARCH models and regime-switching stochastic volatility models in crude oil market. Then these univariate models discussed are extended to multivariate settings and the empirical evaluation provides evidence on the use of the orthogonal GARCH in correlation forecasting and risk management performance when an equally weighted portfolio is considered.

The recent financial turbulence exposed and raised serious concerns about the optimal portfolio selection problem in hedge funds. The dynamic portfolio construction performance of a broad set of multivariate stochastic volatility models is examined in a fund of hedge fund context. It provides further evidence on the use of the orthogonal GARCH in dynamic portfolio constructions and risk management.

Further in this work, a new portfolio optimization model is proposed in order to improve the dynamic portfolio performance. We enhance the safety-first model with standard deviation constraint and derive an analytic formula by filtering the returns with GH skewed t distribution and OGARCH. It is found that the proposed model outperforms

the classical Mean-Variance model and Mean-CVAR model during financial crisis period for a fund of hedge fund.

**To Romeo and Ruoyi**

## Acknowledgements

First and foremost, my deepest gratitude goes to my advisor, Professor Luis Seco. This work would not have been possible without his guidance, patience and limitless support over the years. I am very grateful for his inspiration, motivation, enthusiasm, and immense knowledge in finance research. He has provided continuous help and encouragement in many respects other than research as well. It was an honor to have him as my advisor.

I would like to thank Prof. Robert McCann, Prof. Adrian Nachman, Prof. Mary Pugh, and Prof. Catherine Sulem and Prof. Anatoliy Swishchuk for their insightful comments and questions helped me improve this thesis.

I would like to thank the graduate administrators in the Department of Mathematics, Ida Bulat and Jemima Merisca, for all their patience, kindness and help during these years.

I am grateful to my friends who supported and encouraged me in any respect during my graduate study. In particular I would like to mention Julio Hernandez, Aurthur Huang, Mohammad Shakourifar, Zhansaya Tleuliyeva, Charles Tsang, Bill Wu, Dexter Wu and Weichi Wu.

Last but not least, to my parents, thank you for all your unconditional love, support and sacrifices at every stage of my personal and academic life. I owe my deepest gratitude to my wonderful children: Romeo and Ruoyi, for always making me smile and standing by me through the long hours, days, weeks, months, and years. This dissertation is dedicated to my kids who give me courage.

# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
1.1	Background . . . . .	1
1.2	Outlines and Contributions . . . . .	2
1.3	Future Research . . . . .	4
<b>2</b>	<b>Risk Modeling in Crude Oil Market</b>	<b>5</b>
2.1	Introduction . . . . .	5
2.2	Volatility Models . . . . .	7
2.2.1	Historical Volatility . . . . .	7
2.2.2	Regime switching models . . . . .	10
2.3	Data . . . . .	10
2.3.1	Data and Sample Description . . . . .	10
2.3.2	Distribution Analysis . . . . .	12
2.4	Results . . . . .	12
2.4.1	GARCH modeling . . . . .	12
2.4.2	Markov Regime Switching modeling . . . . .	15
2.5	Conclusion . . . . .	16
<b>3</b>	<b>Stochastic correlation in risk analytics: a financial perspective</b>	<b>25</b>
3.1	Introduction . . . . .	25
3.2	Stochastic Correlation Models . . . . .	27
3.2.1	Riskmetrics EWMA . . . . .	27
3.2.2	Multivariate GARCH Models . . . . .	28
3.2.3	Orthogonal GARCH Model . . . . .	29
3.2.4	Generalized Orthogonal GARCH Model . . . . .	30
3.3	Model Evaluation Measures . . . . .	31
3.3.1	Mean Absolute Error . . . . .	32
3.3.2	Model Confidence Set Approach . . . . .	33
3.3.3	Backtesting . . . . .	34
3.4	Empirical Results . . . . .	35

3.4.1	First Subsample Period . . . . .	37
3.4.2	Second Subsample Period . . . . .	38
3.5	Conclusion . . . . .	38
<b>4</b>	<b>Portfolio Optimization in Hedge Funds by OGARCH and Markov Switching Model</b>	<b>40</b>
4.1	Introduction . . . . .	40
4.2	Methodology . . . . .	41
4.2.1	OGARCH Model . . . . .	42
4.2.2	Markov Switching Model . . . . .	44
4.2.3	EWMA Model . . . . .	45
4.2.4	Portfolio Optimization Model . . . . .	46
4.3	Empirical Results . . . . .	47
4.4	Asset Weight Sensitivities . . . . .	50
4.4.1	Explicit Calculation . . . . .	51
4.5	Conclusion . . . . .	54
<b>5</b>	<b>Portfolio Optimization Under OGARCH with GH Skewed t Distribution</b>	<b>55</b>
5.1	Introduction . . . . .	55
5.2	Model . . . . .	57
5.2.1	Generalized hyperbolic distributions . . . . .	57
5.2.2	OGARCH . . . . .	61
5.2.3	Portfolio return distribution . . . . .	61
5.3	Portfolio Optimization under GH Skewed t distribution . . . . .	62
5.3.1	VaR . . . . .	63
5.3.2	CVaR . . . . .	64
5.3.3	Mean-Variance Portfolio Optimization . . . . .	65
5.3.4	Safety-first Portfolio Optimization . . . . .	66
5.3.5	Mean-CVaR Portfolio Optimization . . . . .	72
5.4	Parameter estimation for skewed t distribution . . . . .	73
5.5	Numerical Result . . . . .	77
5.5.1	Efficient Frontier . . . . .	79
5.5.2	Portfolio Performance Analysis . . . . .	80
5.6	Conclusion . . . . .	82
	<b>Bibliography</b>	<b>86</b>
	<b>List of Publications</b>	<b>95</b>

# List of Tables

2.1	Statistics on the Daily Crude Oil Index Returns from Feb. 2006 to July 2009	11
2.2	Statistics on the Daily Crude Oil Index logreturn from February 2006 to July 2009 . . . . .	12
2.3	GARCH(1,1) estimation using t-distribution . . . . .	13
2.4	Various GARCH modeling . . . . .	18
2.5	Markov regime switching computation example . . . . .	19
2.6	Markov Rregime switching using Hamilton(1989)s Model . . . . .	20
2.7	Markov regime switching using t-distribution . . . . .	21
2.8	Markov regime switching using GED . . . . .	22
3.1	Test statistics for in sample estimation for the whole period . . . . .	37
3.2	Test statistics for out-of -sample forecast in 2008-2009 . . . . .	38
3.3	Test statistics for out-of -sample forecast in 2013 . . . . .	39
4.1	Optimal portfolio allocation for financial crises periods 32-90. The second row gives the average 10-day returns of each fund index. . . . .	49
4.2	Statistics of realized portfolio returns for the out-of-sample period. Average 10-day portfolio return(Mean), standard deviation(St. Dev), worst-case return(Min) are in percentages. . . . .	50
4.3	Statistics of realized portfolio returns for crisis periods(32-90). Average 10-day portfolio return(Mean), standard deviation(St. Dev), 5% quantile, worst-case return(Min) and Sharpe ratio are in percentages. . . . .	50
4.4	Statistics of realized portfolio returns for normal periods(145-204). Average 10-day portfolio return(Mean), standard deviation (St. Dev), 5% quantile, worst-case return(Min) and Sharpe ratio are in percentages. . . . .	50
4.5	Weight sensitivities for selected values of OGARCH parameters in period 60. The numbers in each column are in percentage. The second row gives the optimal weight of each fund index when gamma equals $\gamma - \Delta$ . . . . .	53
4.6	Weight sensitivities for selected values of OGARCH parameters in period 180. All the numbers are in percentage. . . . .	54



5.1	Statistics of daily returns of fund indices: Mean of daily returns of the fund indices in percentage . . . . .	78
5.2	Covariance of daily returns of the fund indices ( $10^{-5}$ ) . . . . .	78
5.3	Estimated parameters . . . . .	84
5.4	Statistics of realized portfolio returns for the out-of-sample period(2007-2008). average portfolio return(Mean), standard deviation(St.Dev), worst-case return(Min) are in percentages. . . . .	85
5.5	Statistics of realized portfolio returns for the out-of-sample period (2013-2014). mean, standard deviation(St. Dev), worst-case return(Min) are in percentages. . . . .	85

# List of Figures

2.1	Daily price movement of crude oil from Feb. 2006-July 2009. . . . .	11
2.2	Daily returns of crude oil from Feb. 2006-July 2009. . . . .	11
2.3	Distribution fit: Normal Distribution vs. T Distribution. . . . .	13
2.4	Innovation and Standard Deviation of crude oil daily returns by GARCH(1,1). . . . .	14
2.5	simulation and forecasting . . . . .	17
2.6	transitional probabilities in Markov regime switching with GED . . . . .	22
2.7	return dynamics of two regimes from 2006.2-2009.7 . . . . .	23
2.8	price dynamics of two regimes from 2006.2-2009.7 . . . . .	24
3.1	Daily return of S&P 500 and 10-yr Bond Index in 2002-2012 . . . . .	36
3.2	In sample estimation of stochastic correlation between S&P 500 and 10-yr Bond Index, 2002-2013 . . . . .	36
3.3	1-day ahead out-of-sample conditional correlation forecast between S&P 500 and 10-yr Bond Index in 2008-2009 . . . . .	38
3.4	1-day ahead out-of-sample conditional correlation forecast between S&P 500 and 10-yr Bond Index in 2013 . . . . .	39
4.1	10-day realized returns of the optimal portfolio selected by different models. . . . .	48
4.2	MSM: optimal Portfolio Allocation. Each position along the vertical axis represents the weight percentage of each asset in optimal portfolio. . . . .	48
4.3	OGARCH: optimal weight of each asset allocated in Portfolio Allocation. . . . .	48
4.4	EWMA: optimal weight of each asset allocated in Portfolio Allocation. . . . .	49
4.5	GMM: optimal weight of each asset allocated in Portfolio Allocation. . . . .	49
5.1	The correlation of HFRX indices daily returns estimated by OGARCH model. . . . .	78
5.2	QQ-plot versus normal distribution of filtered daily returns of HFRX indices. . . . .	79
5.3	QQ-plot versus GH skewed t distribution of filtered daily returns of HFRX indices. . . . .	80
5.4	Mean-Variance efficient frontier vs 0.95 Mean-CVaR . . . . .	81

5.5	The weekly returns of HFRX indices from 2003-2014 . . . . .	81
5.6	The realized weekly returns of optimal portfolios in 2008. . . . .	82
5.7	The realized weekly returns of portfolio for 2013.4-2014.4 . . . . .	83

# Chapter 1

## Introduction

### 1.1 Background

Modeling time varying volatility and correlation in financial time series is an important element in derivative pricing , risk management and portfolio allocation. Higher volatilities increase the risk of assets, and higher correlations cause an increased risk in portfolios. ARCH and Garch models have been applied to model volatility with a great success to capture some stylized facts of financial time series, such as time-varying volatility and volatility clustering. The Autoregressive Conditional Heteroscedasticity (ARCH) model was first introduced in the seminal paper of Engle (1982). Bollerslev (1986) generalized the ARCH model (GARCH) by modeling the conditional variance to depend on its lagged values as well as squared lagged values of disturbance.

As stock returns evolve, their respective volatilities also tend to move together overtime, across both assets and markets. Following the great success of univariate GARCH model in modeling the volatility, a number of multivariate GARCH models have been developed by Bollerslev et al. (1988), Engle and Kroner (1995) and Silvennoinen and Teräsvirta (2009). The Dynamic Conditional Correlation Multivariate GARCH (Engle, 2002) has been widely used to model the stochastic correlation in energy and commodity market; see Bicchetti and Maystre (2013), Creti and Joëts (2013) and Wang (2012). The correlation in crude oil and natural gas markets has been modeled by the orthogonal GARCH in the paper of Alexander (2004) and the generalized orthogonal GARCH model by Van der Weide (2002) is also developed. Regime switching models have become very popular in financial modeling since the seminal contribution of Hamilton (1989). Hamilton first proposed the Markov switching model (MSM) to model the real GNP in the US. Since then, these models have been widely used to model and forecast business cycles, foreign exchange rates and the volatility of financial time series. The hedge fund industry has

grown rapidly in recent years and has become more and more important as an alternative investment class. The recent financial turbulence exposed and raised serious concerns about the optimal portfolio selection problem in hedge funds. Many papers have examined portfolio optimization in a hedge fund context. The structures of hedge fund return and covariance are crucial in portfolio optimization. The non-normal characteristics of hedge fund returns have been widely described in the literature. Kat and Brooks (2001) find that the hedge fund returns exhibit significant degrees of negative skewness and excess kurtosis. According to Getmansky et al. (2004) and Agarwal and Naik (2004), the returns of hedge fund return are not normal and serially correlated.

The Mean-Variance portfolio optimization model proposed by Markowitz (1952) has become the foundation of modern finance theory. It assumes that asset returns follow the multivariate Gaussian distribution with constant parameters. However, it is well-established that the financial time series have asymmetric returns with fat-tail, skewness and volatility clustering characteristics. It takes standard deviation as a risk measure, which treats both upside and downside payoffs symmetrically.

The safety-first model is first introduced by Roy (1952) and the model is extended by Telser (1955) and Kataoka (1963). In recent years, many safety-first modeled have been developed and discussed due to the growing practical relevance of downside risk. Chiu and Li (2012) develop a modified safety-first model and have studied its application in financial risk management of disastrous events, Norkin and Boyko (2012) have improved the safety-first model by introducing one-sided threshold risk measures. The Kataoka's safety-first model with a constraint of mean return is studied by Ding and Zhang (2009).

## 1.2 Outlines and Contributions

The main objective of this thesis is to investigate the performance of multivariate GARCH model in stochastic volatility and correlation forecast and apply these techniques to develop a new model to enhance the dynamic portfolio performance in several context, including hedge fund portfolio construction.

While the main objective of the thesis is multivariate distributions, Chapter 2 aims to examine the performance of various univariate GARCH models and regime-switching stochastic volatility model in the crude oil market. Using daily return data from NYMX Crude Oil market for the period 13.02.2006 up to 21.07.2009, a number of univariate GARCH models are compared with regime-switching models. In regime-switching models, the oil return volatility follows a dynamic process whose mean is subject to shifts, which

is governed by a two-state first-order Markov process. It is found that GARCH models are very useful in modeling a unique stochastic process with conditional variance, while regime-switching models have the advantage of dividing the observed stochastic behavior of a time series into several separate phases with different underlying stochastic processes. Furthermore, it is shown that the regime-switching models show similar goodness-of-fit result to GARCH modeling, while has the advantage of capturing major events affecting the oil market.

We then extend the empirical evaluation of stochastic correlation modeling in risk analytics from a financial perspective in Chapter 3. It provides evidence on the use of the orthogonal GARCH in correlation forecasting and risk management. The volatilities and correlations of S&P 500 index and US Generic Government 10 Year Yield bond index are investigated based on the exponentially weighted moving average model (EWMA) of RiskMetrics, the Dynamic Conditional Correlation Multivariate GARCH (DCC), the orthogonal GARCH (OGARCH) and the generalized orthogonal GARCH (GOGARCH). The out-of-sample forecasting performances of these models are compared by several methods. It is found that the overall performance of multivariate Garch models is better than EWMA and the out-of-sample sample estimation results show that OGARCH model outperforms the other models in stochastic correlation prediction.

Chapter 4 extends the univariate models discussed in Chapter 2 to multivariate settings. It provides further evidence on the use of the orthogonal GARCH in dynamic portfolio forecasting and risk management. We investigate and compare the performances of the optimal portfolio selected by using the Orthogonal GARCH (OGARCH) Model, Markov Switching Models and the Exponentially Weighted Moving Average (EWMA) Model in a fund of hedge funds. These models are used to calibrate the returns of four HFRX indices from which the optimal portfolio is constructed using the Mean-Variance method. The performance of each optimal portfolio is compared in an out-of-sample period. It is found that OGARCH gives the best-performed optimal portfolio with the highest Sharpe ratio and the lowest risk. Moreover, a sensitivity analysis for the parameters of OGARCH is conducted and it shows that the asset weights in the optimal portfolios selected by OGARCH are very sensitive to slight changes in the input parameters. Although there has been much work done on the comparisons of different multivariate Garch models and Markov Switching Models, this is the first work to compare the performance of optimal portfolios in a hedge fund context.

In Chapter 5, we enhance the safety-first model with standard deviation constraint and derive an analytic formula by filtering the returns with GH skewed t distribution and

OGARCH. The analytical solution to classical mean-variance model is also given in terms of parameters in GH skewed t distribution and OGARCH model. Then the parameters are estimated by EM algorithm. The optimal hedge fund portfolios are selected by Mean-Variance, mean-CVaR and safety-first models in 2008 financial crisis and stable period (2013-2014) and the portfolio performances are compared in risk measurement. The efficient frontier is also presented. In the out-of-sample tests, Mean-CVaR model gives the highest mean return in the post-crisis period, while the modified safety-first model provides some improvements over the other two models during the financial crisis period. The enhanced safety-first model achieves the best risk-adjusted returns in both periods.

### 1.3 Future Research

Stochastic correlation forecast under multivariate GARCH models have been studied by many researchers, but it still needs further research in many directions. One direction is to develop a regime-switching OGARCH model.

More research needs to be carried out on dynamic portfolio optimization for heavy-tailed assets. One direction is to develop a multi-period extension of the single period safety-first portfolio optimization model we have developed in Chapter 5. Another possible direction would be considering multivariate skewed t and independent component analysis (ICA) in estimating the assets distributions directly.

# Chapter 2

## Risk Modeling in Crude Oil Market

### 2.1 Introduction

Risk analysis of crude oil market has always been the core research problem that deserves lots of attention from both the practice and academia. Risks occur mainly due to the change of oil prices. During the 1970s and 1980s there were a great deal increases in oil price. Such price fluctuations came to new peaks in 2007 when the price of crude oil doubled during the financial crisis. These fluctuations of double digit numbers in short periods of time continued between 2007 and 2008, when we see highly volatile oil prices. These fluctuations would not be worrisome if oil wouldn't be such an important commodity in the world's economy. When the oil prices become too high and the volatility increases, it has a direct impact on the economy in general and thus affects the government decisions regarding the market regulation and thus the firm and individual consumer incomes (Bacon and Kojima, 2008).

Price volatility analysis has been a hot research area for many years. Commodity markets are characterized by extremely high levels of price volatility. Understanding the volatility dynamic process of oil price is a very important and crucial way for producers and countries to hedge various risks and to avoid the excess exposures to risks (Bacon and Kojima, 2008).

To deal with different phases of volatility behavior and the dependence of the variability of the time series on its own past, models allowing for heteroscedasticity like ARCH, GARCH or regime-switching models have been suggested by researchers. The former two are very useful in modeling a unique stochastic process with conditional variance; the latter has the advantage of dividing the observed stochastic behavior of a time series into several separate phases with different underlying stochastic processes. Both types of models are widely used in practice.



Hung et al. (2008) employ three GARCH models, i.e., GARCH-N, GARCH-t and GARCH-HT, to investigate the influence of fat-tailed innovation process on the performance of energy commodities VaR estimates. Kazmerchuk et al. (2005) consider a continuous-time limit of GARCH(1,1) model for stochastic volatility with delay and the model fit well for the market with delayed response. Narayan et al. (2008) use the exponential GARCH models to evaluate the impact of oil price on the nominal exchange rate. To validate cross-market hedging and sharing of common information by investors, Malik and Ewing (2009) employ bivariate GARCH models to estimate the relations between five different US sector indexes and oil prices. The continuous time GARCH (1,1) model is also used for volatility and variance swaps valuations in the energy market (Swishchuk, 2013, Swishchuk and Couch, 2010). On the other side, regime-switching has been used a lot in modeling stochastic processes with different regimes. Alizadeh et al. (2008) introduce a Markov regime switching vector error correction model with GARCH error structure and show how portfolio risks are reduced using state dependent hedge ratios. Agnolucci (2009) employs a two regime Markov-switching EGARCH model to analyze oil price change and find the probability of transition across regimes. Klaassen (2002) develops a regime-switching GARCH model to account for the high persistence of shocks generated by changes in the variance process. Oil shocks were found to contribute to a better description of the impact of oil on output growth (Cognigni and Manera, 2009). There is no clear evidence regarding which approach outperforms the other one.

Fan et al. (2008) argue that GED-GARCH-based VaR approach is more realistic and more effective than the well-recognized historical simulation with ARMA forecasts in an empirical study. Aloui and Mabrouk (2010) find that the FIAPARCH model outperforms the other models in the VaR's prediction and GARCH models also perform better than the implied volatility by inverting the Black equation. According to Agnolucci (2009), the GARCH model performs best when assuming GED distributed errors. Clear evidences of regime-switching have been discovered in the oil market. Fong and See (2001) believe that regime switching models provide a useful framework for the evolution of volatility and and forecasts of oil futures with short-term volatility. The regime-switching stochastic volatility model performs well in capturing major events affecting the oil market (Vo, 2009).

This chapter will focus on volatility modeling in crude oil market using both regime-switching stochastic volatility model and GARCH models. The next section will review various types of volatility models. We will then look at crude market data in Section 3. Computation and results analysis are presented in Section 4. The last section concludes the chapter.

## 2.2 Volatility Models

### 2.2.1 Historical Volatility

We suppose that  $\epsilon_t$  is the innovation in mean for log returns. To estimate the volatility at time  $t$  over the last  $N$  days we have

$$V_{H,t} = \left[ \frac{1}{N} \sum_{i=0}^{N-1} \epsilon^2 \right]^{1/2} \quad (2.1)$$

This is actually an  $N$ -day simple moving average volatility, where the historical volatility is assumed to be constant over the estimation period and the forecast period. To involve the long-run or unconditional volatility using all previous returns available at time  $t$ , we have many variations of the simple moving average volatility model (Fama, 1970).

#### ARMA (R,M) model

Given a time series  $X_t$ , the autoregressive moving average (ARMA) model is a very useful for predicting future values in time series where there are both an autoregressive (AR) part and a moving average (MA) part. The model is usually then referred to as the ARMA(R,M) model where  $R$  is the order of the first part and  $M$  is the order of the second part.

The following ARMA(R,M) model contains the AR(R) and MA(M) models:

$$X_t = c + \epsilon_t + \sum_{i=1}^R \varphi_i X_{t-i} + \sum_{j=1}^M \theta_j \epsilon_{t-j} \quad (2.2)$$

where  $\varphi_i$  and  $\theta_j$  for any  $i = 1, \dots, R$ ,  $j = 1, \dots, M$  are parameters for AR and MA parts respectively.

#### ARMAX(R,M, b) model

To include the AR(R) and MA(M) models and a linear combination of the last  $b$  terms of a known and external time series  $d_t$ , one can have a model of ARMAX(R,M, b) with  $R$  autoregressive terms,  $M$  moving average terms and  $b$  exogenous inputs terms.

$$X_t = c + \epsilon_t + \sum_{i=1}^R \varphi_i X_{t-i} + \sum_{j=1}^M \theta_j \epsilon_{t-j} + \sum_{k=1}^b \eta_k d_{t-k} \quad (2.3)$$

where  $\eta_1, \dots, \eta_b$  are the parameters of the exogenous input  $d_t$ .

### ARCH(q)

Autoregressive Conditional Heteroscedasticity (ARCH) type modeling is the predominant statistical technique employed in the analysis of time-varying volatility. In ARCH models, volatility is a deterministic function of historical returns.

The original ARCH(q) formulation proposed by Engle (1982) models conditional variance  $\sigma_t^2$  as a linear function of the first q past squared innovations:

$$\sigma_t^2 = c + \sum_{i=1}^q \alpha_i \epsilon_{t-i}^2 \quad (2.4)$$

This model allows today's conditional variance to be substantially affected by the (large) square error term associated with a major market move (in either direction) in any of the previous q periods. It thus captures the conditional heteroscedasticity of financial returns and offers an explanation of the persistence in volatility.

A practical difficulty with the ARCH(q) model is that in many of the applications a long length q is called for.

### GARCH(p,q)

The Generalized Autoregressive Conditional Heteroscedasticity (GARCH) developed by Bollerslev and Wooldridge (1992) generalizes the ARCH model by allowing the current conditional variance to depend on the past conditional variances as well as past squared innovations.

The GARCH(p,q) model is defined as

$$\sigma_t^2 = L + \sum_{j=1}^q \alpha_j \epsilon_{t-j}^2 + \sum_{i=1}^p \beta_i \sigma_{t-i}^2 \quad (2.5)$$

where L denotes the long-run volatility. L,  $\alpha$ ,  $\beta$  are constants and  $\sum_{j=1}^q \alpha_j + \sum_{i=1}^p \beta_i < 1$ .

By accounting for the information in the lag(s) of the conditional variance in addition to the lagged squared returns, the GARCH model reduces the number of parameters required. In most cases, one lag for each variable is sufficient.

The GARCH(1,1) model is given by:

$$\sigma_t^2 = L + \beta_1 \sigma_{t-1}^2 + \alpha_1 \epsilon_{t-1}^2 \quad (2.6)$$

GARCH can successfully capture thick tailed returns and volatility clustering. It can also be modified to allow for several other stylized facts of asset returns.

### EGARCH

The Exponential Generalized Autoregressive Conditional Heteroscedasticity (EGARCH) model introduced by Nelson (1991) builds in a directional effect of price moves on conditional variance. Large price declines, for instance may have a larger impact on volatility than large price increases.

The general EGARCH(p,q) model for the conditional variance of the innovations, with leverage terms and an explicit probability distribution assumption, is

$$\log \sigma_t^2 = L + \sum_{i=1}^p \beta_i \log \sigma_{t-i}^2 + \sum_{j=1}^q \beta_j \left[ \frac{|\epsilon_{t-j}|}{\sigma_{t-j}} - E \left\{ \frac{|\epsilon_{t-j}|}{\sigma_{t-j}} \right\} \right] + \sum_{j=1}^q L_j \left( \frac{\epsilon_{t-j}}{\sigma_{t-j}} \right) \quad (2.7)$$

where  $E \left\{ \frac{|\epsilon_{t-j}|}{\sigma_{t-j}} \right\} = E \{|z_{t-j}|\} = \sqrt{\frac{2}{\pi}}$  for the normal distribution and  $E \left\{ \frac{|\epsilon_{t-j}|}{\sigma_{t-j}} \right\} = E \{|z_{t-j}|\} = \sqrt{\frac{v-2}{\pi} \frac{\Gamma(\frac{\nu-1}{2})}{\Gamma(\frac{\nu}{2})}}$  for the Student's t distribution with degree of freedom  $\nu > 2$ .

### GJR(p,q)

GJR(p,q) model is an extension of an equivalent GARCH(p,q) model with zero leverage terms. Thus, estimation of initial parameter for GJR models should be identical to those of GARCH models. The difference is the additional assumption with all leverage terms being zero:

$$\sigma_t^2 = L + \sum_{j=1}^q \alpha_j \epsilon_{t-j}^2 + \sum_{i=1}^p \beta_i \sigma_{t-i}^2 + \sum_{j=1}^q L_j S_{t-j} \epsilon_{t-j}^2 \quad (2.8)$$

where  $S_{t-j} = 1$  if  $\epsilon_{t-j} < 0$ ,  $S_{t-j} = 0$  otherwise, with constraints

$$\sum_{j=1}^q \alpha_j + \sum_{i=1}^p \beta_i + \frac{1}{2} \sum_{j=1}^q L_j < 1 \quad (2.9)$$

for any  $\alpha_j \geq 0$ ,  $\alpha_j + L_j \geq 0$ ,  $L_j \geq 0$ ,  $\beta_i \geq 0$  where  $i = 1, \dots, p$  and  $j = 1, \dots, q$ .

## 2.2.2 Regime switching models

Markov regime-switching model has been applied in various fields such as oil and the macroeconomic analysis (Raymond and Rich, 1997), analysis of business cycles (Hamilton, 1989) and modeling stock market and asset returns (Engel, 1994).

We now consider a dynamic volatility model with regime-switching. Suppose a time series  $y_t$  follow an AR (p) model with AR coefficients, together with the mean and variance, depending on the regime indicator  $s_t$ :

$$y_t = \mu_{s_t} + \sum_{j=1}^p \varphi_{j,s_t} y_{t-j} + \epsilon_t \quad (2.10)$$

where  $\epsilon_t \sim i.i.dNormal(0, \sigma_{s_t}^2)$ .  $\mu_{s_t}$  is the mean at state  $s_t$ .

The corresponding density function for  $y_t$  is:

$$f(y_t | s_t, y_{t-1}) = \frac{1}{\sqrt{2\pi}\sigma_{s_t}} \exp \left[ -\frac{\omega_t^2}{2\sigma_{s_t}^2} \right] = f(y_t | s_t, y_{t-1}, y_{t-p}) \quad (2.11)$$

where  $\omega_t = y_t - \mu_{s_t} - \sum_{j=1}^p \varphi_{j,s_t} y_{t-j}$ .

The model can be estimated by use of maximum log likelihood estimation. A more practical situation is to allow the density function of  $y_t$  to depend on not only the current value of the regime indicator  $s_t$  but also the past values of the regime indicator  $s_t$ , which means the density function should takes the form of

$$f(y_t | S_t, y_{t-1}, y_{t-p}) \quad (2.12)$$

where  $S_{t-1} = s_{t-1}, s_{t-2}, \dots, s_1$  is the set of all the past information on  $s_t$ .

## 2.3 Data

### 2.3.1 Data and Sample Description

The data contains the daily closing prices of the NYMEX Crude Oil index from February 2006 to July 2009. Weekends and holidays are not included in our data thus considering those days as non moving price days. Using the logarithm prices changes means that our continuously compounded return is symmetric, preventing us from getting nonstationary level of oil prices which would affect our return volatility. Table 2.1 presents the descriptive

statistics of the daily crude oil price changes. In Figure 2.1 we show a plot of the Crude Oil daily price movement.

