JUMP DETECTION WITH POWER AND BIPOWER VARIATION PROCESSES

HAVVA ÖZLEM DURSUN

SEPTEMBER 2007

JUMP DETECTION WITH POWER AND BIPOWER VARIATION PROCESSES

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF APPLIED MATHEMATICS
OF
THE MIDDLE EAST TECHNICAL UNIVERSITY

BY

HAVVA ÖZLEM DURSUN

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

FOR

THE DEGREE OF MASTER OF SCIENCE

IN

FINANCIAL MATHEMATICS

SEPTEMBER 2007

Approval of the Graduate School of Applied Mathematics.

Prof. Dr. Ersan Akyıldız
Director

I certify that this thesis satisfies all the requirements as a thesis for the degree of Master of Science.

Prof. Dr. Ersan Akyıldız
Head of Department

This is to certify that we have read this thesis and that in our opinion it is fully adequate, in scope and quality, as a thesis for the degree of Master of Science.

Assoc. Prof. Dr. Azize Hayfavi
Supervisor

Examining Committee Members

Assoc. Prof. Dr. Azize Hayfavi

Assist. Prof. Dr. Kasırga Yıldırak

Assist. Prof. Dr. Hakan Öktem

Dr. Coşkun Küçüközmen

Dr. Seza Danışoğlu

"I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work."

Name, Lastname : HAVVA ÖZLEM DURSUN

Signature :

ABSTRACT

JUMP DETECTION WITH POWER AND BIPOWER VARIATION PROCESSES

DURSUN, HAVVA ÖZLEM

M.Sc., Department of Financial Mathematics

Supervisor: Assoc. Prof. Dr. Azize Hayfavi

September 2007, 78 pages.

In this study, we show that realized bipower variation which is an extension of realized power variation is an alternative method that estimates integrated variance like realized variance. It is seen that realized bipower variation is robust to rare jumps. Robustness means that if we add rare jumps to a stochastic volatility process, realized bipower variation process continues to estimate integrated variance although realized variance estimates integrated variance plus the quadratic variation of the jump component. This robustness is crucial since it separates the discontinuous component of quadratic variation which comes from the jump part of the logarithmic price process. Thus, we demonstrate that if the logarithmic price process is in the class of stochastic volatility plus rare jumps processes then the difference between realized variance and realized bipower variation process estimates the discontinuous component of the quadratic variation. So, quadratic variation of the jump component can be estimated and jump detection can be achieved.

keywords: Stochastic volatility, Quadratic variation, Power variation, Bipower

variation, Jump process, Realized variance, Realized volatility, Semimartingale, Integrated variance.

ÖZ

KUVVET VE İKİLİ KUVVET VARYASYON SÜREÇLERİ KULLANILARAK SIÇRAMALARIN YAKALANMASI

DURSUN, HAVVA ÖZLEM

Yüksek Lisans, Finansal Matematik Bölümü

Tez Yöneticisi: Assoc. Prof. Dr. Azize Hayfavi

Eylül 2007, 78 sayfa.

Bu çalışmada, gerçekleşen kuvvet varyasyonunun genelleştirilmiş hali olan gerçekleşen ikili kuvvet varyasyonunun, gerçekleşen varyans gibi, bütünleşik varyansı tahmin eden alternatif bir metod olduğunu gösterdik. Gerçekleşen ikili kuvvet varyasyonunun seyrek sıçramalara karşı dayanıklı olduğu görüldü. Dayanıklılık, bir stokastik oynaklık sürecine seyrek sıçramalar ilave ettiğimiz zaman, gerçekleşen varyans, bütünleşik varyans artı sıçramaya ait ikinci dereceden varyasyonunu tahmin ettiği halde, gerçekleşen ikili kuvvet varyasyonunun sadece bütünleşik varyansı tahmin etmeye devam etmesidir. Bu dayanıklılık önemlidir, çünkü logaritmik fiyat sürecine ait sıçrama bileşkeninden gelen ikinci dereceden varyasyonun süreksiz olan bileşenini ayırmaktadır. Böylece, eğer logaritmik fiyat süreci stokastik volatilite artı seyrek sıçramalar süreci sınıfına aitse, gerçekleşen varyans ve gerçekleşen ikili kuvvet varyasyonu arasındaki fark ikinci dereceden varyasyonun sürekli olmayan parçasını tahmin eder. Sonuç olarak, gerçekleşen sıçramalar yakalanmış olur.

Anahtar Kelimeler: Stokastik oynaklık, İkinci dereceden varyasyon, Kuvvet varyasyonu, İkili kuvvet varyasyonu, Sıçrama süreci, Gerçekleşen varyans, Gerçekleşen oynaklık, Yarımartengal, Bütünleşik varyans.

To my mother and father

ACKNOWLEDGMENTS

I would like to appreciate my supervisor, Assoc. Prof. Dr. Azize Hayfavi, for her excellent guidance and kindly supports throughout this thesis.

I would also like to thank Assist. Prof. Dr. Kasırga Yıldırak for his great assistance in MATLAB calculations and encouragements at the beginning of this study.

I am grateful to my parents for their patience and support from the beginning of my education life.

I am also grateful to Dr. Coşkun Küçüközmen for his precious lectures.

I wish thank to Sibel Korkmaz for her kindly assistance throughout this study. I would also like to thank Murat Akbayrak for his valuable encouragements and supports during this study. Finally, I also thank all my friends for their support.

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

1.1 **Introduction**

Volatility measurement, modelling and forecasting has always been to have important impacts on the risk management and asset allocation in finance literature (Figlewski(1997) [1], Ederington and Guan(2005) [2], Gospodinov, Gavolo and Jiang(2006) [3] ,Poon and Granger(2003) [4]). The recent widespread availability of high frequency data for many financial assets (stocks, currencies, bonds,...) has led to new interesting developments in applied econometrics and statistical research on the estimation of daily and intra-daily volatility of financial returns, particularly on the persistence of the volatility [5], [6], [7]. The search for an efficient framework for the estimation and prediction of the volatility of financial assets returns has offered the analysis of high frequency data [8]. The raw data are often obtained tick by tick with unequal time spacing [9], [10]. Organizing the data on the basis of equal fixed time intervals by interpolation [11], for example every five minutes, enables to perform the most well-known econometric models which are based on the assumption of equally spaced intervals of time.

Merton (1980) [12] has illustrated that the variance of an asset over a fixed time interval can be efficiently estimated as the squares of the intra-daily returns if the frequency of the returns is sufficiently high. Moreover, Andersen and Bollerslev (1998) [13] have demonstrated that the highest optimal frequency to provide accurate volatility forecasts is five minutes frequency.

In the recent finance literature, the most well-known method on modelling the daily volatility using intra-day data is *realized variance* (realized quadratic variation). Realized variance is based on quadratic variation and computed as the sum of squared intra-day returns (five min., ten min.,...) which is used to estimate daily volatility. The empirical and theoretical properties of this model has been popularized recently in a collection of papers such as Barndorff-Nielsen and Shephard (2002) [14], Barndorff-Nielsen and Shephard (2002) [15], Andersen, Bollerslev, Diebold, and Labys (2003) [16], Andersen, Bollerslev, and Meddahi (2005) [17], Bandi and Russell (2003) [18], Maheu (2004) [19], Andersen and Bollerslev (2006) [20].

The fundamental theory of asset pricing offers that the logarithmic price process of an asset is in the class of semimartingale processes [21] and obeys the properties of a semimartingale process. For economics aspects of semimartingale processes, see Back (1991) [22] and for mathematics aspects of semimartingale processes, see Protter (2004) [23]. The realized variance is based on the quadratic variation of logarithmic price processes so it is based on the quadratic variation of semimartingale processes (2004, p. 66-76) [23]. We can simply state that quadratic variation of a logarithmic price process is the probability limit of the realized variance. If the logarithmic price process is continuous, the realized variance estimates the integrated variance and if there exists jump part in the logarithmic price process then the realized variance estimates the integrated variance plus the quadratic variation of the jump component (2005) [24].

The generalized version of realized variance called *r-th order realized power variation* is introduced by Barndorff-Nielsen, Shephard and Graversen (2003) [25]. Realized variance is a special case of r-th order realized power variation when r=2. It is suitable to say that *r-th order power variation* is the probability limit of the r-th order realized power variation. Moreover,the r-th order realized power variation process is a consistent estimator of the r-th power of the integrated variance.

Furthermore, the generalization of r-th order realized power variation is the *(r,s)-order realized bipower variation* process which is defined by Barndorff-Nielsen and Shephard (2004) [26]. Also, *(r,s)-order bipower variation* is the probability limit of the (r,s)-order realized bipower variation process [21]. In the realized bipower variation process, the main difference is the summation of the iterated products of different powers of the returns unlike the realized power variation process. The r-th order realized power variation process and the realized variance are the special cases of the (r,s)-order realized bipower variation process. Namely, the r-th order realized power variation is the (r,0)-order realized bipower variation and the realized variance is the (2,0)-order realized bipower variation process.

The most important case of the bipower variation process is the (1,1)-order realized bipower variation process since this process is a consistent estimator of the integrated variance like the realized variance. Yet, the crucial role of the (1,1)-order realized bipower variation process in finance literature is that this estimator is robust to rare jumps in the logarithmic price process(2003) [25]. Simply, we can express this as; in both conditions whether the price process is continuous or discontinuous, (1,1)-order realized bipower variation estimates the integrated variance. In other words, rare jumps does not affect this estimator.

If the logarithmic price process is in the class of the stochastic volatility semimartingale plus rare jump processes (2003) [27], then the quadratic variation can be decomposed as continuous component (integrated variance) and discontinuous component (quadratic variation of the jumps). Realized variance estimates the quadratic variation and continuous component of the quadratic variation (2003) [27] can be estimated by the (1,1)-order realized bipower variation process. Therefore, the difference of realized variance and (1,1)-order realized bipower variation process estimates the quadratic variation of the jump part of the logarithmic price process (2003) [27]. Thus, jump detection can be achieved.

So, by the introduction of the (1,1)-order bipower variation, we can split up the quadratic variation of the logarithmic price process as the continuous part and the jump part. Quadratic variation has an important role in risk measurement (2003) [16]. By splitting up the quadratic variation into its continuous and discontinuous components, we separate the risk that comes from the diffusion part and the jump part.

In this study, we demonstrate the theoretical aspects of the realized variance, power variation, bipower variation and especially (1,1)-order bipower variation processes and the relations between these processes. We mainly focus on the difference of the realized variance and the (1,1)-order realized bipower variation process to achieve the estimation of the quadratic variation of the jump part, i.e. jump detection.

CHAPTER 2

PRELIMINARIES

2.1 Basic Definitions

Let $\mathcal{F}_t$ be an increasing family of σ-fields that reflects the corresponding information at time t where $\mathcal{F}_s \subseteq \mathcal{F}_t$ for $0 \leq s \leq t \leq T$ and $\mathcal{P}$ be the probability measure on the probability space $(\Omega,\mathcal{F},\mathcal{P})$.

Definition 2.1.1 *A random variable $\tau : \Omega \to [0, \infty]$ is a stopping time if for any t, the event $\{\tau \leq t\} \in \mathcal{F}_t$.*

Definition 2.1.2 *A stochastic process X on the probability space $(\Omega,\mathcal{F},\mathcal{P})$ is a collection of random variables X_t, where $0 \leq t < \infty$.*

Definition 2.1.3 *Let X be a stochastic process and let τ be a random time. The process X is stopped at time τ means that $X_t^\tau = X_{t\wedge\tau}$.*[1]

Definition 2.1.4 *A stochastic process X is adapted to the filtration $\mathcal{F}$ if X_t is $\mathcal{F}_t$ measurable, i.e. $X_t \in \mathcal{F}_t$, for all t.*

Definition 2.1.5 *A stochastic process X is càdlàg*[2] *if each sample path*[3] *is right continuous with left limit.*

---

[1] $X_{t\wedge\tau} = X_t 1_{\{t<\tau\}} + X_\tau 1_{\{t\geq\tau\}}$ where $t \wedge \tau = \min(t, \tau)$.
[2] *càdlàg* is an acronym from the French that is for continu à droite, limites à gauche.
[3] The mappings $t \to X_t$ are called the sample paths of the stochastic process X.

Corollary 2.1.1 *Let the stochastic process X be càdlàg and adapted and τ be a stopping time, then $X_t^\tau = X_{t\wedge\tau}$ is also adapted [23](p.10).*

Definition 2.1.6 *A stochastic process X is a finite variation process if it has sample paths of finite variation.*

Definition 2.1.7 *A stochastic process X is predictable means $X_t \in \mathcal{F}_{t-1}$where $0 \leq t < \infty$.*

Proposition 2.1.2 *(The Markov Inequality) [28] Let $X_n \in L_p$, $p \in [1, \infty)$. Then for all $x \in \Re$ and $\epsilon > 0$, we have*

$$P\{|X_n - X| \geq \epsilon\} \leq \frac{1}{\epsilon^p} E(|X_n - X|^p). \tag{2.1}$$

Proposition 2.1.3 *$p - \lim_{n\to\infty} X_n = X$ if and only if [28] for all $\epsilon > 0$, $P(\{|X_n - X| > \epsilon, i.o.\}) = 0$.*

2.2 Martingales

In this section, we give some basic definitions and theorems about martingales, mostly without the proofs, which will be essential to understand the theoretical aspects of semimartingales in the following chapters.

Definition 2.2.1 *A sequence M which is adapted to the filtration $\mathcal{F}$, with $E|M_t| < \infty$, is called martingale if*

$$E(M_t|\mathcal{F}_s) = M_s \tag{2.2}$$

for all $t \geq s$, where $0 \leq t \leq T$.

Definition 2.2.2 *A family of random variables $(U_\alpha)_{\alpha\geq 0}$ is uniformly integrable [29] if*

$$\lim_{n\to\infty}\sup_{\alpha}\int_{\{|U_\alpha|\geq n\}}|U_\alpha|d\mathcal{P}=0. \tag{2.3}$$

Theorem 2.2.1 ***(optional sampling theorem)****If $(M_t)_{t\geq 0}$ is a continuous martingale with respect to the filtration $(\mathcal{F}_t)_{t\geq 0}$ and if τ_1 and τ_2 are two stopping times such that $\tau_1\leq\tau_2\leq K$, where K is a finite real number [30](p.34), then M_{τ_2} is integrable*[4] *and*

$$E(M_{\tau_2}|\mathcal{F}_{\tau_1})=M_{\tau_1}. \tag{2.4}$$

Theorem 2.2.2 ***(Doob inequality)****If $(M_t)_{0\leq t\leq T}$ is a continuous martingale [30], we have*

$$E(\sup_{0\leq t\leq T}|M_t|^2)\leq 4E(|M_T|^2). \tag{2.5}$$

Definition 2.2.3 *We denote by $L^p(\Omega,\mathcal{F},\mathcal{P})$, $p\in[1,\infty)$, [28]the space of equivalence classes of all real random variables X such that*

$$\int_\Omega |X|^p d\mathcal{P}<\infty. \tag{2.6}$$

As a consequence of the definition, we denote $L^0(\Omega,\mathcal{F},\mathcal{P})$ the space of equivalence classes of all real random variables on $(\Omega,\mathcal{F},\mathcal{P})$.

And we denote $L^2(\Omega,\mathcal{F},\mathcal{P})$ the space of equivalence classes which are consist of all real random variables that have the property

$$\int_\Omega |X|^2 d\mathcal{P}<\infty. \tag{2.7}$$

[4] The random variable X is integrable [29](p.18) if and only if $E|X|<\infty$.

Theorem 2.2.3 ***(Doob's maximal quadratic inequality)****Let X be a positive submartingale. For all $p > 1$, with q conjugate to p[5], we have [23](p.11)*

$$\| \sup_t |X_t| \|_{L^p} \leq q \sup_t \|X_t\|_{L^p}. \tag{2.8}$$

We let X^* denote $\sup_s |X_s|$.

Note that if M is a martingale with[6] $M_\infty \in L^2$, then $|M|$ is a positive submartingale, taking $p = 2$ we have

$$E\{(M^\star)^2\} \leq 4E\{M_\infty^2\}. \tag{2.9}$$

The last inequality is called Doob's maximal quadratic inequality.

Definition 2.2.4 *X_t is said to be a local martingale[7] [31] (with respect to $\{\mathcal{F}_t, t \geq 0\}$) if there are stopping times $T_n \uparrow \infty$ so that $X_t^{T_n}$ is a martingale (with respect to $\{\mathcal{F}_{t \wedge T_n} : t \geq 0\}$). The stopping times T_n are said to reduce X.*

The definition of local martingale is so important because it has a crucial role in the definition of semimartingale. We will see in the following chapters that semimartingales are in the center of this thesis.

Next, we are giving some theorems that will be used later.

