# Volatility Forecasting with Latent Information and Exogenous Variables

Author:
**Chainarong Kesamoon**

Supervisor:
**Joan del Castillo Franquet**

Ph.D. Thesis:
**Doctoral Program in Mathematics**

Departament de Matemàtiques
Universitat Autònoma de Barcelona

Barcelona, February 2015

*แด่แม่ พ่อ และครอบครัวที่รัก*

*To mom, dad and my beloved family.*

# Acknowledgements

I would like to express my special appreciation and thanks to my supervisor Dr. Joan del Castillo Franquet, who has always encouraged and supported me throughout the thesis. His valuable advice have been priceless substance that make this thesis come true. Not only academic advice, his kindness, optimist and positive attitude have made years of endurance become enjoyable. I also would like to thank Juan Pablo Ortega who was my advisor when I visited Besançon. He was very hospitable and inspiring every time we met.

This thesis would not be possible without the financial support from the Department of Mathematics, UAB. Thanks to kind staff in the department for being very responsive and helpful. Special thanks to my colleagues and friends, Isabel Serra and Maria Padilla, who have helped me on analyzing data with CV-plots. They are very supportive and always ready to give help.

In memory of my parents and deepest appreciation to my family, brothers and sister-in-laws, I can not express how I am grateful to these people for all the support and encouragement. Being far away from home in different culture is not anything easy, I also thanks to my lovely girlfriend Piraya, all good friends and colleagues who make this long journey wonderful.

# Contents

**Contents** vii

**List of Figures** x

**List of Tables** xi

**1 Introduction** 1
- 1.1 Motivation . . . 1
- 1.2 Data Set Description . . . 3
- 1.3 Mathematics in Finance . . . 5
- 1.4 Organization . . . 6

**2 Financial Econometrics** 7
- 2.1 Introduction . . . 7
- 2.2 Prices and Returns . . . 8
- 2.3 Stylized Facts for Financial Returns . . . 11
- 2.4 Heavy Tails & the CV-Plots . . . 15
- 2.5 Discrete Time Models . . . 18
  - 2.5.1 ARCH models . . . 20
  - 2.5.2 Stochastic volatility models . . . 22
  - 2.5.3 Volatility estimates . . . 23
- 2.6 Continuous Time Models . . . 24
  - 2.6.1 Brownian motion . . . 25
  - 2.6.2 Lévy process . . . 26

**3 Explanatory Data Analysis with NIG** 29
- 3.1 The NIG Distribution . . . 29
- 3.2 Data Descriptive Statistics . . . 31
- 3.3 Fitting Financial Data with NIG Distribution . . . 31
  - 3.3.1 Skewness . . . 32
  - 3.3.2 Parameter estimation . . . 33
  - 3.3.3 Goodness of fit . . . 34
- 3.4 Summary . . . 39

**4 Volatility Forecasting** 41

4.1 Practical Issues in Volatility Forecasting . . . 41
4.1.1 Volatility proxy . . . 41
4.1.2 Forecast evaluation . . . 42
4.1.3 Forecasting models . . . 43
4.1.4 The role of the log transformation . . . 44
4.1.5 Recent work and forecasting strategy . . . 45
4.2 Parametric Lévy Processes . . . 46
4.2.1 The generalized hyperbolic Lévy processes . . . 46
4.2.2 Scale mixture of normal . . . 48
4.3 NIG-SV Model and HGLM Method . . . 49
4.3.1 The model . . . 49
4.3.2 H-likelihood estimation . . . 49
4.3.3 Multivariate NIG-SV model . . . 53
4.3.4 Optimization . . . 53
4.4 Latent Information and Volatility Forecasting . . . 57
4.4.1 Random effects and log volatility estimates . . . 57
4.4.2 Implementation and empirical results . . . 59
4.5 Summary . . . 61

**5 Range-Based Volatility Models 63**
5.1 Range-Based Volatility Estimators . . . 63
5.2 Simulation Study on Range-Based Volatility Estimators . . . 66
5.2.1 Discretization error . . . 67
5.2.2 Simulation with constant volatility . . . 68
5.2.3 Simulation with stochastic volatility . . . 69
5.3 Discussion on the Results . . . 70
5.4 Summary . . . 74

**6 Dynamic Normal Inverse Gaussian Models 77**
6.1 The DNIG Model . . . 77
6.1.1 Definition . . . 78
6.1.2 Higher order setting . . . 79
6.2 Parameter Estimation . . . 79
6.2.1 Maximum likelihood estimation . . . 79
6.2.2 H-likelihood estimation . . . 80
6.2.3 Empirical results on estimation . . . 81
6.2.4 Residual analysis . . . 84
6.3 Volatility Forecasting . . . 86
6.3.1 Moments of the range . . . 86
6.3.2 DNIG forecasting model . . . 88
6.3.3 Implementation and empirical results . . . 89
6.4 Future Research . . . 91

**Summary and Conclusions of the Thesis 93**

**Index** **98**

**Bibliography** **101**

**A HGLM Method for Multivariate NIG-SV Model** **107**

**B Extensive Simulation Results** **109**

**C The BHHH algorithm for DNIG model** **127**

# List of Figures

1.2.1 Currency Exchange Rate between 2006 and 2010 . . . 4
1.2.2 Low price, Close price and High price of EUR . . . 4

2.2.1 EUR returns . . . 10
2.3.1 Kernel density of EUR returns . . . 12
2.3.2 EUR autocorrelations . . . 14
2.3.3 Autocorrelations of absolute and squared returns; averages across 3 currencies . . . 15
2.4.1 CV-plots of some distributions . . . 18
2.4.2 CV-plots of returns for currency exchanges . . . 19
2.5.1 Volatility estimates of EUR (annualized) . . . 24

3.3.1 The density plots for returns superimposed on the fitted densities of normal and NIG. . . . 36
3.3.2 The Q-Q plots of fitted NIG quantiles against sample quantiles for returns 37

4.4.1 The QL loss series of EUR forecasts . . . 60

6.2.1 The evolution of log-likelihood ratio statistics when the NIG-SV model is tested against the DNIG(1) model. . . . 84
6.2.2 The histograms for returns standardized by the volatility estimates from DNIG models . . . 85
6.2.3 CV-plots of returns standardized by volatilities estimated from DNIG models 87
6.3.1 The plots of QL loss series for JPY from GARCH and DNIG models . . . 91

# List of Tables


3.1 Descriptive statistics for time series of returns . . . 32
3.2 Frequency distributions . . . 33
3.3 Likelihood ratio test . . . 34
3.4 Estimated parameters from MoM and MLE . . . 35
3.5 Goodness-of-fit test . . . 39

4.1 Parameter estimation for NIG-SV models with maximum likelihood (ML), first-order h-likelihood (H1),and second-order h-likelihood (H2) with the corresponding standard errors. . . . 52
4.2 Parameter estimation for multivariate NIG-SV models in five periods. . . . 54
4.3 Volatility forecasting with RW, GARCH, NIG-SV and NIG-SV* . . . 61

5.1 The accuracy and efficiency of range-based volatility estimators when log prices follow Brownian motion with intraday movements $k = 40$, daily variance $\sigma^2 = 10 \times 10^{-5}$. . . . 71
5.2 The effect of discretization when log prices follow Brownian motion with $\sigma^2 = 10$. The averages of variance estimates ($\times 10^5$ ) with 95% confidence intervals are presented. . . . 72
5.3 The accuracy and efficiency of range-based volatility estimators when log prices follow NIG-SV model with intraday movements $k = 40$, expected variance $E[\sigma_i^2] = 10 \times 10^{-5}$, kurtosis $1/\omega = 5$. . . . 73
5.4 The effect of discretization when log prices follow NIG-SV model with $E[\sigma_i^2] = 10 \times 10^5$, kurtosis=5. The averages of variance estimates ($\times 10^5$ ) with 95% confidence intervals are presented. . . . 74
5.5 The accuracy of the range-based estimators when the price processes follow GBM and NIG-SV models. . . . 75

6.1 Parameter Estimates and their 95% confidence intervals for DNIG(1) models with MLE and h-likelihood methods. . . . 82
6.2 Parameter Estimates and their 95% confidence intervals for DNIG(2) models with MLE and h-likelihood methods. . . . 83
6.3 The log-likelihood ratio statistics and the corresponding p-values of the null model vs the alternative model. The series are modeled by NIG-SV, DNIG(1) and DNIG(2) models. . . . 83

6.4 Summary statistics for returns standardized by volatilities estimated from NIG-SV, DNIG(1) and DNIG(2) models. . . . . 85
6.5 Average QL loss of cumulative volatility forecasts . . . . . 90

B.1 The effect of drift on range-based estimators with constant volatility. . . . 110
B.2 The effect of drift on range-based estimators with constant volatility. . . . 111
B.3 The effect of drift on range-based estimators with constant volatility. . . . 112
B.4 The effect of drift on range-based estimators with constant volatility. . . . 113
B.5 The effect of drift on range-based estimators with constant volatility. . . . 114
B.6 The effect of drift on range-based estimators with constant volatility. . . . 115
B.7 The average variance estimates with 95% confidence interval. . . . . 116
B.8 The average variance estimates with 95% confidence interval. . . . . 117
B.9 The effect of drift on range-based estimators with stochastic volatility. . . 118
B.10 The effect of drift on range-based estimators with stochastic volatility. . . 119
B.11 The effect of drift on range-based estimators with stochastic volatility. . . 120
B.12 The effect of drift on range-based estimators with stochastic volatility. . . 121
B.13 The effect of drift on range-based estimators with stochastic volatility. . . 122
B.14 The effect of drift on range-based estimators with stochastic volatility. . . 123
B.15 The average variance estimates with 95% confidence interval. . . . . 124
B.16 The average variance estimates with 95% confidence interval. . . . . 125

# Chapter 1

# Introduction

> What we know about the global financial crisis is that we don't know very much.
>
> Paul Samuelson, Nobel laureate in Economic Sciences, 1970

## 1.1 Motivation

In 2008 the world economy came across its most hazardous crisis since the Great Depression of the 1930s. Some financial service firms collapsed even though those who were regarded as 'too big too fail'. Lehmann Brothers, the fourth largest investment bank in the US, underwent the bailout program by the US government along with AIG, Goldman Sachs, Merrill Lynch, Citigroup Inc., to keep them from bankruptcy; however it eventually went bankrupt in September 2008. The Dow Jones got its largest drop in a single day since the days following the attacks on September 11, 2001. Whereas the S&P500 had a number of days with extreme movements ($\geq 4\%$) greater than its overall 80-year history in that October. The confidence in financial market was shattered and the suspiciousness spread throughout the globe. Central banks in England, China, Canada, Sweden, Switzerland and the European Central Bank (ECB) resorted to rate cuts to relief the world economy but could not stop such a widespread financial meltdown.

The crisis motivated practitioners and academics to reassess the statistical models being used in the financial world and started to question the adequacy of standard models. It urges us to reconsider how well we understand the tools being used to forecast the future and manage the risk. *Volatility*, even though is not the same as risk, is strongly related and being used widely to determine risk. For example, the value-at-risk (VaR) that banks and trading houses use to determine the value of reserve capital to set-aside, is defined as the minimum expected loss with a 1-percent confidence level for a given time horizon (see Poon & Granger, 2005). Volatility is also a key input in the pricing

of derivative securities, whose trading volume has been largely increasing recently. The knowledge of past and current volatility is not yet sufficient to deal with the uncertainty in financial market. For instance, to price an option, it is necessary to know the volatility of the underlying asset throughout the life of the option. Therefore, a crucial task for investors and policy makers who seek rational decisions in risk management and derivatives valuation is to forecast the volatility accurately.

A large number of volatility models have been proposed as a matter of massive interest. The major classes of volatility models that have been extensively investigated are the class of generalized conditional autoregressive heteroskedastic (GARCH) models and the class of stochastic volatility (SV) models. GARCH models are constructed by specifying the dynamics to the variances of standardized residuals of returns conditional on past history. Result in a class of models that is simple both in parameter estimation and volatility forecasting. Nevertheless, GARCH are modeled only in discrete time, while principal theory in option pricing are derived in continuous environment. Therefore option pricing under GARCH framework turns into a particularly complicated task. In contrast with SV models that volatility are modeled as stochastic variables with desired properties, whether or not in continuous conditions. SV models are usually discrete time approximations to continuous time stochastic processes. Consequently, SV models are closely related to the fundamental theory in finance, their properties are easier to find and they are easily generalized to multivariate series in a very natural way (Harvey et al., 1994). However SV models draw less attention from practitioners since parameter estimation can be often complicated.

Volatility forecasting is a critical task in several financial applications. It is even more challenging to forecast volatility in the period of financial crisis. Recently, Brownlees et al. (2012) presented a comprehensive study of volatility forecasting through the period of the crisis of 2008 with five GARCH models. A broad range of practical issues in volatility forecasting was examined among this type of models, the amount of data to use in estimation, the frequency of estimation update, and the relevance of heavy-tailed likelihoods for volatility forecasting. They found that volatility during the 2008 crisis is well approximated by predictions made one day ahead, where one-month-ahead forecasts are deteriorated. This encourages us to investigate in the performances of other models and develop new model with high predictive ability. The model that could explain the current situation and foretell the coming crisis would be desirable.

Latency of volatility causes difficulty to infer from its observed values. Both previously indicated models, GARCH and SV models, use returns information to model and forecast volatility as returns are 'byproduct' of unobservable volatilities. Widely-used variables that have been used to inference about volatility are absolute return and squared return. High-frequency realized volatility that is calculated from intraday prices is also valid but the data are not publicly accessible. Alternative to return, price *range*, which is the difference between highest and lowest prices during the day are also employed as volatility estimator. Range-based volatility models are not vastly investigated as return-based volatility models even though some researchers have claimed that the range is more efficient than return. We extensively investigate in the properties of the

range as the estimator of volatility and finally incorporate it into the new proposed volatility model.

## 1.2 Data Set Description

A financial data set usually consists of a record of trades of financial assets such as stocks, bonds or foreign exchange rates. The data can be obtained from many sources, including websites, commercial vendors and financial markets. Most sources provide daily data that usually consist of open price, close price, high price, low price, traded amount (volume) and number of trades. In fact every single trade of a particular asset is registered with time, amount, settling price, bid price, ask price and setting date. So that some sources also provide *intra-daily* data sets with some expense.

Here we work on daily data that are usually accessible freely in some sources. The data set consists of daily exchange rates of three major currencies between 2006 and 2010 collected from Bloomberg. The domestic currency is the US dollar and the foreign currencies are the euro (EUR), the British pound (GBP) and the Japanese yen (JPY). The prices of EUR, GBP and JPY are quoted in terms of USD. The record of data includes price, high price and low price which in fact refer to close price, daily highest price and lowest price respectively. The time horizon spans from 1 January 2006 to 31 December 2010 that we divide into four periods due to the global financial situation at that time.

The *Pre-Crisis* period is from 1 January to 30 June 2007 that the financial market is calm in general. The *Crisis 1* ranges from 1 July 2007 to 30 June 2008 that some signs of coming crisis being noticed. In 9 August 2007, BNP Paribas is the first major bank to acknowledge the risk of exposure to sub-prime mortgage markets by freezing three of their funds, indicating that they have no way of valuing the complex assets inside them. Adam Applegarth, Northern Rock's chief executive, later says that it was "the day the world changed". The *Crisis 2*, from 1 July 2008 to 30 June 2009, is the period that several financial firms face difficulty. The American bank Lehman Brothers files for bankruptcy in 15 September 2008, prompting worldwide financial panic. And the *Post-Crisis*, from 1 July 2009 to 31 December 2010, is just after the period of turmoil from 2008 to 2009.

The data are plotted in Figure 1.2.1 showing the movements of prices in four specified periods. In this figure, it is clearly seen the fluctuation of all exchange rates during the crisis 2. The prices of JPY starts to rise up significantly in the crisis 1 along with the prices EUR and GPB that have been growing since the pre-crisis period. The pre-crisis period seems to be the most stable period while the post-crisis period still show some fluctuations unless not as much as the crisis 2. Figure 1.2.2 shows the (close) prices, high prices and low prices of EUR in the first month of the data set.

Figure 1.2.1: Currency Exchange Rate between 2006 and 2010

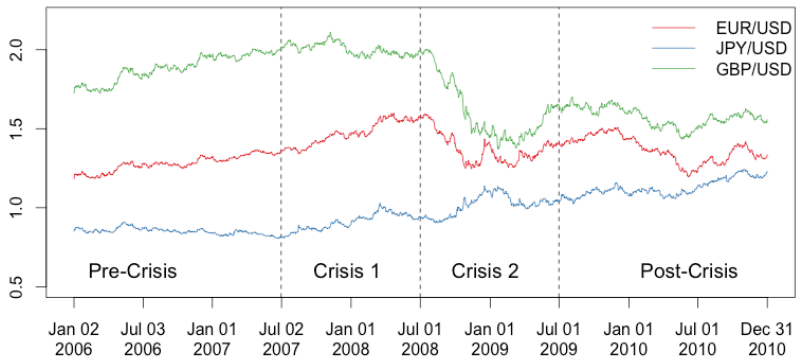
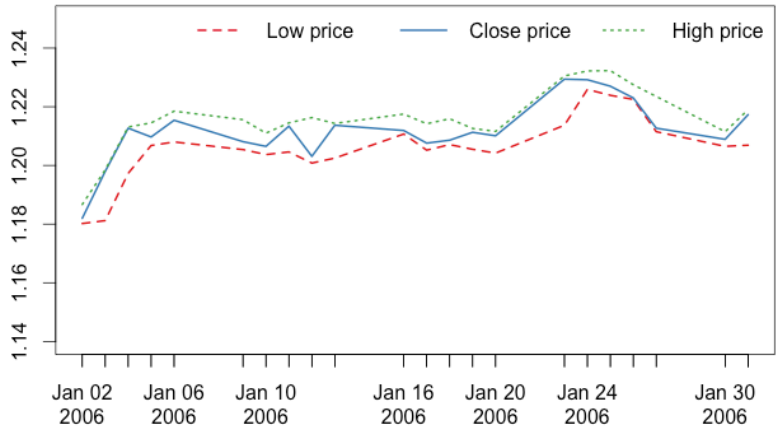



Figure 1.2.2: Low price, Close price and High price of EUR



## 1.3 Mathematics in Finance

In 1900, advanced mathematics was introduced to financial world for the first time by Louis Bachelier, a young French graduate student at the Sorbonne (the historic University of Paris). In his doctoral thesis (Bachelier, 1900), *Théorie de la spéculation* (The theory of speculation), that was advised by Henri Poincaré, he made a remarkable claim that stock prices moved according to a *random walk*. A random walk is that something moves randomly in direction and distance at each increment in time. He presented mathematics of stock price showing that the price evolves away from its initial value as the square root of the time elapsed. The radical implication of Bachelier's claim was that there was no more useful information in the path of a stock price over time than there was in the wanderings of a drunk down the street of Paris. Even though what he had done was recognized by the mathematical community and understood his valuable work by the time he died in 1946, Bachelier's work was not introduced to economists until 1954.

Paul Samuelson, the first American to win the Nobel Memorial Prize in Economic Sciences, was notified about Bachelier's thesis by a statistician name Jimmie Savage and obtained a copy of the thesis from the Sorbonne. The idea of using *stochastic* methods to analyze economic phenomena like the movement of the stock price in Bachelier's thesis was innovative and had a profound influence on Samuelson's work. Bachelier's work had been circulated among economists by Samuelson in 1965(Samuelson, 1965). The term random walk became even more recognizable by the 1973 book of Burton Malkiel's, *A Random Walk Down Wall Street* (Malkiel, 1973). The *random walk hypothesis* asserts that price changes are unpredictable. It is consistent with the *efficient-market hypothesis* developed by Eugene Fama(Fama, 1965) asserting that it is impossible to beat the market, because stock market efficiency causes existing stock prices to always incorporate and reflect all relevant information.

Meanwhile the random walk was being introduced to economics, Harry Markowitz firstly established *risk,* measurement of uncertainty, into financial modeling in his seminal theory of portfolio selection in 1952 (Markowitz, 1952). His mean-variance model emphasized the rule that the investor did (or should) consider expected return a desirable thing and variance of return, which was called risk, an undesirable thing. The introduction of risk to portfolio allocation was novel, prior to his work the emphasis was placed on picking single high-yield stocks without any regard to their effects on portfolios as a whole. In 1973, Fischer Black and Myron Scholes published a famous option pricing model (Black & Scholes, 1973), namely *Black-Scholes formula*, that the price of the stock was assumed to follow a geometric Brownian motion with constant drift and diffusion coefficient. It is used to calculate the theoretical price of European put and call options. The term "Black-Scholes formula" was named by Robert C. Merton, who was the first to publish a paper expanding the mathematical understanding of the options pricing model (Merton, 1973). Afterward, Markowitz won the Nobel Memorial Prize in Economic Sciences in 1990 and so did Merton and Scholes in 1997, but Black was ineligible for the prize because of his death in 1995.

A brief story behind the presence of advanced mathematics that has changed the course of financial engineering has been adequately narrated in Jeremy Bernstein's essay, "Paul Samuelson and the Obscure Origins of the Financial Crisis"[1].

## 1.4 Organization

This thesis is arranged in 6 chapters. This chapter has introduced the motivation, the data set to be investigated and some introduction to financial mathematics. In Chapter 2, the theoretical framework including the definitions, theory, and the literature reviews are provided for further investigation in the following chapters. Then the marginal distributions of returns are investigated with a particular distribution in Chapter 3. This chapter shows how good the normal inverse Gaussian (NIG) distributions are fitted to the data.

Chapter 4 provides the guide to volatility forecasting consisting of all necessary related practical issues in volatility and also propose three forecasting models for volatility. The implementation in the chapter shows how well the forecasting models perform in the real situations. Chapter 5 adds up the contribution of exogenous variables in volatility estimating, especially range-based estimators. Several range-based volatility estimators are investigated by simulations in different scenarios. The information obtained from this chapter is preparatory to construct new volatility models in the last chapter.

Finally, Chapter 6 collects all the ideas and information obtained from previous chapters to introduce the DNIG model. The new stochastic volatility models that are tested to be accurate both in describing the distribution of returns and in volatility forecasting. Related practical information and extensive results are given in the appendix.

[1] Jeremy Bernstein (born December 31, 1929 in Rochester, New York) is an American theoretical physicist and science essayist.

# Chapter 2

# Financial Econometrics

This chapter includes theoretical framework in the financial literature, providing related definitions, theory, and the literature reviews of related works focusing on volatility modeling.

## 2.1 Introduction

In financial markets, there are quantities that we can observed at a curtain frequency such as closing, open, high and low prices, and trading volumes. These values are subject to uncertainty and unknown until they are observed. Mathematically, we regard this information as a real-valued *random variable* whose value is uncertain and can not be determined until it is observed. We denote $X$ a random variable in some *probability space* $(\Omega, \mathcal{F}, P)$ and $x$ its outcome or the observation. If a random variable takes possible values in an interval or a collection of intervals, we call it a continuous random variable. The possible values of a continuous random variable is described by a *cumulative distribution function* (cdf) $F$, where $F(x) \coloneqq P(X \leq x)$ with $P(\cdot)$ referring to the probability of the bracketed event. If the distribution function is differentiable, there exists the *probability density function* $f$, commonly abbreviated to pdf, where $f(x) \coloneqq dF/dx$. The *expectation* or *mean* of a continuous random variable $X$ is defined by $E[X] = \mu \coloneqq \int_{-\infty}^{\infty} xf(x)dx$. And the expectation of a function $g$ of $X$ can be computed by $E[g(X)] = \int_{-\infty}^{\infty} g(x)f(x)dx$. We define the $n^{\text{th}}$ moment of $X$ by $E[X^n]$ and the $n^{\text{th}}$ *central moment* is $\mu_n \coloneqq E[(X-\mu)^n]$. The *variance* of $X$, which measures the expected squared distance from the mean, is defined by $\text{var}(X) \coloneqq E[(X-\mu)^2] = E[X^2] - E[X]^2$. The variance is also denoted by $\sigma^2$ and the square root of variance is called the *standard deviation*. Furthermore, other two important quantities for describing a random variable are the *skewness* and the *kurtosis* defined by $\text{skew}(X) \coloneqq \mu_3/\mu_2^{3/2}$ and $\text{kurt}(X) \coloneqq \mu_4/\mu_2^2 - 3$.

For a pair of random variables $(X, Y)$, the *joint distribution* and the *joint density* functions are defined by $F(x, y) \coloneqq P(X \leq x \text{ and } Y \leq y)$ and $f(x, y) \coloneqq \partial^2 F/\partial x \partial y$ respectively. Provided that a particular outcome $X = x$ occurs, then the density of $Y$ conditional on the event $X = x$, namely the *conditional density* of $Y$ given $X = x$, is

defined by $f(y|x) \coloneqq f(x,y)/f_X(x)$. Here we use subscription to denote the underlying random variable e.g. $f_X$ is the density function of $X$. Consequently, the *conditional expectation* of $Y$ given $x$ is $E[Y|x] \coloneqq \int_{-\infty}^{\infty} yf(y|x)dy$. If the bivariate density $f(x,y)$ equals to the product of the two densities $f_X(x)f_Y(y)$ for all $x$ and $y$, then $X$ and $Y$ are *independent*; otherwise $X$ and $Y$ are *dependent*. The *covariance* is a measure of linear dependence between two random variables defined by $\text{cov}(X,Y) \coloneqq E[(X-\mu_X)(Y-\mu_Y)]$. The *correlation* is the covariance standardized by the standard deviations of the random variables, $\text{cor}(X,Y) \coloneqq \text{cov}(X,Y)/\sigma_X\sigma_Y$. The independence implies zero correlation, but the converse does not hold in general. For several variables, the definitions of related quantities are defined in the same way, see detail in Taylor (2005), Chapter 3.

A sequence of random variables $\{X_t\}$, with $t$ representing time, is called a *stochastic process*. Sometimes we call it the process generating observed data, or simply either a process or a model. A stochastic process can be either discrete or continuous depending on the time domain $t$. For a stochastic process $\{X_t\}$, the *autocovariance* and the *autocorrelation* of $X_t$ at lag $\tau$ are defined respectively by $\gamma_\tau \coloneqq \text{cov}(X_t, X_{t+\tau})$ and $\rho_\tau \coloneqq \text{cov}(X_t, X_{t+\tau})/\gamma_0$. For a stochastic process $\{X_t\}$, the information set available at time $t$ is denoted by $\mathcal{F}_t$. It often contains the history of observations up to time $t$, $\{X_s = x_s, s \leq t\}$ but additional relevant information known at time $t$ are also included. This information set employs the concept of conditioning by a random vector instead of conditioning by event (see Pfeiffer (1989) Chapter 19 and Taylor (2005), Section 3.2 and 9.5 ). The expectation of a random variable $X$ conditional on $\mathcal{F}_t$, denoted by $E_t[X] = E[X|\mathcal{F}_t]$, is called the conditional expectation; in the same way we denote the conditional variance and covariance. A stochastic process $\{X_t\}$ is said to be *stationary* if means, variances and covariances do not depend on time, that is, for all $t$ and $\tau$ we have $E[X_t] = \mu$, $\text{var}(X_t) = \sigma^2$ and $\text{cov}(X_t, X_{t+\tau}) = \gamma_\tau$. The time-ordered set of observations $\{x_1, x_2, ..., x_n\}$ is called a *time series*. The process generating the time series is usually unknown, our task is to exploit, infer and reasonably model the properties of the stochastic process driving the observations. Most interesting observed data in financial market are prices and returns of assets that we will introduce in the next section.

## 2.2 Prices and Returns

Denote by $P_t$ the *price* of an asset at time $t$ and assume that the asset pays no dividends. The return on investment is calculated from the change in price of the asset over a trading period. The *simple net return* $r_t^*$ on the asset between time $t-1$ and $t$ is defined as

$$r_t^* \coloneqq (P_t - P_{t-1})/P_{t-1} \tag{2.2.1}$$

and we call $1 + r_t^* = P_t/P_{t-1}$ the *simple gross return*. It is more convenient when we consider a $k$-period return over most recent $k$ trading periods, $r_{t,k}^*$, in term of simple gross returns

$$1 + r_{t,k}^* = P_t/P_{t-k} = (P_t/P_{t-1})(P_{t-1}/P_{t-2})...(P_{t-k+1}/P_{t-k}) = \prod_{j=0}^{k-1}(1 + r_{t-j}^*)$$

Thus the multi-period simple gross return is the product of single-period gross returns and the simple net return is simply the simple gross return minus one. The simple net return $r_t^*$ is often called a rate of return per a particular time period. The unit of time period in the academic literature are often specified as daily, monthly, or annual. Another definition of return is the *continuously compounded return* or *log return* (for period $t$) defined as

$$r_t = \log(1 + r_t^*) = \log\left(P_t/P_{t-1}\right) = p_t - p_{t-1} \tag{2.2.2}$$

where $p_t = \log(P_t)$ is the *log price* at time $t$. Log returns become preferable when consider multi-period returns because

$$r_{t,k} = \log(1 + r_{t,k}^*) = \sum_{j=1}^{k-1} \log(1 + r_{t-j}^*) = \sum_{j=0}^{k-1} r_{t-j}$$

does not involve multiplicative operation. Practically, simple return and log return are very similar numbers, since the Maclaurin series for $\log(1 + r_t^*)$ is

$$r_t = \log(1 + r_t^*) = r_t^* - \frac{1}{2}r_t^{*2} + \frac{1}{3}r_t^{*3} - ...$$

and daily returns are usually small lying between $-10\%$ to $10\%$. Throughout this thesis, except stated otherwise, return and price are generally referred to log return and log price respectively. Figure 2.2.1 shows the returns of EUR. It can be seen that there are more fluctuations during the crises. Moreover, large changes tend to be followed by large changes, of either sign, and small changes tend to be followed by small changes; this property is called *volatility clustering* that was firstly addressed by Mandelbrot (1963).

As we introduce the random walk hypothesis in Section 1.3, it states that prices wander in an unpredictable manner. There are several definitions of the random walk hypothesis in the literature. They usually incorporate models for price process with conditions expressing the idea of unpredictable movements. Here we give our first definition of random walk hypothesis (RWH1) by assuming that the price process follows a *Gaussian random walk* as the following. For a price process $\{p_t\}$, the dynamics of $\{p_t\}$ are given by the equation:

$$p_t \quad = \quad p_{t-1} + \mu + \sigma\epsilon_t, \ \epsilon_t \sim \text{i.i.d.} N(0,1) \tag{2.2.3}$$

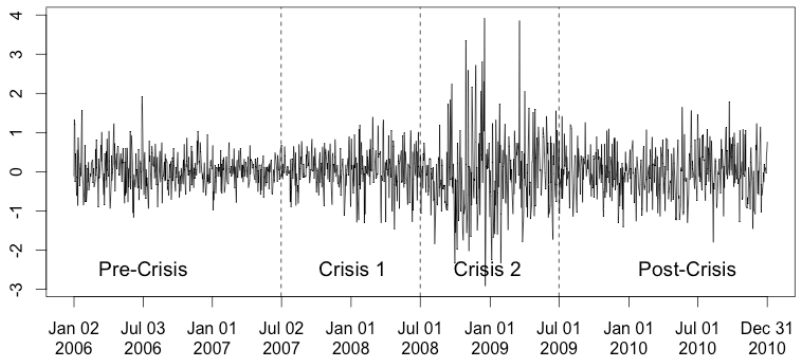
equivalently,

$$r_t = \mu + \sigma\epsilon_t \text{ and } p_t = p_0 + \sum_{i=1}^{t} r_i \tag{2.2.4}$$

where $\mu$ is the expected price change or *drift* and the error terms $\epsilon_t$ are *independent and identically distributed* () as standard normal. Denoting $r_t = p_t - p_{t-1}$ the increment of the process, then RWH1 can be given by the following conditions:

(i) the increments $r_t$ are independent;

Figure 2.2.1: EUR returns




Returns of EUR are more fluctuating during the crises. Moreover, large changes tend to be followed by large changes and small changes tend to be followed by small changes; this property is called volatility clustering.

(ii) the process $\{r_t\}$ is stationary ;

(iii) the increments $r_t$ are normally distributed, $r_t \sim N(\mu, \sigma^2)$.

Without declaring explicitly, it is generally assumed that a Gaussian random walk is driftless ($\mu = 0$), otherwise it is stated as a Gaussian random walk with drift ($\mu \neq 0$). This definition is the restrictive version of the Random Walk 1 model in Campbell et al. (1997). According to Section 2.1 of Campbell et al. (1997), there are three versions of random walk hypothesis conditioned by the dependence that can exist between the increments. A more general version of the random walk hypothesis, corresponding to the Random Walk 3 model in Campbell et al. (1997), is obtained by replacing the independence condition in (i) by uncorrelated increments and omitting the Gaussian condition in (iii). Hence the second definition of the random walk hypothesis () is given by

$$\{r_t\} \text{ is stationary and } \operatorname{cov}(r_t, r_{t+\tau}) = 0 \text{ for all } t \text{ and all } \tau > 0.$$

Remark that in the case of the Gaussian random walk, uncorrelatedness and independence are equivalent. Clearly, the RWH1 implies the RWH2. Widely used tests of the random walk hypothesis such as the Q-test of Box & Pierce (1970) and the variance-ratio test of Lo & MacKinlay (1988) employs sample autocorrelations and hence are tests of RWH2. These tests reject RWH1 whenever they reject RWH2. The uncorrelated hypothesis is of more attention because the i.i.d. hypothesis is not very relevant if we are interested in the predictability of returns. Taylor (2005) also discusses definitions of the random walk hypothesis.

Various kinds of dependence between the increments can be characterized by consid-

ering the covariance

$$\text{cov}\left(f(r_t), g(r_{t+\tau})\right) = 0 \tag{2.2.5}$$

for all $t$ and for $\tau \neq 0$, where $f(\cdot)$ and $g(\cdot)$ are two arbitrary functions. If $f(\cdot)$ and $g(\cdot)$ are restricted to be arbitrary linear functions, then (2.2.5) implies that the increments are *(serially) uncorrelated*. If one of either $f(\cdot)$ or $g(\cdot)$ is restricted to be linear while the other is unrestricted, then (2.2.5) is equivalent to the *martingale hypothesis* stating that tomorrow's price is expected to be equal to today's price, given the asset's entire price history (see Campbell et al. (1997)). The martingale hypothesis is a necessary condition for an efficient market, where the current price fully reflects the information contained in past prices. Finally, if (2.2.5) holds for all arbitrary $f(\cdot)$ and $g(\cdot)$, it implies that the increments are *(mutually) independent*.

## 2.3 Stylized Facts for Financial Returns

Statistical properties of financial returns have been studied and documented broadly across time as well as across markets. The properties that are commonly presented in any set of returns are called *stylized facts* or *stylized features* for financial returns. The statistical features of the distribution of a set of returns can be summarized by four statistics: *sample mean* ($\bar{r}$), *sample variance* ($s^2$), *sample skewness* ($w$), and *sample kurtosis* ($k$). For a set of returns $\{r_1, r_2, ..., r_n\}$, these statistics are defined by

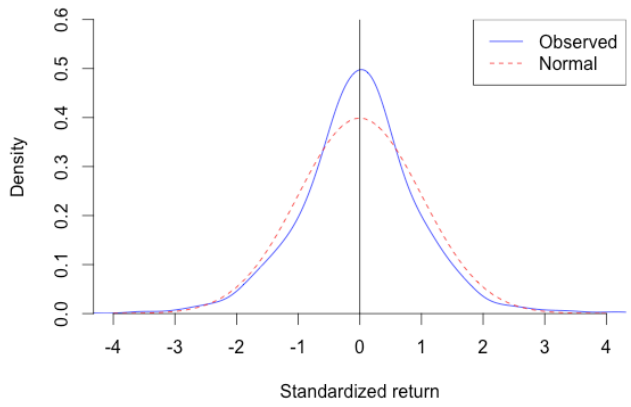
$$\bar{r} = \frac{1}{n}\sum_{t=1}^{n} r_t, \qquad s^2 = \frac{1}{n-1}\sum_{t=1}^{n}(r_t - \bar{r})^2,$$

$$w = \frac{1}{n-1}\sum_{t=1}^{n}\frac{(r_t - \bar{r})^3}{s^3}, \qquad k = \frac{1}{n-1}\sum_{t=1}^{n}\frac{(r_t - \bar{r})^4}{s^4} - 3.$$

Note that the square root of sample variance is the *sample standard deviation*. These statistics are the estimates of population mean, variance, skewness and kurtosis respectively. They are generally used to describe the shape of the inferred distribution. The regular reference is the standard normal distribution; suppose that $X \sim N(0,1)$, then $E[X] = 0$, $\text{var}(X) = 1$, $\text{skew}(X) = 0$ and $\text{kurt}(X) = 0$. A distribution that the kurtosis is positive is said to be *leptokurtic*. Skewness statistics are sometimes used to assess the symmetry of distributions, whereas kurtosis statistics are usually interpreted as a measure of similarity to normal distribution.

Taylor (2005) documents statistical features of twenty daily returns range from January 1991 to December 2000, containing returns from equity investments in indices or individual stocks, currency exchanges, commodity, bill and bond contracts. He found that all twenty sets of returns were leptokurtic and nineteen of the twenty had excess kurtosis more than ten of those standard errors. This is a clear evidence that the returns-generating process is far from normal. However, he argued that the presence of skewness in some sets of returns might be a consequence of very occasional negative outliers. According to Taylor (2005), there are three major stylized facts that are found in almost all sets of daily returns obtained from those prices.

Figure 2.3.1: Kernel density of EUR returns




Kernel density of EUR returns is approximately symmetric, has higher peak and fatter tails than that of normal distribution.

1. First, the distribution of returns is not normal.

2. Second, there is almost no correlation between returns for different days.

3. Third, the correlation between the magnitudes of returns on nearby days are positive and statistically significant.

The incidents of the three major properties are also found across time as well as across markets in Harrison (1998) and Mitchell et al. (2002).The first major stylized fact speaks of the distribution of daily returns that can be said more specifically as: it is approximately symmetric, has a high peak and it has fatter tails than that of a normal distribution. Here we roughly define a tail of a distribution that is fatter than that of a normal distribution as a *heavy tail*. The exact definition of a heavy-tailed distribution and further discussion will be given in Section 2.4. The evidence of a high peak in empirical distributions was shown by the greater number of observations that lied between $\bar{r}-0.5s$ and $\bar{r}+0.5s$ than that of a normal distribution. While the greater numbers of extreme observations below $\bar{r}-3s$ or above $\bar{r}+3s$ than that of a normal distribution corresponded to two heavy tails. The heavy tails indicates that there are more chances that extreme events, so called *outliers*, occurs. In Taylor (2005), outliers such that returns are more than three standard deviations from the mean is about four times the normal figure; the extreme outliers that returns are more than four standard deviations from the mean is approximately sixty times the normal figure.

Figure 2.3.1 compares kernel estimates of the probability distribution for standardized returns, $z_t = (r_t - \bar{r})/s$, with the normal distribution for EUR returns. These density estimates have been calculated in R using a generic function 'density'. The kernel density

estimator, $\hat{f}(z)$, is expressed as

$$\hat{f}(z) = \frac{1}{nB} \sum_{t=1}^{n} \phi \left( \frac{z - z_t}{B} \right)$$

where $\phi(\cdot)$ is the density of the standard normal distribution and $B$ is the bandwidth, a smoothing parameter. For a distribution with unit variance, it is acceptable to use $B = n^{-1/5}$.[1]

The first stylized fact is a principal guideline for modeling a probability distribution of daily returns. A satisfactory probability distribution for daily returns must have high kurtosis and be either exactly or approximately symmetric. Several distributions with these properties have been reviewed in Taylor (2005), including the generalized Student's t, the lognormal-normal, the normal inverse Gaussian and the generalized hyperbolic distributions.

The second stylized fact is of the dependence between the returns for time periods $t$ and $t + \tau$. The dependence is measured by the *sample autocorrelation* at *lag* $\tau$ that estimates the correlation between $\tau$ periods apart returns from $n$ observations;

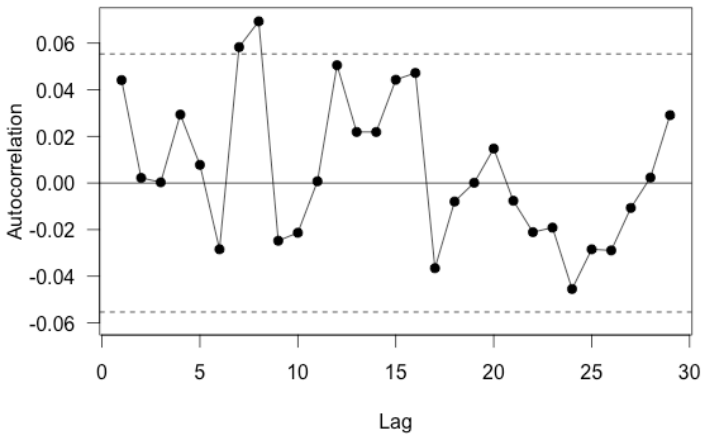
$$\hat{\rho}_{\tau,r} = \sum_{t=1}^{n-\tau} (r_t - \bar{r})(r_{t+\tau} - \bar{r}) \Big/ \sum_{t=1}^{n} (r_t - \bar{r})^2, \ \tau > 0. \tag{2.3.1}$$

The sample autocorrelation $\hat{\rho}$ is the estimator of an autocorrelation parameter $\rho$ in a stationary stochastic process that the autocorrelation between any pair of random variables only depends on the lag. The autocorrelation estimates can be used to test the hypothesis that the process of interest is generated by uncorrelated random variables. The standard error of an autocorrelation estimate is approximately $1/\sqrt{n}$, so that an autocorrelation estimate reject the null hypothesis of zero autocorrelation at 5% level of confidence if it lies outside the confidence interval $(-2/\sqrt{n}, 2/\sqrt{n})$. Figure 2.3.2 shows the autocorrelations for EUR returns up to lag 30 with the 95% confidence interval about zero. It is clear that most of the autocorrelations are not significantly different from zero. Taylor (2005) finds that more than 90% of 600 sample autocorrelations at lag 1 to 30 are between -0.05 and 0.05. Not only that some 99% of the estimates are between -0.1 and 0.1. This is an evidence of the absence of linear dependence in the stochastic process generating daily returns. Taylor also tests the hypothesis that the process generating observed returns is a series of i.i.d. random variables using the portmanteau Q-statistic of Box & Pierce (1970), it results that most of the returns processes, 14 of the 20, are not i.i.d at 5% level of confidence.

Even though the lack of dependence between returns for different day, the dependence between absolute returns, likewise squared returns, on nearby days is positive; this is the third major stylized fact. According to Taylor (2005), he shows that all estimates for the first thirty lags exceed 0.05 and are significant at the 1% level for tests of i.i.d.

[1] Silverman (1986) gives a rule-of-thumb for choosing the bandwidth of a Gaussian kernel density estimator that is expressed as $B = \left(\frac{4S^5}{3n}\right)^{\frac{1}{5}} \approx 1.06Sn^{-\frac{1}{5}}$.

Figure 2.3.2: EUR autocorrelations

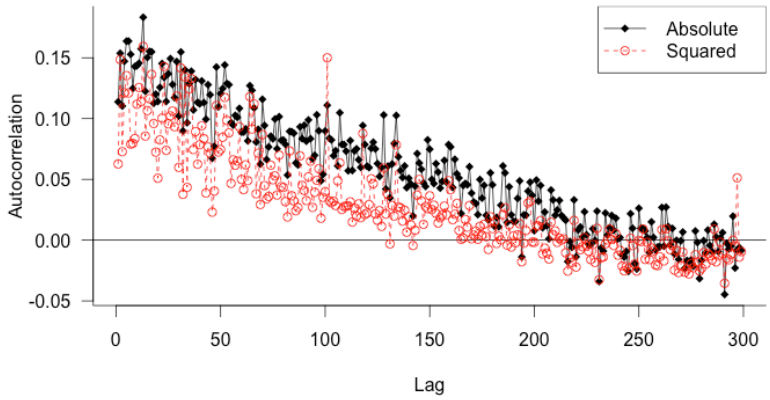



Autocorrelations of returns in the first 30 lags are not significantly different from zero. It can be said that there is almost no correlation between returns for different days.

hypothesis. The linear dependence among absolute returns and among squared returns is evidently far more than that among returns. At the first lag, his results show that a high value of $|r_t|$ tends be followed by a high value of $|r_{t+1}|$. He also demonstrates that there is a considerable number of lags that the across-series averages of the autocorrelations for absolute returns have more dependence than that for squared returns. The averages of the autocorrelations seem to decline slowly as the number of lags increase; Taylor argues that the decline of the averages of the autocorrelations do not give evidence of a long-memory property in the individual series. All the statistical properties of the returns and transformed returns found in Taylor (2005) are also found in our three series of returns.

Figure 2.3.3 shows the averages of the autocorrelations for the absolute and squared returns of the three currencies. The averages of the autocorrelations are positive at nearby lags and decrease at further lags. Most of the averages of the autocorrelations for absolute returns are greater than that for squared returns. In conclusion, the characteristics of returns time series are as the followings: there are very little autocorrelations present in series of returns $\{r_t\}$, the autocorrelations of absolute returns are positive up to several further lags, and the autocorrelations of squared returns are also positive with lesser degree. From the dependence condition (2.2), it is clear that the returns-generating process is (serially) uncorrelated but not (mutually) independent. The incidents of the three major stylized are found in several studies across time and markets; Hsieh (1988),Boothe & Glassman (1987),Campbell et al. (1997).

Figure 2.3.3: Autocorrelations of absolute and squared returns; averages across 3 currencies




Average autocorrelations of absolute and squared returns across the series of three currencies are significantly greater than zero in some substantial lags and decline slowly as the number of lags increases. This indicates the existence of dependency between absolute returns and squared returns.

## 2.4 Heavy Tails & the CV-Plots

The presence of heavy tails in the distribution of returns stated in the first major stylized fact plays major role in this section. In Section 2.3, according to Taylor's results, the heavy tails were addressed by the number of extreme observations that are either below or above the mean by three times the standard deviation. This property is actually refers to leptokurtic distributions whose kurtosis are greater than zero and consequently extreme values are "more probable than normal". A more precise definition of heavy-tailed distribution is given by considering the tail distribution. Letting $F$ be the distribution of a random variable $X$. The tail distribution of $X$, also known as a *survival function* or *reliability function*, is defined as $\bar{F}(x) := 1 - F(x) = P(X > x)$.

**Definition 1.** The distribution $F$ has a (right-) heavy tail[2] if

$$\lim_{x\to\infty} e^{-\lambda x}\bar{F}(x) = \infty, \text{ for all } \lambda > 0.$$

Some authors use the word 'long tail' instead of 'heavy tail' in this definition. The tail distribution $\bar{F}$ of a heavy-tailed distribution is said to be a heavy-tailed function. Some examples of heavy-tailed distributions are the Pareto distribution, the Cauchy distribution, , the Student's t distribution, and the Weibull distribution. The Weibull

[2]See Foss et al. (2013).

distribution has tail distribution $\bar{F}$ given by $\bar{F}(x) = \exp\left(-(x/\kappa)^{\alpha}\right)$ for some scale parameter $\kappa > 0$ and shape parameter $\alpha > 0$. The Weibull distribution is heavy-tailed if and only if $\alpha < 1$. The exponential distribution is a particular case of the Weibull distribution where $\alpha = 1$. We can also say that a heavy-tailed distribution is a distribution that has a tail that is heavier than an exponential. A fundamental theorem in extreme value theory that we will regularly employ in this section is the Pickands–Balkema–de Haan theorem proposed by Pickands (1975); Balkema & de Haan (1974). In first instance, let $X$ be a continuous non-negative random variable with distribution function $F$. For any threshold $u > 0$, the distribution function of *threshold excedances*, $X_u = (X - u|X > u)$, denoted by $F_u$, is defined as

$$1 - F_u(x) = \frac{1 - F(x+u)}{1 - F(u)} \text{ or equivalently, } \bar{F}_u(x) = \frac{\bar{F}(x+u)}{\bar{F}(u)}.$$

**Theorem 2** (Pickands-Balkema-de Haan)**.** *Let $X_u = (X_u - u|X > u)$ with support at $(0, \infty)$. Then, for any distribution, $F(x)$, we have*

$$F_u(x) \to GPD(x; \xi, \psi) \text{ as } u \to \infty$$

*where $GPD(\cdot; \xi, \psi)$ is the generalized Pareto distribution () with parameters $\xi$ and $\psi$.*

The GPD function is defined by

$$GPD(x; \xi, \psi) = \begin{cases} 1 - \left(1 + \frac{\xi x}{\psi}\right)^{-\frac{1}{\xi}} & \text{if } \xi \neq 0 \\ 1 - \exp\left(-\frac{x}{\psi}\right) & \text{if } \xi = 0 \end{cases}.$$

The GPD is the Pareto distribution if $\xi > 0$, it is the exponential distribution if $\xi = 0$ and it is a distribution with compact support if $\xi < 0$. The Pareto distribution has polynomial tails whereas the exponential distribution has exponential tails. Here we introduce a method to distinguish the behavior of tails by comparing to the exponential distribution. The method of CV-plot is proposed by del Castillo et al. (2014), it is a graphical method to show departures from exponentiality in the tails. It relies on the *residual coefficient of variation* () of the conditional excedance over a threshold,$u$, defined by

$$\mathrm{CV}(u) = \mathrm{var}(X - u|X > u)^{1/2} / E[X - u|X > u].$$

The $CV(u)$ is independent of scale parameter. *The empirical CV of of the conditional excedance* for a sample $\{x_j\}$ of size $n$ is given by

$$\mathrm{cv}_n(u) = s_u / \bar{x}_u \tag{2.4.1}$$

where $\bar{x}_u$ and $s_u$ are the mean and the standard deviation of the set of exedances over the threshold $u$, $\{x_j - u \mid x_j > u,\ j = 1, .., n\}$ respectively. The $\mathrm{cv}_n(u)$ is independent of scale and it is a consistent estimator of $\mathrm{CV}(u)$ provided the second moment of $X$ is finite. Let $\{x_{(j)}\}$ be the ordered sample of $\{x_j\}$ such that $x_{(1)} \leq x_{(2)} \leq ... \leq x_{(n)}$,

then the *CV-plot* is the representation of the empirical CV of the conditional excedance (2.4.1) given by $j \to \text{cv}(x_{(j)})$. The CV-plot does not depend on scale parameter so that the CV-plot $\{x_j\}$ and $\{\lambda x_j\}$ are identical. The CV-plot is used to distinguish the tails of the random sample $\{x_j\}$ from that of the exponential tails. del Castillo et al. (2014) also set up theory to determine pointwise error limits for the CV-plots from the null hypothesis of exponentiality. Let $n(u)$ be the number of observations in the set $\{x_j - u \mid x_j > u,\ j = 1, ..., n\}$.

**Proposition 3.** *Let $X$ be a random variable with an exponential distribution with mean $\mu$, then $\sqrt{n(u)}(cv(u) - 1)$ converges to a Gaussian process $\{X_t\}$ with zero mean and covariance function given by*

$$cov(X_s, X_t) = \exp\left(-|t - s|/(2\mu)\right).$$

*This is the covariance function of the Ornstein-Uhlenbeck process, the continuous time version of an AR(1) process. It is a stationary Markov Gaussian process. In particular, for any fixed $u$*

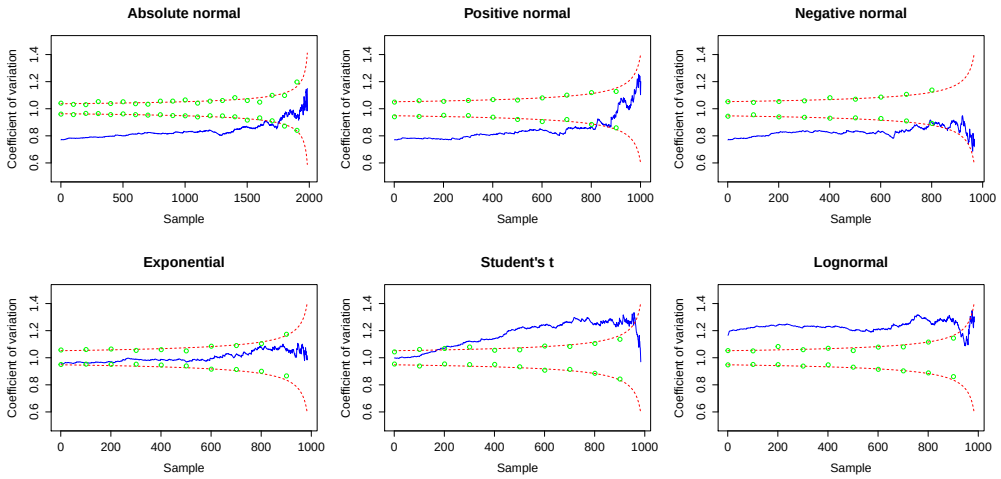
$$n(u)(cv(u) - 1) \xrightarrow{d} N(0, 1). \tag{2.4.2}$$

Therefore, pointwise error limits for the CV-plot are calculated from (2.4.2) (the symbol 'd' refers to the convergence in distribution that will be introduced in Section 2.6).

Figure 2.4.1 shows the CV-plots of samples from different distributions with 95% pointwise limits around $cv = 1$. The set of positive sample is called the positive part and the set of minus the negative sample is called the negative part. In the case of normal distribution, we can see that the *cv* are mostly below the lower limit, and they enter the error limits when the thresholds are sufficiently large. This is because the normal distribution has lighter tails than that of the exponential distribution, and for a sufficiently large threshold the tail distributions converge to the GPD as stated in the Pickands–Balkema–de Haan theorem. The exponential distribution always has its *cv* inside the error limits, this can be used as a reference. In the cases of the Student's t and the lognormal distributions, most of the *cv* are over the upper error limits. This is indicative of heavy-tailed distributions. The *cv* enter the error limits at large thresholds, however, these *cv* are not relevant because the sample size, i.e., the number of excedances over the thresholds are too small.

We apply the CV-plots to the returns for the three currencies. The results shown in Figure 2.4.2 indicate that all the positive tails and also the absolute returns are over the upper error limits while the negative tails are mostly inside the error limits. It is clear that the hypothesis of exponentiality is rejected in all cases. This implies that the returns for these currencies are from distributions with heavy tails which is compatible with the first major stylized fact for returns.

Figure 2.4.1: CV-plots of some distributions




CV-plots of samples from normal distribution, exponential distribution, Student's t distribution and lognormal distribution. The $cv$ of normal distribution are below the lower limit about $cv = 1$ because it has lighter tails than that of exponential distribution. Exponential distribution $cv$ are always in the 95% confidence intervals about $cv = 1$ whereas Student's t distribution and lognormal distribution have heavier tails than that of exponential distribution. At large thresholds, the $cv$ of all distributions enter the confidence intervals about $cv = 1$. However, numbers of excedances over large thresholds are too small and the empirical $cv$ are not relevant.

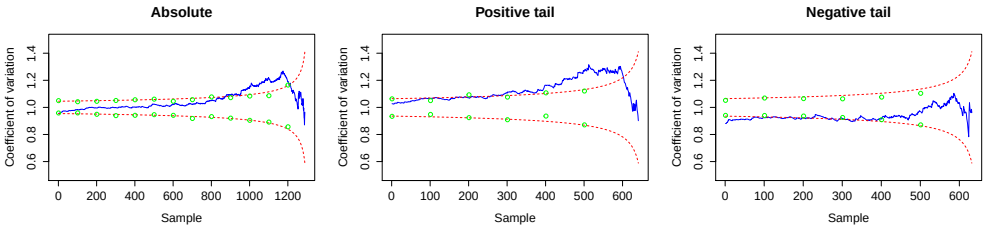
## 2.5 Discrete Time Models

Accordingly, the random walk hypothesis provides the basic idea of price process and the three major stylized facts give a clear-cut direction to model daily returns. Any satisfactory statistical model for daily returns must be consistent with the three major stylized facts that are of prominence. The third major stylized fact for returns is indicative of positive autocorrelations among absolute returns and squared returns; this fact implies the volatility clustering property stated in Section 2.2 such that changes in price are not constant. The measure of price variability over some period of time is called *volatility*. Typically, volatility describes the standard deviation of returns but the definition may vary in different contexts. In the RWH1 model, volatility is the parameter $\sigma$ which describes the standard deviation of returns that is assumed constant for all time $t$. However, in financial markets, it seems that volatility increases during crises and then decrease in at appropriate time. For example, in Figure 2.2.1 the variations of prices are clearly higher during the crises than that of normal periods. Even though there is no complete explanation why volatility changes, it is more relevant to model asset price with changing volatility.