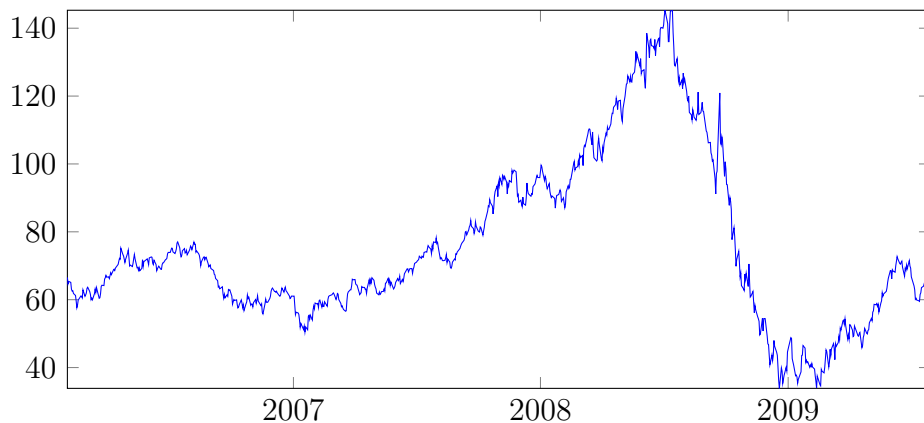


Figure 2.1: Daily price movement of crude oil from Feb. 2006-July 2009.

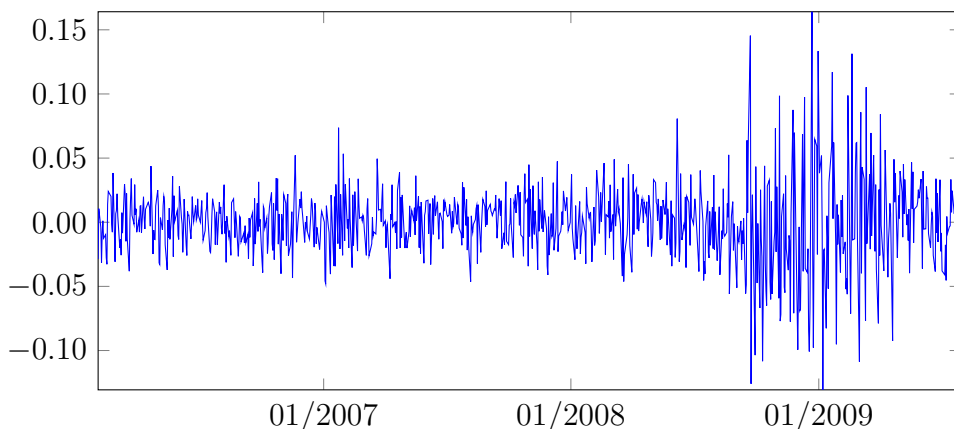


Figure 2.2: Daily returns of crude oil from Feb. 2006-July 2009.

Table 2.1: Statistics on the Daily Crude Oil Index Returns from Feb. 2006 to July 2009

Statistics	Value
Sample Size	881
Mean	75.53
Maximum	145.29
Minimum	33.87
Standard Deviation	24.09
Skewness	0.92
Kurtosis	3.24

To get a preliminary view of volatility change, we show in Table 2.2 the descriptive statistics on the Daily Crude Oil Index logreturn ranging from February 2006 to July 2009. The corresponding plot is given in Figure 2.2 .

Table 2.2: Statistics on the Daily Crude Oil Index logreturn from February 2006 to July 2009

Statistics	Value
Sample Size	880
Mean	6.4692e-006
Maximum	0.1641
Minimum	-0.1307
Standard Deviation	0.0302
Skewness	0.1821
Kurtosis	7.0637

### 2.3.2 Distribution Analysis

The following graph (Figure 2.3) displays a distribution analysis of our data ranging from February 2006 up to July 2009. The data is the log return of the daily crude oil price movements over the time period mentioned above. We can see that the best distribution for our data is a t- Distribution which is shown by the blue line (2.3). The red line represents the normal distribution of our data. So a conditional t- Distribution is preferred to normal distribution in our research. An augmented Dickey-Fuller univariate unit root test yields a resulted p-value of  $1.0 \times 10^{-3}$ ,  $1.1 \times 10^{-3}$  and  $1.1 \times 10^{-3}$  for lags of 0,1 and 2 respectively. All p-values are smaller than 0.05, which indicates that the time series has a trend-stationary property.

## 2.4 Results

### 2.4.1 GARCH modeling

We first estimated the parameters of the GARCH(1,1) model using 865 observations in Matlab, and then tried various GARCH models using different probability distributions with the maximum likelihood estimation technique. In many financial time series the standardized residuals  $z_t = \epsilon_t / \sigma_t$  usually display excess kurtosis which suggests departure from conditional normality. In such cases, the fat-tailed distribution of the innovations driving a dynamic volatility process can be better modeled using the Student's t or the Generalized Error Distribution (GED). Taking the square root of the conditional variance and expressing it as an annualized percentage yields a time-varying volatility estimate. A single estimated model can be used to construct forecasts of volatility over any time horizon. Table 2.3 presents the GARCH(1,1) estimation using t-distribution. The conditional mean process is modeled by use of ARMAX(0,0,0).

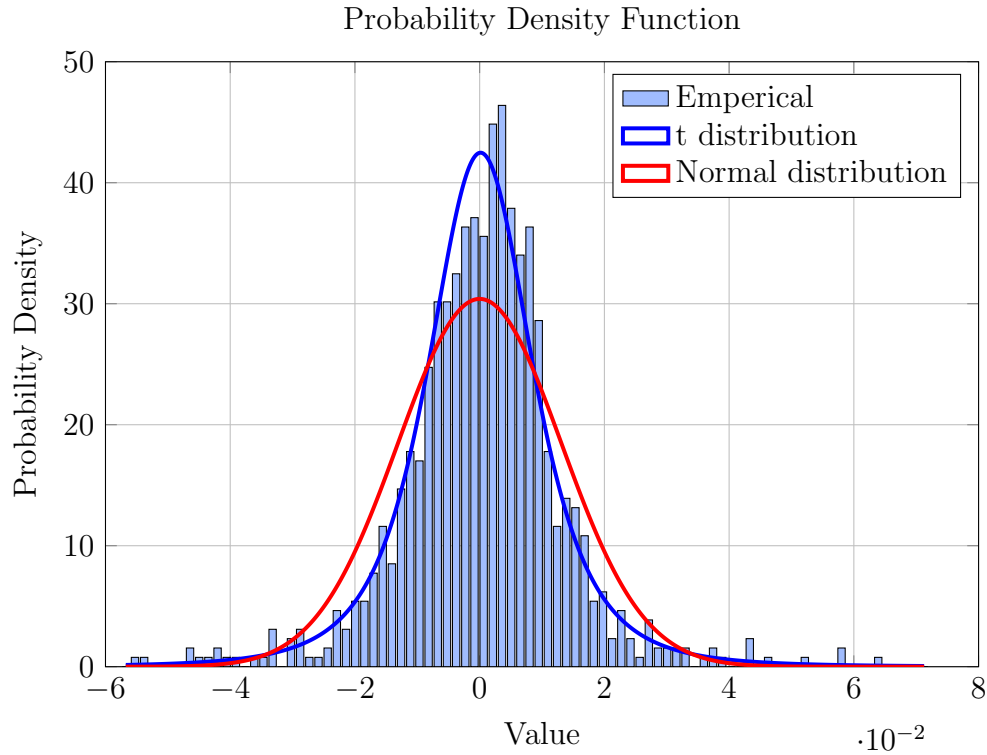


Figure 2.3: Distribution fit: Normal Distribution vs. T Distribution.

Model	AIC	BIC	lnL	Parameter	Value	Standard Error	T Statistic
Mean: AR-MAX(0,0,0); Variance: GARCH(1,1)	-4559.9	-4536.1	2284.97	C	6.819e-4	5.0451e-4	1.3516
				K	2.216e-6	1.306e-6	1.7011
				$\beta_1$	0.9146	0.0174	52.6514
				$\alpha_1$	0.0815	0.0179	4.5539
				DoF	34.603	8.4422e-7	4.0988e+7

Table 2.3: GARCH(1,1) estimation using t-distribution

Substituting these estimated values in the math model, we yield the explicit form as follows:

$$y_t = 6.819e - 4 + \epsilon_t$$

$$\sigma_t^2 = 2.216e - 6 + 0.9146\sigma_{t-1}^2 + 0.0815\epsilon_{t-1}^2$$

Figure 2.4 depicts the dynamics of the innovation and standard deviation using the above estimated GARCH model, i.e., the ARMAX(0,0,0) GARCH(1,1) with the log likelihood value of 2284.97. We want to find a higher log likelihood value for other GARCH modeling, so we use the same data with different models in order to increase the robustness of our model. We now try different combinations of ARMAX and GARCH, EGARCH and GJR models. Computation results are presented in Table 2.4.



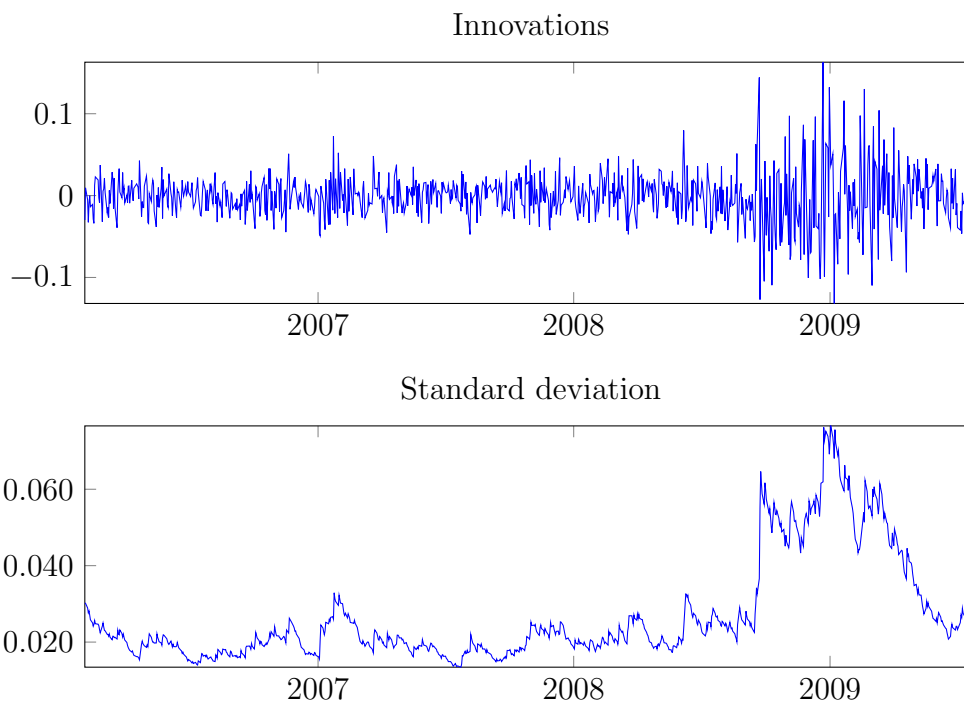


Figure 2.4: Innovation and Standard Deviation of crude oil daily returns by GARCH(1,1).

A general rule for model selection is that we should specify the smallest, simplest models that adequately describe data because simple models are easier to estimate, easier to forecast, and easier to analyze. Model selection criteria such as AIC and BIC penalize models for their complexity when considering best distributions that fit the data. Therefore, we can use log likelihood(LLC), Akaike (AIC) and Bayesian (BIC) information criteria to compare alternative models. Usually, differences in LLC across distributions cannot be compared since distribution functions can have different capabilities for fitting random data, but we can use the minimum AIC and BIC, maximum LLC values as model selection criteria (Cousineau et al., 2004).

As can be seen from Table 2.4, the log likelihood value of ARMAX(1,1,0) GJR(2,1) yields the highest log likelihood value 2292.32 and lowest AIC value -4566.6 among all modeling technique. Thus we conclude that GJR models should be our preferred model.

The forecasting horizon is defined to be 30 days (one month). The simulation uses 20,000 realizations for a 30-day period based on our fitted model ARAMX(1,1,0) GJR(2,1) and the horizon of 30 days from Forecasting. In Figure 2.5 we compare the outputs from forecasting with those derived from Simulation. The first four panels of Figure 5 compare directly each of the outputs from "Forecasting" with the corresponding statistical result obtained from "Simulation". The last two panels of Figure 2.5 illustrate histograms from

which we could compute the approximate probability density functions and empirical confidence bounds. When comparing forecasting with its counterpart derived from the Monte Carlo simulation, we show computation for four parameters in the first four panels of Figure 2.5: the conditional standard deviations of future innovations, the MMSE forecasts of the conditional mean of the NASDAQ return series, cumulative holding period returns and the root mean square errors (RMSE) of the forecasted returns. The fifth panel of Figure 2.5 uses a histogram to illustrate the distribution of the cumulative holding period return obtained if an asset was held for the full 30-day forecast horizon. In other words, we plot the logreturn obtained by investing in NYMEX Crude Oil index today, and sold after 30 days. The last panel of Figure 2.5 uses a histogram to illustrate the distribution of the single-period return at the forecast horizon, that is, the return of the same mutual fund, the 30th day from now.

## 2.4.2 Markov Regime Switching modeling

We now try Markov regime switching modeling in this section. The purpose is twofold: first, check if Markov Switching regressions can beat GARCH models in time series modeling; second, find turmoil regimes in historical time series. We employ a Markov regime switching computation example in Table 2.5 to illustrate our results.

The model in Table 2.5 assumes Normal distribution and allows all parameters to switch. We use  $S = [111]$  to control the switching dynamics, where the first element of S controls the switching dynamic of the mean equation, while the last term controls the switching dynamic of the residual vector, including distribution parameters mean and variance. A value of “1” in S indicates that switching is allowed in the model while a value of “0” in S indicates that parameter is not allowed to change states. Then the model for the mean equation is:

$$\begin{array}{ll} \text{State 1}(S_t = 1) & \text{State 2}(S_t = 2) \\ y_t = -0.0015 - 0.0667y_{t-1} & y_t = 0.0012 - 0.0934y_{t-1} \\ \epsilon_t \sim N(0, 0.0306^2) & \epsilon_t \sim N(0, 0.0115^2) \end{array}$$

where  $\epsilon_t$  is residual vector which follows a particular distribution. The transition matrix,  $P = \begin{bmatrix} 0.99 & 0.01 \\ 0.01 & 0.99 \end{bmatrix}$ , controls the probability of a regime switch from state 1(2) (column 1(2)) to state 2(1) (row 2(1)). The sum of each column in P is equal to one, since they represent full probabilities of the process for each state.

In order to yield the best fitted Markov regime switching models, we now try various parameter settings for traditional Model by Hamilton (1989) and complicated setting using t-distribution and Generalized Error Distribution. We present computational results

in Table 2.6, 2.7 and 2.8. A comparison of log Likelihood values indicate that complicated setting using t-distribution and Generalized Error Distribution usually are preferred. The best fitted Markov Regime Switching models should assume GED and allow all parameters to change states (see Table 2.8).

We now focus on analysis using the best fitted Markov regime switching model, i.e., “MS model,  $S = [11111]$  (GED)” in Table 2.8. Figure 2.6 presents transitional probabilities in Markov regime switching with GED: fitted state probabilities and smoothed state probabilities. Based on such a transitional probability figure, we can classify historical data into two types according to their historical states.

As can be seen from Figure 2.7 and 2.8, the total historical time series are divided into two regimes: a normal one with small change (state 2) and a turmoil one with big risk (state 1). For each state, regime Switching model identifies three periods of data. The normal regime includes two periods: 2006-02-10 to 2006-12-11, and 2007-01-30 to 2007-10-14. The turmoil regime also includes two periods: 2006-12-12 to 2007-01-29, and 2007-10-15 to 2009-07-07. The first turmoil lasts only one and a half months, but the second one covers almost the total financial crisis.

## 2.5 Conclusion

We have examined crude oil price volatility dynamics using daily data for the period 13.02.2006 up to 21.07.2009. To model volatility, we employ the GARCH, EGARCH and GJR models and various Markov regime switching models using the maximum likelihood estimation technique. Codes are written in Matlab language. We have compared several parameter settings in all models. In GARCH models, the ARMAX (1,1,0)/ GJR(2,1) yielded the best fitted result with maximum log likelihood value of 2292.32 when assuming that our data follow a t-distribution. Markov regime switching models generate similar fitted result but with a bit lower log likelihood value. Markov regime switching modeling show interesting results by classifying historical data into two states: a normal one and a turmoil one. This can account for some market stories in financial crisis.

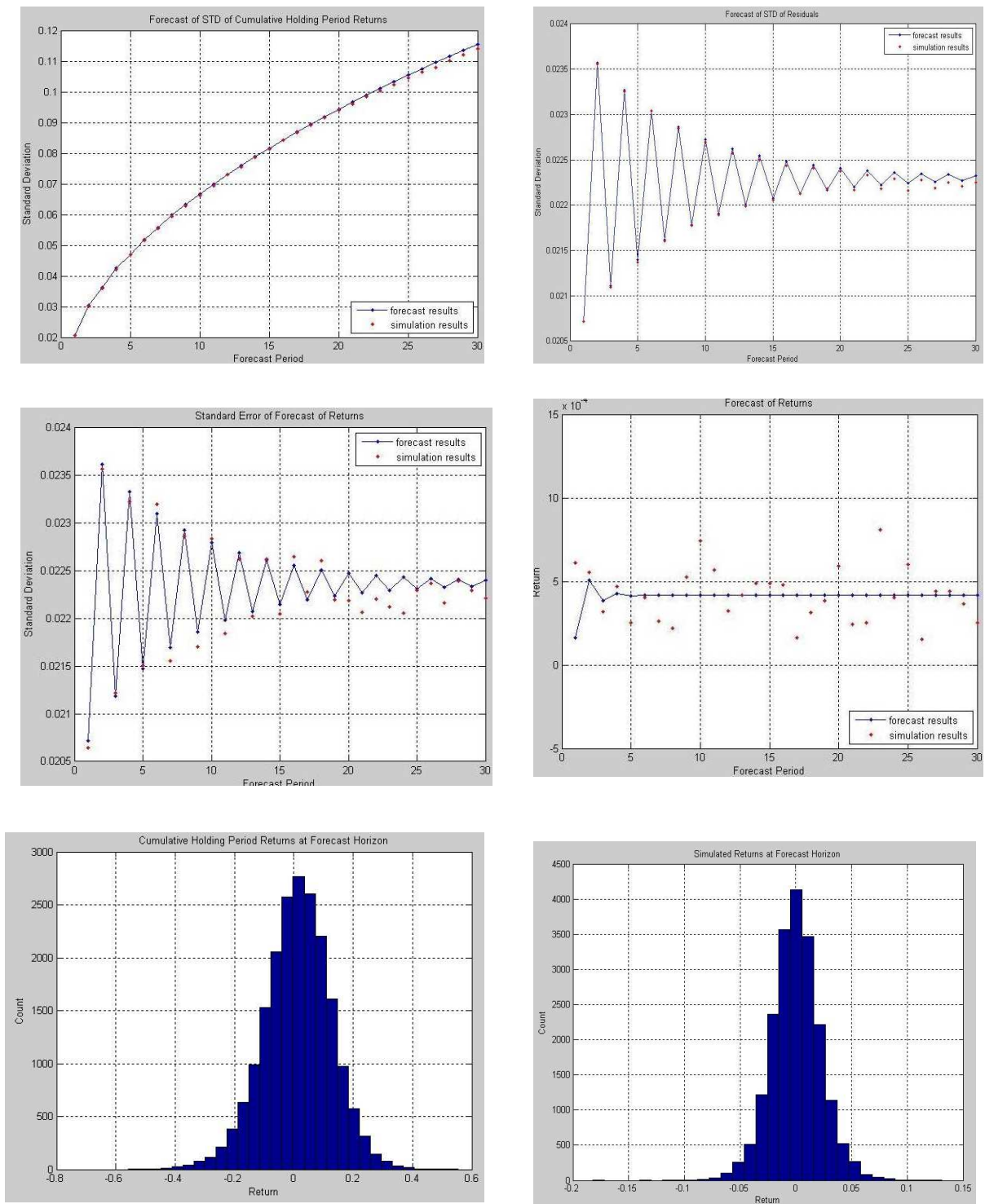


Figure 2.5: simulation and forecasting

Model	AIC	BIC	lnL	Parameter	Value	Standard Error	T Statistic
<i>Mean : ARMAX(1, 1, 0)</i> <i>Variance : GARCH(1, 1)</i>	-4561.0	-4527.7	2287.5	C	8.995e-4	6.6851e-4	1.3455
				$\varphi_1$	-0.3119	0.4386	-0.7111
				$\theta_1$	0.2363	0.4465	0.5292
				K	2.0564e-6	1.2567e-6	1.6363
				$\beta_1$	0.9175	0.0169	54.1607
				$\alpha_1$	0.0790	0.0174	4.5436
				DoF	30.107	1.6771e-4	1.795e+5
<i>Mean : ARMAX(1, 1, 0)</i> <i>Variance : EGARCH(1, 1)</i>	-4557.8	-4524.5	2286.3	C	6.6556e-4	6.2368e-4	1.0672
				$\varphi_1$	-0.3067	0.3895	-0.7874
				$\theta_1$	0.2226	0.3969	0.5607
				K	-0.0396	0.0297	-1.3339
				$\beta_1$	0.9950	3.6255e-3	274.4553
				$\alpha_1$	0.1459	0.0281	5.1980
				L	-0.0316	0.0155	-2.0337
DoF	37.596	48.455	0.7759				
<i>Mean : ARMAX(1, 1, 0)</i> <i>Variance : GJR(1, 1)</i>	-4560.9	-4522.8	2288.4	C	6.912 e-4	6.3923e-4	1.0813
				$\varphi_1$	-0.2967	0.4499	-0.6596
				$\theta_1$	0.2218	0.4567	0.4852
				K	2.1511e-6	1.2682e-6	1.6961
				$\beta_1$	0.9189	0.0168	54.7178
				$\alpha_1$	0.0592	0.0211	2.8779
				L	0.0344	0.0254	1.3541
DoF	38.36	1.1967e-4	3.2054e+4				
<i>Mean : ARMAX(1, 1, 0)</i> <i>Variance : GJR(2, 1)</i>	-4566.6	-4523.8	2292.3	C	5.6469e-4	6.4635e-4	0.8737
				$\varphi_1$	-0.3582	0.4030	-0.8891
				$\theta_1$	0.2843	0.4138	0.6871
				K	3.5044e-6	1.9942e-6	1.7573
				$\beta_1$	0	0.0255	0.0000
				$\beta_2$	0.8682	0.0294	29.5592
				$\alpha_1$	0.0910	0.0255	3.5714
L	0.0677	0.0346	1.9552				
DoF	50.013	6.0689e-6	8.2409e+6				

Table 2.4: Various GARCH modeling

Table 2.5: Markov regime switching computation example

Model Distribution	log Likelihood	Non Switching Parameter	Switching Parameters		Transition Matrix	
			State 1	State 2		
MS Model $S = [111]$ (Normal)	2257.36	N/A	Model's STD Indep column 1 Indep column 2	0.0306 -0.0015 -0.0667	0.0115 0.0012 -0.0934	0.99 0.01 0.01 0.99

Table 2.6: Markov Rregime switching using Hamilton(1989)s Model

Model t Distribution	log Likelihood	Non Switching Parameter	Switching Parameters		Transition Matrix	
			State 1	State 2		
Hamilton(1989)'s Model, $S = [111]$	2212.38	0.0135	DoF	100.00	1.5463	1 0
			Indep column 1	0.0008	-0.0002	0 1
			Indep column	-0.0012	0.0010	
			Indep column	-0.0012	0.0010	
Hamilton(1989)'s Model $S = [111]$	2257.34	N/A	Model's STD	0.0264	0.0113	0.99 0.01
			DoF	7.8238	112.3094	0.01 0.99
			Indep column 1	-0.0012	0.0010	
			Indep column 2	-0.0012	0.0010	

Table 2.7: Markov regime switching using t-distribution

Model t Distribution	log Likelihood	Non Switching		Switching Parameters			Transition Matrix
		Parameter		State 1	State 2	State 2	
MS Model $S = [111]$	2172.41	STD	0.0128	Indep column 1	0.0021	-0.0010	0.45 0.57
		DoF	2.9506	Indep column 2	-0.3925	0.2553	0.55 0.43
$S = [111]$	2212.38	N/A		Model's STD	0.0130	0.0117	0.80 0.98 0.20 0.02
				Indep column 1	0.0013	-0.0034	
				DoF	3.2408	2.3637	
				Indep column 2	-0.2015	0.9080	
$S = [111]$	2212.38	N/A		Model's STD	0.0262	0.0113	0.80 0.98 0.20 0.02
				Indep column 1	-0.0012	0.0011	
				DoF	7.4904	100.000	
				Indep column 2	-0.0736	-0.0915	
				Indep column 3	-0.0121	0.0422	



Table 2.8: Markov regime switching using GED

Model Distribution	log Likelihood	Switching Parameters			Transition Matrix
			State 1	State 2	
MS Model $S = [111]$ (GED)	2257.36	Model's STD	0.0029	0.0094	$\begin{bmatrix} 0.06 & 0.26 \\ 0.94 & 0.74 \end{bmatrix}$
		Value of k	1.4987	0.8011	
		Indep column 1	0.0020	0.0013	
		Indep column 2	0.8905	0.2207	
MS Model $S = [111]$ (GED)  cline3-5	2257.36	Model's STD	0.0029	0.0094	$\begin{bmatrix} 0.99 & 0.01 \\ 0.01 & 0.99 \end{bmatrix}$
		Value of k	0.7122	0.4675	
		Indep column 1	0.0014	0.0010	
		Indep column 2	0.0706	0.0848	
		Indep column 3	0.0287	0.0384	

