Risk Assessment for Intraday Trading

A Dissertation presented

by

Fangfei Dong

to

The Graduate School

in Partial Fulfillment of the

Requirements

for the Degree of

Doctor of Philosophy

in

Applied Mathematics and Statistics

Stony Brook University

December 2016

Stony Brook University

The Graduate School

Fangfei Dong

We, the dissertation committee for the above candidate for the

Doctor of Philosophy degree, hereby recommend

acceptance of this dissertation

Svetlozar Rachev - Dissertation Advisor
Professor, Dept. of Applied Mathematics and Statistics

Arron Kim - Chairperson of Defense
Assistant Professor of Finance, College of Business

Stoyan Stoyanov - Committee Member
Associate Professor of Finance, College of Business

Petar Djuric - External Member
Professor and Department Chair, Department of Electrical and Computer Engineering

This dissertation is accepted by the Graduate School

Charles Taber
Dean of the Graduate School

Abstract of the Dissertation

Risk Assessment in Intraday Trading

by

Fangfei Dong

Doctor of Philosophy

in

Applied Mathematics and Statistics

Stony Brook University

2016

In these days, high frequency hedge funds have developed as a new and successful category of hedge funds. Accordingly, risk management is now obliged to keep pace with this market and takes intraday-risk management into consideration.

To aim to contribute on answering questions on intraday risk management, the dissertation consists of three parts. In first part, an intraday risk assessment model incorporating long-range dependence and heavy-tailness is suggested. Fractional integrated time series model with nearly elliptical distributed innovations are used to compute more accurate intraday level value at risk. Second part investigates the market efficiency by analyzing the relation between market sentiment and price movement. A theoretical consumption-based equilibrium model and empirical analysis are employed to show various behavior under different market sentiment and cross-sectional stocks. The third parts further analyzes the long-range dependence behaviors in equity markets cross-sectionally on different sampling frequencies and various market conditions.

Dedication

This dissertation is dedicated to my husband, Yuzhong, my rock of support. I give my deepest expression of love and appreciation for the encouragement that you gave and the sacrifices you made during this graduate program. To my coming child, Shuchen. You have made me stronger and better. I love you to the moon and back.

Table of Contents

List of Figures

List of Tables

Chapter 1
Risk Assessment Model for Intraday Market

1.1 Introduction and Motivation

Financial industry participants are required by regulations to use mathematical models to forecast the risk associated with portfolios consisting of potentially hundreds and thousands of assets. A standard metric for representing this risk is known as Value-at-Risk (VaR). VaR refers to the maximum expected loss that will not be exceeded under normal market conditions over a predetermined period at a given confidence level (Jorion, 2001, p.xxii). VaR has been largely adopted by financial institutions as a foundation of day-to-day risk measurement since worldwide adoption of the Basel II Accord, beginning in 1999 and nearing completion today. In these days, VaR is probably the most widely used as a foundation of day-to-day risk measurement in financial industry.

As stated in Dionne, Duchesne and Pacurar 2009, financial institutions generally calculate their VaR at the end of the business day, to measure their total risk exposure over the next day. For regulated capital adequacy purposes, banks usually compute the market VaR daily and then re-scale it to a 10-day horizon.

Over the last decades, technology has transformed the way we trade securities and other financial instruments. The speed of trading has been constantly increasing. In many financial markets such as stock markets and future markets, the traditional way, human intermediation via floor trading or the telephone, has been abandoned and substituted by human intermediaries with an electronic limit order book or another automated trading system. This allows orders to buy and sell appearing and matching at a faster rate than ever before. Consequently, "High frequency finance hedge funds" have emerged as a new and successful category of hedge funds. According to the SEC, high-frequency traders are "professional traders acting in a proprietary capacity that engage in strategies that generate a large number of trades on daily basis." (SEC Concept Release on Equity Market Structure, 75 Fed. Reg. 3603, January 21, 2010).

For these high frequency traders, the horizon of their investment is generally less than a day. Therefore, their risk should be evaluated on a shorter time horizon than daily time intervals. In addition, agents who provide service to high frequency traders, such as brokers, should also have the ability to calculate the risk exposure as fast as their clients. In another word,risk management is now obliged to keep pace with this market and takes intraday-risk management into consideration. Moreover, as noted by Gourieroux and Jasiak 2010, and Dionne, Duchesne and Pacurar 2009, banks also use intraday risk analysis for internal control of their trading desk. For example, a trader could be asked at 11 a.m. to give his IVaR for the rest of the day.

Nowadays, high frequency trading takes account for large and larger volume in the market. Hendershott and Riordan (2011) find that HFT accounts for about 42% of (doublecounted) Nasdaq volume in large-cap stocks and 17% of volume in small-cap stocks. Brogaard (2012) similarly finds that 68% of trades have an HFT on at least one side of the transaction. As a consequence, A rapidly increasing attention from both financial industry and academia is focused on intraday risk management these days. Basel III requirements apply to intraday risk and liquidity measures, and pertain to intraday liquidity source and tools, monitoring tools and stress management.

However, existing methods for risk measure calculation are mainly restricted to Gaussian short memory models, which due to their derivation from assumptions with weak correlations, fast decay autocorrelations show rapid decay of probabilities for extreme events, known in the jargon as thin tails and short memory. Such models represent poor fit if directly used on intraday risk management since it is well known that intraday data displays fetures such as long memory, volatility clustering and excess leptokurtosis. Inaccuracies introduced by short memory Gaussian approximation or by dimension reduction techniques require large safety margins to be applied to VaR forecasts, thereby exaggerating the true VaR and requiring excess capital to be held back in reserve in less volatile market conditions, which becomes a even more serious issue if adopted by intraday risk evaluation.

To build an intraday risk assessment system and to minimize the disadvantage discussed in previous paragraph, a bottom-up approach is adopted and we need to model the returns of each asset and the dependence structure among them first, and then to model the probability distribution of portfolio returns. Autoregressive conditional heteroscedastic (ARCH) models introduced by Engle(Engle 1982) and their extensions to generalized ARCH (GARCH) models introduced by Bollerslev (Bollerslev 1986) capture two stylistic features of financial returns: volatility clustering and excess leptokurtosis. Accordingly, ARMA-GARCH models have been widely used in the industry to modell asset returns, and recently more complicated innovation distributions begin to be incorporated into ARMA-GARCH for a more realistic modeling.

A copula is a multivariate probability distribution for which the marginal probability distribution of each variable is uniform. Copulas are used to describe the dependence between random variables. Here we adopt a copula to model the dependence among assets.These types of models have seen as new development in the financial industry.

Intraday data generally displays long memories. To capture this feature, the ARMA-GARCH model is modified and the returns of each asset follow a fractional and fractionally integrated ARMA-GARCH (FARIMA-FIGARCH) process.

To forecast VaR for a portfolio of assets where the assumed asset return process is FARIMA-FIGARCH, the dependence among each asset's FARIMA-FIGARCH innovation is modeled with a copula function. In this way, one accounts for the heavy tails and volatility clustering of the individual assets separately from the dependence structure, among them, with its own heavy tails. This is referred to as a copula FARIMA-FIGARCH model. Quasi-maximum likelihood estimation is adopted here. We first estimates FARIMA-FIGARCH models for each individual asset with maximum likelihood estimation. Then we estimate the parameters of a copula dependence structure for the standardized residuals of each FARIMA-FIGARCH model.

To construction of the VaR of an estimated copula FARIMA-FIGARCH model, Monte Carlo and closed form expressions (for the probability function or for the characteristic function) will be adopted.

Another important component of the system is data cleaning. The data cleaning fills in missing data which are not observed in raw data. The missing data are recovered by the simple and multiple regression methods.

With considering large system which may have hundreds of thousands of assets, a factor analysis are also introduced here. The factor analysis is used for reducing the dimension of portfolio and to find important hidden factors having an effect on the portfolio risk. Two factor analyses are provided, including principal component analysis and GH-distributed factor analysis.

To wrap it up, he high speed risk assessment provided by conventional methods has been generally limited to Gaussian or low dimensional factor models, and has not been previously applied to long memory financial time series exhibiting clustering volatility with heavy tailed innovation process. In this article, we build a model based on multivariate long memory processes with volatility clustering and heavy tails allows accessing the risk of portfolios in intraday level. We believe this will contribute to risk assessment on intraday level and could be used in intraday portfolio optimization.

1.2 Data Cleaning

Before analyzing the data with FARIMA-FIGARCH model with GH innovations, the data is cleaned and missing data filled in. Financial data is inevitable has missing data problem. The common method to deal with this is back filling.

1.2.1 A General Backfill Method

By definition, a time series include two fields: date/time and value. Unfortunately, the data we are using not usually have value at same date/time, or data is not available at some date/time we are interested in. Therefore, we need to fill in the missing data.

To be more precise, denote the returns by $(r_t^{(i)})_{t\in\mathcal{T}}$ for i-th asset. And $\mathcal{T}^{(i)}$ is the set of times that $r^{(i)}$ is accessible, $\overline{\mathcal{T}}^{(i)}$ is the set of times that $r^{(i)}$ is not accessible.

1. $\mathcal{T}^{(i)} \cap \overline{\mathcal{T}}^{(i)} = \varnothing$;

2. $\mathcal{T} = \mathcal{T}^{(i)} \cup \overline{\mathcal{T}}^{(i)}$, for all i.

Factor Model

Assume the following relation:

$$r_t^{(i)} = \alpha + \beta_1 f_t^{(1)} + ... + \beta_d f_t^{(d)} + \varepsilon_t^{(i)},$$

where $f_t^{(j)}$, $j = 1, ..., d$ are the factors, which have no missing data, and $\varepsilon_t^{(i)}$ are i.i.d random variables following normal distribution $N(0, (\sigma^{(i)})^2)$.

OLS Regression

For each asset i, fit the factor model by minimizing the least squared error:

$$\widehat{\boldsymbol{\theta}} = \arg\min \sum_{t \in \mathcal{T}^{(i)}} \widehat{\xi}_t^2 \tag{1.1}$$

$$= \arg\min \sum_{t \in \mathcal{T}^{(i)}} \left(r_t^{(i)} - (\alpha + \beta_1 f_t^{(1)} + ... + \beta_d f_t^{(d)}) \right)^2, \tag{1.2}$$

where $\widehat{\boldsymbol{\theta}} = (\alpha, \beta_1, ..., \beta_d)$, and $\widehat{\xi}_t$ are the sample error.

This optimization gives

$$(\widehat{\alpha}, \widehat{\beta}_1, ...\widehat{\beta}_d)' = (F'F)^{-1} F' \boldsymbol{y} \tag{1.3}$$

Then in order to fill in missing data, two methods can be considered:

1. *Without Bootstrapping*:

   For $t \in \overline{\mathcal{T}}(i)$, let the filled-in data be

$$\widehat{r}_t^{(i)} = \widehat{\alpha} + \widehat{\beta}_1 f_t^{(1)} + ... + \widehat{\beta}_d f_t^{(d)}.$$

2. *With Bootstrapping*:

   For $t \in \overline{\mathcal{T}}(i)$, first generate a sample error $\widehat{\varepsilon}_t^{(i)}$ from $N(0, \widehat{(\sigma^{(i)})^2})$, where $\widehat{(\sigma^{(i)})^2}$ be the sample variance of the error. Then the filled-in data will be
$$\widehat{r}_t^{(i)} = \widehat{\alpha} + \widehat{\beta}_1 f_t^{(1)} + ... + \widehat{\beta}_d f_t^{(d)} + \widehat{\varepsilon}_t^{(i)}.$$

Maximum Likelihood

For each asset i, fit the model by maximizing the log likelihood function
$$\widehat{\boldsymbol{\theta}} = \arg\max l(\boldsymbol{\theta}; \boldsymbol{y}_{\mathcal{T}^{(i)}}^{(i)}, \boldsymbol{y}_{\overline{\mathcal{T}}^{(i)}}^{(i)}), \tag{1.4}$$

where $\boldsymbol{y}_{\mathcal{T}}$ denotes $\{y_t : t \in \mathcal{T}\}$, $\boldsymbol{y}_{\overline{\mathcal{T}}}$ denotes $\{y_t : t \in \overline{\mathcal{T}}\}$, and (to simplify notation, index i is omit)
$$l(\boldsymbol{\theta}; \boldsymbol{y}_{\mathcal{T}}, \boldsymbol{y}_{\overline{\mathcal{T}}}) = -T \ln \sqrt{2\pi} - n \ln \sigma - \frac{1}{2\sigma^2} \sum_{t \in \mathcal{T}^{(i)}} \left(y_t - \alpha - \beta_1 f_t^{(1)} - ... - \beta_d f_t^{(d)} \right)^2 \tag{1.5}$$
$$- \frac{1}{2\sigma^2} \sum_{t \in \overline{\mathcal{T}}} \left(y_t - \alpha - \beta_1 f_t^{(1)} - ... - \beta_d f_t^{(d)} \right)^2. \tag{1.6}$$

In this method, the filled-in data is given by the optimization. However it can be shown that this method gives the same results with OLS regression without bootstrapping when it is assumed that the data is missing or not is independent with factors. In other cases, the log-likelihood function needs to be revised and EM algorithm can be applied to estimate.

Back to our case, we use benchmark indices as regressors and adopt both simple and multiple regressions to do the backfilling.

1.2.2 Data Cleaning with the simple regression method

Assume we have a benchmark index (for example S&P 500 index) return time series data for some given frequency (e.g. 1 minute); the benchmark index return time series data has no missing values; and one has an asset return time series data having missing values. Let $I = \{1, 2, ...T\}$ be thegiven index set for all time steps.

Let $(X(j))_{j \in I}$ be the time series data for S&P 500 index return, and $(Y(j))_{j \in I}$ be the time series data for the given asset return data having missing values. The index for the missing value is defined by $I_{missing}$ and $I_missing \subset I$. That is to say $Y(k)$'s values when $k \in I_{missing}$ are missing. Let $I_{exist} = I/I_{missing}$. If $I = I_{missing}$, it is impossible to fill up the missing data using the proposed method.

Assume that

$$Y(j) = \alpha + \beta X(j) + \varepsilon_j \tag{1.7}$$

where $j \in I_{exist}$, then by OLS

$$\hat{\beta} = \frac{\Sigma_{j \in I_{exist}} (Y(j) - \bar{Y})(X(j) - \bar{X})}{\Sigma_{j \in I_{exist}} (X(j) - \bar{X})^2} \tag{1.8}$$

$$\hat{\alpha} = \frac{\bar{Y}}{\beta \bar{X}} \tag{1.9}$$

where $\bar{X} = \frac{1}{N} \Sigma_{j \in I_{exist}} X(j)$ and $\bar{Y} = \frac{1}{N} \Sigma_{j \in I_{exist}} Y(j)$, and N is the number of element of I_{exist}.

Fill in the missing data by using:

$$Y(k) = \hat{\alpha} + \hat{\beta} X(k) \tag{1.10}$$

where $k \in I_{missing}$.

1.2.3 Data Cleaning with the multiple regression method

Sometime we need to adopt M different benchmark indices with no missing data to backfill an asset. Then

Let $(X_i(j))_{j \in I}$ be the time series of the i-th market index. Use the following multi-variate regression model:

$$Y(j) = \alpha + \Sigma_0^M \beta_i X_i(j) + \varepsilon_j \tag{1.11}$$

The estimate of coefficients $B = (\alpha, \beta_1, ...\beta_M)'$ are

$$\hat{B} = (X'X)^{-1}X'Y \tag{1.12}$$

where $Y = (Y(j_1), Y(j_2), ...Y(j_N))'$ and

$$X = \begin{pmatrix} 1 & X_1(j_1) & \cdots & X_M(j_1) \\ 1 & X_1(j_2) & \cdots & X_M(j_2) \\ \vdots & \vdots & \ddots & \vdots \\ 1 & X_1(j_N) & \cdots & X_M(j_N) \end{pmatrix}$$

for $I_{exist} = \{j_1, j_2, ...j_N\}$.

Fill the missing data by

$$Y(k) = \hat{\alpha} + \Sigma_0^M \hat{\beta}_i X_i(k) \tag{1.13}$$

where $k \in I_{missing}$.

1.3 FARIMA-FIGARCH Model

The asset returns has been recognized not independent through time, and most returns processes tend to exhibit volatility clustering. Autoregressive conditional heteroskedasticity (ARCH) model is first introduce by Engle(1982), and following it, generalized autoregressive conditional heteroskedasticity (GARCH) and exponential generalized autoregressive conditional heteroscedastic (EGARCH) were proposed by Bollerslev (1986) and Nelson (1991). Meanwhile, more and more empirical studies have noted the extreme degree of persistence of shocks to the conditional variance process(Bollerslev, Chou, and Kroner (1992)).

On the other hand, in the 1990s, many empirical studies document the presence of apparent long-memory in the autocorrelations of squared or absolute returns of various financial asset prices, such as Breidt, and Crato (1994), Dacorogna et al. (1993), Ding, Granger, and Engle (1993), and Harvey (1993).

Motivated by these observations, Bailliea, Bollerslev,and Mikkelsen 1996 introduced the Fractionally Integrated Generalized AutoRegressive Conditionally Heteroskedastic, or FIGARCH, class of processes.

As documented in Bailliea, Bollerslev,and Mikkelsen 1996, the FIGARCH process combines many of the features of the fractionally integrated process for the mean together with the regular GARCH process for the conditional variance. In particular, the FIGARCH model implies a slow hyperbolic rate of decay for the lagged squared innovations in the conditional variance function, although the cumulative impulse response weights associated with the influence of a volatility shock on the optimal forecasts of the future conditional variance eventually tend to zero, which is similar to weakly stationary GARCH processes.

For the estimation, an approximate Maximum Likelihood Estimation (MLE) procedure was first discussed in Bailliea, Bollerslev,and Mikkelsen (1996). Chung (1999) suggested a slightly different version and provide detailed discussions about the estimation.

A stochastic process r_t is said to be the fractionally integrated ARMA (FI-ARMA) process of the orders p and q, or the FIARMA(a, d, m) process, if it is defined by

$$\phi_a(L)(1-L)^d(r_t-\mu)=\theta_m(L)\epsilon_t, \tag{1.14}$$

where L is the backward shift operator defined by $L^jX_t = X_{t-j}$, and ϕ_a, θ_m are two polynomials with degree of a, m respectively, ϵ_t are the innovations with mean 0 and variance σ^2.

In order to see how the FARIMA model can be used to describe long-range dependence. We will start with a simple case that $a = m = 0$.

For a model FARIMA$(0, d, 0)$, it can be written as

$$(1-L)^dX_t=\varepsilon_t.$$

Now we can rewrite the model as a moving average(MA) model with infinite lags as follows,

$$X_t=(1-L)^{-d}\epsilon_t, \tag{1.15}$$

or, equivalently when the following power expansion converges,

$$X_t = \sum_{j=0}^{\infty} b_j \epsilon_{t-j}, \tag{1.16}$$

where the coefficients $\{b_j\}$ are given by, $b_0 = 1$, and

$$b_j = \frac{\Gamma(j+d)}{\Gamma(j+1)\Gamma(d)}, \quad j \in \mathbb{N} \tag{1.17}$$

with $\Gamma(p) = \int_0^\infty t^{p-1} e^{-t} dt$ is the gamma function. It follows that when $-\infty < d < 1/2$, the series $\sum_{j=0}^{\infty} b_j \epsilon_{t-j}$ converges in $L^2(\Omega)$. Then we can safely say the MA(∞) representation satisfies FARIMA$(0, d, 0)$.

Here furthermore, we define the basic range for d to be $-1/2 < d < 1/2$,[1] since we can always simply put the integer part of d into the $\phi(L)$ part. Now, we can define the fractional integrated part $(1-L)^d$ by means of expansion,

$$(1-L)^d = \sum_{j=0}^{\infty} \pi_j L^j, \tag{1.18}$$

where $\pi_0 = 1$ and

$$\pi_j = \prod_{k=1}^{j} \frac{k-1+d}{k} = \frac{\Gamma(j-d)}{\Gamma(j+1)\Gamma(-d)}, \quad j \in \mathbb{N}. \tag{1.19}$$

The following proposition shows the relation between FARIMA models and long-range dependence.

Proposition 1 (Brockwell and Davis 1991, pp. 522-523)**.** *If $X = (X_t)_{t\in\mathbb{Z}}$ is* FARIMA$(0, d, 0)$ *with $-1/2 < d/1/2$, $d \neq 0$, then the auto-covariance function $\gamma(k) = \mathbb{E}X_0X_k$ is given as*

$$\gamma(0) = \sigma^2 \frac{\Gamma(1-2d)}{\Gamma^2(1-d)}, \quad \text{and} \tag{1.20}$$

$$\gamma(k) = \sigma^2 \frac{(-1)^k \Gamma(1-2d)}{\Gamma(k-d+1)\Gamma(1-k-d)} \sim c|k|^{2d-1} \tag{1.21}$$

[1] $d = -1/2$ is not included due to the consideration of invertability.

as $k \to \infty$, where

$$c = \sigma^2 \frac{\Gamma(1-2d)}{\Gamma(d)\Gamma(1-d)}. \tag{1.22}$$

Thus, if and only if $0 < d < 1/2$, FARIMA$(0, d, 0)$ has long-range dependence.

Formally, consider the FARIMA model on price returns as in 1.23, we can define r_t as

$$r_t = \mu + \phi_p^{-1}(L)(1-L)^{-d}\theta_m(L)\epsilon_t, , \tag{1.23}$$

if $d < 1/2$ and $\phi_a(z)$ has no roots inside unit circle, the process r_t is well defined, causal and stationary. In addition, if $d > -1/2$ and roots of $Phi_p(z)$ lie outside unit circle, it is invertible.

The long-range dependence is only is not existing in return process itself, but also in the volatility process. Therefore, the same fractional integration is also introduced to model stochastic volatility. A fractional integrated GARCH model (FIGARCH) is expressed as in Chung (1999),

$$\psi_p(L)(1-L)^d(\epsilon_t^2 - \sigma^2) = (1 - \beta(L))v_t, \tag{1.24}$$

$$\epsilon_t = \sqrt{h_t}u_t, \tag{1.25}$$

$$v_t = \epsilon_t^2 - h_t, \tag{1.26}$$

where $u_t, t \in \mathbb{Z}$ are i.i.d. standardized residuals with mean 0 and unit variance. When $0 < d_0 < 1$, FIGARCH model exhibits long-range dependence in variance.

In this work, we will focus only on FIGARCH models with low degrees, that is, $p \leq 1, q \leq 1$. To ensure the positivity of the conditional variances $h_t, t \in \mathbb{Z}$, the following conditions for FIGARCH$(1, d, 1)$ are imposed as in Baillie et al. (1996)

$$\beta - d \leq \psi \leq \frac{1}{3}(2-d), \tag{1.27}$$

$$d(\psi - \frac{1-d}{2}) \leq \beta(d - \beta + \psi). \tag{1.28}$$

In Chung (2001), another condition is use

$$0 \leq \psi \leq \beta \leq d_0 \leq 1. \tag{1.29}$$

A general condition for non-negative conditional variance as given by Conrad and Haag (2006) is shown in the following proposition.

Proposition 2 (Conrad and Haag (2006))**.** *Let the coefficients g_j and f_i be the functions of the fractional differencing parameter d such that $g_j = f_j g_{j-1} = \prod_{i=1}^{j} f_i$, with $f_j = \frac{j-1-d}{j}$ for $j \in \mathbb{N}$ and $g_0 = 1$. Then the conditions are*

$$\kappa_1 = d + \psi_1 - \beta_1, \text{ and } \kappa_i = \beta_1 \kappa_{i-1} + (f_i - \psi_1)(-g_{i-1}), \text{ for } i \geq 2, \tag{1.30}$$

or, alternatively,

$$\kappa_i = \beta_1^2 \kappa_{i-2} + [\beta_1(f_{i-1}\psi_1) + (f_i - \psi_1) f_{i-1}](-g_{i-2}) \text{ for all } i \geq 3. \tag{1.31}$$

Here the FARIMA-FIGARCH structure is summarized as follows

$$\phi_a(L)(1-L)^{d_0}(r_t - \mu) = \theta_m(L)\epsilon_t, \tag{1.32}$$

$$\psi_p(L)(1-L)^d(\epsilon_t^2 - \sigma^2) = (1 - \beta(L))\upsilon_t, \tag{1.33}$$

$$\epsilon_t = \sqrt{h_t} u_t, \tag{1.34}$$

$$\upsilon_t = \epsilon_t^2 - h_t, \tag{1.35}$$

where

$$\phi_a(L) = 1 - \sum_{j=1}^{a} \phi_j L_j, \tag{1.36}$$

$$\theta_m(L) = 1 + \sum_{j=1}^{m} \theta_j L_j, \tag{1.37}$$

$$\psi_p(L) = 1 - \sum_{j=1}^{p} \psi_j L_j, \tag{1.38}$$

$$\beta_q(L) = \sum_{j=1}^{q} \beta_j L_j, \tag{1.39}$$

and μ is the unconditional mean of r_t, σ^2 is the unconditional variance of ϵ_t, ϵ_t are the innovations and u_t are the standardized residuals with mean 0 and unit variance. The stationary conditions is given as

1. $-\frac{1}{2} < d_0 < \frac{1}{2}$, and roots of $\phi_a(z)$ lie outside unit circle.