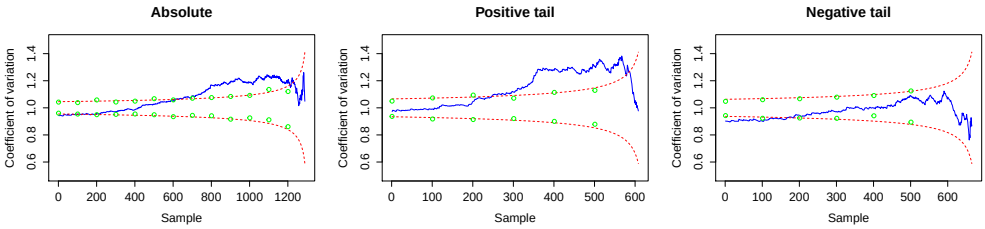
As distinguished from the RHW1 model that assume constant mean and constant variance for returns, the standard formulation of daily returns that have been widely

Figure 2.4.2: CV-plots of returns for currency exchanges

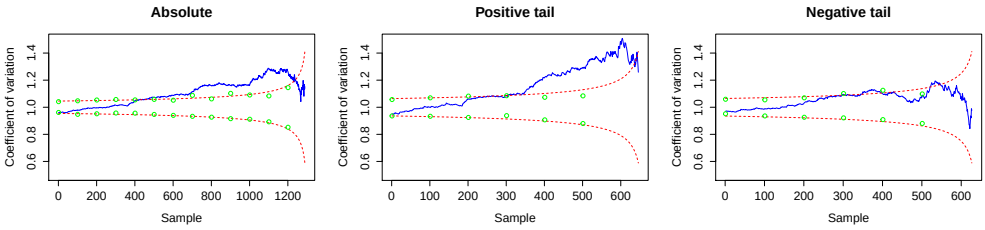
(a) EUR returns




(b) JPY returns




(c) GBP returns




CV-plots of returns from EUR, JPY and GBP show that the distributions of returns are heavy tailed because most of the $cv$ are above the upper limits of $cv = 1$.

accepted currently is given by

$$r_t = \mu + \sigma_t \epsilon_t \tag{2.5.1}$$

where $\mu$ is the expected return, $\epsilon_t \sim$ i.i.d.$N(0,1)$ are random errors and $\sigma_t$ are time-varying volatilities. As a result $r_t$ are normally distributed with constant mean $\mu$ and variance $\sigma_t^2$. Denoting the residual or excess return by $y_t = r_t - \mu$, another common formulation is

$$y_t = \sigma_t \epsilon_t. \tag{2.5.2}$$

There are two main classes of models that use different approaches to model the volatility in (2.5.2), *ARCH* models and *stochastic volatility* (SV) models. ARCH models specify a process for the conditional variance of returns by a linear function of past observations, while SV models specify a stochastic process for volatility.

### 2.5.1 ARCH models

*Autoregressive conditional heteroskedastic* (ARCH) processes have been introduced by Engle (1982). He presented a stochastic process whose variables have conditional mean zero and conditional variance given by a linear function of previous squared variables. In the financial econometric contexts, the variable of interest is the return from an asset. The changes in conditional variance of return give us the word *conditional heteroskesastic* and the word *autoregressive* comes from the autoregressive process of squared residuals in his pioneering research. The simplest specification of ARCH process is ARCH(1). Given that the residuals follow (2.5.2). ARCH(1) process is given by

$$\sigma_t^2 = \omega + \alpha y_{t-1}^2 \tag{2.5.3}$$

where the volatility parameters $\omega > 0$ and $\alpha > 0$ are strictly positive to ensure the positivity of the conditional variance $\sigma_t^2$ and the case that $\alpha = 0$ is out of interest. Therefore the conditional distribution of the return is normal, $r_t|\mathcal{F}_{t-1} \sim N(\mu, \sigma_t^2)$. The conditional changes in the scale variable $\sigma_t$ entitles the *conditional heteroskesastic* (CH) part of the acronym ARCH. This ARCH(1) specification results that the volatility of the return in period $t$ depends only on the previous return. The general formulation of ARCH(q) model is

$$\sigma_t^2 = \omega + \sum_{j=1}^{q} \alpha_j y_{t-j}^2 \tag{2.5.4}$$

with $\omega > 0$ and $\alpha_j \geq 0$. The process is stationary if $\sum_{j=1}^{q} \alpha_j < 1$. Typically, ARCH(p) model can not describe the returns process successfully with low order of $p$ because of the phenomenon of volatility persistence in financial markets (see Section 9.2 in Taylor, 2005 ). This leads to the generalization of ARCH which becomes the best known specification, *GARCH* (generalized ARCH) models proposed by Bollerslev (1986). The GARCH (1,1) model, which is the simplest, yet the most popular model in empirical research, is given by

$$\sigma_t^2 = \omega + \alpha y_{t-1}^2 + \beta \sigma_{t-1}^2 \tag{2.5.5}$$

with $\omega > 0$, $\alpha \geq 0$ and $\beta \geq 0$. The unconditional variance of the return equals $\sigma^2 = \frac{\omega}{1-\alpha-\beta}$. The GARCH(1,1) model is appreciated because it has decent advantages, yet the model is simple with only three parameters. Following Taylor (2005), the major properties of a GARCH(1,1) process, provided $\alpha + \beta < 1$, can be summarized as: the unconditional variance is finite; the unconditional kurtosis is always positive and can be finite; the correlation between the squared return $r_t$ and $r_{t+\tau}$ is zero for all $\tau > 0$; and the correlation between the squared excess returns $y_t^2$ and $y_{t+\tau}^2$ is positive for all $\tau > 0$ and equals $C(\alpha + \beta)^\tau$, with $C$ positive and determined by both $\alpha$ and $\beta$. These properties are adequately consistent with the three major stylized facts for returns. The general formulation of GARCH(p,q) is defined by

$$\sigma_t^2 = \omega + \sum_{i=1}^{p} \alpha_i y_{t-i}^2 + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2.$$

The popularity of ARCH models leads to several specifications, for examples, nonlinear GARCH (NGARCH) from Engle (1990), the exponential GARCH (EGARCH) from Nelson (1991), threshold GARCH (TGARCH) from Glosten et al. (1993) and asymmetric power ARCH (APARCH) from Ding et al. (1993). Some reviews in the literature on ARCH models are in Bollerslev et al. (1992), Bauwens et al. (2006) and Teräsvirta (2009).

Let $\mathcal{F}_{t-1}$ be the information set know at time $t-1$. The distribution of return conditional on past history is

$$r_t|\mathcal{F}_{t-1} \sim N(\mu, \sigma_t^2) \text{ , or equivalently, } y_t|\mathcal{F}_{t-1} \sim N(0, \sigma_t^2). \tag{2.5.6}$$

We also denote $\Theta$ as a vector of parameters. For example, the parameter vector of GARCH(1,1) model is $\Theta = (\mu, \omega, \alpha, \beta)'$. The knowledge of the conditional distributions of returns allows us to form the likelihood function with ease. Given a set of $n$ observed returns $\{r_1, r_2, ..., r_n\}$. The first parameter that could be estimated is the mean $\mu$ that is estimated by the sample mean $\bar{r}$. Then it is more convenient to deal with the excess returns $\{y_1, y_2, .., y_n\}$ where $y_i = r_i - \bar{r}$. Because the conditional distributions of the excess returns are also known but the less number of parameters are to be estimated. The likelihood function is a function of $\Theta$ which is constructed by the product of conditional densities $f(y_t|\mathcal{F}_{t-1})$,

$$L(\Theta) = f(y_1|\mathcal{F}_0) \cdot f(y_2|\mathcal{F}_1) \cdots f(y_n|\mathcal{F}_{n-1}). \tag{2.5.7}$$

Maximizing the likelihood $L(\Theta)$ gives an appropriate estimate of the parameters $\Theta$. The resulting estimate is equivalent to maximizing the logarithm of $L(\Theta)$. The log-likelihood function is

$$l(\Theta) = \log L(\Theta) = \sum_{t=1}^{n} \log f(y_t|\mathcal{F}_{t-1}, \Theta),$$

which is a lot easier to optimize. From (2.5.6), the conditional distributions $y_t|\mathcal{F}_{t-1}$ are normal. Hence, the log-likelihood function $l(\Theta)$ can be explicitly written as

$$l(\Theta) \quad = \quad \sum_{t=1}^{n} \left( -\frac{1}{2}\log(2\pi) - \frac{1}{2}\log(\sigma_t^2) - \frac{y_t^2}{2\sigma_t^2} \right). \tag{2.5.8}$$

Maximization of (2.5.8) provides the maximum likelihood estimate $\hat{\Theta}$.

### 2.5.2 Stochastic volatility models

In contrast to ARCH models that the conditional variance is specified by a function of past observations, *stochastic volatility* (SV) models directly specify a stochastic process for volatility. Therefore the properties of SV models can be designed via the stochastic process generating volatility. SV and ARCH models explain the same stylized facts for returns. While ARCH models are more popular because of their ease of maximum likelihood estimation, SV models arise naturally in derivative pricing theory. The SV literature has its origin in Rosenberg (1970), Clark (1973), Taylor (1982) and Tauchen & Pitts (1983). Recall the formulation of daily returns in (2.5.1), returns in excess of a constant mean $\mu$ is

$$y_t = \sigma_t \epsilon_t, \ \ \epsilon_t \sim \text{i.i.d.} N(0,1).$$

SV models involve two conditions: first the volatilities $\{\sigma_t\}$ follow a positive stationary stochastic process, second the processes $\{\sigma_t\}$ and $\{\epsilon_t\}$ are stochastically independent[3]. The *standard SV model* of Taylor (1986) is given by a Gaussian AR(1) process for its logarithm,

$$\log(\sigma_t) = \alpha + \beta \log(\sigma_{t-1}) + \eta_t. \tag{2.5.9}$$

The parameter $\beta$ represents volatility persistence, with $-1 < \beta < 1$. The volatility residuals $\eta_t$ are i.i.d. normally distributed as $\eta_t \sim \text{i.i.d.} N(0, \sigma_\eta^2)$. The standard SV model has received more attention than any other SV specifications because it holds the following properties: all the moments of returns are finite; the kurtosis of returns is positive; the correlation $\text{cor}(r_t, r_{t+\tau})$ is zero and the correlation of squared excess returns, $\text{cor}(y_t^2, y_{t+\tau}^2)$, is positive for all $\tau > 0$; finally the autocorrelation function of $|y_t|^p$ has approximately the same shape as of $y_t^2$ for all positive $p$. Nevertheless, maximum likelihood estimation for the standard SV model is complicated and hence the parameters are estimated by alternative methods such as *quasi-maximum likelihood* (QML) methods, the *generalized method of moments* (GMM) or the MCMC method. These estimation methods for the standard SV model can be seen in Taylor (2005). Recently, Lee et al. (2011) introduced the *hierarchical-likelihood* approach to estimate the standard SV model.

Other than the standard SV model, various stochastic processes for $\{\sigma_t\}$ have been proposed. If the distribution of $\sigma_t^2$ is assumed properly, then the distribution of returns is a mixture of normal distributions with higher kurtosis than that of normal distribution and their autocorrelations are zero at all positive lags. There are several suggestions about the distribution of $\sigma_t^2$ in the literature. Clark (1973) proposes a lognormal distribution for $\sigma_t^2$, result in a lognormal-normal distribution for returns. Moreover, there are gamma distribution (Madan & Seneta, 1990), *inverse gamma* distribution (Praetz, 1972) and *inverse Gaussian* (IG) distribution(Barndorff-Nielsen, 1997) which are particular cases of the generalized inverse Gaussian (GIG) distribution. Specially, when the

[3]The vector variables $(\sigma_1, \sigma_2, ..., \sigma_n)$ and $(\epsilon_1, \epsilon_2, ..., \epsilon_n)$ are independent for all positive integers $n$.

distribution of $\sigma_t^2$ is IG or GIG, the distribution of returns is normal inverse Gaussian (NIG) or generalized hyperbolic (GH) respectively. The GH and the GIG distributions are both *infinitely divisible*, proven by Barndorff-Nielsen & Halgreen (1977). The merit of this property will be discuss in Section 2.6. Remark that volatility is latent and unobservable, hence the estimation of model parameters is certainly complicated.

### 2.5.3 Volatility estimates

Volatility is latent variable but it is an important input in several financial models. For this reason either the estimate or forecast of volatility is necessary. Denote $\sigma_t$ as the volatility of an asset at time $t$, the squared of the volatility, $\sigma_t^2$, on period $t$ is the *variance rate*. A standard way to estimate the volatility, $\sigma_t$, at the end of period $t$ using the most recent $m$ observations on the return $\{r_t\}$ is

$$\hat{\sigma}_t^2 = \frac{1}{m-1}\sum_{i=1}^{m}(r_{t-i} - \bar{r})^2 \tag{2.5.10}$$

where $\bar{r} = \frac{1}{m}\sum_{i=1}^{m} r_{t-i}$ is the mean of returns of last $m$ observations. This estimator, $\hat{\sigma}_t$, is called *realized volatility* or *historical volatility*. Unambiguously, some authors define realized volatility using intra-daily data, which is not applied in this thesis. Volatility is usually expressed in term of *annualized volatility* representing the volatility per year. The number of trading days per year is regularly assumed to be 252, thus the annualized volatility calculated from daily volatility $\sigma_t$ is approximately $\sqrt{252}\sigma_t$. Suppose that the return process has a constant mean $\mu$, that is estimated by $\bar{r}$. The *excess return* of the process $\{r_t\}$ are $y_t = r_t - \mu$, that can be estimated by $y_t \approx r_t - \bar{r}$. Replacing $\frac{1}{m-1}$ by $\frac{1}{m}$ in 2.5.10, the simplified formula is

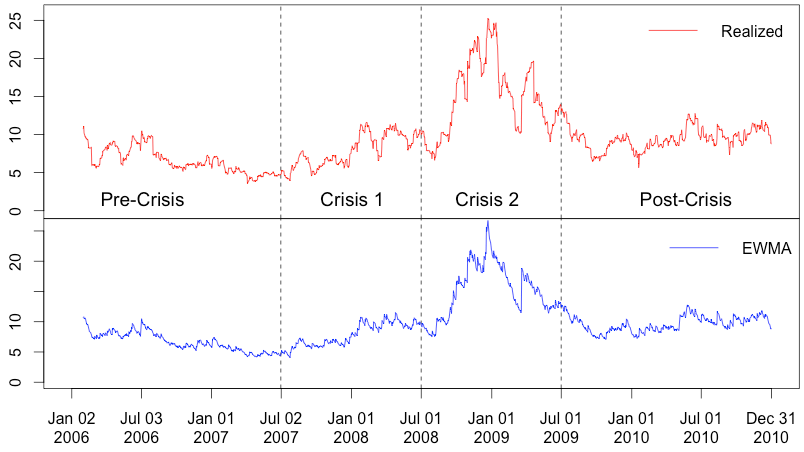
$$\hat{\sigma}_t^2 = \frac{1}{m}\sum_{i=1}^{m} y_{t-i}^2 \tag{2.5.11}$$

which makes very little difference to the variance estimates. The equation (2.5.11) gives equal weigh to all observations. It makes sense to give more weight to recent data to estimate the current level of volatility $\sigma_t$. The weighting scheme is given by

$$\hat{\sigma}_t^2 = \sum_{i=1}^{m} \alpha_i y_{t-i}^2 \tag{2.5.12}$$

The variable $\alpha_i > 0$ is the amount of weight assigned to the observation $i$ days ago. We can design how we weigh each observation properly but the weights must sum to unity, so that $\sum_{i=1}^{m} \alpha_i = 1$. This allows us to assign more significancy to the data that is believed to influence the process. Since we are estimating current level of volatility, it is appropriate that the most recent data are considered with higher weights. So we choose the weights, $\alpha_i$, in such a way that they decrease exponentially as we move back

Figure 2.5.1: Volatility estimates of EUR (annualized)




Volatility estimates of EUR using realized volatility model and EWMA model. The volatility estimates in the Crisis 2 are considerably higher than in other periods.

through time. Given that the weights are exponentially decreasing with rate $0 < \lambda < 1$, that is $\alpha_{i+1} = \lambda\alpha_i$. If $\alpha_1 = 1 - \lambda$, with some simple calculations, then

$$\hat{\sigma}_t^2 = (1-\lambda)y_{t-1}^2 + \lambda\sigma_{t-1}^2 - (1-\lambda)\lambda^m y_{t-1-m}^2$$

The term $(1-\lambda)\lambda^m y_{t-1-m}^2$ is sufficiently small to be ignored for large $m$. Finally, we arrive

$$\hat{\sigma}_t^2 = (1-\lambda)y_{t-1}^2 + \lambda\hat{\sigma}_{t-1}^2. \tag{2.5.13}$$

This is the *exponentially weighted moving average (*EWMA*)* model used in the RiskMetrics database, which was created by J.P. Morgan[4] and made publicly available in 1994. The parameter used in the RiskMetrics is $\lambda = 0.94$ for updating daily volatilities estimates. Figure 2.5.1 shows volatility estimates using 22-day realized volatility and EWMA with $\lambda = 0.94$. Volatility estimates are extremely high during the Crisis 2.

## 2.6 Continuous Time Models

A major breakthrough in financial engineering was made by a continuous time model, when Fischer Black, Myron Scholes, and Robert Merton presented the model for stock price and the formula for option pricing in the early 1970s. They developed probably the most celebrated of all models used in finance that has become known as the *Black-Scholes model*. The model has been greatly influential to practitioners and academics

[4]See J.P. Morgan, *RiskMetrics Monitor*, Fourth Quarter, 1995

on the way to price and hedge options. As a result of the development of the favorable model, Robert Merton and Myron Scholes were awarded the Nobel prize for economics in 1997. Fischer Black should have been awarded the prize as well but he passed away before in 1995. Accordingly, succeeding models in option and derivative pricing theory have been influenced by the Black-Scholes model and mostly rely on continuous time. In this section we present some continuous time models and some important properties that connect continuous time models to discrete time models.

### 2.6.1 Brownian motion

Beginning with the basic idea arising in the RWH1 model, it is sensible to think of the natural continuous-time version of the RWH1 process. *The (standard) Brownian motion*, which is also called the *Wiener process,* is a continuous-time stochastic process $B = \{B_t\}_{t\geq 0}$ satisfying the following properties:

(i) $B_0 = 0$ a.s., that is $P(B_0 = 0) = 1$;

(ii) $\{B_t\}$ has independent increments, that is for $0 \leq t_1 < t_2 \leq t_3 < t_4$, $B_{t_4} - B_{t_3}$ is independent of $B_{t_2} - B_{t_1}$ ;

(iii) $\{B_t\}$ has stationary increments, that is the distribution of $B_t - B_s$ only depend on the time difference $t - s$;

(iv) $B_t$ are Gaussian, that is $B_t = B_t - B_0 \sim N(0, \sigma^2 t)$.

The Brownian motion was named after Robert Brown, an English botanist who firstly observed the irregular motion of pollen grains in water in 1826. Later this motion was described in (plausible) mathematical terms by Bachelier in 1900, by Einstein in 1905, and by von Sinoluchovski in 1906. Nevertheless, Norbert Wiener was the first one who gave a rigorous mathematical derivation of this process in 1923, and so it is also called a Wiener process. The famous Black-Scholes option pricing formula was derived by assuming that the log price process follows a Brownian motion. In consequence, the price process follows the so-called *geometric Brownian motion* (GBM).

Seemingly, the RWH1 model and Brownian motion are intuitively agreeing. One of the formal constructions of Brownian motion employs the concept of weak convergence and the central limit theorem; which asserts that if $\{\xi_j\}_{j=1}^{\infty}$ is a sequence of i.i.d. random variables with mean zero and variance $\sigma^2 < \infty$, then $\{S_n\}$ defined by $S_n = \sum_{j=1}^{n} \xi_j$ converges in distribution to a random variable distributed as $N(0, \sigma^2 n)$. This theorem advised that a properly normalized sequence of random walks will converge in distribution to a Brownian motion. This idea is developed to the invariance principle of Donsker's (1951) which proves the convergence. Let us consider the sequence of partial sums $S = \{S_k\}_{k=0}^{\infty}$ where $S_0 = 0$ and $S_k = \sum_{j=1}^{k} \xi_j$ for $k \geq 1$. From $S$, we obtain a sequence of continuous-time process $X^{(n)} = \{X_t^{(n)}\}_{t\geq 0}$ with scaled linear interpolations

$$X_t^{(n)} = \frac{1}{\sigma\sqrt{n}} \left( S_{\lfloor nt \rfloor} + (nt - \lfloor nt \rfloor)\xi_{\lfloor nt \rfloor + 1} \right), \ t \geq 0 \tag{2.6.1}$$

where $\lfloor t \rfloor$denotes the greatest integer less than or equal to $t$. The following theorem is known as the invariance principle of Donsker's.

**Theorem 4** (Donsker, 1951)**.** *Let $\{\xi_j\}_{j\in\mathbb{N}}$ be a sequence of i.i.d random variables on $(\Omega, \Sigma, P)$ with zero means and finite variances $\sigma^2 > 0$. If $X^{(n)}$ is defined by (2.6.1), then*

$$X^{(n)} \xrightarrow{d} B \text{ as } n \to \infty$$

*where $B = \{B_t\}_{t\geq 0}$ is the standard Brownian motion or the Weiner process with $B_t \sim N(0, \sigma^2 t)$ and 'd' denotes the convergence in distribution.*

The definition of convergence in distribution and the proof of the theorem can be found in Theorem 4.20 of Karatzas & Shreve (2005). The central limit theorem suggests the normal distribution in the Brownian motion. One might believe that normal distribution is the only proper distribution for the increments, however, this is incorrect. A more general continuous-time model is constructed in the following section.

### 2.6.2 Lévy process

As Brownian motion is a limit of the random walk, we may think of constructing a continuous-time process in the same way. Suppose that we wish to design a continuous time process $\{Y_t\}_{t\geq 0}$ such that the value of $Y_1$ at time $t = 1$ has a particular distribution $D$. The time interval is divided into $n$ subintervals of equal length. The corresponding increments $\{\xi_j^{(n)}\}_{j=1}^n$ are assumed to be independent from a common distribution $D^{(n)}$ such that the sum $Y_1 = \sum_{j=1}^n \xi_j^{(n)} \sim D$. When $n$ increases, the distribution of increments $F^{(n)}$ change but the distribution of the sum $D$ stays unchanged. This property of the distribution $D$ leads to the introduction of one of the most important classes as follows:

**Definition 5** (Infinitely divisible)**.** A distribution $D_X$ with *characteristic function* $\varphi_X : u \mapsto E[\exp(iuX)]$ is called *infinitely divisible* if for each $n \in \mathbb{N}$, there is a characteristic function $\varphi_X^{(n)}$, such that $\varphi_X = (\varphi_X^{(n)})^n$.

This is equivalent to saying that a random variable $X$ is infinitely divisible if for each $n \in \mathbb{N}$ there exists $\{X_i^{(n)}\}_{i=1}^n$ of i.i.d. random variables such that

$$X = X_1^{(n)} + X_2^{(n)} + \cdots + X_n^{(n)}.$$

Then if $D_X$ in the above construction is infinitely divisible, it can be used to introduce a continuous-time stochastic process by taking $n \to \infty$. The resulting process is called a *Lévy process* $\{X_t\}_{t\geq 0}$ defined by the following properties:

(i) $X_0 = 0$ a.s.;

(ii) $\{X_t\}$ has independent increments;

(iii) $\{X_t\}$ has stationary increments;

(iv) $X_t$ are continuous in probability, that is for any $\epsilon > 0$ and $t \geq 0$ it holds that $\lim_{s \to t} P(|X_s - X_t| > \epsilon) = 0$.

The last condition (iv) can be relaxed as $X_t$ are right continuous and have limits from the left with probability one. An immediate example of Lévy process is the Brownian motion. Lévy process was introduced by the French mathematician Paul Lévy in the 1930s, much of the theory was developed by himself, A. N. Khintchine, and K. Itô. Recently there has been a great revival of interest in these processes, due to new theoretical developments and also a wealth of novel applications, particularly to option pricing in mathematical finance. Lévy processes are receiving more interest than models based on Brownian motion because they are capable in describing the observations in financial markets in a more accurate way. The applications of Lévy processes in finance can be found in Cont & Tankov (2004) and Kijima (2003). The following theorem allows us to consider Lévy processes in a simple manner (see Raible (2000)).

**Theorem 6.** *If $\{X_t\}_{t \geq 0}$ is a Lévy process, then the marginal distribution of $X_t$ is determined by $X_1$.*

Infinitely divisibility and Lévy process are related by the following theorem (see Sato, 2014).

**Theorem 7.** *If $\{X_t\}_{t \geq 0}$ is a Lévy process, then, for any $t$, the distribution of $X_t$ is infinitely divisible. Conversely, for any infinitely divisible distribution $D$, there uniquely (in the sense of law) exists a Lévy process $\{X_t\}_{t \geq 0}$ such that $X_1$ has distribution $D$.*

Theorem 7 provides us the class of distributions for which Lévy processes exist and it is possible to find the approximation to the Lévy process by a random walk. Examples of infinitely divisible distributions include the normal, Poisson, gamma, inverse Gaussian, hyperbolic, variance gamma, scaled-t, and normal inverse Gaussian. The last four distributions are included in the class of generalized hyperbolic (GH) distributions, while inverse Gaussian distribution is in the class of generalized inverse Gaussian (GIG) distributions. Barndorff-Nielsen & Halgreen (1977) proved that both GH and GIG classes are infinitely divisible. GH distributions are often used to fit financial data since they have tails heavier than the normal distribution. In financial literature, scaled-t distributions were introduced by Praetz (1972), variance gamma distributions by Madan & Seneta (1990), hyperbolic distributions by Eberlein & Keller (1995), and NIG distributions by Barndorff-Nielsen & Shephard (2001).

We are specially interested in the NIG distribution because it is one of only two subclasses of GH distributions that are closed under convolution. The other subclass having this property is the variance gamma distribution. This property asserts that if $X_1$ is has NIG distribution then $X_{1/n}$ also has NIG distribution. In particular, if the observed process at a certain frequency scale follows a NIG distribution, then at lower frequency scales it follows an NIG distribution too. We extensively study financial data with NIG distribution in Chapter 3. test

# Chapter 3

# Explanatory Data Analysis with NIG

In this chapter, financial data is analyzed with a particular distribution stated in previous chapter, the normal inverse Gaussian (NIG) distribution. The properties of NIG distribution are presented and the distribution is fitted to real data. The estimation has been done with three approaches: the method of moments, the maximum likelihood and the h-likelihood. We also show how good the data are fitted with NIG distributions.

## 3.1 The NIG Distribution

The NIG distribution proposed by Barndorff-Nielsen (1997) is the distribution on the whole real line having density function

$$f(x;\alpha,\beta,\mu,\delta) = a(\alpha,\beta,\mu,\delta)q\left(\frac{x-\mu}{\delta}\right)^{-1} K_1\left(\delta\alpha q\left(\frac{x-\mu}{\delta}\right)\right)\exp(\beta x) \tag{3.1.1}$$

where

$$a(\alpha,\beta,\mu,\delta) = \pi^{-1}\alpha\exp\left(\delta\sqrt{(\alpha^2-\beta^2)}-\beta\mu\right) \quad \text{and} \quad q(x) = \sqrt{1+x^2}.$$

$K_\lambda$ is the modified Bessel function of the third kind with index $\lambda$ given by the integral expression

$$K_\lambda(x) = \frac{1}{2}\int_0^\infty y^{\lambda-1}\exp\left(-x(y+y^{-1})/2\right)dy. \tag{3.1.2}$$

The parameters $\alpha,\beta,\mu$ and $\delta$ satisfy $0 \le |\beta| \le \alpha$, $\mu \in \mathbb{R}$ and $\delta > 0$. The distribution is symmetric around $\mu$ provided $\beta = 0$. We shall denote this distribution by $NIG(\alpha,\beta,\mu,\delta)$. The moment generating function $M(t;\alpha,\beta,\mu,\delta)$ of $NIG(\alpha,\beta,\mu,\delta)$ is expressed as

$$M(t;\beta,\mu,\delta) = \exp\left(\delta\left(\sqrt{\alpha^2-\beta^2}-\sqrt{\alpha^2-(\beta+t)^2}\right)+\mu t\right). \tag{3.1.3}$$

Thus all moments of $X \sim NIG(\alpha, \beta, \mu, \delta)$ have simple explicit expression and, in particular, the mean and the variance are

$$E[X] = \mu + \delta\beta/(\alpha^2 - \beta^2)^{1/2} \quad \text{and} \quad \text{var}[X] = \delta\alpha^2/(\alpha^2 - \beta^2)^{3/2}. \tag{3.1.4}$$

It follows from (3.1.3) that the normal inverse Gaussian distributions are infinitely divisible and close under convolution if the parameters $\alpha$ and $\beta$ are fixed. If $X_1, X_2, .., X_m$ are independent normal inverse Gaussian random variables with common parameters $\alpha$ and $\beta$, that is $X_i \sim NIG(\alpha, \beta, \mu_i, \delta_i)$ for $1 \leq i \leq m$, then $X^{(m)} = X_1 + X_2 + \cdots + X_m$ is again distributed as normal inverse Gaussian $X^{(m)} \sim NIG(\alpha, \beta, \sum_{i=1}^m \mu_i, \sum_{i=1}^m \delta_i)$. Remark that the normal distribution $N(\mu, \sigma^2)$ is a limiting case for $\beta = 0$, $\alpha \to \infty$ and $\delta/\alpha = \sigma^2$.

In particular, we aim to employ NIG distribution to fit time series of excess returns $y_t = r_t - \bar{r}$ which have zero mean and are approximately symmetric as we discussed in Section 2.3. Therefore we take special attention to the $NIG(\alpha, \beta, \mu, \delta)$ with $\mu = 0$ and $\beta = 0$. Using the alternative parameterization $\phi = \delta/\alpha > 0$ and $\omega = \alpha\delta > 0$, the *zero-mean symmetric NIG* distribution, denoted by $Y \sim NIG(\phi, \omega)$ has the density function

$$f(y; \phi, \omega) = \frac{\omega \exp(\omega)}{\pi\sqrt{y^2 + \phi\omega}} K_1\left(\sqrt{\omega^2 + \frac{\omega}{\phi}y^2}\right). \tag{3.1.5}$$

And the moment generating function $M(t; \phi, \omega)$ of $NIG(\phi, \omega)$ is simply

$$M(t; \phi, \omega) = \exp\left(\omega - \sqrt{\frac{\omega}{\phi} - t^2}\right).$$

Consequently the variance and the kurtosis are

$$\text{var}[Y] = \phi \text{ and } \text{kurt}[Y] = 3/\omega.$$

Thus, given a sample drawn from $NIG(\phi, \omega)$, the parameters $\phi$ and $\omega$ can be readily estimated from its sample moments. Suppose that $\{y_1, y_2, .., y_n\}$ is a sample of independent observations drawn from the distribution $NIG(\phi, \omega)$. Using the sample variance $s^2$ and the sample kurtosis $k$, the parameters can be estimated by $\hat{\phi} = s^2$ and $\hat{\omega} = 3/k$. Furthermore, the parameters can be estimated by maximizing the log likelihood function

$$l(\phi, \omega) = \sum_{t=1}^{n} \left[\omega + \log(\omega) - \log(\pi) - \frac{1}{2}\log(y_t^2 + \phi\omega) + \log\left(K_1(\sqrt{\omega^2 + \frac{\omega}{\phi}y_t^2})\right)\right]. \tag{3.1.6}$$

The $NIG(\phi, \omega)$ is also close under convolution when the ratio $\omega/\phi$ is fixed. Given $X_1, X_2, ..., X_m$ are independent random variables distributed as $X_i \sim NIG(\phi_i, \omega_i)$ with a common ratio $\alpha^2 = \omega_i/\phi_i$ , then $X^{(m)} = X_1 + X_2 + ... + X_m$ is distributed as $X^{(m)} \sim NIG\left(\sum_{i=1}^m \phi_i, \sum_{i=1}^m \omega_i\right)$.

## 3.2 Data Descriptive Statistics

The data set we are exploring here is the exchange rates of three currencies stated in Section 1.2. From now on we will work with the time series of excess returns times 100

$$y_t = 100(r_t - \bar{r}). \tag{3.2.1}$$

We use the following abbreviations referring to the time series of exchange rates in a specific period: EUR to EUR/USD in the whole period, EUR-pre to EUR/USD in the Pre-Crisis period, EUR-c1 to EUR/USD in the Crisis 1 period, and EUR-post to EUR/USD in the Post-Crisis period. The abbreviations for JPY/USD and GBP/USD are given in the same manner. The number of returns over all period is $n = 1305$, the number of returns in Pre-Crisis, Crisis 1, Crisis 2 and Post-Crisis periods are 390, 261, 261 and 393 respectively.

Table 3.1 shows the summary statistics for all the time series. The average returns $\bar{r}$ over one day are very small thus they are often assumed to be zero. In our case we subtract them from the return series and consider only the series of excess returns $y_t$ in (3.2.1), other statistics are calculated from $y_t$ instead of $r_t$. It is noticeable that the variances during the Crisis 2 are higher than other periods for all currencies. The calmest period of all currencies is the Pre-Crisis period. The skewness statistics are not far from zero and do not provide much evidence of asymmetric distributions. Except for the series EUR-post, all others series have positive kurtosis as pointed out in Section 2.3. The standard error of a kurtosis estimate $k$ is $\sqrt{24/n}$ for a random sample from a normal distribution. In our series the standard errors range from 0.14 to 0.30 depending on $n$. The three series of overall periods EUR, JPY and GBP show significant positive kurtosis because they exceed zero by more than ten times standard errors. Only the EUR-post series has very small negative kurtosis that is insignificantly less than zero.

Table 3.2 shows the relative frequencies for time series of returns within or beyond the number of standard deviation from the mean. The reference distribution is the standard normal distribution. The relative frequencies around the mean in the range from $\bar{y} - 0.5s$ to $\bar{y} + 0.5s$ of all series but GBP-post are higher than that of normal distribution, corresponding to high peaks in empirical distributions. The frequencies of extreme values that are beyond three standard deviations are also more than that of normal distribution in most series. The relative frequencies of extreme values of some series are greater than 1% even beyond 6 standard deviations. The higher frequencies of extreme values corresponds to fat tails. In conclusion, the descriptive statistics for our time series are well agreeing with the stylized facts for financial returns that we have discussed in Section 2.3.

## 3.3 Fitting Financial Data with NIG Distribution

Now we are going to fit the data with NIG distribution. Since we $NIG(\phi, \omega)$ does not involve the skewness parameter $\beta$, so we test weather this parameterization is appropriate

Table 3.1: Descriptive statistics for time series of returns

| series | $n$ | mean ($\bar{r}$) | variance ($s^2$) | skewness ($w$) | kurtosis ($k$) |
|---|---|---|---|---|---|
| EUR | 1305 | 0.01 | 0.42 | 0.46 | 3.70 |
| EUR-pre | 390 | 0.03 | 0.19 | 0.37 | 1.14 |
| EUR-c1 | 261 | 0.06 | 0.27 | -0.26 | 0.30 |
| EUR-c2 | 261 | -0.04 | 1.00 | 0.72 | 2.02 |
| EUR-post | 393 | -0.01 | 0.37 | 0.08 | -0.09 |
| JPY | 1305 | 0.03 | 0.52 | 0.52 | 4.38 |
| JPY-pre | 390 | -0.01 | 0.24 | 0.45 | 1.27 |
| JPY-c1 | 261 | 0.06 | 0.54 | 0.86 | 4.80 |
| JPY-c2 | 261 | 0.04 | 0.97 | 0.57 | 2.77 |
| JPY-post | 393 | 0.04 | 0.48 | -0.11 | 2.01 |
| GBP | 1305 | -0.01 | 0.48 | 0.03 | 5.18 |
| GBP-pre | 390 | 0.04 | 0.21 | 0.25 | 0.79 |
| GBP-c1 | 261 | 0.00 | 0.26 | -0.46 | 0.39 |
| GBP-c2 | 261 | -0.07 | 1.25 | 0.20 | 2.35 |
| GBP-post | 393 | -0.01 | 0.40 | 0.01 | 0.19 |

The descriptive statistics for returns ($\bar{r}$) and excess returns ($s^2$, $w$, $k$) show that the distributions of (excess) returns are approximately symmetric with positive kurtosis.

for financial data. Then we continue analyzing the data with NIG distribution with different methods of estimation and test for the goodness-of-fit.

### 3.3.1 Skewness

The skewness estimates in Table 3.1 shows some non zero skewness indicating that the data may be drawn from asymmetric distribution. If this hypothesis is true then we shall not assume skewness parameter $\beta$ in the NIG distribution to be zero. Here we test if the skewness parameter $\beta$ in the NIG distribution significantly improve the goodness of fit in our data by the likelihood-ratio test. Given the null hypothesis that the excess returns $y_t$ follow a zero mean symmetric NIG distribution, $y_t \sim N(\phi, \omega)$. The alternative hypothesis is that the returns follow the skewed NIG distribution with the density function

$$f_{sk}(y;\phi,\omega,\beta) = \frac{\omega \exp\left(\sqrt{\omega^2 - \phi\omega\beta^2}\right)}{\pi\sqrt{y^2+\phi\omega}} K_1\left(\sqrt{\omega^2 + \frac{\omega}{\phi}y^2}\right)\exp(\beta y).$$

The corresponding log-likelihood function for skewed NIG distribution is

$$l_{sk}(\phi,\omega,\beta) = \sum_{t=1}^{n}\left[\sqrt{\omega^2-\phi\omega\beta^2} + \log(\omega) - \log(\pi) - \frac{1}{2}\log(y_t^2+\phi\omega) + \log\left(K_1(\sqrt{\omega^2+\frac{\omega}{\phi}y_t})\right) + \beta y_t\right].$$

Table 3.2: Frequency distributions

| | Percentage of returns within/beyond the number of standard deviations from the mean | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | with in | | beyond | | | | | | |
| | 0.25 | 0.5 | 1 | 1.5 | 2 | 3 | 4 | 5 | 6 |
| Normal | 19.74% | 38.29% | 31.73% | 13.36% | 4.55% | 0.27% | 0.01% | | |
| EUR | 26.13% | 48.12% | 25.75% | 11.34% | 4.44% | 1.30% | 0.46% | 0.23% | 0.08% |
| EUR-pre | 22.05% | 43.33% | 28.97% | 12.05% | 5.38% | 0.77% | 0.26% | | |
| EUR-c1 | 23.37% | 44.44% | 28.74% | 14.56% | 5.75% | | | | |
| EUR-c2 | 24.52% | 48.28% | 25.29% | 12.64% | 5.75% | 1.15% | | | |
| EUR-post | 22.14% | 40.20% | 32.06% | 15.27% | 4.07% | | | | |
| JPY | 25.98% | 45.82% | 24.90% | 10.04% | 4.60% | 1.23% | 0.38% | 0.15% | 0.15% |
| JPY-pre | 24.87% | 45.38% | 26.92% | 13.33% | 4.87% | 0.77% | 0.26% | | |
| JPY-c1 | 24.52% | 42.91% | 25.67% | 9.20% | 3.45% | 0.77% | 0.38% | 0.38% | 0.38% |
| JPY-c2 | 20.69% | 41.00% | 24.52% | 9.58% | 6.13% | 0.77% | 0.38% | 0.38% | |
| JPY-post | 24.43% | 44.02% | 27.48% | 11.96% | 4.83% | 1.53% | 0.25% | | |
| GBP | 25.52% | 46.44% | 23.75% | 10.57% | 4.44% | 1.38% | 0.46% | 0.23% | 0.08% |
| GBP-pre | 24.36% | 44.10% | 27.69% | 14.36% | 6.15% | 0.26% | 0.26% | | |
| GBP-c1 | 22.22% | 41.38% | 29.12% | 14.18% | 5.36% | 0.38% | | | |
| GBP-c2 | 22.61% | 44.44% | 27.59% | 11.11% | 4.21% | 1.53% | 0.38% | | |
| GBP-post | 18.83% | 34.10% | 28.50% | 11.96% | 4.58% | 0.51% | | | |

The relative frequencies of extreme values of returns for ten returns series show that most of the returns series have greater relative frequencies beyond three standard deviations than that of the standard normal distribution. Evidently, the distribution of returns exhibits fat tails.

The test statistic $D$ is twice the difference between the two log-likelihoods

$$D = 2\left(l_{sk}(\phi, \omega, \beta) - l(\phi, \omega)\right)$$

Then the test statistic $D$ is approximately a chi-square distribution with one degree of freedom. The $95^{th}$ percentile of a chi-square distribution with one degree of freedom is 3.84 that is far greater than the statistics calculated from our data shown in Table 3.3. There is no evidences that the time series in our data set follows a skewed NIG distribution and hence we rationally exclude the skew parameter $\beta$ from our model.

### 3.3.2 Parameter estimation

The parameters of a zero-mean symmetric NIG distribution can be estimated simply either by the method of moments (MoM) or the maximum likelihood estimation (MLE). The method of moments is very convenient and the estimated parameters can be used as initial values for MLE. From (3.1.3) the variance and the kurtosis of $Y \sim NIG(\phi, \omega)$ are $\text{var}(Y) = \phi$ and $\text{kurt}(Y) = 3/\omega$. Thus for a time series of excess returns $\{y_1, y_2, ..., y_n\}$ supposed to follow a symmetric NIG distribution $y_t \sim NIG(\phi, \omega)$, the parameters can be estimated by $\hat{\phi} = s^2$ and $\hat{\omega} = 3/k$, where $s^2$ and $k$ are the sample variance and sample kurtosis. Table 3.4 shows the estimated parameters from MoM and MLE. The values of $\hat{\phi}$ from both method are almost identical in several series, but the estimates of $\omega$ are

Table 3.3: Likelihood ratio test

| Series | log-likelihood | | $D$ |
|---|---|---|---|
| | symmetric NIG | skewed NIG | |
| EUR | -1215.66 | -1215.66 | 2.23E-07 |
| EUR-pre | -220.22 | -220.22 | 5.19E-07 |
| EUR-c1 | -199.25 | -199.25 | 3.68E-06 |
| EUR-c2 | -358.29 | -358.29 | 2.81E-06 |
| EUR-post | -359.33 | -359.33 | -1.98E-05 |
| JPY | -1352.92 | -1352.92 | 1.39E-06 |
| JPY-pre | -267.45 | -267.45 | -1.65E-05 |
| JPY-c1 | -279.59 | -279.59 | 2.23E-07 |
| JPY-c2 | -356.43 | -356.43 | -4.20E-07 |
| JPY-post | -400.07 | -400.07 | -5.08E-06 |
| GBP | -1298.41 | -1298.41 | 1.66E-06 |
| GBP-pre | -241.37 | -241.37 | 1.14E-06 |
| GBP-c1 | -192.94 | -192.94 | 2.65E-07 |
| GBP-c2 | -390.62 | -390.62 | -9.49E-06 |
| GBP-post | -375.61 | -375.61 | 1.09E-05 |

The log-likelihood-ratio statistics $D$ are far smaller than the critical value 3.84. Hence, the null hypotheses stating that the symmetric NIG distribution and the skewed NIG distribution are similarly fitted to the data are not rejected at 95% confidence. Hence we rationally exclude the skewness parameter $\beta$ from our models.

slightly different. Remark that $\hat{\omega}$ from the MoM estimation is satisfactory only when $k > 0$, because $\omega$ must be positive. The EUR-post has negative kurtosis, hence the MoM estimate is not satisfactory. We use an initial value slightly greater than zero, $\hat{\omega}_0 = 0.1$, to obtain the maximum-likelihood estimate that is far greater than other estimates. The more the parameter $\omega$, the less the kurtosis for the distribution. As a consequence, the distribution is approximately normal with variance $\phi$. This result can be seen in the density plot in the next subsection. Hereafter, we take the MoM as the initial guess for true parameter and practically use MLE in application.

### 3.3.3 Goodness of fit

Firstly, we asses the goodness of fit by graphical methods. In Figure 3.3.1, the density plots for most of the series are better with NIG distributions than normal distributions, either with MLE parameters or MoM parameters. The fitted NIG distributions have higher peaks than that of fitted normal distributions agree adequately with the data. In the case of EUR-post that the kurtosis estimate is negative, the MLE fitted NIG distribution is indistinguishable from the fitted normal distribution while the MoM fitted NIG is not satisfactory. Other cases, the MLE fitted NIG distributions and MoM fitted distributions are a bit different, however, clearly that they fit better than normal distributions.

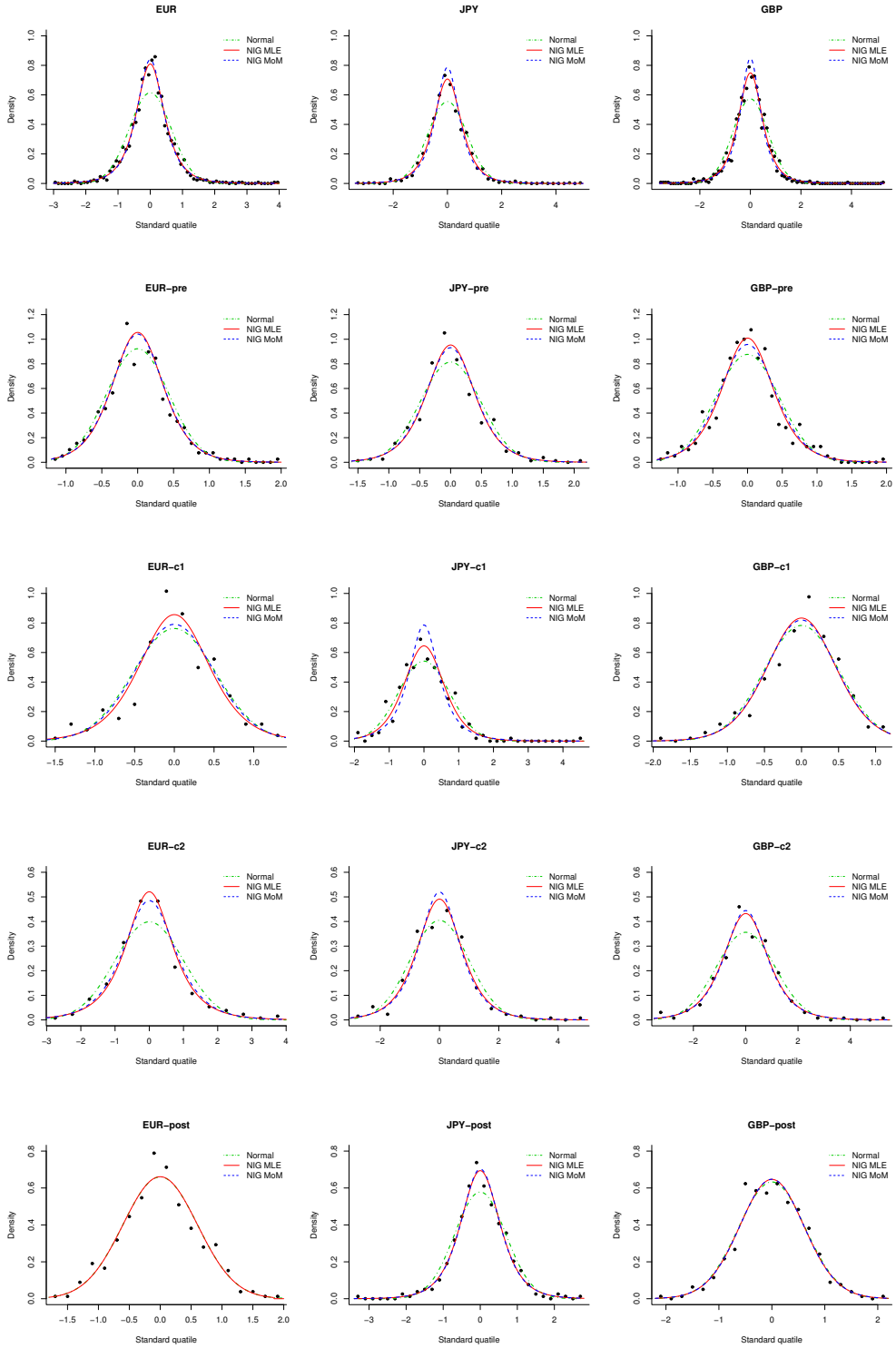
Table 3.4: Estimated parameters from MoM and MLE

| series | $\hat{\phi}$ | | $\hat{\omega}$ | |
|---|---|---|---|---|
| | MoM | MLE | MoM | MLE |
| EUR | 0.421 | 0.418 | 0.811 | 0.977 |
| EUR-pre | 0.187 | 0.187 | 2.637 | 2.346 |
| EUR-c1 | 0.274 | 0.276 | 9.854 | 2.655 |
| EUR-c2 | 1.001 | 1.008 | 1.488 | 0.980 |
| EUR-post | 0.365 | 0.364 | -32.160 | 700.720 |
| JPY | 0.517 | 0.505 | 0.685 | 1.204 |
| JPY-pre | 0.240 | 0.240 | 2.354 | 1.929 |
| JPY-c1 | 0.543 | 0.521 | 0.625 | 1.992 |
| JPY-c2 | 0.972 | 0.950 | 1.083 | 1.633 |
| JPY-post | 0.476 | 0.472 | 1.490 | 1.669 |
| GBP | 0.485 | 0.471 | 0.579 | 1.067 |
| GBP-pre | 0.207 | 0.209 | 3.795 | 2.135 |
| GBP-c1 | 0.259 | 0.259 | 7.707 | 5.695 |
| GBP-c2 | 1.254 | 1.235 | 1.278 | 1.584 |
| GBP-post | 0.397 | 0.396 | 15.956 | 17.724 |

The estimated parameters from MoM and MLE are very similar in most cases. The values of $\hat{\phi}$ from both method are almost identical in several series, but the estimates of $\omega$ are slightly different. Remark that $\hat{\omega}$ from the MoM estimation is satisfactory only when $k > 0$, because $\omega$ must be positive. Practically , we take the MoM as the initial guess for true parameter and use MLE in application.

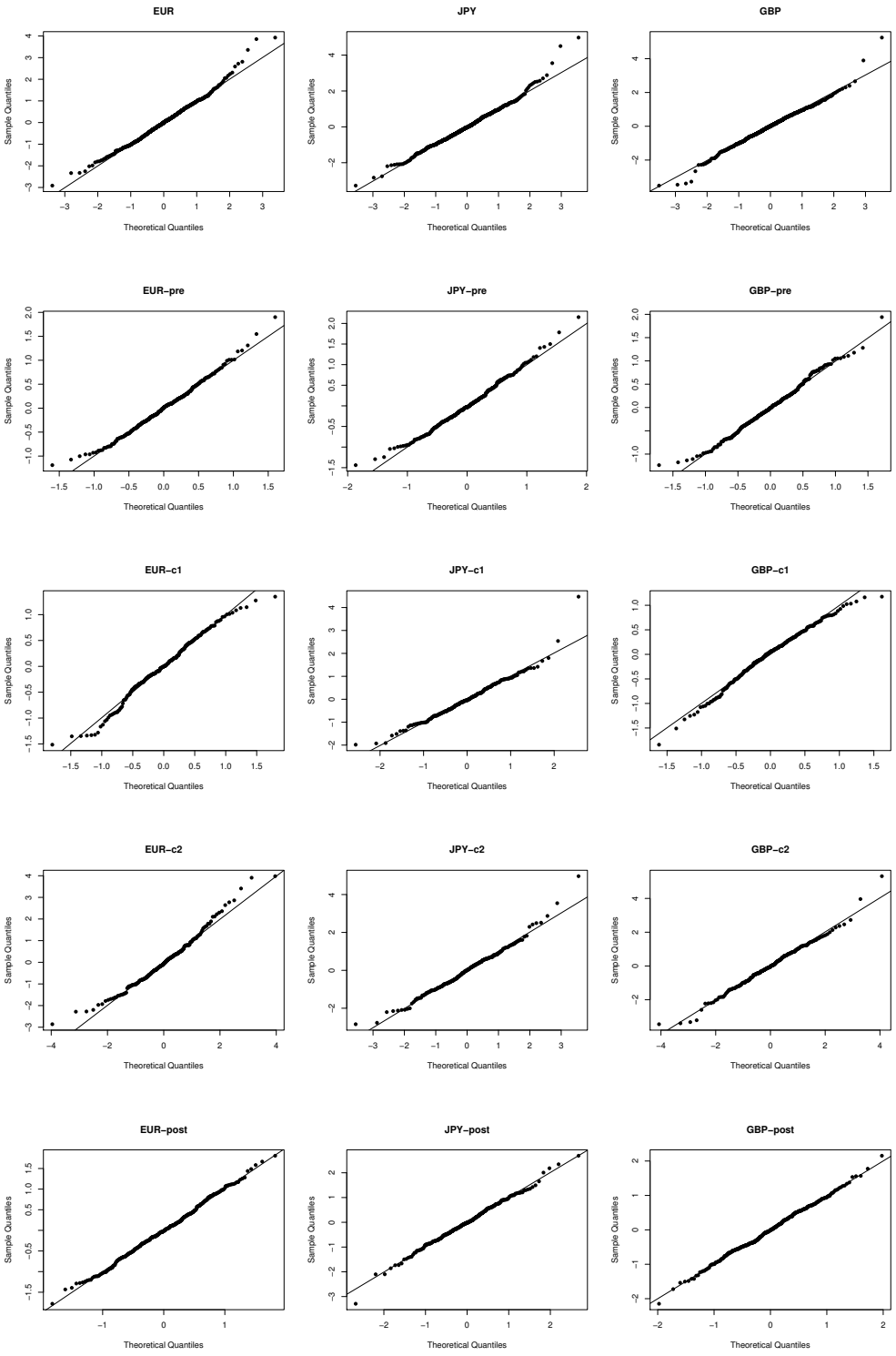
Furthermore, Figure 3.3.2 show the quantile-quantile (Q-Q) plots of MLE fitted NIG quantiles against sample quantiles. Most of the points in each plot lie nicely on the line except for some outliers. The Q-Q plots for the post-crisis period are nearly perfect fit to the lines. These graphics show that the NIG distribution are considerably accurate in describing the distribution of financial returns.

Figure 3.3.1: The density plots for returns superimposed on the fitted densities of normal and NIG.




The density plots for most of the returns series are better with NIG distributions than normal distributions, either with MLE parameters or MoM parameters.

Figure 3.3.2: The Q-Q plots of fitted NIG quantiles against sample quantiles for returns




Most of the points in each plot lie nicely on the line except for some outliers. These graphics show that the NIG distributions are considerably accurate in describing the distributions of financial returns.

For quantitative analysis we make use of the Pearson's $\chi^2$ test for goodness of fit. The null hypothesis is that the observations are consistent with the tested distribution. The test statistic $X^2$ is calculated from categorized data that the partitioning can influence the value of the statistic especially when some categories contain small numbers of observations. Therefore we categorize the data by equal probabilities to have the same expected value in each class. Then the test statistic is calculated by

$$X^2 = \sum_{i=1}^{k} (O_i - E_i)^2 / E_i$$

where $k$ is the number of classes, $O_i$ are the observed frequencies and $E_i$ are the expected frequencies. The expected frequency of NIG distribution is calculated by numerical integration since there is no explicit form of the distribution function. An example of $X^2$ calculation is given in Table 3.5a, we fix the number of classes as $k = 20$. Since the parameters are estimated by maximum likelihood, the asymptotic distribution of $X^2$ can be bounded between chi-square with $k - 1$ and chi-square with $k - p - 1$ degrees of freedom, where $p$ is the number of estimated parameters. The corresponding p-value are reported in Table 3.5b, the true asymptotic p-value lies between p-value1 and p-value2. The least p-value, that is of the JPY-pre, is 0.016 still greater than 0.01, thus it is not rejected at 99% confidence. The other cases are clearly not rejected at 95% confidence. In conclusion, from both graphical and quantitative methods, the NIG distribution is very accurate in fitting financial returns.

Table 3.5: Goodness-of-fit test

(a) $X^2$ statistic calculation

| Class | $O$ | $E$ | |
|---|---|---|---|
| $(-\infty$,-1.03] | 69 | 65.25 | 0.216 |
| (-1.03,-0.735] | 73 | 65.25 | 0.920 |
| (-0.735,-0.565] | 64 | 65.25 | 0.024 |
| (-0.565,-0.444] | 59 | 65.25 | 0.599 |
| (-0.444,-0.347] | 56 | 65.25 | 1.311 |
| (-0.347,-0.265] | 71 | 65.25 | 0.507 |
| (-0.265,-0.193] | 61 | 65.25 | 0.277 |
| (-0.193,-0.126] | 74 | 65.25 | 1.173 |
| (-0.126,-0.062] | 66 | 65.25 | 0.009 |
| (-0.062,0.00] | 55 | 65.25 | 1.610 |
| (0.00,0.062] | 65 | 65.25 | 0.001 |
| (0.062,0.126] | 82 | 65.25 | 4.300 |
| (0.126,0.193] | 65 | 65.25 | 0.001 |
| (0.193,0.265] | 66 | 65.25 | 0.009 |
| (0.265,0.347] | 62 | 65.25 | 0.162 |
| (0.347,0.444] | 57 | 65.25 | 1.043 |
| (0.444,0.565] | 60 | 65.25 | 0.422 |
| (0.565,0.735] | 69 | 65.25 | 0.216 |
| (0.735,1.03] | 67 | 65.25 | 0.047 |
| (1.03, $\infty$) | 64 | 65.25 | 0.024 |
| | | $X^2 =$ | 12.870 |

(b) Pearson's $\chi^2$ test for NIG distribution

| | $X^2$ | p-value1 | p-value2 |
|---|---|---|---|
| EUR | 12.870 | 0.745 | 0.845 |
| EUR-pre | 15.436 | 0.564 | 0.695 |
| EUR-c1 | 18.770 | 0.342 | 0.472 |
| EUR-c2 | 19.690 | 0.290 | 0.413 |
| EUR-post | 16.008 | 0.523 | 0.657 |
| JPY | 24.640 | 0.103 | 0.173 |
| JPY-pre | 31.744 | 0.016 | 0.033 |
| JPY-c1 | 14.785 | 0.611 | 0.736 |
| JPY-c2 | 16.471 | 0.491 | 0.626 |
| JPY-post | 12.547 | 0.766 | 0.861 |
| GBP | 16.119 | 0.515 | 0.649 |
| GBP-pre | 24.462 | 0.107 | 0.179 |
| GBP-c1 | 18.157 | 0.379 | 0.512 |
| GBP-c2 | 12.487 | 0.770 | 0.864 |
| GBP-post | 17.229 | 0.439 | 0.574 |

Table 3.5a show how the statistic $X^2$ is computed. Each class has equal expected frequency, thus the error related to partitioning has been reduced. Table 3.5b shows the test statistics $X^2$ and the estimated p-values. The true p-value lie between p-value1 and p-value2. Clearly, the null hypotheses are not rejected at 99% confidence. The observations are properly fitted to NIG distributions.