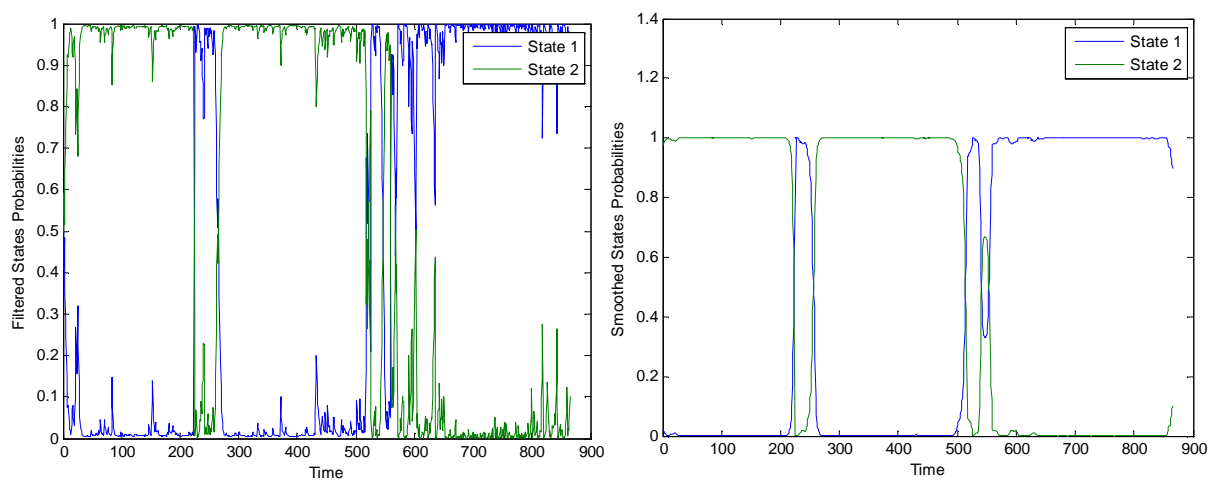


Figure 2.6: transitional probabilities in Markov regime switching with GED

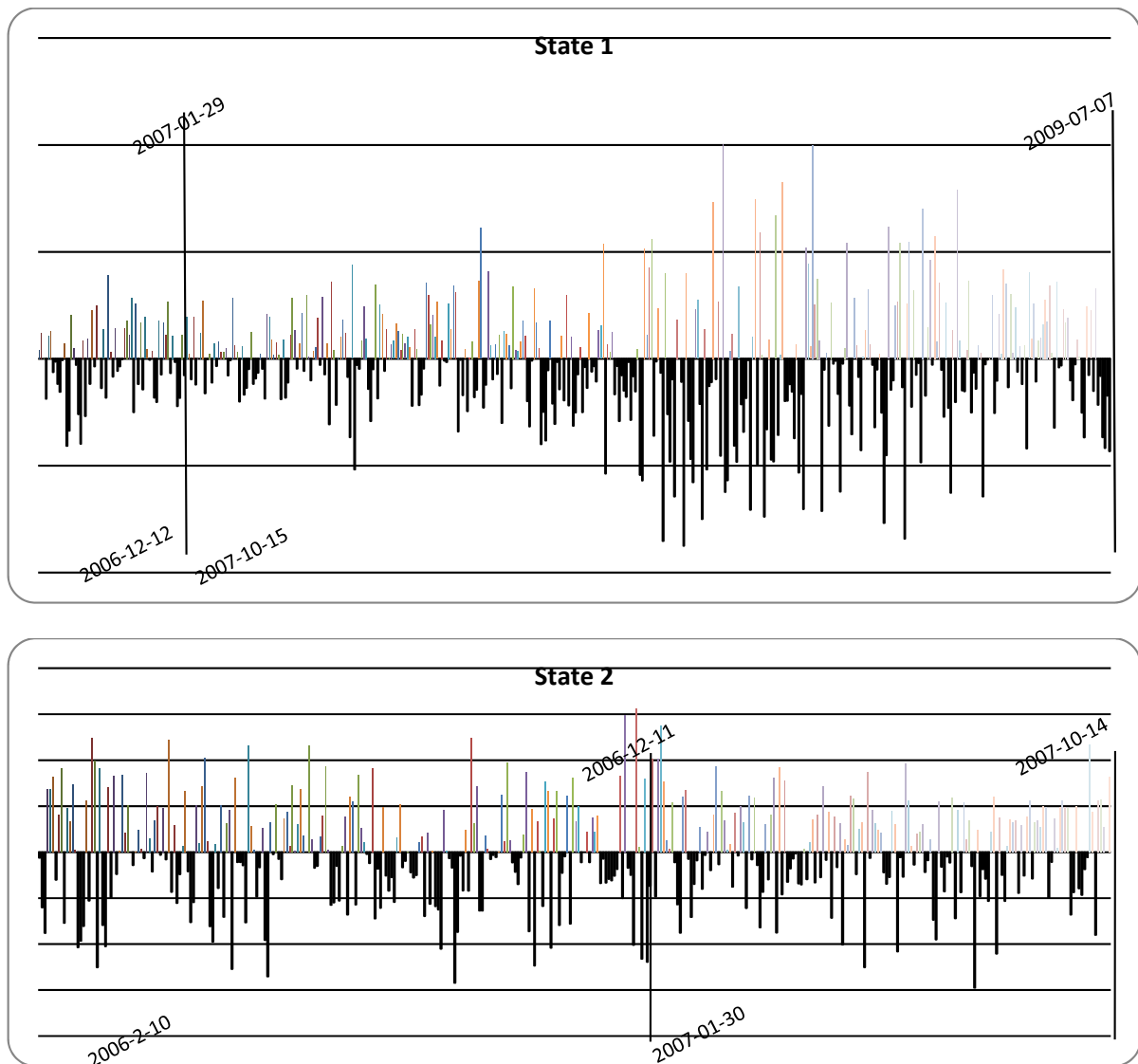


Figure 2.7: return dynamics of two regimes from 2006.2-2009.7

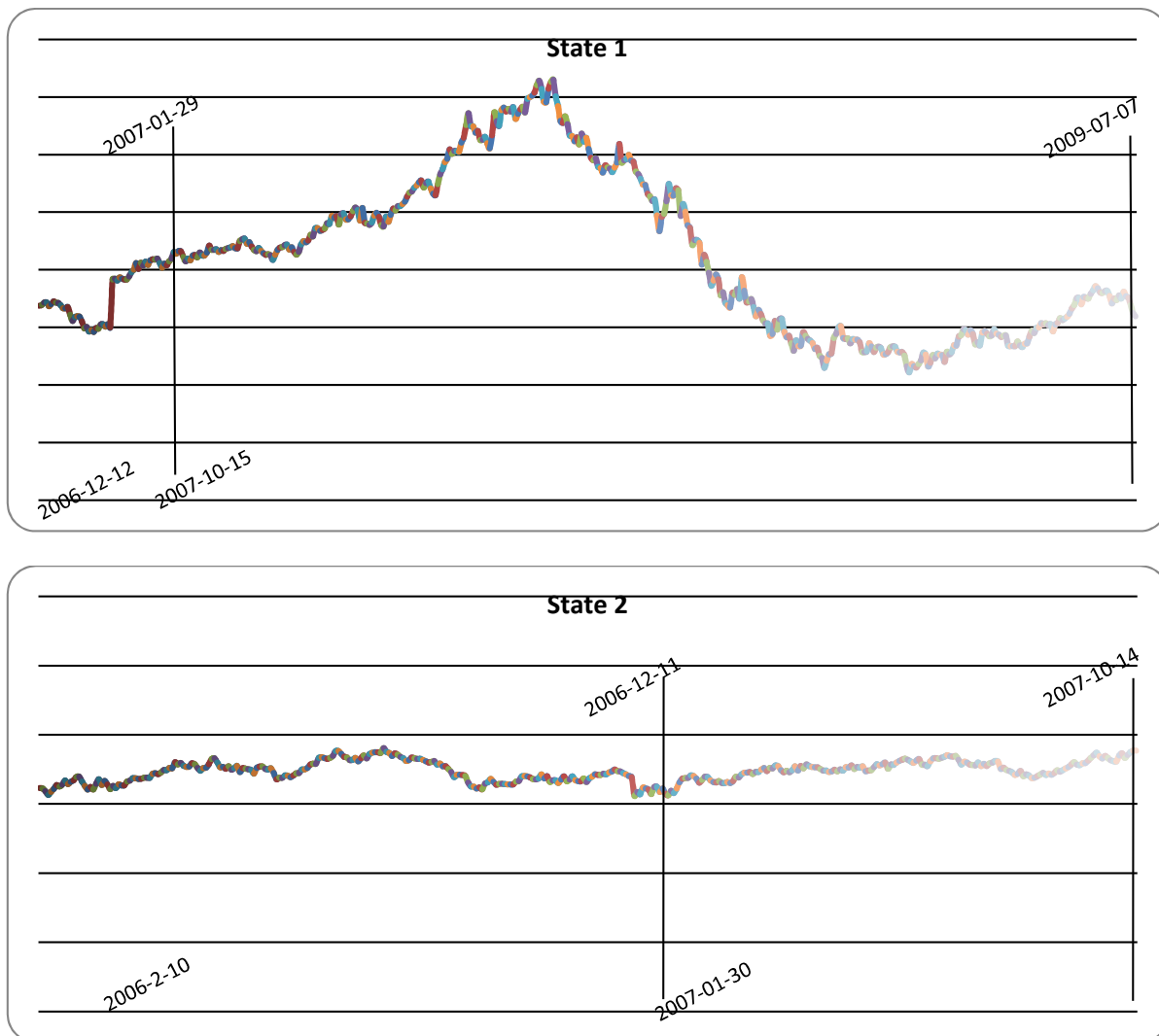


Figure 2.8: price dynamics of two regimes from 2006.2-2009.7

# Chapter 3

## Stochastic correlation in risk analytics: a financial perspective

### 3.1 Introduction

Risk analytics has been popularized by some of today's most successful companies through new theories such as enterprise risk management (Wu and Olson, 2010, Wu et al., 2010). It drives business performance using new sources of data information and advanced modeling tools and techniques. For example, underwriting decisions in the electric power, oil, natural gas and basic-materials industries can be improved by advanced credit-risk analytics so that higher revenues and lower costs are yielded through the analytics of their commodity exposures. By incorporating, we help clients produce models with significantly higher predictive power. Risk analytics can be correlated with the public resources management (Chen et al., 2015).

However, investments from different sources of projects, products and markets can be highly correlated due to the interconnections among these projects, products and markets. Maximizing the benefit from these investments cannot be based on the data and models individually from different sources; it may be more based on the correlation structure dynamically from different sources. For example, value at risk (VaR) has been widely used in financial institutions as a risk management tool after its adoption by the Basel Committee on Banking (1996). Modeling time-varying volatility and correlation for portfolios with a large number of assets is critical and especially valuable in VaR measurement. Significantly higher predictive power has been observed when considering correlation structure in VaR modeling.

Modeling time varying volatility and correlation in financial time series is an important

element in pricing equity, risk management and portfolio management. Many multivariate stochastic correlation models have been proposed to model the time-varying covariance and correlations. Following the great success of univariate GARCH model in modeling the volatility, a number of multivariate GARCH models have been developed; see Bollerslev et al. (1988), Engle and Kroner (1995) and Silvennoinen and Teräsvirta (2009)). The dynamic conditional correlation Multivariate GARCH (Engle, 2002) has been widely used to model the stochastic correlation in energy and commodity market (Bicchetti and Maystre (2013), Creti and Joëts (2013), Wang (2012)). The correlation in crude oil and natural gas markets has been modeled by the orthogonal GARCH in the paper of Alexander (2004) and the generalized orthogonal GARCH model is also developed by Van der Weide (2002). Besides the multivariate GARCH models, the exponentially weighted moving average model (EWMA) of RiskMetrics (1996), which is the simplest matrix generalization of a univariate volatility mode, is also very widely used in variance and covariance forecasting.

Since so many models have been developed over the years, the prediction accuracy of these models becomes a major concern in time series data mining. A number of studies have compared the forecasting performance of the multivariate correlation modes. In the paper of Wong and Vlaar (2003), it shows that the DCC model outperforms other alternatives in modeling time-varying covariance. It is noted that the optimal hedge fund portfolio constructed by dynamic covariance models has lower risk (Giamouridis and Vrontos, 2007). Harris and Mazibas (2010) provide further evidence that the use of multivariate GARCH models in optimal portfolios selection has better performances than static models and also show that exponentially weighted moving average (EWMA) model has the best performance with superior risk-return trade-off and lower tail risk. Engle and Sheppard (2008) compare the performance of some Large-scale multivariate GARCH models using over 50 assets and find that there is value in modeling time-varying covariance of large portfolio by these models. Lu and Tsai (2010) also find that the multivariate GARCH models provide a substantial improvement to the forecast accuracy of the time-varying correlation. The out-of-sample forecasting accuracy of a range of multivariate GARCH models with a focus on large-scale problems is also studied by Caporin (2012) and Laurent and Violante (2012).

This chapter aims to evaluate the forecasting performance of RiskMetrics EWMA, DCC, OGARCH and GOGARCH models for the correlation between S& P 500 index and US Generic Government 10 year yield bond index over 10 years period from 2002 to 2013. First we estimate these models and obtain out-of-sample forecasts of time-varying correlations. Then mean absolute error (MAE) and model confidence set (MCS) approach are applied to assess the prediction abilities. We also compute one-step-ahead out-of-sample VaR of an equally weighted portfolio and perform a backtesting analysis.

The chapter proceeds as follows. Section 2 introduces stochastic correlation models namely, the RiskMetrics EWMA model, DCC, OGARCH and GOGARCH. Section 3 presents the evaluation measures used to compare the forecast performance of different models. Section 4 explains the data involved and presents empirical results on the forecast comparison and section 5 concludes the chapter.

## 3.2 Stochastic Correlation Models

There are many methods to estimate the covariance matrix of a portfolio. In this chapter, we compare the forecasting performance of the models that are widely adopted by market practitioners. In this section, we review these stochastic correlation models.

Let  $y_t$  be a  $k \times 1$  vector multivariate time series of daily log returns on  $k$  assets at time  $t$  :

$$y_t = \mu_t + \epsilon_t \quad (3.1)$$

$$E(y_t | \Omega_{t-1}) = \mu_t \quad (3.2)$$

$$Var(y_t | \Omega_{t-1}) = E(\epsilon_t \epsilon_t' | \Omega_{t-1}) = H_t \quad (3.3)$$

Where  $\Omega_{t-1}$  denotes sigma field generated by the past information until time  $t-1$ .

### 3.2.1 Riskmetrics EWMA

The exponentially weighted moving average (EWMA) models are very popular among market practitioners. The RiskMetrics EWMA model assigns the highest weight to the latest observations and the least weight to the oldest observations in the volatility estimation.

The multivariate form of EWMA model is defined as

$$H_t = \lambda H_{t-1} + (1 - \lambda) y_{t-1} y_{t-1}' \quad (3.4)$$

For each individual element, it is given by

$$\sigma_{i,j,t}^2 = \lambda \sigma_{i,t-1} \sigma_{j,t-1} + (1 - \lambda) y_{i,t-1} y_{j,t-1} \quad (3.5)$$

$\lambda$  is a decay factor, which determines the importance of historical observations used for estimating the covariance matrix. The value of decay factor depends on the sample size

and varies by asset class. J.P. Morgan(1996) suggests that a decay factor of 0.94 is used for daily data set.

Given the decay factor and initial value  $\hat{\Sigma}_0$ , it is very easy to forecast the conditional covariance matrix. The full sample covariance matrix  $\hat{\Sigma}_0$  is defined as

$$\hat{\Sigma}_0 = \frac{1}{T-1} \sum_{t=1}^T (y_t - \bar{y})'(y_t - \bar{y}) \quad (3.6)$$

The EWMA model is very easy to implement for a larger number of assets and the conditional covariance is always semi-definite. The conditional covariance forecasting is also straightforward. However, EWMA model is lack of firm statistical basis and has to estimate decay factor.

### 3.2.2 Multivariate GARCH Models

Following the great success of univariate GARCH model in modeling the volatility, many multivariate GARCH models have been proposed for conditional covariance and correlation modeling. The first direct extension of univariate Garch model is VEC-GARCH model of (Bollerslev et al., 1988) and BEKK models (Engle and Kroner, 1995). When there is a big data set in the portfolio, the large number of covariance is difficult to estimate and evaluation is also complicated. A large number of unknown parameters in these models prevents their successful application in practice.

To solve the larger-scale problems, factor and orthogonal models are introduced. (Bollerslev, 1990) introduced constant conditional correlation (CCC) model with the restriction of constant conditional correlation which reduces the number of parameters. By making the conditional correlation matrix time-dependent, CCC is generalized by the dynamic conditional correlation (DCC) model (Engle, 2002). DCC model captures the advantages of GARCH models and simplifies the computation of multivariate GARCH.

In the DCC model, time series of daily log returns on k assets  $y_t$  is conditionally multivariate normal with mean zero. That is,

$$y_t = \epsilon_t \quad (3.7)$$

$$H_t = D_t R_t D_t \quad (3.8)$$

$$D_t = \text{diag}(\sigma_{11,t}, \dots, \sigma_{kk,t}) \quad (3.9)$$

where  $H_t$  is the conditional covariance matrix and  $R_t$  is the correlation matrix.  $D_t$  is a

diagonal matrix with  $\sigma_{11,t}, \dots, \sigma_{kk,t}$  on main diagonal, which can be estimated from the univariate GARCH models.

As noted by (Engle, 2002),  $R_t$  is also the conditional covariance matrix of the standardized returns. The standardized residuals  $z_t$  is defined as

$$z_t = D_t^{-1} \epsilon_t \quad (3.10)$$

and

$$E(z_t z_t' | \Omega_{t-1}) = R_t \quad (3.11)$$

In this chapter, we consider DCC GARCH (1,1) models by (Engle, 2002). That is, the dynamics for the conditional correlations and the conditional variances follow GARCH-type model. The DCC model can be estimated in two steps.

The first step is to estimate the conditional volatility,  $\sigma_{kk,t}$  by using GARCH (1,1) model

$$\sigma_{kk,t}^2 = \omega_k + \alpha_k \epsilon_{k,t-1}^2 + \beta_k \sigma_{kk,t-1}^2 \quad (3.12)$$

The second step is to estimate the time varying stochastic correlation  $R_t$ .

$$Q_t = (1 - \alpha_0 - \beta_0) \bar{Q} + \alpha_0 z_{t-1} z_{t-1}' + \beta_0 Q_{t-1} \quad (3.13)$$

$$R_t = \text{diag}(Q_t)^{-1} Q_t \text{diag}(Q_t)^{-1} \quad (3.14)$$

where  $Q_t$  is a positive definite matrix defining the structure of the dynamics.  $\bar{Q}$  is the unconditional covariance matrix of  $z_t$  and it can be estimated by

$$\bar{Q} = \frac{1}{T} \sum_{t=1}^T z_t z_t' \quad (3.15)$$

The advantage of DCC model is that it separates the estimation of the volatility for each time series by using single univariate models and the correlation part by imposing the same dynamics to all the correlations. The problem is that  $\alpha_0$  and  $\beta_0$  in DCC are scalars, so all the conditional correlations follow the same dynamics, which is not flexible.

### 3.2.3 Orthogonal GARCH Model

In addition to the DCC model, the orthogonal GARCH (OGARCH) model is introduced by (Alexander and Chibumba, 1997) and (Alexander, 2000) based on univariate GARCH model and principal component analysis (PCA). The OGARCH model is computationally



simpler than the other multivariate GARCH models for a large dimensional covariance matrix because the large number of parameters is reduced by PCA. OGARCH model has achieved outstanding accuracy in correlation forecasting.

In the orthogonal GARCH model, the observed time series are transformed to uncorrelated time series by using PCA. Let  $Y$  be a multivariate time series of daily returns on  $k$  assets with mean zero and length  $T$  and  $y_i$  is the  $i$ th row of  $Y$ . Then the  $T \times k$  matrix  $Y$  can be approximated by

$$Y = PW' \quad (3.16)$$

Where the  $T \times k$  matrix  $P$  is the first  $n$  principal components of the covariance matrix of  $Y$  for  $n \leq k$ .  $W$  is a  $k \times k$  orthogonal matrix of the eigenvectors that are arranged in descending order of the corresponding eigenvalues, so  $W = W^{-1}$ .

Then we can obtain

$$H_t = W\Sigma_tW' \quad (3.17)$$

Where  $\Sigma_t = \text{diag}(\sigma_{1,1,t}^2, \dots, \sigma_{n,n,t}^2)$ .

The conditional variance of the  $i$ th principal component  $p_i$ ,  $i = 1, \dots, N$ , is modeled by GARCH (1,1) model as

$$p_{i,t} = \mu_{i,t} + \epsilon_{i,t} \quad (3.18)$$

$$\sigma_{i,i,t}^2 = \omega_i + \alpha_i \epsilon_{i,t-1}^2 + \beta_i \sigma_{i,i,t-1}^2 \quad (3.19)$$

The OGARCH model reduces unknown parameters significantly and is widely used, but it does not work well when the correlation of time series is very small.

### 3.2.4 Generalized Orthogonal GARCH Model

The GOGARCH model was first proposed by (Van der Weide, 2002) as a natural generalization of the orthogonal GARCH (OGARCH). In OGARCH model, the matrix is assumed to be orthogonal which only contains a very small subset of all possible invertible matrices, while the orthogonal requirement is relaxed in GOGARCH model. (Van der Weide, 2002) argues that OGARCH often underestimate the correlations because of the orthogonal restriction.

Let  $y_t$  be the same time series as defined in OGARCH model and  $y_t$  is transformed to a linear combination of  $n$  uncorrelated factors  $\epsilon_t$  as:

$$y_t = A\epsilon_t \quad (3.20)$$

Where  $A$  is  $k \times k$  non-singular matrix.  $\epsilon_t$  have unconditional unit variance  $\Sigma$  and  $\Sigma = Var(\epsilon_t) = I_n$ .

Here we assume that each unobserved factor  $\epsilon_t$  follows GARCH (1,1) model:

$$\sigma_{i,i,t}^2 = (1 - \alpha_i - \beta_i) + \alpha_i \epsilon_{i,t-1}^2 + \beta_i \sigma_{i,i,t-1}^2 \quad (3.21)$$

$$\Sigma_t = diag(\sigma_{1,1,t}^2, \dots, \sigma_{n,n,t}^2) \quad (3.22)$$

Then the conditional covariance matrix of  $y_t$  is given by

$$H_t = A\Sigma_t A' \quad (3.23)$$

And the unconditional covariance matrix  $H$  of  $y_t$  is defined as  $H = A\Sigma A = AA'$ . Let  $P$  and  $\Lambda$  be orthonormal eigenvectors of  $H$  and a diagonal matrix with the corresponding eigenvalues respectively. Then  $A$  is decomposed by

$$Z = P\Lambda^{1/2}L' \quad (3.24)$$

The parameters can be estimated by two-step approach proposed by (Van der Weide, 2002). First,  $H$  is estimated by sample variance of  $y_t$  and  $\Lambda$  can be estimated after the estimation of  $H$ . Then  $L$ ,  $\alpha$  and  $\beta$  can be estimated by maximizing a multivariate likelihood function.

### 3.3 Model Evaluation Measures

It is very difficult to evaluate the performances of different models because different investors have different concerns. Risk management managers care more about the extreme returns and volatility while portfolio managers may pay more attention to the influence of the correlations on these returns.

Suppose the return of a portfolio return  $y_{p,t}$  at time  $t$  is given by

$$y_{p,t} = \mu_{p,t} + \epsilon_{p,t} \quad (3.25)$$

$$\mu_{p,t} = \sum_{i=1}^n \omega_{i,t} \mu_{i,t} \quad (3.26)$$

$$\sigma_{p,t}^2 = \omega H_t \omega' \quad (3.27)$$

Where  $\mu_{p,t}$  and  $\sigma_{p,t}^2$  are conditional mean and variance of the portfolio at time  $t$  respectively.  $\epsilon_{p,t}$  is Gaussian white noise.  $H_t$  is conditional covariance matrix defined in section 2.  $\omega_{i,t}$  is the portfolio weight of asset  $i$  at time  $t$  and  $\sum_{i=1}^n \omega_{i,t} = 1$ .

The conditional distribution of return for each asset is Gaussian as defined in Section 2 (Engle, 2002), then the portfolio return is also normally distributed because the multivariate normal distribution are closed under linear transformations; see (Pesaran and Zaffaroni, 2005), (Christoffersen, 2009). The portfolio Value-at-Risk (VaR) for 1 day horizon at  $\alpha$  confidence level is

$$VaR_{t+1} = \mu_{p,t} + z_{\alpha} \sigma_{p,t} \quad (3.28)$$

Where  $z_{\alpha}$  is the critical value of the corresponding quantile  $\alpha$ . Here we compute a 95% VaR using  $z_{\alpha} = -1.65$  under normal distribution and we assume  $\mu_{p,t}$  is constant and estimated by the sample mean of the portfolio return (Santos and Moura, 2012).

A number of papers (Plyakha et al., 2012) suggest that equally weighted portfolio strategy (called "1/N") consistently outperforms almost other optimization strategies. Therefore, we use an equally weighted portfolio in our analysis.

### 3.3.1 Mean Absolute Error

We first calculate some popular statistical loss functions in order to evaluate the out of sample forecasting ability of different multivariate models. The loss functions we choose to assess the performance of competing models in volatilities forecasting is mean absolute error (MAE) which is given as follows:

$$MAE = \frac{1}{T} \sum_{t=1}^T \left| \hat{\sigma}_t^2 - h_t \right| \quad (3.29)$$

Where  $h_t$  and  $\hat{\sigma}_t$  are predicted variance and actual volatility at time  $t$  respectively. Since the actual volatility  $\sigma$  is unobservable, we substitute of squared return  $y_t^2$  to the actual conditional variance; see Sadorsky (2006) and Wei et al. (2010).

### 3.3.2 Model Confidence Set Approach

In order to select an optimal model with superior predictive ability in out-of-sample forecasting among these different multivariate models, we consider model confidence set (MCS) approach, the test introduced by Barndorff-Nielsen et al. (2009). The advantage of the MCS is that it performs a joint comparison across a full set of candidate models and does not specify a benchmark model.

Let us denote by  $d_{i,j,t}$  the loss differential between models  $i$  and  $j$  at time  $t$ .  $L_{i,t}$  is a loss function for model  $i$  at time  $t$ , then the null hypothesis is given by

$$H_0: E[d_{i,j,t}] = E[L_{i,t} - L_{j,t}] = 0, \forall j, i \in M \quad (3.30)$$

Where  $M \subset M_0$  and the starting set  $M_0$  contain all the models.

The initial step sets  $M = M_0$ . MCS performs an iterative selection procedure. If the null is rejected at a given confidence level  $\alpha$ , the model with the worst performance is removed from the set at step  $k$ . Repeat this procedure until the null hypothesis cannot be rejected. In order to test  $H_0$ , we use the following t-statistic  $t_{i,j}$ :

$$t_{i,j} = \frac{\bar{d}_{i,j}}{\sqrt{\widehat{Var}(\bar{d}_{i,j})}} \quad (3.31)$$

$$\bar{d}_{i,j} = \frac{1}{T} \sum_{t=1}^T d_{i,j,t} \quad (3.32)$$

where  $\widehat{Var}(\bar{d}_{i,j})$  is an estimate of the variance of average loss differential. p value of the test statistics and  $\widehat{Var}(\bar{d}_{i,j})$  are determined by using a bootstrap approach.

The range test statistics  $t_R = \max_{i,j \in M} |t_{i,j}|$  introduced by (Barndorff-Nielsen et al., 2009) have been used here, because it involves the fewest pairwise comparisons. The elimination rule for the range statistics is  $e_M = \operatorname{argmax}_{j \in M} \sup_{i \in M} t_{i,j}$ . The entire procedure continues to repeat on the smaller set of models until the null hypothesis is not rejected.

In our analysis, the following MAE based loss functions (Patton, 2011) is used:

$$L_{mae} = \left| \hat{\sigma}_t^2 - h_t \right| \quad (3.33)$$

Where the squared return is considered an unbiased proxy of the actual conditional variance.

### 3.3.3 Backtesting

Value at Risk (VaR) is widely used as a measure of risk in portfolio risk management. When we consider conditional stochastic correlations, testing the conditional accuracy of the performance in forecasting becomes important. In order to analyze our results implied by different time-varying volatility models, we compute one-day ahead out-of-sample VaR and perform a backtesting analysis.

The conditional coverage test proposed by (Christoffersen, 1998) is a method to test if the VaR violations are independent and the average number of violations is correct conditional coverage. It is a combination of unconditional test and independence of the violations test.

The sequence of  $VaR_{t,m}$  violations for model m are defined by an indicator function:

$$I_{t,m} = \begin{cases} 1 & \text{if } y_t < -VaR_{t,m}, \\ 0 & \text{if else.} \end{cases}$$

We start by introducing a simple unconditional test for the average probability of a VaR violation. The null hypothesis of correct unconditional coverage test is that  $\pi = \alpha$ , where  $\pi$  is coverage rate in a particular model.

Suppose there are  $T_1$  days in total with  $I_{t,m} = 1$  in a sample period of T, then we get the likelihood function

$$L(\pi) = \pi^{T_1} (1 - \pi)^{T - T_1} \quad (3.34)$$

$$L(\alpha) = \alpha^{T_1} 1 - \alpha^{T - T_1} \quad (3.35)$$

By the likelihood ratio test we get

$$LR_{uc} = -2 \log(L(\alpha)/L(\pi)) \sim \chi^2$$

Under the independence hypothesis, it is assumed that the violation sequence is described as a first order Markov sequence with transition probability matrix

$$\pi = \begin{bmatrix} \pi_{00} & \pi_{01} \\ \pi_{10} & \pi_{11} \end{bmatrix}$$

where  $\pi_{ij} = P(I_{t+1} = j | I_t = i)$  for  $i, j = 0, 1$ . Note that  $1 - \pi_{00} = \pi_{01}$  and  $1 - \pi_{10} = \pi_{11}$ .

The null hypothesis independence test is assumed that  $\pi_{01} = \pi_{11} = \pi$ . The likelihood function for a sample of T observations is

$$L(\pi_1) = \pi_{01}^{T_{01}} (1 - \pi_{01})^{T_{00}} \pi_{11}^{T_{11}} (1 - \pi_{11})^{T_{10}} \quad (3.36)$$

where  $T_{ij}, i, j = 0, 1$  is the number of observations with a j following an i.

Then the independence hypothesis is tested by using a likelihood ratio test

$$LR_{ind} = -2\log(L(\pi)/L(\pi_1)) \sim \chi^2$$

Finally, the likelihood ratio of conditional coverage test is  $L_{cc} = LR_{uc} + LR_{ind}$

### 3.4 Empirical Results

The data used in this chapter is the daily close prices of S&P 500 index and US Generic Government 10 Year Yield bond index. The sample period used here is Sep 9, 2002 through Sep 9, 2013, for a total of 2549 daily observations. We remove common holidays and weekends across these time series to minimize the possibility of inducing spurious correlation. In order to test the robust of the results, the models are applied to two subsample periods. The first period is financial crisis period (2008.4-2009.4) and the second is normal period (2012.10-2013.10). The historical daily returns of S&P 500 and bond index are shown in Figure 3.1 .

First we estimate the dynamic of the conditional correlation for the full sample by using Riskmetrics EWMA, OGARCH, GOGARCH and DCC Garch models. We compare the implied conditional correlations in Figure 3.2 . It shows that the estimated correlation of GOGARCH is quite stable which is around 0.5. However, the correlations estimated by the others are very volatile and EWMA model gives the most volatile correlation curve. The correlations under GOGARCH model have never dropped to 0.4 based on our calculation, even when the colorations estimated by the other model are negative. Furthermore, the

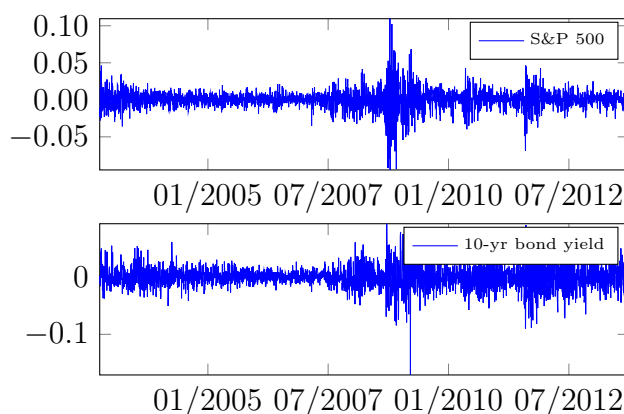


Figure 3.1: Daily return of S&P 500 and 10-yr Bond Index in 2002-2012

correlation estimated by EWMA, OGARCH and DCC Garch models all increase significantly around 2008 except GOGARCH. In Figure 3.2, a very sparse observation appears at the end of 2005 with mainly negative correlations in 2004-2006. A peak also appears in the conditional correlation estimated in 2008. The reason for the large positive correlations is an increase in uncertainty about the economic outlook in 2008 global financial crisis.