2. $0 \leq d < 1$, roots of $\psi_p(z)$ lie outside unit circle.

3. for $p = q = 1$, $0 \leq \psi_1 \leq \beta_1 \leq d \leq 1$.

1.4 Generalized Hyperbolic Innovations

In this section, the generalized hyperbolic (GH) distributions and their characteristics will be summarized as GH distribution will be used to model the innovation distribution.

As in many other literature, the non-Gaussian distributions with fat-tails and excess kurtosis and skewness are haevily used to model financial return. Here we will generalized hyperbolic distribution family as it allows heavy tails and asymmetry. Also GH distributions are infinitely divisible, which is a high desired property since the standardized residuals u_t from FARIMA-FIGARCH model is interpreted as a aggregation of external stationary effects that might be seen in an more frequent sampling or even in continuous time.

A random vector X is said to have a multivariate GH distribution if X can be expressed as a normal mean-variance mixture distribution

$$X \stackrel{d}{=} \mu + W\gamma + \sqrt{W}AZ, \tag{1.40}$$

where $Z \sim \mathcal{N}_k(0, I_k)$ is standard k-dimensional normal distributed random vector, A is a $d \times k$ real matrix, $\mu, \gamma \in \mathbb{R}^d$, and $W \geq 0$ is a scalar-valued random variable independent of Z and having a Generalized Inverse Gaussian distribution $\mathrm{GIG}(\lambda, \chi, \psi)$.

A random variable W is said to have a generalized inverse Gaussian (GIG) distribution if its probability density is given by

$$f_{\mathrm{GIG}}(x; \lambda, \chi, \psi) = \frac{\chi^{-\lambda}(\sqrt{\chi\psi})^{\lambda}}{2K_\lambda(\sqrt{\chi\psi})} x^{\lambda-1} \exp\left(-\frac{1}{2}\left(\chi x^{-1} + \psi x\right)\right), \tag{1.41}$$

for $x > 0$, and where $\chi, \psi > 0$, and K_λ is a modified Bessel function of the third kind with index λ. The parameters satisfy $\chi > 0, \psi \geq 0$ if $\lambda < 0$, $\chi \geq 0, \psi > 0$ if $\lambda > 0$, and $\chi > 0, \psi > 0$ if $\lambda = 0$.

By letting $\Sigma = AA'$, we denote a GH distributed random vector by

$$X \sim \mathrm{GH}_d(\lambda, \chi, \psi, \mu, \gamma, \Sigma). \tag{1.42}$$

The univariate GH distribution is obtain with $d = 1$. The joint density when Σ is not singular is given by, for $x > 0$,

$$f_{\mathrm{GH}}(x) = c \frac{K_{\lambda - \frac{d}{2}}\left(\sqrt{(\chi + (x-\mu)'\Sigma^{-1}(x-\mu))(\psi + \gamma'\Sigma^{-1}\gamma)}\right) e^{(x-\mu)'\Sigma^{-1}\gamma}}{\left(\sqrt{(\chi + (x-\mu)'\Sigma^{-1}(x-\mu))(\psi + \gamma'\Sigma^{-1}\gamma)}\right)^{\frac{d}{2}-\lambda}}, \tag{1.43}$$

where the normalizing constant is given by

$$c = \frac{\left(\sqrt{\chi\psi}\right)^{-\lambda} \left(\psi + \gamma'\Sigma^{-1}\gamma\right)^{(d/2)-\lambda}}{(2\pi)^{d/2}|\Sigma|^{1/2} K_\lambda\left(\sqrt{\chi\psi}\right)}. \tag{1.44}$$

One advantage of GH distribution family is that they are closed under linear transformation.

Theorem 1 (Linear transformation of GH). *If $X \sim GH_d(\lambda, \chi, \psi, \mu, \gamma, \Sigma)$, and*

$$Y = BX + b$$

with $B \in \mathbb{R}^{k\times d}$ and $b \in \mathbb{R}^k$. Then

$$Y \sim GH_d(\lambda, \chi, \psi, B\mu + b, B\gamma, B\Sigma B').$$

Many common-used distributions can be obtained as special cases or limiting cases of GH. Here we describe the skewed Student's t-distribution as the special case of GH distribution, with $\psi = 0$, $\lambda = -\nu/2$, $\chi = \nu$. The skewed t has the following density

$$f_{\text{skewed-t}}(x) = c \frac{K_{\frac{\nu+1}{2}}\left(\sqrt{(\nu + \sigma^{-1}(x-\mu)^2)(\sigma^{-1}\gamma^2)}\right) e^{(x-\mu)\sigma^{-1}\gamma}}{\left(\sqrt{(\nu + \sigma^{-1}(x-\mu)^2)(\sigma^{-1}\gamma^2)}\right)^{-\frac{\nu+1}{2}} \left(1 + \frac{\sigma^{-1}(x-\mu)^2}{\nu}\right)^{\frac{\nu+1}{2}}}, \tag{1.45}$$

where the normalizing constant is

$$c = \frac{2^{\frac{2-(\nu+1)}{2}}}{\Gamma(\nu/2)\sqrt{(\pi\nu)\sigma}}. \tag{1.46}$$

Generalized hyperbolic distribution can capture the heavy-tailedness. The tail behavior of GH is given by

$$f_{\mathrm{GH}}(x;\lambda,\chi,\psi,\mu,\sigma,\gamma) \sim |x|^{\lambda-1}e^{(\mp\alpha+\beta)x}, \quad x \to \pm\infty, \tag{1.47}$$

where $\alpha = \sqrt{\chi\psi}$ and $\beta = \sigma^{-1}\gamma$.

1.5 Estimation

1.5.1 Fraction Differencing Parameter

Considering a FARIMA$(0, d, 0)$ model, the fractional differencing parameter d has a close relation with the Hurst index (self-similarity) H. As in Taqqu (2002), it can be illustrated by the following property. If X is a Gaussian FARIMA(0,d,0) time series with $d \in (0, 1/2)$, then $n \to \infty$,

$$\frac{1}{n^H}\sum_{t=1}^{[ns]} X_t \to B_H(s), \tag{1.48}$$

where $H = d + 1/2$, $\{B_H(s), s \in \mathbb{R}\}$ is an fractional Brownian motion. The auto-covariance function in this case then follows

$$\gamma(n) \triangleq \mathrm{Cov}(X(0), X(n)) \sim cn^{2d-1}, \quad n \to \infty, \tag{1.49}$$

or the spectral density function

$$g(\lambda) \sim C\lambda^{-2d}, \quad \lambda \to 0+, \tag{1.50}$$

where

$$\gamma(n) = \int_{-\pi}^{\pi} g(\lambda)\cos n\lambda d\lambda, n \in \mathbb{Z}. \tag{1.51}$$

For a general FARIMA(a, d, m) model as in (1.23), the spectral density is given by, for $-\pi < \lambda \le \pi$,

$$g(\lambda) = \frac{\sigma^2}{2\pi} \left|1 - e^{i\lambda}\right|^{-2d} \frac{|\phi(e^{i\lambda})|^2}{|\theta(e^{i\lambda})|^2}. \tag{1.52}$$

For Gaussian $X = (X_n)$, an estimation can be considered to maximize the criterion (proportional to the log likelihood after omitting a constant),

$$-\frac{1}{2N} \log |\Sigma(\eta)| - \frac{1}{2N} X'\Sigma(\eta)^{-1} X, \tag{1.53}$$

where η denotes the parameter vector $\eta = (\sigma^2, d, \phi_1, .., \phi_a, \theta_1, .., \theta_m)$, $\Sigma(\eta)$ is the matrix with (i, j) entry being $\gamma(i - j; \eta)$, and $X = (X_1, ..., X_N)'$. Note that the criterion above can be approximated by

$$-\frac{1}{2\pi} \int_{-\pi}^{\pi} \log g(\lambda; \eta) d\lambda - \frac{1}{4\pi} \int_{-\pi}^{\pi} \frac{I(\lambda)}{g(\lambda; \eta)} d\lambda, \tag{1.54}$$

where $I(\lambda)$ is the periodogram

$$I(\lambda) = \frac{1}{2\pi N} \left| \sum_{n=1}^{N} X_n e^{in\lambda} \right|^2, \tag{1.55}$$

and $g(\lambda; \eta^*) > 0$ is assumed for all λ where η^* denote the closed-form solution.

Furthermore, the integrals are approximated by a discrete sum. Thus finally the following approximation is considered,

$$-\frac{1}{2\pi} \sum_{j=1}^{N-1} \left[\log g(\lambda_j; \eta) + \frac{I(\lambda_j)}{g(\lambda_j; \eta)} \right], \tag{1.56}$$

where $\lambda_j = \frac{2\pi j}{N}$. The estimates maximize the approximated criterion (1.56) is called Whittle estimates.

For a simple model as FARIMA$(0, d, 0)$, the following estimated in a closed-form are proposed by Kashyap and Eom (1988)

$$\hat{d} = -\frac{1}{2} \frac{\sum_{j=1}^{N-1} \log |1 - e^{i\lambda}| \log I(\lambda_j)}{\sum_{j=1}^{N-1} (\log |1 - e^{\imath\lambda_j}|)^2}. \tag{1.57}$$

Another local estimator introduced by Geweke and Porter-Hudak (GPH) (1983) is also considered to estimate d. Let us separate the fractional part and ARMA part, and rewrite the FIARMA model as follows,

$$(1-L)^d X_t = w_t \tag{1.58}$$
$$\phi(L) w_t = \theta(L)\epsilon_t. \tag{1.59}$$

Then the spectral density satisfies

$$g_X(\lambda) = |1-e^{i\lambda}|^{-2d} g_w(\lambda). \tag{1.60}$$

Then the pooled periodogram follows the linear form

$$\log I_X(\lambda) = \log g_w(0) - 2d \log \left|2\sin\frac{\lambda}{2}\right| + \log\frac{g_w(\lambda)}{g_w(0)} + \log\frac{I_X(\lambda)}{g_X(0)} \tag{1.61}$$

Given a bandwidth parameter M, the local log-periodogram regression estimator is defined as the minimum of the local least-squares criterion,

$$\hat{d} = \arg\min \sum_{j=1}^{M} \left(\log I(\lambda_j) - 2d\log\left|2\sin\frac{\lambda_j}{2}\right|\right)^2, \tag{1.62}$$

where $\lambda_j = \frac{2\pi j}{N}, j = 1, ..., M$. Thus solve for $\hat{d}$ gives

$$\hat{d}^{\text{GPH}} = -\frac{1}{2}\frac{\sum_{j=1}^{M}(g_j-\bar{g})\log I_j}{\sum_{j=1}^{M}(g_j-\bar{g})^2}, \tag{1.63}$$

where $g_j = \log\left|2\sin\frac{\lambda_j}{2}\right|$, $\bar{g} = M^{-1}\sum_{j=1}^{M} g_j$, and $I_j = I_X(\lambda_j)$. The bandwidth parameter is chosen to be $M = N^{0.65}$.

1.5.2 FARIMA-FIGARCH with Gaussian Innovations

In this section, we consider the FARIMA-FIGARCH model

$$\phi_a(L)(1-L)^{d_0}(r_t-\mu) = \theta_m(L)\epsilon_t, \tag{1.64}$$
$$\psi_p(L)(1-L)^d(\epsilon_t^2-\sigma^2) = (1-\beta(L))v_t, \tag{1.65}$$

$$\epsilon_t = \sqrt{h_t} u_t, \tag{1.66}$$

$$v_t = \epsilon_t^2 - h_t, \tag{1.67}$$

with (u_t) i.i.d standard normal distributed random variables, thus ϵ_t follows $\mathcal{N}(0, h_t)$. Then, as suggested in Chung (1999), the approximated log-likelihood under Gaussian assumption can be written as

$$\log L(\eta; r_{1:T}) = -\frac{T}{2}\log(2\pi) - \frac{1}{2}\sum_{t=1}^{T}\log h_t - \frac{1}{2}\sum_{t=1}^{T}\frac{\epsilon_t^2}{h_t}, \tag{1.68}$$

where parameter vector $\eta = (\mu, d_0, \phi_1, ..., \phi_a, \theta_1, ..., \theta_m, \sigma^2, d, \psi_1, ..., \psi_p, \beta_1, ..., \beta_q)$, $r_{1:T} = (r_1, ..., r_T)$, and (h_t) and (ϵ_t) are the filtered conditional variances and innovations based on the given η. The filtering procedure for (h_t) and (ϵ_t) are simply as follows

$$\epsilon_t = \frac{\phi_a(L)}{\theta_m(L)}(1-L)^{d_0}(r_t - \mu), \tag{1.69}$$

$$v_t = \frac{\psi_p(L)}{1-\beta(L)}(1-L)^{d}(\epsilon_t^2 - \sigma^2), \text{ and,} \tag{1.70}$$

$$h_t = \epsilon_t^2 - v_t. \tag{1.71}$$

In order to get better local maximum for (1.68), the initial point for η is produced by separating fractional integrated part and ARMA part as following,

1. Estimate d_0 based on (r_t) by use of GPH estimator.

2. Filter $w_t = (1-L)^{d_0}(r_t - \mu)$ with μ is the sample mean of (r_t).

3. Estimate (w_t) as ARMA(a, m) model.

4. Filter $\epsilon_t = \frac{\phi_a(L)}{\theta_m(L)} w_t$.

5. Estimate d based on $\epsilon_t^2 - \sigma^2$ with σ^2 is the sample variance of (ϵ_t).

6. Estimate ψ, β by maximizing (1.68) with the given d as estimated from previous step.

1.5.3 GH Innovations

The estimation procedure of generalized hyperbolic innovations is adopted from McNeil, Frey and Embrects (2005). It is a modified EM algorithm so-called MCEMC (multi-cycle expectation conditional estimation), and given as following.

Let $\theta^{[k]}$ denotes the parameter vector at k-th iteration step.

1. Set iteration $k = 1$ and initialize the stating value of $\theta^{[1]}$. For example, set μ, Σ to be sample mean and sample covariance matrix S respectively, and $\gamma = 0$.

2. Calculate the weights $\delta_i^{[k]}$ and $\eta_i^{[k]}$ by

$$\delta_i^{[\cdot]} = \mathbb{E}(W_i^{-1}|X_i; \theta^{[\cdot]}), \tag{1.72}$$
$$\eta_i^{[\cdot]} = \mathbb{E}(W_i|X_i; \theta^{[\cdot]}), \tag{1.73}$$

   where

$$W_i|X_i, \theta \sim \text{GIG}\left(\lambda - d/2, (X_i - \mu)'\Sigma^{-1}(X_i - \mu) + \chi, \psi + \gamma'\Sigma\gamma\right) \tag{1.74}$$

   are the latent subodinator.

3. Computer the average of $\delta_i^{[k]}$ and $\eta_i^{[k]}$

$$\bar{\delta}^{[k]} = \frac{1}{N}\sum_{i=1}^{N} \delta_i^{[k]} \text{ and } \bar{\eta}^{[k]} = \frac{1}{N}\sum_{i=1}^{N} \eta_i^{[k]}. \tag{1.75}$$

4. For a symmetric model set $\gamma^{[k+1]} = 0$. Otherwise set

$$\gamma^{[k+1]} = \frac{1}{N}\frac{\sum_{i=1}^{N} \delta_i^{[k]}(\bar{X} - X_i)}{\bar{\delta}^{[k]}\bar{\eta}^{[k]} - 1}. \tag{1.76}$$

5. Updates the location and dispersion by

$$\mu^{[k+1]} = \frac{1}{N}\frac{\sum_{i=1}^{N} \delta_i^{[k]} X_i - \gamma^{[k+1]}}{\bar{\delta}^{[k]}} \tag{1.77}$$

$$\Sigma^{[k+1]} = \frac{|S|^{1/d}\Psi}{|\Psi|^{1/d}}, \tag{1.78}$$

where

$$\Psi = \frac{1}{N}\sum_{i=1}^{N} \delta_i^{[k]} \left(X_i - \mu^{[k+1]}\right)\left(X_i - \mu^{[k+1]}\right)' - \bar{\eta}^{[k]}\gamma^{[k+1]}(\gamma^{[k+1]})' \tag{1.79}$$

6. Set

$$\theta^{[k,2]} = (\lambda^{[k]}, \chi^{[k]}, \psi^{[k]}, \mu^{[k+1]}, \Sigma^{[k+1]}, \gamma^{[k+1]})'. \tag{1.80}$$

   Calculate $\delta_i^{[k,2]}$, $\eta_i^{[k,2]}$ and $\xi^{[k,2]}$ by (1.72)(1.73) and

$$\xi^{[\cdot]} = \mathbb{E}\left(\log(W_i)|X_i;\theta^{[\cdot]}\right). \tag{1.81}$$

7. Maximize the following function $Q(\lambda, \chi, \psi; \theta^{[k,2]})$ to obtain $\lambda^{[k+1]}, \chi^{[k+1]}, \psi^{[k+1]}$,

$$\begin{aligned} Q(\lambda, \chi, \psi; \theta^{[\cdot]}) =& (\lambda - 1)\sum_{i=1}^{N}\xi^{[\cdot]} - \frac{1}{2}\chi\sum_{i=1}^{N}\delta^{[\cdot]} - \frac{1}{2}\psi\sum_{i=1}^{N}\eta^{[\cdot]} \\ &- \frac{1}{2}N\lambda\log(\chi) + \frac{1}{N}\lambda\log(\psi) - N\log(2K_\lambda(\sqrt{\chi\psi})) \end{aligned} \tag{1.82}$$

8. Increment $k \to k+1$ and go to step 2. Repeat untill convergence.

1.6 Risk Assessment

1.6.1 Dependency Structure

Once the marginal stock process is calibrated in FARIMA-FGARCH model, the filtered innovations of multiple assets then can be coupled by use of the idea of copula. A d-dimensional copula C is a d-dimentional cumulative function on $[0,1]^d$ with standard uniform margins. Sklar's Theorem stats that every multivariate cumulative function F with continuous marginals $F_1, ..., F_d$ can be written as

$$F(x_1, ..., x_d) = C(F_1(x_1), ..., F_d(x_d)), \tag{1.83}$$

for some unique copula function C. It can be seen also in the form

$$C(u_1, ..., u_d) = F(F_1^{-1}(u_1), ..., F_d^{-1}(u_d)), \tag{1.84}$$

where $F_i^{-1}, i = 1, ..., d$ are the inverse (quantile) function of F_i since F_i is continuous thus well-defined.

The process considered here can be obtained by an affine transform of a process with identity dispersion matrix $\Sigma = I_d$ and null location vector $\mu = 0$. The transformed process then can be written in spherical coordinated. In general, the transformed process is not spherically symmetric. The d-dimensional generalized hyperbolic random vector $X = (X - 1, ..., X_d)$ is an example. More generally, let us consider a normal mean-variance mixture in stochastic representation

$$X = m(W) + \sqrt{W} AZ, \tag{1.85}$$

with a measurable function $m : [0, \infty) \to \mathbb{R}^d$ here specified as

$$m(W) = \mu + W\gamma, \tag{1.86}$$

where μ is the location vector, γ is skewness vector, $Z \sim \mathcal{N}_d(0, I_d)$ thus $AZ \sim \mathcal{N}(0, \Sigma)$ the dispersion matrix $\Sigma = AA'$, and W is a positive random variable independent of Z. The transformation to spherical coordinates will simply gives identity dispersion matrix and it will transform W. The independent random variable W, after transformation, introduces non-spherically symmetric parameters into Z.

If the non-spherically symmetric directions for a multivariate random process are confined to a finite dimensional subspace, the process is nearly spherical. By the same terminology, before application of affine transformation, yields the definitions of nearly elliptical if there are a finite number of non-spherical dimensions, and elliptical if all dimensions are spherical after transform.

There several explicit and useful examples of multivariate nearly elliptical distributions. We can estimate the parameters of these distribution, that is, of the copulas.

In the follows, we will look at two copulas: Gaussian and generalized hyperbolic.

Gaussian Copula

The Gaussian copula is obtained from multivariate normal distribution,

$$C_{\Sigma}(u) = N_{\Sigma}(\Phi^{-1}(u_1), ..., \Phi^{-1}(u_d)), \tag{1.87}$$

where Σ is a positive definite correlation matrix, and Φ^{-1} is the inverse distribution function of univariate standard normal distribution. One of the major drawback of this copula is the lack of lower tail dependence.

A multivariate random vector with a standard normal marginal and a Gaussian copula of correlation matrix Σ is Gaussian $\mathcal{N}(0, \Sigma)$. Therefore sampling from the Gaussian copula is a simple process,

1. Compute the Cholesky decomposition A of Σ, i.e., $AA' = \Sigma$.
2. Generate d-dimensional standard normal vector $Z = (Z_1, ..., Z_d)'$ from $\mathcal{N}(0, I_d)$.
3. Compute the Gaussian sample by $X = AZ$.
4. Generate the copula sample by $u_i = \Phi(X_i)$, for $i = 1, ..., d$.

Once the copula sample $U = (u_1, ..., u_d)$ is obtained, each marginal sample can be computed by $F_i^{-1}(u_i)$ for $i = 1, ..., d$.

When the marginal distribution is also Gaussian, the joint distribution will be Gaussian as well. In this case, the VaR can be computed directly. But when the marginal innovation is GH, it can not be computed directly. The simulation-based computation will be applied.

Multivariate Generalized Hyperbolic Copula

From the normal mean-variance mixture construction (1.85) with the subordinator W having the generalized inverse Gaussian (GIG) distribution, X has a multivariate generalized hyperbolic (MGH) distribution. The joint density function is given by

$$f_{\mathrm{GH}}(x) = c\frac{K_{\lambda-\frac{d}{2}}\left(\sqrt{(\chi+(x-\mu)'\Sigma^{-1}(x-\mu))(\psi+\gamma'\Sigma^{-1}\gamma)}\right)e^{(x-\mu)'\Sigma^{-1}\gamma}}{\left(\sqrt{(\chi+(x-\mu)'\Sigma^{-1}(x-\mu))(\psi+\gamma'\Sigma^{-1}\gamma)}\right)^{\frac{d}{2}-\lambda}}, \tag{1.88}$$

where the normalizing constant is given by

$$c = \frac{\left(\sqrt{\chi\psi}\right)^{-\lambda}(\psi+\gamma'\Sigma^{-1}\gamma)^{(d/2)-\lambda}}{(2\pi)^{d/2}|\Sigma|^{1/2}K_\lambda\left(\sqrt{\chi\psi}\right)}. \tag{1.89}$$

When assuming the marginals shares the same subordinator W, the they are joined with the MGH copula, then the innovations follows the multivariate generalized hyperbolic distribution. A MGH scenario can be generated by use of the mean-variance construction,

1. Draw a sample Z from d-dimensional multivariate normal distribution $\mathcal{N}(0,\Sigma)$.

2. Draw a sample W from the generalized inverse Gaussian distribution $\mathrm{GIG}(\lambda,\chi,\psi)$,

3. Generate the sample X from the mean-variance mixture construction (1.85).

1.6.2 Value at Risk

Value at Risk (VaR) is used heavily to measure the risk. Recall that VaR with confidence level α is defined as the smallest number such that the probability that the loss exceed VaR is no larger than $(1-\alpha)$. Formally, given the return X,

$$\mathrm{VaR}_\alpha(X) = \inf(l\in\mathbb{R}:\mathbb{P}(-X>l)\le 1-\alpha) = \inf(l:F(-l)\le 1-\alpha). \tag{1.90}$$

Financial log returns are express as a multivariate probability distribution function. From the FARIMA-FIGARCH model, the joint distribution is transformed to a standard form. It is written as a FARIMA-FIGARCH prediction plus an innovation. For the purpose of VaR, we do not need the full high dimensional stochastic returns. We can reduce the VaR evaluation to a low dimensional space, and operates with this space of the low dimensional marginals, which express the probability of returns within this space only. Recall that a elliptical or nearly elliptical distribution can be reduced to spherical or nearly spherical by an affine transformation. The MGH distribution, considered in details here, is nearly elliptical with an exceptional dimension for the skewness vector. We work in the low dimensional space defined by the (spherically transformed) portfolio and by the exceptional dimensions. Thus for MGH, we have two essential dimensions for the marginal used to evaluate VaR.