## 3.4 Summary

In this chapter, we have analyzed the financial data with NIG distribution. The empirical distributions of the data are not normal, they have high peaks and fat tails. The symmetric NIG distribution has been proved to be equally fitted to the data compared to the skewed NIG distribution. Hence the symmetric NIG distribution is preferred because of less parameters. The (zero-mean symmetric) NIG distribution can be estimated

analogously with method of moments and maximum likelihood estimation. Practically, we use the MoM as initially values for MLE. The goodness-of-fits have been tested and the exploratory data are adequately fitted to NIG distributions. In conclusion, the NIG distribution is very appropriate for describing financial data.

# Chapter 4

# Volatility Forecasting

In this chapter, we develop volatility forecasting models that the volatility is assumed stochastic. The NIG-SV model that the returns follow NIG distributions is specially interested as we have shown in the previous chapter that the marginal distributions of returns are well fitted to NIG distributions. We also discuss on some practical issues in volatility forecasting that practitioners usually encounter. The volatility forecasting strategy including the evaluation measures are also provided. An alternative approach on estimation for stochastic volatility models, especially for the NIG-SV model, is investigated. The latent information estimates are obtained as by-products of the estimation. Consequently, we develop forecasting models based on the latent information that perform better than standard models in some occasions.

## 4.1 Practical Issues in Volatility Forecasting

Volatility forecasting is one of the most challenging fields in financial econometrics. Hence numerous papers studying performance of various models have been published over the last two decades. The investigation in volatility forecasting consists of vast aspects both theoretically and practically including volatility definitions, volatility measurement, volatility models, model's parameter estimation, objectives of volatility forecasting, forecast evaluation and volatility proxies. Poon & Granger (2003) gives a comprehensive review of volatility forecasting covering 93 papers from 1976 to 2002. They also extensively discuss several practical issues in volatility forecasting in Poon & Granger (2005) and . Recently Brownlees et al. (2012) give an informative guide to practically forecast volatility with GARCH models. This section is mainly based on the work of Poon & Granger (2003, 2005) and Brownlees et al. (2012).

### 4.1.1 Volatility proxy

Volatility is unobservable even *ex post.* It is therefore more complicated when we make comparison of forecasting methods. The unknown true volatility is regularly replaced by related observable quantity called *volatility proxy* to be used as a reference when making

comparison. The true volatility is usually estimated by sample standard deviation that is called realized volatility in (2.5.11). This is a result of proxying a daily volatility by a squared daily return. Then the average volatility over $m$ days is proxied by the realized volatility calculated from $m$ observations. Given the excess return $y_t = \sigma_t \epsilon_t$ stated in (2.5.2), Lopez (2001) shows that $y_t^2$ is an unbiased estimator of $\sigma_t^2$. However, Poon & Granger (2003) argues that squared return is very imprecise estimator of volatility. The use of $y_t^2$ as volatility proxy will lead to low coefficient of determination $R^2$ and undermine the inference regarding forecast accuracy.

Other standard volatility proxies are daily range $R_t = \max\{\log P_\tau\} - \min\{\log P_\tau\}, t-1 \leq \tau \leq t$ and daily realized volatility $RV_t^{(k)} = \sum_{j=1}^{k} y_{t,j}^2$, where $y_{t,j}$ are intraday returns. Under the assumption that the log price follows a Brownian motion $y_t = \sigma_t dW_t$, where $\sigma_\tau = \sigma_t$ for $t-1 \leq \tau \leq t$. Parkinson (1980) gives an accurate volatility estimator using daily range by $\hat{\sigma}_t^2 = R_t^2/(4\log(2))$. The mean squared error (MSE) of Parkinson's estimator is approximately one-fifth of the MSE of the squared return. Nevertheless, the range-based volatility estimator depends critically on the assumed price generating process, which is a potential drawback of the range as a volatility proxy. Daily realized volatility is unbiased estimator and has gained much attention recently, see Andersen et al. (2001, 2003) and Barndorff-Nielsen & Shephard (2002, 2004). However, for most assets, high-frequency data are not publicly accessible and it is not easy to obtain reliable high frequency data.

Patton (2011) gives a class of loss functions that is attractively robust in the sense that they asymptotically generate the same ranking of models regardless of the proxy being used as long as the proxy is unbiased and minimal regularity conditions are met. It ensures that model rankings achieved with proxies like squared returns or daily realized volatility correspond to the ranking that would be achieved if forecasts were compared against the true volatility. Hence the squared return is a reasonable and affordable choice of volatility proxy for point forecast evaluation. When a long horizon volatility is forecasted, a point forecast becomes very noisy as the forecast horizon lengthens. Instead, the cumulative volatility over the forecast horizon is more accurate because of error cancellation (see Poon & Granger, 2003). Suppose that the forecast horizon is $k$, the cumulative volatility over the forecast horizon $\sigma_{t+k,t}^2 = \sum_{i=1}^{k} \sigma_{t+i}^2$ is then proxied by the sum of squared returns over the forecast horizon $\hat{\sigma}_{t+k,t}^2 = \sum_{i=1}^{k} y_{t+i}^2$.

### 4.1.2 Forecast evaluation

The performance of forecasting volatility is considered by the choices of models and strategies. It is worth mentioning that we focus only on out-of-sample implementation because it is closer to real applications. A good forecasting model should be one that can withstand the robustness of an out-of-sample test. The forecast performance is evaluated by the average loss achieved by the model that is calculated by a loss function with a proper volatility proxy. The less average loss, the more accuracy. Several loss functions have been employed in the literature on volatility forecast evolution, see Patton (2011). Under our choice of volatility proxy, the squared return, Patton (2011) suggests the MSE

and *quasi likelihood* (QL) loss functions that are robust in the ranking preservation as discussed in the previous subsection. The two loss functions are defined by

$$\begin{aligned} \mathrm{QL}(\hat{\sigma}^2_{t+k}, v_{t+k|t}) &= \frac{\hat{\sigma}^2_{t+k}}{f_{t+k|t}} - \log \frac{\hat{\sigma}^2_{t+k}}{f_{t+k|t}} - 1 \\ \mathrm{MSE}(\hat{\sigma}^2_{t+k}, v_{t+k|t}) &= (\hat{\sigma}^2_{t+k} - f_{t+k|t})^2 \end{aligned}$$

where $\hat{\sigma}^2_{t+t}$ is an unbiased *ex post* proxy of volatility (such as squared return or daily realized volatility ) and $f_{t+k|t}$ is a volatility forecast based on information up to time $t$ and the forecast horizon $k > 0$.

The MSE loss is a usual loss function in the literature, however Brownlees et al. (2012) show that the QL loss is more preferable than the MSE loss for forecast comparison because of two reasons. First, the loss series is iid under the null hypothesis that the forecasting model is correctly specified while MSE contains high levels of serial dependence even under the null. Second, the bias of QL is independent of the volatility level, while MSE has a bias that is proportional to the square of the true volatility. We employ both QL and MSE loss functions in our investigation. Furthermore, the predictive ability of two forecasts if they are equally accurate by the test of Diebold & Mariano (1995). Suppose $f^{(1)}_t$ and $f^{(2)}_t$ are two forecasts of $\sigma^2_t$, we define the forecast loss differential between the two forecast by $e_t = g(\hat{\sigma}^2_t, f^{(1)}_t) - g(\hat{\sigma}^2_t, f^{(2)}_t)$ where $g$ is the loss function. We say that the two forecasts have equal accuracy if and only if the loss differential has zero expectation for all $t$. Assume the loss series $\{e_1, e_2, , e_T\}$, the Diebold-Mariano statistic is

$$DM = \frac{\bar{e}}{\sqrt{2\pi \hat{f}_e(0)/T}}$$

where $\bar{e} = \left(\sum_{\tau=1}^{T} e_\tau\right)/T$ and $\hat{f}_e(0)$ is a consistent estimate of the spectral density of the loss differential at frequency 0. In standard practice $\hat{f}_e(0)$ is given by

$$\hat{f}_e(0) = \frac{1}{2\pi} \sum_{\tau=-(T-1)}^{T-1} I\left(\frac{\tau}{k-1}\right) \hat{\gamma}_e(\tau) \text{ where } \hat{\gamma}_e(\tau) = \frac{1}{T} \sum_{t=|\tau|+1}^{T} (e_t - \bar{e})(e_{t-|\tau|} - \bar{e})$$

and

$$I\left(\frac{\tau}{k-1}\right) = \begin{cases} 1 & \text{for } \left|\frac{\tau}{k-1}\right| \leq 1 \\ 0 & \text{otherwise} \end{cases}.$$

The test statistic $DM$ is asymptotically $N(0,1)$ distributed under the null hypothesis of equal forecast accuracy.

### 4.1.3 Forecasting models

In this subsection, we describe some commonly used models that deploy historical information to formulate volatility forecasts. Base on the information available up to time

$t$, $\mathcal{F}_t$, the forecast for the future volatility $\sigma^2_{t+k}$ is often obtained from the conditional expectation $E[\sigma^2_{t+k}|\mathcal{F}_t]$, alternatively denoted by either $E_t[\sigma^2_{t+k}]$ or $\sigma^2_{t+k|t}$. We denote $f^{\mathrm{M}}_{t+k|t}$ the volatility forecast for $\sigma^2_{t+k}$ formulated at time $t$ with forecasting model M.

The realized volatility in (2.5.11) and the EWMA in (2.5.13) already make one-step-ahead forecasts by the definitions. Assuming that the volatilities follow a random walk, then $\hat{\sigma}^2_t$ is the optimal forecast for $\sigma^2_{t+k}$. If the volatility is estimated at time $t$, for example with realized volatility $\hat{\sigma}^2_t = \frac{1}{m}\sum_{i=0}^{m} y^2_{t-i}$, the forecast for the volatility at time $t+k$ formulated at time $t$ is

$$f^{\mathrm{RW}}_{t+k|t} = \hat{\sigma}^2_t. \tag{4.1.1}$$

A more sophisticated model estimates the current volatility by EWMA model and the forecast obtained from the conditional expectation is also of the form (4.1.1). One of the most popular models is GARCH(1,1), where the $k$-step-ahead forecast at time $t$ is given by

$$f^{\mathrm{GARCH}}_{t+k|t} = \sigma^2 + (\alpha+\beta)^k(\hat{\sigma}^2_t - \sigma^2) \tag{4.1.2}$$

where $\sigma^2 = \frac{\omega}{1-\alpha-\beta}$ is the unconditional variance. It is easily seen that GARCH(1,1) forecasts converge to the unconditional variance as $k \to \infty$. These three forecasting models are used as general benchmarks in the literature. Any newly developed forecasting model should be better or as good as these models in forecasting ability.

### 4.1.4 The role of the log transformation

In the financial literature, some volatility models involve log transformation, for examples, the EGARCH model by Nelson (1991) and the standard SV model by Taylor (1986). When the log transformation involves in the model, the estimates for either volatility or log volatility may be obtained from the estimation procedure. It is worth considering weather we should make forecast based on the original series $\{\sigma_t\}$ or the log-transformed series $\{\log(\sigma_t)\}$. In time-series analysis, the log transformation is considered to stabilize the variance of time series. Hence the time series that is modeled and forecasted under the log transformation is expected to be more accurate. When the series of log volatility $\{\log(\sigma_t)\}$ is modeled and forecasted by a time series model, for example ARMA model, one may directly apply the exponential function to obtain the forecast for the original series $\{\sigma_t\}$. However, instantaneous reverse transformation of optimal forecast for transformed variable does not result in optimal forecast for original variable in general. In other words, if $v_{t+k|t}$ is an optimal forecast for $\log(\sigma_{t+k})$ then $\exp(v_{t+k|t})$ is not an optimal forecast for $\sigma_{t+k}$. Granger & Newbold (1976) propose the optimal forecast for $\sigma_{t+k}$ provided the log-transformed series $\{\log(\sigma_t)\}$ is Gaussian and stationary as

$$f_{t+k|t} = \exp\left\{v_{t+k|t} + \frac{1}{2}\mathrm{var}(v_{t+k|t})\right\}$$

where $v_{t+k|t}$ is the optimal forecast for $\log(\sigma_{t+k})$. This method immediately apply when log-transformed series is modeled by stationary ARMA processes. Further investigation

in the role of the log transformation in forecasting economic variables is also found in Lütkepohl & Xu (2010).

### 4.1.5 Recent work and forecasting strategy

In practice, there are several issues related to volatility forecasting as we have discussed in this section. It is worth to develop from existing discoveries. The recent work from Brownlees et al. (2012) has provided a pragmatic and fruitful guide to volatility forecasting through the period of financial crisis. Brownlees et al. (2012) have tested the forecasting performances of four models in ARCH class including GARCH(1,1), TARCH, EGARCH, NGARCH and APARCH using broad time series of exchange rates, domestic equity indices and international equity indices. The study also takes into account the strategies being used for estimation and different forecast horizons. There findings are summarized as the followings:

(i) models perform best using the longest available data series,

(ii) updating parameter estimates at least weekly counteracts the adverse effects of parameter drift,

(iii) no evidence that the Student t likelihood improves forecasting ability,

(iv) soaring volatility during the crisis of 2008 was well described by short-horizon forecasts,

(v) crisis forecasts deteriorated at long horizons (one-month horizon),

(vi) at the one-month horizon, the difference between asymmetric and symmetric GARCH becomes insignificant.

Based on these findings, we further investigate in volatility forecasting with stochastic volatility models, that are closely related to continuous time models and derivative pricing theory. The objective is to obtain forecasting models that accurately forecast volatility through the crisis of 2008. The ARCH-type models have been proved that they performs well in short-horizon forecasts. Therefore, we forecast only on 22-step-ahead horizon to be compatible with the one-month horizon that deteriorating forecasts were reported in Brownlees et al. (2012). To focus on the forecast performances on different models, we keep the forecasting strategy fix based on the guide of Brownlees et al. (2012).

The parameter estimates are updated every five days using longest available data up to the estimation update. Since point forecast is usually noisy, instead we make cumulative volatility forecast over the horizon $\sigma^2_{t+k,t} = \sum_{i=1}^{k} \sigma^2_{t+i}$ that is proxied by the sum of squared returns over the forecast horizon

$$\hat{\sigma}^2_{t+k,t} = \sum_{i=1}^{k} y^2_{t+i}. \tag{4.1.3}$$

The sum of squared returns $\hat{\sigma}^2_{t+k,t}$ is an unbiased proxy of the cumulative volatility $\sigma^2_{t+k,t}$. Therefore it has been employed as the target variable for the cumulative forecast

$$F_{t+k|t} = \sum_{i=1}^{k} f_{t+i|t} \tag{4.1.4}$$

formulated at time $t$. Forecast the cumulative volatility over the forecast horizon is more accurate than point forecast because of error cancellation, moreover the cumulative volatility forecast is the required input in the pricing model relying on a riskless hedge (Poon, 2005).

The data in the Pre-Crisis period is reserved as the initial sample window for estimation and forecasting. Every five days the sample window is expanded and the related parameters are reestimated with the extended sample. Then the estimation and forecasting run throughout the remaining period using the expanding windows. When the forecasts from different models are made, we evaluate the forecast performances by comparing the average QL losses. The MSE losses are also calculated but we mainly consider the QL losses. The long-run average losses are taken from the whole forecasting period, the average losses from a single period are also considered. When a two forecasting models are compared, the Diebold-Mariano test is applied to test whether they have equal accuracy.

## 4.2 Parametric Lévy Processes

Lévy process is a continuous-time stochastic process that can be constructed from a distribution with infinitely divisible property. Lévy processes have been introduced for analyzing financial returns because the associated distributions can be modeled to capture the heavy tails and other relevant features of returns better than normal distribution. In particular, a class of generalized hyperbolic distributions is very often able to fit the distributions of financial data. This have been established in considerable investigations, such as the variance gamma model by Madan & Seneta (1990), the hyperbolic model by Eberlein & Keller (1995), the NIG model by Barndorff-Nielsen (1997) and the generalized hyperbolic (GH) Lévy processes by Barndorff-Nielsen & Shephard (2001).

### 4.2.1 The generalized hyperbolic Lévy processes

The GH class of distributions was introduced by Barndorff-Nielsen & Shephard (2001) consisting of the hyperbolic distributions, the NIG distributions, the scaled-t distributions and the variance-gamma distributions. The GH distribution is characterized by five parameters, written as $X \sim GH(\lambda, \alpha, \beta, \delta, \mu)$. The probability density function of a GH distribution is given by

$$f(x; \lambda, \alpha, \beta, \mu, \delta) = \frac{(\gamma/\delta)^\lambda}{\sqrt{2\pi K_\lambda(\delta\gamma)}} \cdot \frac{K_{\lambda-\frac{1}{2}}\left(\alpha\sqrt{\delta^2+(x-\mu)^2}\right)}{\left(\sqrt{\delta^2+(x-\mu)^2}/\alpha\right)^{\frac{1}{2}-\lambda}} \cdot e^{\beta(x-\mu)} \tag{4.2.1}$$

where $\gamma^2 = \alpha^2 - \beta^2$ and $K_\lambda$ is the modified Bessel function of the third kind with index $\lambda$ given in (3.1.2). The distribution is symmetric if $\beta = 0$, $\delta$ is the scale parameter and $\mu$ is the location parameter. The subclass of hyperbolic distributions is obtained by letting $\lambda = 1$, the subclass of NIG distributions is obtained by letting $\lambda = -1/2$, the subclass of variance-gamma distributions is obtained when $\delta = 0$ and finally the subclass of asymmetric-scaled t distributions is obtained when $\alpha = |\beta|$. See Bibby & Sørensen (2003) for details and properties of GH class of distributions. As we have shown in Chapter 3, the zero-mean symmetric distributions are preferable in many cases. Again we set $\mu$ and $\beta$ to zeros and reparametrize $\phi = \delta/\gamma = \delta/\alpha$ and $\omega = \delta\gamma = \delta\alpha$, this parameterization is useful because the parameters $\phi, \omega$ and $\lambda$ are invariant under affine transformations (see Bibby & Sørensen, 2003). The zero-mean symmetric GH distribution $GH(\lambda, \phi, \omega)$ has the density

$$f(x; \lambda, \phi, \omega) = \frac{(\phi^2 + \frac{\phi}{\omega}x^2)^{(2\lambda-1)/4}}{\phi^\lambda \sqrt{2\pi} K_\lambda(\omega)} K_{\lambda-1/2}\left(\sqrt{\omega^2 + \frac{\omega}{\phi}x^2}\right) \tag{4.2.2}$$

The GH distribution can be interpreted as scale-location mixture of normal distribution where the mixing distribution is a *generalized inverse Gaussian* (GIG) distribution. The GIG distributions, denoted as $X \sim GIG(\lambda, \phi, \omega)$, are described by three parameters and defined on the positive half axis. The probability density function of the GIG distribution is given by

$$f(x; \lambda, \phi, \omega) = \frac{x^{\lambda-1}}{2\phi^\lambda K_\lambda(\omega)} . \exp\left(-\frac{1}{2}\omega(\phi x^{-1} + \phi^{-1}x)\right), \; x > 0. \tag{4.2.3}$$

The class of GIG distributions consists of the subclasses of gamma distributions, inverse gamma distributions and inverse Gaussian (IG) distributions. The class was first proposed by Étienne Halphen in 1946 to model the distribution of the monthly flow of water in hydroelectric stations (see Bibby & Sørensen, 2003). The subclass of IG distributions is obtained from GIG distributions where $\lambda = -1/2$, that is $IG(\phi, \omega) = GIG(-1/2, \phi, \omega)$. Since $K_{-1/2}(\omega) = \sqrt{\pi/2\omega}\exp(-\omega)$, it follows that the IG distribution does not involve the Bessel function,

$$IG(x; \phi, \omega) = \frac{x^{-3/2}\sqrt{\phi\omega}}{\sqrt{2\pi}} \cdot \exp(\omega) \cdot \exp\left(-\frac{1}{2}\omega(\phi x^{-1} + \phi^{-1}x)\right), \; x > 0. \tag{4.2.4}$$

The relationship between GH distributions and GIG distributions was originally given by Barndorff-Nielsen & Halgreen (1977) as the derivation of GH distribution from GIG distribution. The relationship can be stated as if

$$X|W \sim N(0, w) \text{ and } W \sim GIG(\lambda, \phi, \omega)$$

then the marginal distribution of $X$ is GH, $X \sim GH(\lambda, \phi, \omega)$. GH distribution is then interpreted as a scale mixture of normal distribution where the mixing distribution is GIG. As special cases, the NIG distribution arises when the mixing distribution is an IG

distribution, and the variance-gamma distribution appears when the mixing distribution is a gamma distribution. Hence, their names come from their mixing distributions. When the mixing distribution is an inverse gamma distribution, the mixture becomes the asymmetric-scaled t distribution and the t distribution arises when $\beta = 0$ as a *scale mixture of normal*. In particular, the symmetric NIG distribution and the Student-t distribution have scale mixture of normal representations as in the following definition.

### 4.2.2 Scale mixture of normal

**Definition 8.** Let $X$ be a continuous random variable with location $\mu$ and scale $\sigma$. The probability density function of $X$ is said to have a scale mixtures of normal (SMN) representation if it can be expressed as

$$f_X(x;\mu,\sigma) = \int_0^\infty N\left(x;\mu,\kappa(\lambda)\sigma^2\right)\pi(\lambda)d\lambda \tag{4.2.5}$$

where $N(x;\cdot,\cdot)$ is the normal density function, $\kappa(\lambda)$ is a positive function of $\lambda$, and $\pi(\cdot)$ is a density function defined on $\mathbb{R}^+$.

We refer to $\lambda$ and $\pi(\cdot)$, respectively, as the *mixing parameter* and *mixing density* of this scale mixture representation.

**The Student-t distribution** with location $\mu$, scale $\sigma$ and degrees of freedom $\nu$, $X \sim t_\nu(\mu,\sigma)$ can be expressed as

$$t_\nu(x;\mu,\sigma) = \int_0^\infty N(x;\mu,\frac{\sigma^2}{\lambda})Ga(\lambda;\frac{\nu}{2},\frac{\nu}{2})d\lambda$$

where $Ga(\cdot;a,b)$ is the gamma density function of the form $Ga(x;a,b) = \frac{b^a}{\Gamma(a)}\lambda^{a-1}e^{-b\lambda}$. It is equivalent to the hierarchical form

$$X|(\mu,\sigma^2,\nu,\lambda) \sim N(\mu,\frac{\sigma^2}{\lambda}) \text{ and } \lambda|\nu \sim Ga(\frac{\nu}{2},\frac{\nu}{2}).$$

**The symmetric NIG distribution** with zero mean $Y \sim NIG(\phi,\omega)$ can be represented by

$$f_Y(y;\phi,\omega) = \int_0^\infty N(y;0,\phi u)IG(u;1,\omega)du$$

where $IG(1,\omega)$ denotes the inverse Gaussian distribution with mean 1. It can be expressed hierarchically as

$$Y|(\phi,u) \sim N(0,\phi u) \text{ and } u|\omega \sim IG(1,\omega).$$

Barndorff-Nielsen & Halgreen (1977) shows that GH distribution is infinitely divisible. Thus there is a Lévy process $\{Y_t\}_{t\geq 0}$ uniquely defined by the fact that the law of $Y_1$ has GH density. We call this process the *GH Lévy process* with parameters $(\lambda,\alpha,\delta,\mu)$. The GH Lévy processes considered by Barndorff-Nielsen & Shephard (2001) becomes

a popular alternative to Brownian motion for modeling financial processes. In general, the GH densities of $Y_t$ for $t \neq 1$ are unknown, and they can not be simulated from the process in a non-intensive manner.

As stated in Barndorff-Nielsen & Shephard (2012), this model is so general that it is typically difficult to manipulate mathematically and so is not often used empirically. Instead special cases are usually employed. For examples, the *variance gamma (VG) process* by Madan & Seneta (1990) which has extensively used in the financial literature and the *hyperbolic Lévy process* by Eberlein & Keller (1995). We pay special attention to the *NIG Lévy process* by Barndorff-Nielsen (1997) which we have shown in Chapter 3 that the NIG distributions are adequately fitted to financial data. Moreover, the class of NIG distributions is closed under convolution, this implies that the distribution of $Y_t$ in the NIG Lévy process $\{Y_t\}_{t\geq 0}$ is NIG at all time point.

The classes of NIG distributions and VG distributions are only two subclasses of GH distributions that have this property. Hence the NIG and VG Lévy processes are more natural GH Lévy processes than the other GH Lévy processes. The discrete-time model for NIG Lévy process is also proposed in Barndorff-Nielsen (1997) that allows us to model stochastic volatility for daily returns, denoted by *NIG-SV* model. Note that the close-under-convolution property also asserts that if the observed process at a certain frequency scale follows a NIG distribution, then at lower frequency scales it follows a NIG distribution.

## 4.3 NIG-SV Model and HGLM Method

### 4.3.1 The model

For the excess returns $\{y_t\}_{t=1}^n$, the NIG-SV model is defined by

$$\begin{cases} y_t & = \sigma_t \epsilon_t \\ \sigma_t^2 & \sim IG(\phi, \omega) \end{cases} \tag{4.3.1}$$

where $\epsilon_t \sim$ i.i.d.$N(0,1)$. It follows that $y_t \sim NIG(\phi, \omega)$, the marginal distribution of $\{y_t\}$ and the log-likelihood function are given by (3.1.5) and (3.1.6) respectively. Moreover, the variance and the kurtosis of $y_t$ are

$$\text{var}(y_t) = \phi \text{ and } \text{kurt}(y_t) = 3/\omega.$$

### 4.3.2 H-likelihood estimation

The NIG-SV model and maximum likelihood estimation have been implemented in Chapter 3. Not only the MLE and MoM are available for estimation in the NIG-SV model, del Castillo & Lee (2008) give another approach to estimate parameters for a large class of SV models so called the *hierarchical generalized linear model (HGLM) method*. del Castillo & Lee (2008) view all the GH Lévy models in the previous section as general random effects models introduced by Lee & Nelder (2006), and therefore the associating

*hierarchical likelihood (h-likelihood)* are applied to estimate the model's parameters (see Lee & Nelder, 1996 and Lee & Nelder, 2006).

Given the NIG-SV model in (4.3.1), we firstly supposes that

$$\sigma_t^2 = \phi u_t$$

where $u_t$ are random effects from a positive infinite divisible distribution, that represent the news arriving to the markets. In this case $u_t \sim IG(1, \omega)$. This makes $\phi$ be the scale parameter whereas $E[u_t] = 1$ and $\sigma_t^2 \sim IG(\phi, \omega)$ as specified in (4.3.1). This constraint is useful in multivariate setting that we will clarify later. del Castillo & Lee (2008) gives the log-link function

$$\log(\sigma_t^2) = \alpha + b_t \tag{4.3.2}$$

where $\alpha = \log(\phi)$[1] and $b_t = \log(u_t)$. Then NIG-SV model in (4.3.1) together with (4.3.2) becomes an HGLM model with random effects in the dispersion. In particular, it is a special case of the double HGLM (DHGLM) by Lee & Nelder (2006) where the volatility involves the random effect appearing in the dispersion structure (see del Castillo & Lee, 2008). Using the joint distribution $(y_t, b_t) \sim f(y_t|b_t)f_\Theta(b_t)$, it follows that the h-likelihood is

$$h = \sum_{t=1}^{n} \{\log f(y_t|b_t) + \log f_\Theta(b_t)\} \tag{4.3.3}$$

where $f(y_t|b_t)$ and $f_\Theta(b_t)$ are the conditional density functions of $y_t$ given $b_t$ and of $b_t$ with some parameters $\Theta$. The h-likehood carries all the information in the data about the unobserved quantity $b_t$ and the fixed parameters $\Theta$ when the random effect $b_t$ is additive (see Lee & Nelder, 2006). This is the reason for writing the model with the log-link function in (4.3.2). In general the log-likelihood is obtained by $l = \log \int \exp(h) db$ that is usually difficult. Lee & Nelder (2001) propose the use of Laplace approximation to $l$, so-called the adjusted profile h-likelihood,

$$p_b(\phi, \omega) = h - 2 \log \left( -\frac{1}{2\pi} \frac{\partial^2 h}{\partial b^2} |_{b=\hat{b}} \right) \tag{4.3.4}$$

where $\hat{b}$ solves $\partial h / \partial b = 0$. Maximizing $p_b$ gives the estimates for fixed parameters $(\phi, \omega)$ and the unobserved random effects $b_t$. An advantage of the h-likelihood approach is that it does not require the integration, necessary to obtain ordinary (marginal) likelihood, and hence no analytic formula for the probability density function is needed. del Castillo & Lee (2008) report that the first-order adjusted profile h-likelihood in (4.3.4) is substantially accurate in many cases, however it can lead to non-negligible biases when apply to financial data. This problem can be solved by using the second-order improvement

$$S_b(h) = p_b(h) - F/24$$

[1] The parameter $\alpha$ in this setting is distinct from the parameter of GH distribution.

where

$$F = \text{trace}\left[\left\{3\left(-\frac{\partial^4 h}{\partial b^4}\right) - 5\left(-\frac{\partial^3 h}{\partial b^3}\right) D^{-1}\left(-\frac{\partial^3 h}{\partial b^3}\right)\right\} D^{-2}\right]_{|b=\hat{b}}$$

and $D = \text{diag}\left[\left(\partial^2 h/\partial b^2\right)\right]$. The standard errors for the estimation which are approximated from the Hessian matrices in both approximations are satisfactory (Lee & Nelder, 2006).

Particularly for the NIG-SV model where the parameter vector $\Theta$ is $(\phi, \omega)'$, the conditional density for $y_t|b_t$ is $f(y_t|b_t) = N(0, \phi u(b_t))$ and the density function for $b_t$ is given by $f_\Theta(b_t) = f\left(u(b_t)\right)\dot{u}(b_t) = \exp(b_t) IG(1, \omega)$. The explicit expressions for $\log f(y_t|b_t)$ and $\log f_\Theta(b_t)$ are

$$\log f(y_t|b_t) = -\frac{1}{2}\left\{\log(2\pi) + \log(\phi) + b_t + \frac{y_t^2 e^{-b_t}}{\phi}\right\} \tag{4.3.5}$$

$$\log f_\Theta(b_t) = -\frac{1}{2}\left\{\log(2\pi) - \log(\omega) + 3\log(b_t) - 2\omega + \omega(b_t^{-1} + b_t) - 2b_t\right\} \tag{4.3.6}$$

Hence the h-likelihood in (4.3.3) has the explicit expression as the summation of (4.3.5) and (4.3.6). The random effects can be estimated by solving $\partial h/\partial b_t = 0$, that we have

$$\hat{b}_t = \log\left(\frac{2(\omega_t - 1)}{\omega}\right) \tag{4.3.7}$$

where $\omega_t = \sqrt{1 + \omega^2 + \omega y_t^2/\phi}$. Consequently, the adjusted profile h-likelihood is expressed as

$$p_b(\phi, \omega) = n\left\{-\frac{1}{2}\log(2\pi) - \frac{1}{2}\log(\phi) + \omega + \frac{3}{2}\log(\omega)\right\} - \frac{1}{2}\sum_{t=1}^{n}\log\left[\omega_t\left(\omega_t - 1\right)^2\right] - \sum_{t=1}^{n}\omega_t \tag{4.3.8}$$

and the second-order approximation is

$$S_b(\phi, \omega) = p_b(\phi, \omega) - \frac{1}{24}\sum_{t=1}^{n}\frac{3\omega_t^2 - 5}{\omega_t^3}. \tag{4.3.9}$$

Remark that the adjusted profile h-likelihoods in both (4.3.8) and (4.3.9) do not depend on Bessel functions because the mixing distribution IG does not involve Bessel function. This is a reason why NIG-SV model is preferable than other GH models. Moreover, h-likelihood method does not need integration to find the marginal distribution as required in the maximum likelihood method.

Table 4.1 shows the estimates of parameters for NIG-SV models fitted to twenty time series with first-order, second-order h-likelihoods and maximum likelihood methods. Most estimates of $\phi$ are almost similar with the three methods of estimation. In the case of $\omega$, the first-order h-likelihood estimates and the maximum likelihood estimates are concordant but the estimates from the second-order h-likelihood show some biases. First-order h-likelihood estimations always converge, but second-order h-likelihood and maximum likelihood estimations diverge in the EUR-post series.

Table 4.1: Parameter estimation for NIG-SV models with maximum likelihood (ML), first-order h-likelihood (H1),and second-order h-likelihood (H2) with the corresponding standard errors.

| Series | ML | | H1 | | H2 | |
|---|---|---|---|---|---|---|
| | $\hat{\phi}$ | $\hat{\omega}$ | $\hat{\phi}$ | $\hat{\omega}$ | $\hat{\phi}$ | $\hat{\omega}$ |
| EUR | 0.4176 | 0.9775 | 0.4010 | 1.3232 | 0.4996 | 0.5032 |
| | (0.026) | (0.191) | (0.020) | (0.145) | (0.034) | (0.060) |
| JPY | 0.5048 | 1.2044 | 0.4961 | 1.4019 | 0.6156 | 0.5449 |
| | (0.029) | (0.238) | (0.025) | (0.155) | (0.042) | (0.067) |
| GBP | 0.4706 | 1.0675 | 0.4562 | 1.3484 | 0.5680 | 0.5275 |
| | (0.028) | (0.202) | (0.023) | (0.146) | (0.039) | (0.063) |
| EUR-pre | 0.1868 | 2.3456 | 0.1932 | 1.6627 | 0.2352 | 0.6422 |
| | (0.017) | (1.180) | (0.018) | (0.377) | (0.030) | (0.162) |
| JPY-pre | 0.2403 | 1.9291 | 0.2456 | 1.5892 | 0.3010 | 0.5960 |
| | (0.023) | (0.917) | (0.023) | (0.354) | (0.038) | (0.143) |
| GBP-pre | 0.2089 | 2.1346 | 0.2147 | 1.6111 | 0.2625 | 0.5995 |
| | (0.020) | (1.150) | (0.020) | (0.366) | (0.033) | (0.144) |
| EUR-c1 | 0.2762 | 2.6550 | 0.2869 | 1.6560 | 0.2715 | **4.6630** |
| | (0.030) | (2.125) | (0.032) | (0.472) | (0.027) | (5.715) |
| JPY-c1 | 0.5207 | 1.9923 | 0.5340 | 1.6278 | 0.6511 | 0.6540 |
| | (0.060) | (0.972) | (0.060) | (0.430) | (0.101) | (0.203) |
| GBP-c1 | 0.2587 | 5.6948 | 0.2759 | 1.9077 | 0.2573 | **8.8277** |
| | (0.025) | (5.932) | (0.030) | (0.584) | (0.024) | (10.171) |
| EUR-c2 | 1.0078 | 0.9803 | 0.9677 | 1.3220 | 1.2050 | 0.5095 |
| | (0.138) | (0.422) | (0.111) | (0.324) | (0.185) | (0.135) |
| JPY-c2 | 0.9496 | 1.6326 | 0.9600 | 1.5367 | 1.1782 | 0.6335 |
| | (0.114) | (0.734) | (0.108) | (0.389) | (0.181) | (0.185) |
| GBP-c2 | 1.2354 | 1.5836 | 1.2454 | 1.5178 | 1.5318 | 0.6002 |
| | (0.149) | (0.752) | (0.140) | (0.389) | (0.235) | (0.172) |
| EUR-post | 0.3645 | **700.72** | 0.3969 | 2.1658 | 0.3645 | **2397.30** |
| | (0.026) | (262.14) | (0.036) | (0.608) | (0.026) | NA |
| JPY-post | 0.4722 | 1.6686 | 0.4777 | 1.5406 | 0.5871 | 0.5961 |
| | (0.046) | (0.682) | (0.044) | (0.328) | (0.073) | (0.141) |
| GBP-post | 0.3963 | 17.72 | 0.4297 | 2.3435 | 0.3961 | 23.30 |
| | (0.029) | (28.87) | (0.038) | (0.663) | (0.029) | (37.98) |

Parameter estimates for NIG-SV models with three methods of estimation are almost similar with three methods. The estimates from first-order h-likelihood method and maximum likelihood method are concordant but the estimates from second-order h-likelihood show notably biases in the estimates of $\omega$. The first-order h-likelihood estimations converge in all cases while the other methods diverge in some cases written in boldfaces.

### 4.3.3 Multivariate NIG-SV model

The NIG-SV model with HGLM method can be extended to multivariate model naturally by applying the same random effects to all entries. For $\mathbf{y}_t = (y_{1,t}, y_{2,t}, ..., y_{d,t})'$, the multivariate NIG-SV model is given by

$$\begin{cases} y_{i,t} = \sigma_{i,t}\epsilon_{i,t} \\ \sigma^2_{i,t} \sim IG(\phi_i, \omega) \\ \log(\sigma^2_{i,t}) = \alpha_i + b_t \end{cases} \tag{4.3.10}$$

for $i = 1, .., d$, $t = 1, .., n$ and $\boldsymbol{\epsilon}_t = (\epsilon_{1,t}, ..., \epsilon_{d,t})' \sim N(0, \Omega)$. The random effect is introduced to the model by $\sigma^2_{i,t} = u_t\phi_i$, the $\alpha_i = \log(\phi_i)$ and $b_t = \log(u_t)$. Here $\phi_i = \sigma^2_i$ is a scale parameter in the sense that the random effect $u_t = \sigma^2_{i,t}/\phi_i$ is distributed as $IG(1, \omega)$ and the mixing distribution $\sigma^2_{i,t} = u_t\phi_i \sim IG(\phi_i, \omega)$. The random effect $u_t$ represents the news arriving to the markets that has the same effect to all indices. Suppose that $\text{var}(\boldsymbol{\epsilon}_t) = \Omega = (\rho_{ij})$, then we have

$$E(\mathbf{y}_t|u_t) = 0 \text{ and } \text{var}(\mathbf{y}_t|u_t) = u_t\Sigma$$

where $\Sigma = (\sigma_{ij})$ and $\sigma_{ij} = \sigma_i\sigma_j\rho_{ij}$. Here $\Sigma = \Lambda'\Omega\Lambda$, where the correlation matrix, $\Omega$, is assumed to be known and $\Lambda$ is the diagonal matrix of the standard deviations $\sqrt{\phi_i} = \sigma_i$. It is clearly seen that $\Sigma$ does not depend on $t$, and the distributions of $y_t$ conditional on $u_t$ are multivariate normal. Both MLE and HGLM methods for estimation are applied in the multivariate NIG-SV model. The expressions for the log likelihood, h-likelihood and its adjusted profile likelihoods are provided in the appendix.

Table 4.2 shows the parameter estimates of three currency exchange rates for different periods. Most of the estimates with three methods are almost similar. These results demonstrate the promising extensive utility of the HGLM approach. Other SV models in the class of GH models such as the VG model and scaled-t model are also extendable in this manner.

### 4.3.4 Optimization

An important step when the likelihood function is defined, either the maximum likelihood or the h-likelihood is to find the value that maximize the likelihood function. This step is called *optimization* which includes the algorithm to locate the optimal value that satisfies the criteria. Here we present two common maximization methods in financial econometrics.

#### 4.3.4.1 Method of scoring

A general way to find the value that maximize a function is done by considering the first and the second derivatives of the function. A local maximum $x_0$ of a differentiable function $f$ must satisfy the necessary conditions such that $f'(x_0) = 0$ and the $f''(x_0) < 0$.

Table 4.2: Parameter estimation for multivariate NIG-SV models in five periods.

| period | method | $\hat{\phi}_{EUR}$ | $\hat{\phi}_{JPY}$ | $\hat{\phi}_{GBP}$ | $\hat{\omega}$ |
|---|---|---|---|---|---|
| All | ML | 0.4137 | 0.5198 | 0.4671 | 0.9357 |
| | | (0.021) | (0.029) | (0.024) | (0.097) |
| | H1 | 0.3896 | 0.4877 | 0.4401 | 1.2262 |
| | | (0.018) | (0.025) | (0.021) | (0.118) |
| | H2 | 0.4197 | 0.5276 | 0.4742 | 0.8916 |
| | | (0.022) | (0.030) | (0.025) | (0.088) |
| Pre-Crisis | ML | 0.1909 | 0.2357 | 0.2080 | 2.9766 |
| | | (0.013) | (0.018) | (0.015) | (0.871) |
| | H1 | 0.1913 | 0.2362 | 0.2083 | 2.9278 |
| | | (0.013) | (0.018) | (0.014) | (0.665) |
| | H2 | 0.1927 | 0.2376 | 0.2096 | 2.6814 |
| | | (0.014) | (0.019) | (0.015) | (0.837) |
| Crisis 1 | ML | 0.2717 | 0.5046 | 0.2717 | 2.3023 |
| | | (0.026) | (0.052) | (0.027) | (0.677) |
| | H1 | 0.2708 | 0.5022 | 0.2696 | 2.4670 |
| | | (0.025) | (0.051) | (0.026) | (0.603) |
| | H2 | 0.2748 | 0.5118 | 0.2758 | 2.0606 |
| | | (0.027) | (0.054) | (0.028) | (0.619) |
| Crisis 2 | ML | 1.0078 | 0.9625 | 1.2305 | 1.2914 |
| | | (0.109) | (0.112) | (0.132) | (0.321) |
| | H1 | 0.9696 | 0.9249 | 1.1868 | 1.6173 |
| | | (0.097) | (0.100) | (0.117) | (0.358) |
| | H2 | 1.0264 | 0.9813 | 1.2522 | 1.1871 |
| | | (0.113) | (0.117) | (0.137) | (0.282) |
| Post-Crisis | ML | 0.3766 | 0.4493 | 0.4050 | 5.0156 |
| | | (0.027) | (0.035) | (0.030) | (1.741) |
| | H1 | 0.3824 | 0.4519 | 0.4110 | 4.0188 |
| | | (0.028) | (0.036) | (0.030) | (1.028) |
| | H2 | 0.3766 | 0.4494 | 0.4050 | 5.0276 |
| | | (0.027) | (0.035) | (0.030) | (1.867) |

Parameter estimates for multivariate NIG-SV models with three methods of estimation are consistent. This results demonstrate the promising extensive utility of the HGLM approach.

This is the principle idea to construct an algorithm to find that value. The first derivative of the log-likelihood function with respect to the vector of parameters $\Theta$,

$$G(\Theta) = \frac{\partial \log L(\Theta)}{\partial \Theta},$$

is known as the *gradient* or the *score*. In the case of iid, where the vector of parameters is fixed $\Theta = (\theta_1, \theta_2, \ldots, \theta_k)'$, the gradient is

$$G(\Theta) = (\partial \log L(\Theta)/\partial\theta_1, \partial \log L(\Theta)/\partial\theta_2, \ldots, \partial \log L(\Theta)/\partial\theta_k)'.$$

The maximum likelihood estimate, denoted by $\hat{\Theta}$, requires

$$G(\hat{\Theta}) = G(\Theta)\big|_{\Theta=\hat{\Theta}} = 0 \tag{4.3.11}$$

to satisfy the first-order condition for a maximum. The second-order derivative of $\log L(\Theta)$, so called the *Hessian*, is given by

$$H(\Theta) = \begin{bmatrix} \frac{\partial^2 \log L(\Theta)}{\partial\theta_1\partial\theta_1} & \frac{\partial^2 \log L(\Theta)}{\partial\theta_1\partial\theta_2} & \cdots & \frac{\partial^2 \log L(\Theta)}{\partial\theta_1\partial\theta_k} \\ \frac{\partial^2 \log L(\Theta)}{\partial\theta_2\partial\theta_1} & \frac{\partial^2 \log L(\Theta)}{\partial\theta_2\partial\theta_2} & \cdots & \frac{\partial^2 \log L(\Theta)}{\partial\theta_2\partial\theta_k} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 \log L(\Theta)}{\partial\theta_k\partial\theta_1} & \frac{\partial^2 \log L(\Theta)}{\partial\theta_k\partial\theta_2} & \cdots & \frac{\partial^2 \log L(\Theta)}{\partial\theta_k\partial\theta_k} \end{bmatrix}$$

and denote $H(\Theta_0) = H(\Theta)|_{\Theta=\Theta_0}$. It is necessary for a maximum that Hessian is *negative definite*, i.e. $\mathbf{x}'H(\hat{\Theta})\mathbf{x} < 0$, for all non-zero vector $\mathbf{x}$. To solve the equation (4.3.11), we consider the first-order Taylor series expansion of the gradient function around the true parameter $\Theta_0$,

$$G(\Theta) \simeq G(\Theta_0) + G'(\Theta_0)(\Theta - \Theta_0) \tag{4.3.12}$$

where $G'(\Theta) = H(\Theta)$. Then equation (4.3.12) can be written as

$$G(\Theta) \simeq G(\Theta_0) + H(\Theta_0)(\Theta - \Theta_0).$$

Hence the maximum likelihood estimate $\hat{\Theta}$ satisfies

$$G(\hat{\Theta}) = 0 \simeq G(\Theta_0) + H(\Theta_0)(\hat{\Theta} - \Theta_0). \tag{4.3.13}$$

When (4.3.13) is expressed as an equality, the solution is

$$\hat{\Theta} = \Theta_0 - H^{-1}(\Theta_0)G(\Theta_0). \tag{4.3.14}$$

The true parameter $\Theta_0$ is unknown and it can not be used to evaluate $\hat{\Theta}$, however equation (4.3.14) suggests the route to reach the solution. Typically, numerical procedure to obtain the solution is given by setting an initial value that lies in the plausible region, namely $\Theta_1$. Then the solution is iteratively updated by a particular scheme. The

*Newton-Raphson algorithm* uses equation (4.3.14) directly to update the $k^{th}$ solution $\Theta_{(k)}$ by

$$\Theta_{(k)} = \Theta_{(k-1)} - H^{-1}(\Theta_{(k-1)})G(\Theta_{(k-1)}) \tag{4.3.15}$$

where $H_{(k-1)} = H(\Theta)|_{\Theta=\Theta_{(k-1)}}$, $G_{(k-1)} = G(\Theta)|_{\Theta=\Theta_{(k-1)}}$ and $\Theta_{(k-1)}$ is the solution obtained by the $(k-1)^{th}$ step. The algorithm proceeds until $\Theta_{(k)} \simeq \Theta_{(k-1)}$, that the equation (4.3.15) leads to

$$\Theta_{(k)} - \Theta_{(k-1)} = -H^{-1}(\Theta_{(k-1)})G(\Theta_{(k-1)}) \simeq 0,$$

which is satisfied if $G(\Theta_{(k-1)} \simeq 0$ since $H^{-1}(\Theta_{(k-1)})$ is negative definite. Therefore the final solution $\Theta_{(k)}$ satisfies the condition that defines the maximum likelihood estimator. That is $\hat{\Theta} = \Theta_{(k)}$.

The *method of scoring* employs the information matrix

$$I(\Theta) = E\left[-\frac{\partial^2 \log L(\Theta)}{\partial\Theta\partial\Theta'}\right]$$

by replacing it to $-H^{-1}(\Theta_0)$ in (4.3.14). The updating scheme of the method of scoring is given by

$$\Theta_{(k)} = \Theta_{(k-1)} + I^{-1}(\Theta_{(k-1)})G(\Theta_{(k-1)})$$

where $I(\Theta_{(k-1)}) = I(\Theta)|_{\Theta=\Theta_{(k-1)}}$. The main problem of method of scoring arises from the difficulty in calculation of the information matrix.

#### 4.3.4.2 BHHH algorithm

The BHHH algorithm was proposed by Berndt, Hall, Hall and Hausman in 1974 (Berndt et al. (1974)). The information matrix is replaced by

$$I(\Theta) = E\left[\frac{\partial \log L(\Theta)}{\partial\Theta}\frac{\partial \log L(\Theta)}{\partial\Theta'}\right] = E\left[-\frac{\partial^2 \log L(\Theta)}{\partial\Theta\partial\Theta'}\right] \tag{4.3.16}$$

which is a result from the information matrix equality

$$E\left[\frac{\partial \log L(\Theta)}{\partial\Theta}\frac{\partial \log L(\Theta)}{\partial\Theta'}\right] + E\left[\frac{\partial^2 \log L(\Theta)}{\partial\Theta\partial\Theta\prime}\right] = 0.$$

Equation (4.3.16) also holds for the log-likelihood function at the $t^{th}$ observation,$\log L_t(\Theta)$, therefore

$$E\left[\frac{\partial \log L_t(\Theta)}{\partial\Theta}\frac{\partial \log L_t(\Theta)}{\partial\Theta'}\right] = E\left[-\frac{\partial^2 \log L_t(\Theta)}{\partial\Theta\partial\Theta'}\right].$$

The expectation $E\left[\frac{\partial \log L_t(\Theta)}{\partial\Theta}\frac{\partial \log L_t(\Theta)}{\partial\Theta'}\right]$ is estimated by

$$E\left[\frac{\partial \log L_t(\Theta)}{\partial\Theta}\frac{\partial \log L_t(\Theta)}{\partial\Theta'}\right] = \lim_{n\to\infty}\frac{1}{n}\sum_{t=1}^{n} G_t G_t'$$

where $G_t$ is the gradient evaluated at the $t^{th}$ observation, that is $G_t = \frac{\partial \log L_t(\Theta)}{\partial \Theta}$. Hence the expectation of the outer product of gradient for the full log-likelihood function is

$$E\left[\frac{\partial \log L(\Theta)}{\partial \Theta}\frac{\partial \log L(\Theta)}{\partial \Theta'}\right] = \lim_{n\to\infty}\sum_{t=1}^{n} G_t G_t'.$$

For finite observations, the expectation is estimated by

$$B(\Theta) = \sum_{t=1}^{n} G_t G_t' = \sum_{t=1}^{n} \frac{\partial \log L_t(\Theta)}{\partial \Theta}\frac{\partial \log L_t(\Theta)}{\partial \Theta'}.$$

Consequently, the updating scheme for the BHHH algorithm is

$$\Theta_{(k)} = \Theta_{(k-1)} + B^{-1}(\Theta_{(k-1)})G(\Theta_{(k-1)})$$

which does not require the second derivative to implement.

## 4.4 Latent Information and Volatility Forecasting

Volatility is latent variable, we cannot observe this quantity directly from the market. It is usually estimated from related observable quantities such as return or price range. In the previous section we have applied the h-likelihood method to estimate the model parameters for NIG-SV models. A consequential product of using h-likelihood estimation is the series of random effect estimates and hence the estimates of log volatilities.

### 4.4.1 Random effects and log volatility estimates

The h-likelihood estimation for NIG-SV model yields the estimates of the random effects $\hat{b}_t$ from equation (4.3.7). This variable is modeled to represent the news arriving to the market and the log-volatility estimate can be recovered from

$$\log(\hat{\sigma}_t^2) = \log(\hat{\phi}) + \hat{b}_t \tag{4.4.1}$$

and the estimator for volatility is consequently

$$\hat{\sigma}_t^2 = \exp\left[\log(\hat{\phi}) + \hat{b}_t\right]. \tag{4.4.2}$$

Therefore after the NIG-SV model is estimated by h-likelihood method, we yield the series of log-volatility estimates $\{\log(\hat{\sigma}_t^2)\}_{t=1}^n$. A common way to forecast volatility using (transformed) volatility estimates is to assign a proper model to the time series. The class of ARMA models is a basic class of time series models that has a vast contribution in the financial literature. Several financial time series models are constructed on the basis of ARMA models. For example the standard SV model by Taylor (1986) is set up by an AR(1) process in the log of volatilities, the GARCH(1,1) model is founded by ARMA(1,1) process. Taylor (1987) uses price ranges and returns as volatility estimates

and makes volatility forecasts by fitting the estimates to ARMA(1,1) models. Gallant et al. (1999) and Alizadeh et al. (2002) show that the sum of two AR(1) processes is capable in modeling volatility, the resulting process is ARMA(2,1). The EGARCH model advocated by Nelson (1991) is founded by assigning ARMA process to $\{\log(\sigma_t^2)\}$. Pong et al. (2004) apply ARMA(2,1) and ARFIMA models to the series $\{\log(\sigma_t)\}$ that is estimated from daily realized volatilities. The ARFIMA model is a generalization to ARMA model with fractional integration. Even though ARMA models can be extended in several ways, such as ARFIMA, and higher orders ARMA(p,q), we tend to keep our models simple and tractable with less number of parameters. Hence the ARMA(1,1) model becomes our preference. Modeling the series $\{\log(\hat{\sigma}_t^2)\}$ by ARMA(1,1) is also comparable to the EGARCH(1,1) model. A stationary ARMA(1,1) model for $\{\log(\hat{\sigma}_t^2)\}$ is given by

$$\log(\hat{\sigma}_t^2) = c + \phi_1 \log(\hat{\sigma}_{t-1}^2) + \epsilon_t + \theta_1 \epsilon_{t-1} \tag{4.4.3}$$

where $\epsilon_t \sim WN(0, \sigma_\epsilon^2)$ are *white noise* (uncorrelated random variables with zero mean and finite variance) and $|\phi_1| < 1$. The infinite MA representation of (4.4.3) is

$$\log(\hat{\sigma}_t^2) = c + \sum_{j=0}^{\infty} \psi_j \epsilon_{t-j}$$

where $\psi_0 = 1$, $\psi_1 = \phi_1 + \theta_1$ and $\psi_j = \phi_1^{j-1}(\phi_1 + \theta_1)$ for $j \geq 2$. It follows that $E[\log(\hat{\sigma}_t^2)] = c$ and $\text{var}\left(\log(\hat{\sigma}_t^2)\right) = \sigma_\epsilon^2 \sum_{j=0}^{\infty} \psi_j^2$. The parameter estimation can be achieved by minimizing the conditional sum of squares function $S(\Theta) = \sum e_t^2$ where $e_t = \epsilon_t(\Theta)$ and $\Theta = (\phi_1, \theta_1)'$. If $\log(\hat{\sigma}_1^2)$ is taken to be fixed, the prediction errors $e_t$ can be computed from the recursion

$$e_t = \log(\hat{\sigma}_t^2) - \phi_1 \log(\hat{\sigma}_{t-1}^2) - \theta_1 e_{t-1}, \quad t = 2, \ldots, n$$

with $e_1 = 0$. The minimization can be accomplished by the optimization algorithm in Subsection 4.3.4. The forecast is given by the minimum mean square estimator (MMSE) of $\log(\hat{\sigma}_{t+k}^2)$. That is equal to

$$\log(\hat{\sigma}_{t+k|t}^2) = c + \sum_{j=0}^{\infty} \psi_{k+j} \epsilon_{t-j} \tag{4.4.4}$$

and

$$\text{var}\left(\log(\hat{\sigma}_{t+k|t}^2)\right) = \text{MSE}\left(\log(\hat{\sigma}_{t+k|t}^2)\right) = (1 + \psi_1^2 + \cdots + \psi_{k-1}^2)\sigma_\epsilon^2.$$

In practice, the forecast is usually be carried out with $\epsilon_t$ replaced by $e_t = \epsilon_t(\Theta)$. If $n$ is large, the difference between $\epsilon_t$ and $e_t$ is negligible (see Harvey, 1981). The MMSE in (4.4.4) is the forecast for log volatility. To obtain the volatility forecast $\sigma_{t+k|t}^2$, applying the exponential function to the log forecast $\log(\hat{\sigma}_{t+k|t}^2)$ results in biased forecast for $\sigma_{t+k}^2$. Granger & Newbold (1976) have discussed on this issue and give the correction

$$f_{t+k|t}^{\text{NIG}} = \exp\left\{\log(\hat{\sigma}_{t+k|t}^2) + \text{var}(\log(\hat{\sigma}_{t+k|t}^2)\right\}. \tag{4.4.5}$$

### 4.4.2 Implementation and empirical results

Here we forecast volatility with the model given in (4.4.4) and (4.4.5) employing the strategy given in Section 4.1.5. The target variable is the cumulative volatility over the forecast horizon $\sigma^2_{t+k,t} = \sum_{i=1}^{k} \sigma^2_{t+i}$ that is proxied by sum squared returns $y^2_{t+k,t} = \sum_{i=1}^{k} y^2_{t+i}$. The benchmark forecasts are obtained from the random walk model

$$F^{\text{RW}}_{t+k|t} = \sum_{i=0}^{k-1} y^2_{t-i}$$

and GARCH(1,1) model

$$F^{\text{GARCH}}_{t+k|t} = \sum_{i=i}^{k} f^{\text{GARCH}}_{t+i|t}$$

where $f^{\text{GARCH}}_{t+i|t}$ is $i$-step GARCH(1,1) forecast formulated at time $t$ given by (4.1.2). Our forecasting models obtained from NIG-SV models and h-likelihood are

$$F^{\text{NIG-SV}}_{t+k|t} = \sum_{i=i}^{k} f^{\text{NIG-SV}}_{t+i|t}$$

where $f^{\text{NIG}}_{t+i|t}$ is modeled by (4.4.3).