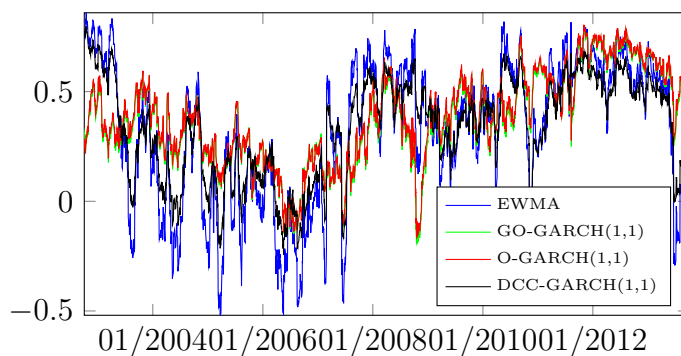


Figure 3.2: In sample estimation of stochastic correlation between S&P 500 and 10-yr Bond Index, 2002-2013

Now let us compare these models by investigating the forecasting performance of the 95% VaR for an equally weighted portfolio by using MAE, MCS and conditional coverage test. We also set the number of bootstrap samples up to 10,000 in order to obtain the distribution under the null. The critical value of  $LR_{c,c}$  statistic with two degrees of freedom used throughout the backtesting process is 5.99 at 95% confidence level. Jorion (2001) suggests that 95% confidence level fits well for backtesting purposes. We construct MCS test at 10% confidence level using range statistics.

Table 3.1 illustrates some test statistics of conditional covariance estimated by EWMA, DCC GARCH (1,1), OGARCH (1,1) and GOGARCH (1,1) for the whole period. The

best performance in forecasting under the in-sample analysis is OGARCH with the least MAE. GOGARCH has largest MAE because of the flat curve in Figure 3.2. The 10% MCS consist solely of OGARCH model and all these models except that EWMA fails the conditional coverage test for 95% VaR. OGARCH(1,1) outperforms the other models according to all evaluation criteria by in sample analysis.

	MAE	MCS	$LR_{cc}(0.95)$
EWMA	4.13E-05	0.0096	0.945
DCC	4.10E-05	0.0834	5.805
OGARCH	3.62E-05	1	Inf
GOGARCH	4.22E-05	0.0834	10.6589

Table 3.1: Test statistics for in sample estimation for the whole period

### 3.4.1 First Subsample Period

In order to examine the time-varying nature of the conditional correlations, as well as to investigate the forecasting performance of these models in a turbulent period, we split the data sample into 2 subsamples. In the first subsample, we only consider the first 1777 observations in order to examine the performance of the models during 2008 financial crisis period. We split the subsample into two parts, 5.5-year estimation period which is used for in sample estimation and the subsequent 1.5-year out of sample forecast periods. We use a rolling window size of 1413 days to forecast the conditional covariance in financial crisis.

Figure 3.3 presents the forecasting performance of the four models mentioned above during 2008 financial crisis. Compared with Figure 3.2, the estimated correlation by GOGARCH is more volatile and the correlation reaches a peak in September and October 2008 when the subprime mortgage crisis hit its peak. Lehman Brothers filed the largest bankruptcy in U.S. history in Sep. 2008. The correlation curves are very close to each other always positive between April 2008 and April 2009.

Table 3.2 reports some test statistics of out of sample prediction performance of the 95% VaR for an equally weighted portfolio by using MAE, MCS and conditional coverage test. All these four models pass the conditional coverage test for 95% VaR during the financial crisis period and all these models are included in MSC except EWMA. That means, all these models perform well during the crisis period. We conclude OGARCH has the best performance based on all the test statistics in this period.



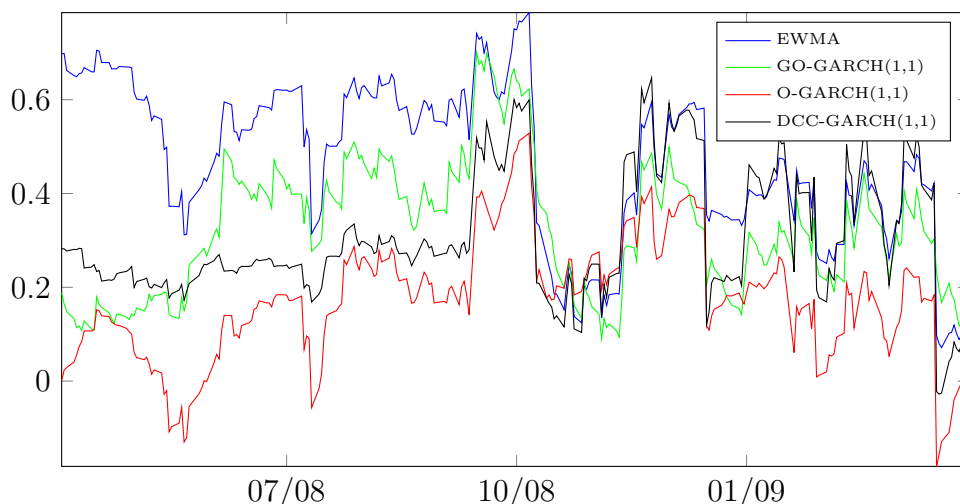


Figure 3.3: 1-day ahead out-of-sample conditional correlation forecast between S&P 500 and 10-yr Bond Index in 2008-2009

	MAE	MCS	$LR_{cc}(0.95)$
EWMA	1.04E-04	0.0066	0.016
DCC	9.80E-05	0.0139	1.444
OGARCH	9.61E-05	1	1.119
GOGARCH	9.71E-05	0.0165	1.119

Table 3.2: Test statistics for out-of -sample forecast in 2008-2009

### 3.4.2 Second Subsample Period

In the second subsample, we estimate the parameters by using the 1777 observations in first 7 years and do one step ahead forecasts for the last 3 years with rolling window of 1777 days. We apply the proposed models to obtain out-of-sample one-step-ahead forecasts of the conditional covariance matrix of all assets.

From Figure 3.4 we can see the curves are overlapping and very close to each other. All the models pass conditional coverage test based on table 3.3. EWMA and DCC both have very volatile correlations. The 10% MCS consists OGARCH and GOGARCH. All the test statistics shown in Table 3.3 suggest that OGARCH and GOGARCH models both perform well and OGARCH achieves the best performance.

## 3.5 Conclusion

In this chapter, we investigate the forecasting ability in different market conditions of four multivariate stochastic correlation models: EWMA, DCC, OGARCH and GOGARCH.

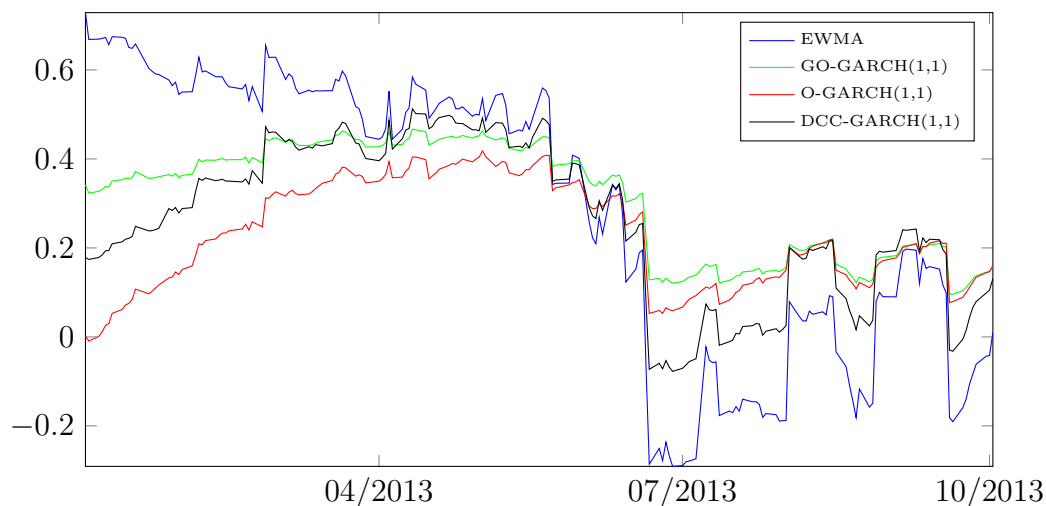


Figure 3.4: 1-day ahead out-of-sample conditional correlation forecast between S&P 500 and 10-yr Bond Index in 2013

	MAE	MCS	$LR_{cc}(0.95)$
EWMA	5.11E-05	0	0.0734
DCC	5.09E-05	0.001	1.0371
OGARCH	4.90E-05	1	0.0135
GOGARCH	4.92E-05	0.161	0.2164

Table 3.3: Test statistics for out-of -sample forecast in 2013

All these models have quite good forecasting performances during 2008 financial crisis and after 2008 crisis periods. GOGARCH and OGARCH models outperform the others. OGARCH achieves the best performance during these two sub-periods. During the stable market conditions, the out of sample conditional correlations are very different among these models. OGARCH model outperforms all the other three models under stable market conditions. We conclude that the out-of-sample forecast results show that OGARCH model has the best performance. The overall performance of multivariate GARCH models is better than EWMA.

Our empirical findings provide a better understanding of the dynamics in the correlations between stock index and bond yield, which can be applied by portfolio managers, risk management managers, policy makers and researchers. The use of OGARCH model in estimating time-varying correlations may reduce market risk and improve the performance of the portfolios.

# Chapter 4

## Portfolio Optimization in Hedge Funds by OGARCH and Markov Switching Model

### 4.1 Introduction

The hedge fund industry has grown rapidly in recent years and has become more and more important in alternative investment. According to the twelfth annual Alternative Investment Survey in 2014 by Deutsche Bank, hedge funds are estimated to manage assets of more than 3 trillion dollars by the end of 2014. The diversification benefits of hedge funds which allow investors to hold portfolios of other investment funds rather than investing directly in stocks, bonds and other securities are the main factor driving its success (Martellini et al., 2007). However, the 2008 financial crisis caused assets under management to fall sharply because of trading losses and the withdrawal of assets from funds by investors, resulting in a decline of assets by nearly 30% in 2008 (Maslakov, 2009).

The recent financial turbulence exposed and raised serious concerns about the optimal portfolio selection problem in hedge funds. Many papers have examined portfolio optimization in a hedge fund context. The structures of hedge fund return and covariance are crucial in portfolio optimization. The non-normal characteristics of hedge fund returns have been widely described in the literature. Kat and Brooks (2001) find that the hedge fund returns exhibit significant degrees of negative skewness and excess kurtosis. According to Getmansky et al. (2004) and Agarwal and Naik (2004) the returns of hedge fund return are not normal and serially correlated.

Meanwhile, a number of empirical studies show that the correlations of hedge fund return

time series are time-varying. Billio et al. (2012) and Blazsek and Downarowicz (2008) have proposed more Regime-switching models to measure dynamic risk exposures of hedge funds. A non-linear Markov switching GARCH (MS-GARCH) model is proposed by Blazsek and Downarowicz (2013) to forecast idiosyncratic hedge fund return volatility. In the other direction, multivariate Garch models are employed to estimate the time-varying covariances/correlations of hedge fund returns. Giamouridis and Vrontos (2007) show that the optimal hedge fund portfolio constructed by dynamic covariance models has lower risk and higher out-of-sample risk-adjusted realized return. Harris and Mazibas (2010) provide further evidence that the use of multivariate GARCH models optimal portfolios selected by using multivariate GARCH models performance better than static models and also show that exponentially weighted moving average (EWMA) model improves the portfolio performance. Saunders et al. (2013) apply Markov-switching model in hedge fund portfolio optimization and show that Markov-switching model outperforms Black-Scholes Model and Gaussian Mixture Model.

In this chapter, we extend the results of Saunders et al. (2013) to compare the optimal portfolio performances selected by the OGARCH model, the two-state Regime-Switching models and the EWMA Model, using daily observations of HFRX indices for the period 2003-2014. We first detect the exact 2008 financial crisis period by using the Regime-Switching model. Then the out-of-sample portfolio performance in the 2007-2009 financial crisis period and the whole sample period are analyzed based on the Sharpe ratio and the mean realized return. Another contribution of the chapter is that we calculate the asset weight sensitivities in optimal Mean-Variance portfolio to the estimated parameters in OGARCH model.

The rest of this chapter is organized as follows. Section 2 introduces the OGARCH model, the Regime-Switching models and the EWMA model. The data set is described and the out-of-sample performance of optimal portfolios is analyzed in Section 3. Section 4 discusses the asset-weight sensitivities with respect to the parameters in the OGARCH Model. Section 5 concludes the chapter.

## 4.2 Methodology

In this section, we first introduce the different models that can be used to estimate the mean and variance of hedge fund index returns. Then we describe the Mean-Variance portfolio optimization model used for optimal portfolio selection.

### 4.2.1 OGARCH Model

GARCH models have been applied to model volatility with a good success to capture some stylized facts of financial time series, such as fat tails and volatility clustering. One good extension from a univariate GARCH model to a multivariate case is the orthogonal GARCH (OGARCH) model introduced by Ding (1994), Alexander and Chibumba (1997) and Alexander (2000) which is based on univariate GARCH model and principal component analysis. Thereafter, OGARCH has become very popular to model the conditional covariance of financial time series. The OGARCH model is computationally simpler than the multivariate GARCH models for a large dimensional covariance matrix and has achieved outstanding accuracy in forecasting correlation.

In the OGARCH model, the observed time series are linearly transformed to a set of independent uncorrelated time series by using principal component analysis. The principal component approach is first used in a GARCH type context by Ding (1994). The OGARCH model by Alexander (2000) is described as follows.

Let  $Y_t$  be a multivariate time series of daily returns with mean zero on  $k$  assets with length  $T$  with columns  $y_1, \dots, y_k$ . Then the  $T \times K$  matrix  $X_t$  whose columns  $x_1, \dots, x_k$  are given by the equation

$$x_t = \frac{y_t}{\sqrt{v_i}}$$

where  $V = \text{diag}(v_1, \dots, v_m)$  with  $v_i$  being the sample variance of the  $i$ th column of  $Y_t$ .

Let  $L$  denote the matrix of eigenvectors of the population correlation of  $x_t$  and by  $l_m = (l_{1,m}, \dots, l_{k,m})$  its  $m$ th column.  $l_m$  is the  $k \times 1$  eigenvector corresponding to the eigenvalue  $\lambda_m$ . The column labelling of  $L$  has been chosen so that  $\lambda_1 > \lambda_2 > \dots > \lambda_k$ . Let  $D$  be the diagonal matrix of eigenvalues. The  $m$ th principal component of the system is defined by

$$p_m = x_1 l_{1,m} + x_2 l_{2,m} + \dots + x_k l_{k,m}$$

If each vector of principal components  $p_m$  is placed as the columns of a  $T \times k$  matrix  $P$ , then,

$$P = XL$$

The principal component columns in modeled by GARCH(1,1)

$$p_t | \Psi_{t-1} \sim N(0, \Sigma_t)$$

$$p_{i,t} = \epsilon_{i,t}$$

$$\sigma_{i,t}^2 = \omega_i + \alpha_i \epsilon_{i,t-1}^2 + \beta_i \sigma_{i,t-1}^2$$

where  $\Sigma_t$  is a diagonal matrix of the conditional variances of the principal components  $P$ .  $\Psi_{t-1}$  contains all the information available up to time  $t-1$ . The Conditional Covariance Matrix of  $X_n$  is  $D_t = L\Sigma_tL_n^T$  and the conditional covariance matrix of  $Y$  is given by

$$H_t = \sqrt{V}D_t\sqrt{V}$$

The estimation procedure in detail is illustrated in the following section.

### Model Estimation in OGARCH

Let  $Y$  be a  $T \times k$  matrix of daily returns on  $k$  assets for  $T$  days. We obtain the daily returns from the daily closing prices by taking the natural log of the quotient of consecutive observations; in other words

$$y_i = \log \frac{P_{i+1}}{P_i}$$

where  $P_i$  is the daily closing price at time  $i$ .

Step 1: We standardize the data into a  $T \times k$  matrix  $X$  with the estimated variance and mean for each  $y_i$  and find the correlation matrix  $XX'$ .

Step 2: Then principal components analysis is performed on  $XX^t$  to get the eigenvectors and the eigenvalues. We denote the matrix of eigenvectors by  $L$  and its  $m$ th column by  $l_m = (l_{1,m}, \dots, l_{k,m})$ , the  $k \times 1$  eigenvector corresponding to eigenvalue  $\lambda_m$ . The column labeling has been chosen so that  $\lambda_1 > \lambda_2 > \dots > \lambda_k$ .

Step 3: We decide how many principal components we should use. If first  $n$  principal components is chosen, the  $m$ th principal component of the system is

$$p_m = x_1l_{1,m} + x_2l_{2,m} + \dots + x_nl_{n,m}$$

where  $x_i$  is the  $i$ th column of  $X_n$ , the  $T \times n$  matrix extracted from  $X$ . Then the principal components matrix  $P$  is given by a  $T \times n$  matrix and we have  $P = X_nW_n$ .

Step 4: The conditional variance of the  $i$ th principal component  $p_i$ ,  $i = 1, \dots, N$ , is estimated by GARCH(1,1):

$$\begin{aligned} p_{i,t} &= \epsilon_{i,t} \\ \sigma_{i,t}^2 &= \omega_i + \alpha_i \epsilon_{i,t-1}^2 + \beta_i \sigma_{i,t-1}^2 \\ \Sigma_t &= \text{diag}(\sigma_{1,t}^2, \dots, \sigma_{n,t}^2) \end{aligned}$$

Step 5: The Conditional Covariance Matrix of  $X_n$  is  $D_t = W_n \Sigma_t W_n^t$  and the conditional covrariance matrix of  $Y$  is given by

$$H_t = \sqrt{V} D_t \sqrt{V}$$

where  $W_n = L_n \text{diag}(\sqrt{\lambda_1}, \dots, \sqrt{\lambda_n})$ .

The accuracy of the conditional covariance matrix  $V_t$  of the original returns, is determined by how many components  $n$  are selected to represent the system.

## 4.2.2 Markov Switching Model

Regime switching models have become very popular in financial modeling since the seminal contribution of Hamilton (1989). Hamilton first proposed the Markov switching model (MSM) to model the real GNP in the US. Since then, these models have been widely used to model and forecast business cycles, foreign exchange rates and the volatility of financial time series.

Suppose the return of  $I$  hedge fund indices  $i \in \{1, \dots, I\}$  and one global stock index  $i = 0$  follow a discrete-time Markov switching process. There exists an unobservable process  $\{S_t\}_{t=1}^T$  with state space  $\{0, 1\}$ . Each index has a drift and volatility parameter. The log-returns of each index are given by  $R_t^i = \mu_{S_t}^i + \sigma_{S_t}^i \varepsilon_t^i$  where  $(\varepsilon_t^0, \dots, \varepsilon_t^I)$  is a multivariate normal with zero mean, unit standard deviation and correlation matrix  $C_{S_t}$ . Then the state process  $S$  is modelled as a time-homogeneous Markov chain with transition matrix

$$M = \begin{pmatrix} p & 1-p \\ 1-q & q \end{pmatrix}$$

where  $p = \mathbb{P}(S_t = 1 | S_{t-1} = 0)$  and  $q = \mathbb{P}(S_t = 0 | S_{t-1} = 1)$  with initial distribution of  $S$  as  $(r, 1-r)$  with  $r = \mathbb{P}(S_1 = 0)$ . The parameters in MSM is denoted by the vector:  $\theta^{MSM} = (p, q, \mu_0, \mu_1, \Sigma_0, \Sigma_1, r)$ , for any  $p, q \in [0, 1]$ ,  $\mu_0, \mu_1 \in \mathbb{R}^{I+1}$ ,  $\Sigma_0, \Sigma_1 \in \mathbb{R}^{(I+1) \times (I+1)}$ ,  $r \in [0, 1]$ .

The Gaussian mixture model (GMM) is considered as a special case of the Markov-switching model where state process  $S_t$  is an iid. random sequence. In this case, the transition probabilities are given by  $p = \mathbb{P}(S_t = 1) = 1-r$  and  $q = \mathbb{P}(S_t = 0) = r$  with  $p+q = 1$ . Therefore, the state process  $S$  is reduced from a Markov chain to a Bernoulli process. The transition

matrix becomes

$$M = \begin{pmatrix} p & 1 - p \\ p & 1 - p \end{pmatrix}$$

The log-returns of each time series are iid. Gaussian mixture random vectors have two components with component weights  $(p, 1 - p)$ . The parameters are denoted by  $\theta^{GMM} = (p, \mu_0, \mu_1, \Sigma_0, \Sigma_1)$ . For the interested reader, the model estimation procedure is described as follows.

### Model Estimation in Regime Switching

The state process  $S$  indicates whether the market is in a normal or distressed regime and it is not directly observable based on market data. Therefore it must be estimated indirectly using market observables as proxies to make inferences regarding its history and parameters.

We focus on the returns of the hedge fund indices, in particular, a major stock index as the crisis indicator to identify distressed market regimes. We use MSCI World Index as the crisis indicator rather than the indices of various hedge funds for two reasons. First, the MSCI World Index is the best representative of the market as a whole. The second reason is that the high frequent data of MSCI World Index is available for longer periods and more reliable than hedge fund data. The estimator obtained by the following steps:

Step 1: Estimate the model parameters for the asset's log-returns for  $i = 0$  and the parameters for the state process  $S$  when only historical data of the global stock index are considered. Then compute the most likely sequence of states.

Step 2: After the market is separated into two regimes by step 1, the mean and variance  $\theta^{MSM}, \theta^{GMM}$  can be calculated. We follow the estimation procedure described in Saunders et al. (2013). The RHmm Package (Taramasco et al., 2013) and Matlab code by Perlin (2010) are used to finish the calculation.

### 4.2.3 EWMA Model

The exponentially weighted moving average (EWMA) models are very popular among market practitioners. The EWMA model assigns the highest weight to the latest observations and the least to the oldest observations in the volatility estimate.



The variance  $\Sigma_t$  in multivariate EWMA model is defined as

$$\Sigma_t = (1 - \lambda)\Sigma_{t-1} + \lambda y'_{t-1} y_{t-1}$$

For each individual element, it is given by

$$\sigma_{ij}^2, j, t = (1 - \lambda)\sigma_{i,t-1}\sigma_{j,t-1} + \lambda y_{i,t-1}y_{j,t-1}$$

where  $\lambda$  is the decay factor which determines the importance of historical observations used for estimating the covariance matrix. The value of the decay factor depends on the sample size and varies by asset class. J.P. Morgan(1996) suggests that a decay factor of 0.94 is used for the daily data set.

Given the decay factor and initial value  $\hat{\Sigma}_0$ , it is very easy to forecast the conditional covariance matrix.  $\hat{\Sigma}_0$  is usually the full sample covariance matrix, which is defined as

$$\hat{\Sigma}_0 = \frac{1}{T-1} \sum_{t=1}^T (y_t - \bar{y})'(y_t - \bar{y})$$

The EWMA model is very easy to implement for a larger number of assets since the conditional covariance is always semi-definite. The forecast of the conditional covariance is also straightforward. The downside of this model is a lack of firm statistical basis and also the need to estimate the decay factor.

#### 4.2.4 Portfolio Optimization Model

The optimal portfolio is selected by the mean variance model(MV) introduced by Markowitz (1952). Suppose that the portfolio consists of k assets with the return  $r_{i,t}$  for asset i at time t and  $R_t = (r_{1,t}, r_{2,t}, \dots, r_{k,t})$ . Let  $w_{i,t}$  be the weight of the portfolio invested in asset i at time t and  $w_t \in \mathfrak{R}^n$  is the vector weight of the portfolio at time t. Given a risk aversion coefficient  $\lambda$ , the optimal portfolio weight  $w_t$  is selected by minimizing the following utility function at the beginning of time t by:

$$\text{minimize} \{ \lambda w_t^T \text{Var}_{t-1}(R_t) w_t - w_t^T E_{t-1}(R_{n,t}) \} \quad (4.1)$$

subject to

$$w_t^T \mathbf{1} = 1 \quad (4.2)$$

where  $Var_{t-1}(R_t)$  is the conditional covariance of the asset returns at  $t$ ,  $E_{t-1}(R_{n,t})$  is the conditional expected returns at  $t$ . In the following example, the same risk aversion coefficient  $\lambda = 11.39$  is used.

### 4.3 Empirical Results

This section presents the results of the performance of the optimal portfolios selected by using OGARCH, MSM, GMM and EWMA models in an out-of-sample period. The weight sensitivities with respect to the parameters in the OGARCH model are also analyzed.

The daily observations of HFRX indices for the period from April 1, 2003 to May 12 2014 and MSCI World index from January 1, 1982 to May 12, 2014 are used to estimate the models and test the out-of-sample performance. The MSCI World Index time series is used to identify crisis period in MSM and GMM models. Four HFRX indices are used to construct the optimal portfolio: Equity Market Neutral (EMN), Equity Hedge (EH), Macro (M) and Merger Arbitrage (MA). These fund indices are the same as the paper (Saunders et al., 2013) for details).

First we use the daily returns of MSCI World Index over 30 year period (Jan.1, 1982 to May 12, 2014) to identify the crisis. The period (32-90) is identified as the financial crisis period (7/17/2007 to 11/2/2009) by using Viterbi path.

The parameters are estimated based on the data available in-sample-period (4/1/2003 to 4/7/2006) and the first optimal portfolio is selected on April 8, 2006. The portfolio is rebalanced every two weeks after the first investment period based on the information available up to the investment date. The realized portfolio returns based on the different models in this out-of-sample period are reported in Figure 4.1. As shown in Figure 4.1, the portfolio returns are very volatile during the financial crisis period and the portfolio returns selected by OGARCH is less volatile than the others.

In Figures 4.2, 4.3, 4.4 and 4.5, we illustrate the optimal weight allocations among EMN, EH, M and MA during the out-of-sample period based on OGARCH, EWMA, MSM and GMM. As shown in Figures 4.2 and 4.5, MSM and GMM investors put largest weights in EMN and EH during financial crisis period. More wealth is invested in EH by MSM and GMM before 2008. Figure 4.4 shows that EWMA investor mainly invest in MA over the entire out-of-sample period. The average weight allocated in MA by EWMA investor during the financial crisis is more than 67%, while GMM investor puts almost 50% of the wealth in EMN (see Table 4.1). Since MA fund index has the best performance during

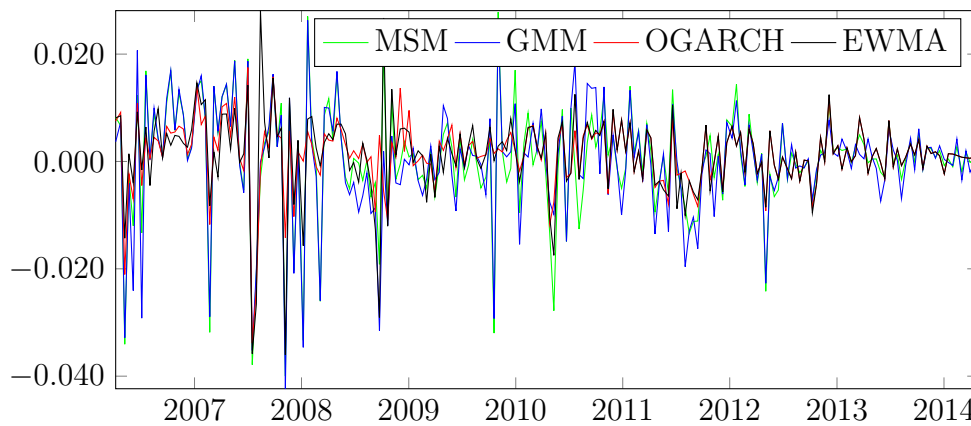


Figure 4.1: 10-day realized returns of the optimal portfolio selected by different models.

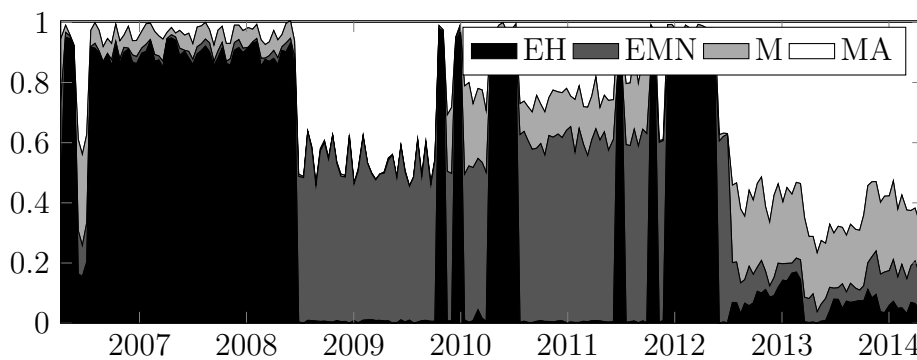


Figure 4.2: MSM: optimal Portfolio Allocation. Each position along the vertical axis represents the weight percentage of each asset in optimal portfolio.

financial crisis period, so OGARCH and EWMA has better cumulative returns during this period. After financial crisis, all the investors except GMM allocate more wealth on EH. After 2011, EWMA investor invests all the wealth in MA.

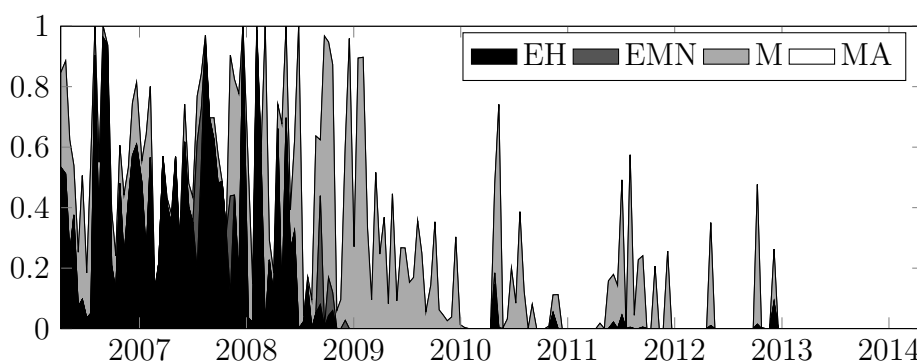


Figure 4.3: OGARCH: optimal weight of each asset allocated in Portfolio Allocation.