1.6.3 Direct Computation From Density

Suppose the multivariate random vector X follows $\mathrm{GH}_d(\lambda, \chi, \psi, \mu, \Sigma, \gamma)$, recall that X can be written as

$$X \overset{d}{=} \mu + \gamma W + \sqrt{(W)} Z, \tag{1.91}$$

where $W \sim \mathrm{GIG}(\lambda, \chi, \psi)$ and $Z \sim \mathcal{N}(0, \Sigma)$ are independent.

Consider X_t is the log return at time t and generated by the FARIMA-FIGARCH filter. Thus, we can write

$$X_t = \mu_{t-1} + D_{t-1} U_t, \tag{1.92}$$

where μ_{t-1} is the conditional mean given by FARIMA, D_{t-1} is the conditional std given by FIGARCH, and U_t is standardized residual follows $U_t \sim \mathrm{GH}(\lambda, \chi, \psi, \mu_U, \Sigma, \gamma)$ with

$$\mu_U = -\gamma \sqrt{\frac{\chi}{\psi}} \frac{K_{\lambda+1}(\sqrt{\chi\psi})}{K_\lambda(\sqrt{\chi\psi})}. \tag{1.93}$$

It ensures the mean of U_t is 0. Then we have

$$X_t \sim \mathrm{GH}(\lambda, \chi, \psi, \mu_{t-1} + D_{t-1}\mu_U, D_{t-1}\Sigma D'_{t-1}, D_{t-1}\gamma). \tag{1.94}$$

Furthermore, the portfolio return with weight w, $Y_t = \langle w, X_t \rangle$ follows univariate GH distribution,

$$Y_t \sim \mathrm{GH}(\lambda, \chi, \psi, \mu_Y, \sigma_Y^2, \gamma_Y), \tag{1.95}$$

where

$$\mu_Y = w' u_{u-1} + w' D_{t-1} \mu_U = w' u_{u-1} - w' D_{t-1} \gamma \sqrt{\frac{\chi}{\psi}} \frac{K_{\lambda+1}(\sqrt{\chi\psi})}{K_\lambda(\sqrt{\chi\psi})}, \tag{1.96}$$

$$\sigma_Y^2 = w' D_{t-1} \Sigma D_{t-1}' w, \text{ and,} \tag{1.97}$$

$$\gamma_Y = w' D_{t-1} \gamma \tag{1.98}$$

The VaR at confidence level α is the solution of the equation

$$\alpha = \mathbb{P}(Y_t < -\mathrm{VaR}_\alpha) = \int_{-\infty}^{-\mathrm{VaR}_\alpha} f_{Y_t}(y) dy, \tag{1.99}$$

where f_{Y_t} is the probability density function of Y_t.

In order to numerically solve it for VaR_α, we can approxiate the integral by

$$\int_{-\infty}^{-\mathrm{VaR}_\alpha} f_{Y_t}(y) dy \approx \int_{\underline{y}}^{-\mathrm{VaR}_\alpha} f_{Y_t}(y) dy \tag{1.100}$$

$$\approx \sum_{j=1}^{N} f_{Y_t}(y_j) \Delta y, \tag{1.101}$$

where $\underline{y}$ is set to $\underline{y} = -5 \cdot \mathrm{std}(Y_t)$, and $y_j = \underline{y} + j\Delta y$, $j = 1, .., N$. the standard deviation of Y_t is given by

$$\left\{ \sqrt{\frac{\chi}{\psi}} \frac{K_{\lambda+1}(\sqrt{\chi\psi})}{K_\lambda(\sqrt{\chi\psi})} \sigma_Y^2 + \left(\frac{\chi}{\psi} \frac{K_{\lambda+2}(\sqrt{\chi\psi})}{K_\lambda(\sqrt{\chi\psi})} - \left(\sqrt{\frac{\chi}{\psi}} \frac{K_{\lambda+1}(\sqrt{\chi\psi})}{K_\lambda(\sqrt{\chi\psi})} \right)^2 \right) \gamma_Y^2 \right\}^{\frac{1}{2}}. \tag{1.102}$$

For the case that the characteristic function is known in closed-form rather that density function, the process can be done similarly, but depends on an FFT to generate density function form the characteristic function.

1.6.4 Monte Carlo Computation

Suppose there are N i.i.d d-dimensional multivariate GH random vectors $X^{(1)}, ..., X^{(N)}$ generated by the algorithm in section 1.6.1. Then we will have N scenarios of our portfolio returns

$$Y^{(N)} = \langle w, X^{(k)} \rangle, \quad k = 1, .., N \tag{1.103}$$

where $\langle \rangle$ denotes the the inner product i.e. $\langle w, x \rangle = w'x$. Then the empirical distribution of the portfolio return Y is given by

$$\widehat{F}(y) = \frac{1}{N} \sum_{i=1}^{N} \mathbb{1}_{\{Y^{(i)} < y\}}, \tag{1.104}$$

where $\mathbb{1}$ is the indicator function. The VaR with confidence level α is the solution of

$$\widehat{F}(-\mathrm{VaR}_\alpha(Y)) - \alpha = 0 \tag{1.105}$$

Thus the above equation can solved by use of the order statistics,

1. Sort the random numbers $Y^{(1)}, ..., Y^{(N)}$ and get $Y^{(s_1)} \leq Y^{(s_2)} \leq ... \leq Y^{(s_N)}$.

2. Pick $k = [\alpha N]$, and $\mathrm{VaR}_\alpha(Y) \approx Y^{(s_k)}$.

The sorting algorithm usually take $O(N \log N)$.

One can also smooth the empirical distribution function by for example the Gaussian kernel as follows,

$$\widehat{F}(y) = \frac{1}{N} \sum_{i=1}^{N} \Phi \left(\frac{y - Y^{(i)}}{\sigma_w} \right), \tag{1.106}$$

where Φ is the CDF of standard normal distribution and $\sigma_w^2 = w' \Sigma_X w$ with Σ_X the covariance matrix of the MGH random vector X. By solving (1.105), it gives a smooth approximation of VaR.

1.6.5 Computation by Affine Transformation

For general nealy elliptical distribution, we can calculation VaR efficiently by use of affine transformation to reduce the dimension. Suppose that the density function is of the form

$$f(x) = g((x-\mu)'\Sigma^{-1}(x-\mu), (x-\mu)'\gamma). \tag{1.107}$$

The goal is to compute the probability distribution of the portfolio return $Y = \langle w, x\rangle$,

$$F(y) = \int_{\langle w,x\rangle \le y} f(x)dx = \int_{\langle w,x\rangle \le y} g((x-\mu)'\Sigma^{-1}(x-\mu), (x-\mu)'\gamma)dx. \tag{1.108}$$

First we change the parameters into the spherical coordinates $z = A^{-1}(x-\mu)$ where A is the Cholesky decompostion of Σ. Then we obtain

$$F(y) = \int_{\langle w_s,z\rangle \le y-\langle w,\mu\rangle} c_1 g\left(\|z\|^2, \langle z, \gamma_s\rangle\right) dz, \tag{1.109}$$

where $w_s = A'w$, $\gamma_s = A'\gamma$ and c_1 is the constant given by the determinant of the Jacobian of the transformation. It is hard to calculate c_1 directly, but we know the fact that

$$\int c_1 g\left(\|z\|^2, \langle z, \gamma_s\rangle\right) dz = 1, \tag{1.110}$$

thus we can set

$$c_1 = \left(\int g\left(\|z\|^2, \langle z, \gamma_s\rangle\right) dz\right)^{-1} \tag{1.111}$$

Without loss of generality, we replace the dummy variable z by x,

$$F(y) = \int_{\langle w_s,x\rangle \le y-\langle w,\mu\rangle} c_1 g\left(\|x\|^2, \langle x, \gamma_s\rangle\right) dx. \tag{1.112}$$

Next we perform the orthogonal transformation $z = Ux$, where U is a unitary matrix with the first row given by $\frac{w_s'}{\|w_s\|}$. Therefore, we have

$$z_1 = \frac{\langle w_s, x\rangle}{\|w_s\|}, \text{ and } \|z\| = \|x\|, \tag{1.113}$$

since the unitary transformation does not change the length. Then we obtain

$$F(y) = \int_{z_1 \leq \frac{y - \langle w, \mu \rangle}{\|w_s\|}} c_1 c_2 g\left(\|z\|^2, \langle z, U\gamma_s \rangle\right) dz, \tag{1.114}$$

where c_2 is the determinant of the Jacobian of the unitary transformation and can be expressed as before as

$$c_2 = \left(\int c_1 g\left(\|z\|^2, \langle z, U\gamma_s \rangle\right) dz \right)^{-1}. \tag{1.115}$$

Without loss of generality, we replace the dummy variable z again by x,

$$F(y) = \int_{x_1 \leq \frac{y - \langle w, \mu \rangle}{\|w_s\|}} c_1 c_2 g\left(\|x\|^2, \langle x, U\gamma_s \rangle\right) dx. \tag{1.116}$$

Then we use the orthogonal transformation again, $z = Vx$. V is a unitary matrix such that the first row of which is $e_1' = (1, 0, ..., 0)$ and the second row is given by $\frac{(U\gamma_s)'}{\|\gamma_s\|}$. Therefore we have

$$z - 1 = x_1, \text{ and } z_2 = \frac{\langle x, U\gamma_s \rangle}{\|\gamma_s\|}. \tag{1.117}$$

Finally, we obtain

$$F(y) = \int_{z_1 \leq \frac{y - \langle w, \mu \rangle}{\|w_s\|}} c_1 c_2 c_3 g\left(\|z\|^2, \|\gamma_s\| z_2\right) dz \tag{1.118}$$

$$= \int_{z_1 \leq \frac{y - \langle w, \mu \rangle}{\|w_s\|}} \int_{z_2} \int_{z^\perp} c_1 c_2 c_3 g\left(z_1^2 + z_2^2 + \|z^\perp\|^2, \|\gamma_s\| z_2\right) dz^\perp dz_2 dz_1, \tag{1.119}$$

where $z^\perp = (z_3, z_4, ..., z_d)'$, and

$$c_3 = \left(\int c_1 c_2 g\left(\|z\|^2, \|\gamma_s\| z_2\right) dz \right)^{-1} \tag{1.120}$$

Thus our goal is to compute

$$h(z_1, z_2) \triangleq \int_{z^\perp} c_1 c_2 c_3 g\left(z_1^2 + z_2^2 + \|z^\perp\|^2, \|\gamma_s\| z_2\right) dz^\perp. \tag{1.121}$$

Let $c = c_1c_2c_3$, then $cg(\cdot)$ is a density of the original nearly elliptical distribution with parameters $\mu = 0$, $\Sigma = I$ and $\gamma = (0, \|\gamma_s\|, 0, ..., 0)'$. and h is the two dimensional marginal distribution of the nearly elliptical distribution. To get h numerically, we can use a Monte Carlo to generate N random vectors from the original elliptical distribution, $X^{(1)}, ..., X^{(N)}$. Then the function h can be apprixated by

$$\hat{h}(x_1, x_2) = \frac{1}{N} \sum_{i=1}^{N} \mathbb{1}_{\left\{X_1^{(i)} \in (x_1 - dx, x_1 + dx) \cap X_2^{(i)} \in (x_2 - dx, x_2 + dx)\right\}}. \tag{1.122}$$

Once the function h is obtained, the VaR at confidence level α then can be computed by solving

$$F(-\text{VaR}_\alpha(Y)) - \alpha = 0, \tag{1.123}$$

where F is given by

$$F(y) = \int_{z_1 \le \frac{y - \langle w, \mu \rangle}{\|w_s\|}} \int_{z_2} h(z_1, z_2) dz_2 dz_1. \tag{1.124}$$

1.7 Factor Analysis

For large number of assets in portfolio, the principle component analysis is commonly used to reduce dimension and find hidden statistical risk factors.

1.7.1 Basic Principle Component Analysis (PCA)

Suppose we have the returns vector from a portfolio $R = (R_1, ..., R_d)'$. And let Σ denotes the variance-covariance matrix of R. Then Σ is be decomposed by

$$\Sigma = V \Lambda V', \tag{1.125}$$

where V is an orthogonal matrix $VV' = I$, and

$$\Lambda = \text{diag}(\lambda_1, ..., \lambda_d) = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_d \end{pmatrix}, \tag{1.126}$$

with the eigenvalues $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_d$, and $V = (V_1, V_2, ..., V_d)$ with V_k the eigenvector corresponding to λ_k for $k = 1, ..., d$.

When $\lambda_i > \lambda_j$, the eigenvector V_i has more effect on the matrix Σ than V_j. Hence, Σ can be approximated by using only first largest K eigenvectors,

$$\Sigma \approx \widetilde{\Sigma} = (V_1, ..., V_K) \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_K \end{pmatrix} (V_1, ..., V_K)'. \tag{1.127}$$

For example, if $\frac{\lambda_1}{\sum \lambda} = 0.5$, then V_1 explains 50% of the portfolio variance. If $\frac{\lambda_2}{\sum \lambda} = 0.25$, the V_2 explains 25% of the portfolio variance additional to V_1 since V_1, V_2 are orthogonal to each other. Thus suppose for some $K \leq d$,

$$\frac{\sum_{k=1}^{K} \lambda_k}{\sum_{n=1}^{d} \lambda_n} \geq 0.95. \tag{1.128}$$

Then $V_1, ..., V_K$ explains more than 95% of the portfolio variance.

The eigenvector V_k can be reviewed as the weights on R so that the variance is given by $V_k' \Sigma V_k = \lambda_k$. The k-th principle component factor is referred as

$$F_k = V_k' R = \sum_{n=1}^{d} v_{n,k} R_n, \quad k = 1, 2, ..., K. \tag{1.129}$$

We have $\text{Cov}(F_i, F_j) = 0$ for $i \neq j$ since eigenvectors are orthogonal.

Applying the regress, we can obtain the following

$$R_n = \beta_{n,0} + \sum_{k=1}^{K} \beta_{n,k} F_k + \epsilon_n, \quad n = 1, 2, ..., d \tag{1.130}$$

where $\mathbb{E}\epsilon_n = 0$ for all $n = 1, ..., d$. Let $R_{P(w)}$ be the portfolio return with a capital allocation weight $w = (w_1, w_2, ..., w_d)'$. Then we have

$$R_{P(w)} = \beta_{P,0} + \sum_{k=1}^{K} \beta_{P,k} F_k + \sum_{n=1}^{d} w_n \epsilon_n, \tag{1.131}$$

where

$$\beta_{P,k} = \sum_{n=1}^{d} w_n \beta_{n,k}, \quad k = 0, 1, 2, ..., K. \tag{1.132}$$

And the portfolio variance is given by

$$\mathrm{Var}(R_{P(w)}) = \sum_{k=1}^{K} \beta_{P,k}^2 \,\mathrm{Var}(F_k) + \sum_{n=1}^{d} w_n^2 \,\mathrm{Var}(\epsilon_n). \tag{1.133}$$

1.7.2 PCA under FARIMA-FIGARCH with GH innovation

Let $R_n(t)$ be the observed n-th asset resturn at time $t \in \{0, 1, 2, ...\}$ and τ be the current time. Assuming R_n follows FARIMA-FIGARCH model, then we have, for $n = 1, ..., d$,

$$R_n(\tau + 1) = \mu_n(\tau + 1) + \sigma_n(\tau + 1) X_n(\tau + 1), \tag{1.134}$$

where $\mu_n(\tau + 1)$ is the forecasted mean by FARIMA, and $\sigma_n(\tau + 1)$ is the forecasted std by FIGARCH.

Let $X(t) = (X_1(t), ..., X_d(t))'$ and

$$X = (X(\tau), X(\tau - 1), ..., X(1)) = \begin{pmatrix} X_1(\tau) & X_1(\tau - 1) & \cdots & X_1(\tau) \\ X_2(\tau) & X_2(\tau - 1) & \cdots & X_2(\tau) \\ \vdots & \vdots & \ddots & \vdots \\ X_d(\tau) & X_d(\tau - 1) & \cdots & X_d(\tau) \end{pmatrix}. \tag{1.135}$$

Note that the random variable $X_n(t)$ has zero mean and unit variance for $n = 1, 2, ..., d$. Then PCA can be easily conducted. Let $\lambda_1, ...\lambda_d$ be the eigenvalues and $V_1, ..., V_d$ the corresponding eigenvectors of the covariance of X, that is,

$$\mathrm{Cov}(X) = \mathbb{E}XX' = V\Lambda V', \tag{1.136}$$

where $\Lambda = \mathrm{diag}(\lambda_1, ..., \lambda_d)$ and $V = (V_1, ..., V_d)$ with $\lambda_1 \geq \cdots \geq \lambda_d$ and $V_n = (v_{1,n}, ..., v_{d,n})'$. And we select K by

$$K = \min\left\{ k \,\middle|\, \frac{\sum_{k=1}^{K} \lambda_k}{\sum_{n=1}^{d} \lambda_n} \geq \alpha \right\}, \tag{1.137}$$

where α is a given accuracy level. Denote $\widehat{V} = (V_1, ..., V_K)$.

Let the factor sequence $F_k(t) = V_k' X(t)$ for $k = 1, .., K$ and $F(t) = (F_1(t), ..., F_K(t))'$, that is, $F(t) = \widehat{V}' X(t)$. Let

$$F = (F(\tau), F(\tau-1), ..., F(1)) = \begin{pmatrix} F_1(\tau) & F_1(\tau-1) & \cdots & F_1(\tau) \\ F_2(\tau) & F_2(\tau-1) & \cdots & F_2(\tau) \\ \vdots & \vdots & \ddots & \vdots \\ F_d(\tau) & F_d(\tau-1) & \cdots & F_d(\tau) \end{pmatrix}. \tag{1.138}$$

Then we have $F = \widehat{V}' X$.

By multiple regression, we have

$$X(t) = \beta' F(t) + \epsilon(t), \tag{1.139}$$

where $\epsilon(t) = (\epsilon_1(t), ..., \epsilon_d(t))'$, and

$$\beta = (FF')^{-1} F X' = (\widehat{V}' X X' \widehat{V})^{-1} \widehat{V}' X X' = \widehat{V}'. \tag{1.140}$$

Hence we have

$$X(t) = \widehat{V} F(t) + \epsilon(t), \quad t = 1, ..., \tau, \tag{1.141}$$

or in a matrix form,

$$X = \widehat{V} F + \epsilon, \tag{1.142}$$

where

$$\epsilon = (\epsilon(\tau), \epsilon(\tau-1), ..., \epsilon(1)) = \begin{pmatrix} \epsilon_1(\tau) & \epsilon_1(\tau-1) & \cdots & \epsilon_1(\tau) \\ \epsilon_2(\tau) & \epsilon_2(\tau-1) & \cdots & \epsilon_2(\tau) \\ \vdots & \vdots & \ddots & \vdots \\ \epsilon_d(\tau) & \epsilon_d(\tau-1) & \cdots & \epsilon_d(\tau) \end{pmatrix}. \tag{1.143}$$

Now we can extract ϵ as

$$\epsilon(t) = X(t) - \widehat{V}\widehat{V}'X(t) = (I - \widehat{V}\widehat{V}')X(t), \ t = 1, ..., \tau, \tag{1.144}$$

or in matrix form

$$\epsilon = X - \widehat{V}\widehat{V}'X = (I - \widehat{V}\widehat{V}')X. \tag{1.145}$$

Here we assume the innovations are from GH distribution,

$$F(t) \sim \mathrm{GH}(\lambda, \chi, \psi, \mu_F, \gamma_F, \Sigma_F), \tag{1.146}$$
$$\epsilon(t) \sim \mathrm{GH}(\lambda, \chi, \psi, \mu_\epsilon, \gamma_\epsilon, \Sigma_\epsilon). \tag{1.147}$$

Then we have that

$$\begin{pmatrix} F(t) \\ \epsilon(t) \end{pmatrix} \sim \mathrm{GH}\left(\lambda, \chi, \psi, \begin{pmatrix} \mu_F \\ \mu_\epsilon \end{pmatrix}, \begin{pmatrix} \gamma_F \\ \gamma_\epsilon \end{pmatrix}, \begin{pmatrix} \Sigma_F & \Sigma_{F\epsilon} \\ \Sigma'_{F\epsilon} & F_\epsilon \end{pmatrix}\right). \tag{1.148}$$

In addition, we assume that $F(t)$ and $\epsilon(t)$ are uncorrelated. It implies the covariance of $(F(t), \epsilon(t))'$ is

$$\begin{pmatrix} C_F & 0 \\ 0 & C_\epsilon \end{pmatrix} = \mathbb{E}(W) \begin{pmatrix} \Sigma_F & \Sigma_{F\epsilon} \\ \Sigma'_{F\epsilon} & F_\epsilon \end{pmatrix} + \mathrm{Var}(W) \begin{pmatrix} \gamma_F\gamma'_F & \gamma_F\gamma_\epsilon \\ \gamma_\epsilon\gamma'_F & \gamma_\epsilon\gamma'_\epsilon \end{pmatrix}, \tag{1.149}$$

where $W \sim \mathrm{GIG}(\lambda, \chi, \psi)$. Thus we have

$$\Sigma_{F\epsilon} = -\frac{\mathrm{Var}(W)}{\mathbb{E}(W)}\gamma_F\gamma'_\epsilon. \tag{1.150}$$

Furthermore, we assume $\epsilon(t)$ is uncorrelated. It implies that C_ϵ is a diagonal matrix. Thus we have

$$\Sigma_\epsilon = \mathrm{diag}(\Sigma_\epsilon) - \frac{\mathrm{Var}(W)}{\mathbb{E}(W)}\left(\gamma_\epsilon\gamma'_\epsilon - \mathrm{diag}(\gamma_\epsilon\gamma'_\epsilon)\right). \tag{1.151}$$

Then the distribution of $X(t)$ is then given by

$$X(t) \sim \begin{pmatrix} \widehat{V} & I \end{pmatrix} \begin{pmatrix} F(t) \\ \epsilon(t) \end{pmatrix} \sim \mathrm{GH}(\lambda, \chi, \psi, \mu_X, \gamma_X, \Sigma_X), \tag{1.152}$$

where

$$\mu_X = \widehat{V}\mu_F + \mu_\epsilon, \tag{1.153}$$

$$\gamma_X = \widehat{V}\gamma_F + \gamma_\epsilon, \tag{1.154}$$

and

$$\Sigma_X = \widehat{V}\Sigma_F\widehat{V}' + \widehat{V}\Sigma_{F\epsilon} + \Sigma_{F\epsilon}'\widehat{V}' + \Sigma_\epsilon \tag{1.155}$$

$$= \widehat{V}\Sigma_F\widehat{V}' + \mathrm{diag}(\Sigma_\epsilon) - \frac{\mathrm{Var}(W)}{\mathbb{E}(W)}\left(\widehat{V}\gamma_F\gamma_\epsilon' + \gamma_\epsilon\gamma_F'\widehat{V}' + \gamma_\epsilon\gamma_\epsilon' - \mathrm{diag}(\gamma_\epsilon\gamma_\epsilon')\right). \tag{1.156}$$

If we further assume that $\epsilon(t)$ follows elliptical GH, that is, $\gamma_\epsilon = 0$, then we can rewrite Σ_X as

$$\Sigma_X = \widehat{V}\Sigma_F\widehat{V}' + \mathrm{diag}(\Sigma_\epsilon). \tag{1.157}$$

Moreover, if we assume $\epsilon(t)$ has zero mean, then Σ_ϵ can be obtained as

$$\Sigma_\epsilon = \frac{\mathrm{diag}(\mathrm{Cov}(\epsilon(t)))}{\mathbb{E}(W)}. \tag{1.158}$$

1.7.3 GH-distributed Factor Analysis

Here we will discuss the factor analysis under GH distribution assumption (GHFA). Consider the following model

$$X = m_X + BF + E, \tag{1.159}$$

where

1. X is the $d \times 1$ random vector for asset return or for the innovations of FARIMA-FIGARCH model,

2. m_X is the $d \times 1$ real vector for the intercept,

3. B is the $d \times k$ loading matrix for coefficients,

4. F is the $k \times 1$ random vector for the latent factors,

5. E is the $d \times 1$ random vector for the noise with diagonal covariance matrix $\Xi = \text{diag}(\xi_1, \xi_2, ..., \xi_d)$.