Theorem 2.2.4 *Let X be a local martingale such that $E\{X_t^*\} < \infty$ [8] for every $t \geq 0$. Then X is a martingale. If $E\{X^\star\} < \infty$[9] ,then X is a uniformly integrable martingale [23](p.38).*

[5] q conjugate to p means $\frac{1}{p} + \frac{1}{q} = 1$.

[6] Let $(X_t)_{0 \leq t \leq \infty}$ be a martingale, if $Y = \lim_{t \to \infty} X_t$ exists a.s., $E\{|Y|\} < \infty$, then $X_\infty = Y$.

[7] In general, we say that a process Y is locally A if there is a sequence of stopping time $T_n \uparrow \infty$ so that the stopped process Y_t^T has property A [31].

[8] Recall that $X_t^* = \sup_{s \leq t} |X_s|$.

[9] Recall that $X^\star = \sup_s |X_s|$.

Theorem 2.2.5 *Let X be a locally square integrable local martingale, and let H be an adapted process with càdlàg paths. Then the stochastic integral $\int HdX$ is also a locally square integrable local martingale [23](p.63).*

Theorem 2.2.6 *Let M be a local martingale. Then M is a martingale with $E\{M_t^2\} < \infty$, all $t \geq 0$, if and only if $E\{[M]_t\} < \infty$[10], all $t \geq 0$. If $E\{[M]_t\} < \infty$, then [23](p.73)*

$$E\{M_t^2\} = E\{[M]_t\}. \tag{2.10}$$

2.3 Brownian motions

Definition 2.3.1 *A Brownian motion is a real-valued, continuous stochastic process $(X_t)_{t\geq 0}$, with independent and stationary increments [23](p.31). In other words:*

- *continuity: $\mathcal{P}$ a.s.[11] the map $s \rightarrow X_s(w)$ is continuous.*
- *independent increments: If $s \leq t$, $X_t - X_s$ is independent of $\mathcal{F}_s = \sigma(X_u, u \leq s)$.*
- *stationary increments: If $s \leq t$, $X_t - X_s$ and $X_{t-s} - X_0$ have the same probability law[12].*

Theorem 2.3.1 *If $(X_t)_{t\geq 0}$ is a Brownian motion[13], then $X_t - X_0$ is a normal random variable with mean rt and variance $\sigma^2 t$, where r and σ are constant real*

[10] We denote the quadratic variation of the process X as $[X]_t$.

[11] If a set $A \in \mathcal{F}$ satisfies $\mathcal{P}(A) = 1$, we say that A occurs almost surely denoted as a.s. [29](p.7)

[12] This statement is denoted as $X_t - X_s =^d X_{t-s} - X_0$.

[13] Recall that brownian motion is a martingale [29].

numbers [30](p.31).

Definition 2.3.2 *A Brownian motion is standard [30](p.31) if $X_0 = 0$ $\mathcal{P}$ a.s. $E(X_t) = 0$, $E(X_t^2) = t = var(X_t)$. In that case, the distribution of X_t is the following:*

$$\frac{1}{\sqrt{2\Pi t}} exp(-\frac{x^2}{2t})dx, \tag{2.11}$$

where dx is the Lebesgue measure on $\mathcal{R}$.

2.4 Introduction to semimartingales

Semimartingales are important processes in the stochastic integration. Firstly, we will construct the stochastic integral for the simple processes. Let us define the simple processes.

Suppose that $(W_t)_{t\geq 0}$ is a standard $\mathcal{F}_t$-Brownian motion defined on a filtered probability space $(\Omega, (\mathcal{F}_t)_{t\geq 0}, \mathcal{F}, \mathcal{P})$.

Definition 2.4.1 *$(H_t)_{0\leq t\leq T}$ is called a simple process [30](p.36) if it can be written as*

$$H_t(w) = \sum_{i=1}^{p} \phi_i(w) 1_{(t_{i-1},t_i]}(t) \tag{2.12}$$

where $0 = t_0 < t_1 < ... < t_p = T$ and ϕ_i is $\mathcal{F}_{t_{i-1}}$-measurable and bounded.

Definition 2.4.2 *The stochastic integral of a simple process H [30](p.36) is the* *continuous process* $(I(H)_t)_{0\leq t\leq T}$ *defined for any* $t \in (t_k, t_{k+1}]$ *as*

$$
\begin{aligned}
I(H)_t = &\sum_{1\leq i\leq k} \phi_i(W_{t_i} - W_{t_{i-1}}) + \phi_{k+1}(W_t - W_{t_k}) \\
= &\sum_{1\leq i\leq p} \phi_i(W_{t_i\wedge t} - W_{t_{i-1}\wedge t}) \\
= &\int_0^t H_s dW_s.
\end{aligned}
\tag{2.13}
$$

Proposition 2.4.1 *If* $(H_t)_{0\leq t\leq T}$ *is a simple process [30](p.36) then:*

- $(\int_0^t H_s dW_s)_{0\leq t\leq T}$ *is a continuous* $\mathcal{F}_t$*-martingale,*
- $E((\int_0^t H_s dW_s)^2) = E(\int_0^t H_s^2 ds),$
- $E(\sup_{t\leq T} |\int_0^t H_s dW_s|^2) \leq 4E(\int_0^T H_s^2 ds).$

Now, we come to the definition of semimartingale processes.

Definition 2.4.3 *A process X is said to be a total semimartingale if X is càdlàg, adapted, and* $I(X)$ *(given by eqn.(2.13)) is continuous [23](p.52).*

Definition 2.4.4 *A process X is called semimartingale if, for each* $t \in [0,\infty)$, X^t *is*[14] *a total semimartingale [23](p.52).*

Theorem 2.4.2 *Some examples of semimartingales are as the following [23]:*

- *Each adapted process with càdlàg paths of finite variation on compacts is a semimartingale.*
- *Each* L^2 *martingale with càdlàg paths is a semimartingale.*
- *Each càdlàg, locally square integrable local martingale is a semimartingale.*
- *The Wiener process, i.e. Brownian motion, is a semimartingale.*

[14] Recall that for a process X and a stopping time T, the notation X^T denotes the process $(X_{t\wedge T})_{t\geq 0}$.

Definition 2.4.5 *We will say an adapted process X with càdlàg paths is decomposable*[15] *if it can be decomposed*

$$X_t = M_t + A_t, \tag{2.14}$$

where $M_0 = A_0 = 0$, M is a locally square integrable martingale, and A is càdlàg, adapted, with paths of finite variation.

Definition 2.4.6 *A process Y_t is said to be a classical semimartingale if it can be decomposed into two adapted, càdlàg processes*

$$Y_t = M_t + A_t, \text{ (eqn. 2.14)}$$

where A_t is a locally finite variation process and M_t is a local martingale where $A(0) = M(0) = 0$.

Theorem 2.4.3 *Let X be an adapted, càdlàg process. The following are equivalent [23](p.102):*

(i) X is a semimartingale;

(ii) X is decomposable;

(iii) given $\beta > 0$, there exist M, A with $M_0 = A_0 = 0$, M is a local martingale with jumps bounded by β, A a finite variation process, such that $X_t = M_t + A_t$;

(iv) X is a classical semimartingale.

2.5 The quadratic variation of a semimartingale

We are giving some properties of the quadratic variation of a semimartingale.

[15] A decomposable process is a semimartingale [23](p.55).

Definition 2.5.1 *Let X, Y be semimartingales. The quadratic variation process of X, denoted $[X,X] = ([X,X])_{t\geq 0}$, is defined by [23](p.66)*

$$[X,X] = X^2 - 2\int X_- dX \tag{2.15}$$

(recall that $X_{0-} = 0$).

The quadratic covariation of X and Y, also called the bracket process of X and Y, is defined by

$$[X,Y] = XY - \int X_- dY - \int Y_- dX. \tag{2.16}$$

Corollary 2.5.1 *The bracket process $[X,Y]$ of two semimartingales has paths of finite variation, and it is also a semimartingale [23].*

Definition 2.5.2 *For a semimartingale X, the process $[X,X]^c$ denotes the path-by-path continuous part of $[X,X]$.*

We can then write [23],

$$\begin{aligned} [X,X]_t = & \; [X,X]^c_t + X_0^2 + \sum_{0<s\leq t} (\Delta X_s)^2 \\ = & \; [X,X]^c_t + \sum_{0\leq s\leq t} (\Delta X_s)^2. \end{aligned} \tag{2.17}$$

Observe that $[X,X]^c_0 = 0$. Analogously, $[X,Y]^c$ denotes the path-by-path continuous part of $[X,Y]$.

2.6 Counting Processes

A random point process is a mathematical model for a physical phenomenon characterized by highly localized events distributed randomly in a continuum.

Each localized event is represented in the model by an idealized point to be conceived of as identifying the position of the event in the continuum. If $\aleph$ denotes the continuum space, then a realization of a random point process on $\aleph$ is a set of points having coordinates in $\aleph$ [32].

It is often of interest in practice to count numbers of points in subsets of the space $\aleph$ on which a point process is defined. A counting process is introduced for this purpose and can be associated with every point process. The idea is as follows. We assume that a realization, call it ω, of a stochastic point process on $\aleph$ is a denumerable point set of $\aleph$. This means that ω can be enumerated as $\omega = \{X_1, X_2, ...\}$, where each X_i denotes the coordinate of a point in $\aleph$. Now, let A be a subset of $\aleph$, and denote by $N(A;\omega)$ the number of points in ω that lie in A. Formally,

$$N(A;\omega) = \sum_i 1_A(X_i) \tag{2.18}$$

where each X_i is the coordinate of a point in ω. When viewed as a function of A and ω, $N(A;\omega)$ defines a nonnegative, integer-valued random process on $\aleph$. A process constructed this way is called a counting process [32].

Definition 2.6.1 *Let N(t) be a counting process that counts the jumps for the interval from 0 up to t. N(t) is said to be finite activity simple counting process if there exist finite jumps in finite interval of time*[16].

[16] i.e. $N(t) < \infty$ for all $0 < t < \infty$.

CHAPTER 3

REALIZED VARIANCE AND QUADRATIC VARIATION

3.1 Arbitrage-free, frictionless price processes

Consider an arbitrage-free price process P_t where T is a positive integer and $t \in [0, T]$. Let $\mathcal{F}_t$ be an increasing family of σ-fields that reflects the corresponding information at time t where $\mathcal{F}_s \subseteq \mathcal{F}_t$ for $0 \leq s \leq t \leq T$ and $\mathcal{P}$ be the probability measure on the probability space (Ω,$\mathcal{F}$,$\mathcal{P}$). Assume that $\mathcal{F}_t$ satisfies $\mathcal{P}$-completeness[1] and right-continuity[2] and P_t is $\mathcal{F}_t$-measurable (P_t is adapted) which means that asset prices through time t are included in the information filtration $\mathcal{F}_t$.

In the fundamental theory of asset pricing, each logarithmic price process Y_t at time t is a semimartingale [22]. From the definition of a semimartingale, Y_t can be decomposed into two adapted, c*à*dl*à*g processes

$$Y_t = A_t + M_t, \qquad \text{(eqn.(2.14))}$$

where A_t is a locally finite variation process and M_t is a local martingale where $A(0) = M(0) = 0$.

[1] Let (Ω,$\mathcal{F}$,$\mathcal{P}$) be a probability space. A subset N of Ω is said to be negligible if there is a set Fin $\mathcal{F}$ such that $N \subset F$ and $\mathcal{P}(F) = 0$. If $\mathcal{F}$ contains all the negligible sets then (Ω,$\mathcal{F}$,$\mathcal{P}$) is said to be complete [28].

[2] $\mathcal{F}_t = \cap_{u>t}\mathcal{F}_u$, all t, $0 \leq t < \infty$; that is, the filtration $\mathcal{F}$ is right continuous [23].

Now, let us introduce some notation. Suppose that $s > 0$ is some fixed time period (e.g. a month or a day) and as mentioned above the logarithmic price of an asset is written as Y_t for $t \in [0, T]$. Then the i-th s (e.g. i-th day or i-th month) *low frequency return* is

$$y(i) = Y(is) - Y((i-1)s), \quad i = 1, 2, ... \tag{3.1}$$

Suppose that we have the intra-period prices at n equally spaced time points during the i-th s (e.g. i-th day or i-th month). Then, we can define *high frequency return* as

$$y(i,j) = Y((i-1)s + sjn^{-1}) - Y((i-1)s + s(j-1)n^{-1}), \quad j = 1, 2, ..., n. \tag{3.2}$$

Here, $y(i,j)$ is the j-th intra-s return for the i-th s. Let us illustrate the notation, suppose that s is a trading day and you have 5-min intra-day returns for the i-th trading day. So, $n = 60$ since we consider only the five hours of the trading day and $y(i) = \sum_{j=1}^{60} y(i,j)$ (i-th day return is the sum of the 5-min intra-day returns).

In figure 3.1, we can see the price data for euro/FX cross rate data in different frequencies. For 5 minutes data, we get $n = 60$. In the same way, for 30 minutes data, we have $n = 10$. For hourly data, $n = 5$ and for the daily data, $n = 1$.

The i-th realized variance is based on squares of these high frequency returns in the interval from $s(i-1)$ to si. To understand realized variance explicitly, we will define quadratic variation (QV) process of Y_t. From now on, we will write the definitions for the interval from time 0 to t, $t \in [0, T]$.

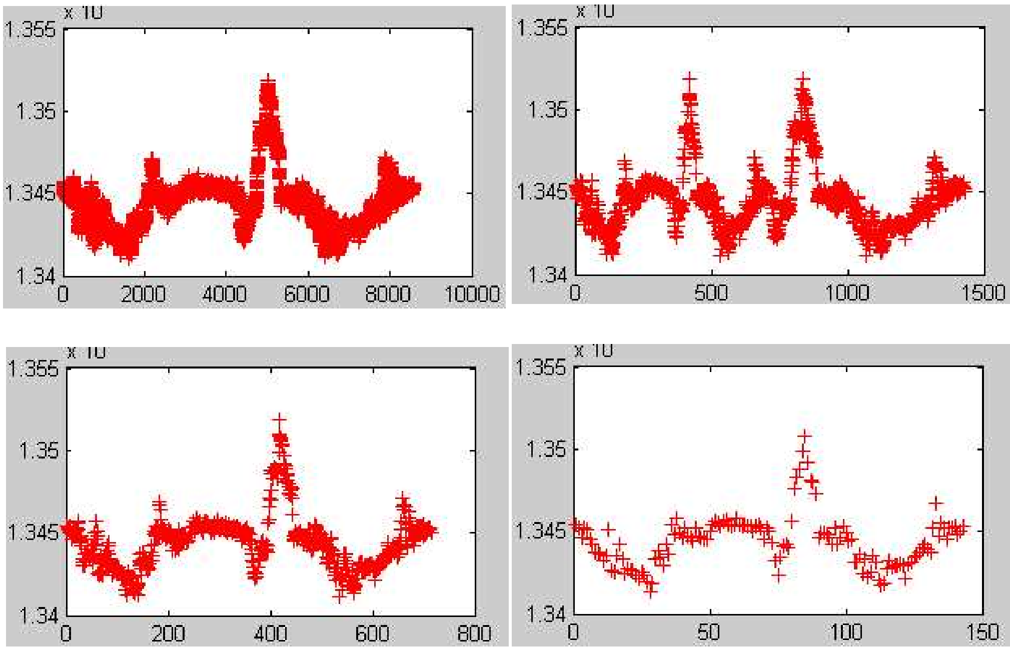

Figure 3.1: Euro/FX cross rate dataset. Top left: raw 5 minute price data. Top right: raw 30 minutes price data. Bottom left: raw hourly price data. Bottom right: daily raw price data.

3.2 The Relation between Realized Variance and Quadratic Variation

Definition 3.2.1 *The quadratic variation (QV) process of the logarithmic price process Y_t is defined as*

$$
\begin{aligned}
[Y]_t &= p - \lim_{n\to\infty} \sum_{j=1}^{n} (Y(t_j) - Y(t_{j-1}))^2 \\
&= p - \lim_{n\to\infty} \sum_{j=1}^{n} (y(t_j))^2
\end{aligned}
\qquad (3.3)
$$

for any sequence of partitions $t_0 = 0 < t_1 < ... < t_n = t$ with $\sup_j \{t_j - t_{j-1}\} \to 0$ for $n \to \infty$.

Suppose that we record the prices for the time intervals of the same length $h > 0$ between 0 and t. In other words, we divide the interval $[0, t]$ into equidistant intervals of length h, $0 = t_0 < t_1 < ... < t_{\lfloor \frac{t}{h} \rfloor} = t$, recalling that $\lfloor \frac{t}{h} \rfloor$ is the integer part of $\frac{t}{h}$. Then, the j-th h-return is shown as

$$y(j) = Y(jh) - Y((j-1)h), \quad j = 1, 2, ..., \lfloor \frac{t}{h} \rfloor, \tag{3.4}$$

these returns are used to construct realized quadratic variation process (realized variance) which is defined by

$$[Y_h]_t = \sum_{j=1}^{\lfloor \frac{t}{h} \rfloor} (y(j))^2, \tag{3.5}$$

the sum of the squares of the intra-day returns for the period from 0 to t.