Table 4.1 presents descriptive statistics for each optimal portfolio returns selected by using EWMA, GMM, MSM and OGARCH for the period April 2006 to May 2014. For 10-day

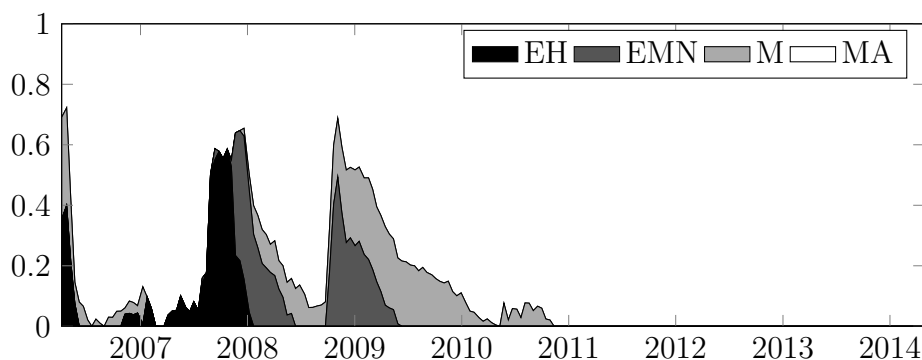


Figure 4.4: EWMA: optimal weight of each asset allocated in Portfolio Allocation.

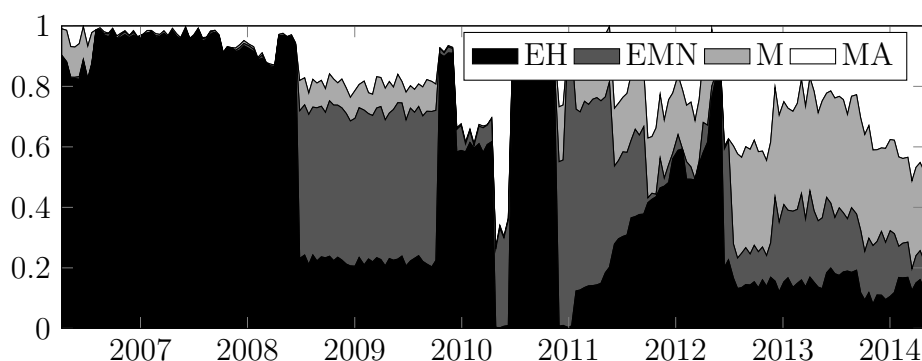


Figure 4.5: GMM: optimal weight of each asset allocated in Portfolio Allocation.

	EH(%)	EMN (%)	M(%)	MA(%)	Cum Return(%)
Average return	-23.69	- 10.59	-3.95	8.93	
OGARCH- average weight	18.10	2.56	29.94	49.40	5.51
EWMA- average weight	7.47	11.97	13.00	67.57	3.58
GMM- average weight	22.46	49.61	8.66	19.27	-5.17
MSM- average weight	0.63	53.59	0.43	45.34	-2.21

Table 4.1: Optimal portfolio allocation for financial crises periods 32-90. The second row gives the average 10-day returns of each fund index.

returns of each portfolio, mean, standard deviation, 5% quantile, minimum and Sharpe ratio are reported. As can be seen from Table 4.1, OGARCH has the lowest risk because the standard deviation in the 5% quantile and worst case return are the smallest. EWMA and OGARCH models outperform from regime switching models across all the measures. Based on the Sharpe ratio, the optimal portfolio selected by OGARCH model has the highest risk-adjusted return. Tables 4.3 and 4.4 also show that the optimal portfolio selected by OGARCH model has the best performance during both the financial crisis period and normal period.

	EWMA	GMM	MSM	OGARCH
Mean(%)	0.12	-0.01	0.03	0.13
St. Dev (%)	0.77	1.07	1.03	0.61
5% Quantile	-1.02	-2.31	-2.04	-0.88
Min (%)	-3.60	-4.23	-4.00	-3.38
Sharpe Ratio (%)	0.159	-0.010	0.032	0.217

Table 4.2: Statistics of realized portfolio returns for the out-of-sample period. Average 10-day portfolio return(Mean), standard deviation(St. Dev), worst-case return(Min) are in percentages.

	EWMA	GMM	MSM	OGARCH
Mean (%)	0.06	-0.27	-0.15	0.09
St. Dev (%)	1.16	1.34	1.33	0.79
5% Quantile	-2.91	-3.46	-3.33	-1.42
Worst Case(%)	-3.60	-4.23	-4.00	-3.38
Sharpe Ratio (%)	0.052	-0.203	-0.113	0.118

Table 4.3: Statistics of realized portfolio returns for crisis periods(32-90). Average 10-day portfolio return(Mean), standard deviation(St. Dev), 5% quantile, worst-case return(Min) and Sharpe ratio are in percentages.

	EWMA	GMM	MSM	OGARCH
Mean (%)	0.10	0.03	0.06	0.10
St. Dev (%)	0.36	0.48	0.51	0.35
5% Quantile	-0.42	-0.71	-0.64	-0.42
Min (%)	-0.96	-2.27	-2.42	-0.92
Sharpe Ratio (%)	0.271	0.068	0.118	0.273

Table 4.4: Statistics of realized portfolio returns for normal periods(145-204). Average 10-day portfolio return(Mean), standard deviation (St. Dev), 5% quantile, worst-case return(Min) and Sharpe ratio are in percentages.

## 4.4 Asset Weight Sensitivities

It is worth investigating the weight sensitivities of the optimal portfolio with respect to the parameters in OGARCH. Chopra and Ziemba (1993) show that even small changes to the estimated expected returns or variance can produce vastly different optimal portfolios selected by Mean-variance model. The sensitivities analysis can help to understand the relationship between the input estimates and output optimal portfolio. The weight sensitivities measure the weight change of each asset in the optimal portfolio to the change of the parameters in the OGARCH model. Risk managers would want to use sensitivity analysis to adjust the optimal portfolio weights to keep the portfolio optimized. Another application is that it could be used to evaluate the impact that each individual or group of assets has on the portfolio variance. This would help risk managers find the major

sources of risk and allow them to evaluate the estimation errors.

The asset weight sensitivities can be inferred and approximated using finite differences. For example, we try to calculate the value of the weight sensitivities of  $\alpha = [\alpha_1, \alpha_2, \alpha_3, \alpha_4]$ . First we find the estimated parameters for  $\alpha = [\alpha_1, \alpha_2, \alpha_3, \alpha_4]$  and get the optimal weights of each assets  $w(\alpha)$ . Then we take the  $\Delta = 0.01\alpha$  and find the new weight  $w(\alpha + \Delta)$  when  $\alpha$  increases by 1%. The weight sensitivities with respect to  $\alpha$  can be approximated by  $\frac{w(\alpha+\Delta)-w(\alpha)}{\Delta}$ , where  $w$  is the optimal weight. The results by using both methods are the same as shown in Table 5.5 and 4.6.

#### 4.4.1 Explicit Calculation

Alternatively, the value of sensitivities can be calculated directly by differentiating the optimal conditions. We assume that the portfolio is optimized based on the Mean-Variance model by Markowitz(1952) and short selling is allowed. If we want to know the rate of change in the weights with respect to the estimated parameters, then we need to calculate the sensitivities. To simply the notation, let

$$C_t = Var_{t-1}(R_t) = Var[R_t|\Psi_{t-1}] = H_t$$

$$\mu_n = E_{t-1}(R_{n,t}) = E[R_t|\Psi_{t-1}] = W(\mu_1, \dots, \mu_k)^T$$

Let  $\theta_i = (\mu_i, \gamma_i, \alpha_i, \beta_i)$ , then the partial derivative of the variance for each asset  $i$  is

$$\frac{\partial \sigma_{i,t}^2}{\partial \theta_i} = \begin{bmatrix} \frac{\partial \sigma_{i,t}^2}{\partial \mu_i} \\ \frac{\partial \sigma_{i,t}^2}{\partial \gamma_i} \\ \frac{\partial \sigma_{i,t}^2}{\partial \alpha_i} \\ \frac{\partial \sigma_{i,t}^2}{\partial \beta_i} \end{bmatrix} = \begin{bmatrix} -2\alpha_i \epsilon_{i,t-1} \\ 1 \\ \epsilon_{i,t-1}^2 \\ \sigma_{i,t-1}^2 \end{bmatrix} + \beta \frac{\partial \sigma_{i,t-1}^2}{\partial \theta_i}$$

$$\frac{\partial \mu_{i,t}}{\partial \theta_i} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

Given the Lagrange multiplier  $\rho$ , the optimal condition by solving the Lagrangian for the constraint in Markowitz's models given by

$$2\lambda C_t w_t - \mu_n + \rho \mathbf{1} = 0 \tag{4.3}$$

$$w^T \mathbf{1} = 1 \tag{4.4}$$

Because the Lagrange multiplier and the optimal weight both depend on the parameter  $\theta$ , we differentiate equations above

$$2\lambda C \nabla_{\theta} w + 2\lambda \frac{\partial C}{\partial \theta} w - \frac{\partial \mu_n}{\partial \theta} + \frac{\partial \rho}{\partial \theta} \mathbf{1} = 0 \quad (4.5)$$

$$\mathbf{1}^T \nabla_{\theta} w = 0 \quad (4.6)$$

We take the parameters  $\mu_1, \gamma_1, \alpha_1$  and  $\beta_1$  for example and the derivatives of the other parameters can be solved in the same way. Given that there are four assets in this portfolio, we let  $\mu_t = (\mu_1, \mu_2, \mu_3, \mu_4)$  and  $\Sigma_t = \text{diag}(\sigma_{1,t}^2, \dots, \sigma_{4,t}^2)$

$$\frac{\partial \mu_n}{\partial \mu_1} = W \frac{\partial \mu_t}{\partial \mu_1} = W \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\frac{\partial \mu_n}{\partial \gamma_1} = \frac{\partial \mu_n}{\partial \alpha_1} = \frac{\partial \mu_n}{\partial \beta_1} = 0$$

$$\frac{\partial C_t}{\partial \mu_1} = W \left( \begin{bmatrix} -2\alpha_1 \epsilon_{1,t-1} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} + \beta \frac{\partial C_{t-1}}{\partial \mu_1} \right) W^T$$

$$\frac{\partial C_t}{\partial \gamma_1} = W \left( \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} + \beta \frac{\partial C_{t-1}}{\partial \mu_1} \right) W^T$$

$$\frac{\partial C_t}{\partial \alpha_1} = W \left( \begin{bmatrix} \epsilon_{1,t-1}^2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} + \beta \frac{\partial C_{t-1}}{\partial \mu_1} \right) W^T$$

$$\frac{\partial C_t}{\partial \beta_1} = W \left( \begin{bmatrix} \sigma_{1,t-1}^2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} + \beta \frac{\partial C_{t-1}}{\partial \mu_1} \right) W^T$$

In order to compare the sensitivities in normal and crisis periods, we consider both Period 60(normal) and 180(crisis) respectively.

The values of weight sensitivities are given in Table 4.5 in financial crisis period (period 60). The asset weight is very sensitive to the changes of the estimated parameters  $\beta$ ,  $\gamma$  and  $\alpha$ . The different sign of the sensitivities shows that different assets respond differently to the changes. For example, when  $\gamma$  increases, the weight of EH and M decrease while the weight of EMN and MA increase. The weight sensitivities of EH, EMN and MA are largest with respect to  $\gamma$ . That is, the optimal weights are very sensitive to  $\gamma$ . When the parameters are estimated, we should pay attention to estimation errors.

Compared with the asset weight sensitivities in normal economic conditions shown in Table 4.6, we conclude that the weights of EMN and MA are more sensitive to the parameters during the financial crisis period. This is reasonable because the optimal weights are more volatile in times of financial crisis. The weight sensitivities are zero for EMN because the weight of EMN during this period is 0 and does not change at all.

	EH(%)	EMN (%)	M(%)	MA(%)
$w(\gamma - \Delta)$	0.0786	0.0155	0.4926	0.4133
$w(\gamma)$	0.0783	0.0157	0.4926	0.4134
$w(\gamma + \Delta)$	0.0780	0.0160	0.4926	0.4134
$\frac{w(\gamma) - w(\gamma - \Delta)}{\Delta}$	-0.8266	6.0328	-0.0062	3.8991
$\frac{w(\gamma + \Delta) - w(\gamma)}{\Delta}$	-0.8244	6.0114	-0.0053	3.8769
$w(\alpha - \Delta)$	0.0785	0.0156	0.4926	0.4133
$w(\alpha)$	0.0783	0.0157	0.4926	0.4134
$w(\alpha + \Delta)$	0.0781	0.0158	0.4927	0.4134
$\frac{w(\alpha) - w(\alpha - \Delta)}{\Delta}$	-0.5861	0.6677	0.1303	0.5240
$\frac{w(\alpha + \Delta) - w(\alpha)}{\Delta}$	-0.5853	0.6658	0.1303	0.5239
$w(\beta - \Delta)$	0.0765	0.0170	0.4928	0.4138
$w(\beta)$	0.0783	0.0157	0.4926	0.4134
$w(\beta + \Delta)$	0.0801	0.0145	0.4925	0.4130
$\frac{w(\beta) - w(\beta - \Delta)}{\Delta}$	0.1941	-0.1259	-0.0185	-0.0416
$\frac{w(\beta + \Delta) - w(\beta)}{\Delta}$	0.1865	-0.1258	-0.0154	-0.0401

Table 4.5: Weight sensitivities for selected values of OGARCH parameters in period 60. The numbers in each column are in percentage. The second row gives the optimal weight of each fund index when gamma equals  $\gamma - \Delta$



	EH(%)	EMN (%)	M(%)	MA(%)
$\frac{w(\gamma)-w(\gamma-\Delta)}{\Delta}$	-1.0244	0	0.0594	-0.2115
$\frac{w(\gamma+\Delta)-w(\gamma)}{\Delta}$	-0.9973	0	0.0575	-0.1918
$\frac{w(\alpha)-w(\alpha-\Delta)}{\Delta}$	-0.1582	0	-inf	0.3402
$\frac{w(\alpha+\Delta)-w(\alpha)}{\Delta}$	-0.1494	0	-inf	0.3329
$\frac{w(\beta)-w(\beta-\Delta)}{\Delta}$	0.2179	0	-1.5047	0.0019
$\frac{w(\beta+\Delta)-w(\beta)}{\Delta}$	0.2099	0	-1.4782	0.0060

Table 4.6: Weight sensitivities for selected values of OGARCH parameters in period 180. All the numbers are in percentage.

## 4.5 Conclusion

In this chapter, we have compared the out-of-sample performances of the optimal portfolios selected by OGARCH, two-state Regime-Switching models (MSM and GMM) and the EWMA Model by using four fund indices during the period 2003-2014. The mean and variance of each fund index are estimated by the four models and then Mean-Variance portfolio optimization model is applied to find the optimal weights. We have found that the optimal portfolio selected by OGARCH model outperforms EWM, MSM and GMM models with respect to the risk and returns during the whole investment period. The portfolio obtained by using EWMA model also has better return and smaller risk than the two Regime-Switching models.

We have also examined the weight sensitivities of the optimal portfolio with respect to the estimated parameters in OGARCH model. The numerical results show that the weight is very sensitive to the parameters in financial crisis period. This would indicate that the estimation errors are also important when the mean and variance are estimated in Mean-Variance portfolio optimization.

# Chapter 5

## Portfolio Optimization Under OGARCH with GH Skewed $t$ Distribution

### 5.1 Introduction

The Mean-Variance portfolio optimization model proposed by Markowitz (1952) has become the foundation of modern portfolio theory. It assumes that asset returns follow the multivariate Gaussian distribution with constant parameters. It takes standard deviation as a risk measure, which treats both upside and downside payoffs symmetrically. Chapter 2 shows that the distribution of crude oil is skewed and asymmetric with fat-tail and the volatility of the asset is time-varying. Chapter 3 provides further evidence that financial time series have asymmetric returns and the correlation between the time series is very volatile, especially in the financial crisis period.

In order to model skewness in conditional distributions of financial time series, Hansen (1994) propose the first skewed extension to student  $t$ . Aas (2006) show that the generalized hyperbolic (GH) skew student  $t$ -distribution fit the financial returns better than the other distributions.

Value-at-Risk (VaR) has been widely used as a market risk measure by practitioners. Under the Basel II Accord, banks are required to use VaR in calculating minimum capital requirements for market risk. Motivated by these facts, the addition of VaR constraint to mean-variance model has been studied by many researchers. Alexander and Baptista (2004) analyze the impact of adding VaR constraint to a single period mean-variance model with a VaR constraint by assuming that asset returns have a multinormal distribu-

tion. Alexander et al. (2007) then examine the impact of VaR and CVaR constraints in the Mean-Variance portfolio optimization model when the asset returns have a discrete distribution with finitely many jump points.

Bauwens (2005) combine a new class of multivariate skew-Student distributions with a multivariate generalized autoregressive conditional heteroscedasticity (GARCH) model. The empirical result shows that multivariate skew-Student improves the performances of VaR forecasts for several portfolios. Hu and Kercheval (2010) also show that the skewed  $t$  distribution fits the equity return better and exam the efficient frontiers computed from the mean-VaR and mean-CVaR portfolio optimization models. Liu (2012) introduces a modified version of the mean-CVaR optimization model by filtering stock returns with CCC-GARCH and skewed  $t$  distributions. Our work differs from these papers in that we filter the fund returns by OGARCH model.

The safety-first model is first introduced by Roy (1952) and the model is extended by Telser (1955) and Kataoka (1963). In recent years, many safety-first models have been developed and discussed due to the growing practical relevance of downside risk. Chiu and Li (2012) develop a modified safety-first model and have studied its application in financial risk management of disastrous events, Norkin and Boyko (2012) have improved the safety-first model by introducing one-sided threshold risk measures. The Kataoka's safety-first model with constraint of mean return is studied by Ding and Zhang (2009).

The first contribution of this chapter is that we filter the data by the orthogonal GARCH model with the combination of multivariate GH skew  $t$  distributions. The second contribution is that we derive an analytic solution to the enhanced safety-first model with standard deviation constraint. We also give the analytic solution to the Mean-Variance optimization problem in terms of parameters under GH skewed  $t$  and OGARCH model. Furthermore, we compare the optimal portfolio performances under Mean-CVaR, enhanced safety-first and classic Mean-Variance model.

The remainder of this chapter is organized as follows. In Section 2 we introduce the Orthogonal Garch model (O-GARCH) and Generalized Hyperbolic skewed  $t$  distribution (GH skewed  $t$ ). Then, we present the Mean-Variance portfolio optimization model, enhanced safety-first model and Mean-CVaR model in Section 3. Section 4 describes the parameter estimation method, the data used in this chapter and summary statistics of the estimated parameters. In Section 5, we compare the optimal portfolio performances under Mean-CVaR, enhanced safety-first and classic Mean-Variance model. Section 6 concludes the chapter.

## 5.2 Model

Let  $Y$  be a  $T \times k$  matrix of daily returns and  $Y$  is obtained by taking the natural log of the quotient of consecutive the daily closing prices; in other words,  $y_{i,t+1} = \log \frac{P_{i,t+1}}{P_{i,t}}$  where  $P_i$  is daily closing price of asset  $i$  at time  $t$ .

Assume that

$$Y_t = \boldsymbol{\mu} + H_t^{\frac{1}{2}} z_t \quad (5.1)$$

where  $Y_t, \boldsymbol{\mu} \in \mathbb{R}^k$  for  $t = 1, \dots, T, i = 1, \dots, k$  and  $H_t$  is the conditional covariance matrix and we assume  $z_t$  follows a generalized hyperbolic (GH) skewed t distribution.

### 5.2.1 Generalized hyperbolic distributions

The generalized hyperbolic (GH) distribution is first introduced by Barndorff-Nielsen (1977) and the skewed t distribution is a special case of GH distribution with fat tails and asymmetry characteristics. In this section, the GH distributions can be defined by using mean-variance mixture method as presented in McNeil et al. (2010).

**Definition 1. Normal Mean-Variance Mixture Distribution** *The random vector  $z_t$  has a  $k$  dimensional multivariate normal mean-variance mixture if*

$$z_t \stackrel{k}{=} \mu_z + \gamma W + \sqrt{W} A Z_t \quad (5.2)$$

Where  $Z \sim N_k(0, I_d)$  is standard  $d$ -dimension multivariate normal distribution.  $W \geq 0$  is a nonnegative, scalar-valued random variable and independent of  $Z$ .  $\mu_z \in \mathbb{R}^k, \gamma \in \mathbb{R}^k$  and  $A \in \mathbb{R}^{k \times d}$ .

It follows from equation (5.2) that

$$z_t | W \sim N(\mu_z + \gamma W, W \Sigma)$$

where  $\Sigma = A A'$  and it is easy to calculate

$$E[z_t] = E[E[z_t | W]] = \mu_z + \gamma E[W] \quad (5.3)$$

$$Cov[z_t] = E[Cov[z_t | W]] + Cov[E[z_t | W]] = E[W] \Sigma + Var[W] \gamma \gamma' \quad (5.4)$$

**Definition 2. Modified Bessel Function** *The modified Bessel function of the second kind with index  $\lambda$  is defined as given by Barndorff-Nielsen and Blaesild (1981):*

$$K_\lambda(x) = \frac{1}{2} \int_0^\infty y^{\lambda-1} e^{-\frac{x}{2}(y+y^{-1})} dy, \quad x > 0$$

If  $\lambda > 0$ , the following asymptotic property of  $K_\lambda$  is satisfied

$$K_\lambda(x) \sim \Gamma(\lambda)2^{\lambda-1}x^{-\lambda} \quad \text{as } x \rightarrow 0^+$$

Where  $\Gamma(x)$  is the gamma function.

The modified Bessel function also has the following important properties

$$K_\lambda(x) = K_{-\lambda}(x)$$

The GH skewed t distribution can be represented as a mean-variance mixture where the mixing variable  $W$  follows the generalized inverse Gaussian (GIG) distribution.

**Definition 3. The Generalized Inverse Gaussian Distribution (GIG).** A random variable  $X$  has a generalized inverse Gaussian (GIG) distribution,  $X \sim GIG(\lambda, \chi, \psi)$  if its density is given by

$$f(x) = \frac{(\sqrt{\chi/\psi})^\lambda}{2K_\lambda(\sqrt{\chi\psi})} x^{\lambda-1} e^{-\frac{1}{2}(\chi x^{-1} + \psi x)} \quad (5.5)$$

$K_\lambda$  is the modified Bessel function of the third kind with index  $\lambda$  and the domain of the parameters  $(\lambda, \chi, \psi)$  is given by

$$\begin{aligned} \chi > 0, \psi &\geq 0 & \text{if } \lambda < 0 \\ \chi > 0, \psi &> 0 & \text{if } \lambda = 0 \\ \chi \geq 0, \psi &> 0 & \text{if } \lambda > 0 \end{aligned}$$

**Definition 4. The Generalized Hyperbolic Distribution (GH)** A random vector  $X$  follows the generalized hyperbolic distribution (GH),  $X \sim GH_k(\lambda, \chi, \psi, \mu_z, \Sigma, \gamma)$  if  $X$  can be represented as mean-variance mixture where the mixing variable  $W$  follows the generalized inverse Gaussian (GIG) distribution,  $W \sim GIG(\lambda, \chi, \psi)$  and its density is given by

$$f(x) = c \frac{K_{\lambda-\frac{k}{2}}(\sqrt{(\chi + (x - \mu_z)^T \Sigma^{-1} (x - \mu_z))(\psi + \gamma^T \Sigma^{-1} \gamma)}) e^{(x - \mu_z)^T \Sigma^{-1} \gamma}}{(\sqrt{(\chi + (x - \mu_z)^T \Sigma^{-1} (x - \mu_z))(\psi + \gamma^T \Sigma^{-1} \gamma)})^{\frac{k}{2} - \lambda}}$$

where the constant  $c$  is given by

$$c = \frac{(\sqrt{\chi\psi})^{-\lambda} \psi^\lambda (\psi + \gamma^T \Sigma^{-1} \gamma)^{\frac{k}{2} - \lambda}}{(2\pi)^{\frac{k}{2}} |\Sigma|^{\frac{1}{2}} K_\lambda(\sqrt{\chi\psi})}$$

where  $|\Sigma|$  is the determinant of  $\Sigma$ .

The characteristic function of the GH random variable  $X \sim GH_k(\lambda, \chi, \psi, \mu_z, \Sigma, \gamma)$  is obtained from the mean-variance mixture definition

$$\begin{aligned}\phi_X(t) &= E[e^{it'X}] \\ &= E[E[e^{it'X} | W]] \\ &= E[e^{it'\mu_z + it'\gamma W - \frac{1}{2}Wt'\Sigma t}] \\ &= e^{it'\mu_z} e^{(it'\gamma - \frac{1}{2}t'\Sigma t)W}\end{aligned}$$

Suppose the density function of  $W$  is  $H$  and the Laplace-Stieltjes transform of  $H$  is defined by

$$\hat{H}(\theta) = E[e^{-\theta W}] = \int_0^\infty e^{-\theta u} dH(u)$$

Then

$$\phi_X(t) = e^{it'\mu_z} \hat{H}\left(\frac{1}{2}t'\Sigma t - it'\gamma\right)$$

An important property of the GH class is that it is closed under linear operations.

**Proposition 1.** *If  $X \sim GH_k(\lambda, \chi, \psi, \mu_z, \Sigma, \gamma)$ , and  $Y = BX + b$  where  $B \in \mathbb{R}^{dxk}$  and  $b \in \mathbb{R}^d$ , then  $Y \sim GH_d(\lambda, \chi, \psi, B\mu_z + b, B\Sigma B', B\gamma)$ .*

*Proof.*

$$\begin{aligned}\phi_Y(t) &= E[e^{it'(BX+b)}] \\ &= e^{it'b} E[e^{it'BX}] \\ &= e^{it'b} E[e^{it' B\mu_z + it' B\gamma W - \frac{1}{2}Wt' B\Sigma B' t}] \\ &= e^{it'b} e^{it' B\mu_z} e^{(it' B\gamma - \frac{1}{2}t' B\Sigma B' t)W} \\ &= e^{it'(B\mu_z + b)} e^{(it' B\gamma - \frac{1}{2}t' B\Sigma B' t)W} \\ &= e^{it'(B\mu_z + b)} \hat{H}\left(-it' B\gamma + \frac{1}{2}t' B\Sigma B' t\right)\end{aligned}$$

□

More specially, if  $B = w' = (w_1, w_2, \dots, w_k)'$  and  $b = 0$ , then  $y = (w_1, w_2, \dots, w_k)'X$  is

$$y \sim GH_1(\lambda, \chi, \psi, w'\mu_z, w'\Sigma w, w'\gamma)$$

The generalized hyperbolic skewed t distribution is a special case of generalized hyperbolic distribution. Suppose  $W \sim GIG(\lambda, \chi, \psi)$ . If  $\psi = 0$ ,  $\chi = \nu$ ,  $\lambda = -\frac{1}{2}\nu$  for  $\nu > 2$ , then  $W$  follows an inverse gamma distribution (McNeil et al., 2015)

**Definition 5. Gamma distribution** The random variable  $X$  has a gamma distribution,  $X \sim \text{Gamma}(\alpha, \beta)$ , if its density is

$$f(x) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x} \quad x > 0, \alpha > 0, \beta > 0$$

Moreover,  $\frac{1}{X}$  has an inverse gamma distribution,  $\frac{1}{X} \sim \text{Igamma}(\alpha, \frac{1}{\beta})$  with the density

$$f(x) = \frac{\frac{1}{\beta}^\alpha}{\Gamma(\alpha)} x^{-(\alpha+1)} e^{-\frac{1}{\beta x}} \quad x > 0, \alpha > 0, \beta > 0 \quad (5.6)$$

The moments of  $X \sim \text{Gamma}(\alpha, \beta)$  given  $\alpha > n$  is

$$E[x^n] = \frac{\beta^n}{(\alpha-1) \cdots (\alpha-n)} \quad (5.7)$$

**Definition 6. GH Skewed  $t$  Distribution** The random variable  $X$  has a generalized hyperbolic skewed  $t$  distribution,  $X \text{ Skewt} \sim (\nu, \mu_z, \Sigma, \gamma)$  if  $W \sim \text{Igamma}(\frac{\nu}{2}, \frac{\nu}{2})$  and  $\nu > 2$ . The multivariate joint density is

$$f(x) = c \frac{K_{\frac{\nu+k}{2}}(\sqrt{(\nu + (x - \mu_z)^T \Sigma^{-1} (x - \mu_z)) \gamma^T \Sigma^{-1} \gamma}) e^{(x - \mu_z)^T \Sigma^{-1} \gamma}}{(\sqrt{(\nu + (x - \mu_z)^T \Sigma^{-1} (x - \mu_z)) \gamma^T \Sigma^{-1} \gamma})^{-\frac{\nu+k}{2}} + (1 + \frac{(x - \mu_z)^T \Sigma^{-1} (x - \mu_z)}{\nu})^{\frac{\nu+k}{2}}}$$

where the constant  $c$  is given by

$$c = \frac{2^{1-\frac{\nu+k}{2}}}{\Gamma(\frac{\nu}{2}) (\pi + \nu)^{\frac{k}{2}} |\Sigma|^{\frac{1}{2}}}$$

where  $|\Sigma|$  is the determinant of  $\Sigma$ .