Parameter estimation for the GHFA is constructed by MLE with the following model,

$$\begin{pmatrix} X \\ F \end{pmatrix} \sim \text{GH}\left(\lambda, \chi, \psi, \begin{pmatrix} m_X \\ m_F \end{pmatrix}, \begin{pmatrix} BB' + \Xi & B \\ B' & I \end{pmatrix}, \begin{pmatrix} \gamma_X \\ \gamma_F \end{pmatrix}\right), \tag{1.160}$$

where I_k is the k-dimensional identity matrix, γ_X and $\gamma - F$ are $d \times 1$ and k real vectors respectively, m_F is $k \times 1$ real vector, and λ, χ, ψ are real numbers. It can be written in normal mean-variance mixture form,

$$\left.\begin{pmatrix} X \\ F \end{pmatrix}\right|_W \sim \mathcal{N}\left(\begin{pmatrix} m_X + \gamma_X W \\ m_F + \gamma_F W \end{pmatrix}, \frac{1}{W}\begin{pmatrix} BB' + \Xi & B \\ B' & I_k \end{pmatrix}\right), \tag{1.161}$$

with $W \sim \text{GIG}(\lambda, \chi, \psi)$.

To simplify the model , we assume that $m_F = -\gamma_F \mathbb{E}(W)$. Then given n i.i.d samples $x_1, ..., x_n$, the EM algorithm can be applied to estimate the parameters $\theta = (B, \Xi, \gamma_X, \gamma_F, \lambda, \chi, \psi)$,

$$\max_{\theta, F} l(\theta, F|X) = \max_{\theta, F} \sum_{i=1}^{N} \log(f_{\text{GH}}(x_i, F|\theta)), \tag{1.162}$$

where f_{GH} denotes the density function of GH distribution.

In the special case that $\lambda = \chi = \frac{\nu}{2}$ and $\psi = 0$, $W \sim \text{Gamma}(\nu/2, \nu/2)$, we have

$$\left.\begin{pmatrix} X \\ F \end{pmatrix}\right|_W \sim \mathcal{N}\left(\begin{pmatrix} m_X \\ 0 \end{pmatrix}, \frac{1}{W}\begin{pmatrix} BB' + \Xi & B \\ B' & I_k \end{pmatrix}\right), \tag{1.163}$$

when assuming $m_F = 0$, $\gamma_X = 0$, and $\gamma_F = 0$.

1.7.4 VaR with Factor Analysis

Here let $R_{P(w)}$ denotes the forecasted portfolio return at $\tau+1$ with a captical allocation weight $w = (w_1, ..., w_d)'$. Then we have

$$R_{P(w)}(\tau+1) = \sum_{n=1}^{d} w_n \mu_n(\tau+1) + \sum_{n=1}^{d} w_n \sigma_n(\tau+1) X_n(\tau+1) \tag{1.164}$$

$$= w'\mu(\tau+1) + w' \operatorname{diag}(\sigma(\tau+1)) X(\tau+1), \tag{1.165}$$

where

$$\mu(\tau+1) = (\mu_1(\tau+1), ..., \mu_d(\tau+1)), \text{ and} \tag{1.166}$$

$$\sigma(\tau+1) = (\sigma_1(\tau+1), ..., \sigma_d(\tau+1)). \tag{1.167}$$

By the portfolio property of GH distribution family, we have

$$Y \triangleq w' \operatorname{diag}(\sigma(\tau+1)) X(\tau+1) \sim \mathrm{GH}(\lambda, \chi, \psi, \mu_Y, \gamma_Y, \Sigma_Y), \tag{1.168}$$

where

$$\mu_Y = w' \operatorname{diag}(\sigma(\tau+1)) \mu_X, \tag{1.169}$$

$$\gamma_Y = w' \operatorname{diag}(\sigma(\tau+1)) \gamma_X, \tag{1.170}$$

$$\Sigma_Y = w' \operatorname{diag}(\sigma(\tau+1)) \Sigma_X \operatorname{diag}(\sigma(\tau+1)) w. \tag{1.171}$$

Therefore, the portfolio VaR under information F_τ can be written as

$$\mathrm{VaR}_\alpha(R_{P(w)(\tau+1)} | F_\tau) = -w'\mu(\tau+1) + \mathrm{VaR}_\alpha(Y). \tag{1.172}$$

1.8 Empirical Results

In this section pick 5 stocks to empirical test FARIMA-FIGARCH with GH innovation models, including MSFT, C, PFE, GE and GIS. The sampling frequency is 1 minute. And the time period is from 2013-01-01 to 2013-1-31. Thus the data set is 1-min log returns of one-month.

Summary Statistics

Table 1.1: Basic Statistics of Intraday 1-min Log-returns

	Mean (10^{-5})	Std (10^{-3})	Skewness	Kurtosis
MSFT	0.4145	0.5741	-0.1080	15.5225
C	0.0685	0.6464	-0.1383	11.7905
PFE	0.2062	0.5297	-0.1376	16.6694
GE	0.0249	0.4984	-0.0913	13.9847
GIS	0.1723	0.4466	-0.2163	30.5321

The basic statistics are summarized in Table 1.1. As documented in many literatures, the asset returns show negative-skewness, heavy-tailedness and volatility clustering.

The auto-correlation function (ACF) of the asset return series and its magnitude are shown in Figure 1.1 and Figure 1.2. The auto-correlation of log-return and absolute log-return is almost flat in the log-log plot when lag is large, which indicates a power decay rather than exponential. It implies that the auto-correlations especially of the absolute log-return have very slow decays when lag is large, which suggests the possibility of long-range dependence especially in volatility.

FARIMA-FIGARCH

The estimated model are listed in Table 1.2. After filtering by FARIMA-FIGARCH model, the residuals has almost zero auto-correlation and as well as its absolute value. It is illustrated as an example in Figure 1.3.

Value at Risk

The VaR are computed for each asset with a rolling window of size one week ($390 \times 5 = 1950$ data points). The results are illustrated in Figure 1.4-1.5.

For a equally weighted portfolio composed with theses 5 assets, the VaR is computed and report in Figure 1.6.

The violation number for each assets and equally-weighted portfolio of the whole period are summarized in Table 1.3.

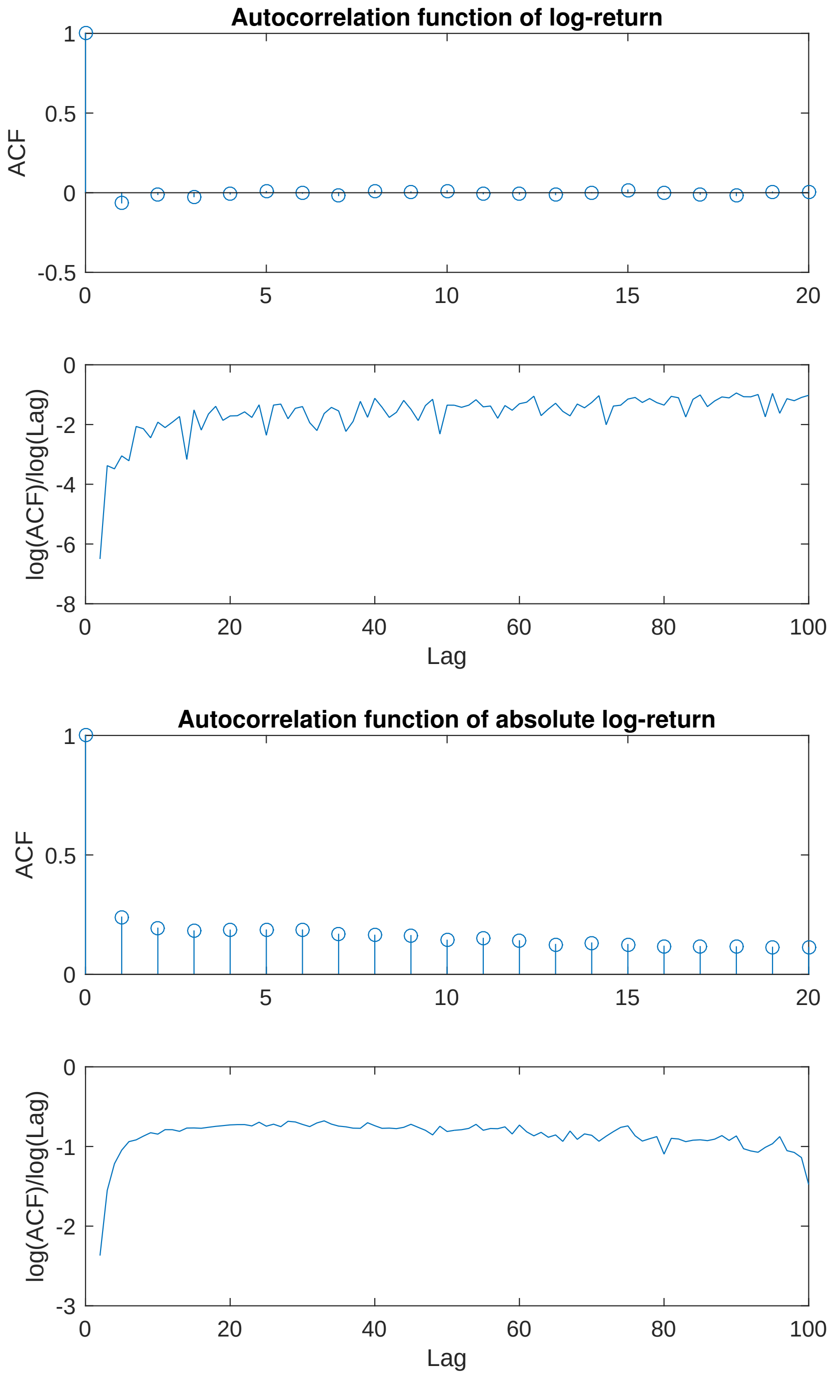

Figure 1.1: MSFT Autocorrelation Function.

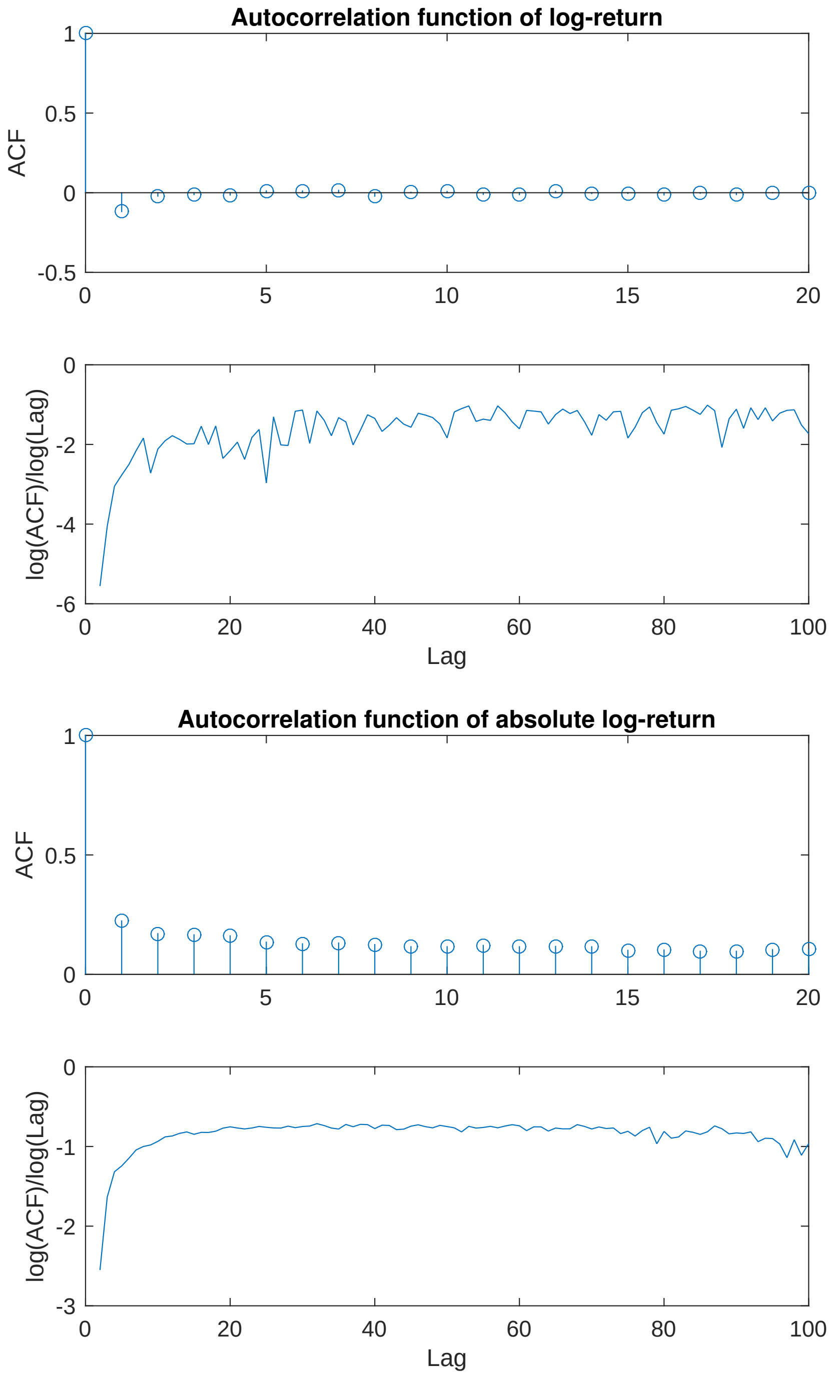

Figure 1.2: GE Autocorrelation Function.

Table 1.2: FARIMA-FIGARCH Estimation Results Illustration

	FARIMA				FIGARCH			
	d_0	ϕ	θ	μ (10^{-4})	d	ψ	β	σ^2 (10^{-8})
MSFT	-0.0255	0.0087	-0.0685	-0.1676	0.2328	0.2300	0.0250	2.9358
	(0.0000)	(0.6363)	(0.5414)	(0.0031)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
C	0.0024	0.0436	-0.2040	0.2473	0.9568	0.8521	0.0560	0.4879
	(0.2603)	(0.6660)	(0.0000)	(0.0008)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
PFE	-0.1406	0.0184	0.0093	0.1185	0.4717	0.6610	0.2683	0.5478
	(0.0000)	(0.0567)	(0.0000)	(0.2013)	(0.0000)	(0.0000)	(0.2194)	(0.0000)
GE	-0.1656	0.0153	-0.0007	-0.1626	0.3204	0.2461	0.0250	2.9879
	(0.0000)	(0.4627)	(0.0000)	(0.0001)	(0.0000)	(0.0000)	(0.0715)	(0.0000)
GIS	-0.1022	0.0510	-0.0442	0.0073	0.3193	0.2388	0.0325	1.7136
	(0.0000)	(0.3081)	(0.1847)	(0.0285)	(0.0000)	(0.0000)	(0.7221)	(0.0000)

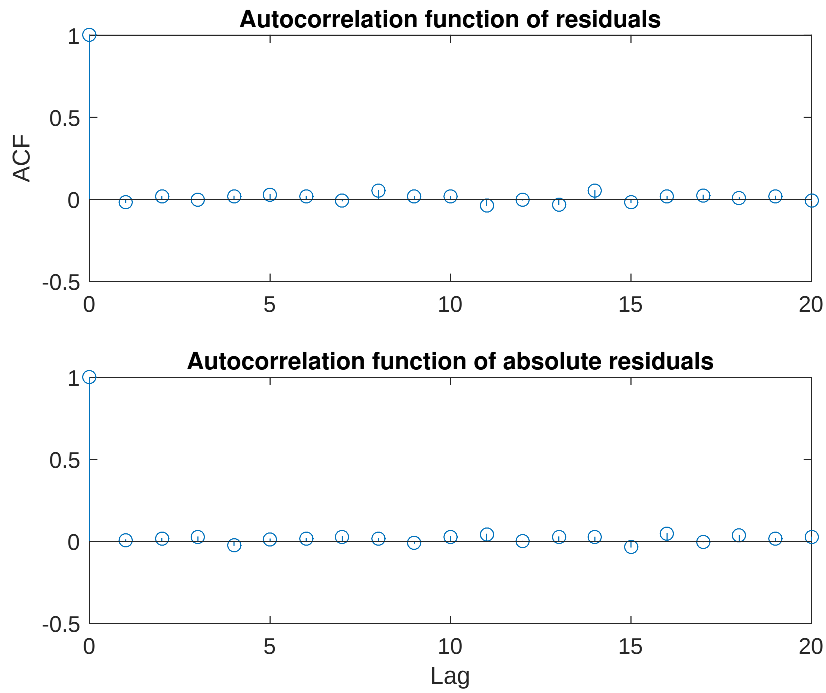

Figure 1.3: MSFT Autocorrelation Function after FARIMA-FIGARCH Filtering.

Table 1.3: Number of Violations of Value-at-Risk.

	VaR(99%) GH	VaR(99%) Normal	VaR(95%) GH	VaR(95%) Normal
MSFT	119 (1.45%)	212 (2.59%)	493 (6.02%)	482 (5.89%)
C	116 (1.42%)	191 (2.23%)	518 (6.32%)	515 (6.29%)
PFE	131 (1.60%)	199 (2.43%)	513 (6.26%)	491 (6.00%)
GE	127 (1.55%)	206 (2.52%)	494 (6.03%)	455 (5.56%)
GIS	75 (0.92%)	137 (1.67%)	436 (5.32%)	361 (4.41%)
Portfolio	85 (1.04%)	120 (1.47%)	497 (6.07%)	446 (5.45%)

Time period is one-month ($390 \times 21 = 8910$ minutely returns).

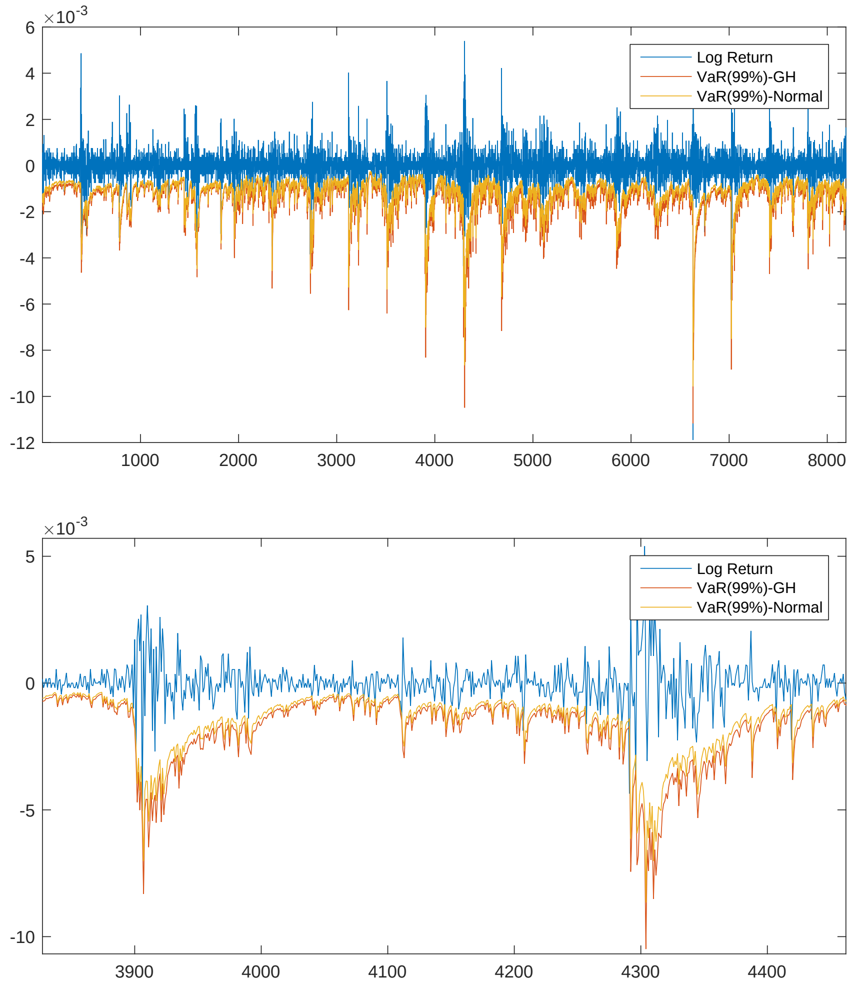

Figure 1.4: VaR of MSFT. Upper plot is for whole period. Lower is a zoomed-in plot.

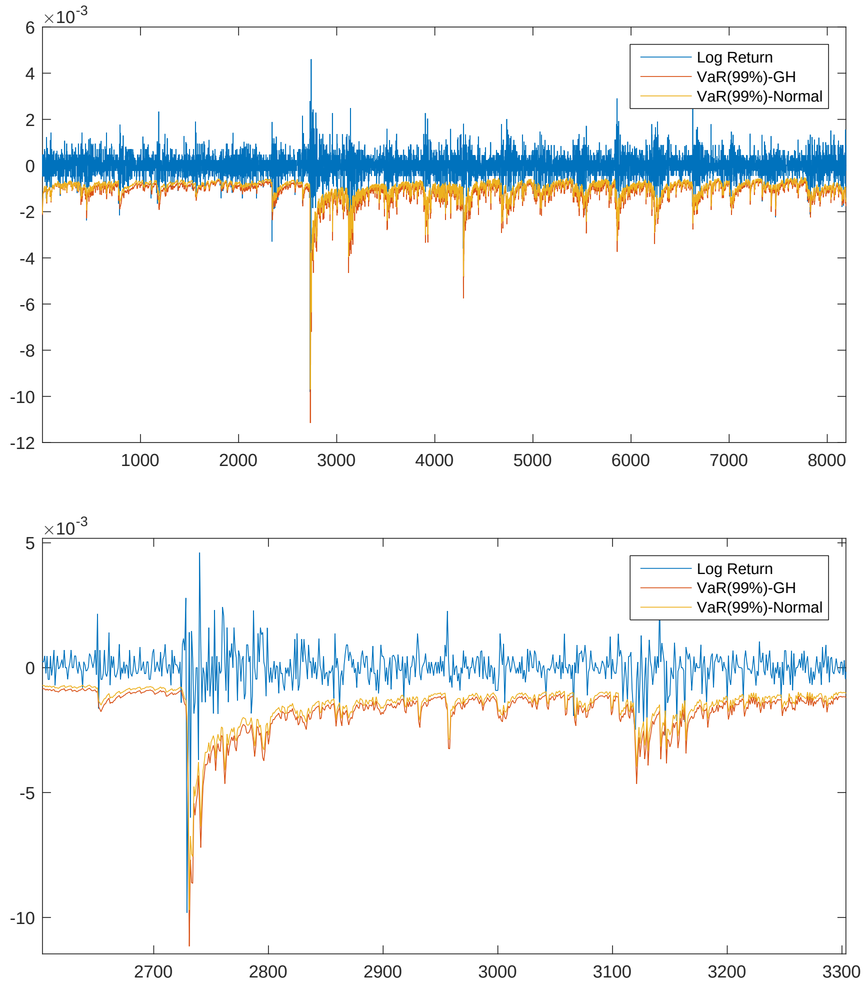

Figure 1.5: VaR of GE. Upper plot is for whole period. Lower is a zoomed-in plot.