Preliminary experiments show that forecasting the volatility with ARMA(1,1) model in (4.4.3) does not produce good results because the averages of forecasts are substantially smaller than the volatility proxies. For instance in the case of EUR series, the long-run average of the 22-day forecasts is 5.269, while the average sum of squared returns is 11.631. That is more than twice the average of forecasts. To adjust the level of forecasts to match the level of volatility proxies, we apply the simple linear regression between the volatility estimates from (4.4.2) and the squared returns. Suppose that in-sample estimates of volatility $\{\hat{\sigma}^2_1, .., \hat{\sigma}^2_n\}$ are fitted to the corresponding squared returns as

$$y^2_t = a\hat{\sigma}^2_t, \ t = 1, ..., n,$$

then the NIG-SV* forecasting model is given by

$$f^{\text{NIG-SV*}}_{t+k|t} = a f^{\text{NIG-SV}}_{t+i|t}. \tag{4.4.6}$$

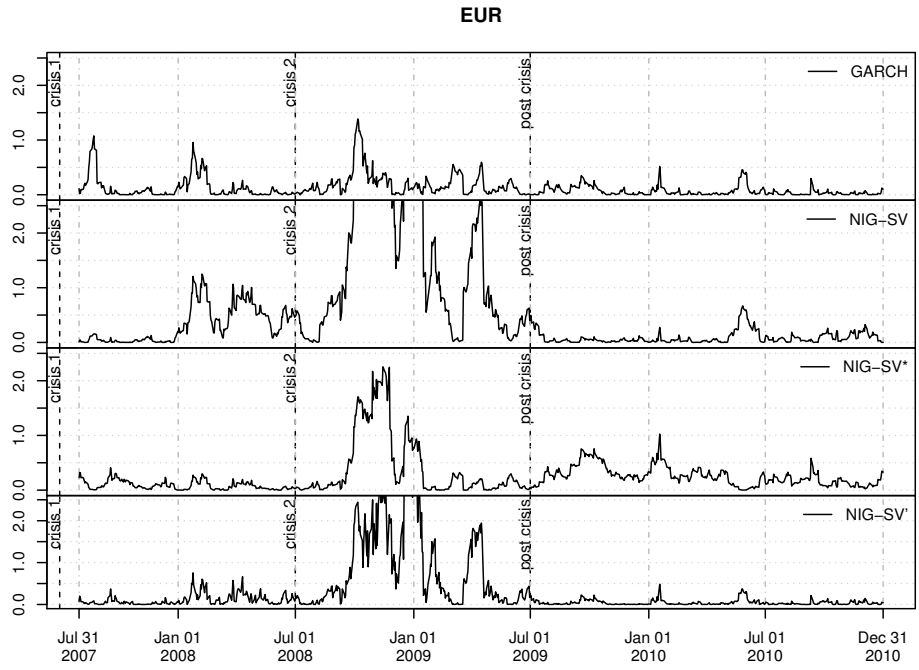
Another forecasting model that is found to considerably improve the forecast performances in our data set is obtained by fitting the log volatility to the ARMA(1,1) process with zero mean

$$\log(\hat{\sigma}^2_t) = \phi'_1 \log(\hat{\sigma}^2_{t-1}) + \epsilon_t + \theta'_1 \epsilon_{t-1}. \tag{4.4.7}$$

If the corresponding infinite MA representation is $\log(\hat{\sigma}^2_t) = \sum_{j=0}^{\infty} \psi'_j$, then the ad hoc forecasting model, denoted by NIG-SV$'$, is given by

$$f^{\text{NIG-SV}'}_{t+k|t} = \exp\left\{\log(\hat{\sigma}^2_{t+k|t}) + \text{var}(\log(\hat{\sigma}^2_{t+k|t})\right\} \tag{4.4.8}$$

Figure 4.4.1: The QL loss series of EUR forecasts




The QL loss series from four forecasting models for EUR, from top: GARCH, NIG-SV, NIG-SV* and NIG-SV′ respectively. The less the loss values, the more the accuracy of the forecasting model.

Table 4.3: Volatility forecasting with RW, GARCH, NIG-SV and NIG-SV*

| | QL loss | | | | | MSE loss | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | RW | GARCH | NIG-SV | NIG-SV* | NIG-SV′ | RW | GARCH | NIG-SV | NIG-SV* | NIG-SV′ |
| Long-run average | | | | | | | | | | |
| EUR | 0.127 | **0.120** | 0.639 | 0.259 | 0.307 | **55.61** | 55.78 | 132.13 | 108.62 | 106.11 |
| JPY | 0.300 | 0.247 | 0.546 | **0.236** | 0.310 | 129.21 | **109.12** | 141.13 | 126.91 | 115.27 |
| GBP | **0.157** | **0.157** | 0.671 | 0.324 | 0.363 | **105.70** | 107.16 | 223.42 | 189.84 | 189.38 |
| Crisis 1 average | | | | | | | | | | |
| EUR | 0.122 | 0.130 | 0.320 | **0.091** | 0.115 | 6.74 | **6.17** | 14.54 | 6.27 | 7.70 |
| JPY | 0.583 | 0.474 | 0.815 | **0.254** | 0.408 | 137.69 | 112.97 | 112.14 | **77.64** | 86.10 |
| GBP | 0.172 | 0.110 | 0.182 | **0.061** | 0.069 | 6.52 | 4.83 | 8.23 | **4.04** | 4.31 |
| Crisis 2 average | | | | | | | | | | |
| EUR | 0.211 | **0.201** | 1.762 | 0.423 | 0.881 | **168.85** | 171.30 | 423.93 | 242.83 | 347.47 |
| JPY | 0.246 | 0.257 | 0.954 | **0.196** | 0.568 | 262.91 | 230.06 | 346.77 | **186.05** | 289.55 |
| GBP | **0.227** | 0.328 | 2.008 | 0.524 | 1.080 | **328.14** | 341.00 | 738.62 | 407.56 | 628.88 |
| Post-Crisis average | | | | | | | | | | |
| EUR | 0.074 | 0.060 | 0.087 | 0.252 | **0.044** | 10.24 | 9.36 | 10.14 | 82.00 | **5.92** |
| JPY | 0.162 | 0.101 | 0.111 | 0.252 | **0.079** | 35.25 | 26.46 | 22.26 | 117.71 | **17.34** |
| GBP | 0.101 | 0.072 | 0.082 | 0.352 | **0.066** | 18.54 | 14.35 | 12.68 | 158.71 | **10.52** |

The average losses taken from different periods. The NIG-SV* is favorite in the Crisis 1 whereas the NIG-SV′ is favorite in the Crisis 2. There is no clear winner in the other periods.

where $\log(\hat{\sigma}^2_{t+k|t}) = \sum_{j=0}^{\infty} \psi'_{k+j}\epsilon_{t-j}$ and $\text{var}\left(\log(\hat{\sigma}^2_{t+k|t})\right) = (1 + \psi'^2_1 + \cdots + \psi'^2_{k-1})\sigma^2_\epsilon$.

Figure 4.4.1 shows the loss series of cumulative volatility forecasts for EUR from the beginning of Crisis 1 to the end of Post-Crisis. It can be seen graphically that the losses of forecast are higher in the Crisis 1 and extremely high in the Crisis 2 for all forecasting models. The NIG-SV model and NIG-SV′ models perform poorly in the Crisis 2 comparing to GARCH model. The losses of NIG-SV* are approximately in the middle of the other three models. And the NIG-SV′ performs nicely in the Post-Crisis period.

Numerically, the average losses taking over different periods are shown in Table4.3. First, we can see that the MSE losses are extremely high in the Crisis 2. This reflects the high dependence between the loss and the level of true volatility as we have discussed in 4.1.2. Therefore we take more attention to the QL losses. The NIG-SV* model is favorite in the Crisis 1 whereas the NIG-SV′ model is favorite in the Post-Crisis. In the Crisis 2, the average losses are relatively high and there is no clear winner in this period, GARCH, NIG-SV* and RW models win the competitions for EUR, JPY and GBP respectively.

## 4.5 Summary

In this chapter, we have presented the route to volatility forecasting from the beginning with related issues to the new models for forecasting. Self-contained guide for volatility forecasting in practice is given. Stochastic volatility models in the class of GH

models, especially the NIG-SV model, have been discussed as promising in describing and forecasting volatility. The HGLM approach for estimation and the associating h-likelihood method have been studied and implemented to the real data. The empirical results show that h-likelihood method is comparable to maximum likelihood method but h-likelihood method is simpler because the integration does not involve. Moreover, applying h-likelihood method yield the estimates of latent variables that can be use to forecast volatility. The new models for volatility forecasting based on NIG-SV model have been presented. Finally, the new forecasting models have been implemented to the real data and they perform better than the standard models in many cases.

# Chapter 5

# Range-Based Volatility Models

Volatility is unobservable in the market and the dynamics of volatility is unknown. To measure the volatility we need to exploit observable variables. In the previous chapter, returns have played a key part in volatility modeling. Nevertheless, there are other observable variables available in financial markets such as trading volume, daily high, low, open, and close prices. In this chapter, we show how other relevant variables are exploited to estimate volatility provided a particular process for prices is assumed. Price range, which is the difference between daily high and low prices, plays an important role in this section. Although the daily high and low are only two numbers, they provide a lot of information about volatility since they are extracted from prices record of the whole day.We introduce some range-based volatility models which assume geometric Brownian motions to price processes and also propose the new estimators that correct the discretization error of the existing models. The accuracy and the efficiency of range-based volatility estimators are investigated by simulations.

## 5.1 Range-Based Volatility Estimators

Data on the daily range are widely available in board assets over long historical span. In general, it does not cost the practitioners to obtain the data as in the case of intraday prices. Daily range is not only a single number, but it does represents the price movements over the whole trading day. Thus some investigators utilize the information contained in the range to estimate volatility rather than the return. Based on the assumption that intraday log prices follow a Brownian motion, Parkinson (1980) provides one of the first estimators of volatility from price range. He asserts that the estimator is five times more accurate than squared return which is the naive estimator and it is unbiased when expected returns are zero. Garman & Klass (1980) improve Parkinson's estimator by incorporating open and close prices to the estimator, results in a more accurate estimator, that is as accurate as the sum of eight squared intraday returns. Brunetti & Lildholdt (2002) compliment the results of Parkinson (1980) providing the range-based volatility estimator subject to *minimum mean squared error* (MMSE) criterion and also the range-based estimator of the standard deviation.

All these estimators only consider the special case that the expected returns are zero. Rogers & Satchell (1991) provide an estimator that is unbiased whatever the drift is, this estimator incorporates high, low, open and close prices. Alizadeh et al. (2002) discover an accurate probability link between volatility and range data when intraday prices follow a driftless GBM, that the distribution of log range given volatility is approximately normal. Further results on range-based volatility estimators appear in Beckers (1983),Gallant et al. (1999),Yang & Zhang (2000).

The estimators of Parkinson (1980), Garman & Klass (1980),Rogers & Satchell (1991) and Yang & Zhang (2000) have been tested by Shu & Zhang (2006) using both simulations and real data. They claim that the variances estimated with range-based estimators are quite close to the daily integrated variance. These estimators, except for the estimator of Yang & Zhang (2000), have been further analyzed by Molnár (2012) with numerous simulations and empirical data. He concludes that the estimator of Garman & Klass (1980) is the best estimator that the returns normalized by this estimator are approximately Gaussian. This result is consistent with the results obtained from high frequency data. The empirical results support the use of estimators for actual market data.

For range-based volatility modeling and forecasting, Taylor (1987) finds that the ranges have higher autocorrelations than that of the absolute returns. This indicates that the ranges are more likely to be forecasted accurately and becomes a satisfactory option in volatility modeling and forecasting. Range-based volatility models include the two-factor SV model of Alizadeh et al. (2002), the conditional autoregressive range model (CARR) of Chou (2005), the range-based EGARCH model of Brandt & Jones (2006), the range-based threshold conditional autoregressive (TARR) model by Chen et al. (2008) and the asymmetric smooth transition dynamic range models of Lin et al. (2012).

Range-based volatility models are developed by common assumptions : the trading is continuous, the volatility is constant during the day, and the log-price process is driftless. Hence, we firstly assume that price movements during the period of time $\tau \in [t-1, t]$, $\{P_\tau\}_{t-1 \leq \tau \leq t}$ follow geometric Brownian motion written in the differential form

$$dP_\tau = \mu_d P_\tau d\tau + \sigma P_\tau dB_\tau \tag{5.1.1}$$

where the drift rate $\mu_d$ and the diffusion coefficient $\sigma$ are assumed constants, and $B_\tau$ is a standard Brownian motion given in Section 2.6.1. Equation (5.1.1) has the well-known solution

$$P_\tau = P_0 \exp\left\{\left(\mu_d - \frac{\sigma^2}{2}\right)\tau + \sigma\left(B_\tau - B_0\right)\right\}$$

that the log price $p_t = \log(P_t)$ over the period $[t-1, t]$ follow a random walk with drift

$$p_t - p_{t-1} = \mu + \sigma\epsilon_t \tag{5.1.2}$$

where $\mu = \mu_d - \frac{\sigma^2}{2}$ and $\epsilon_t = \sigma B_t \sim N(0, \sigma^2)$. The price process is driftless if $\mu = 0$, i.e.,

$$p_t - p_{t-1} = \sigma\epsilon_t. \tag{5.1.3}$$

The highest price and the lowest price in the interval $[t-1, t]$ are $H_t = \sup\{S_\tau : t-1 \leq \tau \leq t\}$ and $L_t = \inf\{S_\tau : t-1 \leq \tau \leq t\}$. Denote the price at the beginning of the day by $O_t$ (open), the price in the end of the day by $C_t$ (close), the highest price of the the day $H_t$, and the lowest price of the day $L_t$. Then we consider the following variables

$$c_t^* = \log(C_t) - \log(O_t),\ h_t^* = \log(H_t) - \log(O_t),\ \text{and}\ l_t^* = \log(L_t) - \log(O_t).$$

Typically the price on the beginning of the day differs slightly from the closing price on the previous day. This setting is valuable when the opening jumps are concerned. However the *opening jumps* are typically small in comparison to daily volatility. Thus we assume that the opening jump does not exist, i.e. $O_t = C_{t-1}$ and consider the following quantities

$$r_t = c_t = \log(C_t) - \log(C_{t-1}),\ h_t = \log(H_t) - \log(C_{t-1}),\ \text{and}\ l_t = \log(L_t) - \log(C_{t-1}).$$

The difference between log-high and log-low is called price range, denoted by $R_t = h_t - l_t$. The estimators of Parkinson (1980), Garman & Klass (1980), Rogers & Satchell (1991) and Brunetti & Lildholdt (2002) are investigated in this chapter. For simplicity, we consider GBM in the unit interval $\tau \in [0, 1]$ and drop out the time subscript $t$.

**The Parkinson Estimator** Under the assumption that the prices follow a drift-less Brownian motion in(5.1.3), Parkinson (1980) proposes one of the first range-based volatility models,

$$\hat{\sigma}_P^2 = \frac{1}{4\log(2)}(h-l)^2. \tag{5.1.4}$$

This estimator is derived in which $\hat{\sigma}^2$ is an unbiased estimator of $\sigma^2$. It is a result from the distribution function of the range of the standard Brownian motion derived by Feller (1951). Parkinson (1980) claims that the estimator is more efficient than the square of return $r^2$ that is a very common estimator of $\sigma^2$. Parkinson showed that the relative MSE calculated by $\text{MSE}\left(\hat{\sigma}_P^2, \sigma^2\right) / \text{MSE}\left(r^2, \sigma^2\right)$ is 0.20367. This means that Parkinson's estimator is approximately five times more efficient than squared return.

**The Brunetti-Lildholdt Estimator** In addition to Parkinson's range-based estimator for $\sigma^2$ that is derived subject to unbiasness. Brunetti & Lildholdt provide the estimator for $\sigma^2$ subject to MMSE criterion.

$$\hat{\sigma}_{BL}^2 \quad = \quad \frac{4\log(2)}{9\xi(3)}\,(h-l)^2 = 0.25628\,(h-l)^2 \tag{5.1.5}$$

$(\,\cdot\,)$ denotes the Riemann zeta function. In the case of $\sigma^2$, the relative MSE increases from 0.20367 to 0.43416 when the estimator is constructed subject to the MMSE criterion.

**The Garman-Klass Estimator** Not only utilize the information contained in the range, Garman & Klass (1980) seek the minimum variance estimator based on $c$, $h$ and $l$ that can be expressed as an analytical function of $c$, $h$ and $l$. Their estimator is given by the formula

$$\hat{\sigma}^2_{GK^*} = 0.511(h-l)^2 - 0.019(c(h+l) - 2hl) - 0.383c^2.$$

The second term is very small and thus they recommend the more practical estimator neglecting the cross-product term,

$$\hat{\sigma}^2_{GK} = 0.5(h-l)^2 - (2\log(2) - 1)c^2. \tag{5.1.6}$$

**The Roger-Satchell Estimator** The previous three estimators are derived from the driftless Brownian motion. Rogers & Satchell (1991) provide the estimator that allows arbitrary drift given by

$$\hat{\sigma}^2_{RS} = h\,(h-c) + l\,(l-c) \tag{5.1.7}$$

They prove that $E\left[h\,(h-c) + l\,(l-c)\right] = \sigma^2$ , therefore $\hat{\sigma}^2_{RS}$ is an unbiased estimator of $\sigma^2$ . This estimator $\hat{\sigma}^2_{RS}$ is independent of drift, it is unbiased even though $\mu \neq 0$.

## 5.2 Simulation Study on Range-Based Volatility Estimators

The properties of some range-based volatility estimators are tested in Shu & Zhang (2006) and Molnár (2012). The common estimators that have been analyzed in both investigations are the estimators of Parkinson, Garmann and Klass, and Rogers and Satchell. Shu & Zhang (2006) focus on the effects of nonzero drift, and opening jump on the accuracy and efficiency of each estimator. The impact of changing volatility on each estimator is also tested. They conclude that if stock prices follow a GBM with small drift and with no opening jump, the three range-based estimators all provide good estimation of the true variance. If the drift term is large, the Parkinson estimator and the Garman-Klass estimator will significantly overestimate the true variance, whereas the Rogers and Satchell estimator estimator is drift independent. They also find that when the volatility is time varying, the average estimation error is smaller than the constant volatility case when the volatility is modeled by a deterministic function of price. This result shows that the range-based estimators are able to capture the short-run dynamics of volatility variation.

However, they did not take into account the effects of discretization. Molnár (2012) neither considers the discretization effects but he takes a great number of simulations that a continuous Brownian motion is approximated by a random walk with 100,000 steps repeated 500,000 simulations. He studies the performance of the estimators when all the assumptions of these estimators hold perfectly and rates the Garman–Klass estimator as the best volatility estimator based on daily (open, high, low and close) data. In Molnár (2012), the empirical results show that the returns standardized by the Garman-Klass

volatility estimates are close to normal. That is consistent with the results obtained from high-frequency data studied by Andersen et al. (2001). Therefore, in the absence of high-frequency data, further development of volatility models based on open, high, low and close prices is promising.

In this section, we study by simulations the properties of range-based volatility estimators in realistic situations that the assumptions of the estimators do not hold perfectly. We focus on the effects of nonzero drift and the changing volatility. In Shu & Zhang (2006), the volatility is simulated with three different models: the constant volatility model, the deterministic volatility model, and the jump volatility model. Here we simulate the volatility by the NIG-SV model. We also take into account the impact of discretization, hence we simulate intraday price paths according to practical record of transactions. A 5-minute prices record over 9 hours of trading, i.e., from 8:00 to 17:00, consists only 108 transactions data per day. Hence we take relatively low numbers of steps comparing to the simulations study in Shu & Zhang (2006) and Molnár (2012) that the number of intraday movements are 500 and 100,000 respectively.

The simulations are directed to examine the performance of each estimator under two assumptions: the changing drift effects and the changing volatility effects. On each day, the price is assumed to move 20, 40 and 100 steps that the numbers of movements correspond to approximately every 30, 15 and 5 minutes prices record over 9 trading hours respectively. For the constant volatility simulations, the price paths are simulated by Gaussian random walks. For the stochastic volatility simulations, the price paths are simulated by NIG-SV models.

### 5.2.1 Discretization error

In simulations, the Brownian motion is modeled by a random walk with Gaussian steps. It follows that the maximum of the random walk is in general smaller than the maximum of the Brownian motion and the minimum of the random walk is greater than the minimum of the Brownian motion. Consequently, all those range-based volatility estimators will generally underestimate true volatility in simulations. The downward bias is identified by Garman & Klass (1980) and also Beckers (1983). Rogers & Satchell (1991) give a correction of the downward bias to their estimators when the number of steps taken by the random walk in the time interval $[0, 1]$ is known. Assume that we simulate a log-price path,$\{p_0, p_{1,}, p_2, \ldots, p_k\}$, by taking $k$ steps of $1/k$ step size to represent a Brownian motion, $p_\tau$, in the time interval $\tau \in [0, 1]$. Thus the simulated high and low prices are given by $h_k = \max\{s_j : 0 \leq j \leq k\}$ and $l_k = \min\{l_j : 0 \leq j \leq k\}$ and $p_k = c_k$. They can be written as

$$h = h_k + \Delta \text{ and } l = l_k - \tilde{\Delta}$$

where $\Delta$ and $\tilde{\Delta}$ have the same law. Then the Rogers-Satchell estimator is

$$\begin{aligned}\hat{\sigma}^2_{RS} &= (h_k + \Delta)((h_k + \Delta) - p_k) + \left(l_k - \tilde{\Delta}\right)\left(\left(l_k - \tilde{\Delta}\right) - p_k\right) \\ &= \left(\Delta^2 + \tilde{\Delta}^2\right) + \Delta(2h_k - p_k) - \tilde{\Delta}(2l_k - p_k) + h_k(h_k - p_k) + l_k(l_k - p_k).\end{aligned}$$

They show that expectations $E[\Delta] = \frac{a}{\sqrt{k}}\sigma$ and $E[\Delta^2] = \frac{b}{k}\sigma^2$, where $a = \sqrt{2\pi}\left(0.25 - \left(\sqrt{2}-1\right)/6\right)$ and $b = (1 + (3\pi/4))/12$. The estimator is then corrected by replacing $\Delta, \tilde{\Delta}, \Delta^2$ and $\tilde{\Delta}^2$ by their expectations. Then the corrected estimator $\hat{\sigma}_{RSk}$ is the positive root of the equation

$$Q(\hat{\sigma}_{RSk}) \quad = \quad (1 - \frac{2b}{k})\hat{\sigma}^2_{RSk} - \frac{2a}{\sqrt{k}}(h_k - l_k)\,\hat{\sigma}_{RSk} - h_k\,(h_k - p_k) - l_k\,(l_k - p_k) = 0.$$

The solution always exists for $k > 2b$, since $Q(\sigma) \to \infty$ as $|\sigma| \to \infty$ and $Q(0) < 0$. Eventually the corrected estimator $\tilde{\sigma}_{RSk}$ is given by

$$\hat{\sigma}_{RSk} = \frac{-B + \sqrt{B^2 - 4AC}}{2A} \tag{5.2.1}$$

where $A = 1 - \frac{2b}{k}$, $B = -\frac{2a}{\sqrt{k}}(h_k - l_k)$ and $C = -h_k\,(h_k - p_k) - l_k\,(l_k - p_k)$. Remark that the final price $p_k$ is the closing price $c$.

The knowledge of $E[\Delta]$ and $E[\Delta^2]$ is very useful since we can also apply this method to correct the other estimators. The corrected Parkinson's estimator is $\hat{\sigma}_{Pk}$ where $\hat{\sigma}_{P\mathrm{k}}$ solves

$$\left(4\log(2) - \frac{4b}{k}\right)\hat{\sigma}^2_{Pk} - \frac{4a}{\sqrt{k}}(h_k - l_k)\hat{\sigma}_{Pk} - (h_k - l_k)^2 = 0 \tag{5.2.2}$$

The corrected Brunetti-Lildholdt estimator solves

$$\left(\frac{1}{0.25628} - \frac{4b}{k}\right)\hat{\sigma}^2_{BLk} - \frac{4a}{\sqrt{k}}(h_k - l_k)\hat{\sigma}_{BLk} - (h_k - l_k)^2 = 0 \tag{5.2.3}$$

And the corrected Garman-Klass estimator solves

$$\left(2 - \frac{4b}{k}\right)\hat{\sigma}^2_{GKk} - \frac{4a}{\sqrt{k}}(h_k - l_k)\hat{\sigma}_{GKk} - (h_k - l_k)^2 + (4\log(2) - 2)p_k^2 = 0. \tag{5.2.4}$$

### 5.2.2 Simulation with constant volatility

Suppose that we simulate a daily log-prices by a $k$-step Gaussian random walk, then the following day prices also follow another $k$-step Gaussian random walk starting from the previous day last price. The process continues until day $n^{th}$, then we will have a matrix of size $n \times k$ that represents the simulated log prices following a Gaussian random walk with constant volatility $\sigma$.

$$p_{i,j+1} - \quad p_{i,j} = \quad \frac{1}{k}\mu + \sigma\epsilon_{i,j}$$

where $i = \{1, \ldots, n\}$, $j = \{1, \ldots, k\}$ and $\epsilon_{i,j} \sim N(0,1)$. We take $p_{i,1} = p_{i-1,k}$ hence opening jumps are not allowed. Each $p_i = p_{i,k}$ is the closing price, $h_i^* = \max\{s_{i,j} : 1 \le j \le k\}$ and $l_i^* = \min\{s_{i,j} : 1 \le j \le k\}$ are high price and low price of day $i^{th}$ respectively. The normalized prices at day $i^{th}$ are given by

$$c_i = p_i - p_{i-1},\ h_i = h_i^* - p_{i-1} \text{ and } l_i = l_i^* - p_{i-1}.$$

The simulations are taken subject to variation in drifts, volatilities and numbers of intraday price movements. Our empirical data has average daily return of 0.0001 and average daily variance of $4.7\times10^{-5}$. Therefore the drifts are given by $\mu = 0$, 0.0005, 0.001, 0.005 and 0.01. The daily variances are taken from $\sigma^2 = 2.5\times10^{-5}$, $10\times10^{-5}$, $40\times10^{-5}$ and $90\times10^{-5}$. The numbers of intraday movements are 20, 40 and 100. In each simulation, we simulate price path for 100 days and repeat for 1000 simulations.

The major concerns on an estimator are the accuracy and the efficiency. The accuracy corresponds to the difference between the true value and the estimate. The efficiency corresponds to the uncertainty of the estimator. Here we measure the accuracy by loss functions given in Section 4.1.2, MSE and QL. In each simulation, given the volatility $\sigma$ constant over $n$ days, the accuracy of the estimator $\hat{\sigma}^2$ is measured by the average losses

$$\overline{\text{MSE}}(\hat{\sigma}^2,\sigma^2) = \frac{1}{n}\sum_{i=1}^{n}(\hat{\sigma}_i^2-\sigma^2)^2 \text{ and } \overline{\text{QL}}(\hat{\sigma}^2,\sigma^2) = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{\sigma^2}{\hat{\sigma}_i^2}-\log\frac{\sigma^2}{\hat{\sigma}_i^2}-1\right).$$

The absolute errors, $|\hat{\sigma}^2-\sigma^2|$, are also computed. The absolute error is useful when the bias is concerned. For the efficiency, we measure by the variance ratio between the Parkinson estimator and the estimator of interest. That is

$$\text{Eff}(\hat{\sigma}^2) = \text{var}(\hat{\sigma}_P^2)\big/\text{var}(\hat{\sigma}^2).$$

In each simulation, we evaluate the performance of each estimator with the indicated measures. The simulations are taken 1000 times, then we report the average values of the measurements in the final results.

### 5.2.3 Simulation with stochastic volatility

In contrast to Brownian motion that the volatility is constant, the volatility in the NIG-SV model is stochastic

$$\begin{aligned} p_{i,j+1}-p_{i,j} &= \frac{1}{k}\mu+\sigma_{i,j}\epsilon_{i,j} \\ \sigma_{i,j}^2 &\sim IG(\frac{\phi}{k},\frac{\omega}{k}) \end{aligned} \tag{5.2.5}$$

where $i=\{1,\ldots,n\}$, $j=\{1,\ldots,k\}$ and $\epsilon_{i,j}\sim N(0,1)$. As the consequence of the close-under-convolution property of NIG distribution, we have the following properties.

(i) If $y\sim IG(\phi,\omega)$ then $\lambda y\sim IG(\lambda\phi,\omega)$.

(ii) If $y_i\sim IG(\phi,\omega)$ for $i=1,2,\ldots,n$, then $\sum_{i=1}^n y_i\sim IG(n\phi,n\omega)$.

(iii) If $y\sim NIG(\phi,\omega)$ then $\lambda y\sim NIG(\lambda^2\phi,\omega)$.

(iv) If $y_i\sim NIG(\phi,\omega)$ for $i=1,2,\cdots,n$, then $\sum_{i=1}^n y_i\sim NIG(n\phi,n\omega)$.

**Proposition 9.** *Suppose $\epsilon\sim N(0,1)$ and $\sigma^2\sim IG(\phi,\omega)$, then $y=\sigma\epsilon\sim NIG(\phi,\omega)$.*

We simulate price paths by assuming that intraday volatility $\sigma^2_{i,j}$ of the day $i^{th}$ follow an inverse Gaussian distribution $\sigma^2_{i,j} \sim IG(\frac{\phi}{k}, \frac{\omega}{k})$, where $k$ is the number of intraday observations. Then the daily volatility follows $\sigma^2_i \sim IG(\phi, \omega)$. By Proposition 9, we multiply the intraday volatility to the Gaussian random variable to obtain the intraday return that follows $y_{i,j} \sim NIG(\frac{\phi}{k}, \frac{\omega}{k})$ and finally the summation of intraday returns is the daily return that follows $y_i \sim NIG(\phi, \omega)$.

The IG random variables $\sigma^2_{i,j}$ are simulated following del Castillo & Lee (2008) employing the method of Rydberg (1997). If $u$ is distributed as $IG(1, \omega')$ then $v = \omega'(u-1)^2/u$ is distributed as a chi-square distribution with degree of freedom one, $\chi^2_1$ . Therefore, we begin with simulating $v$ with distribution $\chi^2_1$ (the square of a standard normal distribution), then we consider

$$u_1 = 1 + \frac{v}{2\omega'} - \frac{\sqrt{4\omega' v + v^2}}{2\omega'}, \text{ and } u_2 = \frac{1}{u_1}.$$

The desired $u$ is chosen from $u_1$ and $u_2$ with probabilities $1/(1+u_1)$ and $u_1/(1+u_1)$ taking $u = u_1$ or $u = u_2$ respectively, it turns out that $u$ is $IG(1, \omega')$ distributed. Given $\omega' = \omega/k$ and multiply u by $1/k$, we obtain the intraday volatility $\sigma^2_{i,j} = u/k \sim IG(\phi/k, \omega/k)$. The daily volatility is simply the summation of intraday volatilities that turn out to be $\sigma^2_i \sim IG(\phi, \omega)$. In simulations, we assume that our data are generated by (5.2.5). The volatility is stochastic with fixed expected daily variance $\phi = E[\sigma^2_i] = 10 \times 10^{-5}$. In stead of vary the volatility that is already stochastic, we vary the kurtosis of the daily returns. In our empirical data, the average kurtosis is 4.42. Accordingly, we vary the kurtosis by 1, 3, 5 and 7 that results in varying the parameter $\omega = 1,\ 1/3,\ 1/5,$ and $1/7$ respectively. The drifts are taken from the same list as in the case of constant volatility simulations.

## 5.3 Discussion on the Results

Table 5.1 shows partial results from the simulations with constant volatility. In panel (a), the log prices follow driftless Brownian motion. The Garman-Klass estimator is the most accurate estimator when the accuracy is measured by MSE. It is also the most efficient estimator in term of relative variance. The corrected Garman-Klass estimator is the most accurate estimator in term of QL measure. The squared return and the (corrected) Parkinson estimators are less bias but the large MSE in the squared return show that it is very noisy estimator. The Brunetti-Lildholdt esimator has smallest standard deviation as it was constructed for.

All range-based estimators are less bias with the correction of discretization error. The QL losses of the squared returns is 4502910, that we replace any number greater 1000 by 'inf'. The extremely high level of QL measure is the result of near zero divisors. Since the QL function involves the ratio between the true value and the estimate, the QL loss increases highly if the estimate is close to zero. The naive estimator suffers the most by the QL measure because of the zero returns. In the case of the Rogers-Satchell,

Table 5.1: The accuracy and efficiency of range-based volatility estimators when log prices follow Brownian motion with intraday movements $k = 40$, daily variance $\sigma^2 = 10 \times 10^{-5}$.

(a) Log prices follow driftless Brownian motion $\mu = 0$

| | $r_i^2$ | P | $P_k$ | BL | $BL_k$ | GK | $GK_k$ | RS | $RS_k$ |
|---|---|---|---|---|---|---|---|---|---|
| Mean | **9.94** | 8.09 | 9.72 | 5.75 | 6.70 | 7.38 | 9.59 | 7.34 | 9.45 |
| Absolute error | **0.06** | 1.91 | 0.28 | 4.25 | 3.30 | 2.62 | 0.41 | 2.66 | 0.55 |
| Standard error | 1.41 | 0.57 | 0.69 | **0.41** | 0.48 | 0.43 | 0.56 | 0.52 | 0.63 |
| MSE | 19.83 | 3.76 | 4.90 | 3.52 | 3.41 | **2.69** | 3.37 | 3.48 | 4.18 |
| QL | inf | 0.46 | 0.33 | 0.88 | 0.66 | 0.48 | **0.27** | 26.77 | 0.89 |
| Efficiency | 0.20 | 1.00 | 0.78 | 1.06 | 1.09 | **1.40** | 1.13 | 1.10 | 0.93 |

(b) Log prices follow Brownian motion with drift $\mu = 0.01$

| | $r_i^2$ | P | $P_k$ | BL | $BL_k$ | GK | $GK_k$ | RS | $RS_k$ |
|---|---|---|---|---|---|---|---|---|---|
| Mean | 19.89 | 11.61 | 13.94 | 8.25 | **9.61** | 8.41 | 11.25 | 7.02 | 9.45 |
| Absolute error | 9.89 | 1.61 | 3.94 | 1.75 | **0.39** | 1.59 | 1.25 | 2.98 | 0.55 |
| Standard error | 2.46 | 0.95 | 1.14 | 0.68 | 0.79 | **0.56** | 0.75 | 0.57 | 0.72 |
| MSE | 69.63 | 8.69 | 13.71 | 4.56 | 5.79 | **2.96** | 5.04 | 3.90 | 4.68 |
| QL | inf | 0.37 | 0.30 | 0.63 | 0.49 | 0.40 | **0.24** | 12.26 | 1.08 |
| Efficiency | 0.13 | 1.00 | 0.63 | 1.88 | 1.49 | **2.94** | 1.75 | 2.26 | 1.92 |

In panel (a), the corrected Garman-Klass estimator ($GK_k$) is the most accurate estimator in term of QL measure. The squared return ($r_i^2$) and the Parkinson estimator and its correction (P and $P_k$) are less bias. The large MSE in the squared return show that it is very noisy estimator. The Brunetti-Lildholdt esimator has smallest standard deviation. All range-based estimators are less bias with the correction of discretization error.
In panel (b), when the drift is very large, the Garman-Klass estimator (GK) still performs properly with small QL and MSE. The Rogers-Satchell (RS) and its correction ($RS_k$) are robust under the change in drift as they are expected. The Parkinson estimator (P) and the Brunetti-Lildholdt estimator (BL) including their corrections ($P_k$ and $BL_k$) deteriorate by the increasing drift when measure by MSE.
Note: Each simulation has 100 days with $k$ movements per day. The simulation is repeated 1000 times. The MSE and QL measure the accuracy of the estimator. The smaller the value of MSE (QL), the greater the accuracy of the estimator. The efficiency is relative to the Parkinson estimator. The higher value of efficiency is preferred. The absolute error measures the average distant from the true volatility. The values of Mean, Absolute error and Standard error are scaled by $10^5$, the values of MSE are scaled by $10^9$.

Table 5.2: The effect of discretization when log prices follow Brownian motion with $\sigma^2 = 10$. The averages of variance estimates ($\times 10^5$ ) with 95% confidence intervals are presented.

(a) Log prices follow driftless Brownian motion $\mu = 0$

| $k$ | $r_i^2$ | P | $P_k$ | BL | $BL_k$ | GK | $GK_k$ | RS | $RS_k$ |
|---|---|---|---|---|---|---|---|---|---|
| 20 | 9.97 | 7.46 | 9.73 | 5.30 | 6.61 | 6.49 | 9.57 | 6.41 | 9.32 |
| | (2.73) | (1.06) | (1.39) | (0.76) | (0.94) | (0.78) | (1.14) | (0.95) | (1.26) |
| 40 | 9.94 | 8.09 | 9.72 | 5.75 | 6.70 | 7.38 | 9.59 | 7.34 | 9.45 |
| | (2.77) | (1.13) | (1.35) | (0.80) | (0.93) | (0.85) | (1.10) | (1.01) | (1.23) |
| 100 | 10.01 | 8.75 | 9.80 | 6.22 | 6.84 | 8.26 | 9.69 | 8.24 | 9.62 |
| | (2.77) | (1.17) | (1.31) | (0.83) | (0.91) | (0.92) | (1.08) | (1.06) | (1.21) |

(b) Log prices follow driftless Brownian motion $\mu = 0.01$

| $k$ | $r_i^2$ | P | $P_k$ | BL | $BL_k$ | GK | $GK_k$ | RS | $RS_k$ |
|---|---|---|---|---|---|---|---|---|---|
| 20 | 20.14 | 10.96 | 14.30 | 7.79 | 9.72 | 7.42 | 11.45 | 5.95 | 9.34 |
| | (4.78) | (1.71) | (2.23) | (1.21) | (1.51) | (0.88) | (1.37) | (0.99) | (1.33) |
| 40 | 19.89 | 11.61 | 13.94 | 8.25 | 9.61 | 8.41 | 11.25 | 7.02 | 9.45 |
| | (4.83) | (1.87) | (2.24) | (1.33) | (1.54) | (1.09) | (1.46) | (1.12) | (1.40) |
| 100 | 20.02 | 12.35 | 13.83 | 8.78 | 9.65 | 9.39 | 11.20 | 8.01 | 9.58 |
| | (4.72) | (1.78) | (1.99) | (1.27) | (1.39) | (1.03) | (1.24) | (1.09) | (1.24) |

The original estimators poorly estimate the true variance when the number of intraday movements is as small as 20. The corrections significantly improve the bias both when the drift is zero and the drift is large.

the estimator is zero if the intraday prices move in one direction that the close price equals to either the high or the low price.

In Table 5.1 panel (b), the log prices follow Brownian motion with (very large) drift. The Garman-Klass is impressively robust. It has small QL, the MSE increases slightly but the bias is even smaller. The Rogers-Satchell and its correction are robust under the change in drift as they are expected. The Parkinson estimator and the Brunetti-Lildholdt estimator including their corrections deteriorate by the increasing drift when we measure the MSE. Nevertheless, they perform better when the QL measure is employed. In fact, all estimators improve in term of QL losses when the drift is increased. The full results of all sixty simulation schemes in the end of this chapter show that the Garman-Klass estimator is the best performing estimator. It wins 59 of 60 in the MSE competitions while its correction wins all the QL competitions.

Table 5.2 show the effect of discretization error. In the case of driftless Brownian motion, all the range-based estimator underestimate the true volatility if the number of intraday movements is as small as 20. The true volatility is not in 95% confidence interval of any uncorrected estimator. The corrected estimators significantly improve the downward bias. Increasing the intraday movements to 100 does not guarantee that

Table 5.3: The accuracy and efficiency of range-based volatility estimators when log prices follow NIG-SV model with intraday movements $k = 40$, expected variance $E[\sigma_i^2] = 10 \times 10^{-5}$, kurtosis $1/\omega = 5$.

(a) Log prices follow NIG-SV model $\mu = 0$, average variance $= 9.8 \times 10^{-5}$

| | $r_t^2$ | P | $P_k$ | BL | $BL_k$ | GK | $GK_k$ | RS | $RS_k$ |
|---|---|---|---|---|---|---|---|---|---|
| Mean | 9.85 | 5.53 | 6.64 | 3.93 | 4.58 | 3.86 | 5.19 | 3.23 | 4.36 |
| Absolute error | **0.05** | 4.27 | 3.16 | 5.87 | 5.22 | 5.93 | 4.61 | 6.57 | 5.44 |
| Standard error | 2.48 | 1.07 | 1.28 | 0.76 | 0.88 | 0.63 | 0.85 | **0.57** | 0.74 |
| MSE | 85.61 | 14.21 | 15.55 | 14.46 | 14.05 | 14.24 | 12.65 | 17.09 | **4.60** |
| QL | inf | 2.07 | **1.60** | 3.33 | 2.70 | 2.47 | 1.61 | 12.78 | 3.59 |
| Efficiency | 0.26 | 1.00 | 0.95 | 0.98 | 1.00 | 1.03 | 1.15 | 0.85 | **3.25** |

(b) Log prices follow NIG-SV model $\mu = 0.01$, average variance $= 9.72 \times 10^{-5}$

| | $r_t^2$ | P | $P_k$ | BL | $BL_k$ | GK | $GK_k$ | RS | $RS_k$ |
|---|---|---|---|---|---|---|---|---|---|
| Mean | 19.68 | 9.21 | 11.06 | 6.54 | 7.63 | 5.17 | 7.16 | 3.47 | 4.88 |
| Absolute error | 9.96 | **0.51** | 1.34 | 3.18 | 2.10 | 4.56 | 2.56 | 6.26 | 4.85 |
| Standard error | 3.26 | 1.28 | 1.54 | 0.91 | 1.06 | 0.70 | 0.95 | **0.67** | 0.85 |
| MSE | 85.61 | 15.78 | 19.65 | 13.66 | 14.03 | 12.71 | 11.60 | 16.75 | **4.60** |
| QL | inf | 0.77 | **0.66** | 1.14 | 0.95 | 1.18 | 0.73 | 51.81 | 4.56 |
| Efficiency | 0.29 | 1.00 | 0.82 | 1.16 | 1.12 | 1.33 | 1.42 | 0.99 | **3.59** |

In both cases, the corrected Rogers-Satchell estimator is significantly better than the others in term of MSE measure, whereas the corrected Parkinson estimator performs best with QL measure. The corrections improve the downward bias only slightly. When the drift is introduced, the MSE change very little but the QL decrease in most cases.

the true volatility will be met. In the case of Brownian motion with drift, all but the Parkinson estimator have average estimates lower than the true volatility at 20 and 40 intraday movements. The Parkinson estimator becomes overestimating when the number of movements increase to 100 but the other estimators improve with the increasing number of movements and the corrections.

Table 5.3 show the partial results when the log price is assumed to follow the NIG-SV model. The best performing estimators are almost identical in both cases with large drift and without drift. The corrected Rogers-Satchell estimator is significantly better than the others in term of MSE measure, whereas the corrected Parkinson estimator performs best with QL measure. The corrections improve the downward bias only slightly. Among the uncorrected estimators, all but the Rogers-Satchell estimators have almost similar MSE when the drift is zero. When the large drift is introduced, the Garman-Klass and the Brunetti-Lildholdt estimators have smaller MSE whereas the QL losses decrease in all uncorrected estimator except for the Rogers-Satchell estimator.

Table 5.4 show that the number of intraday movements in simulation has minor effect on the bias. In the case of driftless simulations, all but squared returns underestimate the

Table 5.4: The effect of discretization when log prices follow NIG-SV model with $E[\sigma_i^2] = 10 \times 10^5$, kurtosis=5. The averages of variance estimates ($\times 10^5$ ) with 95% confidence intervals are presented.

(a) Log prices follow NIG-SV model $\mu = 0$

| $k$ | $r_i^2$ | P | $P_k$ | BL | $BL_k$ | GK | $GK_k$ | RS | $RS_k$ |
|---|---|---|---|---|---|---|---|---|---|
| 20 | 9.60 | 5.26 | 6.86 | 3.74 | 4.66 | 3.59 | 5.53 | 2.96 | 4.62 |
| | (5.50) | (2.29) | (2.99) | (1.63) | (2.03) | (1.26) | (1.99) | (1.09) | (1.62) |
| 40 | 9.85 | 5.53 | 6.64 | 3.93 | 4.58 | 3.86 | 5.19 | 3.23 | 4.36 |
| | (4.86) | (2.09) | (2.51) | (1.48) | (1.73) | (1.23) | (1.67) | (1.12) | (1.45) |
| 100 | 9.78 | 5.65 | 6.33 | 4.02 | 4.42 | 4.06 | 4.86 | 3.45 | 4.14 |
| | (5.33) | (2.26) | (2.53) | (1.60) | (1.76) | (1.30) | (1.57) | (1.15) | (1.35) |

(b) Log prices follow NIG-SV model $\mu = 0.01$

| $k$ | $r_i^2$ | P | $P_k$ | BL | $BL_k$ | GK | $GK_k$ | RS | $RS_k$ |
|---|---|---|---|---|---|---|---|---|---|
| 20 | 19.47 | 8.88 | 11.58 | 6.31 | 7.87 | 4.79 | 7.75 | 3.07 | 5.20 |
| | (6.17) | (2.45) | (3.19) | (1.74) | (2.17) | (1.37) | (2.13) | (1.35) | (1.92) |
| 40 | 19.68 | 9.21 | 11.06 | 6.54 | 7.63 | 5.17 | 7.16 | 3.47 | 4.88 |
| | (6.39) | (2.51) | (3.02) | (1.79) | (2.08) | (1.37) | (1.86) | (1.31) | (1.66) |
| 100 | 19.99 | 9.52 | 10.66 | 6.76 | 7.44 | 5.48 | 6.69 | 3.77 | 4.61 |
| | (6.63) | (2.63) | (2.95) | (1.87) | (2.06) | (1.44) | (1.75) | (1.36) | (1.58) |

The number of intraday movements in simulation has minor effect on the bias. In the case of driftless simulations, all but squared returns underestimate the expected variance even though the number of intraday movements is increased to 100. When the drift is added, the Parkinson estimator is considered to be the best estimator because of small bias and high efficiency.

expected variance even though the number of intraday movements is increased to 100. The squared return is unbiased estimator if the drift does not involve, however it is very noisy because of high standard deviation. The Parkinson estimator has downward bias if the drift is zero but it is more accurate when the large drift is added. The Parkinson estimator is considered to be the best estimator because the small bias is compensated by the high efficiency.

Table 5.5 show the performance comparison of the estimators when the price processes follow different models. Clearly, most of the estimators are less accurate when the volatility is stochastic. Nevertheless, the range-based estimators are still a lot better than the squared return. With the MSE measure, the corrected Rogers-Satchell is the most robust estimator under the change in volatility whatever the drift is.

## 5.4 Summary

The simulation study in this chapter shows that the range-based volatility estimators are very efficient comparing to the squared return in both cases that the true volatility

Table 5.5: The accuracy of the range-based estimators when the price processes follow GBM and NIG-SV models.

| | Model | $r_t^2$ | P | $P_k$ | BL | $BL_k$ | GK | $GK_k$ | RS | $RS_k$ |
|---|---|---|---|---|---|---|---|---|---|---|
| Panel A : $\mu = 0$ | | | | | | | | | | |
| MSE | GBM | 19.83 | 3.76 | 4.90 | 3.52 | 3.41 | 2.69 | 3.37 | 3.48 | 4.18 |
| | NIG-SV | 85.61 | 14.21 | 15.55 | 14.46 | 14.05 | 14.24 | 12.65 | 17.09 | 4.60 |
| QL | GBM | inf | 0.46 | 0.33 | 0.88 | 0.66 | 0.48 | 0.27 | 26.77 | 0.89 |
| | NIG-SV | inf | 2.07 | 1.60 | 3.33 | 2.70 | 2.47 | 1.61 | 12.78 | 3.59 |
| Panel B: $\mu = 0.01$ | | | | | | | | | | |
| MSE | GBM | 69.63 | 8.69 | 13.71 | 4.56 | 5.79 | 2.96 | 5.04 | 3.90 | 4.68 |
| | NIG-SV | 85.61 | 15.78 | 19.65 | 13.66 | 14.03 | 12.71 | 11.60 | 16.75 | 4.60 |
| QL | GBM | inf | 0.37 | 0.30 | 0.63 | 0.49 | 0.40 | 0.24 | 12.26 | 1.08 |
| | NIG-SV | inf | 0.77 | 0.66 | 1.14 | 0.95 | 1.18 | 0.73 | 51.81 | 4.56 |

Most of the estimators are less accurate when the volatility is stochastic but they are a lot better than the squared return. With the MSE measure, the corrected Rogers-Satchell is the most robust estimator under the change in volatility regardless the size of the drift.

is constant and stochastic. Even though the price process involves the drift, the range-based estimators are still accurate. The number of intraday movements can be used to increase the accuracy of the estimators. This number can be replaced by the number of intraday transactions if it is available. If the number of intraday movements is unknown, the Garman-Klass estimator performs relatively accurate in both constant and stochastic volatility models. The Parkinson estimator is also a proper estimator when the volatility is stochastic. Therefore, the use of range and exogenous variables is very satisfactory in modeling the volatility. We take this advantage to create a new stochastic volatility model incorporating exogenous variables in the next chapter.

# Chapter 6

# Dynamic Normal Inverse Gaussian Models

In this chapter, we incorporate all the information obtained from previous chapters to construct new stochastic volatility models. The models are defined and their properties are given in the first section. Then the methods for estimation and forecasting are proposed. The estimation methods are implemented to the real data and the parameters are tested for the signification. The forecast performances are compared to the standard models and the NIG-SV models given in Chapter 4.

## 6.1 The DNIG Model

In previous chapters, several volatility models have been studied. We found that the NIG-SV models are well fitted to the data and the HGLM method of estimation allows us to make forecast with NIG-SV models. Moreover, the simulations in Chapter 5 illustrate the relevance of using the price range as volatility estimator. In this chapter, we propose new SV models that take advantages of the previous models by incorporating the range in to the NIG-SV model and applying the HGLM approach to estimate the model's parameters.

The new volatility model is developed by combining the ideas of three volatility models: the standard SV model, the NIG-SV model and the range-based volatility estimators. The standard SV model (Taylor, 1986) in (2.5.9) is defined by a Gaussian AR(1) process for the logarithm of volatility. It receives a lot of attention because of the capability in describing volatility and return. However, the parameter estimation for the standard SV model is complicated when the MLE is applied to estimate the parameters because the likelihood function is difficult to compute. In another way, the NIG-SV model that the volatility is specified by a random variable is capable of explaining the distribution of returns and the estimation can be easily done by MLE or H-likelihood. We also find that the variance and the kurtosis of returns that are related to the parameters of the NIG-SV model significantly change during the crisis. It motivates us to provide the dynamics that drives the volatility in the NIG-SV model as in the standard SV

model.

The difficulty of estimating the standard SV model arises from the inability to observe the past volatility in the AR(1) process, therefore the estimation requires excessive computation. Substituting the unobserved variable $\log(\sigma_{t-1}^2)$ by a proper volatility estimates would simplify the estimation procedure. Here we choose the range as the volatility estimator as we have proven its relevance in the previous chapter. The combination of these ideas lead us to the *dynamic NIG stochastic volatility (DNIG) model.*

### 6.1.1 Definition

Suppose that the volatility are constant over a unit-time period and the intraday prices follow

$$\begin{aligned} dp_\tau &= \sigma_\tau dB_\tau \\ \sigma_\tau &= \sigma_t \ \forall \tau \in (t-1, t]. \end{aligned}$$

Define the daily log-range by $R_t := \max_{\tau\in(t-1,t]}\{\log S_\tau\} - \min_{\tau\in(t-1,t]}\{\log S_\tau\} = \log(H_t) - \log(L_t)$. The dynamic NIG stochastic volatility model of order 1, denoted by DNIG(1), is defined by

$$\begin{cases} y_t & = \sigma_t \epsilon_t, \ \epsilon_t \sim \text{i.i.d.} N(0,1) \\ \log(\sigma_t^2) & = \alpha + \beta \log(R_{t-1}^2) + b_t \\ u_t & = \exp(b_t) \sim IG(1, \omega). \end{cases} \tag{6.1.1}$$

There are three parameters, namely $\alpha, \beta$ and $\omega$. This specification defines the dynamics for the logarithm of volatility as the combination of the observation $\alpha + \beta \log(R_{t-1}^2)$ and the random effect $b_t$. The distribution of volatility conditional on past history is inverse Gaussian with a constant parameter $\omega$ and a time-varying parameter $\phi_t$,

$$\sigma_t^2 | \mathcal{F}_{t-1} = \exp\left(\alpha + \beta \log(R_{t-1}^2)\right) u_t \sim IG(\phi_t, \omega),$$

where $\phi_t = \exp\left(\alpha + \beta \log(R_{t-1}^2)\right)$ and $\mathcal{F}_t$ is the set of information up to time $t$. Then the distribution of the return conditional on past information is normal inverse Gaussian,

$$y_t | \mathcal{F}_{t-1} \sim NIG(\phi_t, \omega). \tag{6.1.2}$$

Thus the variance and the kurtosis of the return conditional on past history is

$$\text{var}(y_t | \mathcal{F}_{t-1}) = \phi_t = \exp\left(\alpha + \beta \log(R_{t-1}^2)\right) \text{ and } \text{kurt}(y_t | \mathcal{F}_{t-1}) = \omega.$$

It is easily seen that the DNIG(1) model is the NIG-SV model if the parameter $\beta$ equals to zero. The DNIG model is constructed by assigning the dynamics to the volatility and the returns are conditionally distributed as NIG, hence we name this model the dynamic NIG-SV model.

### 6.1.2 Higher order setting

The DNIG(1) model simply considers the volatility driven by the previous log-range. The model can be extended to incorporate higher order of past history. The DNIG model of order $k$, denoted by DNIG(p), is defined by

$$\begin{cases} y_t & = \sigma_t \epsilon_t, \ \epsilon_t \sim \text{i.i.d.} N(0,1) \\ \log(\sigma_t^2) & = \alpha + \sum_{i=1}^{p} \beta_i \log(R_{t-i}^2) + b_t \\ u_t & = \exp(b_t) \sim IG(1,\omega). \end{cases} \tag{6.1.3}$$

This extension is one of most advantages of the DNIG model comparing to the standard SV model because the standard SV model has never been extended to incorporate higher order of autoregression. One of the reasons that the standard SV models with higher orders have not been investigated because the estimation is cumbersome even in the case of the first order standard SV model. The DNIG model allows us to investigate more in the short-term dynamics of volatility. In the later section, we mainly employ the DNIG(1) and the DNIG(2) models to explore their properties.

## 6.2 Parameter Estimation

Another advantage of the DNIG model is that the dynamics involves observable variables, the ranges, instead of the past volatility that is unobservable. The conditional distributions of observed returns are known and thus the estimation is uncomplicated. The parameter estimation is done in two approaches, the maximum likelihood estimation and the h-likelihood estimation. Here we give the explicit expression for estimation in the case of DNIG(1) model. The methods also apply to the higher order models.

### 6.2.1 Maximum likelihood estimation

The knowledge of the conditional distribution of the return allows us to construct the likelihood function for parameters estimation.Suppose that the returns $\{y_t\}_{t=1}^n$ follow the DNIG(1) model with the set of parameters $\Theta = (\alpha, \beta, \omega)'$. Given the observed information set $\mathcal{F}_n$ including the returns $\{y_1, y_2, ..., y_n\}$ and the ranges $\{R_1, \ldots, R_n\}$, the distribution of the return $y_t$ conditional on past history is normal inverse Gaussian as in (6.1.2). The probability density function of the NIG$(\phi, \omega)$ distribution has been given in (3.1.5).Hence the log-likelihood function is explicitly written in terms of $\alpha, \beta$ and $\omega$ as

$$\begin{aligned} l(\alpha,\beta,\omega) &= \sum_{t=1}^{n} \log(f(y_t|\mathcal{F}_{t-1};\alpha,\beta,\omega)) \\ &= n\left[\log(\omega) + \omega - \log(\pi)\right] + \sum_{t=1}^{n} \left\{ -\frac{1}{2} \log\left(y_t^2 + \omega \exp(\alpha + \beta \log(R_{t-1}^2)\right) \right\} \\ &\quad + \sum_{t=1}^{n} \log\left[ K_1\left( \sqrt{\frac{\omega y_t^2}{\exp\left(\alpha + \beta \log(R_{t-1}^2)\right)} + \omega^2 } \right) \right]. \end{aligned} \tag{6.2.1}$$

Maximizing the log-likelihood function in (6.2.1) provides an appropriate estimate of the parameter $\Theta$ from the observed information.