**Proposition 2.** The mean and covariance of a GH skewed  $t$  distributed random variable  $z_t \text{ Skewt} \sim (\nu, \mu_z, \Sigma, \gamma)$ , is  $E[z_t] = \mu_z + \gamma \frac{\nu}{\nu-2}$  and  $\text{Cov}[z_t] = \frac{\nu}{\nu-2} \Sigma + \gamma \gamma' \frac{2\nu^2}{(\nu-2)^2(\nu-4)}$

*Proof.* By applying formulas (5.3) and (5.7), we obtain

$$\begin{aligned} E[z_t] &= \mu_z + \gamma \frac{\frac{\nu}{2}}{\frac{\nu}{2} - 1} \\ &= \mu_z + \gamma \frac{\nu}{\nu - 2} \\ \text{Cov}[z_t] &= \frac{\nu}{\nu - 2} \Sigma + \gamma \gamma' \frac{2\nu^2}{(\nu - 2)^2(\nu - 4)} \end{aligned}$$

□

## 5.2.2 OGARCH

Suppose  $H_t$  can be estimated by orthogonal GARCH model. Then let  $Y$  be a multivariate time series of daily returns with mean zero on  $k$  assets with length  $T$  with columns  $y_1, \dots, y_k$ . We define the  $T \times k$  matrix  $X_t$  whose columns  $x_1, \dots, x_k$  are given by the equation

$$x_t = \frac{y_t - \mu_i}{\sqrt{v_i}}$$

where  $V = \text{diag}(v_1, \dots, v_m)$  with  $v_i$  being the sample variance of  $y_t$  and  $\mu_i$  is the sample mean of  $y_t$ .

Let  $L$  denote the matrix of eigenvectors of the population correlation of  $x_t$  and by  $l_m = (l_{1,m}, \dots, l_{k,m})$  its  $m$ th column.  $l_m$  is the  $k \times 1$  eigenvector corresponding to the eigenvalue  $\lambda_m$ . The column labelling of  $L$  has been chosen so that  $\lambda_1 > \lambda_2 > \dots > \lambda_k$ . Let  $D$  be the diagonal matrix of eigenvalues. The  $m$ th principal component of the system is defined by

$$p_m = x_1 l_{1,m} + x_2 l_{2,m} + \dots + x_k l_{k,m}$$

If each vector of principal components  $p_m$  is placed as the columns of a  $T \times k$  matrix  $P$ , then,

$$P = XL$$

The principal component columns is modeled by GARCH(1,1)

$$p_t | \Psi_{t-1} \sim N(0, \Sigma_t)$$

$$p_{i,t} = \epsilon_{i,t}$$

$$\sigma_{i,t}^2 = \omega_i + \alpha_i \epsilon_{i,t-1}^2 + \beta_i \sigma_{i,t-1}^2$$

where  $\Sigma_t$  is a diagonal matrix of the conditional variances of the principal components  $P$ .  $\Psi_{t-1}$  contains all the information available up to time  $t-1$ . The conditional covariance Matrix of  $X_n$  is  $D_t = L \Sigma_t L^T$  and the conditional covariance matrix of  $Y$  is given by

$$H_t = E[Y_t | \Psi_{t-1}] = \sqrt{V} D_t \sqrt{V} = \sqrt{V} L \Sigma_t L_n^T \sqrt{V}$$

## 5.2.3 Portfolio return distribution

Assume that the portfolio weight  $w = [w_1, w_2, \dots, w_k]^T$ , then the portfolio return at time  $t$  is  $R_t = w_1 y_{1,t} + w_2 y_{2,t} + \dots + w_k y_{k,t} = w^T Y_t = w^T (\boldsymbol{\mu} + H_t^{\frac{1}{2}} z_t)$ .

**Proposition 3. Conditional Distribution of the portfolio** *The conditional distribution of the portfolio return  $R_t$  is one dimensional skewed  $t$  distribution, denoted by  $R_t | \Psi_{t-1} \sim$*



$Skewt(\nu, w^T(H_t\mu_z + \boldsymbol{\mu}), w^T H_t^T \Sigma H_t w, w^T H_t^T \gamma)$ , if the assets  $z_t$  follow a multivariate skewed t distribution,  $z_t \sim Skewt(\nu, \mu_z, \Sigma, \gamma)$ .

*Proof.* Substitute  $B = H_t^{\frac{1}{2}}$  and  $b = \boldsymbol{\mu}$  in Proposition 1, we get the conditional distribution of the asset returns

$$Y_t | \Psi_{t-1} \sim Skewt(\nu, H_t^{\frac{1}{2}} \mu_z + \boldsymbol{\mu}, H_t^{\frac{1}{2}T} \Sigma H_t^{\frac{1}{2}}, H_t^{\frac{1}{2}T} \gamma)$$

Then we apply Proposition 1 again by letting  $B = w^T$ ,  $b = 0$  and get the conditional return of the portfolio follows a one-dimensional skewed t distribution

$$R_t | \Psi_{t-1} \sim Skewt(\nu, w^T(H_t^{\frac{1}{2}} \mu_z + \boldsymbol{\mu}), w^T H_t^{\frac{1}{2}T} \Sigma H_t^{\frac{1}{2}} w, w^T H_t^{\frac{1}{2}T} \gamma)$$

□

**Proposition 4.** *The conditional mean of the return  $Y_t$  is  $\mu_{y,t} = (H_t^{\frac{1}{2}} \mu_z + \boldsymbol{\mu}) + H_t^{\frac{1}{2}T} \gamma \frac{\nu}{\nu-2}$  and the conditional variance is  $\sigma_{y,t}^2 = H_t^{\frac{1}{2}T} \Sigma H_t^{\frac{1}{2}} \frac{\nu}{\nu-2} + H_t^{\frac{1}{2}T} \gamma \gamma^T H_t^{\frac{1}{2}} \frac{2\nu^2}{(\nu-2)^2(\nu-4)}$*

*Proof.* According to Aas (2006), the conditional mean and standard deviation of the asset returns at time t is

$$\mu_{y,t} = E[Y_t | \Psi_{t-1}] = (H_t^{\frac{1}{2}} \mu_z + \boldsymbol{\mu}) + H_t^{\frac{1}{2}T} \gamma \frac{\nu}{\nu-2} \quad (5.8)$$

According to equation 5.3, the covariance of the asset returns at time t is

$$\begin{aligned} \sigma_{y,t}^2 &= Var[Y_t | \Psi_{t-1}] \\ &= H_t^{\frac{1}{2}T} \Sigma H_t^{\frac{1}{2}} \frac{\nu}{\nu-2} + H_t^{\frac{1}{2}T} \gamma \gamma^T H_t^{\frac{1}{2}} \frac{2\nu^2}{(\nu-2)^2(\nu-4)} \end{aligned} \quad (5.9)$$

□

### 5.3 Portfolio Optimization under GH Skewed t distribution

In this section, we find the optimal portfolio weight by assuming the conditional portfolio return follows an univariate GH skewed t distribution.

Suppose there are n risky assets and an investor can invest in all the assets. Assume the conditional return of n risky asset at time t  $Y_t = (y_{1,t}, y_{2,t}, \dots, y_{n,t})$  follows a n dimension skewed t distribution. Now the investor tries to find the optimal weight at time t  $w_t = [w_{t,1}, w_{t,2}, \dots, w_{t,n}]^T$ . Then the portfolio return at time t+1 is  $R_{t+1} = w_t^T Y_{t+1}$ .

By Proposition 3, the conditional portfolio return is skewed t distributed. Given that  $R_{t+1} = w_t^T Y_{t+1}$ , by Proposition 4 we have

$$\mu_{p,t+1} = E[R_{t+1}|\Psi_t] = w_t^T \left[ (H_{t+1}^{\frac{1}{2}} \mu_z + \boldsymbol{\mu}) + H_{t+1}^{\frac{1}{2}T} \gamma \frac{\nu}{\nu - 2} \right] \quad (5.10)$$

and

$$\sigma_{p,t+1}^2 = Var[R_{t+1}|\Psi_t] = w_t^T \left[ H_t^{\frac{1}{2}T} \Sigma H_t^{\frac{1}{2}} \frac{\nu}{\nu - 2} + H_t^{\frac{1}{2}T} \gamma \gamma^T H_t^{\frac{1}{2}} \frac{2\nu^2}{(\nu - 2)^2(\nu - 4)} \right] w \quad (5.11)$$

### 5.3.1 VaR

VaR has played an important role in risk management. Instead of considering both upside and downside payoffs, VaR measures only the downside risk of the expected return.

**Definition 7. Value at Risk (VaR)** *Given a confidence level  $p \in (0, 1)$ , the VaR of a portfolio at a confidence level  $p$  over the time period  $t$  is given by the smallest number  $l_0$  such that the probability of a loss  $L_s$  over a time interval  $t$  greater than  $l_0$  is  $1-p$ .*

That is,

$$\begin{aligned} VaR(L) &= \inf\{l_0 : P(L_s > l_0) \leq 1 - p\} \\ &= \inf\{F_L(l_0) \geq p\} \\ &= F_L^{-1}(p) \end{aligned}$$

Where  $F(\cdot)$  is the cumulative function of  $L_s$  and  $F^{-1}$  is the inverse function of  $F_L$ .

Most applications of VaR are used to control the risk over short horizons, so it's useful to introduce the Value at Risk that employs information up to time  $t-1$  to produce a VaR for time period  $t$ . Let  $(VaR_t)$  denote the one-day-ahead forecast of VaR which is defined as

$$P(R_t \leq -VaR_{t-1}|\Psi_{t-1}) = p \quad (5.12)$$

Where  $R_t$  is the portfolio return at time  $t$  and  $\Psi_{t-1}$  is the available information at  $t-1$ .

**Proposition 5.** *It follows from direct computations that  $VaR_{t-1} = -\mu_{p,t} + k_p \sigma_{p,t}$  if the returns follow normal or student t distributions.*

Where  $k_p$  is the  $p$  % left tail of the normal or student t distribution.

Unfortunately, for GH skewed t distribution,  $k_p$  is complicated and generally has no closed-form solution.

Even though VaR is very popular, it has some drawbacks and undesirable properties such as lack of sub-additivity and non-convexity. As mentioned in Artzner et al. (2002), it is not a coherent measure of risk.

### 5.3.2 CVaR

An alternative risk measure called conditional Value-at-Risk(CVaR) is proposed by Rockafellar and Uryasev (2000). CVaR is defined as follows:

**Definition 8. *Contional Value at Risk (CVaR)*** Given a confidence level  $p \in (0, 1)$  and a random variable loss  $L_s$

$$CVaR_p(L_s) = E[L_s | L_s \geq VaR_p] = \frac{1}{1-p} \int_p^1 VaR_x dx$$

If the cumulative function of  $L_s$   $F_L$  is continuous, then we have

$$\begin{aligned} CVaR_p(L) &= \frac{1}{1-p} \int_p^1 VaR_x dx \\ &= \frac{1}{1-p} \int_p^1 F_L^{-1}(x) dx \\ &= \frac{1}{1-p} \int_{F_L^{-1}(p)}^{\infty} F_L(x) dx \\ &= \frac{1}{1-p} \int_{VaR_p}^{\infty} x dF_L(x) \end{aligned}$$

Artzner et al. (2002) show that CVaR is a coherent risk measure and it has the following properties

**Proposition 6.** Consider the space of random variables representing portfolio losses over some time interval  $\mathcal{G}$  and  $L_1, L_2 \in \mathcal{G}$ . CVaR satisfies the following properties:

1. *Subadditivity:*  $CVaR(L_1 + L_2) \leq CVaR(L_1) + CVaR(L_2)$ .
2. *Positive Homogeneity:* For any  $\lambda > 0$ ,  $CVaR(\lambda L) = \lambda CVaR(L)$
3. *Translation Invariance:* For any constant  $a \in \mathbb{R}$ ,  $CVaR(a + L_s) = a + CVaR(L_s)$
4. *Monotonicity:* If  $L_1 \leq L_2$ , then  $CVaR(L_1) \leq CVaR(L_2)$
5. *Convexity:* For any  $\lambda \in [0, 1]$ ,  $CVaR(\lambda L_1 + (1 - \lambda)L_2) \leq \lambda CVaR(L_1) + (1 - \lambda) CVaR(L_2)$

### 5.3.3 Mean-Variance Portfolio Optimization

Suppose that the investor wants to maximize the utility function, the classical mean-variance portfolio optimization (**MV1**) is given by

$$\text{Maximize } w_t^T E[Y_{t+1}|\Psi_t] - \frac{\lambda}{2} w_t^T \text{Var}(Y_{t+1}|\Psi_t) w_t \quad (5.13)$$

subject to

$$w_t^T \mathbf{1} = 1$$

Define  $A = \mathbf{1}^T \bar{\Sigma}^{-1} \mathbf{1} > 0$ ,  $B = \mathbf{1}^T \bar{\Sigma}^{-1} \bar{\mu} > 0$  and  $C = \bar{\mu}^T \bar{\Sigma}^{-1} \bar{\mu} > 0$

**Theorem 1.** *The classical mean-variance portfolio optimization (**MV1**) has a unique global optimal solution*

$$w_{*t} = \frac{1}{\lambda_t \bar{\Sigma}} \left( \bar{\mu} - \frac{(\bar{\mu} \bar{\Sigma}^{-1} \mathbf{1}^T - \lambda_t) \mathbf{1}}{\mathbf{1}^T \bar{\Sigma}^{-1} \mathbf{1}} \right) = \frac{1}{\lambda_t \bar{\Sigma}} \left( \bar{\mu} - \frac{(B - \lambda_t) \mathbf{1}}{A} \right)$$

where  $\lambda \geq 0$  is a risk aversion coefficient and  $w_t$  denotes the portfolio weight vector.

*Proof.* The Lagrangian of the optimization problem **MV1** subject to variance constraints is expressed as follows:

$$\begin{aligned} L(w_t|\Psi_t) &= w_t^T E[Y_{t+1}|\Psi_t] - \frac{\lambda}{2} w_t^T \text{Var}(Y_{t+1}|\Psi_t) w_t - \eta_t (w_t^T \mathbf{1} - 1) \\ &= w_t^T \left[ (H_{t+1}^{\frac{1}{2}} \mu_z + \boldsymbol{\mu}) + H_{t+1}^{\frac{1}{2}T} \gamma \frac{\nu}{\nu - 2} \right] \\ &\quad - \frac{\lambda}{2} w_t^T \left[ H_{t+1}^{\frac{1}{2}T} \Sigma H_{t+1}^{\frac{1}{2}} \frac{\nu}{\nu - 2} + H_{t+1}^{\frac{1}{2}T} \gamma \gamma^T H_{t+1}^{\frac{1}{2}} \frac{2\nu^2}{(\nu - 2)^2(\nu - 4)} \right] w_t \\ &\quad - \eta_t (w_t^T \mathbf{1} - 1) \end{aligned}$$

where  $\eta_t$  is the Lagrange multipliers for the constraint.

Using this Lagrange function, the KKT (Karush-Kuhn-Tucker) optimality conditions (Floudas and Visweswaran, 1995) are:

$$\begin{aligned} \frac{\partial L(w_t|\Psi_t)}{\partial w_t} &= (H_{t+1}^{\frac{1}{2}} \mu_z + \boldsymbol{\mu}) + H_{t+1}^{\frac{1}{2}T} \gamma \frac{\nu}{\nu - 2} \\ &\quad - \lambda \left[ H_t^{\frac{1}{2}T} \Sigma H_t^{\frac{1}{2}} \frac{\nu}{\nu - 2} + H_t^{\frac{1}{2}T} \gamma \gamma^T H_t^{\frac{1}{2}} \frac{2\nu^2}{(\nu - 2)^2(\nu - 4)} \right] w - \eta_t \mathbf{1} \\ &= 0 \end{aligned} \quad (5.14)$$

$$\frac{\partial L(w_t|\Psi_t)}{\partial \eta_t} = w_t^T \mathbf{1} - 1 = 0 \quad (5.15)$$

The conditions (5.14) and (5.15) have the matrix representation:

$$\begin{bmatrix} \lambda \left[ H_t^{\frac{1}{2}T} \Sigma H_t^{\frac{1}{2}} \frac{\nu}{\nu-2} + H_t^{\frac{1}{2}T} \gamma \gamma^T H_t^{\frac{1}{2}} \frac{2\nu^2}{(\nu-2)^2(\nu-4)} \right] & \mathbf{1} \\ \mathbf{1}^T & 0 \end{bmatrix} \begin{bmatrix} w_t \\ \eta_t \end{bmatrix} = \begin{bmatrix} (H_{t+1}^{\frac{1}{2}} \mu_z + \boldsymbol{\mu}) + H_{t+1}^{\frac{1}{2}T} \gamma \frac{\nu}{\nu-2} \\ 1 \end{bmatrix}$$

Then the optimal solution is given by

$$\eta_t = \frac{\bar{\mu} \Sigma^{-1} \mathbf{1}^T - \lambda_t}{\mathbf{1}^T \bar{\Sigma}^{-1} \mathbf{1}}$$

and

$$w_{*t} = \frac{1}{\lambda_t \bar{\Sigma}} \left( \bar{\mu} - \frac{(B - \lambda_t) \mathbf{1}}{A} \right) \quad (5.16)$$

□

Where  $\bar{\Sigma} = H_t^{\frac{1}{2}T} \Sigma H_t^{\frac{1}{2}} \frac{\nu}{\nu-2} + H_t^{\frac{1}{2}T} \gamma \gamma^T H_t^{\frac{1}{2}} \frac{2\nu^2}{(\nu-2)^2(\nu-4)}$  and  $\bar{\mu} = (H_{t+1}^{\frac{1}{2}} \mu_z + \boldsymbol{\mu}) + H_{t+1}^{\frac{1}{2}T} \gamma \frac{\nu}{\nu-2}$

### 5.3.4 Safety-first Portfolio Optimization

The safety-first model aims to maximize a insured lower level of portfolio return at  $t$   $R_p$  for a given the tolerance level of investor to risk is  $p$ . A safety-first portfolio optimization problem is described by

$$\text{maximize } R_p(w) \quad (5.17)$$

subject to

$$w_t^T \mathbf{1} = 1$$

$$P(R_{t+1} < R_p | \Psi_t) \leq p$$

By Tchebycheff inequality, we have

$$\begin{aligned} P(P(R_{t+1} < R_p | \Psi_t)) &= P(R_{t+1} - E(R_{t+1}) < R_p - E(R_{t+1}) | \Psi_t) \\ &= P(E(R_{t+1}) - R_{t+1} > E(R_{t+1}) - R_p | \Psi_t) \\ &\leq \frac{\text{Var}(R_{t+1} | \Psi_t)}{(E(R_{t+1} | \Psi_t) - R_p)^2} \end{aligned}$$

Then the probabilistic constraint in **MV2** is replaced by a stronger constraint

$$\frac{\text{Var}(R_{t+1}|\Psi_t)}{(E(R_{t+1}|\Psi_t) - R_p)^2} \leq p$$

Therefore, the portfolio optimization problem (5.17) is simplified as

$$\text{maximize } R_p(w) \tag{5.18}$$

subject to

$$w_t^T \mathbf{1} = 1$$

$$\text{Var}(R_{t+1}|\Psi_t) \leq p(E(R_{t+1}|\Psi_t) - R_p)^2 \tag{5.19}$$

or equivalently

$$\text{maximize } w_t^T E[Y_{t+1}|\Psi_t] - \frac{1}{\sqrt{p}} \sqrt{w_t^T \text{Var}(Y_{t+1}|\Psi_t) w_t} \tag{5.20}$$

subject to

$$w_t^T \mathbf{1} = 1$$

The classic safety-first approach only expects a higher lower insured threshold and discourages the risk diversification sometimes. Therefore, a modified safety-first model(**MV2**) is proposed by adding a variance constraint :

$$\text{maximize } w_t^T E[Y_{t+1}|\Psi_t] - \frac{1}{\sqrt{p}} \sqrt{w_t^T \text{Var}(Y_{t+1}|\Psi_t) w_t} \tag{5.21}$$

subject to

$$w_t^T \mathbf{1} = 1$$

$$\sqrt{w_t^T \text{Var}(Y_{t+1}|\Psi_t) w_t} \leq c \tag{5.22}$$

**Proposition 7.** *The function  $f(w) = w_t^T E[Y_{t+1}|\Psi_t] - \frac{1}{\sqrt{p}} \sqrt{w_t^T \text{Var}(Y_{t+1}|\Psi_t) w_t}$  is a concave function for  $w \in R^n$ .*

*Proof.* Let  $w_1, w_2 \in R^n$ , for any  $\lambda \in [0, 1]$ , we have

$$\begin{aligned}
 & \lambda f(w_1) + (1 - \lambda)f(w_2) - f(\lambda w_1 + (1 - \lambda)w_2) \\
 = & \lambda w_1^T E[Y_{t+1}|\Psi_t] - \lambda \frac{1}{\sqrt{p}} \sqrt{w_1^T \text{Var}(Y_{t+1}|\Psi_t) w_1} + (1 - \lambda) w_2^T E[Y_{t+1}|\Psi_t] - \\
 & (1 - \lambda) \frac{1}{\sqrt{p}} \sqrt{w_2^T \text{Var}(Y_{t+1}|\Psi_t) w_2} - (\lambda w_1 + (1 - \lambda)w_2) E[Y_{t+1}|\Psi_t] \\
 & + \frac{1}{\sqrt{p}} \sqrt{(\lambda w_1 + (1 - \lambda)w_2)^T \text{Var}(Y_{t+1}|\Psi_t) (\lambda w_1 + (1 - \lambda)w_2)} \\
 = & \frac{1}{\sqrt{p}} \sqrt{(\lambda w_1 + (1 - \lambda)w_2)^T \text{Var}(Y_{t+1}|\Psi_t) (\lambda w_1 + (1 - \lambda)w_2)} - \\
 & \frac{1}{\sqrt{p}} \left( \lambda \sqrt{w_1^T \text{Var}(Y_{t+1}|\Psi_t) w_1} + (1 - \lambda) \sqrt{w_2^T \text{Var}(Y_{t+1}|\Psi_t) w_2} \right) \\
 & = \frac{1}{\sqrt{p}} (m - n)
 \end{aligned}$$

Where  $m = \sqrt{(\lambda w_1 + (1 - \lambda)w_2)^T \text{Var}(Y_{t+1}|\Psi_t) (\lambda w_1 + (1 - \lambda)w_2)}$   
 and  $n = \lambda \sqrt{w_1^T \text{Var}(Y_{t+1}|\Psi_t) w_1} + (1 - \lambda) \sqrt{w_2^T \text{Var}(Y_{t+1}|\Psi_t) w_2}$

Then  $f(w)$  is a concave function if  $m < n$ .  $\iff m^2 < n^2$ , so we only need to prove that

$$w_1^T \text{Var}(Y_{t+1}|\Psi_t) w_2 < \sqrt{w_1^T \text{Var}(Y_{t+1}|\Psi_t) w_1} \sqrt{w_2^T \text{Var}(Y_{t+1}|\Psi_t) w_2}$$

Note that

$$\begin{aligned}
 (w_1 + w_2)^T \text{Var}(Y_{t+1}|\Psi_t) (w_1 + w_2) &= w_1^T \text{Var}(Y_{t+1}|\Psi_t) w_1 \\
 &+ 2w_1^T \text{Var}(Y_{t+1}|\Psi_t) w_2 + w_2^T \text{Var}(Y_{t+1}|\Psi_t) w_2 \geq 0
 \end{aligned}$$

we have

$$\sqrt{w_1^T \text{Var}(Y_{t+1}|\Psi_t) w_1} \sqrt{w_2^T \text{Var}(Y_{t+1}|\Psi_t) w_2} \geq w_1^T \text{Var}(Y_{t+1}|\Psi_t) w_2$$

Therefore,

$$\lambda f(w_1) + (1 - \lambda)f(w_2) \leq f(\lambda w_1 + (1 - \lambda)w_2)$$

□

**Theorem 2.** *The modified safety-first portfolio optimization problem (MV2) has a global optimal solution.*

(1) When  $c > \frac{1}{\sqrt{A}}$  and  $0 < p < \min(1, \frac{A}{AC-B^2})$ , or when  $c \leq \frac{1}{\sqrt{A}}$  and  $\frac{A-\frac{1}{2}}{AC-B^2} < p < \min(1, \frac{A}{AC-B^2})$ . The modified safety-first model(MV2) has a unique global optimal solution

$w$ , and

$$w = \frac{1}{\sqrt{B^2 - A(C - \frac{1}{p})}} \bar{\Sigma}^{-1} \left( \bar{\mu} + \frac{-B + \sqrt{B^2 - A(C - \frac{1}{p})}}{A} \mathbf{1} \right)$$

(2) When  $c < \frac{1}{\sqrt{C}}$  and  $0 < p < \left( \frac{c\sqrt{C}-1}{c\sqrt{C}+1} \right)^2$  there exists a unique global optimal solution

$$w = A^{-1} \bar{\Sigma}^{-1} \mathbf{1} + \frac{c(\bar{\Sigma}^{-1} \bar{\mu} - BA^{-1} \bar{\Sigma}^{-1} \mathbf{1})}{\frac{1}{\sqrt{p}} + \frac{c\sqrt{C}+1}{c\sqrt{C}-1}}$$

(3) For any other values of  $c$  and  $p$ , the solution does not exist.

*Proof.* The Lagrangian function of the optimization problem (5.21) is

$$L(w_t | \Psi_t) = w_t^T \bar{\mu} - \frac{1}{\sqrt{p}} \sqrt{w_t^T \bar{\Sigma} w_t} + \eta (w_t^T \mathbf{1} - 1) + \lambda (\sqrt{w_t^T \bar{\Sigma} w_t} - c)$$

where  $\eta$  and  $\lambda$  are the Lagrange multipliers for the constraints.

According to Proposition 7,  $f(w)$  is a differentiable concave function. Therefore, the problem has a global optimal solution if and only if the following KKT conditions are satisfied

$$\frac{\partial L(w_t | \Psi_t)}{\partial w_t} = \bar{\mu} - \frac{1}{\sqrt{p}} (w_t^T \bar{\Sigma} w_t)^{-\frac{1}{2}} \bar{\Sigma} w_t + \eta \mathbf{1} + \lambda (w_t^T \bar{\Sigma} w_t)^{-\frac{1}{2}} \bar{\Sigma} w_t = 0 \quad (5.23)$$

$$\frac{\partial L(w_t | \Psi_t)}{\partial \eta} = w_t^T \mathbf{1} - 1 = 0 \quad (5.24)$$

$$\frac{\partial L(w_t | \Psi_t)}{\partial \lambda} = w_t^T \bar{\Sigma} w_t - c \leq 0 \quad (5.25)$$

$$\lambda (\sqrt{w_t^T \bar{\Sigma} w_t} - c) = 0 \quad (5.26)$$

$$\lambda \geq 0 \quad (5.27)$$

Rearrange equation (5.23), we have

$$w = \frac{(\bar{\Sigma}^{-1} \bar{\mu} + \eta \bar{\Sigma}^{-1} \mathbf{1}) \sqrt{w_t^T \bar{\Sigma} w_t}}{\frac{1}{\sqrt{p}} - \lambda} \quad (5.28)$$



and

$$\frac{1}{p} = \lambda^2 + 2\lambda \frac{(B + \eta A)}{\sqrt{C}} + \frac{(B + \eta A)^2}{C} \quad (5.29)$$

From equation (5.24) and (5.28), we get

$$\frac{1}{\sqrt{p}} = \lambda + (B + \eta A) \sqrt{w_t^T \bar{\Sigma} w_t} \quad (5.30)$$

and

$$\sqrt{w_t^T \bar{\Sigma} w_t} = \frac{\frac{1}{\sqrt{p}} - \lambda}{B + \eta A} \quad (5.31)$$

Substitute equation (5.30) into (5.28), we obtain

$$w = \frac{(\bar{\Sigma}^{-1} \bar{\mu} + \eta \bar{\Sigma}^{-1} \mathbf{1})}{B + \eta A} \quad (5.32)$$

(1) If the variance constraint (5.22) is not binding, then we have  $\lambda = 0$  and

$$\bar{\mu} - \frac{1}{\sqrt{p}} (w_t^T \bar{\Sigma} w_t)^{-\frac{1}{2}} \bar{\Sigma} w + \eta \mathbf{1} = 0 \quad (5.33)$$

From (5.24), we get

$$w_t = \frac{1}{\mathbf{1}^T}$$

Substitute  $w_t = \frac{1}{\mathbf{1}^T}$  and rearrange equation (5.28), we get the

$$\begin{aligned} \sqrt{p} &= -\mathbf{1}^T \bar{\Sigma}^{-1} (\bar{\mu} + \eta \mathbf{1}) \sqrt{w_t^T \bar{\Sigma} w_t} \\ &= -(B + \eta A) \sqrt{w_t^T \bar{\Sigma} w_t} \end{aligned} \quad (5.34)$$

$$A\eta^2 + 2B\eta + C + \frac{1}{p} = 0 \quad (5.35)$$

The discriminant of the quadratic equation (5.31)  $\Delta = B^2 - A(C - \frac{1}{p})$ .

If  $\Delta \leq 0$ , the solution of (5.31) does not exist.

If  $\Delta = 0$ ,  $B^2 = A(C - \frac{1}{p})$ , there is a unique solution  $\eta = -\frac{B}{A}$ . However, in this case,  $\sqrt{p} = 0$  which contradicts with our assumption that  $p > 0$ . Hence,  $\eta = \frac{B}{A}$  is rejected.

If  $\Delta > 0$ ,  $B^2 > A(C - \frac{1}{p})$ , there are two distinct solutions  $\eta_1 = \frac{-B+\Delta}{A}$  and  $\eta_2 = \frac{-B-\Delta}{A}$ .  $\eta_2$  is rejected because  $\sqrt{p}$  is negative if  $\eta_2$  is substituted into (5.29).