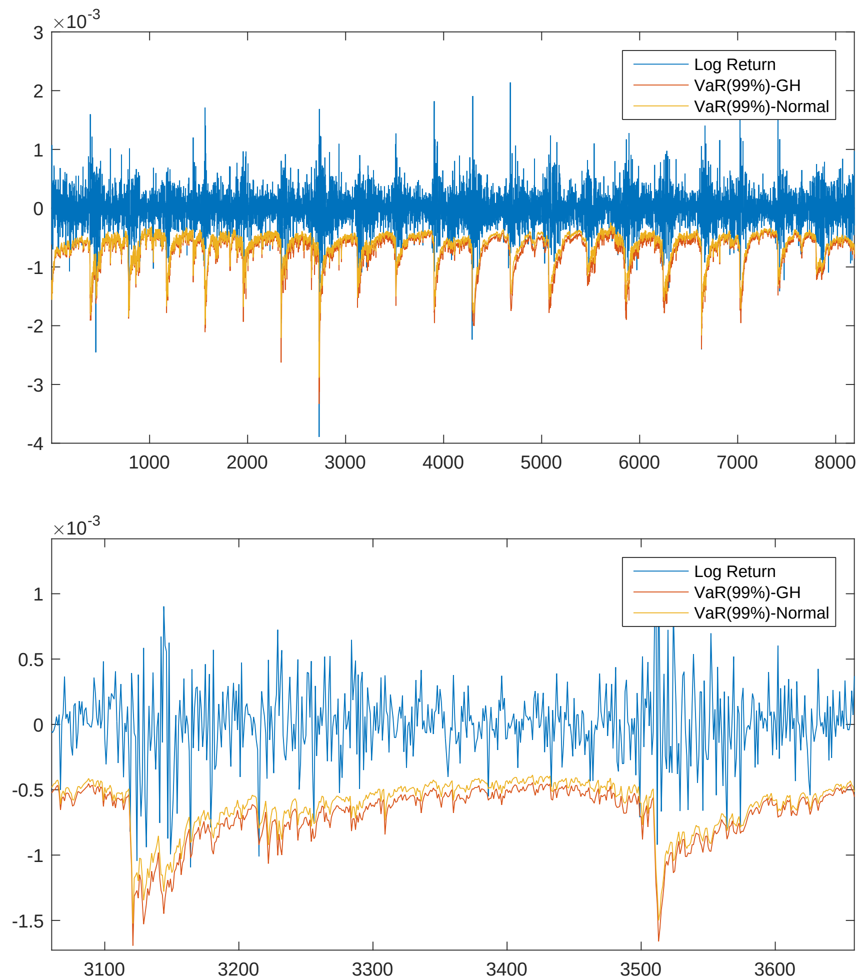

Figure 1.6: VaR of equally weighted portfolio. Upper plot is for whole period. Lower is a zoomed-in plot.

Chapter 2
Market Efficiency in Different Market Sentiment

2.1 Introduction and Motivation

Market efficiency basically says that the price in the stock market reflects all the available information. If the Market Efficiency Hypothesis(EMH) holds, when all the information about the investments is known, it is not possible for anyone to beat the market and expect returns that are above average. EMH views market prices as random thus serial correlations especially long range dependence between observations cannot exist. Therefore, if we already observe the long range dependence in the market as documented in large volume of literatures such as Sensoy 2013, Cajueiro and Tabak 2008, Hammoudeh and Yoon 2014, then EMH should not hold. Meanwhile, the behavior of break of EMH should in certain level related to the behavior of long range dependance.

Based on this, we explore a well documented break of EMH, the effect on market price brought from investor sentiment.

2.2 A Survey on Investor Sentiment

2.2.1 Definition

Investor sentiment can be defined broadly as a belief about future market dynamic and investment risks not justified by the facts at hand.[1] That means investors could be optimistic or pessimistic on the market. As a consequence, investors will make their decisions based on not only the information accessible

[1]Notice this definition is from *Baker and Wurgler (2007)*. Although the sentiment is a belief not justified by the facts, it still could be related to the facts at hand, see overconfidence model.

but also the believes they are holding.

The market efficiency hypothesis states that security prices always fully reflect available information. Over the last decade that paradigm has come under attack. Shleifer (2000)summarizes the related strands of literature. First, theoretical work argues that arbitrage has limited effectiveness. Second, experimental evidence shows that agents hold beliefs that are not completely correct and/or make choices that are normatively questionable. In another word, the question why the investor sentiment will influence the market and price behavior has been well answered by these two aspects. It is because that the arbitrage, which in traditional asset pricing theory, is considered able to offset influences brought by sentiment [*Fama (1970)*], nowadays, is more and more taken as risky, limited and not be able to offset the investor's sentiment/ irrationality.

A number of researchers, such as *Grossman and Stiglitz (1980)*, *Black (1986)*, em DeLong et al. (1990), *Campbell and Kyle (1993)*, *Barberis et al. (1998)*, *Daniel et al. (1998)*, and *Hong and Stein (1999)* have more formally modeled the role of sentiment or investor behavior. However, how to measure the effect of investor sentiment is still an open problem. In this article, we are mainly focusing to answer how investor sentiment influences the market.

2.2.2 Measurement

Authors have been seeking for a good measure of investor sentiment for a long time. Some measures are constructed and/or extracted directly from the financial market, such as closed-end fund discount, IPO first-day returns and trading volume among others. There are also other sentiment measures coming from media sources, such as professional journals and corporate accouterments. The following list summaries the popular measures in existing literature:

- Several sentiment proxies are summarized or proposed in *Baker and Wurgler (2007), p135-p138*, namely, investor surveys, investor mood, retail investor trades, mutual fund flows, trading volume, dividend premium, closed-end fund discount, option implied volatility, IPO first-day returns, IPO volume, equity issues over total new issues and insider trading.

- In behavioral finance world, sentiment is synonymous with error. Thus, in some literature, scientists use mispricing factor such as discretionary accruals, netequity issuances/repurchases, and price momentum as sentiment factor,as in *Polk and Sapienza(2002).*

- In *Baker and Wurgler (2006)*, sentiment index is a linear combination of six selected proxies: "trading volume as measured by NYSE turnover; the dividend premium; the closed-end fund discount; the number and first-day returns on IPOs; and the equity share in new issues". This sentiment index is commonly used in recent literature.

- In *Kaplanski and Levy (2010)*, aviation disasters are used as a indicator of bad mood of investors.

- In *Brown and Cliff (2005)*, bull-bear spread is constructed by tracking the number of market newsletters. This is also a popular measurement of investor sentiment.

- *Rosen (2006)* uses merge announcement as a signal of investor becoming optimistic.

- In *Tetlock (2007)*, so-called pessimism media factor is used, which counts the number of key words in WSJ column.

- *Edmans, García and Norli (2007)* proxies investor mood by international soccer results.

- *Mian and Sankaaraguruswamy (2010)* uses firm-specific earning news as the trigger of investor sentiment reaction.

- *Wang 2001* classifies traders into three classes: large speculators, large hedgers, and small traders. And for each type of traders, its sentiment is measured from its aggregate position.

- Conference Board Consumer Confidence Index (CCI), Consumer Sentiment Index (CSI)

- *News sentiment* that is extracted from news publications.
- *Barron's Confidence Index*is confidence indicator calculated by dividing the average yield on high-grade bonds by the average yield on intermediate-grade bonds. The discrepancy between the yields is indicative of investor confidence.A rising ratio indicates investors are demanding a lower premium in yield for increased risk and so are showing confidence in the economy.

Among these measures, most of them (expect the event-based) are constructed as indexes. Moreover, these indexes, as proxies of investor sentiment, are regarded as risk factors. That means the mechanism and interplay between market and investors believes are put into black-box.

2.2.3 Empirical Effects

There are at least two difficulties to model how investor sentiment affects market prices: 1. It is hard to describe sentiment, a belief investor are holding. Probabilistically speaking, it is the probability measure that investor is using to decide trading strategy for making profits and reducing risks. 2. It is hard to quantitatively describe the mechanism how investor sentiment moves market. When investors are irrational, i.e. subject to their own sentiments, the demand and supply relation changes. And the investors interplay with each other, then finally the market gets to an equilibrium and gives prices. Only a few literature work on quantify and model this mechanism.

However, literature have done rich empirical analysis with considering sentiment proxies as risk factors, that means the mechanism and interplay between market and investors believes are put into black-box. The important empirical effects are listed as follows,

- Short-horizon positive relation and Event effect, that is, stock will be overpriced/underpriced when investor sentiment becoming high/low (triggered by events). *Kaplanski and Levy (2010)* gives evidences that "aviation disaster negatively affect stock prices for a short period of a few days". *Rosen (2006)* shows stock prices tend to increase in short-run scale after a merger announcement. Similar evidences can also be found in *Brown and Cliff (2005)*, *Tetlock (2007)*, *Brown and Cliff (2004)*, *Hengelbrock, Theissen, and Westheide (2010)*.

- Long-run Reversal effect, that is, stock returns will revert back to average slowly after a sentiment shock, or stock returns in long horizon has a negative relation with investor sentiment. *Kaplannski and Levy (2010)* gives an example that market takes about 10 days to revert to average after a decline triggered by aviation disaster. Similar results can also be found in *Rosen (2006)*, *Brown and Cliff (2005)*.

- Cross-sectional effects, that is, the impact of investor sentiment on the cross-section of stock returns. *Chung et al. (2012)* documents the asymmetric predictive power across portfolios formed on size, book-to-market, dividend-yield, P/E ratio, age, sigma, R&D expense/assets ratio, fixed assets sales growth and external finance/asset ratio. *Neal and Wheatley (1998)* finds the differences of prediction pattern between small and large firm. *Kumar and Lee (2006)* also constructs portfolios based on firm size. *Baker and Wurgler (2006)* examines the expected return of characteristic-based portfolios conditioned on sentiment level. See also in *Polk and Sapienza (2002)*, *Chung, Hung and Yeh (2012)*.

- Subject to market regime. *Chung, Hung and Yeh (2012)* shows prediction power of investor sentiment under different market regimes (described by NBER recession index or in Markov-switching model).

- Effects on implied volatility. *Han (2007)* shows implied volatility smile is steeper when investor sentiment is low.

2.2.4 Related Theories

Figure 2.1 shows a logic framework of related researches following the development of Efficient Market Hypothesis (EMH). The statements in black are the assumptions and deductions of EMH. The evidences in blue contradict their corresponding statement in black. The red stars show where investor sentiment is playing a role. The literature corresponding to the red stars are summaries in following.

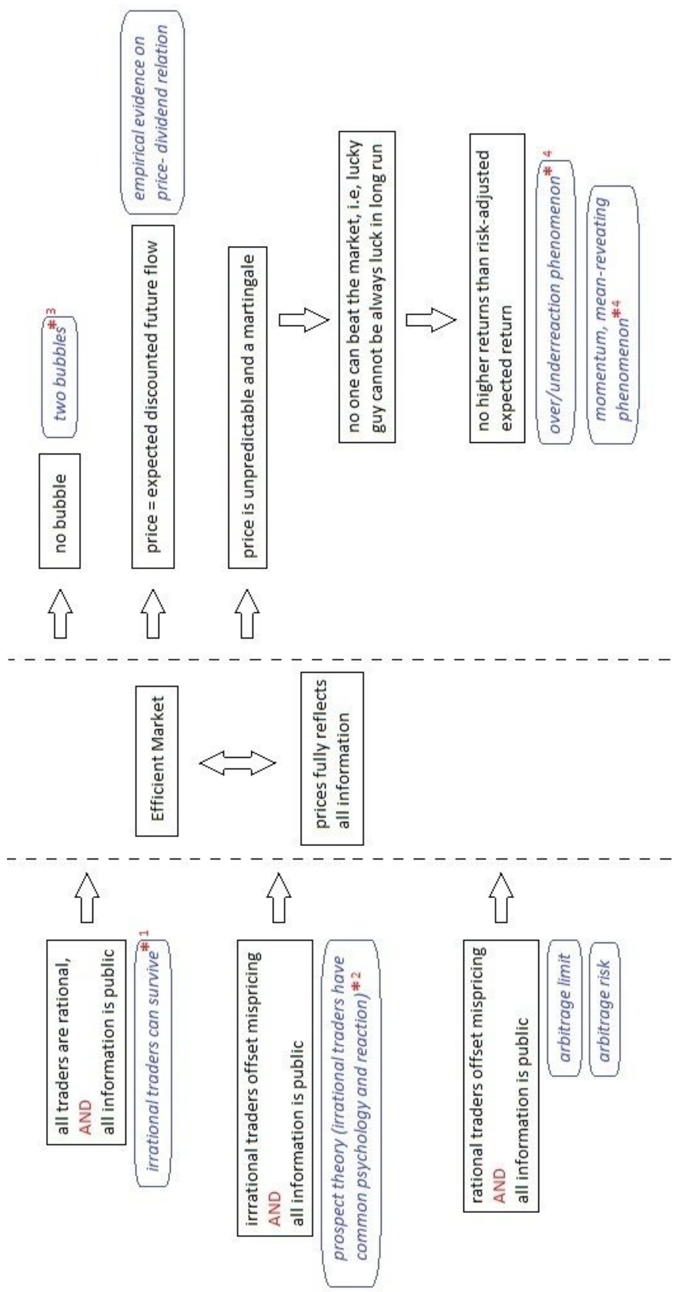

Figure 2.1: Related theories

1* *DeLong, Shleifer, Summers and Waldmann (1991)* build the model for portfolio choice of irrational(noise) and rational trader, and show that noise trader, who has no price impact, can achieve even higher expected return. *Wang (2001)* build an evolutionary game model to examine whether irrational investor can survive in a large economy. Their conclusion claims moderate overconfident or optimistic investors can survive and even dominate. *Kogan, Ross, Wang, and Westerfield (2006)* claims irrational trader can affect market significantly under a competitive equilibrium setting.

2* *Tversky and Kanheman (1979)* develop the Prospect Theory giving a understanding of how investor make decision under uncertainty and risk. *Gallimore and Gary (2002)* investor sentiment affects decision-making for both rational and irrational traders. *Wang, Yan, Yu (2012)* apply prospect theory and provide cross-sectional empirical evidence.

3* *Scheinkman and Xiong (2003)* propose an equilibrium with agents holding overconfident sentiment to illustrate the behavior of many equilibrium vairiables of interest in bubble.

4* *Barberis, Shlefier and Vishny (1997)* propose a regime-switching Markov model to formalize investor over/underreation behavior. Empirical works can be found in many literatures[2].

2.2.5 Discussion

On sentiment measure

In short word, it is not a simple causality relation between market performance and investor sentiment, but interplay. Therefore, for sentiment measures that are extracted directly form the market, such as trading volume, dividend premium, option implied volatility among others, it is hard to distinguish how much sentiment is measured and how much market response is measured.

To construct the good measure, a nature way is to follow the way how investors

[2]See *Brown and Cliff (2004)*, *Rosen (2006)*, *Tetlock (2007)*, *Kaplannski and Levy (2010)* among others.

form their believes. Investors' decision-making subject to their believes, and their believes are moved by not only market performance but also exogenous information (such as news, big social events), their own historical investment performance and even their own personal character. Although to model the whole mechanism seems not feasible, it is suggested that exogenous information should be considered as risk resources.

On time horizon of over/underreation

Time horizon in existing literature are quite vary, from minutely to yearly[3]. It reminds us that investors process news information in different scale. For long-term portfolio management, investors tend to pay more attention on earning announcement, accounting report and other more fundamental and low-frequent news. The daily traders tends to take more high-frequent news into account. For some occasionally happened event, such as aviation disaster and senior management changes, investors could react in vary time horizons. As the consequence, the over/underreaction behavior of prices to news of different frequency could be very differently in terms of time horizon.

On firms who are more sensitive to investors sentiment

Many analysis on cross-section of stocks have been done to identify what type of firms is more sensitive to investor sentiment. A straightforward guess could be that the firms which irrational traders prefer to invest are more sensitive. In other words, the prices of firms which are held by more rational investors are more stable. Therefore, it is reasonable to state the hypothesis, the firms held by more institutional investors are more sensitive to sentiment if assuming institutional investors are more rational.

However, it has never been fully proved that a small amount of irrational traders with extreme or mediate sentiment cannot move the the prices in a considerable scale.

[3]See *Kaplanski and Levy (2010)*, *Brown and Cliff (2005)*, *Tetlock (2007)*, *Brown and Cliff (2004)*, *Hengelbrock, Theissen, and Westheide (2010)*.

2.3 A Simple General Equilibrium Model

We consider a parsimonious model in general equilibrium setting with news sentiment and irrational belief. Our purpose is to find what is the price impact of irrational traders with incorrect belief with is influenced by exogenous news sentiment process.

2.3.1 The Economy

Information structure

We consider a continuous-time economy with finite time horizon $[0, T]$. The only uncertainty source is a one-dimensional, standard Brownian motion $(W_t, 0 \leq t \leq T)$[4], defined on a complete filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$, where $\mathbb{F}$ is the augmented filtration generated by the Brownian motion W_t.

The financial market

In the economy, there exists one risky asset, the stock, which pays aggregate dividend process (D_t) given by

$$\frac{dD_t}{D_t} = \mu dt + \sigma dW_t.$$

The stock price process is denoted by (S_t). Also, the economy allows investment on zero-coupon bond in zero net supply. Each of the bond mature at time T and pays the face value of one[5].

[4] The assumption of only one uncertainty source can be improved by importing other uncertainty source which simultaneously drives news sentiment process.

[5] Note that here no interest rate structure is considered. In other words, risk-free rate is always 1. This condition can be relaxed by assuming the price process of bond B_t follows $dB_t = B_t r_t dt$.

Trading strategy

Assume the market is frictionless and trading can be made in continuous time. Investors' trading strategy (θ_t) satisfies the integrability condition

$$\int_0^T \theta_t^2 S_t^2 dt < \infty. \text{ a.s.}$$

Endowments

Assume there are two investors, one rational and one irrational, denoted by $1, 2$ respectively. Rational and irrational investors are endowed with $q^{(1)} = q$ and $q^{(2)} = 1 - q$ share of the stock at time 0 respectively, where $0 < q < 1$.

News sentiment

Assume there exist two news sentiment processes η_t^+, η_t^- follows mean-reverting dynamic

$$\begin{aligned} d\eta_t^+ &= \theta^+(\bar{\eta}^+ - \eta_t^+)dt + \sigma_\eta^+ dW_t, \\ d\eta_t^- &= \theta^-(\bar{\eta}^- - \eta_t^-)dt - \sigma_\eta^- dW_t, \end{aligned}$$

where $\theta^+, \theta^-, \bar{\eta}^+, \bar{\eta}^-, \sigma_\eta^+, \sigma_\eta^-$ are some positive parameters. These two process denote the positive and negative sentiment processes respectively. The reason to model positive and negative new sentiment separately is that empirical evidences has been shown that positive sentiment and negative sentiment have different momentum properties and also market response to them are different.

Both of rational and irrational traders observe this news sentiment process. But only the irrational one changes his or her belief according to it.

Preferences and beliefs

Assume there are two investors $\{1, 2\}$ and both of them have the utility function

$$u^{(1)}(c) = u^{(2)}(c) = \frac{c^{1-\gamma}}{1-\gamma},\ \gamma > 0.^6$$

Investors reaction process to news sentiment (η_t) is given by

$$\begin{aligned}\delta_t^{(1)} &= 0,\ \text{for rational trader and}\\ \delta_t^{(2)} &= \xi^+\eta_t^+ - \xi^-\eta_t^-, \text{for irrational trader,}\end{aligned}$$

where ξ^+, ξ^- are some positive numbers denoting the strength of reaction to positive and negative news sentiment respectively. Investors' beliefs are the probability measures $\mathbb{P}^{(i)}, i = 1, 2$ under which

$$W_t^{(i)} = W_t - \int_0^t \delta_s^{(i)} ds$$

is a standard Brownian motion. Thus for the rational trader 1, $\mathbb{P}^{(1)}$ is same with $\mathbb{P}$ and dividend process follows the same dynamic as in Section 4.2. For the irrational trader 2, dividend process follows

$$\frac{dD_t}{D_t} = (\mu + \delta_t)dt + \sigma dW_t^{(2)}.$$

When δ_t is positive, the irrational trader is optimistic on the expected growth rate of the aggregate dividend. Conversely, δ_t is negative when he or she is pessimistic. Both of the rational and irrational traders maximize expected utility under their own beliefs:

$$\mathbb{E}^{(i)}\left[\int_0^T e^{-\rho t} u_i(c_t) dt\right],$$

where $\mathbb{E}^{(i)}[\cdot]$ denotes the expectation under probability measure $\mathbb{P}^{(i)}$. Denote by

$$\Delta_t^{(i)} = \left(\frac{d\mathbb{P}^{(i)}}{d\mathbb{P}}\right)_t,\ i = 1, 2$$

[6] In case $\gamma = 1$, $u(c) = \ln c$.

the Radon-Nikodym derivative of probability measure $\mathbb{P}^{(i)}$ to $\mathbb{P}$, then investors objective can be written under probability measure $\mathbb{P}$,

$$\max \mathbb{E}\left[\int_0^T e^{-\rho t}\Delta_t^{(i)} u_i(c_t)dt\right],$$

subject to budget constraint

$$\mathbb{E}\left[\int_0^T p_t c_t dt\right] \leq q^{(i)}\mathbb{E}\left[\int_0^T p_t D_t dt\right],$$

where p_t is the price kernel under $\mathbb{P}$ and $\Delta_t^{(i)}$ follows

$$\begin{aligned}&\Delta_t^{(1)} = 1, \text{ and}\\ &d\Delta_t^{(2)} = \Delta_t^{(2)}\delta_t dt.\end{aligned}$$

To simplify the notion, we use $\Delta_t := \Delta_t^{(2)}$. [7]

The individual optimization problem above is equivalent to

$$\max \mathbb{E}\left[\left(\int_0^T e^{-\rho t}\Delta_t^{(i)} u_i(c_t) - \lambda^{(i)} p_t c_t\right) dt\right],$$

with

$$\mathbb{E}\left[\int_0^T p_t c_t dt\right] = q^{(i)}\mathbb{E}\left[\int_0^T p_t D_t dt\right],$$

for some Lagrange multiplier $\lambda^{(i)} > 0$.

2.3.2 The Equilibrium

Definition 1. A *competitive equilibrium* is a collection of $\left\{(B_t, S_t), (c_t^{(i)}), (\theta_t^{(i)})\right\}$ satisfying:

1. each of the investors optimizes his or her trading strategy and maximizes the individual expected utility.

[7]The choice of the form of η and δ is partially coming from avoiding non-smoothness.

2. all investors are price taker, that is, they take the same price process.

3. all markets are clear.

The following proposition shows the solution of the equilibrium allocation by using the standard approach as in *Duffie 2001*.

Proposition 3. *The equilibrium allocation between the two traders is given by*

$$c_t^{(1)} = \frac{1}{1 + \left(\frac{\lambda^{(1)}}{\lambda^{(2)}}\Delta_t\right)^{\frac{1}{\gamma}}} D_t$$

$$c_t^{(2)} = \frac{\left(\frac{\lambda^{(1)}}{\lambda^{(2)}}\Delta_t\right)^{\frac{1}{\gamma}}}{1 + \left(\frac{\lambda^{(1)}}{\lambda^{(2)}}\Delta_t\right)^{\frac{1}{\gamma}}} D_t$$

and the price kernel can be expressed by

$$p_t = e^{-\rho t}\left[\left(\frac{1}{\lambda^{(1)}}\right)^{\frac{1}{\gamma}} + \left(\frac{\Delta_t}{\lambda^{(2)}}\right)^{\frac{1}{\gamma}}\right]^{\gamma} D_t^{-\gamma},$$

where $\lambda^{(1)}, \lambda^{(2)}$ are the Lagrange multiplier of the budget constraints for each of the investors respectively.

2.4 Empirical Study

2.4.1 Data

The empirical analysis in this paper utilizes the RavenPack News Analytics (Dow Jones Edition). RavenPack News Analytics delivers a company-level record for each news story analyzed. Each record contains fields including time stamp, company identifier, relevance, event novelty, event sentiment and event category among others (See Table 2.1 for a description). The data is available for the period from January 2002 through December 2011.