As mentioned above, in the definition of quadratic variation, subintervals between 0 and t are of different lengths and the limit is taken as the supremum of the intervals goes to 0. But in the definition of the realized quadratic variation process, subintervals between 0 and t are of the same length. Now, as seen from the definitions of quadratic variation and realized quadratic variation, quadratic variation is the limit of the realized quadratic variation as $h \to 0$. That is,

$$[Y]_t = p - \lim_{h \to 0} [Y_h]_t. \tag{3.6}$$

To simplify the calculations, we assume $\frac{t}{h}$ is an integer and equal to n. So, the realized quadratic variation is

$$\begin{aligned} [Y_h]_t &= \sum_{j=1}^{n} (y(j))^2 \\ &= \sum_{j=1}^{n} (y(t_j))^2 \\ &= \sum_{j=1}^{n} (Y(t_j) - Y(t_{j-1}))^2, \end{aligned} \tag{3.7}$$

and the quadratic variation is

$$[Y]_t = p - \lim_{n\to\infty}[Y_h]_t \qquad \text{(eqn. 3.6)}$$

since $h \to 0$ is equivalent to $n \to \infty$.

In general, the logarithmic price process Y_t can exhibit jumps and since Y_t is a semimartingale process, we can decompose Y_t as

$$Y_t = Y_t^{ct} + Y_t^d, \tag{3.8}$$

where Y_t^{ct} and Y_t^d are respectively purely continuous and discontinuous components of Y_t. As mentioned before, Y_t is decomposable as two adapted and c*àdlà*g processes A_t and M_t which are respectively locally finite variation process and a local martingale. Then we can decompose the continuous and the discontinuous parts of Y_t as

$$Y_t^{ct} = A_t^{ct} + M_t^{ct} \tag{3.9}$$

$$Y_t^d = A_t^d + M_t^d, \tag{3.10}$$

where the finite variation components A_t^{ct} and A_t^d are respectively continuous and jump processes, while the local martingales M_t^{ct} and M_t^d are respectively continuous and jump processes.

It is convenient to consider the economics aspects of semimartingales and quadratic variation. Let us introduce some theorems to understand the theory of the semimartingales and quadratic variation of semimartingales.

Theorem 3.2.1 *The logarithmic price process Y_t is a semimartingale and the quadratic variation of Y_t denoted as $[Y]_t$ can be decomposed into the quadratic* *variation of the continuous part and the discontinuous part*

$$\begin{aligned}[Y]_t = \ & [Y^{ct}]_t + [Y^d]_t \\ = \ & [Y^{ct}]_t + \sum_{0 \le u \le t} (\Delta Y_u)^2\end{aligned} \tag{3.11}$$

where $\Delta Y_u = Y_u - Y_{u^-}$ *is the jump at time u.*

Proof: Y_t can be decomposed as $Y_t = Y_t^{ct} + Y_t^d$ (from the equation(3.8)) and $Y_t^d = \sum_{0 \le u \le t} \Delta Y_u$ where $\Delta Y_u = Y_u - Y_{u^-}$ is the jump at time u.

The definition of the quadratic variation for the equally spaced intervals is given as follows

$$\begin{aligned}
[Y]_t = \ & p - \lim_{n \to \infty} \sum_{j=1}^{n} (Y(t_j) - Y(t_{j-1}))^2 \quad (eqn.(3.3)) \\
= \ & p - \lim_{n \to \infty} \sum_{j=1}^{n} (Y^{ct}(t_j) + Y^d(t_j)) - (Y^{ct}(t_{j-1}) + Y^d(t_{j-1}))^2 \quad (eqn.(3.8)) \\
= \ & p - \lim_{n \to \infty} \sum_{j=1}^{n} ((Y^{ct}(t_j) - Y^{ct}(t_{j-1})) + (Y^d(t_j) - Y^d(t_{j-1})))^2 \\
= \ & p - \lim_{n \to \infty} \sum_{j=1}^{n} (Y^{ct}(t_j) - Y^{ct}(t_{j-1}))^2 \\
+ \ & p - \lim_{n \to \infty} 2 \sum_{j=1}^{n} (Y^{ct}(t_j) - Y^{ct}(t_{j-1}))(Y^d(t_j) - Y^d(t_{j-1})) \\
+ \ & p - \lim_{n \to \infty} \sum_{j=1}^{n} (Y^d(t_j) - Y^d(t_{j-1}))^2.
\end{aligned}$$

For the sum of the squares of the continuous part, we have the limit as

$$p - \lim_{n \to \infty} \sum_{j=1}^{n} (Y^{ct}(t_j) - Y^{ct}(t_{j-1}))^2 = [Y^{ct}]_t. \tag{3.12}$$

For the sum of the squares of the discontinuous part, we have the limit as

$$p - \lim_{n \to \infty} \sum_{j=1}^{n} (Y^d(t_j) - Y^d(t_{j-1}))^2 = p - \lim_{n \to \infty} \sum_{j=1}^{n} (\Delta Y_u)^2 = [Y^d]_t. \tag{3.13}$$

For the sum of the cross products,

$$(Y^{ct}(t_j) - Y^{ct}(t_{j-1})) \leq \max_{1 \leq j \leq n} |Y^{ct}(t_j) - Y^{ct}(t_{j-1})| \tag{3.14}$$

for all j=1,...,n then we get

$$\sum_{j=1}^{n} (Y^{ct}(t_j) - Y^{ct}(t_{j-1}))(Y^{d}(t_j) - Y^{d}(t_{j-1}))$$

$$\begin{aligned} &\leq \max_{1 \leq j \leq n} |Y^{ct}(t_j) - Y^{ct}(t_{j-1})| \sum_{j=1}^{n} (Y^{d}(t_j) - Y^{d}(t_{j-1})) \quad (eqn.(3.14)) \\ &= \max_{1 \leq j \leq n} |Y^{ct}(t_j) - Y^{ct}(t_{j-1})| \sum_{j=1}^{n} (\Delta Y_u). \end{aligned}$$

And also since Y_t^{ct} is continuous, we will use the property

$$p - \lim_{n \to \infty} \max_{1 \leq j \leq n} |Y^{ct}(t_j) - Y^{ct}(t_{j-1})| = 0 \tag{3.15}$$

and $\sum_{j=1}^{n} (\Delta Y_u)$ is finite. Thus, the result comes

$$p - \lim_{n \to \infty} 2 \sum_{j=1}^{n} (Y^{ct}(t_j) - Y^{ct}(t_{j-1}))(Y^{d}(t_j) - Y^{d}(t_{j-1})) = 0. \quad (eqn.(3.15))$$

Hence,by the equations (3.12), (3.13) and (3.15), we have

$$[Y]_t = [Y^{ct}]_t + [Y^d]_t. \quad (eqn.(3.11))$$

Q.E.D.

Theorem 3.2.2 *The logarithmic price process Y_t at time t is a semimartingale and can be decomposed into two adapted, càdlàg processes*

$$Y_t = A_t + M_t \qquad (eqn.(2.14))$$

where A_t is a locally finite variation process and M_t is a local martingale process where $A(0) = M(0) = 0$. In addition to these settings, if A_t is continuous, i.e., $A_t^d = 0$ and if M_t is a martingale then

$$E(dY_t|\mathcal{F}_t) = dA_t \tag{3.16}$$

and the drift part can be considered as

$$A_t = \int_0^t E(dY_u|\mathcal{F}_u)du \tag{3.17}$$

so we can say that A_t is the integral of the expectation of the instantaneous returns from 0 to t.

Proof: Let us start with the decomposition of semimartingale Y_t

$$Y_t = A_t + M_t \qquad \text{(eqn.(2.14))}$$

and this implies the instantaneous logarithmic price process dY_t can also be decomposed as

$$dY_t = dA_t + dM_t \tag{3.18}$$

Now, let us take the expectation of both sides

$$E(dY_t|\mathcal{F}_t) = E[(dA_t|\mathcal{F}_t) + (dM_t|\mathcal{F}_t)]$$

and since conditional expectation is a linear operator, we have

$$E(dY_t|\mathcal{F}_t) = E(dA_t|\mathcal{F}_t) + E(dM_t|\mathcal{F}_t) \tag{3.19}$$

and since M_t is a martingale we have

$$E(dM_t|\mathcal{F}_t) = 0. \tag{3.20}$$

Then, we can write the above equation as

$$E(dY_t|\mathcal{F}_t) = E(dA_t|\mathcal{F}_t) \tag{3.21}$$

by the equations (3.19), (3.20) and since A_t is adapted,

$$E(dA_t|\mathcal{F}_t) = dA_t. \tag{3.22}$$

Thus, by the equations (3.21) and (3.22), we have the result

$$E(dY_t|\mathcal{F}_t) = dA_t \quad \text{(eqn.(3.16))}$$

which proves the theorem. Q.E.D.

Theorem 3.2.3 *The logarithmic price process Y_t at time t is a semimartingale that can be decomposed into two adapted and càdlàg processes*

$$Y_t = A_t + M_t \qquad \textit{(eqn.(2.14))}$$

where A_t is a locally finite variation process and M_t is a local martingale, $A(0) = M(0) = 0$. In addition to these settings, if A_t is continuous, i.e., $A_t^d = 0$ and if M_t is a continuous martingale which is locally square integrable, i.e., $E\{M_t^2\} < \infty$ for all $t \geq 0$ and A_t is independent of the martingale process M_t then

$$[Y]_t = [M]_t \tag{3.23}$$

and

$$\begin{aligned} [Y]_t &= \int_0^t var(dY_u|\mathcal{F}_u)du \\ &= \int_0^t var(dM_u|\mathcal{F}_u)du \end{aligned} \tag{3.24}$$

that is the quadratic variation of the logarithmic price process equals to the integrated variance of the instantaneous returns.

Proof: Let us start with the decomposition of Y_t

$$Y_t = A_t + M_t \qquad \text{(eqn.(2.14))}$$

and this implies the instantaneous logarithmic price process dY_t can also be decomposed as

$$dY_t = dA_t + dM_t \qquad \text{(eqn.(3.18))}$$

Now, from the definition of variance,

$$var(dY_t|\mathcal{F}_t) = E((dY_t)^2|\mathcal{F}_t) - (E(dY_t|\mathcal{F}_t))^2. \tag{3.25}$$

By the equation(3.16), we have

$$\begin{aligned}
var(dY_t|\mathcal{F}_t) = \;& E((dY_t)^2|\mathcal{F}_t) - (dA_t)^2 \\
= \;& E((dA_t + dM_t)^2|\mathcal{F}_t) - (dA_t)^2 \quad (eqn.(3.18)) \\
= \;& E(((dA_t)^2 + (dM_t)^2 + 2dA_t dM_t)|\mathcal{F}_t) - (dA_t)^2 \\
= \;& E((dA_t)^2|\mathcal{F}_t) + E((dM_t)^2|\mathcal{F}_t) + 2E((dA_t dM_t)|\mathcal{F}_t) - (dA_t)^2
\end{aligned}$$

A_t is adapted so that $(dA_t)^2$ is also adapted, i.e.,

$$E((dA_t)^2|\mathcal{F}_t) = (dA_t)^2. \tag{3.26}$$

Now, by the equation (3.26),

$$\begin{aligned}
var(dY_t|\mathcal{F}_t) = \;& (dA_t)^2 + E((dM_t)^2|\mathcal{F}_t) + 2E((dA_t dM_t)|\mathcal{F}_t) - (dA_t)^2 \\
= \;& E((dM_t)^2|\mathcal{F}_t) + 2E((dA_t dM_t)|\mathcal{F}_t).
\end{aligned}$$

In the assumptions of the theorem, we have that dA_t is independent of dM_t then

$$E((dA_t dM_t)|\mathcal{F}_t) = E((dA_t)|\mathcal{F}_t)E((dM_t)|\mathcal{F}_t). \tag{3.27}$$

By the equation (3.20), the equation (3.27) becomes

$$E((dA_t dM_t)|\mathcal{F}_t) = 0.$$

Then the conditional variance of the instantaneous price process can be written as

$$var(dY_t|\mathcal{F}_t) = E((dM_t)^2|\mathcal{F}_t). \tag{3.28}$$

Now, to see the equation (3.24), let us write the conditional variance for the martingale part of the instantaneous price process

$$\begin{aligned} var(dM_t|\mathcal{F}_t) = & \; E((dM_t)^2|\mathcal{F}_t) - (E((dM_t)|\mathcal{F}_t)^2 \\ = & \; E((dM_t)^2|\mathcal{F}_t) \end{aligned} \tag{3.29}$$

from the equation (3.20). Thus, we have

$$var(dY_t|\mathcal{F}_t) = var(dM_t|\mathcal{F}_t). \tag{3.30}$$

Now, our assumptions on M_t are that M_t is a local martingale and locally square integrable so we can apply theorem (2.2.6) for the martingale part of the instantaneous price process. Then we have the equation

$$E\{(dM_t^2)|\mathcal{F}_t\} = E\{d[M]_t|\mathcal{F}_t\}. \tag{3.31}$$

When we combine the equations (3.28), (3.29), (3.30) and (3.31), we get

$$\begin{aligned} var(dY_t|\mathcal{F}_t) = & \; var(dM_t|\mathcal{F}_t) \\ = & \; E((dM_t)^2|\mathcal{F}_t) \\ = & \; E\{d[M]_t|\mathcal{F}_t\}. \end{aligned} \tag{3.32}$$

Since $var(dY_t|\mathcal{F}_t)$ is $\mathcal{F}_t$ measurable, we have

$$var(dY_t|\mathcal{F}_t) = E(var(dY_t|\mathcal{F}_t)|\mathcal{F}_t). \tag{3.33}$$

The equations (3.33) and (3.32), we get

$$E(var(dY_t|\mathcal{F}_t)|\mathcal{F}_t) = E\{d[M]_t|\mathcal{F}_t\}. \tag{3.34}$$

Thus, we have

$$var(dY_t|\mathcal{F}_t) = d[M]_t. \tag{3.35}$$

Now, take the integral of both sides from 0 to t

$$[M]_t = \int_0^t var(dY_u|\mathcal{F}_u)du. \tag{3.36}$$

Finally, to finish the proof of the theorem, we will show the equation (3.23), i.e.,

$$[Y]_t = [M]_t.$$

The quadratic variation of Y_t is defined as

$$[Y]_t = p - \lim_{n\to\infty} \sum_{j=1}^{n} (Y(t_j) - Y(t_{j-1}))^2 \qquad \text{(eqn.(3.3))}$$

for any sequence of partitions $t_0 = 0 < t_1 < ... < t_n = t$ with equidistant intervals of length $h > 0$ and with the assumption that $\lfloor \frac{t}{h} \rfloor = n$ where $h \to 0$ as $n \to \infty$. Consider the sum (the realized variance),

$$\begin{aligned}
\sum_{j=1}^{n}(Y(t_j) - Y(t_{j-1}))^2 = & \sum_{j=1}^{n}((A(t_j) + M(t_j)) - (A(t_{j-1}) + M(t_{j-1})))^2 \\
= & \sum_{j=1}^{n}((A(t_j) - A(t_{j-1})) + (M(t_j) - M(t_{j-1})))^2 \\
= & \sum_{j=1}^{n}(A(t_j) - A(t_{j-1}))^2 \\
+ & \ 2\sum_{j=1}^{n}(A(t_j) - A(t_{j-1}))(M(t_j) - M(t_{j-1})) \\
+ & \sum_{j=1}^{n}(M(t_j) - M(t_{j-1}))^2.
\end{aligned}$$

Now, the first sum

$$\sum_{j=1}^{n}(A(t_j) - A(t_{j-1}))^2 \leq \max_{1\leq j\leq n} |A(t_j) - A(t_{j-1})| \sum_{j=1}^{n}(A(t_j) - A(t_{j-1}))$$

and since A_t is of finite variation process with continuous paths, we have

$$p - \lim_{n\to\infty} \max_{1\leq j\leq n} |A(t_j) - A(t_{j-1})| = 0. \tag{3.37}$$

Thus, the limit of the first sum

$$p - \lim_{n\to\infty} \sum_{j=1}^{n}(A(t_j) - A(t_{j-1}))^2 = 0. \tag{3.38}$$

Now, the second sum

$$\begin{gathered}
\sum_{j=1}^{n}(A(t_j) - A(t_{j-1}))(M(t_j) - M(t_{j-1})) \leq \\
\max_{1\leq j\leq n} |M(t_j) - M(t_{j-1})| \sum_{j=1}^{n}(A(t_j) - A(t_{j-1})).
\end{gathered}$$

We can apply the equation (3.37) to the martingale part since M_t is also continuous and since $\sum_{j=1}^{n}(A(t_j) - A(t_{j-1}))$ is finite, the limit of the second sum

$$p - \lim_{n\to\infty} \sum_{j=1}^{n}(A(t_j) - A(t_{j-1}))(M(t_j) - M(t_{j-1})) = 0. \tag{3.39}$$

Finally, by the equations (3.38) and (3.39), the quadratic variation of Y_t becomes

$$\begin{aligned} [Y]_t = & \ p - \lim_{n\to\infty} \sum_{j=1}^{n}(Y(t_j) - Y(t_{j-1}))^2 \\ = & \ p - \lim_{n\to\infty} \sum_{j=1}^{n}(M(t_j) - M(t_{j-1}))^2 \\ = & \ [M]_t. \end{aligned}$$

So, we come to the equation (3.23), i.e.

$$[Y]_t = [M]_t$$

and by the equations (3.36) and (3.30), we have the result

$$\begin{aligned} [Y]_t = & \int_0^t var(dY_u|\mathcal{F}_u)du \\ = & \int_0^t var(dM_u|\mathcal{F}_u)du \end{aligned}$$

which proves the theorem. Q.E.D.