### 6.2.2 H-likelihood estimation

The model (6.1.1) can be viewed as HGLM with random effects in the dispersion. Then the associated h-likelihood is applied

$$h(b;\Theta) = \sum_{t=1}^{n} \{\log f(y_t|b_t;\alpha,\beta) + \log f(b_t;\omega)\}$$

where $b = \{b_t\}_{t=1}^n$ is the vector of random effects. Since $y_t|b_t$ is normally distributed with zero mean and variance $\sigma_t^2$ , then

$$\begin{aligned}\log f(y_t|b_t;\alpha,\beta) &= -\frac{1}{2}\log(\sigma_t^2) - \frac{1}{2}\log(2\pi) - \frac{1}{2}\frac{y_t^2}{\sigma_t^2} \\ &= -\frac{1}{2}\left\{\left(\alpha + \beta\log(R_{t-1}^2) + b_t\right) + \log(2\pi)\right\} \\ &\quad -\frac{1}{2}y_t^2\exp\left[-\left(\alpha + \beta\log(R_{t-1}^2) + b_t\right)\right] \qquad (6.2.2)\end{aligned}$$

The pdf of $b_t$ is $f(b_t;\omega) = |\partial u(b_t)/\partial b_t| \cdot f_u\left(u(b_t);\omega\right)$ where $u(b_t) = \exp(b_t) \sim IG(1,\omega)$. The pdf of $u \sim IG(1,\omega)$ has been given in (4.2.4) , therefore

$$\begin{aligned}\log f(b_t;\omega) &= \log\left(\exp(b_t)\cdot f(\exp(b_t);\omega)\right) \\ &= \frac{1}{2}\log(\omega) - \frac{1}{2}\log(2\pi) + \omega - \frac{b_t}{2} - \frac{\omega}{2}\left(\exp(-b_t) + \exp(b_t)\right) \qquad (6.2.3)\end{aligned}$$

Hence the h-likelihood is (6.2.2) plus(6.2.3). At a fix time $t$, we use the notation $h_t(b_t;\Theta) = \log f(y_t|b_t;\alpha,\beta) + \log f(b_t;\omega)$, then

$$\begin{aligned}h_t(b_t;\Theta) &= \omega + \frac{1}{2}\log(\omega) - \frac{1}{2}\alpha - \log(2\pi) - \frac{1}{2}\beta\log(R_{t-1}^2) \\ &\quad -\frac{\omega}{2}\left(\exp(-b_t) + \exp(b_t)\right) - \frac{1}{2}y_t^2\exp\left[-\left(\alpha + \beta\log(R_{t-1}^2) + b_t\right)\right] - b_t\end{aligned}$$

Consequently we reach

$$\begin{aligned}h(b;\Theta) &= n\left[\omega + \frac{1}{2}\log(\omega) - \frac{1}{2}\alpha - \log(2\pi)\right] - \frac{1}{2}\beta\sum_{t=1}^{n}\log(R_{t-1}^2) \\ &\quad -\frac{1}{2}\omega\sum_{t=1}^{n}\left(\exp(-b_t) + \exp(b_t)\right) - \frac{1}{2}\sum_{t=1}^{n}\left\{y_t^2\exp\left[-\left(\alpha + \beta\log(R_{t-1}^2) + b_t\right)\right] + 2b_t\right\}\end{aligned}$$

The marginal log-likelihood $l = \int \exp(h)db$ is approximated by the Laplace approximation (Lee & Nelder, 2001)

$$p_b(h) = h - \frac{1}{2}\log\det\left\{D(h,b)/(2\pi)\right\}|_{b=\hat{b}} \qquad (6.2.4)$$

where $D(h,b) = -\partial^2 h/\partial b^2$ and $\hat{b}$ solves $\partial h/\partial b = 0$. Consider

$$\frac{\partial h}{\partial b_t} = -1 + \frac{1}{2} y_t^2 \exp\left[-\left(\alpha + \beta \log(R_{t-1}^2) + b_t\right)\right] - \frac{1}{2}\omega\left(-\exp(-b_t) + \exp(b_t)\right)$$

and

$$\frac{\partial^2 h}{\partial b_s \partial b_t} = \begin{cases} -\frac{1}{2} y_t^2 \exp\left[-\left(\alpha + \beta \log(R_{t-1}^2) + b_t\right)\right] - \frac{1}{2}\omega\left(\exp(-b_t) + \exp(b_t)\right), & s = t \\ 0 & s \neq t \end{cases}$$

then $\hat{b}_t$ that solves $\partial h/\partial b_t = 0$ is

$$\hat{b}_t = \log\left(\frac{-1 + \sqrt{\omega^2 + \omega y_t^2 \exp\left[-\left(\alpha + \beta \log(R_{t-1}^2)\right)\right] + 1}}{\omega}\right) \tag{6.2.5}$$

The model parameter $\Theta = (\alpha, \beta, \omega)'$ is directly obtained by maximizing (6.2.4) and the estimates of random effects $\hat{b} = (\hat{b}_1, ..., \hat{b}_n)$ are extracted by (6.2.5). The standard errors ,$std.err$, can be computed from the hessian $\hat{H} = [\partial^2 p_b/\partial\theta^2]|_{\theta=\hat{\theta}}$, $std.err_i = \sqrt{\hat{I}_{i,i}}$ where $\hat{I} = -\hat{H}^{-1}$. Eventually, the volatility is estimated by

$$\hat{\sigma}_t^2 = \exp\left(\alpha + \beta \log(R_{t-1}^2) + \hat{b}_t\right). \tag{6.2.6}$$

### 6.2.3 Empirical results on estimation

Table 6.1 shows that the estimated parameters with both maximum likelihood and h-likelihood methods when the returns follow the DNIG(1) model. The estimated parameter from both methods are very similar, especially the estimates of $\alpha$ and $\beta$. The estimates of $\omega$ are also lie in 95% confidence intervals of each other in most cases even though they are slightly different.

The estimates of $\omega$ in GBP-c1, EUR-post and GBP-post are extremely high and have large standard errors with the maximum likelihood method whereas the estimates from h-likelihood method are more consistent with the estimates from other series. The higher estimated values of $\omega$ in EUR-post, GBP-post, EUR-c1 and GBP-c1 correspond to the lower kurtosis in the marginal distributions presented in Section 3.2. That the kurtosis are 0.09, 0.19, 0.30 and 0.39 respectively. When the time series are taken from the whole period of EUR, JPY and GBP, the parameter $\beta$ are significantly different from zero. The estimates of $\beta$ range from 0.319 to 0.469 which are close to the constant of the Parkinson estimator 0.361. Table 6.2 shows similar results when the returns follow the DNIG(2) models.

Table 6.3 show the log-likelihood ratio statistics when the goodness-of-fit are compared among the NIG-SV, the DNIG(1) and the DNIG(2) models. The null hypothesis is that the null model which is the restricted case of the alternative model are similar fitted to the data. The alternative hypothesis is that the alternative model is better

Table 6.1: Parameter Estimates and their 95% confidence intervals for DNIG(1) models with MLE and h-likelihood methods.

| Series | Maximum likelihood | | | H-likelihood | | |
|---|---|---|---|---|---|---|
| | $\alpha$ | $\beta$ | $\omega$ | $\alpha$ | $\beta$ | $\omega$ |
| EUR | -0.846 | 0.356 | 1.461 | -0.843 | 0.357 | 1.495 |
| | (0.057) | (0.046) | (0.336) | (0.052) | (0.046) | (0.176) |
| JPY | -0.704 | 0.319 | 1.460 | -0.701 | 0.319 | 1.493 |
| | (0.055) | (0.049) | (0.307) | (0.051) | (0.049) | (0.171) |
| GBP | -0.788 | 0.467 | 2.168 | -0.758 | 0.469 | 1.659 |
| | (0.052) | (0.041) | (0.537) | (0.051) | (0.043) | (0.202) |
| EUR-pre | -1.786 | -0.106 | 2.349 | -1.729 | -0.098 | 1.653 |
| | (0.135) | (0.099) | (1.190) | (0.139) | (0.102) | (0.374) |
| JPY-pre | -1.484 | -0.080 | 2.047 | -1.456 | -0.090 | 1.629 |
| | (0.122) | (0.108) | (1.011) | (0.119) | (0.109) | (0.370) |
| GBP-pre | -1.629 | -0.066 | 2.173 | -1.606 | -0.079 | 1.642 |
| | (0.131) | (0.096) | (1.175) | (0.132) | (0.099) | (0.378) |
| EUR-c1 | -1.169 | 0.206 | 2.838 | -1.133 | 0.211 | 1.707 |
| | (0.134) | (0.115) | (2.462) | (0.137) | (0.122) | (0.504) |
| JPY-c1 | -0.691 | 0.263 | 2.308 | -0.664 | 0.258 | 1.737 |
| | (0.112) | (0.113) | (1.177) | (0.111) | (0.117) | (0.481) |
| GBP-c1 | -1.250 | 0.176 | 7.361 | -1.194 | 0.171 | 1.998 |
| | (0.125) | (0.128) | (9.711) | (0.142) | (0.142) | (0.644) |
| EUR-c2 | -0.152 | 0.213 | 1.077 | -0.194 | 0.223 | 1.393 |
| | (0.166) | (0.140) | (0.487) | (0.145) | (0.136) | (0.354) |
| JPY-c2 | -0.220 | 0.243 | 1.611 | -0.219 | 0.245 | 1.560 |
| | (0.140) | (0.117) | (0.719) | (0.135) | (0.116) | (0.400) |
| GBP-c2 | -0.288 | 0.483 | 2.238 | -0.264 | 0.483 | 1.712 |
| | (0.145) | (0.109) | (1.177) | (0.146) | (0.111) | (0.471) |
| EUR-post | -1.013 | -0.043 | 700.883 | -0.928 | -0.067 | 2.214 |
| | (0.072) | (0.090) | (262.144) | (0.089) | (0.108) | (0.639) |
| JPY-post | -0.747 | 0.062 | 1.724 | -0.736 | 0.061 | 1.580 |
| | (0.097) | (0.107) | (0.717) | (0.092) | (0.108) | (0.343) |
| GBP-post | -0.925 | 0.078 | 16.537 | -0.848 | 0.080 | 2.447 |
| | (0.075) | (0.118) | (23.336) | (0.089) | (0.131) | (0.723) |

The parameters estimates for DNIG(1) models from MLE and h-likelihood methods are consistent. The h-likelihood estimation always converges but the maximum likelihood estimation does not converge or converges with extremely high standard error when the kurtosis are close to zero.

fitted to the data. In our cases, the NIG-SV model is the restricted case of the DNIG(1) model. And the DNIG(1) model is the restricted case of the DNIG(2) model. The p-values less than 0.1 indicate that the null hypothesis are rejected at 90% confidence. It is clear that the DNIG(2) model improves the goodness-of-fit when it is compared to the NIG-SV model in most cases. It also improves the goodness-of-fit in many cases when it is compared to the DNIG(1) model.

Table 6.2: Parameter Estimates and their 95% confidence intervals for DNIG(2) models with MLE and h-likelihood methods.

| Series | Maximum likelihood | | | | H-likelihood | | | |
|---|---|---|---|---|---|---|---|---|
| | $\alpha$ | $\beta_1$ | $\beta_2$ | $\omega$ | $\alpha$ | $\beta_1$ | $\beta_2$ | $\omega$ |
| EUR | -0.850 | 0.184 | 0.362 | 2.064 | -0.825 | 0.182 | 0.362 | 1.614 |
| | (0.054) | (0.051) | (0.050) | (0.543) | (0.053) | (0.052) | (0.052) | (0.197) |
| JPY | -0.713 | 0.213 | 0.255 | 1.553 | -0.707 | 0.213 | 0.256 | 1.511 |
| | (0.055) | (0.054) | (0.055) | (0.327) | (0.051) | (0.054) | (0.055) | (0.173) |
| GBP | -0.799 | 0.299 | 0.305 | 2.906 | -0.754 | 0.301 | 0.302 | 1.761 |
| | (0.049) | (0.050) | (0.050) | (0.822) | (0.051) | (0.053) | (0.053) | (0.223) |
| EUR-pre | -1.627 | -0.101 | 0.183 | 2.702 | -1.591 | -0.109 | 0.187 | 1.715 |
| | (0.166) | (0.097) | (0.100) | (1.440) | (0.174) | (0.101) | (0.104) | (0.396) |
| JPY-pre | -1.352 | -0.109 | 0.263 | 2.195 | -1.315 | -0.105 | 0.268 | 1.641 |
| | (0.135) | (0.107) | (0.102) | (1.063) | (0.134) | (0.110) | (0.105) | (0.370) |
| GBP-pre | -1.632 | -0.046 | -0.011 | 2.179 | -1.606 | -0.047 | -0.013 | 1.618 |
| | (0.159) | (0.097) | (0.094) | (1.178) | (0.162) | (0.100) | (0.098) | (0.369) |
| EUR-c1 | -1.115 | 0.188 | 0.109 | 3.429 | -1.069 | 0.192 | 0.105 | 1.695 |
| | (0.143) | (0.114) | (0.111) | (3.486) | (0.152) | (0.124) | (0.121) | (0.496) |
| JPY-c1 | -0.694 | 0.229 | 0.122 | 2.420 | -0.657 | 0.228 | 0.118 | 1.725 |
| | (0.111) | (0.119) | (0.140) | (1.229) | (0.112) | (0.124) | (0.144) | (0.471) |
| GBP-c1 | -1.212 | 0.152 | 0.084 | 8.315 | -1.157 | 0.142 | 0.069 | 1.965 |
| | (0.140) | (0.130) | (0.134) | (11.761) | (0.160) | (0.145) | (0.148) | (0.619) |
| EUR-c2 | -0.314 | 0.032 | 0.357 | 1.368 | -0.320 | 0.033 | 0.358 | 1.460 |
| | (0.165) | (0.151) | (0.137) | (0.657) | (0.152) | (0.150) | (0.136) | (0.375) |
| JPY-c2 | -0.242 | 0.200 | 0.079 | 1.632 | -0.233 | 0.202 | 0.080 | 1.537 |
| | (0.145) | (0.131) | (0.122) | (0.730) | (0.141) | (0.131) | (0.123) | (0.390) |
| GBP-c2 | -0.395 | 0.340 | 0.249 | 2.672 | -0.351 | 0.346 | 0.241 | 1.754 |
| | (0.152) | (0.126) | (0.131) | (1.530) | (0.156) | (0.131) | (0.135) | (0.489) |
| EUR-post | -1.019 | -0.068 | 0.118 | 688.119 | -0.937 | -0.102 | 0.162 | 2.119 |
| | (0.072) | (0.092) | (0.098) | NaN | (0.090) | (0.110) | (0.118) | (0.584) |
| JPY-post | -0.743 | 0.056 | 0.040 | 1.724 | -0.729 | 0.054 | 0.040 | 1.554 |
| | (0.097) | (0.108) | (0.103) | (0.722) | (0.092) | (0.109) | (0.105) | (0.334) |
| GBP-post | -0.930 | 0.067 | 0.133 | 27.629 | -0.847 | 0.070 | 0.104 | 2.377 |
| | (0.073) | (0.117) | (0.112) | (67.153) | (0.089) | (0.133) | (0.129) | (0.685) |

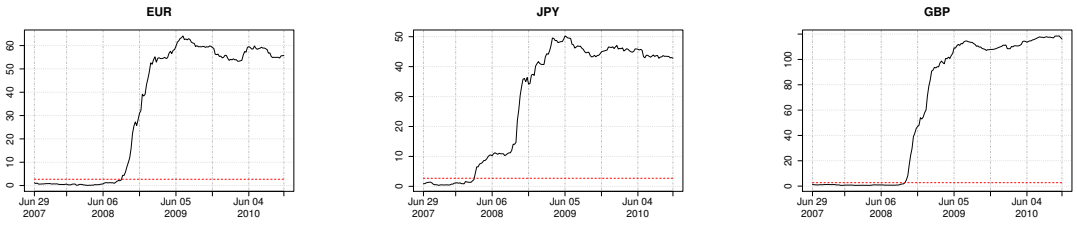
The parameters estimates for DNIG(2) models from MLE and h-likelihood methods are consistent as same as the results for DNIG(1) models. The h-likelihood is preferred because the estimation always converses in our data set.

Table 6.3: The log-likelihood ratio statistics and the corresponding p-values of the null model vs the alternative model. The series are modeled by NIG-SV, DNIG(1) and DNIG(2) models.

| | NIG-SV vs DNIG(1) | | NIG-SV vs DNIG(2) | | DNIG(1) vs DNIG(2) | |
|---|---|---|---|---|---|---|
| | $D$ | p-value | $D$ | p-value | $D_{\beta_1}$ | p-value |
| EUR | 56.34 | **6.11E-14** | 111.45 | **6.29E-25** | 55.11 | **1.14E-13** |
| JPY | 42.29 | **7.87E-11** | 69.06 | **1.01E-15** | 26.77 | **2.29E-07** |
| GBP | 117.07 | **2.77E-27** | 159.06 | **2.89E-35** | 41.98 | **9.20E-11** |
| EUR-pre | 1.16 | 0.280 | 12.64 | **0.002** | 11.48 | **0.001** |
| JPY-pre | 0.55 | 0.457 | 13.22 | **0.001** | 12.67 | **3.72E-04** |
| GBP-pre | 0.48 | 0.490 | 5.92 | **0.052** | 5.45 | **0.020** |
| EUR-c1 | 3.11 | **0.078** | 4.54 | **0.103** | 1.43 | 0.231 |
| JPY-c1 | 5.29 | **0.022** | 6.99 | **0.030** | 1.70 | 0.192 |
| GBP-c1 | 1.81 | 0.179 | 2.52 | 0.284 | 0.71 | 0.400 |
| EUR-c2 | 2.29 | 0.130 | 10.26 | **0.006** | 7.97 | **0.005** |
| JPY-c2 | 4.45 | **0.035** | 6.42 | **0.040** | 1.97 | 0.160 |
| GBP-c2 | 18.96 | **1.33E-05** | 23.26 | **8.90E-06** | 4.29 | **0.038** |
| EUR-post | 0.23 | 0.634 | 5.59 | **0.061** | 5.36 | **0.021** |
| JPY-post | 0.34 | 0.562 | 1.78 | 0.411 | 1.44 | 0.230 |
| GBP-post | 0.45 | 0.504 | 3.66 | 0.160 | 3.22 | **0.073** |

The boldfaced p-values indicate that the null hypothesis of similar likelihood are rejected at 90% confidence. The test statistic $D$ are distributed as chi-square distributions with $k$ degree of freedom, where $k$ is the difference between the numbers of models' parameters. Most cases the DNIG(2) models are better fitted to the tested series than the NIG-SV and the DNIG(1) models. The DNIG(1) models are better fitted to the series in the crises than the NIG-SV models. Overall, both DNIG(1) and DNIG(2) models are better fitted to the series than the NIG-SV models.

Figure 6.2.1: The evolution of log-likelihood ratio statistics when the NIG-SV model is tested against the DNIG(1) model.




The values of the statistic $D$ are plotted at the final day of each estimation period using all available information up to that day and reestimate every five days. The red dashed lines indicate the critical values where the statistics are different from zero at 90% confidence. The test statistics of EUR, JPY and GBP first cross the critical lines at 15 Aug 2008, 14 Mar 2008 and 3 Oct 2008 respectively.

It is remarkable that in the pre-crisis and the post-crisis periods, the null hypothesis holds for the NIG-SV model against the DNIG(1) model. This indicates that the parameter $\beta$ of the DNIG(1) model is close to zero in the regular period, but it becomes more significative in the crises. This may be used as an indicator that a crisis is occurring in the period of estimation. Figure 6.2.1 shows the evolution of the log-likelihood ratio statistics when the models are estimated using all available data up to that day. The red dashed lines indicate the critical values where the DNIG(1) model is fitted to the data better than the NIG-SV model with 90% confidence. The test statistics of EUR, JPY and GBP first cross the critical lines at 15 Aug 2008, 14 Mar 2008 and 3 Oct 2008 respectively.

### 6.2.4 Residual analysis

The empirical study of returns with high-frequency data by Andersen et al. (2001) shows that the standardized returns (returns divided by their standard deviations) are normally distributed. This results correspond to the first line in (2.5.9). Theoretically, the distribution of returns standardized by their true volatility $y_t/\sigma_t = \epsilon_t \sim N(0,1)$ are standard normal.Thus we expect that returns standardized by proper volatility estimates are normally distributed with zero mean and unit variance. Therefore the volatility model is able capture the information that causes heavy-tailed distribution in the returns into its dynamics and thus the standardized returns have no heavy tails.

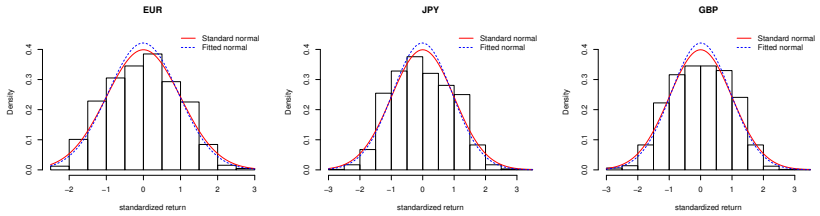
In this section we study the volatility estimated by DNIG models and consider the distributions of standardized returns. The h-likelihood method allows us to estimate the volatility by (6.2.6) that involves the observed ranges and the estimates of random effects. After we estimate the model parameters by h-likelihood for each time series, we compute the volatility estimates and the standardized returns.

Table 6.4: Summary statistics for returns standardized by volatilities estimated from NIG-SV, DNIG(1) and DNIG(2) models.

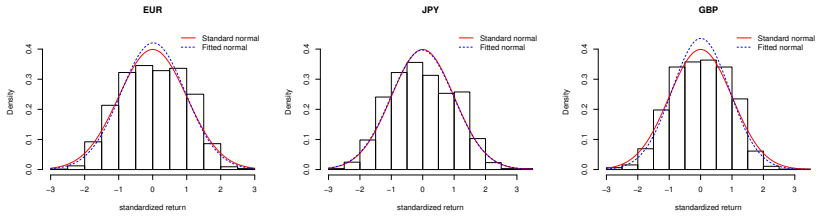
| | NIG-SV | | | | DNIG(1) | | | | DNIG(2) | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | mean | SD | skewness | kurtosis | mean | SD | skewness | kurtosis | mean | SD | skewness | kurtosis |
| EUR | -0.01 | 0.95 | 0.04 | -0.63 | 0.00 | 0.90 | 0.07 | -0.67 | 0.00 | 0.95 | 0.06 | -0.72 |
| JPY | -0.01 | 0.95 | 0.10 | -0.63 | -0.01 | 1.01 | 0.11 | -0.61 | -0.01 | 0.95 | 0.13 | -0.66 |
| GBP | 0.01 | 0.95 | -0.06 | -0.62 | 0.01 | 0.92 | -0.06 | -0.43 | 0.02 | 0.95 | -0.04 | -0.75 |

Figure 6.2.2: The histograms for returns standardized by the volatility estimates from DNIG models

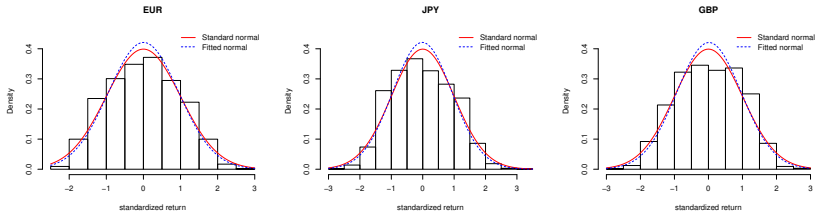
(a) The distributions of returns standardized by NIG-SV volatility estimates.




(b) The distributions of returns standardized by DNIG(1) volatility estimates.




(c) The distributions of returns standardized by DNIG(2) volatility estimates




The returns standardized by volatility estimates from DNIG models are very well fitted to the standard normal distribution. This implies that the information about the returns has been captured into the dynamics of volatility, left only the Gaussian noises in the standardized returns.

The summary statistics for standardized returns are shown in Table 6.4. The histogram for each time series is drawn along with the densities of the standard normal

distribution and the fitted normal distribution.

The standardized returns are very well described by the standard normal distribution both in the summary statistics and the histograms shows in Figure 6.2.2b and Figure 6.2.2c. The presences of skewness and kurtosis are very small in all cases, the means are almost zero and the standard deviations are only slightly different from one. It is clear that the standardized returns are approximately standard normal.

The heavy-tailed information of standardized returns are also captured by the CV-plots of absolute returns explained in Chapter 2. All the empirical cv lie between the 90% confidence intervals about $cv = 1$ provided the number of samples are greater than some thresholds. It follows that the heavy tails do not presence in the distributions of the standardized returns. All these results show that our volatility models are capable in capturing the information from the data into the dynamics of volatility and thus the residuals are normally distributed.

## 6.3 Volatility Forecasting

In Chapter 4, the ARMA(1,1) process is implemented to the log-volatility estimated from the NIG-SV model to make forecast since there is no natural dynamics in the volatility process. In contrast, the DNIG model is constructed with the dynamic in the log-volatility process. Hence the forecast is carried out in a more natural way than the NIG-SV model. The DNIG model involves the range in the dynamics, thus some properties of the range are necessary.

### 6.3.1 Moments of the range

The assumption of constant volatility over a single time period is useful since the asymptotic distribution of range is computable. Feller (1951) provides the probability density function of the range in form of infinite series

$$f(R_t|\sigma_t) = 8\sum_{k=1}^{\infty}(-1)^{k-1}\frac{k^2}{\sigma_t}\phi\left(\frac{kR_t}{\sigma_t}\right)$$

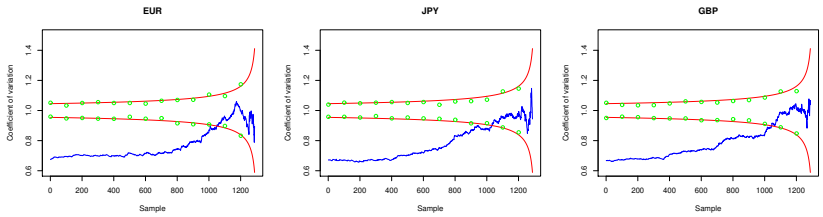
where $\phi$ is the standard normal probability density function. For practical use, the infinite summation is truncated. The cumulative distribution function of the range and a formula for calculating moments have been later provided by Parkinson (1980),

$$\begin{aligned} F(R_t|\sigma_t) &= \sum_{k=1}^{\infty}(-1)^{k-1}k\left\{\text{erfc}\left(\frac{(k+1)R_t}{\sigma_t\sqrt{2}}\right) - 2\text{erfc}\left(\frac{kR_t}{\sigma_t\sqrt{2}}\right) + \text{erfc}\left(\frac{(k-1)R_t}{\sigma_t\sqrt{2}}\right)\right\} \\ E[R_t^p|\sigma_t] &= \frac{4}{\sqrt{\pi}}\Gamma\left(\frac{p+1}{2}\right)\left(2^{p/2} - 2^{2-p/2}\right)\zeta(p-1)\sigma_t^p, \text{ for } p \geq 1 \end{aligned}$$

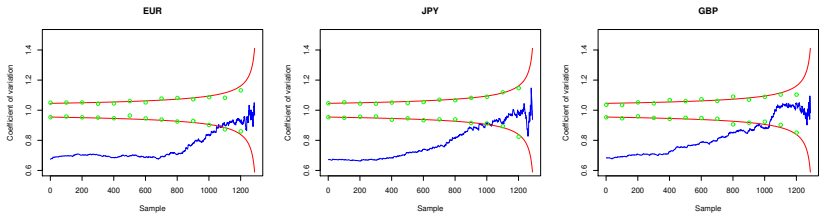
where $\text{erfc}(x) := 1 - \text{erf}(x)$, $\text{erf}(x)$ is the 'error function': $\text{erf}(x) := 2/\sqrt{\pi}\int_0^x e^{-t^2}dt$ and $\zeta(x)$ is the Riemann zeta function. Particularly, we have $E[R_t|\sigma_t] = \sqrt{8/\pi}\sigma_t$ and

Figure 6.2.3: CV-plots of returns standardized by volatilities estimated from DNIG models

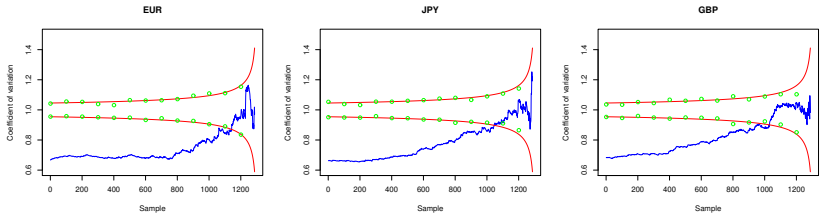
(a) NIG-SV models




(b) DNIG(1) models




(c) DNIG(2) models




The CV-plots show that the $cv$ of absolute standardized returns are below the lower limits and enter the 90% confidence intervals about $cv = 1$ at curtain thresholds. This means that the tails of the standardized returns lighter than exponential, that are not heavy tails.

$E[R_t^2|\sigma_t] = 4\log(2)\sigma_t^2$. The first four moments of log-range are given by Alizadeh et al. (2002),

$$\begin{aligned} E[\log(R_t)|\sigma_t] &= 0.43 + \log(\sigma_t) \\ \text{var}(\log(R_t)|\sigma_t) &= 0.29^2 \\ \text{skew}(\log(R_t)|\sigma_t) &= 0.17 \\ \text{kurt}(\log(R_t)|\sigma_t) &= 2.80. \end{aligned}$$

The results are consistent with Brandt & Jones (2006) , who uses quadrature and ordinary least squares to obtain the expectation of the log-range

$$E[\log(R_t)|\sigma_t] = 0.4257 + \log(\sigma_t).$$

Consequently, we also have

$$E[\log(R_t^2)|\sigma_t] = 2E[\log(R_t)|\sigma_t] = 0.8514 + \log(\sigma_t^2). \tag{6.3.1}$$

These properties are very useful for constructing DNIG forecasting models in the next subsection.

### 6.3.2 DNIG forecasting model

The one-step forecast of the logarithm of volatility for DNIG(1) model based on the information up to time $t$ is given by the conditional expectation

$$\log(\sigma_{t+1|t}^2) = E[\log(\sigma_{t+1}^2)|\mathcal{F}_t] = \alpha + \beta\log(R_t^2) + E[b_{t+1}|\mathcal{F}_t]$$

The two-step forecast involves the conditional expectation of the log-range

$$\log(\sigma_{t+2|t}^2) = \alpha + \beta E[\log(R_{t+1}^2)|\mathcal{F}_t] + E[b_{t+2}|\mathcal{F}_t].$$

Since we know that $E[\log(R_{t+1}^2)|\sigma_{t+1}] = c + \log(\sigma_{t+1}^2)$, therefore we replace $\log(\sigma_{t+1}^2)$ by the first step forecast and write $E[b_{t+k}|\mathcal{F}_t]$ in the compact notation $E_t[b_{t+k}]$,

$$\log(\sigma_{t+2|t}^2) = \alpha + \beta(c + \log(\sigma_{t+1|t}^2) + E_t[b_{t+2}]$$

Then further forecasts can be computed recursively

$$\log(\sigma_{t+k|t}^2) = \alpha + \beta(c + \log(\sigma_{t+k-1|t}^2) + E_t[b_{t+k}].$$

We assume that the random effects $b_t$ are independent, then the conditional expectation of the future random effects are identical to $E_t[b_{t+1}]$. If $\beta < 1$, the $k$-step forecast can be written as

$$\log(\sigma_{t+k|t}^2) = \frac{1-\beta^k}{1-\beta}\alpha + \frac{1-\beta^{k-1}}{1-\beta}\beta c + \beta^k\log(R_t^2) + \frac{1-\beta^k}{1-\beta}E_t[b_{t+1}]. \tag{6.3.2}$$

The conditional expectation of the random effect $E_t[b_{t+1}]$ can be estimated by the mean of the random effects obtained from h-likelihood. Let the forecast horizon $k$ in (6.3.2) run to infinity, the long-run forecast converges to

$$\lim_{k\to\infty} \log(\sigma^2_{t+k|t}) = \frac{\alpha + \beta c + E_t[b_{t+1}]}{1-\beta}.$$

The DNIG(2) forecasting model can be obtained in the same manner. The first and second step forecasts are

$$\log(\sigma^2_{t+1|t}) = \alpha + \beta_1 \log(R^2_t) + \beta_2 \log(R^2_{t-1}) + E_t[b_{t+1}]$$

and

$$\begin{aligned}\log(\sigma^2_{t+2|t}) &= \alpha + \beta_1 E[\log(R^2_{t+1})|\mathcal{F}_t] + \beta_2 \log(R^2_t) + E_t[b_{t+2}] \\ &= \alpha + \beta_1 \left(c + \log(\sigma^2_{t+1|t})\right) + \beta_2 \log(R^2_t) + E_t[b_{t+2}].\end{aligned}$$

Then the further forecasts can be achieved recursively by

$$\log(\sigma^2_{t+k|t}) = \alpha + \beta_1 \left(c + \log(\sigma^2_{t+k-1|t})\right) + \beta_2 \left(c + \log(\sigma^2_{t+k-2|t})\right) + E_t[b_{t+k}]. \quad (6.3.3)$$

This method is also applied for DNIG models with higher orders.

### 6.3.3 Implementation and empirical results

#### 6.3.3.1 Implementation

In the DNIG(1) model, the forecasts for log volatility in (6.3.2) is realized by taking the conditional expectation of the random effect as the average of past $k$ random effect estimates. That is

$$\hat{E}_t[b_{t+k}] = \frac{1}{k}\sum_{i=0}^{k-1} \hat{b}_{t-i}.$$

with the estimated parameters $\hat{\Theta}$ and the random effect estimates $\{\hat{b}_t\}$ computed by (6.2.5). Remark that in each recursion, we apply the same estimate $\hat{E}_t[b_{t+k}]$ to all future forecasts $\log(\sigma^2_{t+1|t}), ..., \log(\sigma^2_{t+k|t})$. Then the $k$-step forecast for the volatility formulated at time $t$ is

$$f^{\text{DNIG(1)}}_{t+k|t} = \exp\left[\log(\hat{\sigma}^2_{t+k|t})\right]. \quad (6.3.4)$$

Preliminary implementations show that the DNIG(1) forecasting model in (6.3.4) generally produce volatility forecasts substantially lower the corresponding squared returns which are the volatility proxies as it happens in the case of NIG-SV model in Chapter 4. Hence the simple linear regression between the volatility estimates from the DNIG(1) model and the squared returns is applied to improve the forecast performance. Suppose

Table 6.5: Average QL loss of cumulative volatility forecasts

| | Average QL loss | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | RW | GARCH | NIG-SV | NIG-SV* | NIG-SV′ | DNIG(1) | DNIG(1)* | DNIG(2) | DNIG(2)* |
| Long-run average | | | | | | | | | |
| EUR | 0.127 | **0.120** | 0.639 | 0.259 | 0.307 | 0.609 | 0.184 | 0.628 | 0.177 |
| JPY | 0.300 | 0.247 | 0.546 | 0.236 | 0.310 | 0.603 | 0.200 | 0.670 | **0.197** |
| GBP | **0.157** | **0.157** | 0.671 | 0.324 | 0.363 | 0.671 | 0.225 | 0.662 | 0.214 |
| Crisis 1 average | | | | | | | | | |
| EUR | 0.122 | 0.130 | 0.320 | 0.091 | 0.115 | 0.289 | 0.079 | 0.302 | **0.078** |
| JPY | 0.583 | 0.474 | 0.815 | **0.254** | 0.408 | 0.799 | 0.271 | 0.884 | 0.282 |
| GBP | 0.172 | 0.110 | 0.182 | 0.061 | 0.069 | 0.186 | 0.060 | 0.191 | **0.059** |
| Crisis 2 average | | | | | | | | | |
| EUR | 0.211 | **0.201** | 1.762 | 0.423 | 0.881 | 1.601 | 0.358 | 1.600 | 0.413 |
| JPY | 0.246 | 0.257 | 0.954 | **0.196** | 0.568 | 1.081 | **0.194** | 1.163 | 0.235 |
| GBP | **0.227** | 0.328 | 2.008 | 0.524 | 1.080 | 1.946 | 0.461 | 1.885 | 0.483 |
| Post-Crisis average | | | | | | | | | |
| EUR | 0.074 | 0.060 | 0.087 | 0.252 | **0.044** | 0.147 | 0.133 | 0.182 | 0.080 |
| JPY | 0.162 | 0.101 | 0.111 | 0.252 | **0.079** | 0.166 | 0.161 | 0.211 | 0.120 |
| GBP | 0.101 | 0.072 | 0.082 | 0.352 | **0.066** | 0.120 | 0.169 | 0.137 | 0.129 |

The average QL losses over different periods for all forecasting models including the models from Chapter 4. The target variable is the sum of squared return that is the proxy for cumulative volatility. The forecast horizon is $k = 22$, that correspond to one-month forecast. The parameters are reestimated every five days.

that in-sample volatility estimates $\{\hat{\sigma}_1^2, .., \hat{\sigma}_n^2\}$ obtained from DNIG(1) model by (6.2.6) are fitted to the corresponding squared returns as

$$y_t^2 = a\hat{\sigma}_t^2, \ t = 1, ..., n,$$

then the DNIG(1)* forecasting model is given by

$$f_{t+k|t}^{\text{DNIG(1)*}} = a f_{t+k|t}^{\text{DNIG(1)}}.$$

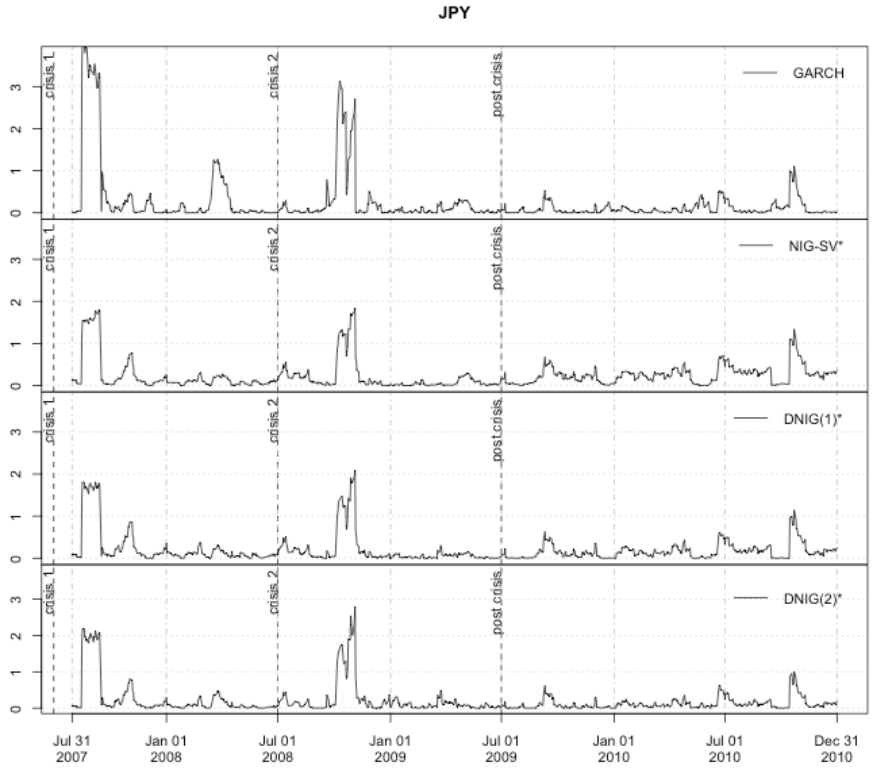
The DNIG(2) and DNIG(2)* are also implemented in the same manner.

#### 6.3.3.2 Results

The QL average losses over different periods from the DNIG forecasting models including the results from the forecasting models in Chapter 4 are shown in Table6.5. The long-run average QL losses show that GARCH, DNIG(2)* and GARCH forecasting models are favorites for EUR, JPY and GBP respectively. Among the DNIG models, the DNIG models with higher orders have less average QL losses. The DNIG(2)* and NIG-SV* are also favorites in the Crisis 1 period but in the Crisis 2 there is no clear winner. In the Post-Crisis, the NIG-SV′ has impressive performances. It is notable that the DNIG forecasting models are favorites for JPY in all cases.

Figure 6.3.1 show the plots of QL loss series for JPY forecasted with GARCH, NIG-SV*, DNIG(1)* and DNIG(2)*. The less value of the QL loss, the more the accuracy of the forecasting model. Graphically, the DNIG(2)* has relatively low QL losses than

Figure 6.3.1: The plots of QL loss series for JPY from GARCH and DNIG models




The QL loss series are taken from different forecasting models with the same $xy$ -scale. The less the QL loss, the more the accuracy. The graphs show that the DNIG(2)* is very accurate relative to the other models. The accuracy can be measured by the average loss over a particular period, however, this plots show that a forecasting model can be favorite if a proper period is taken.

NIG-SV* and DNIG(1)* and it is comparable to GARCH. The forecast accuracy that is taken from the average QL loss depends on the periods of measurement. GARCH model exhibits extremely high losses in the crises, that the peaks of the losses from DNIG models are considerably smaller. In the Post-Crisis, GARCH and DNIG(2)* have almost similar losses.

## 6.4 Future Research

The DNIG model has great potential to develop in several ways. Here are some ideas for future research. First, multivariate DNIG model can be obtained in the same manner of multivariate NIG-SV model shown in Chapter 4 that the HGLM method of estimation is readily available. The multivariate model will help us understand the co-movement of several asset returns simultaneously. Second, other exogenous variables such as trading

volume, open and close prices might be incorporated into the dynamics of volatility in addition to the range. Moreover, high-frequency realized volatility is also promising, thanks to Christian Brownlees for this suggestion. Relevant variables added into the model might result in more accurate estimator of volatility as we have seen in Chapter 5. Finally, other GH models presented in Chapter 4 such as variance gamma model can be estimated by HGLM method as same as NIG-SV model. Hence it is possible to incorporate the dynamics into GH models for more general results. These ideas have not been comprehensively investigated but they are very promising for the future research.

## Summary

The new stochastic volatility model has been proposed. It is constructed by incorporating the ideas from the NIG-SV model, the standard SV model and the range-based volatility estimators. Results in the stochastic volatility model with dynamics in log volatility. The new proposed DNIG model can be easily extended to higher order setting. It can be estimated simply by maximum likelihood or h-likelihood. Moreover, the h-likelihood method also provides the volatility estimates. It is also remarkable from the estimates of the parameter $\beta$ that the ranges become more significative in the crises.

In most cases the DNIG(2) models fit better to the series than the NIG-SV and the DNIG(1) models. This tells us that the AR(2) information, that has not been used by other researchers, is relevant. The residual analysis shows that the returns standardized by DNIG volatility estimates are approximately standard normal. That heavy-tailed information in the returns is captured into the dynamics of volatility perfectly. The DNIG models are also employed to forecast the volatility at one-month horizon. The DNIG forecasting models can be implemented in a simple way and the results are better than GARCH models in many cases.

# Summary and Conclusions of the Thesis

This thesis has presented the insight into volatility forecasting covering from the basic ideas, required theory, simulation study, practical implementation, to the new proposed models. We have analyzed the data with alternative tools that can make us aware of different things that have been observed with other common methods. We aim to bring new light, rather than replace models settled. The ideas have been developed from the preliminary theory in Chapter 2. The properties of return have been investigated based on existing distribution in Chapter 3. We conclude that the NIG distribution is capable of describing the marginal distribution of return during the crisis and it can be estimated plainly with either the method of moments or the maximum likelihood estimation. The results in Chapter 3 show that the NIG distribution has attractive potential to model the financial data. It grants an alternative way of modeling financial data in such a way that GARCH does not supply (the marginal distribution).

Chapter 4 provides the pragmatic guide to volatility forecasting including all necessary information. The ideas from Chapter 3 have been developed to introduce the NIG-SV model, a stochastic volatility model proposed by Barndorff-Nielsen (1997). We also introduce the HGLM method for estimation that is comparable to the maximum likelihood method but the complicated integration is avoided. The empirical results show that the HGLM method is as accurate as the maximum likelihood method. Moreover, the key role of the h-likelihood in the HGLM method, it provides us the estimates of random effects that are latent in the market. We consequently apply the random effects to estimate and forecast volatility. The new forecasting models in this chapter overcome the standard forecasting models in some occasions.

Chapter 5 is investigated separately from previous chapters. Rather than the return, exogenous variables such as open, close, high and low prices play the most important role in this chapter. Several range-based volatility estimators have been introduced and we also correct the bias generated from discretization. We test by simulations whether these estimators are relevant in different scenarios when the theoretical conditions do not hold perfectly. It turns out that the Garman-Klass estimator perform impressively in many occasions. Other estimator also provides proper estimates for volatility when the conditions are close to their theoretical settings. We conclude that the range contains substantial information and it is relevant to incorporate into a model.

In the end, all information obtained from Chapter 2 to Chapter 5 are summarized into the new model in Chapter 6. The DNIG model is a stochastic volatility model that is based the NIG-SV model with the dynamics driven by the range. The DNIG model can be easily extended to higher orders, that has never been done in the standard SV model and it can be estimated by the HGLM method. In most cases the DNIG(2) models fit better to the series than the NIG-SV and the DNIG(1) models. The relevant information of AR(2) shown in this thesis has never been discovered by other researchers. It is also remarked that the parameter $\beta$ that is the coefficient of the range might be an indication of the crisis. Estimate the DNIG model with the HGLM method yields the random effects estimates and consequently the estimates for volatility which are latent information. It has been tested that the returns standardized by volatilities estimated from DNIG models do not exhibit heavy tails. This result shows that DNIG models are capable in capturing the relevant information from the returns. The DNIG model with the HGLM method also allow us to forecast volatility with ease. The DNIG forecasting models have been tested with the real data in comparisons with the standard models and they perform nicely. In many cases, the results are better than GARCH(1,1). Last but not least, the DNIG model can be developed in many ways such as multivariate modeling, exogenous variables incorporating and generalization to GH models. The further research is promising.

# Nomenclature

| | |
|---|---|
| $\bar{F}(\cdot)$ | survival function, reliability function |
| $\mathcal{F}_t$ | information set available at time $t$ |
| $\gamma_\tau$ | autocovariance at lag $\tau$ |
| $\mu_n$ | $n^{th}$ central moment |
| $\rho_\tau$ | autocorrelation at lag $\tau$ |
| $\sigma^2$ | (unconditional) variance |
| $\text{cor}(X, Y)$ | correlation between $X$ and $Y$ |
| $\text{cov}(X, Y)$ | covariance between $X$ and $Y$ |
| $\text{kurt}(\cdot)$ | kurtosis |
| $\text{skew}(\cdot)$ | skewness |
| $\text{var}(\cdot)$ | variance |
| $\{X_t\}$ | stochastic process |
| $D$ | distribution |
| $E[\cdot]$ | expectation |
| $F$ | cumulative distribution function |
| $f$ | probability density function or a function in general |
| $F_u$ | distribution function of threshold excedances |
| $f_X$ | density function of rancom variable $X$ |
| $f_{t+k\|t}$ | volatility forecast |

$G(\Theta)$ gradient of $\log L(\Theta)$

$GH(\lambda, \phi, \omega)$ zero-mean symmetric generalized hyperbolic distribution

$H(\Theta)$ hessian of $\log L(\Theta)$

$I(\Theta)$ information matrix of $\log L(\Theta)$

$k$ forecast horizon

$k$ sample kurtosis, forecast horizon

$P(\cdot)$ probability

$P_t$ price of an asset at time $t$

$r_t^*$ simple net return in Chapter 2, logarithm of price range in later Chapters

$r_{t,k}^*$ simple net return over most recent $k$ trading period

$R^2$ coefficient of determination

$s^2$ sample variance

$w$ sample skewness

$WN(0, \sigma^2)$ white noise (uncorrelated random variable with zero mean and finite variance).

ARCH autoregressive conditional heteroskedastic

ARFIMA autoregressive fractionally integrated moving average

CV residual coefficient of variation

DNIG dynamic NIG-SV

EGARCH exponetial GARCH

EWMA exponentially weighted moving average

GARCH generalized autoregressive conditional heteroskedastic

GBM geometric Brownian motion

GH generalized hyperbolic

GIG generalized inverse Gaussian

GMM generalized method of moments

GPD generalized Pareto distribution

HGLM hierachical generalized linear model

i.i.d. independent and identically distributed

IG inverse Gaussian

MLE maximum likelihood estimation

MMSE minimum mean square estimator

MoM method of moments

MSE mean squared error

NIG normal inverse Gaussian

NIG-SV normal inverse Gaussian stochastic volatility

pdf probability density function

QL quasi likelihood

QML quasi-maximum likelihood

RWH1 random walk hypothesis 1 : Gaussian random walk

RWH2 random walk hypothesis 2: uncorrelated and stationary increments

SV stochastic volatility

# Index


**A**
APARCH, 21, 45
autocorrelation, 8, 13
    sample, 13
autocovariance, 8
Autoregressive conditional heteroskedastic, 20

**B**
Black-Scholes formula, 5
Black-Scholes model, 24
Brownian motion, 25

**C**
central moment, 7
characteristic function, 26
coefficient of variation, 16
conditional density, 7
conditional expectation, 8
continuous, 7
correlation, 8
covariance, 8
cumulative distribution function, 7
CV-plot, 16

**D**
dependent, 8
drift, 9
dynamic NIG stochastic volatility, 78

**E**
efficient-market hypothesis, 5
EGARCH, 21, 45
EWMA, 24
expectation, 7
exponentially weighted moving average, 24

**G**
gamma
    distribution, 27
GARCH, 2, 20
GBM, 25
generalized hyperbolic
    distribution, 22, 27
generalized inverse Gaussian, 47
    distribution, 22, 27
generalized method of moments, 22
generalized Pareto distribution, 16
geometric Brownian motion, 25
GH Lévy
    process, 48
GIG, 22
GMM, 22
gradient, 53

**H**
heavy tails, 12
Hessian, 53
HGLM method, 49
hierarchical-likelihood, 22
hyperbolic
    distribution, 27
hyperbolic Lévy
    process, 48

**I**
independent, 8
independent and identically distributed, 9
Infinitely divisible, 26
infinitely divisible, 22, 26

information set, 8
inverse gamma
    distribution, 22
inverse Gaussian
    distribution, 22, 27

**J**
joint density, 7
joint distribution, 7

**K**
kurtosis, 7

**L**
lag, 13
Lévy process, 26
leptokurtic, 11

**M**
martingale hypothesis, 11
MCMC, 22
mean, 7, 11
method of scoring, 55
minimum mean squared error, 63
mixing density, 48
mixing parameter, 48

**N**
negative definite, 55
Newton-Raphson algorithm, 55
NGARCH, 21, 45
NIG Lévy
    process, 48
NIG-SV, 49
normal
    distribution, 27
normal inverse Gaussian
    distribution, 22, 27

**O**
opening jump, 65, 66
optimization, 53
outliers, 12

**P**
Poisson
    distribution, 27
price, 8
    log, 9
probability density function, 7
probability space, 7

**Q**
QML, 22
quasi likelihood, 42
quasi-maximum likelihood, 22

**R**
random variable, 7
random walk, 5
    Gaussian, 9
random walk hypothesis, 5
range, 2
reliability function, 15
return
    continuously compounded, 9
    log, 9
    simple gross, 8
    simple net, 8
risk, 5
RiskMetrics, 24
RWH1, 9, 25
RWH2, 10

**S**
sample
    kurtosis, 11
    skewness, 11
    standard deviation, 11
    variance, 11
scale mixture of normal, 47
scaled-t
    distribution, 27
score, 53
serially uncorrelated, 11
skewness, 7
standard deviation, 7
standard SV model, 22
stationary, 8, 13
stochastic process, 8

stochastic volatility, 2, 20, 22
stylized facts, 11
stylized features, 11
survivor function, 15
SV, 2, 22

**T**
tail distribution, 15
TARCH, 45
TGARCH, 21
The empirical CV of of the conditional exceedance, 16
threshold excedances, 16
time series, 8

**V**
variance, 7
variance gamma
    distribution, 27
    process, 48
variance rate, 23
volatility, 1, 18
    annualized, 23
    historical, 23
    realized, 23
volatility clustering, 9
volatility proxy, 41

**W**
white noise, 57
Wiener process, 25

# Bibliography


Alizadeh, S., Brandt, M. W., & Diebold, F. X. (2002). Range-based estimation of stochastic volatility models. *The Journal of Finance*, 57(3), 1047–1091.

Andersen, T. G., Bollerslev, T., Diebold, F. X., & Ebens, H. (2001). The distribution of realized stock return volatility. *Journal of Financial Economics*, 61, 43–76.

Andersen, T. G., Bollerslev, T., Diebold, F. X., & Labys, P. (2003). Modeling and forecasting realized volatility. *Econometrica*, 71, 579–625.

Bachelier, L. (1900). Théorie de la spéculation. *Annales Scientifiques de l'École Normale Supérieure*, 3(17), 21–86.

Balkema, A. A. & de Haan, L. (1974). Residual Life Time at Great Age. *The Annals of Probability*, 2(5), 792–804.

Barndorff-Nielsen, O. E. (1997). Normal Inverse Gaussian Distributions and Stochastic Volatility Modelling. *Scandinavian Journal of statistics*, 24(1), 1–13.

Barndorff-Nielsen, O. E. & Halgreen, C. (1977). Infinite divisibility of the hyperbolic and generalized inverse Gaussian distributions. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 38(4), 309–311.

Barndorff-Nielsen, O. E. & Shephard, N. (2001). Modelling by Lévy processes for financial econometrics. In S. I. Resnick, O. E. Barndorff-Nielsen, & T. Mikosch (Eds.), *Lévy processes : Theory and applications* (pp. 283–318). Boston: Birkhäuser.

Barndorff-Nielsen, O. E. & Shephard, N. (2002). Econometric analysis of realized volatility and its use in estimating stochastic volatility models. *Journal of the Royal Statistical Society. Series B: Statistical Methodology*, 64, 253–280.

Barndorff-Nielsen, O. E. & Shephard, N. (2004). Econometric analysis of realized covariation: High frequency based covariance, regression, and correlation in financial economics. *Econometrica*, 72, 885–925.

Barndorff-Nielsen, O. E. & Shephard, N. (2012). Basics of Lévy processes. Draft Chapter from a book by the authors on Lévy Driven Volatility Models.

Bauwens, L., Laurent, S., & Rombouts, J. V. K. (2006). Multivariate GARCH models: A survey. *Journal of Applied Econometrics*, 21(1), 79–109.

Beckers, S. (1983). Variances of Security Price Returns Based on High, Low, and Closing Prices. *The Journal of Business*, 56(1), 97.

Berndt, E. K., Hall, B. H., Hall, R. E., & Hausman, J. a. (1974). Estimation and Inference in Nonlinear Structural Models. *Annals of Economic and Social Measurement*, 3(4), 653–665.

Bibby, B. M. & Sørensen, M. (2003). Hyperbolic Processes in Finance. In S. T. Rachev (Ed.), *Handbook of Heavy Tailed Distributions in Finance* chapter 6, (pp. 211–248). Amsterdam: Elsevier B.V.

Black, F. & Scholes, M. (1973). The Pricing of Options and Corporate Liabilities. *Journal of Political Economy*, 81(3), 637.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of econometrics*, 31(3), 307–327.

Bollerslev, T., Chou, R. Y.-T., & Kroner, K. (1992). ARCH Modelling in Finance: A Review of the Theory and Empirical Evidence. *Journal of Econometrics*, 52, 5–59.

Boothe, P. & Glassman, D. (1987). The statistical distribution of exchange rates: Empirical evidence and economic implications. *Journal of International Economics*, 22, 297–319.

Box, G. E. P. & Pierce, D. A. (1970). Distribution of Residual Autocorrelations in Autoregressive-Integrated Moving Average Time Series Models. *Journal of the American Statistical Association*, 65, 1509–1526.

Brandt, M. W. & Jones, C. S. (2006). Volatility Forecasting With Range-Based EGARCH Models. *Journal of Business and Economic Statistics*, 24(4), 470–486.

Brownlees, C., Engle, R. F., & Kelly, B. (2012). A practical guide to volatility forecasting through calm and storm. *Journal of Risk*, 14(2), 3–22.

Brunetti, C. & Lildholdt, P. M. (2002). Return-based and range-based (co) variance estimation - with an application to foreign exchange markets. *Available at SSRN 296875*.

Campbell, J. Y., Lo, A. W., & MacKinlay, A. C. (1997). *The Econometrics of Financial Markets*. Princeton N.J.: Princeton University Press.

Chen, C. W. S., Gerlach, R., & Lin, E. M. H. (2008). Volatility forecasting using threshold heteroskedastic models of the intra-day range. *Computational Statistics & Data Analysis*, 52(6), 2990–3010.

Chou, R. Y.-T. (2005). Forecasting financial volatilities with extreme values: the conditional autoregressive range (CARR) model. *Journal of Money, Credit and Banking*, 37(3), 561–582.

Clark, P. K. (1973). A Subordinated Stochastic Process Model with Finite Variance for Speculative Prices. *Econometrica*, 41(1), 135–155.