The news items those are used in my analysis are filtered by the following conditions: ENS=100, G_ENS=100 and Relevance is larger than 90. That

Table 2.1: RavenPack News Annalytics Data Field Description

Field	Description
Time Stamp	The Data/Time (with millisecond accuracy) at which the news item was recorded.
Company	The company identifier related to the news item.
Relevance	A score (0-100) that indicates how strongly related the company is to the underlying news story.
Category	An element or tag representing a company-specific news announcement or formal event. Relevant stories about companies are classified into a set of predefined event categories following the RavenPack taxonomy.
ESS - Event Sentiment Score	A granular score (0-100) that represents the news sentiment for a given company by measuring various proxies sampled from the news.
ENS - Event Novelty Score	A score between 0 and 100 that represents how ”new” or novel a news story is within a 24-hour time window across all news stories.
G_ENS - Global Event Novelty Score	A score (0-100) that represents how ”new” or novel a news story is within a 24-hour time window across all news providers covered by RavenPack.

is, only the news stories of high novelty and high relevance are chosen for my analysis. The reason to choose new items of high novelty is to avoid additional noises brought by random citation and reprints. The trade-off is that it is possible that the dimension of reprints reflects how much media cares about the news and contains valuable information. The reason to filter the news items by relevance is that 80% of news items in our dataset hold low relevance (< 85) which could be adding noises .

Our universe includes a cross-section of 330 firms. The total number of news items used for the analysis is 504480.

Another data source is Bloomberg. Individual equity characteristics (market size, PE ratio, dividend, volatility, sector) are obtained from Bloomberg.

2.4.2 Empirical Analysis

Construction of News Sentiment Measure

Investor sentiment is measured by the sentiment of news. Individual sentiment for stock i is constructed by

$$Sent_t^{(i)} = \sum_{\text{news } j} \frac{ESS_j - 50}{50} \mathbf{1}_{t_j \in (t-1,t]} \mathbf{1}_{Relevance(i,j)>90},$$

where $Relevance(i, j)$ is the relevance between news j and stock i.

For daily-based sentiment, the time interval $(t-1, t]$ is from 4:00pm of day $t-1$ to 4:00pm of day t. As in Figure 2.2 and Figure 2.3, the seasonality pattern of the number of news items can be observed. In these figures, we can see, most news firstly revealed during 7:00pm- 10:00pm, which is surprising, given the firm decisions/announcements should be made and revealed in the working hours. To this phenomena, our explanation is, some reviews to the daytime market performance maybe considered as innovated news, however, these news have been revealed/even priced before that. To decrease this effect(can not remove), we might need to consider the categories in the future work.

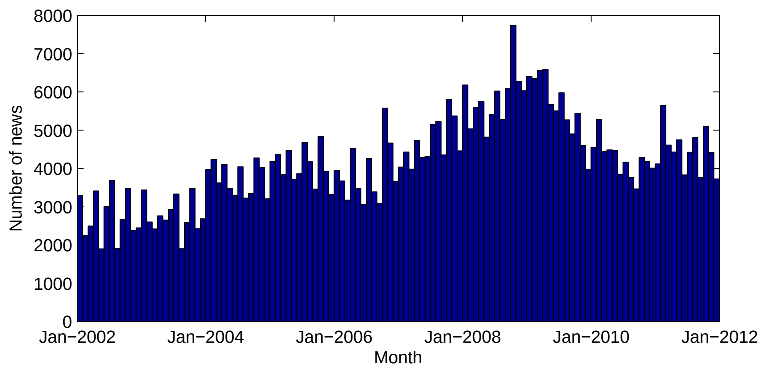

Figure 2.2: The number of news items in each month from Jan-2002 through Dec-2011

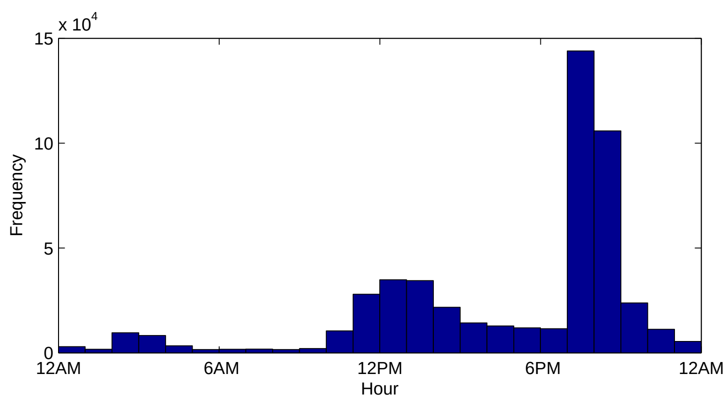

Figure 2.3: The number of news items during 24 hours

Portfolio sentiment is the sum of capital size weighted sentiments of portfolio components,

$$Sent_t^{(p)} = \sum_{i \in \text{portfolio } p} w_t^{(i)} Sent_t^{(i)},$$

where weights $w_t^{(i)}$'s are rebalanced monthly, at the beginning of each month.

Market sentiment $Sent_t^{(M)}$ is the sum of capital size weighted sentiment of all stocks in my universe.

In addition, positive sentiment and negative sentiment are also be constructed separately. That is,

$$\begin{aligned} Sent_t^{+,(i)} &= \sum_{\text{news } j} \frac{ESS_j - 50}{50} \mathbf{1}_{ESS_j > 50} \mathbf{1}_{t_j \in (t-1,t]} \mathbf{1}_{Relevance(i,j) > 90}, \\ Sent_t^{-,(i)} &= \sum_{\text{news } j} \frac{50 - ESS_j}{50} \mathbf{1}_{ESS_j < 50} \mathbf{1}_{t_j \in (t-1,t]} \mathbf{1}_{Relevance(i,j) > 90}. \end{aligned}$$

Similarly, positive sentiment and negative sentiment of market and portfolio, $Sent_t^{+,(M)}, Sent_t^{-,(M)}, Sent_t^{+,(p)}, Sent_t^{-,(p)}$, are also constructed.

The monthly market sentiments are shown in Figure 2.4. In my analysis, positive sentiment and negative sentiment are both investigated. From Figure 2.4, roughly speaking, positive sentiment and negative sentiment are showing very different features. Positive sentiment is keeping at a certain level around 3.5. Negative sentiment increases greatly during September 2008. There are always be relatively stable volume of positive news across the whole time period from 2002 to 2011. It is showing that low level of negative sentiment dose not imply high level of positive sentiment. In addition, the correlation between positive sentiment and negative sentiment is -0.06 with p-value 0.50. Thus, it is reasonable to believe positive sentiment and negative sentiment have different characteristics and should be assumed to have different relation with market.

Sorts

Portfolios are constructed based on 5 criterions: Capital Size, PE ratio, dividend, volatility and Sector. On each criterion, 10 portfolios are formed

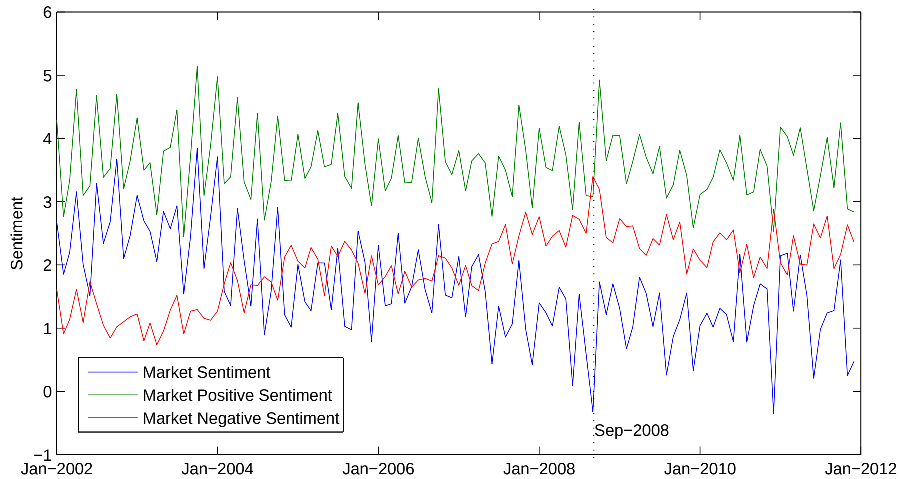

Figure 2.4: Monthly Market Sentiment Measures

with equally weight. Thus, in total, 50 portfolios are analyzed in this paper. Table 2.2 and Table 2.3 shows the conditional characteristics effects in a simple and non parametric way.

The first rows of Table 2.2 show the effect of size conditioned on market sentiment. The rows reveals that the size effect appears in both of the months with high and low sentiment. Small size firms tend to have higher return than large size firm. The average return is higher when sentiment is high than it is when sentiment is low, across all of the cross-section. However the difference tends to be large for small size firms.

The effect of PE ratio does not show any strong pattern, as in Table 2.2.

As shown in Table 2.2, the effect of dividend is not strong conditioned on high sentiment periods. However in the months with low sentiment, the average return tends to be high for the firms of high dividend.

Table 2.2 also shows the effect of volatility shows strong patterns during the month is of high and low sentiment. When sentiment is high, high volatility firms tend to have high returns. But when sentiment is low, high volatility firms tend to have low returns on contrary.

It also can be told by Table 2.2 that for some sectors, the difference between average return in high and low sentiment is quite large (≥ 3.2). These sectors includes Energy, Financials, and Technology. For sector like Health Care and Utilities, the difference is small (< 1.5).

Noticing that Table 2.2 shows the return and sentiment of the same month, and it can not indicate any predictability, we present the monthly return and the previous moth sentiment in Table 2.3. The first rows of Table 2.3 show the effect of size conditioned on market negative sentiment. The rows reveals that the size effect appears not so obvious as in Table 2.2. Although it still holds that at the beginning of the month low negative sentiment firms will be followed by higher return, there is no pattern show that small firms will be ore sensitive to sentiment.

The second rows of Table 2.3 show the sentiment effect on different PE ratio firms. The effect of PE ratio does not appear in the months of high negative sentiment. However, in the months of low negative sentiment, the average return of high PE ratio firms tends to be smaller than the return of low PE ratio firms.

The effect of dividend conditioned on negative sentiment, as shown in Table 2.3, is not found.

The effect of volatility does not show a strong pattern when the previous month is of high negative sentiment. However, it appears during the months with low negative sentiment,high volatility firms tend to have a higher return.

It also can be told by Table 2.3 that for some sectors, the difference between average subsequent return in high and low negative sentiment is quite large (< -1.5). These sectors includes Communications, Materials, and Utilities. For sector like Consumer Staples, Health Care and Technology, the difference is small (> -0.8).

Table 2.2: Same Month Return by Market Sentiment and Firm Characteristics

		Portfolio (High to Low)										Comparison		
	$Sent_t^{(M)}$	1	2	3	4	5	6	7	8	9	10	10 - 1	10 - 5	5 - 1
Firm Size	>Median	1.51	1.94	1.87	1.73	2.20	2.36	2.33	2.41	2.53	3.66	2.15	1.46	0.69
	<Median	-0.53	-0.40	-0.14	-0.25	-0.21	-0.23	-0.26	-0.08	-0.19	-0.01	0.52	0.20	0.32
	Difference	2.04	2.34	2.01	1.98	2.41	2.59	2.58	2.50	2.72	3.67	1.63	1.26	0.36
PE Ratio	>Median	2.13	1.70	1.76	1.82	2.06	2.12	2.11	2.40	2.13	3.81	1.68	1.75	-0.07
	<Median	-0.30	-0.38	-0.01	-0.33	-0.06	0.04	-0.24	-0.42	-0.48	-0.24	0.06	-0.18	0.24
	Difference	2.43	2.08	1.77	2.15	2.12	2.08	2.35	2.83	2.61	4.05	1.62	1.93	-0.31
Dividend	>Median	2.52	2.28	2.14	2.08	2.11	2.40	2.20	2.28	2.30	2.19	-0.32	0.08	-0.41
	<Median	-0.10	-0.09	-0.12	-0.08	-0.15	-0.34	-0.26	-0.49	-0.55	-0.45	-0.36	-0.30	-0.06
	Difference	2.62	2.36	2.25	2.16	2.26	2.74	2.46	2.77	2.85	2.65	0.03	0.38	-0.35
Volatility	>Median	4.13	3.26	2.68	2.27	2.25	2.23	1.77	1.62	1.33	0.81	-3.33	-1.45	-1.88
	<Median	-0.39	-0.56	-0.47	-0.64	-0.29	-0.19	-0.34	0.05	0.14	0.04	0.43	0.33	0.10
	Difference	4.52	3.82	3.15	2.91	2.54	2.42	2.11	1.57	1.19	0.77	-3.75	-1.78	-1.98
Sector	>Median	1.83	2.21	1.40	3.30	2.53	1.58	2.42	2.70	2.49	1.60			
	<Median	-0.63	-0.31	0.19	0.09	-0.76	0.15	-0.02	-0.35	-0.71	0.32			
	Difference	2.46	2.52	1.21	3.21	3.29	1.43	2.44	3.04	3.20	1.28			

Table 2.3: Next Month Return by Market Negative Sentiment and Firm Characteristics

		Portfolio (High to Low)										Comparison		
	$Sent_{t-1}^{-,(M)}$	1	2	3	4	5	6	7	8	9	10	10 - 1	10 - 5	5 - 1
Firm Size	>Median	0.26	0.27	0.44	0.27	0.34	0.25	0.56	0.32	0.38	1.23	0.97	0.89	0.08
	<Median	0.72	1.27	1.29	1.21	1.65	1.88	1.51	2.01	1.96	2.42	1.70	0.77	0.93
	Difference	-0.46	-1.00	-0.86	-0.94	-1.32	-1.63	-0.95	-1.69	-1.58	-1.19	-0.73	0.12	-0.85
PE Ratio	>Median	0.46	0.29	0.40	0.10	0.35	0.45	0.24	0.30	0.12	1.20	0.73	0.85	-0.12
	<Median	1.37	1.03	1.35	1.40	1.66	1.70	1.63	1.68	1.53	2.37	1.00	0.72	0.29
	Difference	-0.91	-0.73	-0.95	-1.30	-1.31	-1.25	-1.40	-1.38	-1.42	-1.18	-0.27	0.13	-0.40
Dividend	>Median	0.56	0.48	0.42	0.46	0.63	0.45	0.53	0.13	0.28	0.27	-0.29	-0.36	0.08
	<Median	1.87	1.71	1.60	1.55	1.32	1.61	1.42	1.65	1.47	1.47	-0.39	0.15	-0.54
	Difference	-1.31	-1.22	-1.18	-1.09	-0.69	-1.16	-0.90	-1.52	-1.19	-1.20	0.11	-0.51	0.62
Volatility	>Median	1.48	0.58	0.43	0.06	0.18	0.52	0.11	0.20	0.32	0.07	-1.41	-0.11	-1.30
	<Median	2.27	2.12	1.77	1.57	1.78	1.52	1.31	1.47	1.14	0.77	-1.50	-1.01	-0.49
	Difference	-0.79	-1.53	-1.34	-1.51	-1.60	-1.00	-1.20	-1.27	-0.82	-0.70	0.09	0.90	-0.81
Sector	>Median	-0.18	0.35	0.54	1.18	0.31	0.55	0.61	0.26	0.52	0.12			
	<Median	1.39	1.56	1.06	2.20	1.47	1.18	1.79	2.09	1.27	1.80			
	Difference	-1.57	-1.21	-0.52	-1.02	-1.16	-0.63	-1.18	-1.82	-0.75	-1.67			

2.4.3 Long-Short Portfolio Regressions

Based on the previous discussion, we can see the sentiment might have predictability on the return of a long-short portfolio. To incorporate continuous features of data, and perform significant test, we run the following regression models.

$$R_t^{p(X=High)} - R_t^{p(X=Low)} = c + \beta MRP_t + sSMB_t + hHML_t + mMOM_t + d_+ Sent_{t-1}^{+(M)} + \epsilon_t. \quad (2.1)$$

$$R_t^{p(X=High)} - R_t^{p(X=Low)} = c + \beta MRP_t + sSMB_t + hHML_t + mMOM_t + d_- Sent_{t-1}^{-(M)} + \epsilon_t. \quad (2.2)$$

The dependent variable on LHS is monthly return of a long-short portfolio based on firm size, PE ratio, dividend and volatility, respectively. The regressors are widely used comovement factors, plus sentiment factor.

The variable MRP_t is excess return of SP 500 over the risk-free rate. The variable SMB_t is the return on portfolios of small and big ME stocks that is separate from returns on HML_t, where HML_t is constructed to isolate the difference between high and low BE/ME portfolios. The variable MOM_t is the return on high-momentum stocks minus the return on low-momentum stocks.The variable $Sent_{t-1}^{-(M)}$ is the negative market sentiment factor of last month. And the variable $Sent_{t-1}^{+(M)}$ is the positive market sentiment factor of last month.

Table 2.4 shows the results. The results provide formal support to our preliminary impressions from the sorts. From Table 2.4 , we can see positive news and negative news have different effect on the market. Positive sentiment have significant effect on long-short portfolio based on firm size and dividend. Negative sentiment have significant effect on long-short portfolio based on PE ratio and volatility.

As show in first rows of Table 2.4, small firms are more sensitive to positive news. When the positive sentiment of the beginning of the month is strong, small firms is tend to have a better market performance than large firms. However, there is no similar conclusion on negative sentiment.

In the second rows of Table 2.4, we can see, firms with a lower P/E ratio, tend to have a low return when the negative sentiment at the beginning of the month is strong. Since PE ratio reflect the investors' expectation of the firms future performance, this effect may be caused by that during the pessimistic period, investor tend to invest on firms they consider will grow in the future. However, there is no similar conclusion on positive sentiment.

In the third rows, low dividend firms tend to have a higher return if the positive sentiment is high at the beginning of the month. This maybe reflect the under reaction of market. There is no similar conclusion on positive sentiment.(Considering the dividend may be influence by company strategies so that it can not reflect the profitability of a company. In future work, we may also consider decile portfolios of earnings and compare the result with dividend decile portfolios.)

In the fourth rows of Table 2.4, high volatility firms may have a higher return when the beginning of the month has a strong negative market sentiment. Positive news doesn't have similar effect.

From second rows and fourth rows, we can see, firms with high P/E ratio and high volatility tend to have a higher return during the low sentiment market. This is consistent with our intuition, that when the market is bad, investor expect higher risk premium. This is to say, investors might be more risk averse during the bearish and these effect can not be only priced in market portfolio betas and the HML factor. If this is the reason, it is not hard to understand why these two long-short portfolios are not significant effected by positive sentiment.

However, small firms do not show the similar results. Another phenomena maybe related to this phenomena is that, small firms have very little news.

Table 2.4: Long-Short Regressions

		$Sent_{t-1}^{+,(M)}$		$Sent_{t-1}^{-,(M)}$	
		$d_+ \times 10^2$	p-value	$d_- \times 10^2$	p-value
Firm Size	High - Low	-1.01	(0.03)	0.08	(0.87)
	High - Medium	-0.24	(0.49)	0.22	(0.52)
	Medium - Low	-0.77	(0.05)	-0.14	(0.73)
PE Ratio	High - Low	0.53	(0.36)	1.01	(0.09)
	High - Medium	-0.27	(0.10)	1.10	(0.01)
	Medium - Low	0.81	(0.16)	-0.09	(0.88)
Dividend	High - Low	-1.64	(0.02)	-0.54	(0.47)
	High - Medium	-0.81	(0.17)	-0.92	(0.13)
	Medium - Low	-0.83	(0.06)	0.38	(0.40)
Volatility	High - Low	-0.63	(0.37)	1.43	(0.05)
	High - Medium	-0.80	(0.24)	1.72	(0.01)
	Medium - Low	0.16	(0.71)	-0.29	(0.07)

Chapter 3
Long Range Dependence for Different Market Period and Cross-sectional Assets

3.1 Introduction and Motivation

We see from last chapter that market efficiency varies in different market situation, say in positive market sentiment, or in negative market sentiment. A natural hypothesis would be, the LRD parameters may also change in different market situation.

To better understand the LRD component to the return process, in this chapter, we will focus on S&P 500 Index, and to explore the LRD behavior of the index returns in different market situation with different sampling frequency. This is a key component of determining the optimal investment strategies and portfolio management because of its relevance to market efficiency

Some relevant researches can be found in the literature. Researchers has used several methods such as the rescaledrange (R/S), the modified R/S test, the Geweke and Porter-Hudak (GPH) method, the Gaussian semi parametric (GSP) approach and the exact maximum likelihood (EML) method, among others to test the long-memory hypothesis.

For example, Sensoy 2013 studies the time-varying efficiency of nineteen members of the Federation of Euro-Asian Stock Exchanges by generalized Hurst exponent(GHE) analysis of daily data with a rolling window technique. Cajueiroa and Tabakb 2008 also uses GHE to test for long-range dependence in equity returns and volatility, ranking stock market indices in terms of weak form efficiency. Walid, Hammoudeh, and Yoon 2014 analyses long memory properties of four major foreign exchange markets of the world oil exporter Saudi Arabia, using the FARIMA-FIGARCH model under several global events. Kang, Cheong, and Yoon (2011) investigate the impacts of structural changes on volatility persistence, and then incorporated these impacts into the bivariate estimation in order to understand the information flow and volatility transmission in two crude oil markets. Following the same line, Arouri, Hammoudeh, Lahiani, and Nguyen (2012) examine the potential of structural changes and long memory properties in returns and volatility of four major precious metal commodities (gold, silver, platinum and palladium) traded on the COMEX markets and show that dual long memory is found to be adequately captured by an FARIMA-FIGARCH model. In addition, evidence shows that conditional volatility of precious metals is better explained by long memory than by structural breaks.

The most recognized model used to examine LRD in the conditional mean is the FARIMA model, which is introduced in Chapter 1. With the seminar work of Engle (1982) and Bollerslev (1986), the GARCH-family processes become the most popular processes to capture persistence and volatility clustering. Following this research line, the analysis of LRD has been extended from focusing on persistence in the conditional mean to also examining persistence in the conditional volatility of financial time series, which is conducted by employing FIGARCH model.

In this study, we first focus on S&P 500 Index(SPY),to test the relevance of the FARIMA-FIGARCH-GH model for intra-day index returns for different frequencies(1-minute, 5-minute, 10-minute, and 30-minute). We employ two periods: a turbulent period (from 07/2008 to 12/2008) and a calm period (from 01/2013 to 06/2013). First we test the FARIMA-FIGARCH-GH against the two nested models ARIMA-IGARCH-GH and ARIMA-IGARCH-Gaussian. Then we use the likelihood-ratio test to see which model should be employed. The first comparison will answer the question if the fractional component is statistically significant. The second one is about the distribution of the residual. Finally we test the residuals for fractionality (apply the Hurst index test). This is to see if there is any residual LRD in the residuals (to the extent that the Hurst index can capture it). Later in this chapter we move to stock level to see for individual stocks with different capital size and P/E ratio, whether the findings in index level still hold.

3.2 Data Preparation

The dataset is based on the TAQ CD-ROMs of the NYSE and contains intraday trade and quote data for SPY. The finest resolution of the data is 1 second. We choose two period data: a turbulent period (from 07/2008 to 12/2008) and a calm period (from 01/2013 to 06/2013). We clean the data following the procedure described in Chapter 1 and construct the 1-minute frequency return data data $\{y_t^1\}, t = 1, \dots .N$.

Here to avoid the crossing-day problem, we only calculate 390 one-minute returns in one day, instead of 391 minutes.

Further more, intraday data exhibits intraday seasonality, well know as U-shape. Figure 3.1 and Figure 3.2 show the intraday pattern of the 1-min return magnitude and 1-min return of variance, of SPY.

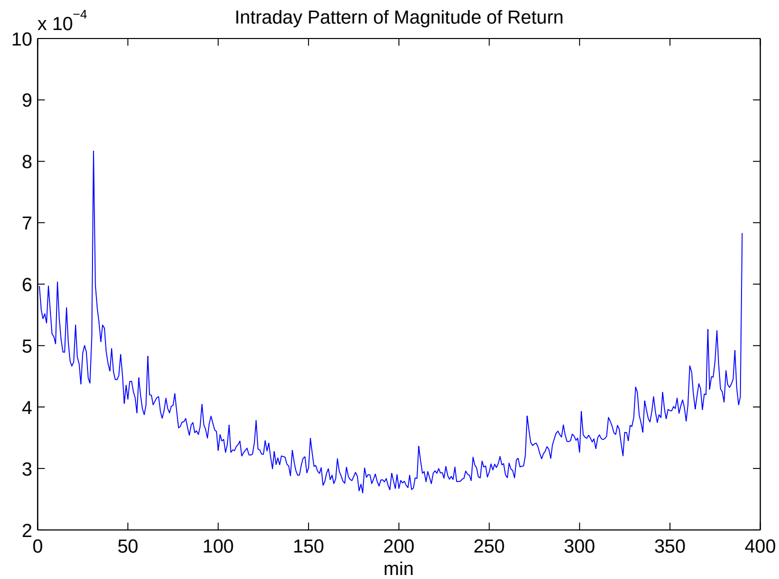

Figure 3.1: Intraday Pattern of Return Magnitude

To avoid such 'time of the day effect', we employ the method in Giot (2005), by assuming a deterministic seasonality in the intraday volatility.