After replacing the expression for  $\eta$  by  $\eta_1 = \frac{-B+\Delta}{A}$  in (5.29), we obtain

$$\sqrt{w_t^T \bar{\Sigma} w_t} = \frac{1}{\Delta \sqrt{p}} \leq c \quad (5.36)$$

From equation (5.36), we have

$$p \geq \frac{A - \frac{1}{c^2}}{AC - B^2}$$

If  $c \leq \frac{1}{\sqrt{A}}$ , then the unique optimal solution of (5.28) is obtained

$$w = \frac{1}{\sqrt{B^2 - A(C - \frac{1}{p})}} \bar{\Sigma}^{-1} \left( \bar{\mu} + \frac{-B + \sqrt{B^2 - A(C - \frac{1}{p})}}{A} \mathbf{1} \right) \quad (5.37)$$

when  $\frac{A - \frac{1}{c^2}}{AC - B^2} < p < \min(1, \frac{A}{AC - B^2})$ .

If  $c > \frac{1}{\sqrt{A}}$ , solution (5.37) exists when  $0 < p < \min(1, \frac{A}{AC - B^2})$  and variance constraint (5.22) is binding, then we have  $\lambda > 0$ . We get  $\sqrt{w_t^T \bar{\Sigma} w_t} = c$  from equation (5.26) and  $\lambda < \frac{1}{\sqrt{p}}$ . Substituting  $\sqrt{w_t^T \bar{\Sigma} w_t} = c$  into equation (5.34), we have

$$\lambda = \frac{1}{\sqrt{p}} - (B + \eta A)c \quad (5.38)$$

Solve equation (5.29) and (5.38), we get

$$\frac{C}{p} = \left( \sqrt{C} \lambda + \frac{\frac{1}{\sqrt{p}} - \lambda}{c} \right)^2 \quad (5.39)$$

Solve equation (5.39), we have  $\lambda_1 = \frac{1}{\sqrt{p}}$  which contradicts with  $\lambda < \frac{1}{\sqrt{p}}$  and  $\lambda_2 = -\frac{c\sqrt{C}+1}{c\sqrt{C}-1}$  if  $c < \frac{1}{\sqrt{C}}$  and  $p < \left( \frac{c\sqrt{C}-1}{c\sqrt{C}+1} \right)^2$

In this case, the unique solution is

$$w = A^{-1} \bar{\Sigma}^{-1} \mathbf{1} + \frac{c(\bar{\Sigma}^{-1} \bar{\mu} - BA^{-1} \bar{\Sigma}^{-1} \mathbf{1})}{\frac{1}{\sqrt{p}} + \frac{c\sqrt{C}+1}{c\sqrt{C}-1}}$$

□

**Corollary 1.** *If  $\lambda_t = \sqrt{B^2 - A(C - \frac{1}{p})}$ , the modified safety-first portfolio optimization problem (MV2) has the same solution as the Mean-Variance problem (MV1) when  $c > \frac{1}{\sqrt{A}}$  and  $0 < p < \min(1, \frac{A}{AC-B^2})$ , or when  $c \leq \frac{1}{\sqrt{A}}$  and  $\frac{A-\frac{1}{c^2}}{AC-B^2} < p < \min(1, \frac{A}{AC-B^2})$ .*

### 5.3.5 Mean-CVaR Portfolio Optimization

For a given expected return of portfolio  $r$  with weight vector  $w$ , the Mean-CVaR portfolio optimization problem (MV3) is given

$$\text{minimize } CVaR_t(w) \quad (5.40)$$

subject to

$$\begin{aligned} w_t^T \mathbf{1} &= 1 \\ w_t^T E[Y_{t+1} | \Psi_t] &= r \end{aligned}$$

Since the calculation of CVaR as a function of portfolio weights  $w$  and the loss distribution FL is difficult, the following auxiliary function (Rockafellar and Uryasev, 2000) is considered

$$F_p(w, \xi) = \xi + \frac{1}{1-p} \int_{x \in \mathbb{R}} (w^T x - \xi)^+ f_L(x) dx \quad (5.41)$$

where  $f_L(x)$  is the probability density function of the loss distribution function  $F_L(x)$  and  $[a]^+ = \max(0, a)$ .  $w$  denotes the portfolio weight.  $F_p(w, \xi)$  is a convex function of  $\xi$  and continuously differentiable.

**Theorem 3.** *Minimizing the CVaR of the loss of portfolio with weight  $w$  is equivalent to minimizing  $F_p(w, \xi)$ , that is,*

$$\min CVaR(w) = \min F_p(w, \xi)$$

*Proof.* See Rockafellar and Uryasev (2000) □

The integral in (5.41) is very difficult to calculate, so we consider the approximations using Monte Carlo simulation. Assume that  $x_j$  are independent samples for  $j = 1, \dots, n$ .  $F_p(w, \xi)$  can be approximated by using the weighted sum of all the scenarios

$$F_p(\hat{w}, \xi) = \xi + \frac{1}{(1-p)n} \sum_{j=1}^n (w^T x_j - \xi)^+ \quad (5.42)$$

The optimization problem(MV3) becomes

$$\text{minimize } \xi + \frac{1}{(1-p)n} \sum_{j=1}^n (w^T x_j - \xi)^+ \quad (5.43)$$

subject to

$$w_t^T \mathbf{1} = 1$$

$$w_t^T \left[ (H_{t+1}^{\frac{1}{2}} \mu_z + \boldsymbol{\mu}) + H_{t+1}^{\frac{1}{2}T} \gamma \frac{\nu}{\nu-2} \right] = R$$

Let  $(w^T x_j - \xi)^+$  be replaced by auxiliary variables  $z_i$  with two linear constraints  $z_i \geq 0$  and  $z_i \geq w^T x_j - \xi$  for  $i = 1, \dots, n$ . Then mean-CVaR problem is equivalent to

$$\text{Maximize } w_t^T E[Y_{t+1} | \Psi_t] - \lambda \left( \xi + \frac{1}{(1-p)n} \sum_{j=1}^n z_j \right) \quad (5.44)$$

subject to

$$\begin{aligned} z_j &\geq 0 \quad j = 1, \dots, n \\ z_j &\geq w^T x_j - \xi \quad j = 1, \dots, n \\ \sum_{j=1}^n z_j &= 1 \end{aligned}$$

Therefore, the mean-CVaR problem becomes a linear programming problem with respect to portfolio weight  $w$ .

## 5.4 Parameter estimation for skewed t distribution

Given a return time series  $Y_t$  in 5.2, the parameters in OGARCH model are estimated by the procedure in appendix.  $\boldsymbol{\mu}$  is estimated by sample mean with  $\mu_i = \frac{1}{T} \sum_{t=1}^T Y_{i,t}$  for  $i = 1, 2, 3, 4$ . After the conditional covariance  $H_t$  in OGARCH model is determined, the filter data  $z_t$  is obtained by

$$z_t = (Y_t - \boldsymbol{\mu}) H_t^{-\frac{1}{2}}$$

Then the multivariate skewed t distribution is calibrated to  $z_t$  by applying EM Algorithm.

### EM algorithm

The EM algorithm developed by Dempster et al. (1977) is a very useful tool for the iterative computation of maximum likelihood estimates to deal with the missing or incomplete data problems. In each iteration of the EM algorithm, there are two steps called the expectation step (E-step) and the maximization step (M-step). We use similar estimation method as described in the paper by Aas (2006).

Let  $\theta = (\nu, \mu_z, \Sigma, \gamma)$  and the a log-likelihood function for  $\theta$  is given by

$$\log L(\theta, z) = \sum_{i=1}^T f_Z(z_i, \theta) \quad (5.45)$$

Where  $f_Z(\cdot)$  is GH skewed  $t$  distribution.

Note that  $W \sim Igamma(\frac{\nu}{2}, \frac{\nu}{2})$  and  $z_t|W \sim N(\mu_z + \gamma W, W\Sigma)$ . The joint distribution of  $Z$  and  $W$  is

$$f_{Z,W}(z, w, \theta) = f_{Z/W}(z/w, \mu_z, \Sigma, \gamma)g_W(w, \nu)$$

where  $g_W$  is the density of inverse gamma  $(\frac{\nu}{2}, \frac{\nu}{2})$ .  $f_{Z/W}$  is the density of conditional normal distribution  $N(\mu_z + \gamma W, W\Sigma)$ .

From (5.6),  $g_W$  can be expressed as

$$g_W(w) = \frac{\left(\frac{\nu}{2}\right)^{\frac{\nu}{2}}}{\Gamma\left(\frac{\nu}{2}\right)} w^{-(\frac{\nu}{2}-1)} e^{-\frac{\nu}{2w}} \quad (5.46)$$

$f_{Z/W}$  can be written as

$$f_{Z/W}(z/w) = \frac{e^{(z-\mu_z)^T \Sigma^{-1} \gamma - \frac{(z-\mu_z)^T \Sigma^{-1} (z-\mu_z) - \frac{w}{2} \gamma^T \Sigma^{-1} \gamma}{2w}}}{(2\pi)^{\frac{d}{2}} \sqrt{|\Sigma|} w^{\frac{d}{2}}} \quad (5.47)$$

The augmented log-likelihood is

$$\begin{aligned} \log \tilde{L}(\theta, z, W) &= \sum_{i=1}^T f_{Z,W}(z_i, w_i, \theta) \\ &= \sum_{i=1}^T \log f_{Z/W}(z_i/w_i, \mu_z, \Sigma, \gamma) \sum_{i=1}^T \log g_W(w_i, \nu) \\ &= L_1(\mu_z, \Sigma, \gamma, z/w) + L_2(\nu, w) \end{aligned}$$

$L_1$  and  $L_2$  could be maximized separately if the latent mixing variables are observable. From (5.46) and (5.4), we get

$$\begin{aligned}
 L_1(\mu_z, \Sigma, \gamma, z/w) &= \sum_{i=1}^T (z_i - \mu_z)^T \Sigma^{-1} \gamma - \sum_{i=1}^T \frac{(z_i - \mu_z)^T \Sigma^{-1} (z_i - \mu_z)}{2w_i} \\
 &- \sum_{i=1}^T \frac{w_i}{2} \gamma^T \Sigma^{-1} \gamma - \frac{d}{2} \sum_{i=1}^T \log(w_i) - \frac{T}{2} \log |\Sigma| \\
 &- T \frac{d}{2} \log 2\pi
 \end{aligned} \tag{5.48}$$

$$L_2(\nu, w) = T \frac{\nu}{2} \log \left( \frac{\nu}{2} \right) - \left( \frac{\nu}{2} - 1 \right) \sum_{i=1}^T \log w_i - \frac{\nu}{2} \sum_{i=1}^T \frac{1}{w_i} - T \log \Gamma \left( \frac{\nu}{2} \right) \tag{5.49}$$

### E Step

The E-step computes the conditional expectation of the sufficient statistics of the log-likelihood  $L_1$  and  $L_2$  given the current values of the parameters, which are  $W$ ,  $W^{-1}$  and  $\log W$ . This involves the calculation of the following conditional expectations.

$$e_i^{[k]} = E[W_i | z_i, \theta^{[k]}]$$

$$\zeta_i^{[k]} = E[W_i^{-1} | z_i, \theta^{[k]}]$$

$$\xi_i^{[k]} = E[\log W_i | z_i, \theta^{[k]}]$$

It can be shown that the conditional density function of  $W$  is GIG distribution. That is,

$$W|z \sim GIG\left(-\frac{d+\nu}{2}, (z - \mu_z)^T \Sigma^{-1} (z - \mu_z) + \nu, \gamma^T \Sigma^{-1} \gamma\right)$$

The moments of the GIG(a, b, c) distribution are given by Karlis (2002)

$$E(x^r) = \frac{b^r}{c} \frac{K_{a+c}(bc)}{K_a(bc)}$$

Define  $A = (z - \mu_z)^T \Sigma^{-1} (z - \mu_z) + \nu$  and  $B = \gamma^T \Sigma^{-1} \gamma$ , then we have

$$e_i^{[k]} = \left( \frac{A^{[k]}}{B^{[k]}} \right)^{\frac{1}{2}} \frac{K_{\nu+d-2} \sqrt{A^{[k]} + B^{[k]}}}{K_{\nu+d} \sqrt{A^{[k]} + B^{[k]}}} \tag{5.50}$$

$$\zeta_i^{[k]} = \left( \frac{A^{[k]}}{B^{[k]}} \right)^{-\frac{1}{2}} \frac{K_{\frac{\nu+d+2}{2}} \sqrt{A^{[K]} + B^{[K]}}}{K_{\frac{\nu+d}{2}} \sqrt{A^{[K]} + B^{[K]}}} \quad (5.51)$$

According to  $E(\log W) = \left. \frac{\partial E(W^r)}{\partial r} \right|_{r=0}$  (Mencía and Sentana, 2005)

$$\xi_i^{[k]} = \frac{1}{2} \log \left( \frac{A^{[k]}}{B^{[k]}} \right) \frac{\left. \frac{\partial K_{-\frac{\nu+d}{2}+\alpha} \sqrt{A^{[K]}+B^{[K]}}}{\partial \alpha} \right|_{\alpha=0}}{K_{\frac{\nu+d}{2}} \sqrt{A^{[K]} + B^{[K]}}} \quad (5.52)$$

### M Step

In M-step, the parameters are computed by maximizing the log-likelihood function (5.48) and (5.49). Take partial derivative of  $L_1$  with respect to  $\gamma$ ,  $\Sigma$ ,  $\mu_z$ ,

$$\frac{\partial L_1}{\partial \gamma} = \sum_{i=1}^T (z_i - \mu_z)^T \Sigma^{-1} - 2 \sum_{i=1}^T \frac{w_i}{2} \gamma^T \Sigma^{-1} = 0 \quad (5.53)$$

$$\frac{\partial L_1}{\partial \mu_z} = - \sum_{i=1}^T \Sigma^{-1} \gamma + \sum_{i=1}^T \frac{\Sigma^{-1} (z_i - \mu_z)}{w_i} = 0 \quad (5.54)$$

$$\begin{aligned} \frac{\partial L_1}{\partial \Sigma^{-1}} &= \sum_{i=1}^T (z_i - \mu_z)^T \gamma - \sum_{i=1}^T \frac{(z_i - \mu_z)(z_i - \mu_z)^T}{2w_i} \\ &- \sum_{i=1}^T \frac{w_i}{2} \gamma \gamma^T + \frac{T}{2} \Sigma^T = 0 \end{aligned} \quad (5.55)$$

Solve equation (5.54) we have

$$\mu_z = \frac{\sum_{i=1}^T \frac{z_i}{w_i} - \sum_{i=1}^T \gamma}{\sum_{i=1}^T \frac{1}{w_i}} \quad (5.56)$$

Substitute (5.56) into (5.53) and solve  $\gamma$

$$\gamma = \frac{\sum_{i=1}^T w_i^{-1} (z_i - \bar{z}_i)}{T^{-1} \sum_{i=1}^T w_i w_i^{-1} + T} \quad (5.57)$$

$$\Sigma = \frac{1}{T} \left( \sum_{i=1}^T \frac{(z_i - \mu_z)(z_i - \mu_z)^T}{2w_i} - \sum_{i=1}^T \frac{w_i}{2} \gamma \gamma^T \right) \quad (5.58)$$

The parameter  $\nu$  is given by solving the equation below numerically.

$$\frac{\partial L_2}{\partial \nu} = \frac{T}{2} \log \left( \frac{\nu}{2} \right) + \frac{T}{2} - \frac{1}{2} \sum_{i=1}^T \log w_i - \frac{1}{2} \sum_{i=1}^T \frac{1}{w_i} - \frac{T \Gamma'(\frac{\nu}{2})}{\Gamma(\frac{\nu}{2})} = 0 \quad (5.59)$$

Let  $\bar{e} = \frac{1}{T} \sum_{i=1}^T e_i$ ,  $\bar{\zeta} = \frac{1}{T} \sum_{i=1}^T \zeta_i$ , and  $\bar{\xi} = \frac{1}{T} \sum_{i=1}^T \xi_i$ . The initial value for  $\mu_z$ ,  $\Sigma$  and  $\gamma$  are the sample mean, sample covariance and zero vector respectively.

Then according to equation (5.50), (5.51) and (5.52), we can compute  $e_i^{[k]}$ ,  $\zeta_i^{[k]}$  and  $\xi_i^{[k]}$  at step  $k$ .

Next the parameters are updated by

$$\begin{aligned} \gamma^{[k+1]} &= \frac{T^{-1} \sum_{i=1}^T e_i^{[k]} (\bar{z} - z_i)}{\bar{e}^{[k]} \bar{\zeta}^{[k]} + 1} \\ \mu_z^{[k+1]} &= \frac{T^{-1} \sum_{i=1}^T e_i^{[k]} z_i - \gamma^{[k]}}{\bar{e}^{[k+1]}} \\ \Sigma^{[k+1]} &= \frac{1}{T} \left( \sum_{i=1}^T e_i^{[k]} (z_i - \mu_z^{[k+1]})(z_i - \mu_z^{[k+1]})^T - \bar{\zeta}^{[k]} \gamma^{[k+1]} \gamma^{[k+1]T} \right) \end{aligned}$$

$\nu^{[k+1]}$  is given as the solution to the following equation

$$\log \left( \frac{\nu}{2} \right) + 1 - \bar{\xi}^{[k]} - \bar{e}^{[k]} - \frac{\Gamma'(\frac{\nu}{2})}{\Gamma(\frac{\nu}{2})} = 0$$

## 5.5 Numerical Result

The data used in this section is the daily returns of four HFRX indices (EH, EMN, M and MA) covering the period from April 1, 2003 to May 12, 2014. We consider portfolios composed of these four indices.

In Table 5.1 and 5.2, we present summary statistics. Note that the average daily return of MA is the highest, while the average return on EMN is negative. The large kurtosis indicates that all the return time series have fat tails. Table 5.2 shows that the correlations of these returns are positive and the covariance of M is largest.



	mean	skewness	kurtosis
EH	0.005	-0.8406	8.63
EMN	-0.001	-0.1857	20.36
M	0.004	-1.0434	10.58
MA	0.016	1.5449	70.64

Table 5.1: Statistics of daily returns of fund indices: Mean of daily returns of the fund indices in percentage

	EH	EMN	M	MA
EH	1.673	0.1633	0.420	0.6086
EMN	0.163	0.6644	0.1623	0.0606
M	0.420	0.1623	1.705	0.1085
MA	0.607	0.0606	0.1085	0.8063

Table 5.2: Covariance of daily returns of the fund indices ( $10^{-5}$ )

Then the daily return data are filtered by OGARCH model with Gaussian innovations and the standardized filtered time series  $z_t$  is calibrated by the multivariate GH skewed  $t$  distribution. Figure 5.1 shows the stochastic correlation of HFRX indices of daily returns estimated by OGARCH model.

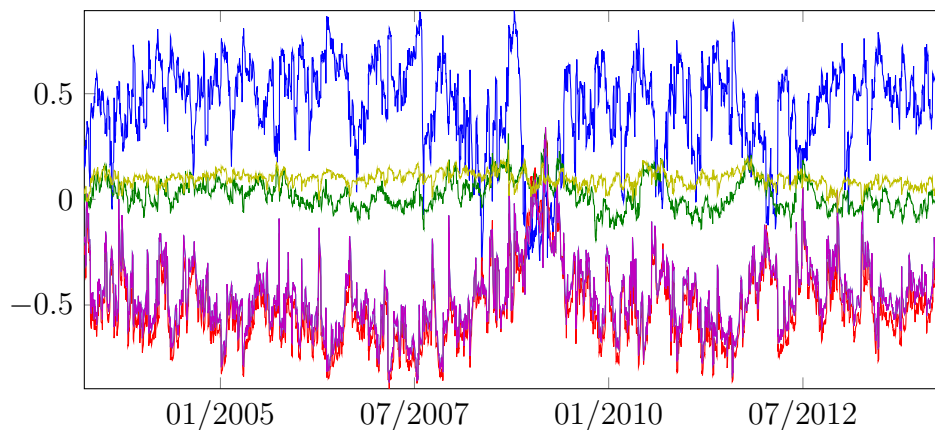


Figure 5.1: The correlation of HFRX indices daily returns estimated by OGARCH model.

Figure 5.2 displays normal QQ plots of the filtered returns. The marginal densities of the filtered series are skewed and have fatter tails than Gaussian distribution. Therefore, we fit GH skew  $t$ , student  $t$ , NIG, GH and Variance Gamma distributions to the filtered data. Table 5.3 presents parameters estimated by various distributions and log likelihood statistics. The GH skew  $t$  distribution has the largest log likelihood, so it fits the best, while Gaussian distribution has the smallest log likelihood. From Figure 5.3 we can see that GH skew  $t$  distribution has a better fit to the data.

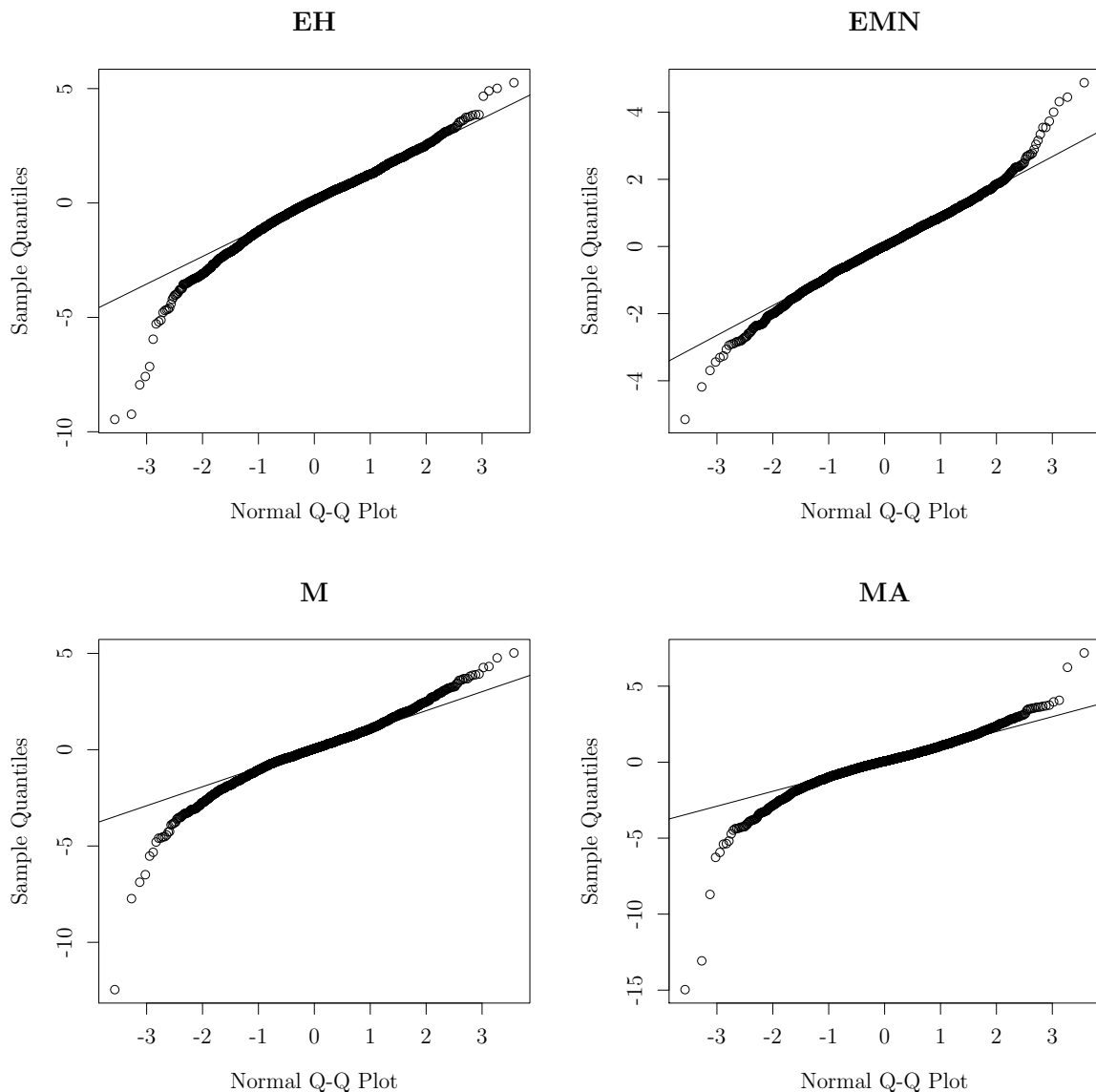


Figure 5.2: QQ-plot versus normal distribution of filtered daily returns of HFRX indices.

### 5.5.1 Efficient Frontier

In this section, we solve the Mean-Variance, enhanced safety first and Mean-CVaR portfolio optimization problems by assuming GH skewed  $t$  distribution and plot efficient frontiers versus different risk measures. First, we calibrate the mean and covariance of a multivariate GH skewed  $t$  to our filtered data, and forecast one week ahead covariance and mean of weekly return on May 12, 2014. Then we calculate the optimal portfolios by using different optimization models.

Figure 5.4 displays the efficient frontiers for Mean-Variance and Mean-CVaR portfolio optimization models. According to Figure 5.4, for a given return, the Mean-CVaR portfolio

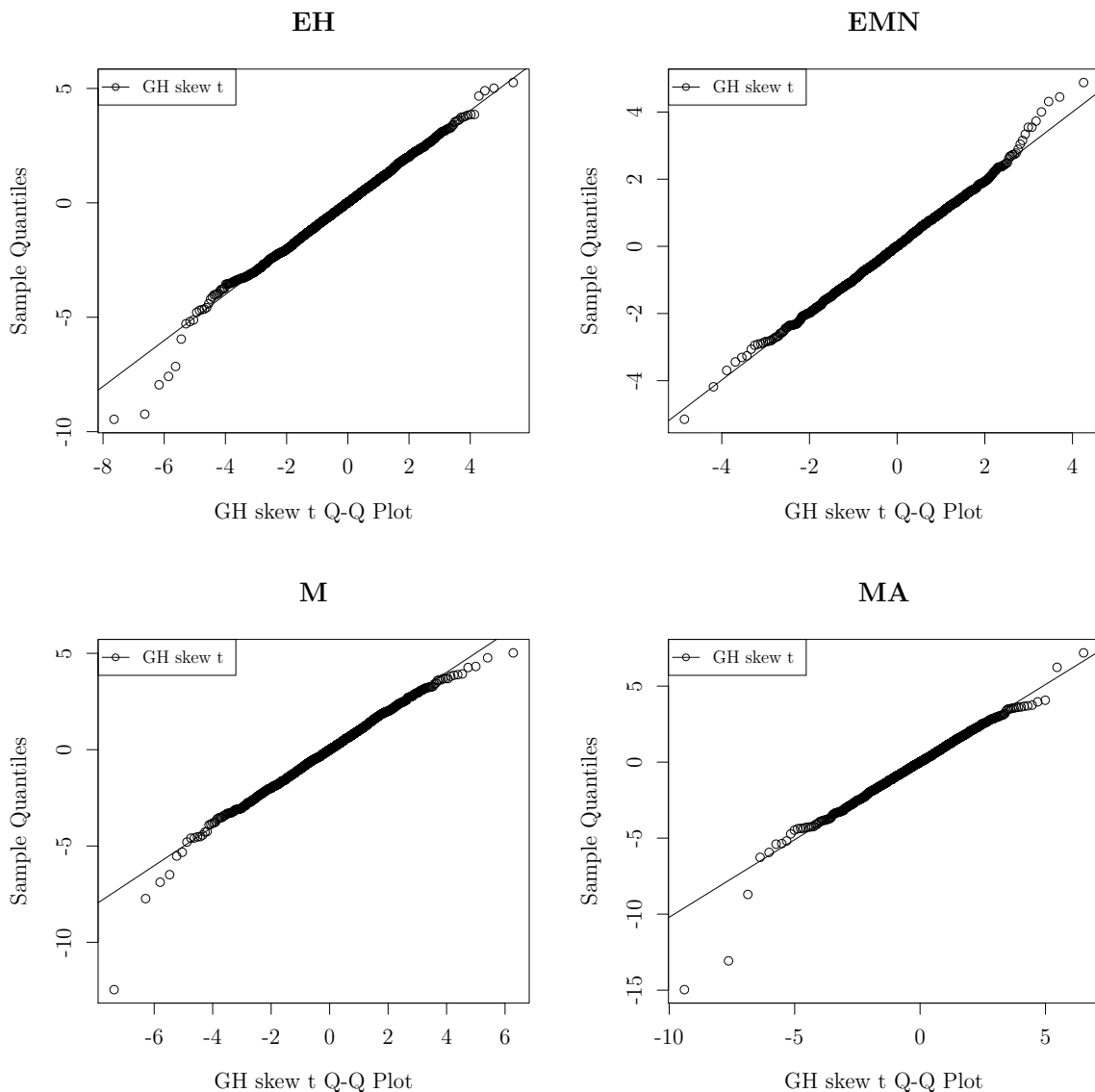


Figure 5.3: QQ-plot versus GH skewed  $t$  distribution of filtered daily returns of HFRX indices.

has a higher standard deviation than the Mean-Variance portfolio. That means that Mean-CVaR portfolios can protect against the situations of extreme loss better.