In addition to Theorem (3.2.3), if we allow discontinuities in the martingale component of prices and continue to assume that the path of finite variation process A_t is continuous, we have the quadratic variation of price process as

$$\begin{aligned} [Y]_t = & \ [M^{ct}]_t + [M^d]_t \\ = & \int_0^t var(dM_u^{ct}|\mathcal{F}_u)du + \sum_{u=1}^{N(t)} \Delta M_u{}^2 \end{aligned} \tag{3.40}$$

where $N(t)$ is a simple counting process which is assumed to be finite for all t and ΔM_u are non-zero random variables which are the size of the jumps.

We can summarize what we have stated as

$$[Y_h]_t = \sum_{j=1}^{n} (y_j)^2, \quad (\text{eqn.}(3.7))$$

where $0 = t_0 < t_1 < ... < t_n = t$ and

$$\lim_{n \to \infty} [Y_h]_t = [Y]_t, \quad (\text{eqn.}(3.6))$$

i.e. realized variance estimates the quadratic variation of the price process. If we assume that Y_t is continuous then we have

$$[Y]_t = [M]_t \quad (\text{eqn.}(3.23)).$$

In this case,

$$\lim_{n \to \infty} [Y_h]_t = [M]_t, \tag{3.41}$$

i.e. realized variance estimates the quadratic variation of the martingale component of the price process. If we decompose $[M]_t$ in its continuous and discontinuous parts, we have

$$\begin{aligned} \lim_{n \to \infty} [Y_h]_t = & \; [M^{ct}]_t + [M^d]_t \\ = & \int_0^t var(dM_u^{ct} | \mathcal{F}_u) du + \sum_{u=1}^{N(t)} \Delta M_u{}^2, \end{aligned} \tag{3.42}$$

i.e. realized variance estimates the integrated variance of the instantaneous returns of the continuous part of M_t plus the quadratic variation of the jump part of M_t.

3.3 Stochastic Volatility Semimartingales

Recall that the logarithmic price process is a semimartingale and can be decomposed into two adapted and *càdlàg* processes

$$Y_t = A_t + M_t \qquad \text{(eqn. 2.14)}$$

where A_t is a locally finite variation process and M_t is a local martingale, $A(0) = M(0) = 0$.

In this section, we have some additional assumptions on A_t and M_t. Firstly, we assume M_t is continuous and it is an Ito integral of spot volatility process, $\sigma_t > 0$, with respect to a standard Brownian motion relative to the filtration $\mathcal{F}_t$ then

$$M_t = \int_0^t \sigma_u dW(u) \tag{3.43}$$

where the spot volatility process, $\sigma_t > 0$, is an adapted, *càdlàg* and locally bounded away from zero and also W_t is the standard Brownian motion. In this setting, we can define the integrated variance as

$$\vartheta_t^2 = \int_0^t \sigma_u^2 du, \tag{3.44}$$

where $\vartheta_t^2 < \infty$ and this equality is defined for all $t < \infty$. Secondly, we assume A_t is continuous and it is the Riemann integral of a_t where a_t is an adapted and *càdlàg* process with paths of finite variation then

$$A_t = \int_0^t a_u du. \tag{3.45}$$

The semimartingales with these assumptions are in the class called continuous stochastic volatility semimartingales or continuous Brownian semimartingales.

Now, with the assumptions above if the logarithmic price process Y_t is in the class of continuous stochastic volatility semimartingales then Y_t can be decomposed as

$$Y_t = \int_0^t a_u du + \int_0^t \sigma_u dW(u) \tag{3.46}$$

where $A_t = \int_0^t a_u du$ is the drift term and $M_t = \int_0^t \sigma_u dW(u)$ is the stochastic term.

Theorem 3.3.1 *Let Y_t be the logarithmic price process which is in the class of continuous stochastic volatility semimartingales and can be decomposed into*

$$Y_t = \int_0^t a_u du + \int_0^t \sigma_u dW(u) \qquad \textit{(eqn. (3.46))}$$

where $A_t = \int_0^t a_u du$ (eqn. (3.45)) is the drift term and $M_t = \int_0^t \sigma_u dW(u)$ (eqn. (3.43)) is the stochastic volatility part. Then we have the quadratic variation of Y_t as

$$[Y]_t = \int_0^t \sigma_u^2 du = \vartheta^2. \tag{3.47}$$

Proof:First of all, $A_t = \int_0^t a_u du$ (eqn. (3.45)) is the finite variation process with continuous paths and $M_t = \int_0^t \sigma_u dW(u)$ (eqn. (3.43)) is a continuous martingale with $E(M_t)^2 < \infty$ so by Theorem (2.2.6), M_t is a local martingale with $E[M]_t < \infty$. By Theorem (3.2.3), since the assumptions on A_t and M_t in the theorem are the same for the class of continuous stochastic volatility semimartingales, we know that $[Y]_t = [M]_t$ (eqn. (3.23)).

Now, we will concentrate on $[M]_t$. We know that simple processes are dense in the class of adapted and square integrable martingales, i.e. $M_t \in \mathcal{F}_t$ and

$E(M_t)^2 < \infty$, so it is enough to show for simple processes.

Let us construct the stochastic integral for simple processes,

$$\int_0^t \sigma_u dW(u) = \sum_{i=1}^{k} \phi_i (W(t_i) - W(t_{i-1})) \qquad \text{(eqn. (2.13))}$$

is the stochastic integral of the simple process σ_t where

$$\sigma_t = \sum_{i=1}^{k} \phi_i 1_{(t_{i-1},t_i]} \qquad \text{(eqn. (2.12))}$$

where $0 = t_0 < t_1 < ... < t_k = t$ and ϕ_i is $\mathcal{F}(t_{i-1})$-measurable[3] and bounded.

Recall that, in the decomposition of Y_t we have

$$\begin{aligned} M_t = & \int_0^t \sigma_u dW(u) \quad (eqn.(3.43)) \\ = & \sum_{i=1}^{k} \phi_i (W(t_i) - W(t_{i-1})). \quad (eqn.(2.13)) \end{aligned}$$

Then,

$$E\{M_t^2\} = E\{(\phi_i(W(t_i) - W(t_{i-1}))^2)\} \tag{3.48}$$

and the square of this sum consists of the squares and a constant times cross products, i.e.

$$\begin{aligned} (\phi_i(W(t_i) - W(t_{i-1}))^2) = & \; C \sum_{i=1}^{k} \sum_{j=1}^{k} \phi_i \phi_j (W(t_i) - W(t_{i-1}))(W(t_j) - W(t_{j-1})) \\ + & \sum_{i=1}^{k} (\phi_i)^2 (W(t_i) - W(t_{i-1}))^2 \end{aligned}$$

where C is some constant and $i \neq j$.

Let us look at the sum of the cross products,

$$E(\sum_{i=1}^{k} \sum_{j=1}^{k} \phi_i \phi_j (W(t_i) - W(t_{i-1}))(W(t_j) - W(t_{j-1})))$$

[3] ϕ_i is predictable.

$$= {}^4E(E(\sum_{i=1}^{k}\sum_{j=1}^{k}\phi_i\phi_j(W(t_i)-W(t_{i-1}))(W(t_j)-W(t_{j-1}))|\mathcal{F}(t_{j-1})))$$

$$= E(\sum_{i=1}^{k}\sum_{j=1}^{k}\phi_i\phi_j E((W(t_i)-W(t_{i-1}))(W(t_j)-W(t_{j-1}))|\mathcal{F}(t_{j-1})))$$

$$=^5 E(\sum_{i=1}^{k}\sum_{j=1}^{k}\phi_i\phi_j E((W(t_j)-W(t_{j-1}))|\mathcal{F}(t_{j-1}))E((W(t_i)-W(t_{i-1}))|\mathcal{F}(t_{j-1}))).$$

And we have

$$E(W(t_i)-W(t_{i-1})|\mathcal{F}(t_{j-1})) =^6 E(W(t_i)-W(t_{i-1})) = 0$$

since W_t is a standard Brownian motion. Thus, we have that

$$E(\sum_{i=1}^{k}\sum_{j=1}^{k}\phi_i\phi_j(W(t_i)-W(t_{i-1}))(W(t_j)-W(t_{j-1}))) = 0.$$

Then, remains only the sum of the squares

$$\begin{aligned} E\{M_t^2\} = &\ E(\sum_{i=1}^{k}(\phi_i)^2(W(t_i)-W(t_{i-1}))^2) \quad (eqn.(3.48)) \\ = &\ \sum_{i=1}^{k}E((\phi_i)^2(W(t_i)-W(t_{i-1}))^2)) \\ = &\ \sum_{i=1}^{k}E((\phi_i)^2E((W(t_i)-W(t_{i-1}))^2)) \end{aligned}$$

since ϕ_i is $\mathcal{F}(t_{i-1})$-measurable and $(W(t_i)-W(t_{i-1}))$ is independent of $\mathcal{F}(t_{i-1})$.

Then we have the equality

$$\sum_{i=1}^{k}E((\phi_i)^2E((W(t_i)-W(t_{i-1}))^2)) =^7 \sum_{i=1}^{k}E((\phi_i)^2(t_i-t_{i-1}))$$

[4] Let $X \in (\Omega, \mathcal{F}, P)$, if $E(X|C)$ is defined then $E(E(X|C)) = E(X)$ where $\mathcal{C}$ is a sub σ-algebra of $\mathcal{F}$.

[5] $(W(t_j)-W(t_{j-1}))$ is independent of $(W(t_i)-W(t_{i-1}))$.

[6] $W(t_i)-W(t_{i-1})$ is independent of $\mathcal{F}(t_{j-1})$.

[7] $var((W(t_i)-W(t_{i-1}))) = E((W(t_i)-W(t_{i-1}))^2) = (t_i-t_{i-1})$.

and we know that

$$E(\sum_{i=1}^{k}(\phi_i)^2(t_i - t_{i-1})) = E(\int_0^t \sigma_u^2 du). \tag{3.49}$$

Thus, we get the equality,

$$E\{M_t^2\} = E(\int_0^t \sigma_u^2 du) \tag{3.50}$$

And from the theorem (2.2.6), $E\{M_t^2\} = E\{[M]_t\}$, if we combine these equalities, we get the result

$$E\{[M]_t\} = E\{M_t^2\} = E(\int_0^t \sigma_u^2 du)$$

Finally,

$$[M]_t = \int_0^t \sigma_u^2 du \tag{3.51}$$

and by the equations (3.51) and (3.23),

$$[Y]_t = \int_0^t \sigma_u^2 du \quad (\text{eqn.}(3.47))$$

which proves the theorem. Q.E.D.

Finally, the results of this chapter can be summarized as follows. The realized variance estimates the quadratic variation as $h \to 0$ or $n \to \infty$, i.e. the realized variance estimates the integrated variance. That is,

$$\begin{aligned} [Y]_t &= \lim_{n\to\infty} [Y_h]_t \quad (eqn.(3.6)) \\ &= \lim_{n\to\infty} \sum_{j=1}^{n} (Y(t_j) - Y(t_{j-1}))^2 \quad (eqn.(3.7)) \\ &= \int_0^t \sigma_u^2 du. \quad (eqn.(3.47)) \end{aligned}$$

CHAPTER 4

POWER VARIATION AND BIPOWER VARIATION PROCESSES

4.1 Power Variation Processes

The generalization of the quadratic variation process is the power variation process which was first stated by Barndorff-Nielsen and Shephard (2004) [26].

In this chapter, again we work for the interval from 0 to t and we assume that we record the prices for equally spaced intervals of length h, $0 = t_0 < t_1 < ... < t_{\lfloor \frac{t}{h} \rfloor} = t$, i.e. we have $\lfloor \frac{t}{h} \rfloor$ observations for the interval from 0 to t. Before the definition of power variation process, recall that j-th h-return is

$$y(j) = Y(jh) - Y((j-1)h), \quad j = 1, 2, ..., \lfloor \tfrac{t}{h} \rfloor \quad \text{(eqn.(3.4))}$$

where Y_t is the logarithmic price process at time t. And also, again assume that $\frac{t}{h}$ is an integer and $\lfloor \frac{t}{h} \rfloor = n$ to simplify the calculations.

In the last chapter, quadratic variation is defined for unequally spaced intervals but in this chapter power variation is defined for equally spaced intervals.

The r-th order power variation process is the probability limit of the realized power variation process as $h \to 0$. Let us first define the r-th order realized power variation.

Definition 4.1.1 *The r-th order realized power variation of Y_t is defined as*

$$\begin{aligned}\{Y_h\}_t^{[r]} = &\ h^{(1-\frac{r}{2})}\sum_{j=1}^{n}|y(t_j)|^r \\ = &\ h^{(1-\frac{r}{2})}\sum_{j=1}^{n}|Y(t_j)-Y(t_{j-1})|^r\end{aligned} \tag{4.1}$$

where $h^{(1-\frac{r}{2})}$ is the normalization in power variation and $0 = t_0 < t_1 < ... < t_n = t$.

Definition 4.1.2 *The r-th order power variation process is defined as the probability limit of the r-th order realized power variation process*

$$\begin{aligned}\{Y\}_t^{[r]} = &\ p - \lim_{n\to\infty}\{Y_h\}_t^{[r]} \\ = &\ p - \lim_{n\to\infty} h^{(1-\frac{r}{2})}\sum_{j=1}^{n}|y(t_j)|^r \\ = &\ p - \lim_{n\to\infty} h^{(1-\frac{r}{2})}\sum_{j=1}^{n}|Y(t_j)-Y(t_{j-1})|^r\end{aligned} \tag{4.2}$$

where $h^{(1-\frac{r}{2})}$ is the normalization in power variation and $0 = t_0 < t_1 < ... < t_n = t$.

Theorem 4.1.1 *Let the logarithmic price process Y_t be in the class of continuous stochastic volatility semimartingales which is defined in the last chapter. Then*

$$Y_t = \int_0^t a_u du + \int_0^t \sigma_u dW(u). \quad \textit{(eqn. (3.46))}$$

Now, additionally assume that $a_t = 0$ and σ_t is independent of W_t, then

$$\{Y\}_t^{[r]} = \mu_r \int_0^t \sigma_u^r du \tag{4.3}$$

where $\mu_r = E|u|^r$, $(r > 0)$ and $u \sim N(0,1)$.

Proof: In the proof, we consider only the martingale part, $M_t = \int_0^t \sigma_u dW(u)$, since we assume that $a_t = 0$. Now, it is enough to prove that

$$\{M\}_t^{[r]} = \mu_r \int_0^t \sigma_u^r du \tag{4.4}$$

where $\mu_r = E|u|^r$ $(r > 0)$ and $u \sim N(0,1)$.

In the proposition (2.1.2), choose $p = 1$ since $M_t \in L_1$. If we show that $E(|X_n - X|)$ goes to 0 as $n \to \infty$ then from the proposition (2.1.2), since

$$P\{|X_n - X| \geq \epsilon\} \leq \tfrac{1}{\epsilon^p} E(|X_n - X|^p) \qquad \text{(eqn. (2.1))},$$

we get that $P\{|X_n - X| \geq \epsilon\}$ goes to 0. Then $P\{|X_n - X| > \epsilon\}$ goes to 0 so, by proposition (2.1.3), we conclude that $p - \lim_{n\to\infty} X_n = X$.

So,it is enough to show that

$$\lim_{n\to\infty} E|\{M_h\}_t^{[r]} - \mu_r \int_0^t \sigma_u^r du| = 0. \tag{4.5}$$

In addition, it is enough to show this limit for simple processes since simple processes are dense in the class of adapted and square integrable semimartingales, i.e. $M_t \in \mathcal{F}_t$, $E\{M_t^2\} < \infty$ and $M_t = \int_0^t \sigma_u dW(u)$ satisfies these conditions.