Cont, R. & Tankov, P. (2004). *Financial Modelling With Jump Processes*. Boca Raton Fla.: Chapman & Hall/CRC.

del Castillo, J., Daoudi, J., & Lockhart, R. (2014). Methods to Distinguish Between Polynomial and Exponential Tails. *Scandinavian Journal of Statistics*, 41(2), 382–393.

del Castillo, J. & Lee, Y. (2008). GLM-methods for volatility models. *Statistical Modelling*, 8(3), 263–283.

Diebold, F. X. & Mariano, R. S. (1995). Comparing predictive accuracy. *Jurnal of Business & Economic Statistics*, 13(3), 253–263.

Ding, Z., Granger, C. W. J., & Engle, R. F. (1993). A long memory property of stock market returns and a new model. *Journal of Empirical Finance*, 1, 83–106.

Eberlein, E. & Keller, U. (1995). Hyperbolic distributions in finance. *Bernoulli*, 1(3), 281–299.

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

Engle, R. F. (1990). Stock Volatility and the Crash of '87: Discussion. *The Review of Financial Studies*, 3(1), 103–106.

Fama, E. F. (1965). The Behavior of Stock-Market Prices. *The Journal of Business*, 38(1), 34–105.

Feller, W. (1951). The Asymptotic Distribution of the Range of Sums of Independent Random Variables. *The Annals of Mathematical Statistics*, 22(3), 427–432.

Foss, S., Korshunov, D., & Zachary, S. (2013). *An Introduction to Heavy-Tailed and Subexponential Distributions*. New York: Springer, 2nd edition.

Gallant, A. R., Hsu, C., & Tauchen, G. (1999). Using Daily Range Data to Calibrate Volatility Diffusions and Extract the Forward Integrated Variance. *Review of Economics and Statistics*, 81(4), 617–631.

Garman, M. B. & Klass, M. J. (1980). On the Estimation of Security Price Volatilities from Historical Data. *The Journal of Business*, 53(1), 67–78.

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

Granger, C. W. J. & Newbold, P. (1976). Forecasting Transformed Series. *Journal of the Royal Statistical Society. Series B (Methodological)*, 38(2), 189–203.

Harrison, P. (1998). Similarities in the Distribution of Stock Market Price Changes between the Eighteenth and Twentieth Centuries. *The Journal of Business*, 71(1), 55–79.

Harvey, A. (1981). *Time series models*. New York: Wiley.

Harvey, A., Ruiz, E., & Shephard, N. (1994). Multivariate stochastic variance models. *The Review of Economic Studies*, 61(2), 247–264.

Hsieh, D. A. (1988). The statistical properties of daily foreign exchange rates: 1974–1983. *Journal of International Economics*, 24(1-2), 129–145.

Karatzas, I. & Shreve, S. E. (2005). *Brownian motion and stochastic calculus*. New York: Springer, 2nd edition.

Kijima, M. (2003). *Stochastic processes with applications to finance*. Boca Raton Fla.: Chapman & Hall/CRC.

Lee, W., Lim, J., Lee, Y., & del Castillo, J. (2011). The hierarchical-likelihood approach to autoregressive stochastic volatility models. *Computational Statistics & Data Analysis*, 55(1), 248–260.

Lee, Y. & Nelder, J. (1996). Hierarchical Generalized Linear Models. *Journal of the Royal Statistical Society. Series B (Methodological)*, 58(4), 619–678.

Lee, Y. & Nelder, J. A. (2001). Hierarchical generalised linear models: A synthesis of generalised linear models, random-effect models and structured dispersions. *Biometrika*, 88, 987–1006.

Lee, Y. & Nelder, J. A. J. A. (2006). Double hierarchical generalized linear models. *Journal of the Royal Statistical Society.*, 55(2), 139–185.

Lin, E. M. H., Chen, C. W. S., & Gerlach, R. (2012). Forecasting volatility with asymmetric smooth transition dynamic range models. *International Journal of Forecasting*, 28(2), 384–399.

Lo, A. W. & MacKinlay, C. a. (1988). Stock Market Prices Do Not Follow Random Walks: Evidence from a Simple Specification Test. *The Review of Financial Studies*, 1(1), 41–66.

Lopez, J. A. (2001). Evaluating the predictive accuracy of volatility models. *Journal of Forecasting*, 20(2), 87–109.

Lütkepohl, H. & Xu, F. (2010). The role of the log transformation in forecasting economic variables. *Empirical Economics*, 42(3), 619–638.

Madan, D. B. & Seneta, E. (1990). The Variance Gamma (V.G.) Model for Share Market Returns. *The Journal of Business*, 63(4), 511–524.

Malkiel, B. G. (1973). *A random walk down Wall Street*. New York: Norton.

Mandelbrot, B. B. (1963). The variation of certain speculative prices. *The Journal of Business*, 36(4), 394–419.

Markowitz, H. (1952). Portfolio selection. *The Journal Of Finance*, 7(1), 77–91.

Merton, R. C. (1973). Theory of rational option pricing. *The Bell Journal of Economics and Management*, 4(1), 141–183.

Mitchell, H., Brown, R., & Easton, S. (2002). Old volatility - ARCH effects in 19th century consol data. *Applied Financial Economics*, 12(4), 301–307.

Molnár, P. (2012). Properties of range-based volatility estimators. *International Review of Financial Analysis*, 23, 20–29.

Nelson, D. B. (1991). Conditional Heteroskedasticity in Asset Returns: A New Approach. *Econometrica*, 59(2), 347–370.

Parkinson, M. (1980). The Extreme Value Method for Estimating the Variance of the Rate of Return. *Journal of Business*, 53(1), 61–65.

Patton, A. J. (2011). Volatility forecast comparison using imperfect volatility proxies. *Journal of Econometrics*, 160(1), 246–256.

Pfeiffer, P. (1989). *Probability for applications*. New York: Springer-Verlag.

Pickands, J. (1975). Statistical Inference Using Extreme Order Statistics. *the Annals of Statistics*, 3(1), 119–131.

Pong, S., Shackleton, M. B., Taylor, S. J., & Xu, X. (2004). Forecasting currency volatility: A comparison of implied volatilities and AR(FI)MA models. *Journal of Banking & Finance*, 28(10), 2541–2563.

Poon, S.-H. (2005). *A practical guide to forecasting financial market volatility*. Chichester: Wiley.

Poon, S.-H. & Granger, C. W. J. (2003). Forecasting Volatility in Financial Markets : A Review. *Journal of Economic Literature*, 41(2), 478–539.

Poon, S.-H. & Granger, C. W. J. (2005). Practical issues in forecasting volatility. *Financial Analysts Journal*, 61(1), 45–56.

Praetz, P. D. (1972). The Distribution of Share Price Changes. *The Journal of Business*, 45(1), 49.

Raible, S. (2000). *Lévy Processes in Finance : Theory , Numerics , and Empirical Facts*. PhD thesis, Freiburg University, Freiburg.

Rogers, L. C. G. & Satchell, S. E. (1991). Estimating variance from high, low and closing prices. *The Annals of Applied Probability*, 1(4), 504–512.

Rosenberg, B. (1970). The distribution of the mid-range–A comment. *Econometrica*, 38, 176–177.

Rydberg, T. H. (1997). The normal inverse gaussian lévy process: simulation and approximation. *Communications in Statistics. Stochastic Models*, 13(4), 887–910.

Samuelson, P. A. (1965). Proof that properly anticipated prices fluctuate randomly. *Industrial management review*, 6, 41–49.

Sato, K.-I. (2014). Stochastic integrals with respect to Lévy processes and infinitely divisible distributions. *Sugaku Expositions*, 27(1), 161–181.

Shu, J. & Zhang, J. E. (2006). Testing range estimators of historical volatility. *Journal of Futures Markets*, 26(3), 297–313.

Silverman, B. (1986). *Density estimation for statistics and data analysis*. New York: Chapman and Hall.

Tauchen, G. & Pitts, M. (1983). The Price Variability-Volume Relationship on Speculative Markets. *Econometrica: Journal of the Econometric Society*, 51(2), 485–505.

Taylor, S. J. (1982). Financial returns modelled by the product of two stochastic processes, a study of daily sugar prices 1961-79. In O. Anderson (Ed.), *Time series analysis : theory and practice 1* (pp. 203–226). Amsterdam: North-Holland.

Taylor, S. J. (1986). *Modelling financial time series*. Chichester: Wiley.

Taylor, S. J. (1987). Forecasting the volatility of currency exchange rates. *International Journal of Forecasting*, 3(June 1986), 159–170.

Taylor, S. J. (2005). *Asset price dynamics, volatility, and prediction*. Princeton (N.J.); Oxford: Princeton University Press.

Teräsvirta, T. (2009). An introduction to univariate GARCH models. In T. Mikosch, J.-P. Kreiß, R. A. Davis, & T. G. Andersen (Eds.), *Handbook of Financial Time Series* (pp. 17–42). Berlin: Springer.

Yang, D. & Zhang, Q. (2000). Drift Independent Volatility Estimation Based on High, Low, Open, and Close Prices. *The Journal of Business*, 73(3), 477–492.

# Appendix A

# HGLM Method for Multivariate NIG-SV Model

For multivariate NIG-SV model given by (4.3.10), suppose that $\{\mathbf{y}_t\}_{t=1}^n$ where $\mathbf{y}_t = (y_{1,t}, y_{2,t}, ..., y_{d,t})'$ is the sample of $d$ indices of size $n$. The parameter $\Theta = (\phi, \omega)$ with $\phi = (\phi_1, \phi_2, .., \phi_d)$ can be estimated by maximizing the log likelihood

$$\begin{aligned} l(\phi,\omega) &= n\left\{\log(2) - \frac{d+1}{2}\log(2\pi) - \frac{1}{2}\sum_{i=1}^{d}\log(\phi_i) + \omega + \frac{d+3}{4}\log(\omega) - \frac{1}{2}\log(|\Omega|)\right\} \\ &\quad -\frac{d+1}{4}\sum_{t=1}^{n}\log\left(\mathbf{y}_t'\Sigma^{-1}\mathbf{y}_t + \omega\right) + \sum_{t=1}^{n}\log\left[K_{\frac{d+1}{2}}\left(\sqrt{\omega\mathbf{y}_t'\Sigma^{-1}\mathbf{y}_t + \omega^2}\right)\right] \end{aligned}$$

Alternatively, the h-likelihood is expressed as

$$h = \sum_{t=1}^{n}\left\{\log f(\mathbf{y}_t|b_t) + \log f_\Theta(b_t)\right\}$$

where the explicit expressions for $\log f(\mathbf{y}_t|b_t)$ and $\log f_\Theta(b_t)$ are

$$\begin{aligned} \log f(\mathbf{y}_t|b_t) &= -\frac{1}{2}\left\{d\log(2\pi) + \log|\Sigma| + db_t + \mathbf{y}_t'\Sigma^{-1}\mathbf{y_t}e^{-b_t}\right\} \\ \log f_\Theta(b_t) &= -\frac{1}{2}\left\{\log(2\pi) - \log(\omega) + 3\log(b_t) - 2\omega + \omega(b_t^{-1} + b_t) - 2b_t\right\}. \end{aligned}$$

The random effects can be estimated by solving $\partial h/\partial b_t = 0$, that we have

$$\hat{b}_t = \log\left(\frac{2\omega_t - (d+1)}{2\omega}\right)$$

where $\omega_t = \frac{1}{2}\sqrt{(d+1)^2 + 4(\omega \mathbf{y}_t' \Sigma^{-1} \mathbf{y}_t + \omega^2)}$ . Consequently, the adjusted profile h-likelihood is expressed as

$$\begin{aligned} p_b(\phi, \omega) =& n \left\{ -\frac{d}{2}\log(2\pi) - \frac{1}{2}\sum_{i=1}^{d}\log(\phi_i) - \frac{1}{2}\log(|\Omega|) + \omega + \frac{d+2}{2}\log(\omega) \right\} \\ & - \frac{1}{2}\sum_{t=1}^{n}\log\left[\omega_t \left(\omega_t - \frac{d+1}{2}\right)^{d+1}\right] - \sum_{t=1}^{n}\omega_t \end{aligned}$$

and the second-order approximation is

$$S_b(\phi, \omega) = p_b(\phi, \omega) - \frac{1}{24}\sum_{t=1}^{n}\frac{3\omega_t^2 - 5(d+1)^2/4}{\omega_t^3}.$$

Practically, $\Sigma = \Lambda'\Omega\Lambda$, where $\Lambda$ is the diagonal matrix of the standard deviations $\sqrt{\phi_i} = \sigma_i$ and we use the sample correlation matrix for the estimation of $\Omega$ to speed up the algorithm.

# Appendix B

# Extensive Simulation Results

Table B.1 to B.8 are the results from the simulations that the price paths are simulated by geometric Brownian motion with constant volatility. Table B.1 to B.6 show the effect of different drifts added to the simulations. The measurements of accuracy and efficiency are reported. B.7 and B.8 show the effect of discretization. The average estimates and their 95% confidence intervals are reported with different numbers of intraday movements. Table B.9 to B.16 show the corresponding results when the price paths are simulated with the NIG-SV model. The values of mean, absolute error and standard error are scaled by $10^5$, the values of MSE are scaled by $10^9$.

Table B.1: The effect of drift on range-based estimators with constant volatility.

| daily variance $10^5 \times \sigma^2$ =2.5 and intraday movements $k = 20$ | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | $r_i^2$ | P | $P_k$ | BL | $BL_k$ | GK | $GK_k$ | RS | $RS_k$ |
| Panel A: $\mu = 0$ | | | | | | | | | |
| Mean | 2.50 | 1.86 | 2.43 | 1.33 | 1.65 | 1.62 | 2.39 | 1.59 | 2.32 |
| abs.err | 0.00 | 0.64 | 0.07 | 1.17 | 0.85 | 0.88 | 0.11 | 0.91 | 0.18 |
| std.err | 0.36 | 0.15 | 0.19 | 0.10 | 0.13 | 0.10 | 0.15 | 0.12 | 0.17 |
| MSE | 1.26 | 0.24 | 0.34 | 0.24 | 0.23 | 0.19 | 0.23 | 0.24 | 0.28 |
| QL | inf | 0.64 | 0.40 | 1.18 | 0.79 | 0.74 | 0.34 | 3.37 | 0.885 |
| Eff | 0.21 | 1.00 | 0.73 | 1.00 | 1.04 | 1.30 | 1.06 | 1.03 | 0.89 |
| Panel B: $\mu = 0.0005$ | | | | | | | | | |
| Mean | 2.51 | 1.87 | 2.44 | 1.33 | 1.66 | 1.62 | 2.39 | 1.60 | 2.32 |
| abs.err | 0.01 | 0.63 | 0.06 | 1.17 | 0.84 | 0.88 | 0.11 | 0.90 | 0.18 |
| std.err | 0.35 | 0.14 | 0.18 | 0.10 | 0.13 | 0.11 | 0.16 | 0.13 | 0.18 |
| MSE | 1.25 | 0.24 | 0.34 | 0.24 | 0.23 | 0.19 | 0.23 | 0.24 | 0.28 |
| QL | inf | 0.63 | 0.39 | 1.17 | 0.79 | 0.73 | 0.33 | 3.25 | 0.87 |
| Eff | 0.20 | 1.00 | 0.72 | 1.00 | 1.04 | 1.29 | 1.05 | 1.02 | 0.88 |
| Panel C: $\mu = 0.001$ | | | | | | | | | |
| Mean | 2.60 | 1.90 | 2.48 | 1.35 | 1.68 | 1.63 | 2.41 | 1.59 | 2.32 |
| abs.err | 0.10 | 0.60 | 0.02 | 1.15 | 0.82 | 0.87 | 0.09 | 0.91 | 0.18 |
| std.err | 0.38 | 0.15 | 0.19 | 0.11 | 0.13 | 0.11 | 0.15 | 0.13 | 0.17 |
| MSE | 1.36 | 0.25 | 0.37 | 0.24 | 0.24 | 0.19 | 0.24 | 0.24 | 0.29 |
| QL | inf | 0.63 | 0.39 | 1.16 | 0.78 | 0.73 | 0.33 | 3.81 | 0.89 |
| Eff | 0.20 | 1.00 | 0.70 | 1.04 | 1.06 | 1.34 | 1.06 | 1.05 | 0.90 |
| Panel D: $\mu = 0.005$ | | | | | | | | | |
| Mean | 5.01 | 2.73 | 3.57 | 1.94 | 2.42 | 1.85 | 2.86 | 1.49 | 2.34 |
| abs.err | 2.51 | 0.23 | 1.07 | 0.56 | 0.08 | 0.65 | 0.36 | 1.01 | 0.16 |
| std.err | 0.62 | 0.23 | 0.30 | 0.16 | 0.20 | 0.12 | 0.19 | 0.13 | 0.17 |
| MSE | 4.37 | 0.51 | 0.97 | 0.29 | 0.40 | 0.19 | 0.37 | 0.27 | 0.32 |
| QL | inf | 0.47 | 0.35 | 0.82 | 0.57 | 0.60 | 0.28 | 5.87 | 0.94 |
| Eff | 0.12 | 1.00 | 0.53 | 1.76 | 1.28 | 2.72 | 1.41 | 1.90 | 1.66 |
| Panel E: $\mu = 0.01$ | | | | | | | | | |
| Mean | 12.47 | 5.31 | 6.93 | 3.78 | 4.71 | 2.55 | 4.26 | 1.24 | 2.41 |
| abs.err | 9.97 | 2.81 | 4.43 | 1.28 | 2.21 | 0.05 | 1.76 | 1.26 | 0.09 |
| std.err | 1.09 | 0.39 | 0.51 | 0.28 | 0.34 | 0.16 | 0.27 | 0.14 | 0.20 |
| MSE | 21.15 | 2.21 | 4.38 | 0.88 | 1.60 | 0.25 | 0.99 | 0.35 | 0.39 |
| QL | inf | 0.37 | 0.43 | 0.38 | 0.36 | 0.34 | 0.25 | 7.05 | 0.83 |
| Eff | 0.10 | 1.00 | 0.50 | 2.53 | 1.38 | 9.16 | 2.25 | 6.42 | 5.96 |

| daily variance $10^5 \times \sigma^2$ =10 and intraday movements $k = 20$ | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | $r_i^2$ | P | $P_k$ | BL | $BL_k$ | GK | $GK_k$ | RS | $RS_k$ |
| Panel A: $\mu = 0$ | | | | | | | | | |
| Mean | 9.97 | 7.46 | 9.73 | 5.30 | 6.61 | 6.49 | 9.57 | 6.41 | 9.32 |
| abs.err | 0.03 | 2.54 | 0.27 | 4.70 | 3.39 | 3.51 | 0.43 | 3.59 | 0.68 |
| std.err | 1.39 | 0.54 | 0.71 | 0.39 | 0.48 | 0.40 | 0.58 | 0.49 | 0.64 |
| MSE | 19.81 | 3.82 | 5.42 | 3.81 | 3.65 | 2.99 | 3.75 | 3.83 | 4.60 |
| QL | inf | 0.63 | 0.39 | 1.17 | 0.79 | 0.73 | 0.33 | 3.88 | 0.872 |
| Eff | 0.21 | 1.00 | 0.72 | 1.00 | 1.04 | 1.28 | 1.04 | 1.01 | 0.86 |
| Panel B: $\mu = 0.0005$ | | | | | | | | | |
| Mean | 10.04 | 7.48 | 9.75 | 5.31 | 6.63 | 6.49 | 9.57 | 6.39 | 9.30 |
| abs.err | 0.04 | 2.52 | 0.25 | 4.69 | 3.37 | 3.51 | 0.43 | 3.61 | 0.70 |
| std.err | 1.45 | 0.58 | 0.76 | 0.41 | 0.52 | 0.42 | 0.62 | 0.50 | 0.67 |
| MSE | 20.22 | 3.87 | 5.51 | 3.83 | 3.68 | 2.99 | 3.76 | 3.82 | 4.57 |
| QL | inf | 0.63 | 0.39 | 1.17 | 0.79 | 0.73 | 0.33 | 8.41 | 0.88 |
| Eff | 0.21 | 1.00 | 0.72 | 1.00 | 1.05 | 1.30 | 1.05 | 1.02 | 0.88 |
| Panel C: $\mu = 0.001$ | | | | | | | | | |
| Mean | 10.01 | 7.45 | 9.72 | 5.29 | 6.60 | 6.46 | 9.53 | 6.37 | 9.27 |
| abs.err | 0.01 | 2.55 | 0.28 | 4.71 | 3.40 | 3.54 | 0.47 | 3.63 | 0.73 |
| std.err | 1.41 | 0.57 | 0.74 | 0.41 | 0.51 | 0.42 | 0.62 | 0.51 | 0.68 |
| MSE | 20.32 | 3.89 | 5.53 | 3.85 | 3.70 | 3.00 | 3.76 | 3.85 | 4.59 |
| QL | inf | 0.64 | 0.40 | 1.18 | 0.80 | 0.75 | 0.34 | 6.07 | 0.89 |
| Eff | 0.21 | 1.00 | 0.72 | 1.00 | 1.04 | 1.30 | 1.06 | 1.02 | 0.88 |
| Panel D: $\mu = 0.005$ | | | | | | | | | |
| Mean | 12.52 | 8.33 | 10.87 | 5.92 | 7.38 | 6.71 | 10.02 | 6.26 | 9.29 |
| abs.err | 2.52 | 1.67 | 0.87 | 4.08 | 2.62 | 3.29 | 0.02 | 3.74 | 0.71 |
| std.err | 1.73 | 0.65 | 0.85 | 0.46 | 0.58 | 0.43 | 0.64 | 0.51 | 0.68 |
| MSE | 30.77 | 4.74 | 7.67 | 3.92 | 4.19 | 3.00 | 4.25 | 3.96 | 4.72 |
| QL | inf | 0.59 | 0.38 | 1.07 | 0.73 | 0.70 | 0.32 | 5.90 | 0.91 |
| Eff | 0.16 | 1.00 | 0.63 | 1.20 | 1.12 | 1.58 | 1.13 | 1.21 | 1.04 |
| Panel E: $\mu = 0.01$ | | | | | | | | | |
| Mean | 20.14 | 10.96 | 14.30 | 7.79 | 9.72 | 7.42 | 11.45 | 5.95 | 9.34 |
| abs.err | 10.14 | 0.96 | 4.30 | 2.21 | 0.28 | 2.58 | 1.45 | 4.05 | 0.66 |
| std.err | 2.44 | 0.87 | 1.14 | 0.62 | 0.77 | 0.45 | 0.70 | 0.51 | 0.68 |
| MSE | 71.09 | 8.26 | 15.75 | 4.61 | 6.42 | 3.02 | 5.86 | 4.37 | 5.12 |
| QL | inf | 0.47 | 0.34 | 0.81 | 0.57 | 0.60 | 0.28 | 13.61 | 0.95 |
| Eff | 0.12 | 1.00 | 0.53 | 1.77 | 1.28 | 2.75 | 1.43 | 1.90 | 1.67 |

Table B.2: The effect of drift on range-based estimators with constant volatility.

| daily variance $10^5 \times \sigma^2 = 40$ and intraday movements $k = 20$ | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | $r_i^2$ | P | $P_k$ | BL | $BL_k$ | GK | $GK_k$ | RS | $RS_k$ |
| Panel A: $\mu = 0$ | | | | | | | | | |
| Mean | 40.1 | 29.9 | 39.0 | 21.3 | 26.5 | 26.0 | 38.3 | 25.6 | 37.3 |
| abs.err | 0.1 | 10.1 | 1.0 | 18.7 | 13.5 | 14.0 | 1.7 | 14.4 | 2.7 |
| std.err | 5.34 | 2.21 | 2.88 | 1.57 | 1.96 | 1.67 | 2.44 | 1.98 | 2.67 |
| MSE | 326 | 62 | 89 | 61 | 59 | 48 | 60 | 61 | 73 |
| QL | inf | 0.63 | 0.39 | 1.17 | 0.79 | 0.73 | 0.33 | 6.22 | 0.87 |
| Eff | 0.20 | 1.00 | 0.72 | 1.01 | 1.05 | 1.31 | 1.06 | 1.03 | 0.88 |
| Panel B: $\mu = 0.0005$ | | | | | | | | | |
| Mean | 40.1 | 29.9 | 39.0 | 21.3 | 26.5 | 26.0 | 38.4 | 25.7 | 37.3 |
| abs.err | 0.1 | 10.1 | 1.0 | 18.7 | 13.5 | 14.0 | 1.6 | 14.3 | 2.7 |
| std.err | 5.79 | 2.26 | 2.95 | 1.61 | 2.01 | 1.66 | 2.42 | 2.05 | 2.72 |
| MSE | 321 | 62 | 89 | 61 | 59 | 48 | 61 | 62 | 75 |
| QL | inf | 0.63 | 0.39 | 1.17 | 0.79 | 0.74 | 0.33 | 3.28 | 0.88 |
| Eff | 0.21 | 1.00 | 0.72 | 1.01 | 1.05 | 1.29 | 1.04 | 1.02 | 0.86 |
| Panel C: $\mu = 0.001$ | | | | | | | | | |
| Mean | 40.2 | 29.9 | 39.0 | 21.2 | 26.5 | 25.9 | 38.2 | 25.5 | 37.2 |
| abs.err | 0.2 | 10.1 | 1.0 | 18.8 | 13.5 | 14.1 | 1.8 | 14.5 | 2.8 |
| std.err | 5.52 | 2.23 | 2.90 | 1.58 | 1.97 | 1.67 | 2.43 | 1.99 | 2.67 |
| MSE | 322 | 62 | 88 | 61 | 59 | 48 | 60 | 62 | 74 |
| QL | inf | 0.63 | 0.39 | 1.18 | 0.79 | 0.74 | 0.33 | 4.06 | 0.88 |
| Eff | 0.20 | 1.00 | 0.72 | 1.00 | 1.04 | 1.29 | 1.05 | 1.01 | 0.87 |
| Panel D: $\mu = 0.005$ | | | | | | | | | |
| Mean | 42.2 | 30.6 | 39.9 | 21.8 | 27.1 | 26.1 | 38.6 | 25.4 | 37.2 |
| abs.err | 2.2 | 9.4 | 0.1 | 18.2 | 12.9 | 13.9 | 1.4 | 14.6 | 2.8 |
| std.err | 5.77 | 2.37 | 3.10 | 1.69 | 2.10 | 1.75 | 2.56 | 2.05 | 2.77 |
| MSE | 357 | 65 | 95 | 62 | 61 | 48 | 62 | 62 | 75 |
| QL | inf | 0.63 | 0.39 | 1.16 | 0.78 | 0.73 | 0.33 | 3.23 | 0.89 |
| Eff | 0.19 | 1.00 | 0.70 | 1.04 | 1.06 | 1.35 | 1.06 | 1.06 | 0.90 |
| Panel E: $\mu = 0.01$ | | | | | | | | | |
| Mean | 50.1 | 33.3 | 43.4 | 23.7 | 29.5 | 26.8 | 40.1 | 25.1 | 37.2 |
| abs.err | 10.1 | 6.7 | 3.4 | 16.3 | 10.5 | 13.2 | 0.1 | 14.9 | 2.8 |
| std.err | 6.77 | 2.60 | 3.39 | 1.85 | 2.30 | 1.74 | 2.58 | 2.04 | 2.74 |
| MSE | 491 | 75 | 122 | 62 | 67 | 48 | 67 | 64 | 76 |
| QL | inf | 0.58 | 0.37 | 1.07 | 0.72 | 0.70 | 0.32 | 4.83 | 0.90 |
| Eff | 0.16 | 1.00 | 0.63 | 1.19 | 1.12 | 1.58 | 1.14 | 1.19 | 1.04 |

| daily variance $10^5 \times \sigma^2 = 90$ and intraday movements $k = 20$ | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | $r_i^2$ | P | $P_k$ | BL | $BL_k$ | GK | $GK_k$ | RS | $RS_k$ |
| Panel A: $\mu = 0$ | | | | | | | | | |
| Mean | 90.1 | 67.1 | 87.5 | 47.7 | 59.5 | 58.2 | 85.9 | 57.4 | 83.6 |
| abs.err | 0.1 | 22.9 | 2.5 | 42.3 | 30.5 | 31.8 | 4.1 | 32.6 | 6.4 |
| std.err | 12.73 | 4.96 | 6.47 | 3.52 | 4.39 | 3.66 | 5.32 | 4.52 | 6.02 |
| MSE | 1616 | 310 | 439 | 309 | 296 | 241 | 301 | 309 | 368 |
| QL | inf | 0.63 | 0.39 | 1.17 | 0.79 | 0.73 | 0.33 | 3.54 | 0.87 |
| Eff | 0.21 | 1.00 | 0.73 | 1.00 | 1.04 | 1.29 | 1.05 | 1.01 | 0.88 |
| Panel B: $\mu = 0.0005$ | | | | | | | | | |
| Mean | 90.2 | 67.3 | 87.9 | 47.8 | 59.7 | 58.5 | 86.3 | 57.7 | 84.0 |
| abs.err | 0.2 | 22.7 | 2.1 | 42.2 | 30.3 | 31.5 | 3.7 | 32.3 | 6.0 |
| std.err | 12.13 | 4.98 | 6.50 | 3.54 | 4.41 | 3.79 | 5.53 | 4.55 | 6.11 |
| MSE | 1610 | 311 | 443 | 309 | 296 | 242 | 306 | 309 | 371 |
| QL | inf | 0.63 | 0.39 | 1.17 | 0.79 | 0.74 | 0.34 | 4.47 | 0.89 |
| Eff | 0.21 | 1.00 | 0.72 | 1.00 | 1.04 | 1.29 | 1.04 | 1.02 | 0.87 |
| Panel C: $\mu = 0.001$ | | | | | | | | | |
| Mean | 90.1 | 67.1 | 87.6 | 47.7 | 59.5 | 58.3 | 86.0 | 57.4 | 83.7 |
| abs.err | 0.1 | 22.9 | 2.4 | 42.3 | 30.5 | 31.7 | 4.0 | 32.6 | 6.3 |
| std.err | 12.99 | 5.29 | 6.90 | 3.76 | 4.69 | 3.80 | 5.60 | 4.44 | 6.00 |
| MSE | 1644 | 316 | 449 | 312 | 300 | 243 | 307 | 311 | 373 |
| QL | 42041.06 | 0.63 | 0.39 | 1.17 | 0.79 | 0.74 | 0.34 | 7.04 | 0.89 |
| Eff | 0.21 | 1.00 | 0.73 | 1.01 | 1.05 | 1.30 | 1.05 | 1.03 | 0.88 |
| Panel D: $\mu = 0.005$ | | | | | | | | | |
| Mean | 92.0 | 67.8 | 88.5 | 48.2 | 60.1 | 58.5 | 86.3 | 57.4 | 83.7 |
| abs.err | 2.0 | 22.2 | 1.5 | 41.8 | 29.9 | 31.5 | 3.7 | 32.6 | 6.3 |
| std.err | 13.17 | 5.35 | 6.99 | 3.80 | 4.75 | 3.81 | 5.63 | 4.35 | 5.90 |
| MSE | 1685 | 316 | 455 | 310 | 299 | 241 | 305 | 310 | 370 |
| QL | inf | 0.63 | 0.39 | 1.17 | 0.79 | 0.74 | 0.34 | 3.27 | 0.90 |
| Eff | 0.20 | 1.00 | 0.72 | 1.01 | 1.05 | 1.31 | 1.05 | 1.03 | 0.88 |
| Panel E: $\mu = 0.01$ | | | | | | | | | |
| Mean | 99.7 | 70.6 | 92.0 | 50.1 | 62.5 | 59.3 | 87.9 | 57.2 | 83.9 |
| abs.err | 9.7 | 19.4 | 2.0 | 39.9 | 27.5 | 30.7 | 2.1 | 32.8 | 6.1 |
| std.err | 13.55 | 5.38 | 7.02 | 3.82 | 4.77 | 3.92 | 5.74 | 4.64 | 6.24 |
| MSE | 1973 | 340 | 515 | 312 | 313 | 243 | 323 | 316 | 379 |
| QL | inf | 0.61 | 0.39 | 1.13 | 0.76 | 0.72 | 0.33 | 6.45 | 0.89 |
| Eff | 0.18 | 1.00 | 0.68 | 1.08 | 1.08 | 1.40 | 1.08 | 1.09 | 0.93 |

Table B.3: The effect of drift on range-based estimators with constant volatility.

| daily variance $10^5 \times \sigma^2 = 2.5$ and intraday movements $k = 40$ | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | $r_i^2$ | P | $P_k$ | BL | $BL_k$ | GK | $GK_k$ | RS | $RS_k$ |
| Panel A: $\mu = 0$ | | | | | | | | | |
| Mean | 2.50 | 2.03 | 2.44 | 1.44 | 1.68 | 1.85 | 2.40 | 1.84 | 2.37 |
| abs.err | 0.00 | 0.47 | 0.06 | 1.06 | 0.82 | 0.65 | 0.10 | 0.66 | 0.13 |
| std.err | 0.34 | 0.14 | 0.17 | 0.10 | 0.12 | 0.11 | 0.14 | 0.13 | 0.16 |
| MSE | 1.24 | 0.23 | 0.31 | 0.22 | 0.21 | 0.17 | 0.21 | 0.21 | 0.26 |
| QL | inf | 0.46 | 0.33 | 0.88 | 0.66 | 0.48 | 0.27 | 2.84 | 0.886 |
| Eff | 0.20 | 1.00 | 0.78 | 1.06 | 1.09 | 1.41 | 1.14 | 1.11 | 0.94 |
| Panel B: $\mu = 0.0005$ | | | | | | | | | |
| Mean | 2.52 | 2.03 | 2.44 | 1.44 | 1.68 | 1.84 | 2.40 | 1.83 | 2.36 |
| abs.err | 0.02 | 0.47 | 0.06 | 1.06 | 0.82 | 0.66 | 0.10 | 0.67 | 0.14 |
| std.err | 0.36 | 0.15 | 0.18 | 0.11 | 0.12 | 0.11 | 0.14 | 0.13 | 0.16 |
| MSE | 1.27 | 0.24 | 0.31 | 0.22 | 0.21 | 0.17 | 0.21 | 0.22 | 0.26 |
| QL | inf | 0.46 | 0.33 | 0.89 | 0.66 | 0.48 | 0.28 | 2.35 | 0.89 |
| Eff | 0.20 | 1.00 | 0.78 | 1.07 | 1.10 | 1.42 | 1.15 | 1.12 | 0.95 |
| Panel C: $\mu = 0.001$ | | | | | | | | | |
| Mean | 2.59 | 2.06 | 2.47 | 1.46 | 1.70 | 1.85 | 2.41 | 1.83 | 2.36 |
| abs.err | 0.09 | 0.44 | 0.03 | 1.04 | 0.80 | 0.65 | 0.09 | 0.67 | 0.14 |
| std.err | 0.37 | 0.15 | 0.18 | 0.11 | 0.13 | 0.11 | 0.14 | 0.13 | 0.16 |
| MSE | 1.34 | 0.24 | 0.32 | 0.22 | 0.22 | 0.17 | 0.21 | 0.22 | 0.26 |
| QL | inf | 0.45 | 0.32 | 0.87 | 0.65 | 0.48 | 0.27 | 3.45 | 0.91 |
| Eff | 0.19 | 1.00 | 0.77 | 1.09 | 1.11 | 1.46 | 1.17 | 1.14 | 0.98 |
| Panel D: $\mu = 0.005$ | | | | | | | | | |
| Mean | 5.00 | 2.90 | 3.49 | 2.06 | 2.40 | 2.10 | 2.80 | 1.74 | 2.35 |
| abs.err | 2.50 | 0.40 | 0.99 | 0.44 | 0.10 | 0.40 | 0.30 | 0.76 | 0.15 |
| std.err | 0.63 | 0.24 | 0.29 | 0.17 | 0.20 | 0.13 | 0.18 | 0.14 | 0.17 |
| MSE | 4.34 | 0.54 | 0.85 | 0.28 | 0.36 | 0.18 | 0.31 | 0.24 | 0.29 |
| QL | inf | 0.37 | 0.30 | 0.63 | 0.49 | 0.40 | 0.24 | 11.26 | 1.10 |
| Eff | 0.13 | 1.00 | 0.63 | 1.88 | 1.49 | 2.94 | 1.75 | 2.24 | 1.91 |
| Panel E: $\mu = 0.01$ | | | | | | | | | |
| Mean | 12.55 | 5.55 | 6.66 | 3.94 | 4.59 | 2.84 | 4.01 | 1.54 | 2.34 |
| abs.err | 10.05 | 3.05 | 4.16 | 1.44 | 2.09 | 0.34 | 1.51 | 0.96 | 0.16 |
| std.err | 1.07 | 0.39 | 0.46 | 0.27 | 0.32 | 0.17 | 0.24 | 0.15 | 0.19 |
| MSE | 21.47 | 2.40 | 3.86 | 0.95 | 1.45 | 0.29 | 0.78 | 0.30 | 0.35 |
| QL | inf | 0.35 | 0.39 | 0.33 | 0.32 | 0.25 | 0.22 | 8.28 | 1.20 |
| Eff | 0.11 | 1.00 | 0.62 | 2.54 | 1.66 | 8.49 | 3.11 | 8.03 | 7.11 |

| daily variance $10^5 \times \sigma^2 = 10$ and intraday movements $k = 40$ | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | $r_i^2$ | P | $P_k$ | BL | $BL_k$ | GK | $GK_k$ | RS | $RS_k$ |
| Panel A: $\mu = 0$ | | | | | | | | | |
| Mean | 9.94 | 8.09 | 9.72 | 5.75 | 6.70 | 7.38 | 9.59 | 7.34 | 9.45 |
| abs.err | 0.06 | 1.91 | 0.28 | 4.25 | 3.30 | 2.62 | 0.41 | 2.66 | 0.55 |
| std.err | 1.41 | 0.57 | 0.69 | 0.41 | 0.48 | 0.43 | 0.56 | 0.52 | 0.63 |
| MSE | 19.83 | 3.76 | 4.90 | 3.52 | 3.41 | 2.69 | 3.37 | 3.48 | 4.18 |
| QL | inf | 0.46 | 0.33 | 0.88 | 0.66 | 0.48 | 0.27 | 26.77 | 0.893 |
| Eff | 0.20 | 1.00 | 0.78 | 1.06 | 1.09 | 1.40 | 1.13 | 1.10 | 0.93 |
| Panel B: $\mu = 0.0005$ | | | | | | | | | |
| Mean | 10.05 | 8.14 | 9.78 | 5.78 | 6.74 | 7.40 | 9.62 | 7.36 | 9.48 |
| abs.err | 0.05 | 1.86 | 0.22 | 4.22 | 3.26 | 2.60 | 0.38 | 2.64 | 0.52 |
| std.err | 1.41 | 0.57 | 0.69 | 0.41 | 0.47 | 0.44 | 0.57 | 0.53 | 0.65 |
| MSE | 20.50 | 3.80 | 4.99 | 3.52 | 3.43 | 2.67 | 3.36 | 3.47 | 4.17 |
| QL | inf | 0.46 | 0.33 | 0.88 | 0.66 | 0.48 | 0.28 | 2.53 | 0.89 |
| Eff | 0.20 | 1.00 | 0.78 | 1.07 | 1.10 | 1.43 | 1.15 | 1.12 | 0.95 |
| Panel C: $\mu = 0.001$ | | | | | | | | | |
| Mean | 10.08 | 8.15 | 9.79 | 5.79 | 6.75 | 7.40 | 9.62 | 7.35 | 9.47 |
| abs.err | 0.08 | 1.85 | 0.21 | 4.21 | 3.25 | 2.60 | 0.38 | 2.65 | 0.53 |
| std.err | 1.44 | 0.59 | 0.70 | 0.42 | 0.49 | 0.45 | 0.58 | 0.52 | 0.64 |
| MSE | 20.63 | 3.84 | 5.05 | 3.54 | 3.46 | 2.69 | 3.39 | 3.44 | 4.14 |
| QL | inf | 0.46 | 0.33 | 0.88 | 0.66 | 0.48 | 0.27 | 2.77 | 0.88 |
| Eff | 0.20 | 1.00 | 0.77 | 1.08 | 1.10 | 1.44 | 1.15 | 1.13 | 0.96 |
| Panel D: $\mu = 0.005$ | | | | | | | | | |
| Mean | 12.52 | 9.00 | 10.81 | 6.40 | 7.45 | 7.64 | 10.01 | 7.25 | 9.45 |
| abs.err | 2.52 | 1.00 | 0.81 | 3.60 | 2.55 | 2.36 | 0.01 | 2.75 | 0.55 |
| std.err | 1.76 | 0.67 | 0.81 | 0.48 | 0.56 | 0.46 | 0.60 | 0.54 | 0.66 |
| MSE | 31.13 | 4.87 | 6.94 | 3.71 | 3.92 | 2.75 | 3.77 | 3.60 | 4.32 |
| QL | inf | 0.43 | 0.32 | 0.81 | 0.61 | 0.46 | 0.27 | 3.78 | 0.96 |
| Eff | 0.16 | 1.00 | 0.71 | 1.30 | 1.23 | 1.77 | 1.31 | 1.38 | 1.17 |
| Panel E: $\mu = 0.01$ | | | | | | | | | |
| Mean | 19.89 | 11.61 | 13.94 | 8.25 | 9.61 | 8.41 | 11.25 | 7.02 | 9.45 |
| abs.err | 9.89 | 1.61 | 3.94 | 1.75 | 0.39 | 1.59 | 1.25 | 2.98 | 0.55 |
| std.err | 2.46 | 0.95 | 1.14 | 0.68 | 0.79 | 0.56 | 0.75 | 0.57 | 0.72 |
| MSE | 69.63 | 8.69 | 13.71 | 4.56 | 5.79 | 2.96 | 5.04 | 3.90 | 4.68 |
| QL | inf | 0.37 | 0.30 | 0.63 | 0.49 | 0.40 | 0.24 | 12.26 | 1.08 |
| Eff | 0.13 | 1.00 | 0.63 | 1.88 | 1.49 | 2.94 | 1.75 | 2.26 | 1.92 |

Table B.4: The effect of drift on range-based estimators with constant volatility.

| daily variance $10^5 \times \sigma^2 = 40$ and intraday movements $k = 40$ | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | $r_i^2$ | P | $P_k$ | BL | $BL_k$ | GK | $GK_k$ | RS | $RS_k$ |
| Panel A: $\mu = 0$ | | | | | | | | | |
| Mean | 40.0 | 32.5 | 39.0 | 23.1 | 26.9 | 29.5 | 38.4 | 29.4 | 37.8 |
| abs.err | 0.0 | 7.5 | 1.0 | 16.9 | 13.1 | 10.5 | 1.6 | 10.6 | 2.2 |
| std.err | 5.77 | 2.43 | 2.92 | 1.73 | 2.01 | 1.79 | 2.33 | 2.05 | 2.52 |
| MSE | 320 | 60 | 79 | 56 | 55 | 43 | 54 | 55 | 66 |
| QL | inf | 0.46 | 0.33 | 0.89 | 0.66 | 0.48 | 0.28 | 2.26 | 0.88 |
| Eff | 0.20 | 1.00 | 0.78 | 1.06 | 1.09 | 1.41 | 1.14 | 1.11 | 0.94 |
| Panel B: $\mu = 0.0005$ | | | | | | | | | |
| Mean | 40.2 | 32.5 | 39.0 | 23.1 | 26.9 | 29.5 | 38.4 | 29.3 | 37.8 |
| abs.err | 0.2 | 7.5 | 1.0 | 16.9 | 13.1 | 10.5 | 1.6 | 10.7 | 2.2 |
| std.err | 5.63 | 2.33 | 2.80 | 1.66 | 1.93 | 1.80 | 2.33 | 2.12 | 2.60 |
| MSE | 321 | 61 | 79 | 56 | 55 | 43 | 54 | 56 | 67 |
| QL | inf | 0.46 | 0.33 | 0.89 | 0.67 | 0.48 | 0.28 | 2.80 | 0.90 |
| Eff | 0.20 | 1.00 | 0.78 | 1.07 | 1.09 | 1.41 | 1.14 | 1.11 | 0.94 |
| Panel C: $\mu = 0.001$ | | | | | | | | | |
| Mean | 40.2 | 32.5 | 39.1 | 23.1 | 26.9 | 29.6 | 38.4 | 29.4 | 37.8 |
| abs.err | 0.2 | 7.5 | 0.9 | 16.9 | 13.1 | 10.4 | 1.6 | 10.6 | 2.2 |
| std.err | 5.71 | 2.38 | 2.86 | 1.69 | 1.97 | 1.80 | 2.33 | 2.08 | 2.55 |
| MSE | 324 | 61 | 79 | 56 | 55 | 43 | 54 | 56 | 67 |
| QL | inf | 0.46 | 0.33 | 0.88 | 0.66 | 0.48 | 0.28 | 2.06 | 0.90 |
| Eff | 0.20 | 1.00 | 0.78 | 1.07 | 1.09 | 1.42 | 1.15 | 1.10 | 0.94 |
| Panel D: $\mu = 0.005$ | | | | | | | | | |
| Mean | 42.6 | 33.4 | 40.1 | 23.7 | 27.6 | 29.8 | 38.8 | 29.3 | 37.9 |
| abs.err | 2.6 | 6.6 | 0.1 | 16.3 | 12.4 | 10.2 | 1.2 | 10.7 | 2.1 |
| std.err | 5.98 | 2.48 | 2.98 | 1.76 | 2.05 | 1.80 | 2.34 | 2.03 | 2.51 |
| MSE | 362 | 64 | 86 | 57 | 56 | 43 | 55 | 56 | 67 |
| QL | inf | 0.45 | 0.32 | 0.86 | 0.65 | 0.47 | 0.27 | 2.70 | 0.91 |
| Eff | 0.19 | 1.00 | 0.76 | 1.12 | 1.13 | 1.50 | 1.18 | 1.17 | 0.99 |
| Panel E: $\mu = 0.01$ | | | | | | | | | |
| Mean | 50.0 | 35.9 | 43.2 | 25.5 | 29.8 | 30.5 | 40.0 | 28.9 | 37.7 |
| abs.err | 10.0 | 4.1 | 3.2 | 14.5 | 10.2 | 9.5 | 0.0 | 11.1 | 2.3 |
| std.err | 6.97 | 2.73 | 3.28 | 1.94 | 2.26 | 1.86 | 2.44 | 2.13 | 2.62 |
| MSE | 486 | 76 | 108 | 59 | 62 | 44 | 60 | 57 | 69 |
| QL | inf | 0.43 | 0.32 | 0.81 | 0.61 | 0.46 | 0.27 | 4.41 | 0.95 |
| Eff | 0.16 | 1.00 | 0.71 | 1.28 | 1.22 | 1.74 | 1.29 | 1.35 | 1.15 |

| daily variance $10^5 \times \sigma^2 = 90$ and intraday movements $k = 40$ | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | $r_i^2$ | P | $P_k$ | BL | $BL_k$ | GK | $GK_k$ | RS | $RS_k$ |
| Panel A: $\mu = 0$ | | | | | | | | | |
| Mean | 90.2 | 73.1 | 87.7 | 51.9 | 60.5 | 66.4 | 86.3 | 66.0 | 85.0 |
| abs.err | 0.2 | 16.9 | 2.3 | 38.1 | 29.5 | 23.6 | 3.7 | 24.0 | 5.0 |
| std.err | 12.53 | 5.13 | 6.16 | 3.64 | 4.25 | 3.89 | 5.03 | 4.65 | 5.68 |
| MSE | 1636 | 306 | 400 | 285 | 277 | 217 | 271 | 281 | 336 |
| QL | inf | 0.46 | 0.33 | 0.88 | 0.66 | 0.48 | 0.28 | 2.47 | 0.90 |
| Eff | 0.20 | 1.00 | 0.78 | 1.06 | 1.09 | 1.41 | 1.14 | 1.10 | 0.94 |
| Panel B: $\mu = 0.0005$ | | | | | | | | | |
| Mean | 90.2 | 73.1 | 87.8 | 52.0 | 60.5 | 66.5 | 86.5 | 66.1 | 85.2 |
| abs.err | 0.2 | 16.9 | 2.2 | 38.0 | 29.5 | 23.5 | 3.5 | 23.9 | 4.8 |
| std.err | 12.56 | 5.21 | 6.25 | 3.70 | 4.31 | 4.07 | 5.25 | 4.84 | 5.92 |
| MSE | 1623 | 306 | 400 | 285 | 277 | 217 | 273 | 281 | 338 |
| QL | inf | 0.46 | 0.33 | 0.88 | 0.66 | 0.48 | 0.27 | 2.50 | 0.88 |
| Eff | 0.20 | 1.00 | 0.78 | 1.06 | 1.09 | 1.41 | 1.14 | 1.10 | 0.94 |
| Panel C: $\mu = 0.001$ | | | | | | | | | |
| Mean | 90.1 | 73.0 | 87.6 | 51.8 | 60.4 | 66.3 | 86.2 | 65.9 | 84.9 |
| abs.err | 0.1 | 17.0 | 2.4 | 38.2 | 29.6 | 23.7 | 3.8 | 24.1 | 5.1 |
| std.err | 13.05 | 5.37 | 6.45 | 3.81 | 4.44 | 4.06 | 5.25 | 4.77 | 5.84 |
| MSE | 1606 | 302 | 395 | 284 | 275 | 216 | 269 | 280 | 334 |
| QL | 82140.60 | 0.46 | 0.33 | 0.89 | 0.66 | 0.48 | 0.28 | 3.49 | 0.89 |
| Eff | 0.20 | 1.00 | 0.78 | 1.06 | 1.09 | 1.40 | 1.14 | 1.10 | 0.94 |
| Panel D: $\mu = 0.005$ | | | | | | | | | |
| Mean | 92.0 | 73.8 | 88.7 | 52.5 | 61.1 | 66.8 | 86.9 | 66.1 | 85.2 |
| abs.err | 2.0 | 16.2 | 1.3 | 37.5 | 28.9 | 23.2 | 3.1 | 23.9 | 4.8 |
| std.err | 13.47 | 5.43 | 6.52 | 3.86 | 4.49 | 4.02 | 5.21 | 4.71 | 5.75 |
| MSE | 1695 | 314 | 415 | 286 | 281 | 219 | 278 | 282 | 340 |
| QL | inf | 0.46 | 0.33 | 0.88 | 0.66 | 0.48 | 0.28 | 2.38 | 0.89 |
| Eff | 0.20 | 1.00 | 0.77 | 1.09 | 1.11 | 1.44 | 1.15 | 1.13 | 0.96 |
| Panel E: $\mu = 0.01$ | | | | | | | | | |
| Mean | 100.9 | 76.8 | 92.3 | 54.6 | 63.6 | 67.6 | 88.2 | 65.7 | 85.1 |
| abs.err | 10.9 | 13.2 | 2.3 | 35.4 | 26.4 | 22.4 | 1.8 | 24.3 | 4.9 |
| std.err | 13.98 | 5.67 | 6.81 | 4.03 | 4.69 | 4.10 | 5.34 | 4.70 | 5.77 |
| MSE | 2017 | 341 | 468 | 289 | 292 | 218 | 285 | 285 | 342 |
| QL | inf | 0.44 | 0.32 | 0.85 | 0.63 | 0.47 | 0.27 | 2.54 | 0.94 |
| Eff | 0.18 | 1.00 | 0.74 | 1.17 | 1.16 | 1.57 | 1.21 | 1.22 | 1.03 |

Table B.5: The effect of drift on range-based estimators with constant volatility.

| daily variance $10^5 \times \sigma^2 = 2.5$ and intraday movements $k = 100$ | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | $r_i^2$ | P | $P_k$ | BL | $BL_k$ | GK | $GK_k$ | RS | $RS_k$ |
| Panel A: $\mu = 0$ | | | | | | | | | |
| Mean | 2.51 | 2.19 | 2.46 | 1.56 | 1.71 | 2.07 | 2.43 | 2.07 | 2.41 |
| abs.err | 0.01 | 0.31 | 0.04 | 0.94 | 0.79 | 0.43 | 0.07 | 0.43 | 0.09 |
| std.err | 0.35 | 0.15 | 0.17 | 0.11 | 0.12 | 0.11 | 0.13 | 0.13 | 0.15 |
| MSE | 1.26 | 0.24 | 0.29 | 0.20 | 0.20 | 0.16 | 0.19 | 0.20 | 0.24 |
| QL | inf | 0.35 | 0.28 | 0.69 | 0.57 | 0.32 | 0.23 | 1.48 | 0.865 |
| Eff | 0.20 | 1.00 | 0.84 | 1.16 | 1.17 | 1.51 | 1.25 | 1.19 | 1.03 |
| Panel B: $\mu = 0.0005$ | | | | | | | | | |
| Mean | 2.51 | 2.19 | 2.45 | 1.56 | 1.71 | 2.07 | 2.42 | 2.06 | 2.40 |
| abs.err | 0.01 | 0.31 | 0.05 | 0.94 | 0.79 | 0.43 | 0.08 | 0.44 | 0.10 |
| std.err | 0.35 | 0.15 | 0.17 | 0.11 | 0.12 | 0.12 | 0.14 | 0.14 | 0.16 |
| MSE | 1.28 | 0.24 | 0.29 | 0.20 | 0.20 | 0.16 | 0.19 | 0.20 | 0.24 |
| QL | inf | 0.35 | 0.28 | 0.69 | 0.57 | 0.32 | 0.23 | 1.36 | 0.91 |
| Eff | 0.20 | 1.00 | 0.84 | 1.15 | 1.17 | 1.52 | 1.26 | 1.21 | 1.05 |
| Panel C: $\mu = 0.001$ | | | | | | | | | |
| Mean | 2.63 | 2.24 | 2.50 | 1.59 | 1.75 | 2.08 | 2.45 | 2.06 | 2.41 |
| abs.err | 0.13 | 0.26 | 0.00 | 0.91 | 0.75 | 0.42 | 0.05 | 0.44 | 0.09 |
| std.err | 0.38 | 0.16 | 0.18 | 0.11 | 0.12 | 0.12 | 0.14 | 0.14 | 0.16 |
| MSE | 1.38 | 0.25 | 0.31 | 0.21 | 0.21 | 0.16 | 0.20 | 0.21 | 0.24 |
| QL | inf | 0.34 | 0.28 | 0.68 | 0.56 | 0.32 | 0.23 | 1.52 | 0.88 |
| Eff | 0.19 | 1.00 | 0.83 | 1.20 | 1.20 | 1.58 | 1.30 | 1.25 | 1.08 |
| Panel D: $\mu = 0.005$ | | | | | | | | | |
| Mean | 4.98 | 3.08 | 3.45 | 2.19 | 2.41 | 2.35 | 2.80 | 2.01 | 2.40 |
| abs.err | 2.48 | 0.58 | 0.95 | 0.31 | 0.09 | 0.15 | 0.30 | 0.49 | 0.10 |
| std.err | 0.64 | 0.24 | 0.27 | 0.17 | 0.19 | 0.14 | 0.16 | 0.14 | 0.16 |
| MSE | 4.36 | 0.59 | 0.78 | 0.29 | 0.34 | 0.19 | 0.28 | 0.23 | 0.27 |
| QL | inf | 0.30 | 0.27 | 0.51 | 0.43 | 0.28 | 0.21 | 1.92 | 1.21 |
| Eff | 0.14 | 1.00 | 0.75 | 2.01 | 1.72 | 3.06 | 2.10 | 2.63 | 2.27 |
| Panel E: $\mu = 0.01$ | | | | | | | | | |
| Mean | 12.47 | 5.74 | 6.43 | 4.08 | 4.49 | 3.14 | 3.87 | 1.85 | 2.37 |
| abs.err | 9.97 | 3.24 | 3.93 | 1.58 | 1.99 | 0.64 | 1.37 | 0.65 | 0.13 |
| std.err | 1.05 | 0.39 | 0.43 | 0.27 | 0.30 | 0.18 | 0.22 | 0.15 | 0.18 |
| MSE | 21.20 | 2.56 | 3.43 | 1.01 | 1.31 | 0.35 | 0.66 | 0.27 | 0.32 |
| QL | inf | 0.34 | 0.37 | 0.30 | 0.30 | 0.20 | 0.19 | 5.89 | 1.65 |
| Eff | 0.12 | 1.00 | 0.74 | 2.54 | 1.95 | 7.38 | 3.93 | 9.56 | 8.30 |

| daily variance $10^5 \times \sigma^2 = 10$ and intraday movements $k = 100$ | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | $r_i^2$ | P | $P_k$ | BL | $BL_k$ | GK | $GK_k$ | RS | $RS_k$ |
| Panel A: $\mu = 0$ | | | | | | | | | |
| Mean | 10.01 | 8.75 | 9.80 | 6.22 | 6.84 | 8.26 | 9.69 | 8.24 | 9.62 |
| abs.err | 0.01 | 1.25 | 0.20 | 3.78 | 3.16 | 1.74 | 0.31 | 1.76 | 0.38 |
| std.err | 1.41 | 0.59 | 0.67 | 0.42 | 0.46 | 0.47 | 0.55 | 0.54 | 0.62 |
| MSE | 19.91 | 3.77 | 4.53 | 3.25 | 3.21 | 2.53 | 3.06 | 3.26 | 3.81 |
| QL | inf | 0.35 | 0.28 | 0.69 | 0.57 | 0.32 | 0.23 | 1.93 | 0.876 |
| Eff | 0.20 | 1.00 | 0.84 | 1.15 | 1.16 | 1.50 | 1.25 | 1.18 | 1.02 |
| Panel B: $\mu = 0.0005$ | | | | | | | | | |
| Mean | 10.07 | 8.78 | 9.83 | 6.24 | 6.86 | 8.28 | 9.71 | 8.25 | 9.63 |
| abs.err | 0.07 | 1.22 | 0.17 | 3.76 | 3.14 | 1.72 | 0.29 | 1.75 | 0.37 |
| std.err | 1.42 | 0.58 | 0.65 | 0.41 | 0.45 | 0.45 | 0.53 | 0.54 | 0.61 |
| MSE | 20.07 | 3.79 | 4.57 | 3.26 | 3.21 | 2.56 | 3.11 | 3.29 | 3.85 |
| QL | inf | 0.35 | 0.28 | 0.69 | 0.57 | 0.33 | 0.23 | 1.65 | 0.90 |
| Eff | 0.20 | 1.00 | 0.84 | 1.15 | 1.17 | 1.49 | 1.24 | 1.18 | 1.02 |
| Panel C: $\mu = 0.001$ | | | | | | | | | |
| Mean | 10.10 | 8.79 | 9.84 | 6.24 | 6.86 | 8.28 | 9.71 | 8.24 | 9.62 |
| abs.err | 0.10 | 1.21 | 0.16 | 3.76 | 3.14 | 1.72 | 0.29 | 1.76 | 0.38 |
| std.err | 1.49 | 0.61 | 0.68 | 0.43 | 0.47 | 0.47 | 0.54 | 0.55 | 0.62 |
| MSE | 20.45 | 3.82 | 4.61 | 3.27 | 3.23 | 2.53 | 3.08 | 3.26 | 3.80 |
| QL | inf | 0.35 | 0.28 | 0.69 | 0.57 | 0.32 | 0.23 | 1.45 | 0.89 |
| Eff | 0.20 | 1.00 | 0.84 | 1.16 | 1.17 | 1.52 | 1.26 | 1.20 | 1.04 |
| Panel D: $\mu = 0.005$ | | | | | | | | | |
| Mean | 12.53 | 9.65 | 10.81 | 6.86 | 7.54 | 8.54 | 10.07 | 8.17 | 9.60 |
| abs.err | 2.53 | 0.35 | 0.81 | 3.14 | 2.46 | 1.46 | 0.07 | 1.83 | 0.40 |
| std.err | 1.79 | 0.71 | 0.79 | 0.50 | 0.55 | 0.47 | 0.56 | 0.53 | 0.60 |
| MSE | 31.03 | 4.97 | 6.29 | 3.49 | 3.63 | 2.63 | 3.36 | 3.32 | 3.87 |
| QL | inf | 0.33 | 0.28 | 0.64 | 0.53 | 0.31 | 0.22 | 1.34 | 0.93 |
| Eff | 0.17 | 1.00 | 0.80 | 1.40 | 1.35 | 1.90 | 1.49 | 1.53 | 1.32 |
| Panel E: $\mu = 0.01$ | | | | | | | | | |
| Mean | 20.02 | 12.35 | 13.83 | 8.78 | 9.65 | 9.39 | 11.20 | 8.01 | 9.58 |
| abs.err | 10.02 | 2.35 | 3.83 | 1.22 | 0.35 | 0.61 | 1.20 | 1.99 | 0.42 |
| std.err | 2.41 | 0.91 | 1.02 | 0.65 | 0.71 | 0.53 | 0.63 | 0.56 | 0.63 |
| MSE | 69.82 | 9.34 | 12.48 | 4.59 | 5.37 | 3.06 | 4.47 | 3.61 | 4.23 |
| QL | inf | 0.30 | 0.27 | 0.50 | 0.42 | 0.28 | 0.21 | 2.11 | 1.25 |
| Eff | 0.14 | 1.00 | 0.75 | 2.02 | 1.73 | 3.08 | 2.11 | 2.64 | 2.29 |