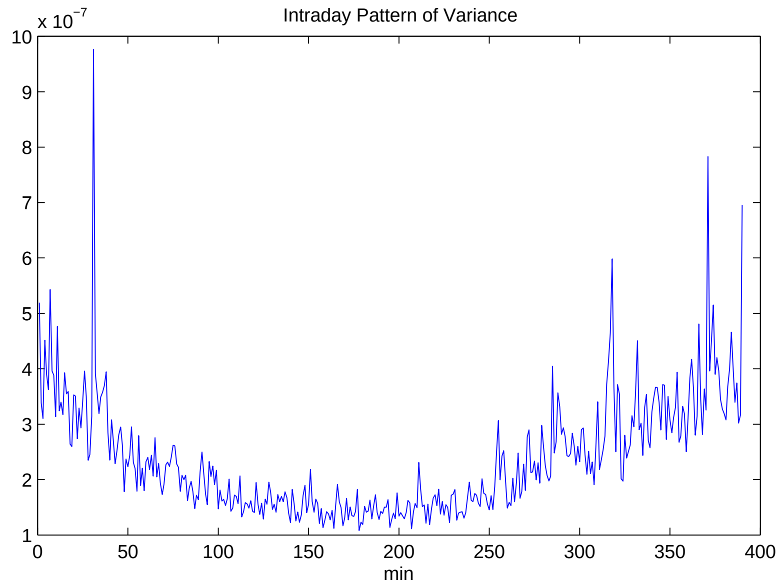

Figure 3.2: Intraday Pattern of Variance

We deseasonalize return $\{y_t^1\}$

$$r_t^1 = \frac{y_t^1}{\phi_i} \tag{3.1}$$

where ϕ_i is the deterministic intraday seasonal component, defined as the expected volatility conditioned on time-of-day, where the expectation is computed by averaging the squared raw returns over the i-th minute of each day, given y_t^1 is the return of i-th minute of the day.

According to Giot (2005), Similar deterministic techniques are also used in Andersen and Bollerslev (1997, 1998, 1999). Beltratti and Morana (1999) define a stochastic seasonality for the intraday volatility (which is much more complicated to estimate) and the results are hardly better than those obtained with the more simple deterministic seasonality.

Then we get 5-min $\{y_t^5\}$, 10-min $\{y_t^{10}\}$ and 30-min $\{r_t^{30}\}$ by simply adding up the 1-minute return in different frequency.

3.3 Models and Tests

We test the FARIMA-FIGARCH-GH against the two nested models ARIMA-IGARCH-GH and ARIMA-IGARCH-Gaussian.

3.3.1 FARIMA-FIGARCH-GH

The FARIMA-FIGARCH structure is summarized as follows

$$\phi_a(L)(1-L)^{d_0}(r_t-\mu)=\theta_m(L)\epsilon_t, \tag{3.2}$$

$$\psi_p(L)(1-L)^{d}(\epsilon_t^2-\sigma^2)=(1-\beta(L))v_t, \tag{3.3}$$

$$\epsilon_t=\sqrt{h_t}u_t, \tag{3.4}$$

$$v_t=\epsilon_t^2-h_t, \tag{3.5}$$

where

$$\phi_a(L)=1-\sum_{j=1}^{a}\phi_j L_j, \tag{3.6}$$

$$\theta_m(L)=1+\sum_{j=1}^{m}\theta_j L_j, \tag{3.7}$$

$$\psi_p(L)=1-\sum_{j=1}^{p}\psi_j L_j, \tag{3.8}$$

$$\beta_q(L)=\sum_{j=1}^{q}\beta_j L_j, \tag{3.9}$$

$$\tag{3.10}$$

and μ is the unconditional mean of r_t, σ^2 is the unconditional variance of ϵ_t, ϵ_t are the innovations and u_t are the standardized residuals with mean 0 and unit variance.

For estimation we preset a, m, p, q all equal to 1.

The generalized hyperbolic (GH) distributions can be written in the following way.

A random vector X is said to have a multivariate GH distribution if X can be expressed as a normal mean-variance mixture distribution

$$X \stackrel{d}{=} \mu + W\gamma + \sqrt{W}AZ, \tag{3.11}$$

where $Z \sim \mathcal{N}_k(0, I_k)$ is standard k-dimensional normal distributed random vector, A is a $d \times k$ real matrix, $\mu, \gamma \in \mathbb{R}^d$, and $W \geq 0$ is a scalar-valued random variable independent of Z and having a Generalized Inverse Gaussian distribution $\mathrm{GIG}(\lambda, \chi, \psi)$.

A random variable W is said to have a generalized inverse Gaussian (GIG) distribution if its probability density is given by

$$f_{\mathrm{GIG}}(x; \lambda, \chi, \psi) = \frac{\chi^{-\lambda}(\sqrt{\chi\psi})^{\lambda}}{2K_\lambda(\sqrt{\chi\psi})} x^{\lambda-1} \exp\left(-\frac{1}{2}\left(\chi x^{-1} + \psi x\right)\right), \tag{3.12}$$

for $x > 0$, and where $\chi, \psi > 0$, and K_λ is a modified Bessel function of the third kind with index λ. The parameters satisfy $\chi > 0, \psi \geq 0$ if $\lambda < 0$, $\chi \geq 0, \psi > 0$ if $\lambda > 0$, and $\chi > 0, \psi > 0$ if $\lambda = 0$.

By letting $\Sigma = AA'$, we denote a GH distributed random vector by

$$X \sim \mathrm{GH}_d(\lambda, \chi, \psi, \mu, \gamma, \Sigma). \tag{3.13}$$

The univariate GH distribution is obtain with $d = 1$. The joint density when Σ is not singular is given by, for $x > 0$,

$$f_{\mathrm{GH}}(x) = c \frac{K_{\lambda-\frac{d}{2}}\left(\sqrt{(\chi + (x-\mu)'\Sigma^{-1}(x-\mu))(\psi + \gamma'\Sigma^{-1}\gamma)}\right) e^{(x-\mu)'\Sigma^{-1}\gamma}}{\left(\sqrt{(\chi + (x-\mu)'\Sigma^{-1}(x-\mu))(\psi + \gamma'\Sigma^{-1}\gamma)}\right)^{\frac{d}{2}-\lambda}}, \tag{3.14}$$

where the normalizing constant is given by

$$c = \frac{\left(\sqrt{\chi\psi}\right)^{-\lambda}\left(\psi + \gamma'\Sigma^{-1}\gamma\right)^{(d/2)-\lambda}}{(2\pi)^{d/2}|\Sigma|^{1/2}K_\lambda\left(\sqrt{\chi\psi}\right)}. \tag{3.15}$$

For comparison, we use ARIMA(1,0,1)-IGARCH(1,0,1)-GH, which is to say for the previous FIARMA-FIGARCH-GH, we specifically set d, d_0 all equal to 0.

And to get ARIMA(1,0,1)-IGARCH(1,0,1)-Gaussian, we simply replace the innovation by Gaussian distribution described in 1.5.2.

3.3.2 Hurst Exponent

The Hurst exponent is employed as a measure of long range dependence of time series, usually noted as H.

As presented by Hurst(1951), H, is defined in terms of the asymptotic behavior of the rescaled range as a function of the time span of a time series as follows,

$$E\left[\frac{R(n)}{S(n)}\right] = Cn^H \text{ as } n \to \infty \tag{3.16}$$

where,

- R(n) is the range of the first n values, and S(n) is their standard deviation
- $E[x]$ is the expected value
- n is the time span of the observation (number of data points in a time series)
- C is a constant.

To estimate Hurst Exponent, we us the R/S analysis of Hurst proposed in Hurst 1951, corrected for small sample bias proposed in Weron (2002), with following steps.

1. We begin with dividing the time series $\{X_i\}, i = 1, 2...L.$, which is of interest, of length L, into d subseries of length n $\{X_j^k\}, j = 1, 2...n, k = 1, 2...d$.

2. Then for each subseries $\{X_j^k\}$, we find its mean E_k and standard deviation S_k.

3. Demean $\{X_j^k\}$ to get $\{\bar{X}_j^k\}$.

4. creat a cumulatie timeseries $\{Y_j^k\}$, $Y_j^k = \sum_{m=1}^{j} \bar{X}_m^k, j = 1, ...n$.

5. find the range of $\{Y_j^k\}$ for each k, $R_k = \max\{Y_1^k, ...Y_n^k\} - \min\{Y_1^k, ...Y_n^k\}$.

6. rescale range R_k by using standard deviation, R_k/S_k

7. calculate the mean value of the rescaled range for all subseries of length n

$$(R/S)_n = \frac{1}{d}\sum_{l=1}^{d} R_l/S_l \tag{3.17}$$

8. Mandelbrot (1975) shows

$$(R/S)_n \sim cn^H. \tag{3.18}$$

Therefore, H can be given by a simple linear regression,

$$\log(R/S)_n = \log c + H \log n \tag{3.19}$$

Weron (2002) also documents, for small n there is a significant deviation from the 0.5 slope. To solve this problem, he suggests to use the estimation from Anis and Lloyd (1975), in which they theoretically estimated the values of the R/S statistic to be:

$$E[R(n)/S(n)] = \begin{cases} \frac{\Gamma(\frac{n-1}{2})}{\sqrt{\pi}\Gamma(\frac{n}{2})} \sum_{i=1}^{n-1} \sqrt{\frac{n-i}{i}}, & \text{for } n \leq 340 \\ \frac{1}{\sqrt{n\frac{\pi}{2}}} \sum_{i=1}^{n-1} \sqrt{\frac{n-i}{i}}, & \text{for } n > 340 \end{cases}, \tag{3.20}$$

where Γ is the gamma function, with a minor adjustment,

$$E[R(n)/S(n)] = \begin{cases} \frac{n-1/2}{n} \frac{\Gamma(\frac{n-1}{2})}{\sqrt{\pi}\Gamma(\frac{n}{2})} \sum_{i=1}^{n-1} \sqrt{\frac{n-i}{i}}, & \text{for } n \leq 340 \\ \frac{n-1/2}{n} \frac{1}{\sqrt{n\frac{\pi}{2}}} \sum_{i=1}^{n-1} \sqrt{\frac{n-i}{i}}, & \text{for } n > 340 \end{cases}, \tag{3.21}$$

where the $\frac{n-1/2}{n}$ term was added by Peters to improve the performance for very small n.

If the process is a long memory process then the slope H is greater than 0.5, that is to say $0.5 < H < 1$; if it is anti-persistent or mean reversion then the slope H is less than 0.5, that is to say $0 < H < 0.5$. A value of $H = 0.5$ can indicate a completely uncorrelated series, but in fact it is the value applicable to series for which the autocorrelations at small time lags can be positive or negative but where the absolute values of the autocorrelations decay exponentially quickly to zero.

3.3.3 Likelihood-ratio Test

A likelihood-ratio test is a statistical test used to compare the goodness of fit of two models, one of which (the null model) is a special case of the other (the alternative model).

A statistical model is often a parametrized family of probability density functions or probability mass functions $f(x|\theta)$. A simple hypotheses are

$$H_0 : \theta = \theta_0, \tag{3.22}$$

$$H_1 : \theta = \theta_1. \tag{3.23}$$

Let

$$\Lambda(x) = \frac{L(\theta_0|x)}{L(\theta_1|x)} = \frac{L(\cup_i x_i|\theta_0)}{L(\cup_i x_i|\theta_1)}, \tag{3.24}$$

where $L(\theta|x)$ is the likelihood function, and Λ is the likelihood ratio function for the hypotheses and $\Lambda(x)$ is the likelihood ratio statistic.

Obviously, the likelihood ratio is small if the alternative model is better than the null model and the likelihood ratio test provides the decision rule as follows:

If $\Lambda > c$, do not reject H_0; If $\Lambda <= c$, reject H_0;

with a significance level α, where $P(\Lambda < c|H_0) = \alpha$.

The Neyman-Pearson Lemma, named for Jerzy Neyman and Egon Pearson, shows that when performing a hypothesis test between two simple hypotheses $H_0 : \theta = \theta_0$ and $H_1 : \theta = \theta_1$, the likelihood-ratio test which rejects H_0 in favor of H_1 when

$$\Lambda(x) = \frac{L(x \mid \theta_0)}{L(x \mid \theta_1)} \leq c \tag{3.25}$$

where

$$P(\Lambda(X) \leq c \mid H_0) = \alpha \tag{3.26}$$

the likelihood-ratio test given above is most powerful at significance level α for a threshold c.

3.4 Empirical Results on Index

Table 3.1, 3.2, 3.3, 3.4 show the fitting result from ARIMA(1, 0, 1) - IGARCH (1, 0, 1)- Gaussian, ARIMA(1, 0, 1) - IGARCH (1, 0, 1)- GH and FARIMA(1, d, 1) - FIGARCH (1, d_0, 1)- GH.

Then we run likelihood-ratio test on ARIMA-IGARCH-GH (H_1) against ARIMA-IGARCH-Gaussian (H_0) and FARIMA-FIGARCH-GH (H_1) against ARIMA-IGARCH-GH (H_0). The output $h = 1$ indicates that there is strong evidence suggesting that the unrestricted model (H_1) fits the data better than the restricted model (H_0). Table 3.5 presents likelihood-ratio test result.

From the table, we can see, in all period and all frequency, ARIMA-IGARCH-GH is better than ARIMA-IGARCH-Gaussian, indicating in intraday level, returns have a non-gaussian behavior for certain. On the other hand, FARIMA-FIGARCH-GH outperforms ARIMA-IGARCH-GH in longer frequency,10 - minutes frequency at turbulent period and 10 - min, 30 - min frequency in calm period. This indicates

- return series is more persistent in calm period than turbulent period.

Table 3.1: ARIMA(1, 0, 1) - IGARCH (1, 0, 1)- Gaussian Parameter

Ticker	Time Horizon	Sampling Frequency	ARIMA Parameters			IGARCH Parameters		
			ϕ	θ	μ	β	ψ	σ^2
SPY	Turbulent Period (2008-July-01 -2008-Dec-31)	1min	-0.0391	0.0171	-0.0025	0.0414	0.9586	0.0027
SPY		5min	-0.8761	0.8692	-0.0115	0.0550	0.9450	0.0055
SPY		10min	0.9130	-0.9066	-0.0007	0.0674	0.9318	0.009
SPY		30min	-0.8994	0.8804	-0.0190	0.0895	0.910	-0.0135
SPY	Calm Period (2013-Jan-02 -2013-June-28)	1min	0.4910	-0.5167	0.0044	0.0423	0.9547	0.0009
SPY		5min	0.2951	-0.3272	0.0128	0.0722	0.9130	0.0044
SPY		10min	-0.4347	0.3883	0.0274	0.0760	0.9084	0.0048
SPY		30min	-0.5486	0.5413	0.0470	0.0839	0.8873	0.0076

Table 3.2: GH Innovation Fitting of ARIMA(1, 0, 1) - IGARCH (1, 0, 1)

Ticker	Time Horizon	Sampling Frequency	GH Fitting(ARIMA-IGARCH)					
			λ	χ	ψ	μ	σ	γ
SPY	Turbulent Period (2008-July-01 -2008-Dec-31)	1min	1.2626	1.1109	4.0438	-0.0271	1	0.0209
SPY		5min	0.1321	2.4001	3.5171	0.0247	1	-0.0241
SPY		10min	-4.2172	6.5186	0.0114	0.0317	1	-0.0399
SPY		30min	0.3355	2.0527	3.4849	0.0074	1	-0.0107
SPY	Calm Period (2013-Jan-02 -2013-June-28)	1min	0.0842	0.7187	1.6872	0.0254	1	-0.0254
SPY		5min	-3.1322	3.8336	0.0138	0.1059	1	-0.1403
SPY		10min	-0.8542	1.5746	1.2743	0.1056	1	-0.1494
SPY		30min	-1.7975	2.3857	0.8087	0.2927	1	-0.3474

Table 3.3: FARIMA(1, d, 1) - FIGARCH (1, d_0, 1)- Gaussian Parameter

Ticker	Time Horizon	Sampling Frequency	FARIMA Parameters				FIGARCH Parameter			
			d	ϕ	θ	μ	d_0	β	ψ	σ^2
SPY	Turbulent Period (2008-July-01 -2008-Dec-31)	1min	-0.0027	-0.0248	0.0048	-0.0036	0.5358	0.7849	0.3334	0.0021
SPY		5min	0.0362	0.2582	0.3108	-0.0067	0.5745	0.7339	0.2054	0.0305
SPY		10min	0.0256	-0.329	1.34e-05	-0.0034	0.5125	0.6091	0.1350	0.0539
SPY		30min	0.0197	-0.0237	0.0144	-0.0071	0.7697	0.8026	0.1217	-0.0258
SPY	Calm Period (2013-Jan-02 -2013-June-28)	1min	-0.211	-0.0120	-0.0067	0.0084	0.4432	0.7415	0.3850	0.0049
SPY		5min	-0.0114	-0.0161	-0.0046	0.0180	0.3589	0.5356	0.2717	0.014
SPY		10min	-0.0048	-0.3978	0.3548	0.0230	0.3227	0.2904	0.0592	0.0262
SPY		30min	0.0052	-0.0299	0.0105	0.0327	0.3371	0.4941	0.2512	0.0200

Table 3.4: GH Innovation Fitting of ARIMA(1, d, 1) - IGARCH (1, d_0, 1)

Ticker	Time Horizon	Sampling Frequency	GH Fitting(FARIMA-FIGARCH)					
			λ	χ	ψ	μ	σ	γ
SPY	Turbulent Period (2008-July-01 -2008-Dec-31)	1min	1.4549	0.9374	4.5587	-0.0252	1	0.0220
SPY		5min	-0.4400	2.8194	3.2492	-0.0029	1	-0.0072
SPY		10min	-4.6842	6.8985	0.01	-0.0087	1	-0.0083
SPY		30min	0.3987	1.9105	3.5152	-0.0363	1	-0.0339
SPY	Calm Period (2013-Jan-02 -2013-June-28)	1min	0.8192	0.7020	13.1738	0.0183	1	-0.0339
SPY		5min	-3.1695	3.8073	0.01	0.1122	1	-0.1550
SPY		10min	-3.1895	3.9112	0.0101	0.1754	1	-0.0226
SPY		30min	-1.92	2.6387	-0.7483	0.2718	1	-0.3373

- for very high frequency, no matter calm or turbulent period, return series shows less persistency.

It maybe able to be explained from market participants' behaviors. For example, the less persistency in higher frequency suggests a larger noise-information ratio in higher frequency, which makes trading signal harder to be detected and market more efficiency. The larger persistency in calm period may come from attention bias and overreaction during calm period.

We then calculate the Hurst exponent for the original return series and the residual of ARIMA-IGARCH and FARIMA-FIGARCH, on both original and square level. From Table 3.6, 3.7, we can see, the persistency of return mainly comes from the variance level, and FIGARCH can greatly removes the persistency in variance.

3.5 Empirical Results on Stocks

For a further investigation , we will look into the long-range dependence of cross-sectional stocks. We group stocks based on two characteristics: Capital size and P/E ratio. Thus, we pick and sort stocks into 4 groups, that is, {Large Capital(20% quantile), Small Capital(20% quantile)}×{High P/E(20% quantile), Low P/E(20% quantile)}. For each group, 9-14 stocks are selected. The results of the likelihood ratio test for long-range-dependence model are summarized in Table 3.8. Crossing all stocks and time periods, generalized hyperbolic innovation assumption is all accepted comparing to Gaussian innovation.

Fractional integrated models are preferred more in large-capital high-PE stocks comparing to small-capital low-PE ones when time period is calm. When time period is turbulent, the fractional integration is not as signification as calm period, and no stock cross-sectionally prefers long-range dependence than others. 5-min and 10-min returns has stronger long-range dependence comparing to 1-min and 30-min samplings.

Table 3.5: Likelihood-ratio Test Result

Ticker	Time Horizon	Sampling Frequency	ARIMA-IGARCH-GH against ARIMA-IGARCH-Gaussian		FARIMA-FIGARCH-GH against ARIMA-IGARCH-GH	
			H	P-value	H	P-value
SPY	Turbulent Period (2008-July-01 -2008-Dec-31)	1min	1	0	0	1
SPY		5min	1	0	0	1
SPY		10min	1	0	1	3.89E-15
SPY		30min	1	1.18E-08	0	1.00E+00
SPY	Calm Period (2013-Jan-02 -2013-June-28)	1min	1	0	0	0.9754
SPY		5min	1	0	0	1
SPY		10min	1	0	1	0
SPY		30min	1	0	1	0

Table 3.6: Hurst Exponent of the Original Return Series and the Residuals

Ticker	Time Horizon	Sampling Frequency	Original Return Series	Residual of ARIMA-IGARCH	Residual of FARIMA-FIGARCH
SPY	Turbulent Period (2008-July-01 -2008-Dec-31)	1min	0.5022	0.5184	0.5214
SPY		5min	0.4918	0.5186	0.4968
SPY		10min	0.4831	0.5160	0.4906
SPY		30min	0.4314	0.4735	0.4740
SPY	Calm Period (2013-Jan-02 -2013-June-28)	1min	0.5092	0.4998	0.5154
SPY		5min	0.5216	0.4978	0.5136
SPY		10min	0.5102	0.4821	0.4998
SPY		30min	0.5376	0.5214	0.5175

Table 3.7: Hurst Exponent of Square of Original Return Series and the Residuals

Ticker	Time Horizon	Sampling Frequency	Square of Original Return Series	Residual Square of ARIMA-IGARCH	Residual Square of FARIMA-FIGARCH
SPY	Turbulent Period (2008-July-01 -2008-Dec-31)	1min	0.9049	0.6312	0.5869
SPY		5min	0.8974	0.7528	0.5542
SPY		10min	0.9009	0.6816	0.5098
SPY		30min	0.8935	0.6816	0.5222
SPY	Calm Period (2013-Jan-02 -2013-June-28)	1min	0.8411	0.6471	0.5853
SPY		5min	0.7798	0.6572	0.5575
SPY		10min	0.7738	0.7234	0.5714
SPY		30min	0.7675	0.6824	0.5322

Table 3.8: Number of rejections of H_0.

	Time Horizon	Sampling Frequency	ARIMA-IGARCH-GH against ARIMA-IGARCH-Gaussian		FARIMA-FIGARCH-GH against ARIMA-IGARCH-GH	
Large Cap High P/E (14 stocks)	Turbulent Period (2008-July-01 -2008-Dec-31)	1min	14	100%	4	28.57%
		5min	14	100%	6	42.86%
		10min	14	100%	11	78.57%
		30min	14	100%	7	50%
	Calm Period (2013-Jan-02 -2013-June-28)	1min	14	100%	12	85.71%
		5min	14	100%	13	92.86%
		10min	14	100%	14	100%
		30min	14	100%	13	92.86%
Small Cap High P/E (9 stocks)	Turbulent Period (2008-July-01 -2008-Dec-31)	1min	16	100%	2	12.5%
		5min	16	100%	8	50%
		10min	16	100%	12	75%
		30min	16	100%	10	62.5%
	Calm Period (2013-Jan-02 -2013-June-28)	1min	16	100%	8	50%
		5min	16	100%	14	97.5%
		10min	16	100%	16	100%
		30min	16	100%	11	68.75%
Large Cap Low P/E (9 stocks)	Turbulent Period (2008-July-01 -2008-Dec-31)	1min	9	100%	3	33.33%
		5min	9	100%	3	33.33%
		10min	9	100%	7	77.78%
		30min	9	100%	7	77.78%
	Calm Period (2013-Jan-02 -2013-June-28)	1min	9	100%	6	66.67%
		5min	9	100%	8	88.89%
		10min	9	100%	9	100%
		30min	9	100%	8	88.89%
Small Cap Low P/E (11 stocks)	Turbulent Period (2008-July-01 -2008-Dec-31)	1min	11	100%	5	45.45%
		5min	11	100%	5	45.45%
		10min	11	100%	8	72.73%
		30min	11	100%	7	63.64%
	Calm Period (2013-Jan-02 -2013-June-28)	1min	11	100%	7	62.64%
		5min	11	100%	10	90.91%
		10min	11	100%	10	90.91%
		30min	11	100%	9	81.82%

References

[1] Basel III: international regulatory framework for banks, March 2015.