Then, let us start with defining simple processes $\sigma_t = \sum_{i=1}^{k} \phi_i 1_{(t_{i-1},t_i]}$ where the interval $[0,t]$ is divided equally as $0 = t_0 < t_1 < ... < t_k = t$ and ϕ_i is $\mathcal{F}(t_{i-1})$ measurable and bounded. The next step is to construct the stochastic integral of the simple process with respect to the standard Brownian motion W_t which is defined as ,

$$M_t = \int_0^t \sigma_u dW(u)$$

$$= \sum_{i=1}^{k} \phi_i(W(t_i) - W(t_{i-1})). \quad (eqn.(2.13))$$

Now, we will write $\{M_h\}_{t_1}^{[r]}$ for the interval $(t_0, t_1]$ then we integrate the result for the interval [0,t] to get $\{M_h\}_t^{[r]}$. $\{M_h\}_{t_1}^{[r]}$ can be defined as

$$\{M_h\}_{t_1}^{[r]} = h^{(1-\frac{r}{2})} \sum_{j=1}^{\lfloor \frac{t_1}{h} \rfloor} |\phi_1(W(t_j) - W(t_{j-1}))|^r. \tag{4.6}$$

It is time to write the expectation for the interval $[0, t]$

$$\begin{aligned} E(\{M_h\}_t^{[r]} - \mu_r \int_0^t \sigma_u^r du) = & \ E(\{M_h\}_t^{[r]}) - E(\mu_r \int_0^t \sigma_u^r du) \\ = & \ E(\{M_h\}_t^{[r]}) - \mu_r \int_0^t \sigma_u^r du \end{aligned} \tag{4.7}$$

To find the expectation $E(\{M_h\}_t^{[r]})$, firstly we will find $E(\{M_h\}_{t_1}^{[r]})$ then take the sum from 0 to t. Now,

$$\begin{aligned} E(\{M_h\}_{t_1}^{[r]}) = & \ E(h^{(1-\frac{r}{2})} \sum_{j=1}^{\lfloor \frac{t_1}{h} \rfloor} |\phi_1(W(t_j) - W(t_{j-1}))|^r) \\ = & \ h^{(1-\frac{r}{2})} E(\sum_{j=1}^{\lfloor \frac{t_1}{h} \rfloor} |\phi_1(W(t_j) - W(t_{j-1}))|^r) \\ = & \ h^{(1-\frac{r}{2})} \sum_{j=1}^{\lfloor \frac{t_1}{h} \rfloor} E|\phi_1(W(t_j) - W(t_{j-1}))|^r \\ = & \ h^{(1-\frac{r}{2})} \sum_{j=1}^{\lfloor \frac{t_1}{h} \rfloor} \phi_1^r E|(W(t_j) - W(t_{j-1}))|^r \end{aligned}$$

Since W_t is a standard Brownian motion, $(W(t_j) - W(t_{j-1}))$ has standard deviation $\sqrt{t_j - t_{j-1}}$ then we can write the r-th power expectation as

$$E|(W(t_j) - W(t_{j-1}))|^r \stackrel{d}{=} E|u\sqrt{t_j - t_{j-1}}|^r \tag{4.8}$$

where $u \sim N(0,1)$. Then, we have the equality

$$\begin{aligned} h^{(1-\frac{r}{2})} \sum_{j=1}^{\lfloor \frac{t_1}{h} \rfloor} \phi_1^r E|(W(t_j) - W(t_{j-1}))|^r &= h^{(1-\frac{r}{2})} \sum_{j=1}^{\lfloor \frac{t_1}{h} \rfloor} \phi_1^r E|u\sqrt{t_j - t_{j-1}}|^r \\ &= h^{(1-\frac{r}{2})} \sum_{j=1}^{\lfloor \frac{t_1}{h} \rfloor} \phi_1^r (\sqrt{t_j - t_{j-1}})^r E|u|^r \\ &= h^{(1-\frac{r}{2})} \sum_{j=1}^{\lfloor \frac{t_1}{h} \rfloor} \phi_1^r |t_j - t_{j-1}|^{\frac{r}{2}} E|u|^r \end{aligned}$$

since $\sqrt{t_j - t_{j-1}}$ is positive, we can get out of the absolute value function and since it is deterministic, its expectation is equal to itself ,i.e. we can get out of the expectation function.

We have the assumption that we record the prices for equally spaced intervals of length h for the interval $[0, t]$ so $|t_j - t_{j-1}| = h$. Then the equation becomes

$$h^{(1-\frac{r}{2})} \sum_{j=1}^{\lfloor \frac{t_1}{h} \rfloor} \phi_1^r |t_j - t_{j-1}|^{\frac{r}{2}} E|u|^r = h^{(1-\frac{r}{2})} \sum_{j=1}^{\lfloor \frac{t_1}{h} \rfloor} \phi_1^r h^{\frac{r}{2}} E|u|^r$$

and we can take $h^{(1-\frac{r}{2})}$ in the sum since it does not depend on j. And we get the result,

$$\begin{aligned} h^{(1-\frac{r}{2})} \sum_{j=1}^{\lfloor \frac{t_1}{h} \rfloor} \phi_1^r |t_j - t_{j-1}|^{\frac{r}{2}} E|u|^r &= \sum_{j=1}^{\lfloor \frac{t_1}{h} \rfloor} \phi_1^r h^{(1-\frac{r}{2})} h^{\frac{r}{2}} E|u|^r \\ &= E|u|^r \sum_{j=1}^{\lfloor \frac{t_1}{h} \rfloor} \phi_1^r h \end{aligned}$$

and we denote $E|u|^r = \mu_r$ so the r-th order realized variance for the interval $(t_0, t_1]$ becomes

$$E\{M_h\}_{t_1}^{[r]} = \mu_r \sum_{j=1}^{\lfloor \frac{t_1}{h} \rfloor} \phi_1^r h \tag{4.9}$$

This expectation is for the first subinterval of the interval $[0, t]$. If we find all of the expectations for the subintervals and take the sum of them then we get the r-th order realized variance for the interval $[0, t]$. Thus,

$$E\{M_h\}_t^{[r]} = \mu_r \sum_{i=1}^{k} \phi_i^r h \tag{4.10}$$

where $E|u|^r = \mu_r$ and $u \sim N(0, 1)$.

Now, if we take the limit of the r-th order realized variance, we get

$$\begin{aligned} \lim_{n\to\infty} E(\{M_h\}_t^{[r]}) = & \lim_{n\to\infty} E(\mu_r \sum_{i=1}^{k} \phi_i^r h) \\ = & \mu_r \int_0^t \sigma_u^r du \end{aligned} \tag{4.11}$$

where $n \to \infty$ is equivalent to $h \to 0$. Thus,

$$\lim_{n\to\infty} E(\{M_h\}_t^{[r]} - \mu_r \int_0^t \sigma_u^r du) = 0. \tag{4.12}$$

Recall that for any process X_n if $X_n \to 0$ then $|X_n| \to 0$. So, we see that expectation of the absolute value of the difference is also zero, that is

$$\lim_{n\to\infty} E|\{M_h\}_t^{[r]} - \mu_r \int_0^t \sigma_u^r du| = 0. \quad \text{(eqn. (4.5))}$$

So,

$$p - \lim_{n\to\infty} \{M_h\}_t^{[r]} = \mu_r \int_0^t \sigma_u^r du.$$

From the definition of the r-th order power variation, we get

$$\begin{aligned} \{M\}_t^{[r]} = & \; p - \lim_{n\to\infty} \{M_h\}_t^{[r]} \\ = & \; \mu_r \int_0^t \sigma_u^r du \end{aligned} \tag{4.13}$$

In the theorem, we have the assumption that $a_t = 0$ so

$$\{Y\}_t^{[r]} = \{M\}_t^{[r]}. \tag{4.14}$$

Finally, when we combine the equations (4.13) and (4.14), we get the result

$$\{Y\}_t^{[r]} = \mu_r \int_0^t \sigma_u^r du. \quad \text{(eqn. (4.3))}$$

This proves the theorem. Q.E.D.

4.2 Bipower Variation Processes

In this section, we will work on the sum of the cross products of the different powers of returns instead of the powers of the returns. The general variation process is the bipower variation process since it contains the power variation process as well as the quadratic variation process. To understand this explicitly, let us give some definitions.

Definition 4.2.1 *The (r,s)-order bipower variation process is defined as*

$$\begin{aligned} \{Y\}_t^{[r,s]} &= p - \lim_{h \to 0} h^{(1-\frac{r+s}{2})} \sum_{j=2}^{n} |y(t_j)|^r |y(t_{j-1})|^s \\ &= p - \lim_{h \to 0} h^{(1-\frac{r+s}{2})} \sum_{j=2}^{n} |Y(t_j) - Y(t_{j-1})|^r |Y(t_{j-1}) - Y(t_{j-2})|^s \end{aligned} \tag{4.15}$$

where $r, s \geq 0$ *and* $j = 1, 2, ..., n$.

After the definition of bipower variation, it is convenient to remark that the special cases of the bipower variation process include the power variation process and also the quadratic variation process since quadratic variation is a special case of the power variation in which $r = 2$. Importantly,

$$\{Y\}_t^{[r,0]} = \{Y\}_t^{[0,r]} = \{Y\}_t^{[r]}.$$

In this section, it is again necessary to define the (r,s)-order realized bipower variation process to understand the (r,s)-order power variation process.

Definition 4.2.2 *The (r,s)-order realized bipower variation process is defined as*

$$\begin{aligned}
\{Y_h\}_t^{[r,s]} &= h^{(1-\frac{r+s}{2})}\sum_{j=2}^{n}|y(t_j)|^r|y(t_{j-1})|^s \\
&= h^{(1-\frac{r+s}{2})}\sum_{j=2}^{n}|Y(t_j)-Y(t_{j-1})|^r|Y(t_{j-1})-Y(t_{j-2})|^s
\end{aligned} \tag{4.16}$$

where $r,s \geq 0$ *and* $j = 1,2,...,n$.

Theorem 4.2.1 *Let the logarithmic price process* Y_t *be in the class of continuous stochastic volatility semimartingales that is defined as*

$$Y_t = \int_0^t a_u du + \int_0^t \sigma_u dW(u). \quad \textit{(eqn. (3.46))}$$

Additionally, assume that $a_t = 0$ *and* σ_t *is independent of* W_t*, then*

$$\{Y\}_t^{[r,s]} = \mu_r\mu_s\int_0^t \sigma_u^{r+s}du \tag{4.17}$$

where $r,s \geq 0$ *and* $\mu_r = E|u|^r$ *where* $u \sim N(0,1)$.

Proof:We consider only the martingale part M_t since we assume the expectation of the instantaneous returns is zero, i.e. $a_t = 0$.

Now, we will show the statement of the theorem for only the simple processes since simple processes are dense in the class of adapted and square integrable martingales, i.e. $M_t \in \mathcal{F}_t$, $E\{M_t^2\} < \infty$ and M_t satisfies these conditions.

Then, σ_t can be defined as the simple process $\sigma_t = \sum_{i=1}^{k}\phi_i 1_{(t_{i-1},t_i]}$ where $0 = t_0 < t_1 < ... < t_k = t$ and ϕ_i is $\mathcal{F}(t_{i-1})$ measurable and bounded. Now let us construct the stochastic integral of this simple process with respect to the standard Brownian motion W_t which is defined as,

$$\begin{aligned} M_t = & \int_0^t \sigma_u dW(u) \\ = & \sum_{i=1}^{k} \phi_i (W(t_i) - W(t_{i-1})). \quad (eqn.(2.13)) \end{aligned}$$

Recall that the limit in probability of the (r,s)-order realized bipower variation process is the (r,s)-order bipower variation process. So, let us start with defining the (r,s)-order realized bipower variation process

$$\{M_h\}_t^{[r,s]} = h^{(1-\frac{r+s}{2})} \sum_{j=2}^{n} |M(t_j) - M(t_{j-1})|^r |M(t_{j-1}) - M(t_{j-2})|^s \quad \text{(eqn. (4.16))}$$

where $r, s \geq 0$ and $j = 1, 2, ..., n$.

We will show

$$\begin{aligned} \{M\}_t^{[r,s]} = & \; p - \lim_{h \to 0} \{M_h\}_t^{[r,s]} \\ = & \; \mu_r \mu_s \int_0^t \sigma_u^{r+s} du \quad (eqn.(4.17)) \end{aligned}$$

where $r, s \geq 0$ and $\mu_r = E|u|^r$ where $u \sim N(0,1)$.

As mentioned, in the proof of the Theorem (4.1.1), it is enough to show

$$\lim_{h \to 0} E|\{M_h\}_t^{[r,s]} - \mu_r \mu_s \int_0^t \sigma_u^{r+s} du| = 0. \tag{4.18}$$

Now, let us determine the following expectation

$$\begin{aligned} E(\{M_h\}_t^{[r,s]} - \mu_r \mu_s \int_0^t \sigma_u^{r+s} du) = & \; E(\{M_h\}_t^{[r,s]}) - E(\mu_r \mu_s \int_0^t \sigma_u^{r+s} du) \\ = & \; E(\{M_h\}_t^{[r,s]}) - \mu_r \mu_s \int_0^t \sigma_u^{r+s} du \end{aligned}$$

since the last term $\mu_r \mu_s \int_0^t \sigma_u^{r+s} du$ is deterministic, we get it out of the expectation function.

Now, we will construct $E(\{M_h\}_{t_1}^{[r,s]})$ which is the (r,s)-order realized bipower variation for the interval $(t_0, t_1]$ and then take the integral for the interval $[0, t]$ to take the total sum and find $E\{M_h\}_t^{[r,s]}$.

$$\begin{aligned}
E\{M_h\}_{t_1}^{[r,s]} &= E(h^{(1-\frac{r+s}{2})} \sum_{j=2}^{\lfloor \frac{t_1}{h} \rfloor} |\phi_1(W(t_j) - W(t_{j-1}))|^r |\phi_1(W(t_{j-1}) - W(t_{j-2}))|^s) \\
&= h^{(1-\frac{r+s}{2})} \sum_{j=2}^{\lfloor \frac{t_1}{h} \rfloor} E(|\phi_1(W(t_j) - W(t_{j-1}))|^r |\phi_1(W(t_{j-1}) - W(t_{j-2}))|^s) \\
&= h^{(1-\frac{r+s}{2})} \sum_{j=2}^{\lfloor \frac{t_1}{h} \rfloor} E|\phi_1(W(t_j) - W(t_{j-1}))|^r E|\phi_1(W(t_{j-1}) - W(t_{j-2}))|^s
\end{aligned}$$

since $|\phi_1(W(t_j) - W(t_{j-1}))|^r$ and $|\phi_1(W(t_{j-1}) - W(t_{j-2}))|^s$ are independent. And also since W is standard Brownian motion, we have the equality

$$E|(W(t_j) - W(t_{j-1}))|^r \stackrel{d}{=} E|u\sqrt{t_j - t_{j-1}}|^r \qquad \text{(eqn. (4.8))}$$

where $u \sim N(0, 1)$. Thus,

$$\begin{aligned}
E\{M_h\}_{t_1}^{[r,s]} &= h^{(1-\frac{r+s}{2})} \sum_{j=2}^{\lfloor \frac{t_1}{h} \rfloor} \phi_1^r E|u\sqrt{t_j - t_{j-1}}|^r \phi_1^s E|u\sqrt{t_{j-1} - t_{j-2}}|^s \\
&= h^{(1-\frac{r+s}{2})} \sum_{j=2}^{\lfloor \frac{t_1}{h} \rfloor} \phi_1^{r+s} (\sqrt{t_j - t_{j-1}})^r E|u|^r (\sqrt{t_{j-1} - t_{j-2}})^s E|u|^s
\end{aligned}$$

since $\sqrt{t_j - t_{j-1}}$ is positive and also deterministic, we can take out of both the expectation and the absolute value function.