Table B.6: The effect of drift on range-based estimators with constant volatility.

| daily variance $10^5 \times \sigma^2 = 40$ and intraday movements $k = 100$ | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | $r_i^2$ | P | $P_k$ | BL | $BL_k$ | GK | $GK_k$ | RS | $RS_k$ |
| Panel A: $\mu = 0$ | | | | | | | | | |
| Mean | 39.9 | 35.0 | 39.1 | 24.8 | 27.3 | 33.0 | 38.7 | 33.0 | 38.5 |
| abs.err | 0.1 | 5.0 | 0.9 | 15.2 | 12.7 | 7.0 | 1.3 | 7.0 | 1.5 |
| std.err | 5.46 | 2.35 | 2.63 | 1.67 | 1.83 | 1.87 | 2.18 | 2.13 | 2.41 |
| MSE | 315 | 60 | 72 | 52 | 51 | 40 | 49 | 52 | 60 |
| QL | inf | 0.35 | 0.28 | 0.69 | 0.57 | 0.32 | 0.23 | 1.27 | 0.90 |
| Eff | 0.20 | 1.00 | 0.84 | 1.14 | 1.16 | 1.50 | 1.25 | 1.18 | 1.03 |
| Panel B: $\mu = 0.0005$ | | | | | | | | | |
| Mean | 39.9 | 35.0 | 39.2 | 24.8 | 27.3 | 33.1 | 38.8 | 33.0 | 38.5 |
| abs.err | 0.1 | 5.0 | 0.8 | 15.2 | 12.7 | 6.9 | 1.2 | 7.0 | 1.5 |
| std.err | 5.75 | 2.44 | 2.74 | 1.74 | 1.91 | 1.93 | 2.26 | 2.22 | 2.52 |
| MSE | 315 | 60 | 72 | 52 | 51 | 40 | 49 | 52 | 61 |
| QL | inf | 0.35 | 0.28 | 0.69 | 0.57 | 0.32 | 0.23 | 1.28 | 0.86 |
| Eff | 0.20 | 1.00 | 0.84 | 1.14 | 1.16 | 1.49 | 1.24 | 1.18 | 1.02 |
| Panel C: $\mu = 0.001$ | | | | | | | | | |
| Mean | 40.3 | 35.1 | 39.3 | 25.0 | 27.4 | 33.2 | 38.9 | 33.1 | 38.6 |
| abs.err | 0.3 | 4.9 | 0.7 | 15.0 | 12.6 | 6.8 | 1.1 | 6.9 | 1.4 |
| std.err | 5.70 | 2.44 | 2.74 | 1.74 | 1.91 | 1.88 | 2.21 | 2.14 | 2.43 |
| MSE | 323 | 61 | 73 | 52 | 51 | 40 | 49 | 52 | 61 |
| QL | inf | 0.34 | 0.28 | 0.68 | 0.57 | 0.32 | 0.23 | 1.26 | 0.88 |
| Eff | 0.20 | 1.00 | 0.84 | 1.15 | 1.17 | 1.52 | 1.26 | 1.20 | 1.04 |
| Panel D: $\mu = 0.005$ | | | | | | | | | |
| Mean | 42.4 | 35.9 | 40.2 | 25.5 | 28.1 | 33.4 | 39.2 | 33.0 | 38.5 |
| abs.err | 2.4 | 4.1 | 0.2 | 14.5 | 11.9 | 6.6 | 0.8 | 7.0 | 1.5 |
| std.err | 5.95 | 2.50 | 2.80 | 1.77 | 1.95 | 1.86 | 2.18 | 2.09 | 2.38 |
| MSE | 361 | 65 | 79 | 53 | 53 | 41 | 50 | 52 | 61 |
| QL | inf | 0.34 | 0.28 | 0.67 | 0.56 | 0.32 | 0.23 | 7.05 | 0.91 |
| Eff | 0.19 | 1.00 | 0.83 | 1.21 | 1.21 | 1.60 | 1.31 | 1.27 | 1.10 |
| Panel E: $\mu = 0.01$ | | | | | | | | | |
| Mean | 49.9 | 38.6 | 43.2 | 27.4 | 30.2 | 34.3 | 40.4 | 32.8 | 38.5 |
| abs.err | 9.9 | 1.4 | 3.2 | 12.6 | 9.8 | 5.7 | 0.4 | 7.2 | 1.5 |
| std.err | 6.95 | 2.84 | 3.18 | 2.02 | 2.22 | 2.00 | 2.36 | 2.21 | 2.52 |
| MSE | 491 | 79 | 100 | 56 | 58 | 43 | 55 | 53 | 63 |
| QL | inf | 0.33 | 0.27 | 0.63 | 0.53 | 0.31 | 0.22 | 1.27 | 0.98 |
| Eff | 0.17 | 1.00 | 0.80 | 1.40 | 1.35 | 1.88 | 1.47 | 1.52 | 1.31 |

| daily variance $10^5 \times \sigma^2 = 90$ and intraday movements $k = 100$ | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | $r_i^2$ | P | $P_k$ | BL | $BL_k$ | GK | $GK_k$ | RS | $RS_k$ |
| Panel A: $\mu = 0$ | | | | | | | | | |
| Mean | 90.7 | 79.0 | 88.5 | 56.1 | 61.7 | 74.5 | 87.4 | 74.3 | 86.7 |
| abs.err | 0.7 | 11.0 | 1.5 | 33.9 | 28.3 | 15.5 | 2.6 | 15.7 | 3.3 |
| std.err | 12.74 | 5.27 | 5.91 | 3.75 | 4.12 | 4.14 | 4.84 | 4.92 | 5.55 |
| MSE | 1631 | 307 | 370 | 263 | 260 | 204 | 248 | 263 | 307 |
| QL | inf | 0.34 | 0.28 | 0.68 | 0.57 | 0.32 | 0.23 | 1.28 | 0.89 |
| Eff | 0.20 | 1.00 | 0.84 | 1.15 | 1.17 | 1.51 | 1.25 | 1.19 | 1.03 |
| Panel B: $\mu = 0.0005$ | | | | | | | | | |
| Mean | 90.7 | 79.1 | 88.6 | 56.2 | 61.8 | 74.6 | 87.5 | 74.4 | 86.8 |
| abs.err | 0.7 | 10.9 | 1.4 | 33.8 | 28.2 | 15.4 | 2.5 | 15.6 | 3.2 |
| std.err | 12.81 | 5.45 | 6.10 | 3.87 | 4.26 | 4.28 | 5.01 | 4.92 | 5.58 |
| MSE | 1649 | 309 | 373 | 264 | 261 | 205 | 250 | 265 | 310 |
| QL | inf | 0.35 | 0.28 | 0.69 | 0.57 | 0.32 | 0.23 | 2.16 | 0.88 |
| Eff | 0.20 | 1.00 | 0.84 | 1.16 | 1.17 | 1.52 | 1.26 | 1.19 | 1.04 |
| Panel C: $\mu = 0.001$ | | | | | | | | | |
| Mean | 90.0 | 79.0 | 88.4 | 56.1 | 61.7 | 74.7 | 87.5 | 74.4 | 86.9 |
| abs.err | 0.0 | 11.0 | 1.6 | 33.9 | 28.3 | 15.3 | 2.5 | 15.6 | 3.1 |
| std.err | 12.57 | 5.39 | 6.04 | 3.83 | 4.21 | 4.16 | 4.88 | 4.70 | 5.34 |
| MSE | 1614 | 308 | 371 | 264 | 261 | 206 | 251 | 263 | 308 |
| QL | 67551524.88 | 0.35 | 0.28 | 0.69 | 0.57 | 0.32 | 0.23 | 1.87 | 0.90 |
| Eff | 0.20 | 1.00 | 0.84 | 1.15 | 1.17 | 1.50 | 1.24 | 1.19 | 1.03 |
| Panel D: $\mu = 0.005$ | | | | | | | | | |
| Mean | 92.3 | 79.6 | 89.1 | 56.5 | 62.2 | 74.7 | 87.6 | 74.2 | 86.7 |
| abs.err | 2.3 | 10.4 | 0.9 | 33.5 | 27.8 | 15.3 | 2.4 | 15.8 | 3.3 |
| std.err | 12.99 | 5.36 | 6.00 | 3.81 | 4.18 | 4.04 | 4.73 | 4.75 | 5.37 |
| MSE | 1726 | 318 | 386 | 267 | 265 | 205 | 250 | 263 | 307 |
| QL | inf | 0.35 | 0.28 | 0.68 | 0.57 | 0.32 | 0.23 | 1.73 | 0.90 |
| Eff | 0.19 | 1.00 | 0.83 | 1.18 | 1.19 | 1.57 | 1.29 | 1.24 | 1.08 |
| Panel E: $\mu = 0.01$ | | | | | | | | | |
| Mean | 99.5 | 82.3 | 92.1 | 58.5 | 64.3 | 75.6 | 88.8 | 74.1 | 86.6 |
| abs.err | 9.5 | 7.7 | 2.1 | 31.5 | 25.7 | 14.4 | 1.2 | 15.9 | 3.4 |
| std.err | 13.98 | 5.71 | 6.39 | 4.06 | 4.46 | 4.19 | 4.92 | 4.77 | 5.41 |
| MSE | 1962 | 343 | 423 | 270 | 272 | 208 | 259 | 266 | 311 |
| QL | inf | 0.34 | 0.28 | 0.66 | 0.55 | 0.32 | 0.23 | 1.52 | 0.96 |
| Eff | 0.18 | 1.00 | 0.82 | 1.26 | 1.24 | 1.66 | 1.34 | 1.32 | 1.14 |

Table B.7: The average variance estimates with 95% confidence interval.

| | daily variance $10^5 \times \sigma^2 = 2.5$ | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| $k$ | $r_i^2$ | P | $P_k$ | BL | $BL_k$ | GK | $GK_k$ | RS | $RS_k$ |
| Panel A: $\mu = 0$ | | | | | | | | | |
| 20 | 2.50 | 1.86 | 2.43 | 1.33 | 1.65 | 1.62 | 2.39 | 1.59 | 2.32 |
| | (0.70) | (0.29) | (0.37) | (0.20) | (0.25) | (0.21) | (0.30) | (0.24) | (0.33) |
| 40 | 2.50 | 2.03 | 2.44 | 1.44 | 1.68 | 1.85 | 2.40 | 1.84 | 2.37 |
| | (0.66) | (0.27) | (0.33) | (0.19) | (0.23) | (0.21) | (0.27) | (0.25) | (0.31) |
| 100 | 2.51 | 2.19 | 2.46 | 1.56 | 1.71 | 2.07 | 2.43 | 2.07 | 2.41 |
| | (0.69) | (0.29) | (0.33) | (0.21) | (0.23) | (0.22) | (0.26) | (0.26) | (0.29) |
| Panel B: $\mu = 0.0005$ | | | | | | | | | |
| 20 | 2.51 | 1.87 | 2.44 | 1.33 | 1.66 | 1.62 | 2.39 | 1.60 | 2.32 |
| | (0.69) | (0.28) | (0.36) | (0.20) | (0.25) | (0.21) | (0.31) | (0.26) | (0.34) |
| 40 | 2.52 | 2.03 | 2.44 | 1.44 | 1.68 | 1.84 | 2.40 | 1.83 | 2.36 |
| | (0.70) | (0.29) | (0.35) | (0.21) | (0.24) | (0.22) | (0.28) | (0.25) | (0.30) |
| 100 | 2.51 | 2.19 | 2.45 | 1.56 | 1.71 | 2.07 | 2.42 | 2.06 | 2.40 |
| | (0.69) | (0.29) | (0.33) | (0.21) | (0.23) | (0.24) | (0.28) | (0.28) | (0.31) |
| Panel C: $\mu = 0.001$ | | | | | | | | | |
| 20 | 2.60 | 1.90 | 2.48 | 1.35 | 1.68 | 1.63 | 2.41 | 1.59 | 2.32 |
| | (0.75) | (0.29) | (0.38) | (0.21) | (0.26) | (0.21) | (0.30) | (0.25) | (0.33) |
| 40 | 2.59 | 2.06 | 2.47 | 1.46 | 1.70 | 1.85 | 2.41 | 1.83 | 2.36 |
| | (0.73) | (0.30) | (0.36) | (0.21) | (0.25) | (0.22) | (0.28) | (0.25) | (0.31) |
| 100 | 2.63 | 2.24 | 2.50 | 1.59 | 1.75 | 2.08 | 2.45 | 2.06 | 2.41 |
| | (0.74) | (0.31) | (0.34) | (0.22) | (0.24) | (0.24) | (0.28) | (0.27) | (0.31) |
| Panel D: $\mu = 0.005$ | | | | | | | | | |
| 20 | 5.01 | 2.73 | 3.57 | 1.94 | 2.42 | 1.85 | 2.86 | 1.49 | 2.34 |
| | (1.22) | (0.45) | (0.59) | (0.32) | (0.40) | (0.24) | (0.37) | (0.25) | (0.34) |
| 40 | 5.00 | 2.90 | 3.49 | 2.06 | 2.40 | 2.10 | 2.80 | 1.74 | 2.35 |
| | (1.23) | (0.47) | (0.56) | (0.33) | (0.39) | (0.26) | (0.35) | (0.27) | (0.34) |
| 100 | 4.98 | 3.08 | 3.45 | 2.19 | 2.41 | 2.35 | 2.80 | 2.01 | 2.40 |
| | (1.25) | (0.47) | (0.53) | (0.33) | (0.37) | (0.27) | (0.32) | (0.28) | (0.32) |
| Panel E: $\mu = 0.01$ | | | | | | | | | |
| 20 | 12.47 | 5.31 | 6.93 | 3.78 | 4.71 | 2.55 | 4.26 | 1.24 | 2.41 |
| | (2.14) | (0.76) | (0.99) | (0.54) | (0.67) | (0.31) | (0.52) | (0.27) | (0.38) |
| 40 | 12.55 | 5.55 | 6.66 | 3.94 | 4.59 | 2.84 | 4.01 | 1.54 | 2.34 |
| | (2.10) | (0.75) | (0.91) | (0.54) | (0.62) | (0.33) | (0.46) | (0.29) | (0.37) |
| 100 | 12.47 | 5.74 | 6.43 | 4.08 | 4.49 | 3.14 | 3.87 | 1.85 | 2.37 |
| | (2.05) | (0.76) | (0.85) | (0.54) | (0.59) | (0.35) | (0.43) | (0.30) | (0.35) |

| | daily variance $10^5 \times \sigma^2 = 10$ | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| $k$ | $r_i^2$ | P | $P_k$ | BL | $BL_k$ | GK | $GK_k$ | RS | $RS_k$ |
| Panel A: $\mu = 0$ | | | | | | | | | |
| 20 | 9.97 | 7.46 | 9.73 | 5.30 | 6.61 | 6.49 | 9.57 | 6.41 | 9.32 |
| | (2.73) | (1.06) | (1.39) | (0.76) | (0.94) | (0.78) | (1.14) | (0.95) | (1.26) |
| 40 | 9.94 | 8.09 | 9.72 | 5.75 | 6.70 | 7.38 | 9.59 | 7.34 | 9.45 |
| | (2.77) | (1.13) | (1.35) | (0.80) | (0.93) | (0.85) | (1.10) | (1.01) | (1.23) |
| 100 | 10.01 | 8.75 | 9.80 | 6.22 | 6.84 | 8.26 | 9.69 | 8.24 | 9.62 |
| | (2.77) | (1.17) | (1.31) | (0.83) | (0.91) | (0.92) | (1.08) | (1.06) | (1.21) |
| Panel B: $\mu = 0.0005$ | | | | | | | | | |
| 20 | 10.04 | 7.48 | 9.75 | 5.31 | 6.63 | 6.49 | 9.57 | 6.39 | 9.30 |
| | (2.83) | (1.14) | (1.49) | (0.81) | (1.01) | (0.83) | (1.22) | (0.98) | (1.32) |
| 40 | 10.05 | 8.14 | 9.78 | 5.78 | 6.74 | 7.40 | 9.62 | 7.36 | 9.48 |
| | (2.77) | (1.12) | (1.35) | (0.80) | (0.93) | (0.87) | (1.12) | (1.04) | (1.26) |
| 100 | 10.07 | 8.78 | 9.83 | 6.24 | 6.86 | 8.28 | 9.71 | 8.25 | 9.63 |
| | (2.79) | (1.14) | (1.27) | (0.81) | (0.89) | (0.89) | (1.04) | (1.06) | (1.19) |
| Panel C: $\mu = 0.001$ | | | | | | | | | |
| 20 | 10.01 | 7.45 | 9.72 | 5.29 | 6.60 | 6.46 | 9.53 | 6.37 | 9.27 |
| | (2.77) | (1.12) | (1.46) | (0.79) | (0.99) | (0.83) | (1.22) | (0.99) | (1.34) |
| 40 | 10.08 | 8.15 | 9.79 | 5.79 | 6.75 | 7.40 | 9.62 | 7.35 | 9.47 |
| | (2.82) | (1.15) | (1.38) | (0.82) | (0.95) | (0.87) | (1.13) | (1.02) | (1.25) |
| 100 | 10.10 | 8.79 | 9.84 | 6.24 | 6.86 | 8.28 | 9.71 | 8.24 | 9.62 |
| | (2.91) | (1.19) | (1.33) | (0.85) | (0.93) | (0.91) | (1.07) | (1.08) | (1.21) |
| Panel D: $\mu = 0.005$ | | | | | | | | | |
| 20 | 12.52 | 8.33 | 10.87 | 5.92 | 7.38 | 6.71 | 10.02 | 6.26 | 9.29 |
| | (3.38) | (1.28) | (1.67) | (0.91) | (1.13) | (0.84) | (1.25) | (1.00) | (1.34) |
| 40 | 12.52 | 9.00 | 10.81 | 6.40 | 7.45 | 7.64 | 10.01 | 7.25 | 9.45 |
| | (3.44) | (1.32) | (1.59) | (0.94) | (1.09) | (0.89) | (1.17) | (1.06) | (1.29) |
| 100 | 12.53 | 9.65 | 10.81 | 6.86 | 7.54 | 8.54 | 10.07 | 8.17 | 9.60 |
| | (3.52) | (1.38) | (1.55) | (0.98) | (1.08) | (0.93) | (1.09) | (1.04) | (1.18) |
| Panel E: $\mu = 0.01$ | | | | | | | | | |
| 20 | 20.14 | 10.96 | 14.30 | 7.79 | 9.72 | 7.42 | 11.45 | 5.95 | 9.34 |
| | (4.78) | (1.71) | (2.23) | (1.21) | (1.51) | (0.88) | (1.37) | (0.99) | (1.33) |
| 40 | 19.89 | 11.61 | 13.94 | 8.25 | 9.61 | 8.41 | 11.25 | 7.02 | 9.45 |
| | (4.83) | (1.87) | (2.24) | (1.33) | (1.54) | (1.09) | (1.46) | (1.12) | (1.40) |
| 100 | 20.02 | 12.35 | 13.83 | 8.78 | 9.65 | 9.39 | 11.20 | 8.01 | 9.58 |
| | (4.72) | (1.78) | (1.99) | (1.27) | (1.39) | (1.03) | (1.24) | (1.09) | (1.24) |

Table B.8: The average variance estimates with 95% confidence interval.

| | daily variance $10^5 \times \sigma^2 = 40$ | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| $k$ | $r_i^2$ | P | $P_k$ | BL | $BL_k$ | GK | $GK_k$ | RS | $RS_k$ |
| Panel A: $\mu = 0$ | | | | | | | | | |
| 20 | 40.1 | 29.9 | 39.0 | 21.3 | 26.5 | 26.0 | 38.3 | 25.6 | 37.3 |
| | (10.5) | (4.3) | (5.6) | (3.1) | (3.8) | (3.3) | (4.8) | (3.9) | (5.2) |
| 40 | 40.0 | 32.5 | 39.0 | 23.1 | 26.9 | 29.5 | 38.4 | 29.4 | 37.8 |
| | (11.3) | (4.8) | (5.7) | (3.4) | (3.9) | (3.5) | (4.6) | (4.0) | (4.9) |
| 100 | 39.9 | 35.0 | 39.1 | 24.8 | 27.3 | 33.0 | 38.7 | 33.0 | 38.5 |
| | (10.7) | (4.6) | (5.2) | (3.3) | (3.6) | (3.7) | (4.3) | (4.2) | (4.7) |
| Panel B: $\mu = 0.0005$ | | | | | | | | | |
| 20 | 40.1 | 29.9 | 39.0 | 21.3 | 26.5 | 26.0 | 38.4 | 25.7 | 37.3 |
| | (11.4) | (4.4) | (5.8) | (3.2) | (3.9) | (3.3) | (4.7) | (4.0) | (5.3) |
| 40 | 40.2 | 32.5 | 39.0 | 23.1 | 26.9 | 29.5 | 38.4 | 29.3 | 37.8 |
| | (11.0) | (4.6) | (5.5) | (3.2) | (3.8) | (3.5) | (4.6) | (4.2) | (5.1) |
| 100 | 39.9 | 35.0 | 39.2 | 24.8 | 27.3 | 33.1 | 38.8 | 33.0 | 38.5 |
| | (11.3) | (4.8) | (5.4) | (3.4) | (3.7) | (3.8) | (4.4) | (4.3) | (4.9) |
| Panel C: $\mu = 0.001$ | | | | | | | | | |
| 20 | 40.2 | 29.9 | 39.0 | 21.2 | 26.5 | 25.9 | 38.2 | 25.5 | 37.2 |
| | (10.8) | (4.4) | (5.7) | (3.1) | (3.9) | (3.3) | (4.8) | (3.9) | (5.2) |
| 40 | 40.2 | 32.5 | 39.1 | 23.1 | 26.9 | 29.6 | 38.4 | 29.4 | 37.8 |
| | (11.2) | (4.7) | (5.6) | (3.3) | (3.9) | (3.5) | (4.6) | (4.1) | (5.0) |
| 100 | 40.3 | 35.1 | 39.3 | 25.0 | 27.4 | 33.2 | 38.9 | 33.1 | 38.6 |
| | (11.2) | (4.8) | (5.4) | (3.4) | (3.7) | (3.7) | (4.3) | (4.2) | (4.8) |
| Panel D: $\mu = 0.005$ | | | | | | | | | |
| 20 | 42.2 | 30.6 | 39.9 | 21.8 | 27.1 | 26.1 | 38.6 | 25.4 | 37.2 |
| | (11.3) | (4.7) | (6.1) | (3.3) | (4.1) | (3.4) | (5.0) | (4.0) | (5.4) |
| 40 | 42.6 | 33.4 | 40.1 | 23.7 | 27.6 | 29.8 | 38.8 | 29.3 | 37.9 |
| | (11.7) | (4.9) | (5.8) | (3.5) | (4.0) | (3.5) | (4.6) | (4.0) | (4.9) |
| 100 | 42.4 | 35.9 | 40.2 | 25.5 | 28.1 | 33.4 | 39.2 | 33.0 | 38.5 |
| | (11.7) | (4.9) | (5.5) | (3.5) | (3.8) | (3.6) | (4.3) | (4.1) | (4.7) |
| Panel E: $\mu = 0.01$ | | | | | | | | | |
| 20 | 50.1 | 33.3 | 43.4 | 23.7 | 29.5 | 26.8 | 40.1 | 25.1 | 37.2 |
| | (13.3) | (5.1) | (6.6) | (3.6) | (4.5) | (3.4) | (5.1) | (4.0) | (5.4) |
| 40 | 50.0 | 35.9 | 43.2 | 25.5 | 29.8 | 30.5 | 40.0 | 28.9 | 37.7 |
| | (13.7) | (5.4) | (6.4) | (3.8) | (4.4) | (3.6) | (4.8) | (4.2) | (5.1) |
| 100 | 49.9 | 38.6 | 43.2 | 27.4 | 30.2 | 34.3 | 40.4 | 32.8 | 38.5 |
| | (13.6) | (5.6) | (6.2) | (4.0) | (4.4) | (3.9) | (4.6) | (4.3) | (4.9) |

| | daily variance $10^5 \times \sigma^2 = 90$ | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| $k$ | $r_i^2$ | P | $P_k$ | BL | $BL_k$ | GK | $GK_k$ | RS | $RS_k$ |
| Panel A: $\mu = 0$ | | | | | | | | | |
| 20 | 90.1 | 67.1 | 87.5 | 47.7 | 59.5 | 58.2 | 85.9 | 57.4 | 83.6 |
| | (24.9) | (9.7) | (12.7) | (6.9) | (8.6) | (7.2) | (10.4) | (8.9) | (11.8) |
| 40 | 90.2 | 73.1 | 87.7 | 51.9 | 60.5 | 66.4 | 86.3 | 66.0 | 85.0 |
| | (24.6) | (10.1) | (12.1) | (7.1) | (8.3) | (7.6) | (9.9) | (9.1) | (11.1) |
| 100 | 90.7 | 79.0 | 88.5 | 56.1 | 61.7 | 74.5 | 87.4 | 74.3 | 86.7 |
| | (25.0) | (10.3) | (11.6) | (7.3) | (8.1) | (8.1) | (9.5) | (9.6) | (10.9) |
| Panel B: $\mu = 0.0005$ | | | | | | | | | |
| 20 | 90.2 | 67.3 | 87.9 | 47.8 | 59.7 | 58.5 | 86.3 | 57.7 | 84.0 |
| | (23.8) | (9.8) | (12.7) | (6.9) | (8.6) | (7.4) | (10.8) | (8.9) | (12.0) |
| 40 | 90.2 | 73.1 | 87.8 | 52.0 | 60.5 | 66.5 | 86.5 | 66.1 | 85.2 |
| | (24.6) | (10.2) | (12.3) | (7.3) | (8.5) | (8.0) | (10.3) | (9.5) | (11.6) |
| 100 | 90.7 | 79.1 | 88.6 | 56.2 | 61.8 | 74.6 | 87.5 | 74.4 | 86.8 |
| | (25.1) | (10.7) | (12.0) | (7.6) | (8.3) | (8.4) | (9.8) | (9.6) | (10.9) |
| Panel C: $\mu = 0.001$ | | | | | | | | | |
| 20 | 90.1 | 67.1 | 87.6 | 47.7 | 59.5 | 58.3 | 86.0 | 57.4 | 83.7 |
| | (25.5) | (10.4) | (13.5) | (7.4) | (9.2) | (7.4) | (11.0) | (8.7) | (11.8) |
| 40 | 90.1 | 73.0 | 87.6 | 51.8 | 60.4 | 66.3 | 86.2 | 65.9 | 84.9 |
| | (25.6) | (10.5) | (12.6) | (7.5) | (8.7) | (7.9) | (10.3) | (9.4) | (11.4) |
| 100 | 90.0 | 79.0 | 88.4 | 56.1 | 61.7 | 74.7 | 87.5 | 74.4 | 86.9 |
| | (24.6) | (10.6) | (11.8) | (7.5) | (8.3) | (8.2) | (9.6) | (9.2) | (10.5) |
| Panel D: $\mu = 0.005$ | | | | | | | | | |
| 20 | 92.0 | 67.8 | 88.5 | 48.2 | 60.1 | 58.5 | 86.3 | 57.4 | 83.7 |
| | (25.8) | (10.5) | (13.7) | (7.5) | (9.3) | (7.5) | (11.0) | (8.5) | (11.6) |
| 40 | 92.0 | 73.8 | 88.7 | 52.5 | 61.1 | 66.8 | 86.9 | 66.1 | 85.2 |
| | (26.4) | (10.6) | (12.8) | (7.6) | (8.8) | (7.9) | (10.2) | (9.2) | (11.3) |
| 100 | 92.3 | 79.6 | 89.1 | 56.5 | 62.2 | 74.7 | 87.6 | 74.2 | 86.7 |
| | (25.5) | (10.5) | (11.8) | (7.5) | (8.2) | (7.9) | (9.3) | (9.3) | (10.5) |
| Panel E: $\mu = 0.01$ | | | | | | | | | |
| 20 | 99.7 | 70.6 | 92.0 | 50.1 | 62.5 | 59.3 | 87.9 | 57.2 | 83.9 |
| | (26.6) | (10.5) | (13.8) | (7.5) | (9.3) | (7.7) | (11.2) | (9.1) | (12.2) |
| 40 | 100.9 | 76.8 | 92.3 | 54.6 | 63.6 | 67.6 | 88.2 | 65.7 | 85.1 |
| | (27.4) | (11.1) | (13.3) | (7.9) | (9.2) | (8.0) | (10.5) | (9.2) | (11.3) |
| 100 | 99.5 | 82.3 | 92.1 | 58.5 | 64.3 | 75.6 | 88.8 | 74.1 | 86.6 |
| | (27.4) | (11.2) | (12.5) | (7.9) | (8.7) | (8.2) | (9.6) | (9.4) | (10.6) |

Table B.9: The effect of drift on range-based estimators with stochastic volatility.

| kurtosis = 1 and intraday movements $k$= 20 | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | $r_i^2$ | P | $P_k$ | BL | $BL_k$ | GK | $GK_k$ | RS | $RS_k$ |
| Panel A: $\mu = 0$ | | | | | | | | | |
| Mean | 9.47 | 6.13 | 8.00 | 4.36 | 5.43 | 4.84 | 7.26 | 4.51 | 6.72 |
| abs.err | 0.06 | 3.40 | 1.53 | 5.17 | 4.10 | 4.69 | 2.27 | 5.02 | 2.81 |
| std.err | 1.67 | 0.72 | 0.94 | 0.51 | 0.64 | 0.49 | 0.73 | 0.53 | 0.73 |
| MSE | 85.61 | 5.52 | 6.93 | 5.68 | 5.41 | 4.91 | 4.70 | 5.89 | 4.60 |
| QL | inf | 1.21 | 0.79 | 2.07 | 1.47 | 1.48 | 0.73 | 7.47 | 1.629 |
| Eff | 0.10 | 1.00 | 0.84 | 0.97 | 1.01 | 1.13 | 1.20 | 0.95 | 1.27 |
| Panel B: $\mu = 0.0005$ | | | | | | | | | |
| Mean | 9.51 | 6.13 | 8.00 | 4.36 | 5.44 | 4.83 | 7.25 | 4.49 | 6.70 |
| abs.err | 0.03 | 3.35 | 1.48 | 5.12 | 4.05 | 4.65 | 2.23 | 4.99 | 2.78 |
| std.err | 1.70 | 0.73 | 0.95 | 0.52 | 0.64 | 0.49 | 0.74 | 0.53 | 0.73 |
| MSE | 85.61 | 5.55 | 7.01 | 5.68 | 5.43 | 4.88 | 4.65 | 5.90 | 4.60 |
| QL | inf | 1.19 | 0.78 | 2.04 | 1.45 | 1.45 | 0.71 | 9.42 | 1.58 |
| Eff | 0.10 | 1.00 | 0.83 | 0.97 | 1.01 | 1.14 | 1.21 | 0.95 | 1.27 |
| Panel C: $\mu = 0.001$ | | | | | | | | | |
| Mean | 9.59 | 6.18 | 8.06 | 4.39 | 5.47 | 4.86 | 7.29 | 4.51 | 6.73 |
| abs.err | 0.06 | 3.36 | 1.48 | 5.15 | 4.06 | 4.68 | 2.24 | 5.02 | 2.80 |
| std.err | 1.80 | 0.75 | 0.98 | 0.53 | 0.67 | 0.50 | 0.75 | 0.53 | 0.73 |
| MSE | 85.61 | 5.58 | 7.05 | 5.70 | 5.45 | 4.90 | 4.70 | 5.90 | 4.60 |
| QL | inf | 1.21 | 0.79 | 2.07 | 1.47 | 1.47 | 0.72 | 6.89 | 1.62 |
| Eff | 0.10 | 1.00 | 0.83 | 0.97 | 1.02 | 1.15 | 1.21 | 0.96 | 1.28 |
| Panel D: $\mu = 0.005$ | | | | | | | | | |
| Mean | 12.02 | 7.03 | 9.17 | 4.99 | 6.23 | 5.10 | 7.76 | 4.47 | 6.79 |
| abs.err | 2.51 | 2.48 | 0.34 | 4.52 | 3.28 | 4.41 | 1.75 | 5.04 | 2.72 |
| std.err | 1.98 | 0.79 | 1.03 | 0.56 | 0.70 | 0.48 | 0.73 | 0.51 | 0.70 |
| MSE | 85.61 | 6.20 | 8.79 | 5.67 | 5.77 | 4.81 | 4.88 | 6.14 | 4.60 |
| QL | inf | 1.02 | 0.67 | 1.73 | 1.23 | 1.31 | 0.63 | 8.04 | 1.65 |
| Eff | 0.11 | 1.00 | 0.74 | 1.08 | 1.06 | 1.30 | 1.29 | 1.02 | 1.41 |
| Panel E: $\mu = 0.01$ | | | | | | | | | |
| Mean | 19.53 | 9.64 | 12.57 | 6.85 | 8.54 | 5.82 | 9.19 | 4.27 | 6.91 |
| abs.err | 10.02 | 0.14 | 3.07 | 2.65 | 0.96 | 3.68 | 0.31 | 5.23 | 2.59 |
| std.err | 2.63 | 0.98 | 1.28 | 0.70 | 0.87 | 0.53 | 0.82 | 0.59 | 0.79 |
| MSE | 85.61 | 8.66 | 14.96 | 5.86 | 7.19 | 4.50 | 5.55 | 6.70 | 4.60 |
| QL | inf | 0.65 | 0.48 | 1.07 | 0.77 | 0.98 | 0.46 | 17.17 | 1.71 |
| Eff | 0.16 | 1.00 | 0.59 | 1.46 | 1.19 | 1.96 | 1.58 | 1.31 | 1.98 |

| kurtosis = 3 and intraday movements $k$= 20 | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | $r_i^2$ | P | $P_k$ | BL | $BL_k$ | GK | $GK_k$ | RS | $RS_k$ |
| Panel A: $\mu = 0$ | | | | | | | | | |
| Mean | 9.50 | 5.51 | 7.18 | 3.91 | 4.88 | 3.96 | 6.05 | 3.44 | 5.27 |
| abs.err | 0.02 | 4.01 | 2.33 | 5.60 | 4.63 | 5.55 | 3.46 | 6.07 | 4.24 |
| std.err | 2.21 | 0.93 | 1.21 | 0.66 | 0.82 | 0.55 | 0.85 | 0.51 | 0.74 |
| MSE | 85.61 | 9.88 | 11.75 | 10.08 | 9.73 | 9.58 | 8.48 | 11.39 | 4.60 |
| QL | inf | 1.86 | 1.26 | 3.04 | 2.22 | 2.32 | 1.21 | 13.92 | 2.475 |
| Eff | 0.18 | 1.00 | 0.90 | 0.97 | 1.00 | 1.05 | 1.18 | 0.89 | 2.25 |
| Panel B: $\mu = 0.0005$ | | | | | | | | | |
| Mean | 9.59 | 5.55 | 7.24 | 3.94 | 4.92 | 3.99 | 6.08 | 3.45 | 5.29 |
| abs.err | 0.05 | 4.00 | 2.31 | 5.60 | 4.63 | 5.56 | 3.46 | 6.10 | 4.26 |
| std.err | 2.12 | 0.90 | 1.18 | 0.64 | 0.80 | 0.55 | 0.84 | 0.53 | 0.76 |
| MSE | 85.61 | 9.96 | 11.83 | 10.16 | 9.80 | 9.60 | 8.46 | 11.43 | 4.60 |
| QL | inf | 1.85 | 1.25 | 3.01 | 2.20 | 2.29 | 1.20 | 13.17 | 2.44 |
| Eff | 0.18 | 1.00 | 0.90 | 0.97 | 1.01 | 1.06 | 1.19 | 0.89 | 2.27 |
| Panel C: $\mu = 0.001$ | | | | | | | | | |
| Mean | 9.46 | 5.48 | 7.14 | 3.89 | 4.85 | 3.94 | 6.01 | 3.41 | 5.23 |
| abs.err | 0.06 | 4.05 | 2.38 | 5.63 | 4.67 | 5.59 | 3.51 | 6.11 | 4.30 |
| std.err | 2.14 | 0.90 | 1.18 | 0.64 | 0.80 | 0.55 | 0.85 | 0.54 | 0.78 |
| MSE | 85.61 | 9.43 | 11.07 | 9.74 | 9.34 | 9.28 | 8.12 | 11.06 | 4.60 |
| QL | inf | 1.84 | 1.24 | 3.00 | 2.19 | 2.29 | 1.20 | 11.99 | 2.48 |
| Eff | 0.17 | 1.00 | 0.91 | 0.96 | 1.00 | 1.04 | 1.18 | 0.87 | 2.16 |
| Panel D: $\mu = 0.005$ | | | | | | | | | |
| Mean | 12.07 | 6.44 | 8.40 | 4.57 | 5.70 | 4.26 | 6.60 | 3.44 | 5.39 |
| abs.err | 2.48 | 3.16 | 1.19 | 5.02 | 3.89 | 5.33 | 2.99 | 6.15 | 4.20 |
| std.err | 2.36 | 0.95 | 1.23 | 0.67 | 0.84 | 0.54 | 0.84 | 0.54 | 0.76 |
| MSE | 85.61 | 10.15 | 13.08 | 9.73 | 9.71 | 9.11 | 8.22 | 11.27 | 4.60 |
| QL | inf | 1.38 | 0.95 | 2.26 | 1.64 | 1.93 | 0.97 | 18.03 | 2.62 |
| Eff | 0.18 | 1.00 | 0.83 | 1.04 | 1.03 | 1.14 | 1.26 | 0.91 | 2.33 |
| Panel E: $\mu = 0.01$ | | | | | | | | | |
| Mean | 19.52 | 9.13 | 11.92 | 6.49 | 8.09 | 5.12 | 8.21 | 3.47 | 5.77 |
| abs.err | 19.52 | 9.13 | 11.92 | 6.49 | 8.09 | 5.12 | 8.21 | 3.47 | 5.77 |
| std.err | 2.95 | 1.15 | 1.50 | 0.82 | 1.02 | 0.63 | 0.99 | 0.64 | 0.90 |
| MSE | 85.61 | 11.80 | 17.98 | 9.49 | 10.48 | 8.44 | 8.30 | 11.62 | 4.60 |
| QL | inf | 0.75 | 0.59 | 1.17 | 0.87 | 1.21 | 0.58 | 19.16 | 2.33 |
| Eff | 0.21 | 1.00 | 0.68 | 1.25 | 1.12 | 1.47 | 1.47 | 1.05 | 2.69 |

Table B.10: The effect of drift on range-based estimators with stochastic volatility.

| kurtosis = 5 and intraday movements $k$= 20 | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | $r_i^2$ | P | $P_k$ | BL | $BL_k$ | GK | $GK_k$ | RS | $RS_k$ |
| Panel A: $\mu = 0$ | | | | | | | | | |
| Mean | 9.60 | 5.26 | 6.86 | 3.74 | 4.66 | 3.59 | 5.53 | 2.96 | 4.62 |
| abs.err | 0.06 | 4.28 | 2.68 | 5.80 | 4.88 | 5.96 | 4.01 | 6.58 | 4.92 |
| std.err | 2.81 | 1.17 | 1.52 | 0.83 | 1.04 | 0.64 | 1.02 | 0.56 | 0.83 |
| MSE | 85.61 | 14.77 | 17.90 | 14.51 | 14.35 | 14.06 | 12.57 | 16.72 | 4.60 |
| QL | inf | 2.35 | 1.62 | 3.74 | 2.77 | 2.97 | 1.60 | 33.85 | 3.17 |
| Eff | 0.26 | 1.00 | 0.92 | 0.98 | 1.00 | 1.06 | 1.15 | 0.89 | 3.38 |
| Panel B: $\mu = 0.0005$ | | | | | | | | | |
| Mean | 9.44 | 5.17 | 6.74 | 3.67 | 4.58 | 3.52 | 5.43 | 2.90 | 4.53 |
| abs.err | 0.11 | 4.38 | 2.81 | 5.88 | 4.97 | 6.03 | 4.12 | 6.65 | 5.02 |
| std.err | 2.58 | 1.07 | 1.40 | 0.76 | 0.95 | 0.59 | 0.94 | 0.52 | 0.77 |
| MSE | 85.61 | 13.48 | 15.64 | 13.81 | 13.34 | 13.69 | 11.77 | 16.40 | 4.60 |
| QL | inf | 2.31 | 1.59 | 3.70 | 2.74 | 2.91 | 1.57 | 13.65 | 3.11 |
| Eff | 0.24 | 1.00 | 0.93 | 0.97 | 1.00 | 1.02 | 1.16 | 0.85 | 3.07 |
| Panel C: $\mu = 0.001$ | | | | | | | | | |
| Mean | 9.62 | 5.24 | 6.84 | 3.72 | 4.64 | 3.55 | 5.48 | 2.91 | 4.56 |
| abs.err | 0.06 | 4.31 | 2.72 | 5.83 | 4.91 | 6.00 | 4.07 | 6.64 | 5.00 |
| std.err | 2.55 | 1.06 | 1.39 | 0.76 | 0.94 | 0.60 | 0.94 | 0.54 | 0.79 |
| MSE | 85.61 | 14.48 | 16.89 | 14.69 | 14.28 | 14.46 | 12.57 | 17.17 | 4.60 |
| QL | inf | 2.33 | 1.61 | 3.71 | 2.75 | 2.93 | 1.58 | 17.71 | 3.13 |
| Eff | 0.26 | 1.00 | 0.93 | 0.97 | 1.00 | 1.03 | 1.16 | 0.86 | 3.33 |
| Panel D: $\mu = 0.005$ | | | | | | | | | |
| Mean | 12.12 | 6.18 | 8.06 | 4.39 | 5.48 | 3.89 | 6.09 | 2.98 | 4.76 |
| abs.err | 2.59 | 3.35 | 1.47 | 5.14 | 4.06 | 5.65 | 3.45 | 6.55 | 4.78 |
| std.err | 2.90 | 1.16 | 1.51 | 0.82 | 1.03 | 0.61 | 0.97 | 0.55 | 0.79 |
| MSE | 85.61 | 14.57 | 18.24 | 14.07 | 14.03 | 13.70 | 12.01 | 16.91 | 4.60 |
| QL | inf | 1.58 | 1.10 | 2.52 | 1.86 | 2.31 | 1.18 | 21.60 | 3.28 |
| Eff | 0.26 | 1.00 | 0.87 | 1.02 | 1.02 | 1.11 | 1.22 | 0.88 | 3.31 |
| Panel E: $\mu = 0.01$ | | | | | | | | | |
| Mean | 19.47 | 8.88 | 11.58 | 6.31 | 7.87 | 4.79 | 7.75 | 3.07 | 5.20 |
| abs.err | 9.96 | 0.63 | 2.07 | 3.20 | 1.64 | 4.72 | 1.76 | 6.44 | 4.31 |
| std.err | 3.15 | 1.25 | 1.63 | 0.89 | 1.11 | 0.70 | 1.09 | 0.69 | 0.98 |
| MSE | 85.61 | 15.62 | 22.22 | 13.37 | 14.27 | 12.58 | 11.82 | 16.58 | 4.60 |
| QL | inf | 0.84 | 0.68 | 1.24 | 0.95 | 1.34 | 0.68 | 48.74 | 2.67 |
| Eff | 0.28 | 1.00 | 0.73 | 1.18 | 1.09 | 1.34 | 1.38 | 0.98 | 3.56 |

| kurtosis = 7 and intraday movements $k$= 20 | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | $r_i^2$ | P | $P_k$ | BL | $BL_k$ | GK | $GK_k$ | RS | $RS_k$ |
| Panel A: $\mu = 0$ | | | | | | | | | |
| Mean | 9.59 | 5.04 | 6.57 | 3.58 | 4.46 | 3.28 | 5.10 | 2.58 | 4.11 |
| abs.err | 0.04 | 4.51 | 2.98 | 5.97 | 5.09 | 6.28 | 4.45 | 6.97 | 5.44 |
| std.err | 3.07 | 1.30 | 1.69 | 0.92 | 1.15 | 0.70 | 1.12 | 0.55 | 0.85 |
| MSE | 85.61 | 18.16 | 20.88 | 18.56 | 17.97 | 19.07 | 16.45 | 22.94 | 4.60 |
| QL | inf | 2.68 | 1.87 | 4.23 | 3.16 | 3.38 | 1.85 | 34.99 | 3.61 |
| Eff | 0.33 | 1.00 | 0.94 | 0.98 | 1.00 | 0.99 | 1.12 | 0.81 | 4.19 |
| Panel B: $\mu = 0.0005$ | | | | | | | | | |
| Mean | 9.52 | 4.98 | 6.50 | 3.54 | 4.42 | 3.23 | 5.04 | 2.53 | 4.04 |
| abs.err | 0.02 | 4.51 | 2.99 | 5.96 | 5.08 | 6.26 | 4.46 | 6.96 | 5.46 |
| std.err | 2.98 | 1.23 | 1.61 | 0.87 | 1.09 | 0.64 | 1.03 | 0.51 | 0.78 |
| MSE | 85.61 | 17.70 | 20.46 | 17.97 | 17.47 | 18.27 | 15.74 | 21.86 | 4.60 |
| QL | inf | 2.67 | 1.86 | 4.20 | 3.14 | 3.41 | 1.86 | 17.08 | 3.69 |
| Eff | 0.32 | 1.00 | 0.94 | 0.97 | 1.00 | 0.99 | 1.13 | 0.83 | 4.05 |
| Panel C: $\mu = 0.001$ | | | | | | | | | |
| Mean | 9.61 | 5.04 | 6.57 | 3.58 | 4.47 | 3.27 | 5.10 | 2.59 | 4.11 |
| abs.err | 0.10 | 4.47 | 2.93 | 5.93 | 5.04 | 6.23 | 4.41 | 6.92 | 5.40 |
| std.err | 2.82 | 1.16 | 1.52 | 0.83 | 1.03 | 0.63 | 1.01 | 0.57 | 0.84 |
| MSE | 85.61 | 17.66 | 20.19 | 18.10 | 17.51 | 18.49 | 15.94 | 22.26 | 4.60 |
| QL | inf | 2.68 | 1.87 | 4.22 | 3.16 | 3.39 | 1.86 | 19.76 | 3.66 |
| Eff | 0.32 | 1.00 | 0.93 | 0.97 | 1.00 | 1.00 | 1.13 | 0.83 | 4.03 |
| Panel D: $\mu = 0.005$ | | | | | | | | | |
| Mean | 11.86 | 5.88 | 7.68 | 4.18 | 5.21 | 3.58 | 5.64 | 2.64 | 4.27 |
| abs.err | 2.38 | 3.59 | 1.80 | 5.30 | 4.26 | 5.90 | 3.84 | 6.84 | 5.20 |
| std.err | 3.10 | 1.25 | 1.63 | 0.89 | 1.10 | 0.64 | 1.02 | 0.53 | 0.79 |
| MSE | 85.61 | 16.76 | 19.75 | 16.95 | 16.48 | 17.42 | 14.81 | 21.47 | 4.60 |
| QL | inf | 1.66 | 1.18 | 2.63 | 1.96 | 2.47 | 1.27 | 23.96 | 3.63 |
| Eff | 0.30 | 1.00 | 0.90 | 0.99 | 1.01 | 1.00 | 1.15 | 0.81 | 3.78 |
| Panel E: $\mu = 0.01$ | | | | | | | | | |
| Mean | 19.40 | 8.71 | 11.36 | 6.19 | 7.72 | 4.58 | 7.45 | 2.82 | 4.85 |
| abs.err | 9.85 | 0.83 | 1.82 | 3.35 | 1.82 | 4.96 | 2.09 | 6.72 | 4.69 |
| std.err | 3.44 | 1.36 | 1.77 | 0.97 | 1.20 | 0.73 | 1.14 | 0.67 | 0.97 |
| MSE | 85.61 | 19.54 | 26.44 | 17.52 | 18.22 | 17.24 | 15.72 | 22.26 | 4.60 |
| QL | inf | 0.90 | 0.75 | 1.29 | 1.01 | 1.41 | 0.74 | 25.37 | 2.85 |
| Eff | 0.35 | 1.00 | 0.78 | 1.13 | 1.07 | 1.25 | 1.32 | 0.92 | 4.46 |

Table B.11: The effect of drift on range-based estimators with stochastic volatility.

| kurtosis = 1 and intraday movements $k$= 40 | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | $r_i^2$ | P | $P_k$ | BL | $BL_k$ | GK | $GK_k$ | RS | $RS_k$ |
| Panel A: $\mu = 0$ | | | | | | | | | |
| Mean | 9.80 | 6.63 | 7.96 | 4.71 | 5.49 | 5.41 | 7.12 | 5.09 | 6.66 |
| abs.err | 0.03 | 3.13 | 1.80 | 5.05 | 4.28 | 4.36 | 2.64 | 4.67 | 3.10 |
| std.err | 1.64 | 0.72 | 0.86 | 0.51 | 0.59 | 0.51 | 0.67 | 0.55 | 0.68 |
| MSE | 85.61 | 5.68 | 6.64 | 5.70 | 5.50 | 4.74 | 4.49 | 5.72 | 4.60 |
| QL | inf | 0.99 | 0.73 | 1.72 | 1.35 | 1.12 | 0.68 | 6.19 | 1.76 |
| Eff | 0.10 | 1.00 | 0.88 | 0.99 | 1.02 | 1.21 | 1.28 | 1.01 | 1.30 |
| Panel B: $\mu = 0.0005$ | | | | | | | | | |
| Mean | 9.72 | 6.58 | 7.90 | 4.68 | 5.45 | 5.37 | 7.07 | 5.06 | 6.61 |
| abs.err | 0.02 | 3.16 | 1.83 | 5.06 | 4.29 | 4.37 | 2.67 | 4.68 | 3.12 |
| std.err | 1.71 | 0.75 | 0.90 | 0.54 | 0.62 | 0.53 | 0.70 | 0.56 | 0.70 |
| MSE | 85.61 | 5.62 | 6.52 | 5.69 | 5.47 | 4.73 | 4.44 | 5.73 | 4.60 |
| QL | inf | 0.99 | 0.73 | 1.72 | 1.35 | 1.11 | 0.67 | 6.76 | 1.73 |
| Eff | 0.10 | 1.00 | 0.89 | 0.98 | 1.01 | 1.19 | 1.27 | 0.99 | 1.28 |
| Panel C: $\mu = 0.001$ | | | | | | | | | |
| Mean | 9.93 | 6.68 | 8.03 | 4.75 | 5.53 | 5.43 | 7.15 | 5.10 | 6.68 |
| abs.err | 0.16 | 3.08 | 1.74 | 5.02 | 4.23 | 4.34 | 2.62 | 4.67 | 3.09 |
| std.err | 1.77 | 0.75 | 0.90 | 0.53 | 0.62 | 0.50 | 0.66 | 0.53 | 0.66 |
| MSE | 85.61 | 5.75 | 6.74 | 5.73 | 5.53 | 4.76 | 4.52 | 5.80 | 4.60 |
| QL | inf | 0.97 | 0.72 | 1.70 | 1.33 | 1.11 | 0.67 | 6.67 | 1.80 |
| Eff | 0.10 | 1.00 | 0.88 | 0.99 | 1.03 | 1.22 | 1.28 | 1.01 | 1.31 |
| Panel D: $\mu = 0.005$ | | | | | | | | | |
| Mean | 12.24 | 7.48 | 8.99 | 5.32 | 6.20 | 5.65 | 7.51 | 5.03 | 6.66 |
| abs.err | 2.49 | 2.26 | 0.76 | 4.43 | 3.55 | 4.10 | 2.24 | 4.71 | 3.08 |
| std.err | 1.93 | 0.79 | 0.95 | 0.56 | 0.66 | 0.52 | 0.69 | 0.56 | 0.70 |
| MSE | 85.61 | 6.25 | 7.85 | 5.66 | 5.66 | 4.67 | 4.57 | 5.97 | 4.60 |
| QL | inf | 0.82 | 0.62 | 1.44 | 1.12 | 1.01 | 0.60 | 20.29 | 2.05 |
| Eff | 0.11 | 1.00 | 0.82 | 1.09 | 1.09 | 1.35 | 1.38 | 1.06 | 1.42 |
| Panel E: $\mu = 0.01$ | | | | | | | | | |
| Mean | 19.84 | 10.19 | 12.23 | 7.24 | 8.43 | 6.46 | 8.79 | 4.93 | 6.76 |
| abs.err | 10.07 | 0.42 | 2.47 | 2.53 | 1.33 | 3.31 | 0.97 | 4.84 | 3.00 |
| std.err | 2.77 | 1.06 | 1.27 | 0.75 | 0.88 | 0.59 | 0.79 | 0.62 | 0.77 |
| MSE | 85.61 | 9.03 | 13.16 | 5.95 | 6.79 | 4.40 | 4.91 | 6.53 | 4.60 |
| QL | inf | 0.56 | 0.45 | 0.92 | 0.73 | 0.77 | 0.45 | 15.15 | 2.43 |
| Eff | 0.16 | 1.00 | 0.70 | 1.49 | 1.31 | 2.09 | 1.86 | 1.41 | 2.07 |

| kurtosis = 3 and intraday movements $k$= 40 | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | $r_i^2$ | P | $P_k$ | BL | $BL_k$ | GK | $GK_k$ | RS | $RS_k$ |
| Panel A: $\mu = 0$ | | | | | | | | | |
| Mean | 9.71 | 5.85 | 7.02 | 4.16 | 4.84 | 4.36 | 5.81 | 3.85 | 5.12 |
| abs.err | 0.06 | 3.92 | 2.75 | 5.61 | 4.93 | 5.41 | 3.96 | 5.92 | 4.65 |
| std.err | 2.10 | 0.92 | 1.10 | 0.65 | 0.76 | 0.58 | 0.78 | 0.56 | 0.72 |
| MSE | 85.61 | 9.70 | 10.76 | 9.90 | 9.57 | 9.21 | 8.27 | 10.95 | 4.60 |
| QL | inf | 1.59 | 1.21 | 2.62 | 2.10 | 1.86 | 1.19 | 10.21 | 2.78 |
| Eff | 0.17 | 1.00 | 0.93 | 0.97 | 1.00 | 1.08 | 1.19 | 0.91 | 2.23 |
| Panel B: $\mu = 0.0005$ | | | | | | | | | |
| Mean | 9.80 | 5.87 | 7.05 | 4.17 | 4.86 | 4.35 | 5.79 | 3.82 | 5.09 |
| abs.err | 0.10 | 3.84 | 2.66 | 5.54 | 4.85 | 5.36 | 3.91 | 5.89 | 4.62 |
| std.err | 2.20 | 0.94 | 1.13 | 0.67 | 0.78 | 0.59 | 0.79 | 0.58 | 0.74 |
| MSE | 85.61 | 10.05 | 11.17 | 10.19 | 9.88 | 9.43 | 8.44 | 11.19 | 4.60 |
| QL | inf | 1.57 | 1.19 | 2.60 | 2.08 | 1.85 | 1.18 | 11.53 | 2.77 |
| Eff | 0.18 | 1.00 | 0.94 | 0.97 | 1.00 | 1.09 | 1.20 | 0.92 | 2.30 |
| Panel C: $\mu = 0.001$ | | | | | | | | | |
| Mean | 9.82 | 5.89 | 7.07 | 4.18 | 4.88 | 4.37 | 5.83 | 3.85 | 5.13 |
| abs.err | 0.12 | 3.81 | 2.63 | 5.52 | 4.83 | 5.33 | 3.87 | 5.85 | 4.57 |
| std.err | 2.22 | 0.95 | 1.14 | 0.67 | 0.78 | 0.58 | 0.78 | 0.57 | 0.73 |
| MSE | 85.61 | 9.81 | 10.83 | 10.08 | 9.73 | 9.35 | 8.33 | 11.15 | 4.60 |
| QL | inf | 1.56 | 1.18 | 2.58 | 2.06 | 1.84 | 1.17 | 13.45 | 2.77 |
| Eff | 0.18 | 1.00 | 0.94 | 0.96 | 1.00 | 1.07 | 1.19 | 0.89 | 2.23 |
| Panel D: $\mu = 0.005$ | | | | | | | | | |
| Mean | 12.22 | 6.77 | 8.13 | 4.81 | 5.61 | 4.67 | 6.28 | 3.87 | 5.21 |
| abs.err | 2.47 | 2.98 | 1.62 | 4.94 | 4.15 | 5.09 | 3.48 | 5.88 | 4.54 |
| std.err | 2.38 | 0.97 | 1.17 | 0.69 | 0.80 | 0.57 | 0.77 | 0.56 | 0.71 |
| MSE | 85.61 | 10.63 | 12.41 | 10.18 | 10.06 | 9.28 | 8.40 | 11.36 | 4.60 |
| QL | inf | 1.20 | 0.92 | 1.99 | 1.59 | 1.59 | 0.98 | 14.67 | 3.43 |
| Eff | 0.19 | 1.00 | 0.89 | 1.03 | 1.04 | 1.17 | 1.27 | 0.96 | 2.43 |
| Panel E: $\mu = 0.01$ | | | | | | | | | |
| Mean | 19.96 | 9.59 | 11.52 | 6.81 | 7.94 | 5.58 | 7.70 | 3.91 | 5.46 |
| abs.err | 10.18 | 0.18 | 1.74 | 2.96 | 1.84 | 4.19 | 2.08 | 5.86 | 4.31 |
| std.err | 2.95 | 1.16 | 1.39 | 0.82 | 0.96 | 0.64 | 0.87 | 0.64 | 0.80 |
| MSE | 85.61 | 12.44 | 16.51 | 9.80 | 10.42 | 8.46 | 8.01 | 11.65 | 4.60 |
| QL | inf | 0.69 | 0.58 | 1.07 | 0.87 | 1.04 | 0.62 | 69.03 | 3.74 |
| Eff | 0.22 | 1.00 | 0.77 | 1.27 | 1.18 | 1.56 | 1.61 | 1.11 | 2.84 |