[2] Anat R. Admati and Paul Pfleiderer. A Theory of Intraday Patterns: Volume and Price Variability. *Review of Financial Studies*, 1(1):3–40, January 1988.

[3] Robert Almgren and Neil Chriss. Optimal execution of portfolio transactions. *The Journal of Risk*, 3(2):5–39, January 2001.

[4] A. A. Annis and E. H. Lloyd. The expected value of the adjusted rescaled Hurst range of independent normal summands. *Biometrika*, 63(1):111–116, January 1976.

[5] Richard T. Baillie, Tim Bollerslev, and Hans Ole Mikkelsen. Fractionally integrated generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 74(1):3–30, September 1996.

[6] Malcolm Baker and Jeffrey Wurgler. Investor Sentiment and the Cross-Section of Stock Returns. *The Journal of Finance*, 61(4):1645–1680, August 2006.

[7] Malcolm Baker and Jeffrey Wurgler. Investor Sentiment in the Stock Market. *The Journal of Economic Perspectives*, 21(2):129–151, April 2007.

[8] Nicholas Barberis, Andrei Shleifer, and Robert Vishny. A model of investor sentiment1. *Journal of Financial Economics*, 49(3):307–343, September 1998.

[9] O. Barndorff-Nielsen and Christian Halgreen. Infinite divisibility of the hyperbolic and generalized inverse Gaussian distributions. *Zeitschrift fr Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 38(4):309–311, 1977.

[10] Ole E. Barndorff-Nielsen and Neil Shephard. Normal modified stable

processes. Economics Paper 2001-W6, Economics Group, Nuffield College, University of Oxford, 2001.

[11] Ole Eiler Barndorff-Nielsen. Exponentially decreasing distributions for the logarithm of particle size. *Royal Society of London. Proceedings. Mathematical, Physical and Engineering Sciences*, 353(1674):401–419, 1977.

[12] Jan Beran. *Statistics for Long-Memory Processes*. Chapman and Hall/CRC, New York, 1 edition edition, October 1994.

[13] Jeremy Berkowitz. Testing Density Forecasts, With Applications to Risk Management. *Journal of Business & Economic Statistics*, 19(4):465–474, October 2001.

[14] Fischer Black. Noise. *The Journal of Finance*, 41(3):529–543, July 1986.

[15] Tim Bollerslev. Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 31(3):307–327, April 1986.

[16] Tim Bollerslev. A Conditionally Heteroskedastic Time Series Model for Speculative Prices and Rates of Return. *The Review of Economics and Statistics*, 69(3):542–547, 1987.

[17] Jean-Philippe Bouchaud, Yuval Gefen, Marc Potters, and Matthieu Wyart. Fluctuations and response in financial markets: the subtle nature of random price changes. *Quantitative Finance*, 4(2):176–190, April 2004.

[18] Svetlana I. Boyarchenko and Sergei Z. Levendorski. Option pricing for truncated lvy processes. *International Journal of Theoretical and Applied Finance*, 03(03):549–552, July 2000.

[19] Brendan O. Bradley and Murad S. Taqqu. Financial Risk and Heavy Tails A2 - Rachev, Svetlozar T. In *Handbook of Heavy Tailed Distributions in Finance*, volume 1 of *Handbooks in Finance*, pages 35–103. North-Holland, Amsterdam, 2003.

[20] Jonathan Brogaard. The Activity of High Frequency Traders. *SSRN Electronic Journal*, 2010.

[21] Jonathan Brogaard, Terrence Hendershott, and Ryan Riordan. High-Frequency Trading and Price Discovery. *Review of Financial Studies*, page hhu032, May 2014.

[22] Gregory W. Brown and Michael T. Cliff. Investor sentiment and the near-term stock market. *Journal of Empirical Finance*, 11(1):1–27, January 2004.

[23] GregoryW. Brown and MichaelT. Cliff. Investor Sentiment and Asset Valuation. *The Journal of Business*, 78(2):405–440, 2005.

[24] Enzo Busseti and Fabrizio Lillo. Calibration of optimal execution of financial transactions in the presence of transient market impact. *Journal of Statistical Mechanics: Theory and Experiment*, 2012(09):P09010, 2012.

[25] Daniel O. Cajueiro and Benjamin M. Tabak. Testing for long-range dependence in world stock markets. *Chaos, Solitons & Fractals*, 37(3):918–927, August 2008.

[26] John Y. Campbell and Albert S. Kyle. Smart Money, Noise Trading and Stock Price Behaviour. *The Review of Economic Studies*, 60(1):1–34, January 1993.

[27] Bidisha Chakrabarty, Pamela C. Moulton, and Andriy Shkilko. Short Sales, Long Sales, and the Lee-Ready Trade Classification Algorithm Revisited. SSRN Scholarly Paper ID 1755025, Social Science Research Network, Rochester, NY, January 2012.

[28] Rong Chen, Zhixin Kang, and Lan Zhang. Forecasting return volatility in the presence of microstructure noise. *Statistics and Its Interface*, 3:145–157, 2010.

[29] Peter F. Christoffersen. Evaluating Interval Forecasts. *International Economic Review*, 39(4):841–862, 1998.

[30] Li-Chung Chu and Urbashi Mitra. Approximate Maximum Likelihood Sequence Detection for DS/CDMA Systems with Tracking Errors. In *in 33rd Annu. Conf. Information Science and Systems*, 1999.

[31] Ching-Fan Chung. Estimating the Fractionally Integrated GARCH Model. *unpublished paper.*

[32] San-Lin Chung, Chi-Hsiou Hung, and Chung-Ying Yeh. When does investor sentiment predict stock returns? *Journal of Empirical Finance*, 19(2):217–240, March 2012.

[33] A. Clauset, C. Shalizi, and M. Newman. Power-Law Distributions in Empirical Data. *SIAM Review*, 51(4):661–703, November 2009.

[34] Rama Cont. Empirical properties of asset returns: stylized facts and statistical issues. *Quantitative Finance*, 1:223–236, 2001.

[35] Kent Daniel, David Hirshleifer, and Avanidhar Subrahmanyam. Investor Psychology and Security Market Under- and Overreactions. *The Journal of Finance*, 53(6):1839–1885, December 1998.

[36] J. Bradford De Long, Andrei Shleifer, Lawrence H. Summers, and Robert J. Waldmann. Noise Trader Risk in Financial Markets. *Journal of Political Economy*, 98(4):703–738, 1990.

[37] J. Bradford Delong, Andrei Shleifer, Lawrence H. Summers, and Robert Waldmann. The Survival of Noise Traders in Financial-Markets. 64:1–19, 1991.

[38] Stefano Demarta and Alexander J. McNeil. The t Copula and Related Copulas. *International Statistical Review / Revue Internationale de Statistique*, 73(1):111–129, 2005.

[39] Francis X. Diebold and Atsushi Inoue. Long Memory and Structural Change. SSRN Scholarly Paper ID 267789, Social Science Research Network, Rochester, NY, May 1999.

[40] Georges Dionne, Pierre Duchesne, and Maria Pacurar. Intraday Value at Risk (IVaR) using tick-by-tick data with application to the Toronto

Stock Exchange. *Journal of Empirical Finance*, 16(5):777–792, December 2009.

[41] Paul Doukhan, George Oppenheim, and Murad Taqqu. *Theory and Applications of Long-Range Dependence*. Springer Science & Business Media, December 2002. Google-Books-ID: xIrDf3DAnKMC.

[42] William H. DuMouchel. On the Asymptotic Normality of the Maximum-Likelihood Estimate when Sampling from a Stable Distribution. *The Annals of Statistics*, 1(5):948–957, September 1973.

[43] Ernst Eberlein and Karsten Prause. The Generalized Hyperbolic Model: Financial Derivatives and Risk Measures. In *Mathematical Finance Bachelier Congress 2000, Geman*, pages 245–267. Springer, 1998.

[44] Alex Edmans, Diego Garca, and yvind Norli. Sports Sentiment and Stock Returns. *The Journal of Finance*, 62(4):1967–1998, August 2007.

[45] Paul Embrechts, Filip Lindskog, and Alexander Mcneil. Modelling Dependence with Copulas and Applications to Risk Management. In *Handbook of Heavy Tailed Distributions in Finance*, pages 329–384. Elsevier, 2003.

[46] Robert F. Engle. Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation. *Econometrica*, 50(4):987–1007, 1982.

[47] Hassan Ezzat. Long Memory Processes and Structural Breaks in Stock Returns and Volatility: Evidence from the Egyptian Exchange, August 2013.

[48] Eugene F. Fama. The Behavior of Stock-Market Prices. *The Journal of Business*, 38(1):34–105, 1965.

[49] Lars Forsberg and Tim Bollerslev. Bridging the gap between the distribution of realized (ECU) volatility and ARCH modelling (of the Euro): the GARCH-NIG model. *Journal of Applied Econometrics*, 17(5):535–548, September 2002.

[50] Paul Gallimore and Adelaide Gray. The role of investor sentiment in property investment decisions. *Journal of Property Research*, 19(2):111–120, January 2002.

[51] Jack E. Gaumnitz and Carlos A. Salabar. The Barrons Confidence Index: An Examination of Its Value as a Market Indicator. *Financial Analysts Journal*, 25(5):16–17, 70, September 1969.

[52] Hlyette Geman, Peter Carr, Dilip B. Madan, and Marc Yor. The Fine Structure of Asset Returns: An Empirical Investigation. SSRN Scholarly Paper ID 305710, Social Science Research Network, Rochester, NY, May 2002.

[53] Ramazan Genay, Michel Dacorogna, Ulrich A. Muller, Olivier Pictet, and Richard Olsen. *An Introduction to High-Frequency Finance*. Academic Press, San Diego, 1 edition edition, May 2001.

[54] John Geweke and Susan Porter-Hudak. The Estimation and Application of Long Memory Time Series Models. *Journal of Time Series Analysis*, 4(4):221–238, July 1983.

[55] Pierre Giot. Market risk models for intraday data. *The European Journal of Finance*, 11(4):309–324, August 2005.

[56] James Glimm, Jimmie Goode, Beiyu Lin, Nicholas Pezolano, and Svetlozar Rachev. Real Time Evaluation of Financial Returns Based on Nearly Elliptical Models, U.S. Patent Application No. 14/163,674, July 2014.

[57] Christian Gourieroux and Joann Jasiak. Local Likelihood Density Estimation and Value-at-Risk. *Journal of Probability and Statistics*, 2010:e754851, June 2010.

[58] C. W. J. Granger. Long memory relationships and the aggregation of dynamic models. *Journal of Econometrics*, 14(2):227–238, October 1980.

[59] C. W. J. Granger and Roselyne Joyeux. An Introduction to Long-Memory Time Series Models and Fractional Differencing. *Journal of*

Time Series Analysis, 1(1):15–29, January 1980.

[60] Sanford J. Grossman and Joseph E. Stiglitz. On the Impossibility of Informationally Efficient Markets. *The American Economic Review*, 70(3):393–408, 1980.

[61] Bing Han. Stochastic Volatilities and Correlations of Bond Yields. *The Journal of Finance*, 62(3):1491–1524, June 2007.

[62] Jrdis Hengelbrock, Erik Theissen, and Christian Westheide. Market Response to Investor Sentiment. *SSRN Electronic Journal*, 2010.

[63] Harrison Hong and Jeremy C. Stein. A Unified Theory of Underreaction, Momentum Trading, and Overreaction in Asset Markets. *The Journal of Finance*, 54(6):2143–2184, December 1999.

[64] J. R. M. Hosking. Fractional differencing. *Biometrika*, 68(1):165–176, April 1981.

[65] H. E. Hurst. Long-Term Storage Capacity of Reservoirs. *Transactions of the American Society of Civil Engineers*, 116(1):770–799, 1951.

[66] Clifford M. Hurvich, Rohit Deo, and Julia Brodsky. The mean squared error of Geweke and Porter-Hudak's estimator of the memory parameter of a long-memory time series. *Journal of Time Series Analysis*, 19(1):19–46, January 1998.

[67] Gwilym M. Jenkins and Donald G. Watts. *Spectral Analysis and Its Applications*. Holden Day, San Francisco, June 1968.

[68] Charles M. Jones. What Do We Know About High-Frequency Trading? SSRN Scholarly Paper ID 2236201, Social Science Research Network, Rochester, NY, March 2013.

[69] Philippe Jorion. *Financial Risk Manager Handbook 2001-2002*. John Wiley & Sons, New York, 2nd edition edition, August 2001.

[70] Daniel Kahneman and Amos Tversky. Prospect Theory: An Analysis of Decision under Risk. *Econometrica*, 47(2):263–291, 1979.

[71] Guy Kaplanski and Haim Levy. Sentiment and stock prices: The case of aviation disasters. *Journal of Financial Economics*, 95(2):174–201, February 2010.

[72] Y. S. Kim, S. T. Rachev, D. M. Chung, and M. L. Bianchi. A modified tempered stable distribution with volatility clustering. *ResearchGate*, January 2008.

[73] Young Shin Kim, Svetlozar Rachev, Michele Leonardo Bianchi, and Frank J. Fabozzi. Computing VAR and AVaR in Infinitely Divisible Distributions. SSRN Scholarly Paper ID 1400965, Social Science Research Network, Rochester, NY, March 2009.

[74] Young Shin Kim, Svetlozar T. Rachev, Michele Leonardo Bianchi, and Frank J. Fabozzi. Financial market models with Lvy processes and time-varying volatility. *Journal of Banking & Finance*, 32(7):1363–1378, July 2008.

[75] Young Shin Kim, Svetlozar T. Rachev, Michele Leonardo Bianchi, Ivan Mitov, and Frank J. Fabozzi. Time series analysis for financial market meltdowns. *Journal of Banking & Finance*, 35(8):1879–1891, August 2011.

[76] Leonid Kogan, Stephen A. Ross, Jiang Wang, and Mark M. Westerfield. The Price Impact and Survival of Irrational Traders. *The Journal of Finance*, 61(1):195–229, February 2006.

[77] Ismo Koponen. Analytic approach to the problem of convergence of truncated Lvy flights towards the Gaussian stochastic process. *Physical Review E*, 52(1):1197–1199, July 1995.

[78] Sebastian Kring, Svetlozar T. Rachev, Markus Hchsttter, and Frank J. Fabozzi. Composed and Factor Composed Multivariate GARCH Models. *ResearchGate*.

[79] Alok Kumar and Charles M.c. Lee. Retail Investor Sentiment and Return Comovements. *The Journal of Finance*, 61(5):2451–2486, October 2006.

[80] Paul H Kupiec. Techniques for Verifying the Accuracy of Risk Mea-

surement Models. *The Journal of Derivatives*, 3(2):73–84, January 1995.

[81] Charles M. C. Lee and Mark J. Ready. Inferring Trade Direction from Intraday Data. *The Journal of Finance*, 46(2):733–746, June 1991.

[82] Michael Lemmon and Evgenia Portniaguina. Consumer Confidence and Asset Prices: Some Empirical Evidence. *Review of Financial Studies*, 19(4):1499–1529, December 2006.

[83] Josef Leydold and Wolfgang Hrmann. Generating generalized inverse Gaussian random variates by fast inversion. *Computational Statistics & Data Analysis*, 55(1):213–217, January 2011.

[84] David X Li. On Default Correlation: A Copula Function Approach. *The Journal of Fixed Income*, 9(4):43–54, March 2000.

[85] Fabrizio Lillo and J. Doyne Farmer. The Long Memory of the Efficient Market. *Studies in Nonlinear Dynamics & Econometrics*, 8(3), January 2004.

[86] Shi-Miin Liu and B. Wade Brorsen. Maximum Likelihood Estimation of a Garch-Stable Model. *Journal of Applied Econometrics*, 10(3):273–85, 1995.

[87] Andrew W. Lo. Long-Term Memory in Stock Market Prices. *Econometrica*, 59(5):1279–1313, 1991.

[88] Julian Michael Lorenz. *Optimal Trading Algorithms: Portfolio Transactions, Multiperiod Portfolio Selection, and Competitive Online Search*. ETH, 2008. Google-Books-ID: WHXMSAAACAAJ.

[89] Benoit Mandelbrot. The Variation of Some Other Speculative Prices. *The Journal of Business*, 40(4):393–413, 1967.

[90] Harry M. Markowitz. *Portfolio Selection*. Yale University Press, October 2013.

[91] Lisana B. Martinez, M. Beln Guercio, Aurelio Fernandez Bariviera, and Antonio Terceo. The impact of the financial crisis on the long-range

memory of European corporate bond and stock markets. *Empirica*, pages 1–15, May 2016.

[92] T. Di Matteo. Multi-scaling in finance. *Quantitative Finance*, 7(1):21–36, February 2007.

[93] T. Di Matteo, T. Aste, and Michel M. Dacorogna. Long-term memories of developed and emerging markets: Using the scaling analysis to characterize their stage of development. Econometrics 0503004, EconWPA, 2005.

[94] J. Huston McCulloch. Measuring Tail Thickness to Estimate the Stable Index : A Critique. *Journal of Business & Economic Statistics*, 15(1):74–81, 1997.

[95] Angus Ian McLeod and Keith William Hipel. Preservation of the rescaled adjusted range: 1. A reassessment of the Hurst Phenomenon. *Water Resources Research*, 14(3):491–508, June 1978.

[96] Alexander J. McNeil, Rdiger Frey, and Paul Embrechts. *Quantitative Risk Management: Concepts, Techniques and Tools*. Princeton University Press, May 2015. Google-Books-ID: SfJnBgAAQBAJ.

[97] Alexander J. McNeil, Rdiger Frey, and Paul Embrechts. *Quantitative Risk Management: Concepts, Techniques and Tools*. Princeton University Press, Princeton, NJ, revised edition edition, May 2015.

[98] Christian Menn and Svetlozar T. Rachev. Smoothly truncated stable distributions, GARCH-models, and option pricing. *Mathematical Methods of Operations Research*, 69(3):411, September 2008.

[99] Walid Mensi, Shawkat Hammoudeh, and Seong-Min Yoon. Structural breaks and long memory in modeling and forecasting volatility of foreign exchange markets of oil exporters: The importance of scheduled and unscheduled news announcements. *International Review of Economics & Finance*, 30:101–119, March 2014.

[100] G. Mujtaba Mian and Srinivasan Sankaraguruswamy. Investor Sentiment and Stock Market Response to Earnings News. *The Accounting Review*,

87(4):1357–1384, March 2012.

[101] S. Mittnik, S. T. rachev, T. Doganoglu, and D. Chenyao. Maximum likelihood estimation of stable Paretian models. *Mathematical and Computer Modelling*, 29(10):275–293, May 1999.

[102] Stefan Mittnik, Marc S. Paolella, and Svetlozar T. Rachev. Diagnosing and treating the fat tails in financial returns data. *Journal of Empirical Finance*, 7(34):389–416, November 2000.

[103] Stefan Mittnik, Marc S. Paolella, and Svetlozar T. Rachev. Stationarity of stable power-GARCH processes. *Journal of Econometrics*, 106(1):97–107, January 2002.

[104] Mervin E. Muller. A Note on a Method for Generating Points Uniformly on N-dimensional Spheres. *Commun. ACM*, 2(4):19–20, April 1959.

[105] A. Nagurney. Finance and variational inequalities*. *Quantitative Finance*, 1(3):309–317, March 2001.

[106] Robert Neal and Simon M. Wheatley. Do Measures of Investor Sentiment Predict Returns? *Journal of Financial and Quantitative Analysis*, 33(4):523–547, December 1998.

[107] Daniel B. Nelson. Stationarity and Persistence in the GARCH(1,1) Model. *Econometric Theory*, 6(3):318–334, 1990.

[108] Daniel B. Nelson and Charles Q. Cao. Inequality Constraints in the Univariate GARCH Model. *Journal of Business & Economic Statistics*, 10(2):229–235, 1992.

[109] J. P. Nolan, A. K. Panorska, and J. H. McCulloch. Estimation of stable spectral measures. *Mathematical and Computer Modelling*, 34(9):1113–1122, November 2001.

[110] John P. Nolan. Multivariate elliptically contoured stable distributions: theory and estimation. *Computational Statistics*, 28(5):2067–2089, January 2013.

[111] A. K. Panorska, S. Mittnik, and S. T. Rachev. Stable GARCH models

for financial time series. *Applied Mathematics Letters*, 8(5):33–37, September 1995.

[112] Edgar E. Peters. *Fractal Market Analysis: Applying Chaos Theory to Investment and Economics*. Wiley, New York, 1 edition edition, January 1994.

[113] Christopher Polk and Paola Sapienza. The Real Effects of Investor Sentiment. Working Paper 10563, National Bureau of Economic Research, June 2004.

[114] S. T. Rachev. *Handbook of Heavy Tailed Distributions in Finance: Handbooks in Finance*. Elsevier, March 2003. Google-Books-ID: sv8jGSVFra8C.

[115] Svetlozar T. Rachev, Young Shin Kim, Michele L. Bianchi, and Frank J. Fabozzi. *Financial Models with Levy Processes and Volatility Clustering*. John Wiley & Sons, February 2011. Google-Books-ID: XKvUUrcS_twC.

[116] Svetlozar T. Rachev and Stefan Mittnik PhD. *Stable Paretian Models in Finance*. Wiley, June 2000. Google-Books-ID: k5OxAAAAIAAJ.

[117] P. M. Robinson. Log-Periodogram Regression of Time Series with Long Range Dependence. *The Annals of Statistics*, 23(3):1048–1072, 1995.

[118] RichardJ. Rosen. Merger Momentum and Investor Sentiment: The Stock Market Reaction to Merger Announcements. *The Journal of Business*, 79(2):987–1017, 2006.

[119] Peter J. Rousseeuw and Geert Molenberghs. Transformation of non positive semidefinite correlation matrices. *Communications in Statistics - Theory and Methods*, 22(4):965–984, January 1993.

[120] Jeffrey R. Russell and Robert F. Engle. Autoregressive Conditional Duration: A New Model for Irregularly Spaced Transaction Data. SSRN Scholarly Paper ID 6809, Social Science Research Network, Rochester, NY, April 1998.

[121] Gennady Samoradnitsky and M. S. Taqqu. *Stable Non-Gaussian Ran-*

dom Processes: Stochastic Models with Infinite Variance. CRC Press, June 1994. Google-Books-ID: wTTUfYwjksAC.

[122] Gennady Samorodnitsky. Long Range Dependence. In *Encyclopedia of Actuarial Science*. John Wiley & Sons, Ltd, 2006.

[123] JosA. Scheinkman and Wei Xiong. Overconfidence and Speculative Bubbles. *Journal of Political Economy*, 111(6):1183–1220, 2003.

[124] A. Sensoy. Time-varying long range dependence in market returns of FEAS members. *Chaos, Solitons & Fractals*, 53:39–45, August 2013.

[125] Ahmet Sensoy and Benjamin M. Tabak. Time-varying long term memory in the European Union stock markets. *Physica A: Statistical Mechanics and its Applications*, 436:147–158, October 2015.

[126] Andrei Shleifer. *Inefficient Markets: An Introduction to Behavioural Finance*. OUP Oxford, March 2000.

[127] Stoyan Stoyanov and Borjana Racheva-Iotova. Univariate stable laws in the field of finance parameter estimation. *ResearchGate*, 2(4), January 2004.

[128] Peter Tankov and Rama Cont. *Financial Modelling with Jump Processes*. Chapman and Hall/CRC, Boca Raton, Fla, 1 edition edition, December 2003.

[129] Paul C. Tetlock. Giving Content to Investor Sentiment: The Role of Media in the Stock Market. *The Journal of Finance*, 62(3):1139–1168, June 2007.

[130] University of Michigan. University of Michigan: Consumer Sentiment, November 1952.

[131] F. Albert Wang. Overconfidence, Investor Sentiment, and Evolution. *Journal of Financial Intermediation*, 10(2):138–170, April 2001.

[132] Huijun Wang, Jinghua Yan, and Jianfeng Yu. Prospect Theory and the Risk-Return Tradeoff. *SSRN Electronic Journal*, 2012.

[133] Rafa Weron. Estimating long-range dependence: finite sample properties and confidence intervals. *Physica A Statistical Mechanics and its Applications*, 312:285–299, September 2002.