Recall that we have the assumption that we record the prices for equally spaced intervals of length h as $0 = t_1 < t_1 < ... < t_n = t$. Then the length of the interval $t_j - t_{j-1} = h$ for all $j = 1, 2, ..., n$. Thus, the equation becomes

$$\begin{aligned}
E\{M_h\}_{t_1}^{[r,s]} &= h^{(1-\frac{r+s}{2})} \sum_{j=2}^{\lfloor \frac{t_1}{h} \rfloor} \phi_1^{r+s} (\sqrt{h})^r E|u|^r (\sqrt{h})^s E|u|^s \\
&= h^{(1-\frac{r+s}{2})} \sum_{j=2}^{\lfloor \frac{t_1}{h} \rfloor} \phi_1^{r+s} h^{\frac{r+s}{2}} E|u|^r E|u|^s \\
&= \sum_{j=2}^{\lfloor \frac{t_1}{h} \rfloor} h^{(1-\frac{r+s}{2})} h^{\frac{r+s}{2}} \phi_1^{r+s} E|u|^r E|u|^s \\
&= E|u|^r E|u|^s \sum_{j=2}^{\lfloor \frac{t_1}{h} \rfloor} \phi_1^{r+s} h.
\end{aligned}$$

So we have the (r,s)-order realized bipower variation as

$$E\{M_h\}_{t_1}^{[r,s]} = \mu_r \mu_s \sum_{j=2}^{\lfloor \frac{t_1}{h} \rfloor} \phi_1^{r+s} h$$

where $\mu_r = E|u|^r$.

Up to now, we have find the (r,s)-order realized bipower variation for the first subinterval and if we take the sum of these realized power variations for all the subintervals then we get the realized power variation for the interval $[0, t]$. Thus,

$$E\{M_h\}_t^{[r,s]} = \mu_r \mu_s \sum_{i=1}^{k} \phi_i^{r+s} h. \tag{4.19}$$

Now let us take the limit as $h \to 0$,

$$\lim_{h \to 0} E\{M_h\}_t^{[r,s]} = \mu_r \mu_s \int_o^t \sigma_u^{r+s} du \tag{4.20}$$

So,

$$\lim_{h \to 0} (E\{M_h\}_t^{[r,s]} - \mu_r \mu_s \int_o^t \sigma_u^{r+s} du) = 0. \tag{4.21}$$

Recall that for any process X_n if $X_n \to 0$ then $|X_n| \to 0$. Then the limit of the absolute value of the difference is also zero, i.e.

$$\lim_{h\to 0} |E\{M_h\}_t^{[r,s]} - \mu_r\mu_s \int_o^t \sigma_u^{r+s} du| = 0. \quad \text{(eqn. (4.18))}$$

And this limit implies the limit in probability so we get the result.

$$\begin{aligned} \{M\}_t^{[r,s]} &= p - \lim_{h\to 0}\{M_h\}_t^{[r,s]} \\ &= \mu_r\mu_s \int_o^t \sigma_u^{r+s} du. \end{aligned} \tag{4.22}$$

where $u \sim N(0,1)$. And since in the theorem we assume that $a_t = 0$,

$$\{Y\}_t^{[r,s]} = \{M\}_t^{[r,s]}. \tag{4.23}$$

And the result comes

$$\{Y\}_t^{[r,s]} = \mu_r\mu_s \int_o^t \sigma_u^{r+s} du. \quad \text{(eqn. (4.17))}$$

This proves the theorem. Q.E.D.

Now, we will consider the special cases of bipower variation processes. The following case is the condition of the equality of r and s.

Corollary 4.2.2 *The (r,r)-order bipower variation process is defined as*

$$\begin{aligned} \{Y\}_t^{[r,r]} &= p - \lim_{h\to 0} h^{(1-r)} \sum_{j=2}^{n} |y(t_j)|^r |y(t_{j-1})|^r \\ &= p - \lim_{h\to 0} h^{(1-r)} \sum_{j=2}^{n} |Y(t_j) - Y(t_{j-1})|^r |Y(t_{j-1}) - Y(t_{j-2})|^r \end{aligned} \tag{4.24}$$

where $r \geq 0$ and $j = 1, 2, ..., n$. Assume that Y_t is in the class of continuous stochastic volatility semiartingales. Additionally, if $a_t = 0$ and σ_t is independent from W_t then

$$\{Y\}_t^{[r,r]} = \mu_r^2 \int_o^t \sigma_u^{2r} du. \tag{4.25}$$

where $\mu_r = E|u|^r$ where $u \sim N(0,1)$.

After this corollary, the next condition that we will discuss what happens if $r+s=2$, i.e. we choose $s=2-r$.

Corollary 4.2.3 *The (r,2-r)-order bipower variation process is defined as*

$$\begin{aligned}\{Y\}_t^{[r,s]} &= p-\lim_{h\to 0}\sum_{j=2}^{n}|y(t_j)|^r|y(t_{j-1})|^{2-r} \\ &= p-\lim_{h\to 0}\sum_{j=2}^{n}|Y(t_j)-Y(t_{j-1})|^r|Y(t_{j-1})-Y(t_{j-2})|^{2-r}\end{aligned} \tag{4.26}$$

where $\mu_r = E|u|^r$ where $u \sim N(0,1)$.Assume that Y_t is in the class of continuous stochastic volatility semiartingales. Additionally, if $a_t = 0$ and σ_t is independent from W_t then

$$\{Y\}_t^{[r,2-r]} = \mu_r\mu_{2-r}\int_o^t \sigma_u^2 du. \tag{4.27}$$

where $\mu_r = E|u|^r$ where $u \sim N(0,1)$.

After this corollary, it is convenient to say that the value of $r+s$ is crucial in the bipower variation process since it determines the normalization identity as well as the power of the integrated variance. As shown in the corollary, in the case of $r+s=2$, the normalization is equal to 1. So, in the calculation, the normalization identity disappears. And $r+s$ also determines the power of the integrated variance. In the last corollary, it is seen that the (r,2-r)-order realized bipower variation process estimates the integrated variance and so gives the same result as the realized variance.

It is time to give a special bipower variation process in which we choose $r=s=1$.

Definition 4.2.3 *The (1,1)-order bipower variation process is defined as*

$$\begin{aligned}
\{Y\}_t^{[1,1]} &= p - \lim_{h\to 0}\sum_{j=2}^{n}|y(t_j)||y(t_{j-1})| \\
&= p - \lim_{h\to 0}\sum_{j=2}^{n}|Y(t_j)-Y(t_{j-1})||Y(t_{j-1})-Y(t_{j-2})|
\end{aligned} \tag{4.28}$$

In this chapter, the most consequential difference is the assumption on the drift part. In the next theorem, we will show that this restriction is not needed in the (1,1)-order bipower variation process.

Theorem 4.2.4 *The logarithmic price process Y_t at time t is a semimartingale and can be decomposed into two adapted, càdlàg processes*

$$Y_t = A_t + M_t \qquad \textit{(eqn.(2.14))}$$

where A_t is a locally finite variation process and M_t is a local martingale where $A(0) = M(0) = 0$. In addition to these settings, if Y_t is continuous, then

$$\{Y\}_t^{[1,1]} = \{M\}_t^{[1,1]}. \tag{4.29}$$

Proof:The logarithmic price process Y_t is decomposable as

$$Y_t = A_t + M_t. \qquad \text{(eqn.(2.14))}$$

Let us write the definition of the (1,1)-order bipower variation process for Y_t

$$\{Y\}_t^{[1,1]} = p - \lim_{h\to 0}\sum_{j=2}^{n}|Y(t_j)-Y(t_{j-1})||Y(t_{j-1})-Y(t_{j-2})|. \quad \text{(eqn. (4.28))}$$

Consider the (1,1)-order realized bipower variation process

$$\begin{aligned}
\{Y_h\}_t^{[1,1]} &= \sum_{j=2}^{n} |Y(t_j) - Y(t_{j-1})||Y(t_{j-1}) - Y(t_{j-2})| \\
&= \sum_{j=2}^{n} |(A(t_j) + M(t_j)) - (A(t_{j-1}) + M(t_{j-1}))| \\
&\quad |(A(t_{j-1}) + M(t_{j-1})) - (A(t_{j-2}) + M(t_{j-2}))| \qquad (4.30) \\
&= \sum_{j=2}^{n} |(A(t_j) - A(t_{j-1})) + (M(t_j) - M(t_{j-1}))| \\
&\quad |(A(t_{j-1}) - A(t_{j-2})) + (M(t_{j-1}) - M(t_{j-2}))|.
\end{aligned}$$

To simplify the calculations, let us introduce some notations

$$a_j = (A(t_j) - A(t_{j-1})) \qquad (4.31)$$

and

$$b_j = (M(t_j) - M(t_{j-1})). \qquad (4.32)$$

Now, the equation (4.30) becomes

$$\begin{aligned}
\{Y_h\}_t^{[1,1]} &= \sum_{j=2}^{n} |(a_j) + (b_j)||(a_{j-1}) + (b_{j-1})| \\
&= \sum_{j=2}^{n} |(a_j + b_j)(a_{j-1} + b_{j-1})| \qquad (4.33) \\
&= \sum_{j=2}^{n} |a_j a_{j-1} + b_j a_{j-1} + a_j b_{j-1} + b_j b_{j-1}|.
\end{aligned}$$

Recall the triangular inequality

$$|a| - |b| \leq ||a| - |b|| \leq |a + b| \leq |a| + |b|. \qquad (4.34)$$

Let us apply triangular inequality to the equation (4.33)),

$$\sum_{j=2}^{n} |a_j a_{j-1}| - |b_j a_{j-1}| - |a_j b_{j-1}| - |b_j b_{j-1}|$$
$$\leq \sum_{j=2}^{n} |a_j a_{j-1} + b_j a_{j-1} + a_j b_{j-1} + b_j b_{j-1}|$$

$$\leq \sum_{j=2}^{n} |a_j a_{j-1}| + |b_j a_{j-1}| + |a_j b_{j-1}| + |b_j b_{j-1}| \tag{4.35}$$

Now,

$$\begin{aligned} p - \lim_{n\to\infty} \sum_{j=2}^{n} |b_j a_{j-1}| = &\ p - \lim_{n\to\infty} \sum_{j=2}^{n} |(M(t_j) - M(t_{j-1}))(A(t_{j-1}) - A(t_{j-2}))| \\ \leq &\ p - \lim_{n\to\infty} \max |M(t_j) - M(t_{j-1})| \sum_{j=2}^{n} |A(t_{j-1}) - A(t_{j-2})|. \end{aligned} \tag{4.36}$$

We have

$$p - \lim_{n\to\infty} \max |M(t_j) - M(t_{j-1})| = 0 \tag{4.37}$$

since M_t is continuous. And also, since A_t is finite variation process, $|A(t_{j-1}) - A(t_{j-2})|$ is finite. So, we conclude that

$$p - \lim_{n\to\infty} \max |M(t_j) - M(t_{j-1})| \sum_{j=2}^{n} |A(t_{j-1}) - A(t_{j-2})| = 0. \tag{4.38}$$

By the equations (4.38) and (4.36), we get

$$p - \lim_{n\to\infty} \sum_{j=2}^{n} |b_j a_{j-1}| = 0. \tag{4.39}$$

By the same way,

$$p - \lim_{n\to\infty} \sum_{j=2}^{n} |a_j b_{j-1}| = 0 \tag{4.40}$$

and

$$p - \lim_{n\to\infty} \sum_{j=2}^{n} |a_j a_{j-1}| = 0. \tag{4.41}$$

Let us take the limits of the sums in the equation (4.35),

$$p - \lim_{n\to\infty} \sum_{j=2}^{n} |a_j a_{j-1}| - |b_j a_{j-1}| - |a_j b_{j-1}| - |b_j b_{j-1}|$$

$$\leq p - \lim_{n\to\infty} \sum_{j=2}^{n} |a_j a_{j-1} + b_j a_{j-1} + a_j b_{j-1} + b_j b_{j-1}|$$

$$\leq p - \lim_{n\to\infty} \sum_{j=2}^{n} |a_j a_{j-1}| + |b_j a_{j-1}| + |a_j b_{j-1}| + |b_j b_{j-1}|.$$

By the equations (4.39), (4.40) and (4.41),

$$\begin{aligned} p - \lim_{n\to\infty} \sum_{j=2}^{n} |b_j b_{j-1}| \quad &\leq p - \lim_{n\to\infty} \sum_{j=2}^{n} |a_j a_{j-1} + b_j a_{j-1} + a_j b_{j-1} + b_j b_{j-1}| \\ &\leq \quad p - \lim_{n\to\infty} \sum_{j=2}^{n} |b_j b_{j-1}|. \end{aligned} \tag{4.42}$$

Since $\{M_h\}_t^{[1,1]} = \sum_{j=2}^{n} |b_j||b_{j-1}|$ and by the equations (4.33) and (4.42), we have

$$p - \lim_{n\to\infty} \{M_h\}_t^{[1,1]} \leq p - \lim_{n\to\infty} \{Y_h\}_t^{[1,1]} \leq p - \lim_{n\to\infty} \{M_h\}_t^{[1,1]}. \tag{4.43}$$

Equation (4.43) implies

$$\{Y\}_t^{[1,1]} = p - \lim_{n\to\infty} \{Y_h\}_t^{[1,1]} = p - \lim_{n\to\infty} \{M_h\}_t^{[1,1]} = \{M\}_t^{[1,1]}. \tag{4.44}$$

So, we get the result

$$\{Y\}_t^{[1,1]} = \{M\}_t^{[1,1]}. \quad \text{(eqn.(4.29))}$$

This proves the theorem.Q.E.D.

Corollary 4.2.5 *Assume that Y_t is in the class of continuous stochastic volatility semimartingales. Additionally,σ_t is independent from W_t then*

$$\{Y\}_t^{[1,1]} = \mu_1^2 \int_o^t \sigma_u^2 du. \tag{4.45}$$

where $\mu_1 = E|u| = \frac{\sqrt{2}}{\sqrt{\Pi}}$ where $u \sim N(0,1)$.

So, clearly the integrated variance $\int_o^t \sigma_u^2 du$ can be estimated by the (1,1)-order realized bipower variation process. That is

$$\int_o^t \sigma_u^2 du = \mu_1^{-2} p - \lim_{h \to 0} \{Y_h\}_t^{[1,1]} \tag{4.46}$$

where $\mu_1 = E|u| = \frac{\sqrt{2}}{\sqrt{\Pi}}$ where $u \sim N(0,1)$.

In the next chapter, we will discuss about the quadratic variation process and (1,1)-order bipower variation processes on discontinuous logarithmic price processes. Accordingly, it is very important to understand the relationship between the quadratic variation and (1,1)-order bipower variation processes for Y_t which is in the class of continuous stochastic volatility semimartingale processes.

In consequence, let us summarize the last two chapters. Recall that the logarithmic price process Y_t is semimartingale and can be decomposed into two adapted, *càdlàg* processes

$$Y_t = A_t + M_t \qquad \text{(eqn.(2.14))}$$

where A_t is a locally finite variation process and M_t is a local martingale where $A(0) = M(0) = 0$. Moreover, we have worked on special class called continuous stochastic volatility semimartingales. In this class, we have some additional assumptions on Y_t. Firstly, we assume M_t is continuous and it is an Ito integral of spot volatility process, $\sigma_t > 0$, with respect to a standard Brownian motion relative to the filtration $\mathcal{F}_t$ then

$$M_t = \int_0^t \sigma_u dW(u) \qquad \text{(eqn.(3.43))}$$

where the spot volatility process, $\sigma_t > 0$, is an adapted, c*àdlà*g and locally bounded away from zero and also W_t is the standard Brownian motion. In this setting, we can define the integrated variance as

$$\vartheta_t^2 = \int_0^t \sigma_u^2 du,$$

where $\vartheta_t^2 < \infty$ and this equality is defined for all $t < \infty$. Secondly, we assume A_t is continuous and it is the Riemann integral of a_t where a_t is an adapted and c*àdlà*g process with paths of finite variation then

$$A_t = \int_0^t a_u du. \qquad \text{(eqn.(3.45))}$$

The semimartingales with these assumptions are in the class called continuous stochastic volatility semimartingales or continuous Brownian semimartingales.

Furthermore, let us remember that we perform on the interval from 0 to t and assume that the prices are recorded for equally spaced intervals of length h and $\frac{t}{h}$ is an integer with $\frac{t}{h} = n$. Thus, we work on equidistant intervals such as $0 = t_0 < t_1 < ... < t_n = t$.