Table B.12: The effect of drift on range-based estimators with stochastic volatility.

| kurtosis = 5 and intraday movements $k$= 40 | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | $r_i^2$ | P | $P_k$ | BL | $BL_k$ | GK | $GK_k$ | RS | $RS_k$ |
| Panel A: $\mu = 0$ | | | | | | | | | |
| Mean | 9.85 | 5.53 | 6.64 | 3.93 | 4.58 | 3.86 | 5.19 | 3.23 | 4.36 |
| abs.err | 0.05 | 4.27 | 3.16 | 5.87 | 5.22 | 5.93 | 4.61 | 6.57 | 5.44 |
| std.err | 2.48 | 1.07 | 1.28 | 0.76 | 0.88 | 0.63 | 0.85 | 0.57 | 0.74 |
| MSE | 85.61 | 14.21 | 15.55 | 14.46 | 14.05 | 14.24 | 12.65 | 17.09 | 4.60 |
| QL | inf | 2.07 | 1.60 | 3.33 | 2.70 | 2.47 | 1.61 | 12.78 | 3.59 |
| Eff | 0.26 | 1.00 | 0.95 | 0.98 | 1.00 | 1.03 | 1.15 | 0.85 | 3.25 |
| Panel B: $\mu = 0.0005$ | | | | | | | | | |
| Mean | 9.64 | 5.44 | 6.54 | 3.87 | 4.51 | 3.82 | 5.13 | 3.21 | 4.33 |
| abs.err | 0.08 | 4.28 | 3.19 | 5.85 | 5.22 | 5.90 | 4.59 | 6.51 | 5.39 |
| std.err | 2.52 | 1.06 | 1.28 | 0.76 | 0.88 | 0.62 | 0.84 | 0.57 | 0.74 |
| MSE | 85.61 | 13.93 | 15.17 | 14.23 | 13.81 | 13.84 | 12.28 | 16.50 | 4.60 |
| QL | inf | 2.02 | 1.56 | 3.26 | 2.64 | 2.42 | 1.57 | 16.26 | 3.59 |
| Eff | 0.25 | 1.00 | 0.96 | 0.98 | 1.00 | 1.05 | 1.17 | 0.88 | 3.21 |
| Panel C: $\mu = 0.001$ | | | | | | | | | |
| Mean | 9.90 | 5.55 | 6.67 | 3.95 | 4.60 | 3.88 | 5.21 | 3.24 | 4.37 |
| abs.err | 0.09 | 4.26 | 3.14 | 5.87 | 5.21 | 5.94 | 4.61 | 6.58 | 5.44 |
| std.err | 2.58 | 1.08 | 1.30 | 0.77 | 0.90 | 0.62 | 0.84 | 0.56 | 0.73 |
| MSE | 85.61 | 14.71 | 16.22 | 14.75 | 14.42 | 14.32 | 12.87 | 16.92 | 4.60 |
| QL | inf | 1.98 | 1.52 | 3.19 | 2.58 | 2.38 | 1.55 | 13.82 | 3.60 |
| Eff | 0.27 | 1.00 | 0.96 | 0.98 | 1.00 | 1.04 | 1.15 | 0.89 | 3.38 |
| Panel D: $\mu = 0.005$ | | | | | | | | | |
| Mean | 12.34 | 6.49 | 7.79 | 4.61 | 5.37 | 4.23 | 5.73 | 3.33 | 4.54 |
| abs.err | 2.55 | 3.30 | 2.00 | 5.18 | 4.42 | 5.56 | 4.06 | 6.46 | 5.25 |
| std.err | 2.88 | 1.19 | 1.43 | 0.84 | 0.98 | 0.68 | 0.92 | 0.62 | 0.80 |
| MSE | 85.61 | 14.67 | 16.73 | 14.29 | 14.08 | 13.81 | 12.38 | 16.90 | 4.60 |
| QL | inf | 1.41 | 1.10 | 2.28 | 1.84 | 1.95 | 1.22 | 26.38 | 4.48 |
| Eff | 0.26 | 1.00 | 0.92 | 1.01 | 1.03 | 1.09 | 1.20 | 0.89 | 3.31 |
| Panel E: $\mu = 0.01$ | | | | | | | | | |
| Mean | 19.68 | 9.21 | 11.06 | 6.54 | 7.63 | 5.17 | 7.16 | 3.47 | 4.88 |
| abs.err | 9.96 | 0.51 | 1.34 | 3.18 | 2.10 | 4.56 | 2.56 | 6.26 | 4.85 |
| std.err | 3.26 | 1.28 | 1.54 | 0.91 | 1.06 | 0.70 | 0.95 | 0.67 | 0.85 |
| MSE | 85.61 | 15.78 | 19.65 | 13.66 | 14.03 | 12.71 | 11.60 | 16.75 | 4.60 |
| QL | inf | 0.77 | 0.66 | 1.14 | 0.95 | 1.18 | 0.73 | 51.81 | 4.56 |
| Eff | 0.29 | 1.00 | 0.82 | 1.16 | 1.12 | 1.33 | 1.42 | 0.99 | 3.59 |

| kurtosis = 7 and intraday movements $k$= 40 | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | $r_i^2$ | P | $P_k$ | BL | $BL_k$ | GK | $GK_k$ | RS | $RS_k$ |
| Panel A: $\mu = 0$ | | | | | | | | | |
| Mean | 9.79 | 5.27 | 6.33 | 3.75 | 4.36 | 3.53 | 4.77 | 2.83 | 3.87 |
| abs.err | 0.02 | 4.50 | 3.44 | 6.02 | 5.41 | 6.24 | 5.00 | 6.94 | 5.90 |
| std.err | 2.82 | 1.18 | 1.41 | 0.84 | 0.97 | 0.65 | 0.90 | 0.57 | 0.74 |
| MSE | 85.61 | 18.83 | 20.29 | 19.28 | 18.74 | 19.56 | 17.39 | 23.38 | 4.60 |
| QL | inf | 2.33 | 1.81 | 3.71 | 3.02 | 2.83 | 1.86 | 18.19 | 4.15 |
| Eff | 0.34 | 1.00 | 0.97 | 0.98 | 1.00 | 1.01 | 1.12 | 0.84 | 4.30 |
| Panel B: $\mu = 0.0005$ | | | | | | | | | |
| Mean | 9.84 | 5.30 | 6.36 | 3.76 | 4.39 | 3.54 | 4.79 | 2.84 | 3.88 |
| abs.err | 0.11 | 4.44 | 3.37 | 5.97 | 5.35 | 6.19 | 4.95 | 6.90 | 5.86 |
| std.err | 2.93 | 1.23 | 1.47 | 0.87 | 1.02 | 0.67 | 0.93 | 0.56 | 0.75 |
| MSE | 85.61 | 18.42 | 20.00 | 18.68 | 18.21 | 18.84 | 16.80 | 22.51 | 4.60 |
| QL | inf | 2.30 | 1.78 | 3.66 | 2.98 | 2.80 | 1.84 | 30.91 | 4.16 |
| Eff | 0.33 | 1.00 | 0.96 | 0.99 | 1.00 | 1.02 | 1.14 | 0.84 | 4.21 |
| Panel C: $\mu = 0.001$ | | | | | | | | | |
| Mean | 9.80 | 5.28 | 6.35 | 3.75 | 4.37 | 3.54 | 4.78 | 2.84 | 3.88 |
| abs.err | 0.08 | 4.44 | 3.38 | 5.97 | 5.35 | 6.18 | 4.94 | 6.88 | 5.84 |
| std.err | 2.80 | 1.16 | 1.39 | 0.82 | 0.96 | 0.63 | 0.86 | 0.54 | 0.71 |
| MSE | 85.61 | 17.92 | 19.32 | 18.33 | 17.82 | 18.55 | 16.50 | 22.18 | 4.60 |
| QL | inf | 2.29 | 1.78 | 3.64 | 2.97 | 2.81 | 1.85 | 22.34 | 4.32 |
| Eff | 0.32 | 1.00 | 0.97 | 0.97 | 1.00 | 1.00 | 1.12 | 0.83 | 4.09 |
| Panel D: $\mu = 0.005$ | | | | | | | | | |
| Mean | 12.08 | 6.16 | 7.39 | 4.37 | 5.10 | 3.87 | 5.27 | 2.94 | 4.04 |
| abs.err | 2.35 | 3.57 | 2.34 | 5.36 | 4.63 | 5.86 | 4.46 | 6.80 | 5.69 |
| std.err | 2.98 | 1.25 | 1.50 | 0.89 | 1.03 | 0.70 | 0.96 | 0.62 | 0.80 |
| MSE | 85.61 | 18.23 | 19.81 | 18.53 | 18.04 | 18.75 | 16.66 | 22.72 | 4.60 |
| QL | inf | 1.53 | 1.20 | 2.44 | 1.98 | 2.19 | 1.38 | 317.13 | 5.39 |
| Eff | 0.33 | 1.00 | 0.95 | 0.99 | 1.01 | 1.01 | 1.14 | 0.83 | 4.21 |
| Panel E: $\mu = 0.01$ | | | | | | | | | |
| Mean | 19.56 | 8.99 | 10.79 | 6.38 | 7.44 | 4.90 | 6.82 | 3.16 | 4.48 |
| abs.err | 9.76 | 0.81 | 1.00 | 3.41 | 2.36 | 4.89 | 2.97 | 6.63 | 5.31 |
| std.err | 3.23 | 1.29 | 1.55 | 0.92 | 1.07 | 0.74 | 1.00 | 0.72 | 0.91 |
| MSE | 85.61 | 19.46 | 23.32 | 17.61 | 17.83 | 16.93 | 15.34 | 21.63 | 4.60 |
| QL | inf | 0.82 | 0.72 | 1.17 | 0.98 | 1.23 | 0.77 | 39.81 | 4.97 |
| Eff | 0.35 | 1.00 | 0.85 | 1.13 | 1.10 | 1.26 | 1.37 | 0.96 | 4.43 |

Table B.13: The effect of drift on range-based estimators with stochastic volatility.

| kurtosis = 1 and intraday movements $k$= 100 | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | $r_i^2$ | P | $P_k$ | BL | $BL_k$ | GK | $GK_k$ | RS | $RS_k$ |
| Panel A: $\mu = 0$ | | | | | | | | | |
| Mean | 9.96 | 6.99 | 7.83 | 4.97 | 5.46 | 5.85 | 6.92 | 5.54 | 6.52 |
| abs.err | 0.06 | 2.91 | 2.07 | 4.93 | 4.44 | 4.05 | 2.98 | 4.36 | 3.37 |
| std.err | 1.68 | 0.75 | 0.84 | 0.53 | 0.58 | 0.53 | 0.63 | 0.55 | 0.64 |
| MSE | 85.61 | 5.68 | 6.20 | 5.68 | 5.52 | 4.63 | 4.41 | 5.56 | 4.60 |
| QL | inf | 0.85 | 0.70 | 1.51 | 1.29 | 0.91 | 0.66 | 5.81 | 2.02 |
| Eff | 0.10 | 1.00 | 0.93 | 0.99 | 1.02 | 1.24 | 1.30 | 1.04 | 1.30 |
| Panel B: $\mu = 0.0005$ | | | | | | | | | |
| Mean | 9.83 | 6.91 | 7.74 | 4.91 | 5.40 | 5.79 | 6.84 | 5.49 | 6.46 |
| abs.err | 0.07 | 2.99 | 2.16 | 4.99 | 4.50 | 4.11 | 3.05 | 4.41 | 3.43 |
| std.err | 1.68 | 0.73 | 0.82 | 0.52 | 0.57 | 0.53 | 0.63 | 0.58 | 0.67 |
| MSE | 85.61 | 5.63 | 6.14 | 5.62 | 5.47 | 4.57 | 4.36 | 5.51 | 4.60 |
| QL | inf | 0.85 | 0.70 | 1.50 | 1.29 | 0.91 | 0.67 | 6.40 | 2.05 |
| Eff | 0.10 | 1.00 | 0.93 | 0.99 | 1.02 | 1.25 | 1.30 | 1.04 | 1.29 |
| Panel C: $\mu = 0.001$ | | | | | | | | | |
| Mean | 9.95 | 6.97 | 7.81 | 4.95 | 5.45 | 5.82 | 6.89 | 5.52 | 6.50 |
| abs.err | 0.01 | 2.97 | 2.13 | 4.99 | 4.49 | 4.12 | 3.05 | 4.42 | 3.44 |
| std.err | 1.75 | 0.78 | 0.88 | 0.56 | 0.61 | 0.56 | 0.67 | 0.60 | 0.69 |
| MSE | 85.61 | 5.81 | 6.36 | 5.73 | 5.58 | 4.64 | 4.44 | 5.58 | 4.60 |
| QL | inf | 0.85 | 0.70 | 1.50 | 1.28 | 0.91 | 0.66 | 6.95 | 2.03 |
| Eff | 0.10 | 1.00 | 0.93 | 1.00 | 1.02 | 1.26 | 1.32 | 1.06 | 1.33 |
| Panel D: $\mu = 0.005$ | | | | | | | | | |
| Mean | 12.39 | 7.85 | 8.80 | 5.58 | 6.14 | 6.10 | 7.26 | 5.50 | 6.52 |
| abs.err | 2.49 | 2.05 | 1.11 | 4.32 | 3.77 | 3.80 | 2.65 | 4.40 | 3.38 |
| std.err | 1.93 | 0.80 | 0.90 | 0.57 | 0.63 | 0.54 | 0.64 | 0.58 | 0.67 |
| MSE | 85.61 | 6.35 | 7.27 | 5.67 | 5.64 | 4.58 | 4.46 | 5.83 | 4.60 |
| QL | inf | 0.73 | 0.61 | 1.28 | 1.10 | 0.84 | 0.61 | 7.89 | 2.64 |
| Eff | 0.11 | 1.00 | 0.89 | 1.11 | 1.11 | 1.41 | 1.44 | 1.11 | 1.46 |
| Panel E: $\mu = 0.01$ | | | | | | | | | |
| Mean | 19.99 | 10.58 | 11.84 | 7.52 | 8.26 | 6.94 | 8.38 | 5.43 | 6.56 |
| abs.err | 10.09 | 0.67 | 1.94 | 2.39 | 1.64 | 2.96 | 1.53 | 4.48 | 3.35 |
| std.err | 2.62 | 1.00 | 1.12 | 0.71 | 0.78 | 0.58 | 0.70 | 0.64 | 0.73 |
| MSE | 85.61 | 9.47 | 11.91 | 6.13 | 6.61 | 4.44 | 4.67 | 6.40 | 4.60 |
| QL | inf | 0.50 | 0.44 | 0.83 | 0.71 | 0.64 | 0.45 | 15.23 | 3.77 |
| Eff | 0.17 | 1.00 | 0.80 | 1.53 | 1.42 | 2.19 | 2.07 | 1.52 | 2.16 |

| kurtosis = 3 and intraday movements $k$= 100 | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | $r_i^2$ | P | $P_k$ | BL | $BL_k$ | GK | $GK_k$ | RS | $RS_k$ |
| Panel A: $\mu = 0$ | | | | | | | | | |
| Mean | 9.88 | 6.10 | 6.83 | 4.33 | 4.76 | 4.64 | 5.53 | 4.13 | 4.91 |
| abs.err | 0.03 | 3.74 | 3.01 | 5.51 | 5.08 | 5.20 | 4.31 | 5.71 | 4.93 |
| std.err | 2.26 | 0.98 | 1.10 | 0.70 | 0.77 | 0.61 | 0.74 | 0.59 | 0.68 |
| MSE | 85.61 | 10.13 | 10.73 | 10.23 | 10.00 | 9.26 | 8.54 | 11.00 | 4.60 |
| QL | inf | 1.41 | 1.18 | 2.35 | 2.04 | 1.59 | 1.20 | 11.96 | 3.36 |
| Eff | 0.18 | 1.00 | 0.97 | 0.98 | 1.00 | 1.12 | 1.20 | 0.95 | 2.31 |
| Panel B: $\mu = 0.0005$ | | | | | | | | | |
| Mean | 9.99 | 6.14 | 6.87 | 4.36 | 4.79 | 4.65 | 5.54 | 4.12 | 4.90 |
| abs.err | 0.09 | 3.77 | 3.03 | 5.54 | 5.11 | 5.26 | 4.36 | 5.79 | 5.00 |
| std.err | 2.24 | 0.96 | 1.07 | 0.68 | 0.75 | 0.59 | 0.71 | 0.56 | 0.65 |
| MSE | 85.61 | 10.20 | 10.82 | 10.27 | 10.04 | 9.34 | 8.63 | 11.06 | 4.60 |
| QL | inf | 1.40 | 1.18 | 2.34 | 2.04 | 1.59 | 1.20 | 75.96 | 3.32 |
| Eff | 0.18 | 1.00 | 0.96 | 0.98 | 1.00 | 1.12 | 1.20 | 0.94 | 2.33 |
| Panel C: $\mu = 0.001$ | | | | | | | | | |
| Mean | 10.05 | 6.16 | 6.90 | 4.38 | 4.81 | 4.66 | 5.56 | 4.13 | 4.91 |
| abs.err | 0.15 | 3.74 | 3.00 | 5.52 | 5.08 | 5.24 | 4.34 | 5.77 | 4.98 |
| std.err | 2.28 | 0.97 | 1.09 | 0.69 | 0.76 | 0.60 | 0.73 | 0.59 | 0.69 |
| MSE | 85.61 | 9.81 | 10.39 | 9.97 | 9.73 | 9.21 | 8.50 | 11.09 | 4.60 |
| QL | inf | 1.39 | 1.17 | 2.32 | 2.02 | 1.58 | 1.19 | 20.68 | 3.45 |
| Eff | 0.18 | 1.00 | 0.97 | 0.97 | 0.99 | 1.09 | 1.17 | 0.91 | 2.25 |
| Panel D: $\mu = 0.005$ | | | | | | | | | |
| Mean | 12.49 | 7.07 | 7.92 | 5.02 | 5.52 | 4.98 | 5.97 | 4.18 | 5.00 |
| abs.err | 2.53 | 2.88 | 2.03 | 4.93 | 4.43 | 4.97 | 3.98 | 5.78 | 4.95 |
| std.err | 2.59 | 1.08 | 1.21 | 0.77 | 0.85 | 0.65 | 0.78 | 0.63 | 0.73 |
| MSE | 85.61 | 10.59 | 11.61 | 10.00 | 9.90 | 9.00 | 8.36 | 11.11 | 4.60 |
| QL | inf | 1.08 | 0.91 | 1.80 | 1.56 | 1.37 | 1.02 | 29.84 | 4.92 |
| Eff | 0.19 | 1.00 | 0.93 | 1.04 | 1.05 | 1.20 | 1.29 | 0.97 | 2.41 |
| Panel E: $\mu = 0.01$ | | | | | | | | | |
| Mean | 19.86 | 9.78 | 10.95 | 6.95 | 7.64 | 5.88 | 7.15 | 4.26 | 5.18 |
| abs.err | 9.96 | 0.13 | 1.04 | 2.96 | 2.27 | 4.02 | 2.76 | 5.65 | 4.72 |
| std.err | 2.85 | 1.12 | 1.25 | 0.80 | 0.88 | 0.66 | 0.79 | 0.70 | 0.80 |
| MSE | 85.61 | 12.53 | 14.70 | 9.98 | 10.25 | 8.54 | 8.08 | 11.61 | 4.60 |
| QL | inf | 0.64 | 0.57 | 1.00 | 0.88 | 0.92 | 0.67 | 48.51 | 6.73 |
| Eff | 0.23 | 1.00 | 0.86 | 1.26 | 1.22 | 1.55 | 1.62 | 1.12 | 2.85 |

Table B.14: The effect of drift on range-based estimators with stochastic volatility.

| kurtosis = 5 and intraday movements $k$= 100 | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | $r_i^2$ | P | $P_k$ | BL | $BL_k$ | GK | $GK_k$ | RS | $RS_k$ |
| Panel A: $\mu = 0$ | | | | | | | | | |
| Mean | 9.78 | 5.65 | 6.33 | 4.02 | 4.42 | 4.06 | 4.86 | 3.45 | 4.14 |
| abs.err | 0.09 | 4.22 | 3.54 | 5.86 | 5.46 | 5.81 | 5.01 | 6.42 | 5.73 |
| std.err | 2.72 | 1.15 | 1.29 | 0.82 | 0.90 | 0.66 | 0.80 | 0.59 | 0.69 |
| MSE | 85.61 | 14.62 | 15.44 | 14.56 | 14.32 | 13.84 | 12.81 | 16.41 | 4.60 |
| QL | inf | 1.80 | 1.53 | 2.93 | 2.57 | 2.09 | 1.60 | 13.34 | 4.27 |
| Eff | 0.27 | 1.00 | 0.98 | 0.98 | 0.99 | 1.08 | 1.15 | 0.90 | 3.37 |
| Panel B: $\mu = 0.0005$ | | | | | | | | | |
| Mean | 9.89 | 5.71 | 6.39 | 4.05 | 4.46 | 4.09 | 4.90 | 3.48 | 4.17 |
| abs.err | 0.05 | 4.23 | 3.55 | 5.88 | 5.48 | 5.85 | 5.04 | 6.46 | 5.77 |
| std.err | 2.66 | 1.13 | 1.26 | 0.80 | 0.88 | 0.65 | 0.78 | 0.58 | 0.68 |
| MSE | 85.61 | 14.49 | 15.18 | 14.70 | 14.41 | 14.17 | 13.08 | 16.86 | 4.60 |
| QL | inf | 1.80 | 1.53 | 2.93 | 2.57 | 2.11 | 1.61 | 13.88 | 4.46 |
| Eff | 0.26 | 1.00 | 0.98 | 0.97 | 0.99 | 1.04 | 1.12 | 0.87 | 3.31 |
| Panel C: $\mu = 0.001$ | | | | | | | | | |
| Mean | 10.05 | 5.75 | 6.44 | 4.09 | 4.49 | 4.09 | 4.90 | 3.45 | 4.14 |
| abs.err | 0.05 | 4.24 | 3.56 | 5.91 | 5.50 | 5.91 | 5.09 | 6.55 | 5.86 |
| std.err | 2.67 | 1.14 | 1.27 | 0.81 | 0.89 | 0.65 | 0.79 | 0.58 | 0.68 |
| MSE | 85.61 | 14.68 | 15.39 | 14.87 | 14.57 | 14.32 | 13.20 | 17.12 | 4.60 |
| QL | inf | 1.80 | 1.53 | 2.92 | 2.56 | 2.07 | 1.58 | 17.36 | 4.31 |
| Eff | 0.27 | 1.00 | 0.98 | 0.97 | 0.99 | 1.05 | 1.13 | 0.87 | 3.34 |
| Panel D: $\mu = 0.005$ | | | | | | | | | |
| Mean | 12.37 | 6.65 | 7.45 | 4.73 | 5.20 | 4.45 | 5.36 | 3.57 | 4.30 |
| abs.err | 2.45 | 3.27 | 2.47 | 5.20 | 4.73 | 5.48 | 4.57 | 6.36 | 5.63 |
| std.err | 2.67 | 1.09 | 1.22 | 0.78 | 0.85 | 0.63 | 0.76 | 0.61 | 0.70 |
| MSE | 85.61 | 14.01 | 14.89 | 13.94 | 13.69 | 13.49 | 12.42 | 16.67 | 4.60 |
| QL | inf | 1.27 | 1.08 | 2.06 | 1.80 | 1.68 | 1.26 | 41.19 | 6.60 |
| Eff | 0.25 | 1.00 | 0.96 | 1.01 | 1.02 | 1.08 | 1.18 | 0.87 | 3.20 |
| Panel E: $\mu = 0.01$ | | | | | | | | | |
| Mean | 19.99 | 9.52 | 10.66 | 6.76 | 7.44 | 5.48 | 6.69 | 3.77 | 4.61 |
| abs.err | 10.04 | 0.42 | 0.72 | 3.18 | 2.51 | 4.47 | 3.26 | 6.18 | 5.34 |
| std.err | 3.38 | 1.34 | 1.51 | 0.96 | 1.05 | 0.74 | 0.89 | 0.69 | 0.81 |
| MSE | 85.61 | 17.07 | 19.58 | 14.18 | 14.47 | 12.89 | 12.23 | 16.74 | 4.60 |
| QL | inf | 0.72 | 0.66 | 1.07 | 0.95 | 1.06 | 0.78 | 46.84 | 8.58 |
| Eff | 0.31 | 1.00 | 0.89 | 1.19 | 1.16 | 1.41 | 1.47 | 1.07 | 3.94 |

| kurtosis = 7 and intraday movements $k$= 100 | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | $r_i^2$ | P | $P_k$ | BL | $BL_k$ | GK | $GK_k$ | RS | $RS_k$ |
| Panel A: $\mu = 0$ | | | | | | | | | |
| Mean | 10.05 | 5.51 | 6.17 | 3.92 | 4.31 | 3.76 | 4.52 | 3.06 | 3.69 |
| abs.err | 0.05 | 4.48 | 3.82 | 6.08 | 5.69 | 6.24 | 5.47 | 6.94 | 6.31 |
| std.err | 3.00 | 1.25 | 1.40 | 0.89 | 0.97 | 0.67 | 0.82 | 0.56 | 0.66 |
| MSE | 85.61 | 18.91 | 19.83 | 19.00 | 18.67 | 18.90 | 17.44 | 22.75 | 4.60 |
| QL | inf | 2.13 | 1.81 | 3.41 | 2.99 | 2.51 | 1.94 | 16.12 | 5.14 |
| Eff | 0.34 | 1.00 | 0.98 | 0.98 | 1.00 | 1.05 | 1.12 | 0.87 | 4.34 |
| Panel B: $\mu = 0.0005$ | | | | | | | | | |
| Mean | 9.94 | 5.46 | 6.12 | 3.88 | 4.27 | 3.74 | 4.49 | 3.05 | 3.68 |
| abs.err | 0.01 | 4.47 | 3.81 | 6.05 | 5.66 | 6.20 | 5.44 | 6.89 | 6.26 |
| std.err | 2.88 | 1.22 | 1.36 | 0.87 | 0.95 | 0.69 | 0.83 | 0.60 | 0.70 |
| MSE | 85.61 | 18.54 | 19.23 | 19.04 | 18.63 | 19.23 | 17.77 | 23.01 | 4.60 |
| QL | inf | 2.14 | 1.83 | 3.43 | 3.01 | 2.51 | 1.94 | 18.95 | 5.06 |
| Eff | 0.33 | 1.00 | 0.98 | 0.97 | 0.99 | 1.00 | 1.08 | 0.84 | 4.23 |
| Panel C: $\mu = 0.001$ | | | | | | | | | |
| Mean | 10.02 | 5.50 | 6.16 | 3.91 | 4.30 | 3.76 | 4.52 | 3.06 | 3.69 |
| abs.err | 0.12 | 4.40 | 3.74 | 5.99 | 5.60 | 6.14 | 5.38 | 6.84 | 6.21 |
| std.err | 2.97 | 1.24 | 1.39 | 0.88 | 0.97 | 0.68 | 0.83 | 0.58 | 0.69 |
| MSE | 85.61 | 18.58 | 19.21 | 19.24 | 18.79 | 19.57 | 18.05 | 23.58 | 4.60 |
| QL | inf | 2.10 | 1.79 | 3.36 | 2.96 | 2.50 | 1.93 | 15.43 | 5.21 |
| Eff | 0.33 | 1.00 | 0.98 | 0.97 | 0.99 | 1.00 | 1.08 | 0.82 | 4.22 |
| Panel D: $\mu = 0.005$ | | | | | | | | | |
| Mean | 12.18 | 6.34 | 7.10 | 4.50 | 4.95 | 4.08 | 4.93 | 3.16 | 3.82 |
| abs.err | 2.32 | 3.52 | 2.76 | 5.36 | 4.91 | 5.78 | 4.93 | 6.70 | 6.04 |
| std.err | 2.91 | 1.21 | 1.36 | 0.86 | 0.95 | 0.69 | 0.83 | 0.62 | 0.72 |
| MSE | 85.61 | 17.84 | 18.69 | 18.05 | 17.70 | 18.10 | 16.69 | 22.00 | 4.60 |
| QL | inf | 1.40 | 1.20 | 2.24 | 1.97 | 1.91 | 1.43 | 33.90 | 8.17 |
| Eff | 0.32 | 1.00 | 0.97 | 0.99 | 1.01 | 1.04 | 1.13 | 0.85 | 4.07 |
| Panel E: $\mu = 0.01$ | | | | | | | | | |
| Mean | 19.82 | 9.23 | 10.34 | 6.56 | 7.21 | 5.14 | 6.30 | 3.40 | 4.18 |
| abs.err | 9.95 | 0.64 | 0.46 | 3.31 | 2.66 | 4.73 | 3.58 | 6.47 | 5.70 |
| std.err | 3.61 | 1.42 | 1.59 | 1.01 | 1.11 | 0.76 | 0.92 | 0.70 | 0.82 |
| MSE | 85.61 | 20.28 | 22.54 | 18.14 | 18.23 | 17.40 | 16.16 | 22.42 | 4.60 |
| QL | inf | 0.79 | 0.72 | 1.12 | 1.00 | 1.14 | 0.85 | 43.54 | 9.80 |
| Eff | 0.37 | 1.00 | 0.91 | 1.13 | 1.12 | 1.28 | 1.35 | 0.97 | 4.64 |

Table B.15: The average variance estimates with 95% confidence interval.

| | expected variance = 10, kurtosis = 1 | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| $k$ | $r_i^2$ | P | $P_k$ | BL | $BL_k$ | GK | $GK_k$ | RS | $RS_k$ |
| Panel A: $\mu = 0$ | | | | | | | | | |
| 20 | 9.47 | 6.13 | 8.00 | 4.36 | 5.43 | 4.84 | 7.26 | 4.51 | 6.72 |
| | (3.27) | (1.40) | (1.83) | (1.00) | (1.25) | (0.96) | (1.44) | (1.03) | (1.43) |
| 40 | 9.80 | 6.63 | 7.96 | 4.71 | 5.49 | 5.41 | 7.12 | 5.09 | 6.66 |
| | (3.22) | (1.40) | (1.69) | (1.00) | (1.16) | (0.99) | (1.31) | (1.07) | (1.34) |
| 100 | 9.96 | 6.99 | 7.83 | 4.97 | 5.46 | 5.85 | 6.92 | 5.54 | 6.52 |
| | (3.29) | (1.46) | (1.64) | (1.04) | (1.14) | (1.03) | (1.23) | (1.08) | (1.25) |
| Panel B: $\mu = 0.0005$ | | | | | | | | | |
| 20 | 9.51 | 6.13 | 8.00 | 4.36 | 5.44 | 4.83 | 7.25 | 4.49 | 6.70 |
| | (3.33) | (1.42) | (1.86) | (1.01) | (1.26) | (0.97) | (1.45) | (1.03) | (1.44) |
| 40 | 9.72 | 6.58 | 7.90 | 4.68 | 5.45 | 5.37 | 7.07 | 5.06 | 6.61 |
| | (3.35) | (1.48) | (1.77) | (1.05) | (1.22) | (1.03) | (1.37) | (1.09) | (1.37) |
| 100 | 9.83 | 6.91 | 7.74 | 4.91 | 5.40 | 5.79 | 6.84 | 5.49 | 6.46 |
| | (3.30) | (1.44) | (1.61) | (1.02) | (1.12) | (1.04) | (1.23) | (1.14) | (1.31) |
| Panel C: $\mu = 0.001$ | | | | | | | | | |
| 20 | 9.59 | 6.18 | 8.06 | 4.39 | 5.47 | 4.86 | 7.29 | 4.51 | 6.73 |
| | (3.52) | (1.47) | (1.92) | (1.05) | (1.31) | (0.97) | (1.47) | (1.03) | (1.44) |
| 40 | 9.93 | 6.68 | 8.03 | 4.75 | 5.53 | 5.43 | 7.15 | 5.10 | 6.68 |
| | (3.47) | (1.47) | (1.76) | (1.04) | (1.21) | (0.98) | (1.30) | (1.03) | (1.29) |
| 100 | 9.95 | 6.97 | 7.81 | 4.95 | 5.45 | 5.82 | 6.89 | 5.52 | 6.50 |
| | (3.44) | (1.53) | (1.72) | (1.09) | (1.20) | (1.10) | (1.30) | (1.18) | (1.35) |
| Panel D: $\mu = 0.005$ | | | | | | | | | |
| 20 | 12.02 | 7.03 | 9.17 | 4.99 | 6.23 | 5.10 | 7.76 | 4.47 | 6.79 |
| | (3.87) | (1.54) | (2.01) | (1.10) | (1.37) | (0.94) | (1.43) | (1.00) | (1.38) |
| 40 | 12.24 | 7.48 | 8.99 | 5.32 | 6.20 | 5.65 | 7.51 | 5.03 | 6.66 |
| | (3.78) | (1.55) | (1.87) | (1.10) | (1.29) | (1.02) | (1.35) | (1.10) | (1.37) |
| 100 | 12.39 | 7.85 | 8.80 | 5.58 | 6.14 | 6.10 | 7.26 | 5.50 | 6.52 |
| | (3.78) | (1.57) | (1.76) | (1.11) | (1.23) | (1.05) | (1.25) | (1.15) | (1.31) |
| Panel E: $\mu = 0.01$ | | | | | | | | | |
| 20 | 19.53 | 9.64 | 12.57 | 6.85 | 8.54 | 5.82 | 9.19 | 4.27 | 6.91 |
| | (5.15) | (1.92) | (2.50) | (1.36) | (1.70) | (1.05) | (1.62) | (1.15) | (1.56) |
| 40 | 19.84 | 10.19 | 12.23 | 7.24 | 8.43 | 6.46 | 8.79 | 4.93 | 6.76 |
| | (5.43) | (2.07) | (2.49) | (1.47) | (1.72) | (1.15) | (1.55) | (1.21) | (1.51) |
| 100 | 19.99 | 10.58 | 11.84 | 7.52 | 8.26 | 6.94 | 8.38 | 5.43 | 6.56 |
| | (5.14) | (1.96) | (2.19) | (1.39) | (1.53) | (1.14) | (1.37) | (1.25) | (1.43) |

| | expected variance = 10, kurtosis = 3 | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| $k$ | $r_i^2$ | P | $P_k$ | BL | $BL_k$ | GK | $GK_k$ | RS | $RS_k$ |
| Panel A: $\mu = 0$ | | | | | | | | | |
| 20 | 9.50 | 5.51 | 7.18 | 3.91 | 4.88 | 3.96 | 6.05 | 3.44 | 5.27 |
| | (4.32) | (1.82) | (2.38) | (1.30) | (1.62) | (1.07) | (1.67) | (1.00) | (1.45) |
| 40 | 9.71 | 5.85 | 7.02 | 4.16 | 4.84 | 4.36 | 5.81 | 3.85 | 5.12 |
| | (4.11) | (1.80) | (2.16) | (1.28) | (1.49) | (1.14) | (1.53) | (1.10) | (1.41) |
| 100 | 9.88 | 6.10 | 6.83 | 4.33 | 4.76 | 4.64 | 5.53 | 4.13 | 4.91 |
| | (4.43) | (1.93) | (2.16) | (1.37) | (1.50) | (1.20) | (1.44) | (1.15) | (1.34) |
| Panel B: $\mu = 0.0005$ | | | | | | | | | |
| 20 | 9.59 | 5.55 | 7.24 | 3.94 | 4.92 | 3.99 | 6.08 | 3.45 | 5.29 |
| | (4.15) | (1.77) | (2.30) | (1.25) | (1.56) | (1.07) | (1.65) | (1.04) | (1.49) |
| 40 | 9.80 | 5.87 | 7.05 | 4.17 | 4.86 | 4.35 | 5.79 | 3.82 | 5.09 |
| | (4.31) | (1.85) | (2.22) | (1.31) | (1.53) | (1.15) | (1.55) | (1.13) | (1.44) |
| 100 | 9.99 | 6.14 | 6.87 | 4.36 | 4.79 | 4.65 | 5.54 | 4.12 | 4.90 |
| | (4.39) | (1.87) | (2.10) | (1.33) | (1.46) | (1.15) | (1.38) | (1.10) | (1.28) |
| Panel C: $\mu = 0.001$ | | | | | | | | | |
| 20 | 9.46 | 5.48 | 7.14 | 3.89 | 4.85 | 3.94 | 6.01 | 3.41 | 5.23 |
| | (4.19) | (1.77) | (2.31) | (1.26) | (1.57) | (1.08) | (1.67) | (1.07) | (1.52) |
| 40 | 9.82 | 5.89 | 7.07 | 4.18 | 4.88 | 4.37 | 5.83 | 3.85 | 5.13 |
| | (4.35) | (1.85) | (2.23) | (1.32) | (1.54) | (1.14) | (1.54) | (1.12) | (1.42) |
| 100 | 10.05 | 6.16 | 6.90 | 4.38 | 4.81 | 4.66 | 5.56 | 4.13 | 4.91 |
| | (4.46) | (1.91) | (2.13) | (1.35) | (1.49) | (1.19) | (1.42) | (1.16) | (1.35) |
| Panel D: $\mu = 0.005$ | | | | | | | | | |
| 20 | 12.07 | 6.44 | 8.40 | 4.57 | 5.70 | 4.26 | 6.60 | 3.44 | 5.39 |
| | (4.63) | (1.85) | (2.42) | (1.32) | (1.64) | (1.06) | (1.64) | (1.06) | (1.49) |
| 40 | 12.22 | 6.77 | 8.13 | 4.81 | 5.61 | 4.67 | 6.28 | 3.87 | 5.21 |
| | (4.67) | (1.90) | (2.29) | (1.35) | (1.58) | (1.12) | (1.51) | (1.10) | (1.39) |
| 100 | 12.49 | 7.07 | 7.92 | 5.02 | 5.52 | 4.98 | 5.97 | 4.18 | 5.00 |
| | (5.08) | (2.12) | (2.38) | (1.51) | (1.66) | (1.28) | (1.54) | (1.24) | (1.44) |
| Panel E: $\mu = 0.01$ | | | | | | | | | |
| 20 | 19.52 | 9.13 | 11.92 | 6.49 | 8.09 | 5.12 | 8.21 | 3.47 | 5.77 |
| | (5.78) | (2.26) | (2.95) | (1.60) | (2.00) | (1.24) | (1.94) | (1.25) | (1.76) |
| 40 | 19.96 | 9.59 | 11.52 | 6.81 | 7.94 | 5.58 | 7.70 | 3.91 | 5.46 |
| | (5.78) | (2.26) | (2.72) | (1.61) | (1.87) | (1.25) | (1.70) | (1.25) | (1.57) |
| 100 | 19.86 | 9.78 | 10.95 | 6.95 | 7.64 | 5.88 | 7.15 | 4.26 | 5.18 |
| | (5.59) | (2.20) | (2.46) | (1.56) | (1.72) | (1.30) | (1.55) | (1.37) | (1.57) |

Table B.16: The average variance estimates with 95% confidence interval.

| | expected variance = 10, kurtosis = 5 | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| $k$ | $r_i^2$ | P | $P_k$ | BL | $BL_k$ | GK | $GK_k$ | RS | $RS_k$ |
| Panel A: $\mu = 0$ | | | | | | | | | |
| 20 | 9.60 | 5.26 | 6.86 | 3.74 | 4.66 | 3.59 | 5.53 | 2.96 | 4.62 |
| | (5.50) | (2.29) | (2.99) | (1.63) | (2.03) | (1.26) | (1.99) | (1.09) | (1.62) |
| 40 | 9.85 | 5.53 | 6.64 | 3.93 | 4.58 | 3.86 | 5.19 | 3.23 | 4.36 |
| | (4.86) | (2.09) | (2.51) | (1.48) | (1.73) | (1.23) | (1.67) | (1.12) | (1.45) |
| 100 | 9.78 | 5.65 | 6.33 | 4.02 | 4.42 | 4.06 | 4.86 | 3.45 | 4.14 |
| | (5.33) | (2.26) | (2.53) | (1.60) | (1.76) | (1.30) | (1.57) | (1.15) | (1.35) |
| Panel B: $\mu = 0.0005$ | | | | | | | | | |
| 20 | 9.44 | 5.17 | 6.74 | 3.67 | 4.58 | 3.52 | 5.43 | 2.90 | 4.53 |
| | (5.06) | (2.10) | (2.74) | (1.49) | (1.86) | (1.16) | (1.83) | (1.02) | (1.51) |
| 40 | 9.64 | 5.44 | 6.54 | 3.87 | 4.51 | 3.82 | 5.13 | 3.21 | 4.33 |
| | (4.95) | (2.08) | (2.50) | (1.48) | (1.73) | (1.21) | (1.64) | (1.13) | (1.45) |
| 100 | 9.89 | 5.71 | 6.39 | 4.05 | 4.46 | 4.09 | 4.90 | 3.48 | 4.17 |
| | (5.21) | (2.21) | (2.47) | (1.57) | (1.73) | (1.27) | (1.54) | (1.13) | (1.32) |
| Panel C: $\mu = 0.001$ | | | | | | | | | |
| 20 | 9.62 | 5.24 | 6.84 | 3.72 | 4.64 | 3.55 | 5.48 | 2.91 | 4.56 |
| | (5.00) | (2.09) | (2.72) | (1.48) | (1.85) | (1.17) | (1.85) | (1.05) | (1.55) |
| 40 | 9.90 | 5.55 | 6.67 | 3.95 | 4.60 | 3.88 | 5.21 | 3.24 | 4.37 |
| | (5.06) | (2.12) | (2.55) | (1.51) | (1.75) | (1.21) | (1.65) | (1.10) | (1.43) |
| 100 | 10.05 | 5.75 | 6.44 | 4.09 | 4.49 | 4.09 | 4.90 | 3.45 | 4.14 |
| | (5.23) | (2.23) | (2.49) | (1.58) | (1.74) | (1.28) | (1.55) | (1.13) | (1.33) |
| Panel D: $\mu = 0.005$ | | | | | | | | | |
| 20 | 12.12 | 6.18 | 8.06 | 4.39 | 5.48 | 3.89 | 6.09 | 2.98 | 4.76 |
| | (5.68) | (2.27) | (2.96) | (1.61) | (2.01) | (1.19) | (1.90) | (1.07) | (1.55) |
| 40 | 12.34 | 6.49 | 7.79 | 4.61 | 5.37 | 4.23 | 5.73 | 3.33 | 4.54 |
| | (5.64) | (2.33) | (2.80) | (1.65) | (1.93) | (1.32) | (1.80) | (1.22) | (1.57) |
| 100 | 12.37 | 6.65 | 7.45 | 4.73 | 5.20 | 4.45 | 5.36 | 3.57 | 4.30 |
| | (5.23) | (2.14) | (2.40) | (1.52) | (1.67) | (1.24) | (1.49) | (1.19) | (1.38) |
| Panel E: $\mu = 0.01$ | | | | | | | | | |
| 20 | 19.47 | 8.88 | 11.58 | 6.31 | 7.87 | 4.79 | 7.75 | 3.07 | 5.20 |
| | (6.17) | (2.45) | (3.19) | (1.74) | (2.17) | (1.37) | (2.13) | (1.35) | (1.92) |
| 40 | 19.68 | 9.21 | 11.06 | 6.54 | 7.63 | 5.17 | 7.16 | 3.47 | 4.88 |
| | (6.39) | (2.51) | (3.02) | (1.79) | (2.08) | (1.37) | (1.86) | (1.31) | (1.66) |
| 100 | 19.99 | 9.52 | 10.66 | 6.76 | 7.44 | 5.48 | 6.69 | 3.77 | 4.61 |
| | (6.63) | (2.63) | (2.95) | (1.87) | (2.06) | (1.44) | (1.75) | (1.36) | (1.58) |

| | expected variance = 10, kurtosis = 7 | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| $k$ | $r_i^2$ | P | $P_k$ | BL | $BL_k$ | GK | $GK_k$ | RS | $RS_k$ |
| Panel A: $\mu = 0$ | | | | | | | | | |
| 20 | 9.59 | 5.04 | 6.57 | 3.58 | 4.46 | 3.28 | 5.10 | 2.58 | 4.11 |
| | (6.03) | (2.54) | (3.31) | (1.80) | (2.25) | (1.36) | (2.19) | (1.09) | (1.67) |
| 40 | 9.79 | 5.27 | 6.33 | 3.75 | 4.36 | 3.53 | 4.77 | 2.83 | 3.87 |
| | (5.52) | (2.31) | (2.77) | (1.64) | (1.91) | (1.28) | (1.76) | (1.12) | (1.46) |
| 100 | 10.05 | 5.51 | 6.17 | 3.92 | 4.31 | 3.76 | 4.52 | 3.06 | 3.69 |
| | (5.88) | (2.45) | (2.74) | (1.74) | (1.91) | (1.31) | (1.60) | (1.09) | (1.29) |
| Panel B: $\mu = 0.0005$ | | | | | | | | | |
| 20 | 9.52 | 4.98 | 6.50 | 3.54 | 4.42 | 3.23 | 5.04 | 2.53 | 4.04 |
| | (5.84) | (2.41) | (3.15) | (1.71) | (2.14) | (1.25) | (2.02) | (1.00) | (1.53) |
| 40 | 9.84 | 5.30 | 6.36 | 3.76 | 4.39 | 3.54 | 4.79 | 2.84 | 3.88 |
| | (5.74) | (2.41) | (2.89) | (1.71) | (1.99) | (1.32) | (1.81) | (1.11) | (1.46) |
| 100 | 9.94 | 5.46 | 6.12 | 3.88 | 4.27 | 3.74 | 4.49 | 3.05 | 3.68 |
| | (5.65) | (2.39) | (2.67) | (1.70) | (1.86) | (1.34) | (1.63) | (1.17) | (1.38) |
| Panel C: $\mu = 0.001$ | | | | | | | | | |
| 20 | 9.61 | 5.04 | 6.57 | 3.58 | 4.47 | 3.27 | 5.10 | 2.59 | 4.11 |
| | (5.52) | (2.28) | (2.98) | (1.62) | (2.02) | (1.24) | (1.97) | (1.12) | (1.64) |
| 40 | 9.80 | 5.28 | 6.35 | 3.75 | 4.37 | 3.54 | 4.78 | 2.84 | 3.88 |
| | (5.49) | (2.27) | (2.72) | (1.61) | (1.88) | (1.23) | (1.69) | (1.07) | (1.39) |
| 100 | 10.02 | 5.50 | 6.16 | 3.91 | 4.30 | 3.76 | 4.52 | 3.06 | 3.69 |
| | (5.83) | (2.44) | (2.73) | (1.73) | (1.90) | (1.34) | (1.63) | (1.14) | (1.34) |
| Panel D: $\mu = 0.005$ | | | | | | | | | |
| 20 | 11.86 | 5.88 | 7.68 | 4.18 | 5.21 | 3.58 | 5.64 | 2.64 | 4.27 |
| | (6.07) | (2.44) | (3.19) | (1.74) | (2.17) | (1.25) | (2.01) | (1.03) | (1.54) |
| 40 | 12.08 | 6.16 | 7.39 | 4.37 | 5.10 | 3.87 | 5.27 | 2.94 | 4.04 |
| | (5.84) | (2.45) | (2.94) | (1.74) | (2.03) | (1.37) | (1.88) | (1.21) | (1.57) |
| 100 | 12.18 | 6.34 | 7.10 | 4.50 | 4.95 | 4.08 | 4.93 | 3.16 | 3.82 |
| | (5.70) | (2.38) | (2.66) | (1.69) | (1.86) | (1.35) | (1.63) | (1.21) | (1.42) |
| Panel E: $\mu = 0.01$ | | | | | | | | | |
| 20 | 19.40 | 8.71 | 11.36 | 6.19 | 7.72 | 4.58 | 7.45 | 2.82 | 4.85 |
| | (6.75) | (2.66) | (3.47) | (1.89) | (2.36) | (1.42) | (2.24) | (1.32) | (1.89) |
| 40 | 19.56 | 8.99 | 10.79 | 6.38 | 7.44 | 4.90 | 6.82 | 3.16 | 4.48 |
| | (6.32) | (2.53) | (3.03) | (1.80) | (2.09) | (1.45) | (1.95) | (1.41) | (1.79) |
| 100 | 19.82 | 9.23 | 10.34 | 6.56 | 7.21 | 5.14 | 6.30 | 3.40 | 4.18 |
| | (7.07) | (2.78) | (3.11) | (1.97) | (2.17) | (1.49) | (1.80) | (1.38) | (1.60) |

# Appendix C

# The BHHH algorithm for DNIG model

The BHHH algorithm is one of the most widely used methods for maximizing the log-likelihood function in financial econometrics. The BHHH requires only the gradient of the log-likelihood function and therefore the implementation is relatively simple.

The log-likelihood function at the $t^{th}$ observation is

$$\begin{aligned} l_t(\alpha,\beta,\omega) &= \log(f(y_t|\mathcal{F}_{t-1};\alpha,\beta,\omega)) \\ &= \log\left[\frac{\omega\exp(\omega)}{\pi\sqrt{y_t^2+\phi_t\omega}}\left(K_1\left(\sqrt{\frac{\omega y_t^2}{\phi_t}+\omega^2}\right)\right)\right] \\ &= \log(\omega)+\omega-\log(\pi)-\frac{1}{2}\log\left(y_t^2+\omega\exp(\alpha+\beta\log(d_{t-1}^2)\right) \\ &\quad +\log\left(K_1\left(\sqrt{\frac{\omega y_t^2}{\exp\left(\alpha+\beta\log(d_{t-1}^2)\right)}+\omega^2}\right)\right) \end{aligned}$$

Thus, the gradient at the $t^{th}$ observation is the vector $G_t = (\partial l_t/\partial\alpha, \partial l_t/\partial\beta, \partial l_t/\partial\omega)'$. Given $z_t = \sqrt{\frac{\omega y_t^2}{\exp(\alpha+\beta\log(R_{t-1}^2))}+\omega^2}$ and $\omega \geq 0$ the entries of the gradient are

$$\begin{aligned} \frac{\partial l_t}{\partial\alpha} &= -\frac{1}{2}\frac{\omega\exp\left(\alpha+\beta\log(R_{t-1}^2)\right)}{y_t^2+\omega\exp\left(\alpha+\beta\log(R_{t-1}^2)\right)}+\left(\frac{1}{z}-\frac{K_2(z_t)}{K_1(z_t)}\right)\frac{\partial z_t}{\partial\alpha} \\ \frac{\partial l_t}{\partial\beta} &= -\frac{1}{2}\frac{\omega\log(R_{t-1}^2)\exp\left(\alpha+\beta\log(R_{t-1}^2)\right)}{y_t^2+\omega\exp\left(\alpha+\beta\log(R_{t-1}^2)\right)}+\left(\frac{1}{z}-\frac{K_2(z_t)}{K_1(z_t)}\right)\frac{\partial z_t}{\partial\beta} \\ \frac{\partial l_t}{\partial\omega} &= \frac{1}{\omega}+1-\frac{1}{2}\frac{\exp\left(\alpha+\beta\log(R_{t-1}^2)\right)}{y_t^2+\omega\exp\left(\alpha+\beta\log(R_{t-1}^2)\right)}+\left(\frac{1}{z}-\frac{K_2(z_t)}{K_1(z_t)}\right)\frac{\partial z_t}{\partial\omega} \end{aligned}$$

where

$$\begin{aligned}
\frac{\partial z_t}{\partial \alpha} &= -\frac{1}{2z_t}\left(\frac{\omega y_t^2}{\exp\left(\alpha+\beta\log(R_{t-1}^2)\right)}\right)\\
\frac{\partial z_t}{\partial \beta} &= -\frac{1}{2z_t}\left(\frac{\omega y_t^2 \log(R_{t-1}^2)}{\exp\left(\alpha+\beta\log(R_{t-1}^2)\right)}\right)\\
\frac{\partial z_t}{\partial \omega} &= \frac{1}{2z_t}\left(\frac{y_t^2}{\exp\left(\alpha+\beta\log(R_{t-1}^2)\right)}+2\omega\right)
\end{aligned}$$

The gradient $G$ is therefore the summation of $G_t$, that is $G=\sum_{t=1}^{n} G_t$. The information matrix $I=[I_{i,j}]$ equals the outer product of gradient $B=[B_{i,j}]$ that is estimated by

$$\hat{I}_{i,j}=\hat{B}_{i,j}=\frac{1}{n}\sum_{t=1}^{n}\frac{\partial l_t}{\partial\Theta_i}.\frac{\partial l_t}{\partial\Theta_j}.$$

Employing the BHHH algorithm, the estimated parameter are updated by

$$\hat{\Theta}_{(k)}=\hat{\Theta}_{(k-1)}+B_{(k-1)}^{-1}G_{(k-1)}.$$