### 5.5.2 Portfolio Performance Analysis

In this section, the performances of the optimal portfolios selected by using Mean-Variance, enhanced safety-first and Mean-CVaR models are compared in 2008 financial crisis period. The weekly returns of HFRX indices for the period from April 1, 2003 to Dec. 31 2007 are used to estimate the parameters and then the first optimal portfolio is selected on

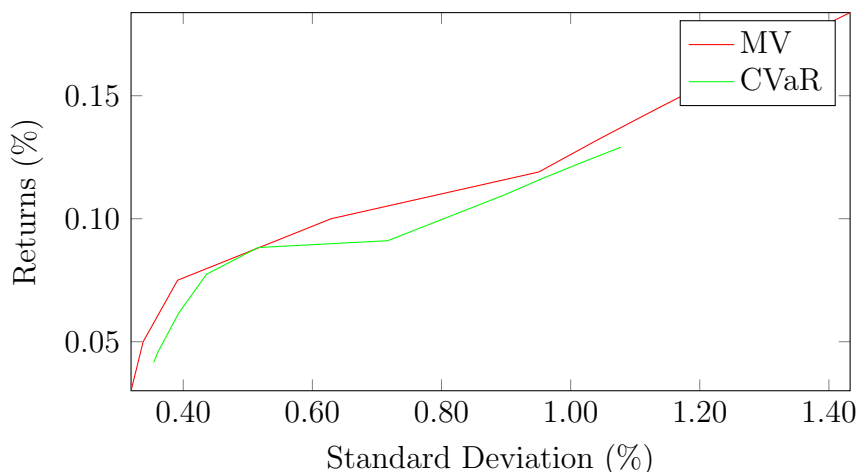


Figure 5.4: Mean-Variance efficient frontier vs 0.95 Mean-CVaR .

Dec. 31, 2007. The portfolio is rebalanced every week after the first investment period based on the information available up to the investment date. For each model, the weight of each fund index is between -1 and 1. We assume  $\lambda$  equals to 1 in Mean-Variance model and the probability  $p$  is 0.05 in enhanced safety-first model. The 0.95-CVaR is considered in the mean-CVaR model. The historical weekly returns are shown in Figure 5.5.

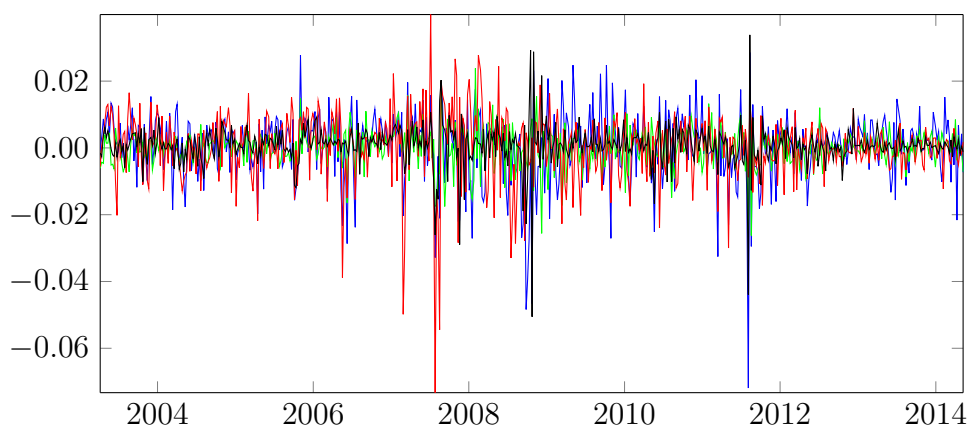


Figure 5.5: The weekly returns of HFRX indices from 2003-2014 .

### Portfolio performance in 2008

The realized portfolio returns based on the different models in the out-of-sample period (2007.12.31-2008.12.31) are plotted in Figure 5.6 and main statistics is reported in Table 5.4. As shown in Figure 5.6, the portfolio returns selected by minimizing 0.95-CVaR are more volatile. The optimal portfolio returns are very close for Mean-Variance and safety-first models. The optimal portfolio returns selected by the safety-first model have the highest mean and smallest standard deviation. The Mean-Variance model gives the maximum return. As can be observed in figures 5.6, the weekly returns do not vary much

between these three portfolios for most weeks. The most significant difference is that the losses are larger for Mean-CVaR and Mean-Variance models in October 2008. For a risk averse investor, the safety-first strategy is preferred in the financial crisis period.

### Portfolio performance in 2013-2014

The second out-of-sample period is chosen from May 2013 to May 2014. From Figure 5.5 we can see that the weekly returns are very stable in this period. Table 5.5 shows that Mean-CVaR model has the highest return and outperforms Mean-Variance model based on all the statistics. Both the portfolio sets selected by the Mean-CVaR and Mean-Variance methods yield a significantly higher return for this period. Both Table 5.5 and Figure 5.7 show that portfolio returns selected by the safety-first model are more stable. Compare Table 5.4 and 5.5, we find that the average return of portfolios selected by the enhanced safety-first model in 2008 is even better than that in 2013-2014. Meanwhile, the difference in rates of mean return in this period is much smaller than during the financial crisis period. According to Table 5.5, the enhanced safety-first has the smallest daily return, standard deviation and return spread between min and max returns, while classic Mean-Variance model yields the best return. The Sharpe ratio of enhanced safety-first model is the largest, which means the enhanced safety-first model achieves the best risk-adjusted return. For a risk averse investor, the enhanced safety-first model is the best choice.

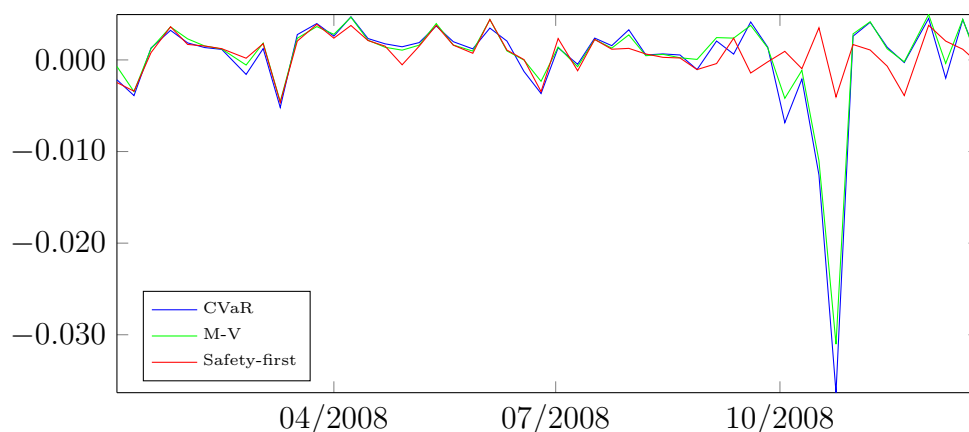


Figure 5.6: The realized weekly returns of optimal portfolios in 2008.

## 5.6 Conclusion

In this chapter, we enhance the safety-first model with standard deviation constraint and derive an analytic formula by filtering the returns with GH skewed  $t$  distribution and OGARCH. The solution is also given for classical mean-variance model in terms of parameters in GH skewed  $t$  distribution and OGARCH model. Then we compare

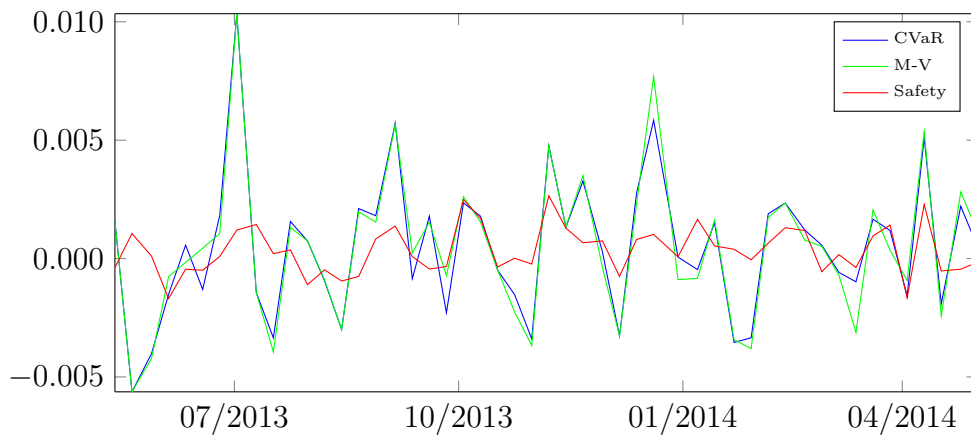


Figure 5.7: The realized weekly returns of portfolio for 2013.4-2014.4 .

the performances of the optimal portfolio selected by Mean-Variance, Mean-CVaR and enhanced safety-first models in 2008 financial crisis and stable period(2013-2014). The efficient frontier is also presented. It shows that Mean-CVaR model gives the highest mean return in the post-crisis period, while the enhanced safety-first model outperforms the other two models during the financial crisis period. Meanwhile, the enhanced safety-first model achieves the best risk-adjusted returns in both periods. For a risk averse investor, the enhanced safety-first model should be considered.

Model	$\nu(\lambda)$	$\mu$	$\gamma$	$\Sigma$						log Likelihood		
				1.84	-0.16	0.33	1.15	-0.16	0.96		0.11	0.28
GH Skewed $t$	6.20	[ 0.33 0.07 0.21 0.27]	[-0.30 -0.07 -0.20 -0.270]	0.33	1.15	0.11	0.28	0.33	1.15	0.11	0.28	-15897.77
Student $t$	6.03	[0.08 0.01 0.05 0.05]	[0 0 0 0]	1.89	-0.16	0.36	1.19	-0.16	0.97	0.11	0.28	-15917.59
NIG	-0.5	[ 0.34 0.07 0.21 0.27]	[-0.32 -0.072 -0.19 -0.27]	1.82	-0.16	0.33	1.14	-0.16	0.95	0.11	0.28	-15907.40
GH	2	[0.15 0.026 0.08 0.11]	[-0.13 -0.03 -0.07 -0.11]	1.88	-0.19	0.35	1.20	-0.15	0.94	0.11	0.29	-15935.91
Variance Gamma	2.59	[0.35 0.07 0.18 0.27]	[-0.33 -0.07 -0.17 -0.27]	1.79	-0.16	0.32	1.13	-0.16	0.93	0.10	0.27	-15932.95
Gaussian		[0.021 -0.002 0.012 0.002]		1.93	-0.13	0.36	1.28	-0.13	0.90	0.10	0.30	-16448.49

Table 5.3: Estimated parameters

	CVAR(0.95)	Mean-Variance	Safety-first
Mean	0	0.038	0.071
St. Dev	0.61	0.53	0.22
5% Quantile	-0.68	-0.46	-0.39
Min	-3.63	-3.10	-0.46
Max	0.47	0.50	0.44
Sharpe Ratio	0	0.072	0.323

Table 5.4: Statistics of realized portfolio returns for the out-of-sample period(2007-2008). average portfolio return(Mean), standard deviation(St.Dev), worst-case return(Min) are in percentages.

	CVAR(0.95)	Mean-Variance	Safety-first
Mean	0.046	0.044	0.033
St. Dev	0.29	0.31	0.10
5% Quantile	-0.35	-0.39	-0.11
Min	-0.56	-0.56	-0.17
Max	1.03	1.03	0.27
Sharpe Ratio	0.158	0.147	0.33

Table 5.5: Statistics of realized portfolio returns for the out-of-sample period (2013-2014). mean, standard deviation(St. Dev), worst-case return(Min) are in percentages.



# Bibliography

- I. H. Aas, K. and Haff. The generalized hyperbolic skew students t-distribution. *Journal of financial econometrics*, 4(2):275–309, 2006.
- V. Agarwal and N. Y. Naik. Risks and portfolio decisions involving hedge funds. *Review of Financial studies*, 17(1):63–98, 2004.
- P. Agnolucci. Volatility in crude oil futures: a comparison of the predictive ability of garch and implied volatility models. *Energy Economics*, 31(2):316–321, 2009.
- C. Alexander. A primer on the orthogonal garch model. *manuscript ISMA Centre, University of Reading, UK*, 2000.
- C. Alexander. Correlation in crude oil and natural gas markets. *Managing energy price risk*, pages 573–606, 2004.
- C. Alexander and A Chibumba. Multivariate orthogonal factor garch. *University of Sussex, Mimeo*, 1997.
- G. J Alexander, A. M Baptista, and S. Yan. Mean–variance portfolio selection with at-risk constraints and discrete distributions. *Journal of Banking & Finance*, 31(12): 3761–3781, 2007.
- G.J Alexander and A. M Baptista. A comparison of var and cvar constraints on portfolio selection with the mean-variance model. *Management Science*, 50(9):1261–1273, 2004.
- A.H Alizadeh, N. K Nomikos, and P. K Pouliasis. A markov regime switching approach for hedging energy commodities. *Journal of Banking & Finance*, 32(9):1970–1983, 2008.
- C. Aloui and S. Mabrouk. Value-at-risk estimations of energy commodities via long-memory, asymmetry and fat-tailed garch models. *Energy Policy*, 38(5):2326–2339, 2010.
- Y. Amihud and A. Barnea. Portfolio selection for managerial control. *Omega*, 2(6): 775–783, 1974.

- Philippe Artzner, Freddy Delbaen, Jean-Marc Eber, and David Heath. Coherent measures of risk1. *Risk management: value at risk and beyond*, page 145, 2002.
- R. Bacon and M. Kojima. Coping with oil price volatility. 2008.
- E. Ballesteros and D. Pla-Santamaria. Selecting portfolios for mutual funds. *Omega*, 32(5): 385–394, 2004.
- Deutsche Bank. Twelfth annual alternative investment survey. URL <https://www.managedfunds.org/wp-content/uploads/2014/04/2014-Deutsche-Bank-AIS.pdf>.
- O. Barndorff-Nielsen. Exponentially decreasing distributions for the logarithm of particle size. *Proceedings of the Royal Society of London. A. Mathematical and Physical Sciences*, 353(1674):401–419, 1977.
- O. Barndorff-Nielsen and P. Blaesild. *Hyperbolic distributions and ramifications: Contributions to theory and application*. Springer, 1981.
- O. E Barndorff-Nielsen, P R. Hansen, and N. Lunde, A.and Shephard. Realized kernels in practice: Trades and quotes. *The Econometrics Journal*, 12(3):C1–C32, 2009.
- S. Bauwens, L.and Laurent. A new class of multivariate skew densities, with application to generalized autoregressive conditional heteroscedasticity models. *Journal of Business & Economic Statistics*, 23(3), 2005.
- D. Bicchetti and N. Maystre. The synchronized and long-lasting structural change on commodity markets: evidence from high frequency data. *Algorithmic Finance*, 2(3): 233–239, 2013.
- M. Billio, M. Getmansky, and L. Pelizzon. Dynamic risk exposures in hedge funds. *Computational Statistics & Data Analysis*, 56(11):3517–3532, 2012.
- S. Blazsek and A. Downarowicz. Regime switching models of hedge fund returns. *Working Papers (Universidad de Navarra. Facultad de Ciencias Económicas y Empresariales)*, (12):1, 2008.
- S. Blazsek and A. Downarowicz. Forecasting hedge fund volatility: a markov regime-switching approach. *The European Journal of Finance*, 19(4):243–275, 2013.
- T. Bollerslev. Generalized autoregressive conditional heteroskedasticity. *Journal of econometrics*, 31(3):307–327, 1986.
- T. Bollerslev. Modelling the coherence in short-run nominal exchange rates: a multivariate generalized arch model. *The Review of Economics and Statistics*, pages 498–505, 1990.

- T. Bollerslev and J. M Wooldridge. Quasi-maximum likelihood estimation and inference in dynamic models with time-varying covariances. *Econometric reviews*, 11(2):143–172, 1992.
- T. Bollerslev, R. F Engle, and J. M Wooldridge. A capital asset pricing model with time-varying covariances. *The Journal of Political Economy*, pages 116–131, 1988.
- Walter Briec and Kristiaan Kerstens. Multi-horizon markowitz portfolio performance appraisals: a general approach. *Omega*, 37(1):50–62, 2009.
- M. Caporin, M. McAleer. Robust ranking of multivariate garch models by problem dimension. *Computational Statistics & Data Analysis*, 2012.
- N. Chan, M. Getmansky, S. M. Haas, and A. W. Lo. Systemic risk and hedge funds. Technical report, National Bureau of Economic Research, 2005.
- M. Chen, L. Wang, J. Wang, S. Sun, and C. Xia. Impact of individual response strategy on the spatial public goods game within mobile agents. *Applied Mathematics and Computation*, 251:192–202, 2015.
- H. Y Chiu, M. C. Wong and D. Li. Roys safety-first portfolio principle in financial risk management of disastrous events. *Risk Analysis*, 32(11):1856–1872, 2012.
- V. K. Chopra and W. T. Ziemba. The effect of errors in means, variances, and covariances on optimal portfolio choice. *The journal of portfolio management*, 19(2):6–11, 1993.
- P. Christoffersen. Value-at-risk models. In *Handbook of Financial Time Series*, pages 753–766. Springer, 2009.
- P. F Christoffersen. Evaluating interval forecasts. *International economic review*, pages 841–862, 1998.
- A. Cologni and M. Manera. The asymmetric effects of oil shocks on output growth: A markov-switching analysis for the g-7 countries. *Economic Modelling*, 26(1):1–29, 2009.
- D. Cousineau, S. Brown, and A. Heathcote. Fitting distributions using maximum likelihood: Methods and packages. *Behavior Research Methods, Instruments, & Computers*, 36(4):742–756, 2004.
- A. Creti and V. Joëts, M. and Mignon. On the links between stock and commodity markets’ volatility. *Energy Economics*, 37:16–28, 2013.
- Arthur P Dempster, Nan M Laird, and Donald B Rubin. Maximum likelihood from incomplete data via the em algorithm. *Journal of the royal statistical society. Series B (methodological)*, pages 1–38, 1977.

- S. Ding, J. Wang, S. Ruan, and C. Xia. Inferring to individual diversity promotes the cooperation in the spatial prisoners dilemma game. *Chaos, Solitons & Fractals*, 71: 91–99, 2015.
- Y. Ding and B. Zhang. Optimal portfolio of safety-first models. *Journal of Statistical Planning and Inference*, 139(9):2952–2962, 2009.
- Z. Ding. *Time series analysis of speculative returns*. PhD thesis, University of California, San Diego, Department of Economics, 1994.
- C. Engel. Can the markov switching model forecast exchange rates? *Journal of International Economics*, 36(1):151–165, 1994.
- R. Engle. Dynamic conditional correlation: A simple class of multivariate generalized autoregressive conditional heteroskedasticity models. *Journal of Business & Economic Statistics*, 20(3):339–350, 2002.
- R. Engle and K. Sheppard. Evaluating the specification of covariance models for large portfolios. *New York University, working paper*, 2008.
- R. F. Engle. Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation. *Econometrica: Journal of the Econometric Society*, pages 987–1007, 1982.
- R. F Engle and K. F Kroner. Multivariate simultaneous generalized arch. *Econometric theory*, 11(01):122–150, 1995.
- R. F Engle and K. Sheppard. Theoretical and empirical properties of dynamic conditional correlation multivariate garch. Technical report, National Bureau of Economic Research, 2001.
- R.F. Engle, V. K. Ng, and M. Rothschild. Asset pricing with a factor-arch covariance structure: Empirical estimates for treasury bills. *Journal of Econometrics*, 45(1): 213–237, 1990.
- E. F Fama. Efficient capital markets: A review of theory and empirical work\*. *The journal of Finance*, 25(2):383–417, 1970.
- Y. Fan, Y. and Zhang, H. Tsai, and Y. Wei. Estimating value at risk of crude oil price and its spillover effect using the ged-garch approach. *Energy Economics*, 30(6):3156–3171, 2008.

- L. Ferruz Agudo and J. L Sarto Marzal. An analysis of spanish investment fund performance: some considerations concerning sharpe's ratio. *Omega*, 32(4):273–284, 2004.
- C. A Floudas and V. Visweswaran. Quadratic optimization. In *Handbook of global optimization*, pages 217–269. Springer, 1995.
- W.M Fong and K. H See. Modelling the conditional volatility of commodity index futures as a regime switching process. *Journal of Applied Econometrics*, 16(2):133–163, 2001.
- A. J Frey, R. and McNeil. Var and expected shortfall in portfolios of dependent credit risks: conceptual and practical insights. *Journal of Banking & Finance*, 26(7): 1317–1334, 2002.
- M. Getmansky, A. W. Lo, and I. Makarov. An econometric model of serial correlation and illiquidity in hedge fund returns. *Journal of Financial Economics*, 74(3):529–609, 2004.
- D. Giamouridis and I. D Vrontos. Hedge fund portfolio construction: A comparison of static and dynamic approaches. *Journal of Banking & Finance*, 31(1):199–217, 2007.
- J. D Hamilton. A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica: Journal of the Econometric Society*, pages 357–384, 1989.
- B. E Hansen. Autoregressive conditional density estimation. *International Economic Review*, pages 705–730, 1994.
- R DF Harris and M. Mazibas. Dynamic hedge fund portfolio construction. *International Review of Financial Analysis*, 19(5):351–357, 2010.
- W. Hu and A. N Kercheval. Portfolio optimization for student t and skewed t returns. *Quantitative Finance*, 10(1):91–105, 2010.
- T. Hui and EK Kwan. International portfolio diversification: A factor analysis approach. *Omega*, 22(3):263–267, 1994.
- J. Hung, M. Lee, and H. Liu. Estimation of value-at-risk for energy commodities via fat-tailed garch models. *Energy Economics*, 30(3):1173–1191, 2008.
- P. Jorion. *Financial Risk Manager Handbook 2001-2002*, volume 96. John Wiley & Sons, 2001.
- D. Karlis. An em type algorithm for maximum likelihood estimation of the normal–inverse gaussian distribution. *Statistics & Probability Letters*, 57(1):43–52, 2002.

- H. M Kat and C. Brooks. The statistical properties of hedge fund index returns and their implications for investors. *Cass Business School Research Paper*, 2001.
- S. Kataoka. A stochastic programming model. *Econometrica: Journal of the Econometric Society*, pages 181–196, 1963.
- Yuriy Kazmerchuk, Anatoliy Swishchuk, and Jianhong Wu. A continuous-time garch model for stochastic volatility with delay. *Canadian Applied Mathematics Quarterly*, 13(2):123–149, 2005.
- F. Klaassen. *Improving GARCH volatility forecasts with regime-switching GARCH*, pages 223–254. Springer, 2002.
- J. VK Laurent, S. and Rombouts and F. Violante. On the forecasting accuracy of multivariate garch models. *Journal of Applied Econometrics*, 27(6):934–955, 2012.
- Yang Liu. Risk forecasting and portfolio optimization with garch, skewed t distributions and multiple timescales. 2012.
- J. Lu and P. Tsai. Forecasting equity market correlation: A comparison of alternative measures. *Journal of International and Global Economic Studies*, 2(2):77–95, 2010.
- F. Malik and B. T Ewing. Volatility transmission between oil prices and equity sector returns. *International Review of Financial Analysis*, 18(3):95–100, 2009.
- S. Manganello. Asset allocation by variance sensitivity analysis. *Journal of Financial Econometrics*, 2(3):370–389, 2004.
- H. Markowitz. Portfolio selection\*. *The journal of finance*, 7(1):77–91, 1952.
- L Martellini, M Vaissié, and V Ziemann. Investing in hedge funds: Adding value through active style allocation decisions. *Journal of Portfolio Management*, 2007.
- M. Maslakovic. Hedge funds 2009. *IFSL Research*, April 2009.
- D. S Matteson and R. S Tsay. Dynamic orthogonal components for multivariate time series. *Journal of the American Statistical Association*, 106(496), 2011.
- A. J McNeil, R. Frey, and P. Embrechts. *Quantitative risk management: concepts, techniques, and tools*. Princeton university press, 2010.
- Alexander J McNeil, Rüdiger Frey, and Paul Embrechts. *Quantitative Risk Management: Concepts, Techniques and Tools: Concepts, Techniques and Tools*. Princeton university press, 2015.

- Javier Mencía and Enrique Sentana. Estimation and testing of dynamic models with generalized hyperbolic innovations. 2005.
- J.P. Morgan. *Riskmetrics: technical document*. Morgan Guaranty Trust Company of New York, 1996.
- P.K Narayan, S. Narayan, and A. Prasad. Understanding the oil price-exchange rate nexus for the fiji islands. *Energy Economics*, 30(5):2686–2696, 2008.
- Daniel B Nelson. Conditional heteroskedasticity in asset returns: A new approach. *Econometrica: Journal of the Econometric Society*, pages 347–370, 1991.
- V.I Norkin and S.V Boyko. Safety-first portfolio selection. *Cybernetics and Systems Analysis*, 48(2):180–191, 2012.
- R Östermark. Portfolio efficiency of univariate time series models. *Omega*, 18(2):159–169, 1990.
- J. Palczewski, A. and Palczewski. Theoretical and empirical estimates of mean–variance portfolio sensitivity. *European Journal of Operational Research*, 234(2):402–410, 2014.
- A. J Patton. Volatility forecast comparison using imperfect volatility proxies. *Journal of Econometrics*, 160(1):246–256, 2011.
- M. Perlin. Ms regress-the matlab package for markov regime switching models. *Available at SSRN 1714016*, 2010.
- M. H. Pesaran and P. Zaffaroni. Model averaging and value-at-risk based evaluation of large multi-asset volatility models for risk management. *CEPR Discussion Paper*, 2005.
- Y. Plyakha, R. Uppal, and G. Vilkov. Why does an equal-weighted portfolio outperform value-and price-weighted portfolios. *Preprint available at [http://dx. doi. org/10.2139/ssrn, 1787045](http://dx.doi.org/10.2139/ssrn.1787045)*, 2012.
- S. Poon and C.WJ Granger. Forecasting volatility in financial markets: A review. *Journal of economic literature*, 41(2):478–539, 2003.
- A Rashid and S Ahmad. Predicting stock returns volatility: An evaluation of linear vs. nonlinear methods. *International Research Journal of Finance and Economics*, 20: 141–150, 2008.
- J. E Raymond and R. W Rich. Oil and the macroeconomy: a markov state-switching approach. *Journal of Money, Credit, and Banking*, pages 193–213, 1997.

- RiskMetrics. Technical document. *J.P.Morgan*, Dec. 1996.
- R T. Rockafellar and S. Uryasev. Optimization of conditional value-at-risk. *Journal of risk*, 2:21–42, 2000.
- A. D. Roy. Safety first and the holding of assets. *Econometrica: Journal of the Econometric Society*, pages 431–449, 1952.
- P. Sadorsky. Modeling and forecasting petroleum futures volatility. *Energy Economics*, 28(4):467–488, 2006.
- A. AP Santos and G. V Moura. Dynamic factor multivariate garch model. *Computational Statistics & Data Analysis*, 2012.
- D. Saunders, L. Seco, C. Vogt, and R. Zagst. A fund of hedge funds under regime switching. *The Journal of Alternative Investments*, 15(4):8–23, 2013.
- T. Sawik. Selection of supply portfolio under disruption risks. *Omega*, 39(2):194–208, 2011.
- A. Silvennoinen and T. Teräsvirta. Multivariate garch models. In *Handbook of Financial Time Series*, pages 201–229. Springer, 2009.
- Anatoliy Swishchuk. Variance and volatility swaps in energy markets. *The Journal of Energy Markets*, 6(1):33, 2013.
- Anatoliy Swishchuk and Matthew Couch. Volatility and variance swaps for the cogarch (1, 1) model. *Wilmott Journal*, 2(5):231–246, 2010.
- G. YN Tang. How efficient is naive portfolio diversification? an educational note. *Omega*, 32(2):155–160, 2004.
- O. Taramasco, S. Bauer, and M. O. Taramasco. Package rhmm. 2013.
- L. G Telsler. Safety first and hedging. *The Review of Economic Studies*, pages 1–16, 1955.
- C.A Valle, N Meade, and J.E Beasley. Absolute return portfolios. *Omega*, 45:20–41, 2014.
- R. Van der Weide. Go-garch: a multivariate generalized orthogonal garch model. *Journal of Applied Econometrics*, 17(5):549–564, 2002.
- Minh T Vo. Regime-switching stochastic volatility: Evidence from the crude oil market. *Energy Economics*, 31(5):779–788, 2009.



- I. D. Vrontos and D. Giamouridis. Hedge fund return predictability in the presence of model uncertainty and implications for wealth allocation. *Available at SSRN 1274201*, 2008.
- C. Wang, Y. and Wu. Forecasting energy market volatility using garch models: Can multivariate models beat univariate models? *Energy Economics*, 34(6):2167–2181, 2012.
- R. I Webb. A note on volatility and pricing of futures options during choppy markets. *Journal of Futures Markets*, 7(3):333–337, 1987.
- Y. Wei, Y. Wang, and D. Huang. Forecasting crude oil market volatility: Further evidence using garch-class models. *Energy Economics*, 32(6):1477–1484, 2010.
- ASK Wong and PJG Vlaar. Modelling time-varying correlations of financial markets. *De Nederlandsche Bank NV Research Memorandum*, 739:0319, 2003.
- D. Wu and D. L Olson. Enterprise risk management: coping with model risk in a large bank. *Journal of the Operational Research Society*, 61(2):179–190, 2010.
- D. Wu, K. Xie, G Chen, and P Gui. A risk analysis model in concurrent engineering product development. *Risk Analysis*, 30(9):1440–1453, 2010.
- C. Xia, S. Meloni, and Y. Moreno. Effects of environment knowledge on agglomeration and cooperation in spatial public goods games. *Advances in Complex Systems*, 15 (supp01):1250056, 2012.
- C. Zhu, S. Sun, L. Wang, S. Ding, J. Wang, and C. Xia. Promotion of cooperation due to diversity of players in the spatial public goods game with increasing neighborhood size. *Physica A: Statistical Mechanics and its Applications*, 406:145–154, 2014.

## List of Publications

- Chapter 2 of this thesis is published in the journal of Kybernetes, with the reference of "C, Luo, L.A. Seco, H. Wang, and D. Wu. "Risk modeling in crude oil market: a comparison of Markov switching and GARCH models." Kybernetes 39, no. 5 (2010): 750-769".
- Chapter 4 of this thesis is published in the journal of Omega, with the reference of "C, Luo, L. Seco, and L. Wu. "Portfolio Optimization in Hedge Funds by OGARCH and Markov Switching Model." Omega (2015), 34-39".
- Chapter 3 of this thesis is accepted to IEEE Systems Journal.