As mentioned in the third chapter, if Y_t is in the continuous stochastic volatility semimartingales then the integrated variance is equal to the quadratic variation and quadratic variation can be estimated by the realized variance. That is

$$\begin{aligned} [Y]_t = \; & p - \lim_{n\to\infty} [Y_h]_t \quad (eqn.(3.6)) \\ = \; & p - \lim_{n\to\infty} \sum_{j=1}^{n} (Y(t_j) - Y(t_{j-1}))^2 \quad (eqn.(3.3)) \\ = \; & \int_0^t \sigma_u^2 du. \quad (eqn.(3.47)) \end{aligned}$$

Besides, in this chapter, if Y_t is in the continuous stochastic volatility semimartingales then the integrated variance can be estimated by the (1,1)-order realized bipower variation process. That is

$$\begin{aligned}
\{Y\}_t^{[1,1]} &= p - \lim_{n\to\infty} \{Y_h\}_t^{[1,1]} \quad (eqn.(4.46)) \\
&= p - \lim_{n\to\infty} \sum_{j=2}^{n} |Y(t_j) - Y(t_{j-1})||Y(t_{j-1}) - Y(t_{j-2})| \quad (eqn.(4.28)) \\
&= \mu_1^2 \int_0^t \sigma_u^2 du. \quad (eqn.(4.45))
\end{aligned}$$

where $\mu_1 = E|u| = \frac{\sqrt{2}}{\sqrt{\Pi}}$ where $u \sim N(0,1)$. Thus, the integrated variance can be estimated as

$$\int_0^t \sigma_u^2 du = \mu_1^{-2} p - \lim_{n\to\infty} \{Y_h\}_t^{[1,1]}. \tag{4.47}$$

In conclusion, if we combine the results of the last two chapters, then we get access to the result that for the continuous stochastic volatility semimartingale logarithmic price processes, both the realized quadratic variation and the (1,1)-order realized bipower variation process are consistent estimators of the integrated variance

$$\begin{aligned}
\int_0^t \sigma_u^2 du = &\ \mu_1^{-2} p - \lim_{n\to\infty} \{Y_h\}_t \\
= &\ p - \lim_{n\to\infty} [Y_h]_t
\end{aligned} \tag{4.48}$$

where $\mu_1 = E|u| = \frac{\sqrt{2}}{\sqrt{\Pi}}$ where $u \sim N(0,1)$.

In the next chapter, we will see what happens to these estimators when we add jump part to the logarithmic price process Y_t.

CHAPTER 5

STOCHASTIC VOLATILITY SEMIMARTINGALE PLUS RARE JUMP PROCESSES

In the last chapter, we have worked on continuous logarithmic price processes and have ensured that both the realized variance and the (1,1)-order realized bipower variation are consistent estimators of the integrated variance.

In this chapter, we will see how this result changes when we work on stochastic volatility semimartingale plus rare jump logarithmic price processes.

Definition 5.0.4 *Let Y_t be the logarithmic price process which have both the continuous part and the discontinuous part*

$$Y_t = Y_t^{ct} + Y_t^d \qquad \textit{(eqn. (3.8))}$$

where Y_t^{ct} is in the class of stochastic volatility semimartingale processes as

$$Y_t^{ct} = \int_0^t a_u du + \int_0^t \sigma_u dW(u) \qquad \textit{(eqn. (3.46))}$$

and Y_t^d is the jump part which can be written as

$$Y_t^d = \sum_{i=1}^{N(t)} c_i \tag{5.1}$$

where N is the number of jumps and is finite activity simple counting process[1] *and*

[1] i.e.Recall that finite activity simple counting process means there exist finite jumps in finite interval of time, such that $N(t) < \infty$ for all $0 < t < \infty$.

c_i are non-zero random variables. Then, we say that Y_t is in the class of stochastic volatility semimartingale plus rare jump processes if it can be decomposed as

$$Y_t = \int_0^t a_u du + \int_0^t \sigma_u dW(u) + \sum_{i=1}^{N(t)} c_i. \tag{5.2}$$

5.1 Rare Jumps and Quadratic Variation

We will deal with the continuous price processes with rare jumps, i.e. finite jumps in finite interval of time. As mentioned in the third chapter, if Y_t has jumps then

$$[Y]_t = [Y^{ct}]_t + [Y^d]_t \qquad \text{(eqn. (3.8))}.$$

Theorem 5.1.1 *Let Y_t be in the class of stochastic volatility semimartingale plus rare jump processes, i.e.*

$$Y_t = \int_0^t a_u du + \int_0^t \sigma_u dW(u) + \sum_{i=1}^{N(t)} c_i \qquad \textit{(eqn. (5.2))}$$

where a_t is adapted, càdlàg process with paths of finite variation, σ_t is adapted, càdlàg and locally bounded away from zero, N is a finite activity simple counting processes for all t and c_i are non-zero random variables, then

$$[Y]_t = \int_0^t \sigma_u^2 du + \sum_{i=1}^{N(t)} c_i^2. \tag{5.3}$$

Proof: Let us start by writing the logarithmic price process which is in the class of the stochastic volatility semimartingale plus rare jump processes

$$Y_t = \int_0^t a_u du + \int_0^t \sigma_u dW(u) + \sum_{i=1}^{N(t)} c_i \qquad \text{(eqn. (5.2))}$$

where

$$Y_t^{ct} = \int_0^t a_u du + \int_0^t \sigma_u dW(u) \qquad \text{(eqn. (3.46))}$$

and

$$Y_t^d = \sum_{i=1}^{N(t)} c_i \qquad \text{(eqn. (5.2))}.$$

Let us recall the below equation,

$$[Y]_t = [Y_t^{ct}] + [Y_t^d] \qquad \text{(eqn. (3.11))}$$

from the Theorem (3.2.1) and also the quadratic variation of the continuous semimartingale process

$$[Y^{ct}]_t = \int_0^t \sigma_u^2 du \qquad \text{(eqn. (3.47))}$$

from the Theorem (3.3.1).

Therefore, we will maintain only the discontinuous part, i.e. jump part of Y_t

$$Y_t^d = \sum_{i=1}^{N(t)} c_i \qquad \text{(eqn. (5.1))}.$$

The quadratic variation of the discontinuous part is

$$[Y^d]_t = p - \lim_{n\to\infty} \sum_{j=1}^{n} \Big(\sum_{i=1}^{N(t_j)} - \sum_{i=1}^{N(t_{j-1})} \Big)^2$$

where $N(t_j)$ denotes the number of jumps for the interval from 0 up to t_j. And if we get the difference of these sums, we will have

$$p - \lim_{n\to\infty} \sum_{j=1}^{n} \Big(\sum_{i=1}^{N(t_j)} - \sum_{i=1}^{N(t_{j-1})} \Big)^2 = p - \lim_{n\to\infty} \sum_{j=1}^{n} \Big(\sum_{N(t_{j-1})}^{N(t_j)} \Big)^2.$$

This term is the sum of the squares of the jumps for the subintervals $(t_{j-1}, t_j]$ so the sum of these squares from 0 to n will give us the sum of the all jumps in the interval $[0, t]$. Thus, this sum will come to be

$$\begin{aligned} p - \lim_{n\to\infty} \sum_{j=1}^{n} (\sum_{N(t_{j-1})}^{N(t_j)})^2 = & \; p - \lim_{n\to\infty} \sum_{j=1}^{N(t)} c_i^2 \\ = & \; \sum_{j=1}^{N(t)} c_i^2. \end{aligned}$$

Consequently,

$$[Y^d]_t = \sum_{j=1}^{N(t)} c_i^2. \tag{5.4}$$

Now, let us put together the results and write the quadratic variation of Y_t if there exist jumps in Y_t,

$$\begin{aligned} [Y]_t = & \; [Y^{ct}]_t + [Y^d]_t \quad (eqn.(3.8)) \\ = & \; \int_0^t \sigma_u^2 du + \sum_{j=1}^{N(t)} c_i^2. \quad (eqns.(3.47), (5.4)) \end{aligned}$$

This proves the theorem. Q.E.D.

This theorem says that if the logarithmic price process Y_t has jump part then the quadratic variation of Y_t has also an additional sum that comes from the jump part.

In the next section, we will see whether the (1,1)-order bipower variation will change or not when we add a jump part to Y_t.

5.2 Rare Jumps and (1,1)-order Bipower Variation Process

As we come to the bipower variation process, we again look the case what happens if Y_t is in the class of stochastic volatility semimartingale plus jump processes, i.e.

$$Y_t = \int_0^t a_u du + \int_0^t \sigma_u dW(u) + \sum_{i=1}^{N(t)} c_i. \qquad \text{(eqn. (5.2))}$$

All the assumptions on a_t, σ_t, $N(t)$ and c_i are the same as in the last section.

Theorem 5.2.1 *Let the logarithmic price process Y_t be in the class of stochastic volatility semimartingale plus rare jump processes, i.e.*

$$Y_t = \int_0^t a_u du + \int_0^t \sigma_u dW(u) + \sum_{i=1}^{N(t)} c_i \qquad \textit{(eqn. (5.2))}$$

where

$$Y_t^{ct} = \int_0^t a_u du + \int_0^t \sigma_u dW(u) \qquad \textit{(eqn. (3.46))}$$

and

$$Y_t^d = \sum_{i=1}^{N(t)} c_i \qquad \textit{(eqn. (5.2))}.$$

Then,

$$\{Y^d\}_t^{[1,1]} = 0. \tag{5.5}$$

Proof: Now, let us write (1,1)-order bipower variation for the discontinuous part (jump part) of Y_t

$$\{Y^d\}_t^{[1,1]} = p - \lim_{h \to 0} \sum_{j=2}^{n} \left| \sum_{i=1}^{N(t_j)} c_i - \sum_{i=1}^{N(t_{j-1})} c_i \right| \left| \sum_{i=1}^{N((t_{j-1})} c_i - \sum_{i=1}^{N(t_{j-2})} c_i \right|$$

where $N(t_j)$ is the number of the jumps for the interval from 0 up to t_j.

The difference of the jumps is equal to the jump in the subinterval. What we want to say is that

$$\sum_{i=1}^{N(t_j)} c_i - \sum_{i=1}^{N(t_{j-1})} c_i = \sum_{i=N(t_{j-1})}^{N(t_j)} c_i.$$

Then the 1,1-order bipower variation comes to be

$$\{Y^d\}_t^{[1,1]} = p - \lim_{h \to 0} \sum_{j=2}^{n} |\sum_{i=N(t_{j-1})}^{N(t_j)} c_i||\sum_{i=N(t_{j-2})}^{N(t_{j-1})} c_i|.$$

As we know, we have the assumption that we record the prices for equally spaced intervals of length h where $0 = t_0 < t_1 < ... < t_n = t$ and $t_j - t_{j-1} = h$ for all $j = 1, 2, ..., n$. Then as $h \to 0$, we have $(t_j - t_{j-1}) \to 0$ for all $j = 1, 2, ..., n$. Thus, as $h \to 0$, both $(t_j - t_{j-1}) \to 0$ and $(t_{j-1} - t_{j-2}) \to 0$. In addition, we have the assumption that we have rare jumps, i.e. we have finite jumps in finite interval of time. Therefore, if Y_t has jump in the subinterval $(t_{j-1}, t_j]$ then Y_t can not have jump in the interval $(t_{j-2}, t_{j-1}]$ since as $h \to 0$, both the subintervals goes to zero. Otherwise, if we have jump in the interval $(t_{j-1}, t_j]$ as well as in the interval $(t_{j-2}, t_{j-1}]$ as $h \to 0$ then we have infinite jumps for the finite interval $[0, t]$ and this contradicts our assumption that we have rare jumps. Consequently, we have

$$\{Y^d\}_t^{[1,1]} = p - \lim_{h \to 0} \sum_{j=2}^{n} |\sum_{i=N(t_{j-1})}^{N(t_j)} c_i||\sum_{i=N(t_{j-2})}^{N(t_{j-1})} c_i| = 0$$

since either

$$p - \lim_{h \to 0} |\sum_{i=N(t_{j-1})}^{N(t_j)} c_i| = 0$$

or

$$p - \lim_{h \to 0} |\sum_{i=N(t_{j-2})}^{N(t_{j-1})} c_i| = 0$$

for all $j = 2, 3, ...n$. Consequently, the product goes to zero as $h \to 0$,

$$\{Y^d\}_t^{[1,1]} = 0. \quad \text{(eqn. (5.5))}$$

This proves the theorem. Q.E.D.

Theorem 5.2.2 *Let the logarithmic price process Y_t be in the class of stochastic volatility semimartingale plus rare jump processes, i.e.*

$$Y_t = \int_0^t a_u du + \int_0^t \sigma_u dW(u) + \sum_{i=1}^{N(t)} c_i \quad \textit{(eqn. (5.2))}$$

then,

$$\{Y\}_t^{[1,1]} = \mu_1^2 \int_o^t \sigma_u^2 du. \quad \textit{(eqn.(4.45))} \tag{5.6}$$

where $\mu_1 = E|u| = \frac{\sqrt{2}}{\sqrt{\Pi}}$ where $u \sim N(0,1)$.

Proof: Let us first write the definition of the (1,1)-order bipower variation process

$$\{Y\}_t^{[1,1]} = p - \lim_{h \to 0} \sum_{j=2}^{n} |Y(t_j) - Y(t_{j-1})||Y(t_{j-1}) - Y(t_{j-2})|. \quad \text{(eqn. (4.28))}$$

Recall that

$$Y_t = Y_t^{ct} + Y_t^d. \quad \text{(eqn. (3.8))}$$

Now, the equation (4.28) becomes

$$\begin{aligned} \{Y\}_t^{[1,1]} = \; & p - \lim_{h \to 0} \sum_{j=2}^{n} |(Y(t_j)^{ct} + Y(t_j)^d) - (Y(t_{j-1})^{ct} + Y(t_{j-1})^d)| \\ & |(Y(t_{j-1})^{ct} + Y(t_{j-1})^d) - (Y(t_{j-2})^{ct} + Y(t_{j-2})^d)| \\ = \; & p - \lim_{h \to 0} \sum_{j=2}^{n} |(Y(t_j)^{ct} - Y(t_{j-1})^{ct}) + (Y(t_j)^d - Y(t_{j-1})^d)| \\ & |(Y(t_{j-1})^{ct} - Y(t_{j-2})^{ct}) + (Y(t_{j-1})^d - Y(t_{j-2})^d)| \end{aligned} \tag{5.7}$$

To simplify the calculations, let us introduce some notations

$$a_j = Y(t_j)^{ct} - Y(t_{j-1})^{ct} \tag{5.8}$$

and

$$b_j = Y(t_j)^{d} - Y(t_{j-1})^{d}. \tag{5.9}$$

By these new notations, the sum in the equation (5.7) becomes

$$\begin{aligned}
\{Y\}_t^{[1,1]} &= p - \lim_{h\to 0}\sum_{j=2}^{n} |(a_j) + (b_j)||(a_{j-1}) + (b_{j-1})| \\
&= p - \lim_{h\to 0}\sum_{j=2}^{n} |(a_j + b_j)(a_{j-1} + b_{j-1})| \qquad (5.10) \\
&= p - \lim_{h\to 0}\sum_{j=2}^{n} |a_j a_{j-1} + b_j a_{j-1} + a_j b_{j-1} + b_j b_{j-1}|.
\end{aligned}$$

We will use the triangular inequality

$$|a| - |b| \leq ||a| + |b|| \leq |a| + |b|. \qquad \text{(eqn. (4.34))}$$

Consider the (1,1)-order realized bipower variation process from the equation (5.10)

$$\{Y_h\}_t^{[1,1]} = \sum_{j=2}^{n} |a_j a_{j-1} + b_j a_{j-1} + a_j b_{j-1} + b_j b_{j-1}|. \tag{5.11}$$

Let us apply the triangular equality in the equation (4.34) to the equation (5.11)

$$\begin{aligned}
&\sum_{j=2}^{n} |a_j a_{j-1}| - |b_j a_{j-1}| - |a_j b_{j-1}| - |b_j b_{j-1}| \\
&\leq \sum_{j=2}^{n} |a_j a_{j-1} + b_j a_{j-1} + a_j b_{j-1} + b_j b_{j-1}| \\
&\leq \sum_{j=2}^{n} |a_j a_{j-1}| + |b_j a_{j-1}| + |a_j b_{j-1}| + |b_j b_{j-1}|.
\end{aligned}$$

We know that

$$\{Y^d\}_t^{[1,1]} = 0 \qquad \text{(eqn. (5.5))}$$

and

$$\begin{aligned} \{Y^d\}_t^{[1,1]} &= p - \lim_{h\to 0} \sum_{j=2}^{n} |Y(t_j)^d - Y(t_{j-1})^d||Y(t_{j-1})^d - Y(t_{j-2})^d| \\ &= p - \lim_{h\to 0} \sum_{j=2}^{n} |b_j b_{j-1}|. \end{aligned} \tag{5.12}$$

So,

$$p - \lim_{n\to\infty} \sum_{j=2}^{n} |b_j b_{j-1}| = 0. \tag{5.13}$$

Consider the cross products

$$\begin{aligned} \sum_{j=2}^{n} |b_j a_{j-1}| &= \sum_{j=2}^{n} |Y(t_j)^d - Y(t_{j-1})^d||Y(t_{j-1})^{ct} - Y(t_{j-2})^{ct}| \\ &\le \max |Y(t_{j-1})^{ct} - Y(t_{j-2})^{ct}| \sum_{j=2}^{n} |Y(t_j)^d - Y(t_{j-1})^d|. \end{aligned} \tag{5.14}$$

Since $Y(t)^{ct}$ is continuous, we have

$$p - \lim_{n\to\infty} \max |Y(t_{j-1})^{ct} - Y(t_{j-2})^{ct}| = 0 \tag{5.15}$$

and $|Y(t_j)^d - Y(t_{j-1})^d|$ is finite, we have the limit as

$$p - \lim_{n\to\infty} \sum_{j=2}^{n} |b_j a_{j-1}| = 0. \tag{5.16}$$

And the same reasons imply that

$$p - \lim_{n\to\infty} \sum_{j=2}^{n} |a_j b_{j-1}| = 0. \tag{5.17}$$

Therefore, by the equations (5.13), (5.16) and (5.17),

$$p - \lim_{n\to\infty} \sum_{j=2}^{n} |a_j a_{j-1}| - |b_j a_{j-1}| - |a_j b_{j-1}| - |b_j b_{j-1}|$$

$$\leq p - \lim_{n\to\infty} \sum_{j=2}^{n} |a_j a_{j-1} + b_j a_{j-1} + a_j b_{j-1} + b_j b_{j-1}|$$

$$\leq p - \lim_{n\to\infty} \sum_{j=2}^{n} |a_j a_{j-1}| + |b_j a_{j-1}| + |a_j b_{j-1}| + |b_j b_{j-1}|$$

becomes

$$\begin{aligned} p - \lim_{n\to\infty} \sum_{j=2}^{n} |a_j a_{j-1}| \leq \; & p - \lim_{n\to\infty} \sum_{j=2}^{n} |a_j a_{j-1} + b_j a_{j-1} + a_j b_{j-1} + b_j b_{j-1}| \\ \leq \; & p - \lim_{n\to\infty} \sum_{j=2}^{n} |a_j a_{j-1}|. \end{aligned} \tag{5.18}$$

So, we have

$$p - \lim_{n\to\infty} \sum_{j=2}^{n} |a_j a_{j-1} + b_j a_{j-1} + a_j b_{j-1} + b_j b_{j-1}| = p - \lim_{n\to\infty} \sum_{j=2}^{n} |a_j a_{j-1}|. \tag{5.19}$$

From the equation (5.11) and (5.19), we know

$$\begin{aligned} \{Y\}_t^{[1,1]} &= p - \lim_{n\to\infty} \sum_{j=2}^{n} |a_j a_{j-1} + b_j a_{j-1} + a_j b_{j-1} + b_j b_{j-1}| \\ &= p - \lim_{n\to\infty} \sum_{j=2}^{n} |a_j a_{j-1}| \\ &= p - \lim_{n\to\infty} \sum_{j=2}^{n} |Y(t_j)^{ct} - Y(t_{j-1})^{ct}||Y(t_{j-1})^{ct} - Y(t_{j-2})^{ct}| \\ &= \{Y^{ct}\}_t^{[1,1]}. \end{aligned} \tag{5.20}$$

And we know from the equation (4.45) that

$$\{Y^{ct}\}_t^{[1,1]} = \mu_1^2 \int_o^t \sigma_u^2 du$$

where $\mu_1 = E|u| = \frac{\sqrt{2}}{\sqrt{\Pi}}$ where $u \sim N(0,1)$. Finally,

$$\{Y\}_t^{[1,1]} = \mu_1^2 \int_o^t \sigma_u^2 du$$

and this proves the theorem. Q.E.D.

In conclusion, if Y_t is in the class of the stochastic volatility semimartingale plus rare jump processes then the quadratic variation of Y_t includes both the integrated variance and the sum of the jump squares but the (1,1)-order bipower variation includes only the integrated variance. So, the difference of these processes gives us the quadratic variation of the jump part. Let us write what we have said in the following corollary.

Corollary 5.2.3 *Let Y_t be in the class of the stochastic volatility semimartingale plus rare jump processes then Y_t can be decomposed as*

$$Y_t = \int_0^t a_u du + \int_0^t \sigma_u dW(u) + \sum_{i=1}^{N(t)} c_i \qquad \textit{(eqn. (5.2))}$$

where the spot volatility process, $\sigma_t > 0$, is an adapted, càdlàg and locally bounded from zero and also W_t is the standard Brownian motion, a_t is an adapted and càdlàg process with paths of finite variation, a_t and σ_t is independent of W_t, N is a finite activity simple counting processes for all t and c_i are non-zero random variables. Then the quadratic variation process of Y_t comes to be

$$[Y]_t = \int_0^t \sigma_u^2 du + \sum_{i=1}^{N(t)} c_i^2 \qquad \textit{(eqn. (5.3))}$$

and the (1,1)-order bipower variation of Y_t comes to be

$$\{Y\}_t^{[1,1]} = \mu_1^2 \int_0^t \sigma_u^2 du \qquad \textit{(eqn. (4.45))}$$

where $\mu_1 = E|u| = \frac{\sqrt{2}}{\sqrt{\Pi}}$ where $u \sim N(0,1)$.

When we come to the problem how we will estimate the integrated variance and the sum that comes from the jump part of Y_t, it is suitable to recall the realized variance and the (1,1)-order realized bipower variation processes. Recall that

$$[Y]_t = p - \lim_{n\to\infty}[Y_h]_t \qquad \text{(eqn. (3.6))}$$

and

$$\{Y\}_t^{[1,1]} = p - \lim_{n\to\infty}\{Y_h\}_t^{[1,1]}. \qquad \text{(eqn. (4.46))}$$

Then the estimation is done by the realized processes as

$$\int_0^t \sigma_u^2 du + \sum_{i=1}^{N(t)} c_i^2 = p - \lim_{n\to\infty}[Y_h]_t \tag{5.21}$$

and

$$\int_0^t \sigma_u^2 du = \mu_1^{-2} p - \lim_{h\to 0}\{Y_h\}_t^{[1,1]} \tag{5.22}$$

where $\mu_1 = E|u| = \frac{\sqrt{2}}{\sqrt{\Pi}}$ where $u \sim N(0,1)$.

Thus, the estimation of the jump part can be performed by looking the difference of the realized variance and the 1,1-order realized bipower variation processes that can be showed as

$$\begin{aligned}\sum_{i=1}^{N(t)} c_i^2 = &\ p - \lim_{n\to\infty}[Y_h]_t - \mu_1^{-2} p - \lim_{h\to 0}\{Y_h\}_t^{[1,1]} \\ = &\ p - \lim_{n\to\infty}([Y_h]_t - \mu_1^{-2}\{Y_h\}_t^{[1,1]}).\end{aligned} \tag{5.23}$$

CHAPTER 6

APPLICATION

Up to now, we give the theoretical aspects of jump detection by using quadratic variation and bipower variation processes. In this chapter, we will discuss some applications on euro/FX price data set.

6.1 **Simple Application of Jump Detection**

In this section, we illustrate an application of what we have theoretically showed in the last chapters. In other words, we demonstrate how we can achieve the quadratic variation of the jump part by the estimation of the difference between realized variance and (1,1)-order bipower variation.

In this application, we have euro/FX cross rate price data for 143 days. The frequency of the data is 5 minutes. The graphs for the raw price data and the return data is in Figure 6.1.

Moreover, the data set contains 8581 observations. In daily trading period, to eliminate the zero returns we get only five hours of the day. So, for a trading day, we have 60 five minutes observations.

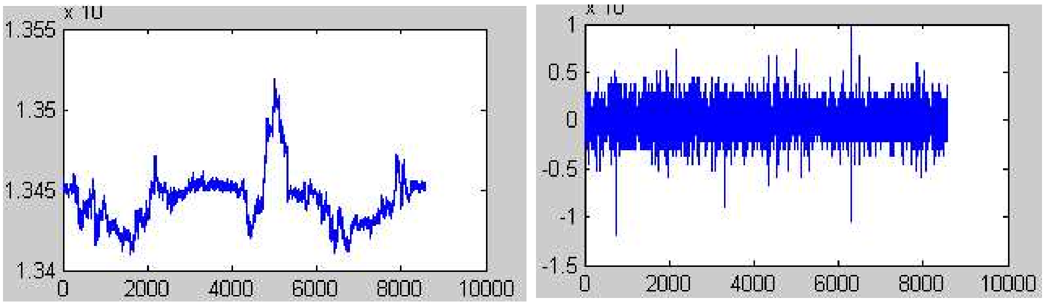

Figure 6.1: Euro/FX cross rate dataset. Price data and the return data for the dataset.

We compute the realized variance by the equation,

$$[Y_h]_t = \sum_{j=1}^{n}(y(j))^2, \quad (eqn.(3.5))$$

where the j-th h-return is shown as

$$y(j) = Y(jh) - Y((j-1)h), \qquad j = 1, 2, ..., n. \qquad \text{(eqn. (3.4))}$$

In this equation, according to our data, we have $n = 60$ and $h = 5minutes$. So, for our application, equation becomes

$$\begin{aligned} [Y_h]_t = & \sum_{j=1}^{60}(y(j))^2 \\ = & \sum_{j=1}^{60}(Y(t_j) - Y(t_{j-1}))^2. \quad (eqn.(3.7)) \end{aligned}$$

Recall from Chapter 5 that the realized variance estimates the integrated variance plus the quadratic variation of the jump part. It has been shown as

$$\begin{aligned} [Y]_t = & \; p - \lim_{n\to\infty}[Y_h]_t \\ = & \int_0^t \sigma_u^2 du + \sum_{i=1}^{N(t)} c_i^2. \quad (eqn.(5.2)) \end{aligned}$$

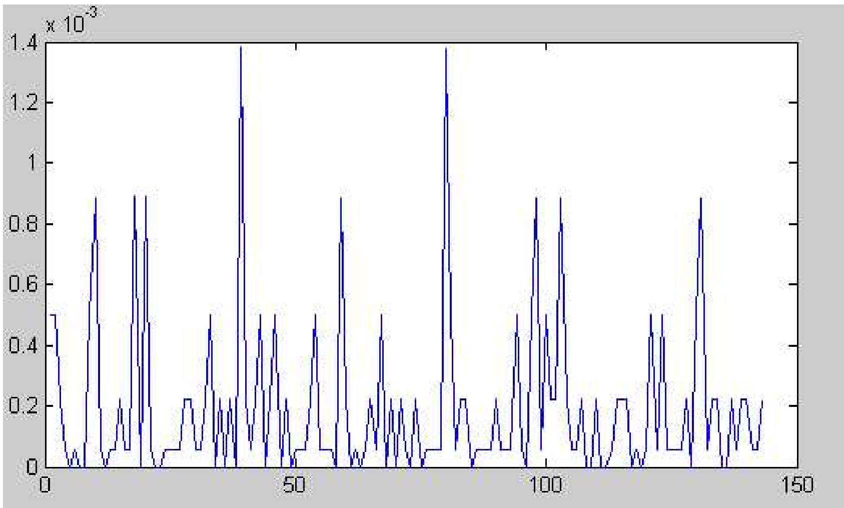

Figure 6.2: The Realized Variance for the data.

The graph for the realized variance is as in the Figure 6.2.

As illustrated in Chapter 4, (1,1)-order bipower variation is estimated by the realized (1,1)-order bipower variation. This can be shown as

$$\begin{aligned}\{Y_h\}_t^{[1,1]} &= \sum_{j=2}^{n} |y(j)||y(j-1)| \\ &= \sum_{j=2}^{n} |Y(t_j) - Y(t_{j-1})||Y(t_{j-1}) - Y(t_{j-2})|. \quad (eqn.(4.28))\end{aligned}$$

The key point of this process is that (1,1)-order bipower variation is robust to rare jumps. This means that this process is not affected from the jumps. Thus, even in the presence of jumps, (1,1)-order bipower variation is always equal to

$$\begin{aligned}\{Y\}_t^{[1,1]} &= p - \lim_{h \to 0} \{Y_h\}_t^{[1,1]} \\ &= \mu_1^2 \int_o^t \sigma_u^2 du \quad (eqn.(4.45))\end{aligned}$$

where $\mu_1 = E|u| = \frac{\sqrt{2}}{\sqrt{\Pi}} \simeq 0.79788$ where $u \sim N(0,1)$. Moreover, the estimation of the integrated variance can be done by the realized (1,1)-order bipower variation by

$$\mu_1^{-2}\{Y\}_t^{[1,1]} = p - \lim_{h\to 0} \mu_1^{-2}\{Y_h\}_t^{[1,1]}$$

$$= \int_o^t \sigma_u^2 du \quad (eqn.(5.22))$$

where $\mu_1 = E|u| =\simeq 0.79788$ where $u \sim N(0,1)$.

The graph for $\mu_1^{-2}\{Y_h\}_t^{[1,1]}$ is as in the Figure 6.3.

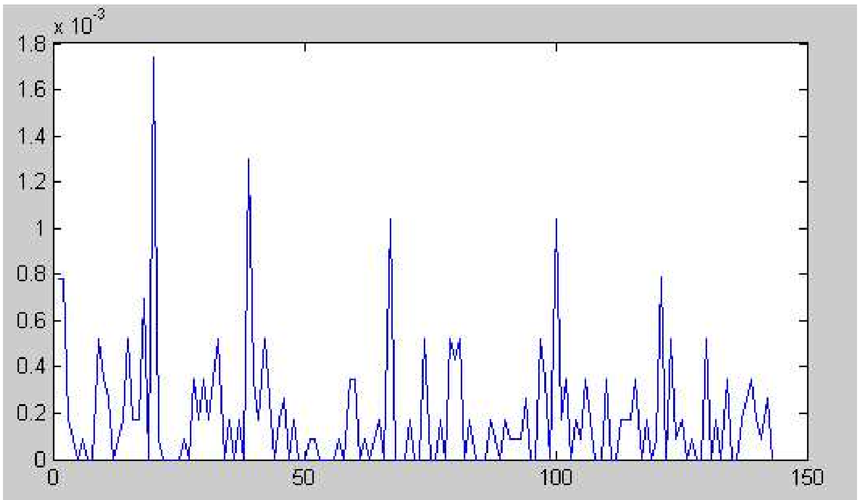

Figure 6.3: μ_1^{-2} times Realized (1,1)-order Bipower Variation for the data.

Thus, the difference between the realized variance and μ_1^{-2} times realized (1,1)-order bipower variation consistently estimates the quadratic variation of the jump process. This can be shown as

$$[Y]_t - \mu_1^{-2}\{Y\}_t^{[1,1]} = p - \lim_{h\to 0}([Y_h]_t - \mu_1^{-2}\{Y_h\}_t^{[1,1]})$$

$$= (\int_0^t \sigma_u^2 du + \sum_{i=1}^{N(t)} c_i^2) - (\int_0^t \sigma_u^2 du)$$

$$= \sum_{i=1}^{N(t)} c_i^2 \quad (eqn.(5.23))$$

where $\mu_1 = E|u| =\simeq 0.79788$ where $u \sim N(0,1)$.

Thus, it is clear that we can estimate the quadratic variation of the jump part by the difference of these two processes.

At this point, we should notice that this difference might be negative so estimation of the squared jump process for a finite sample might be negative. Therefore, it is suitable to truncate this difference at zero(2006) [33], i.e.

$$\sum_{i=1}^{N(t)} c_i^2 = p - \lim_{h \to 0} \max[[Y_h]_t - \mu_1^{-2}\{Y_h\}_t^{[1,1]}, 0].$$

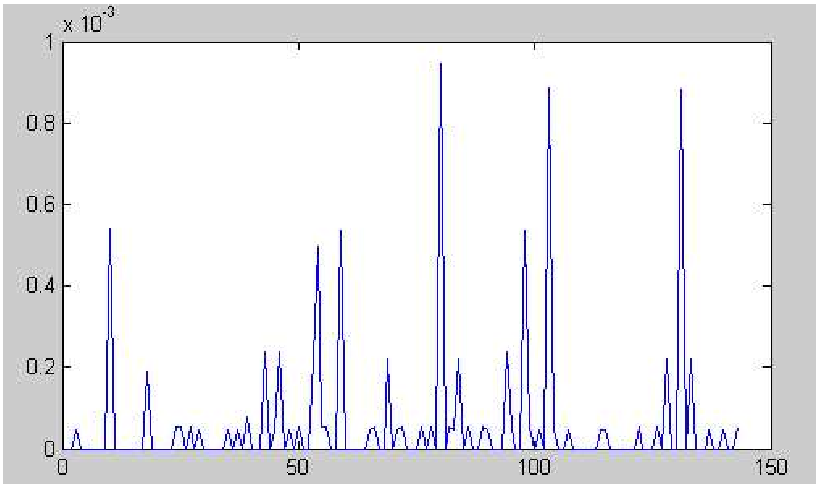

Figure 6.4: The Quadratic Variation of the jump process for the data.

The Figure 6.4 shows the quadratic variation of the jump part for the logarithmic price process for each day.

To see the jumps explicitly, we can plot the graph by dots. The figure 6.5 shows the quadratic variation of the jump part as dots.

Finally, for a finite sample data we have detected the jumps for each day using 5 minutes frequency data.

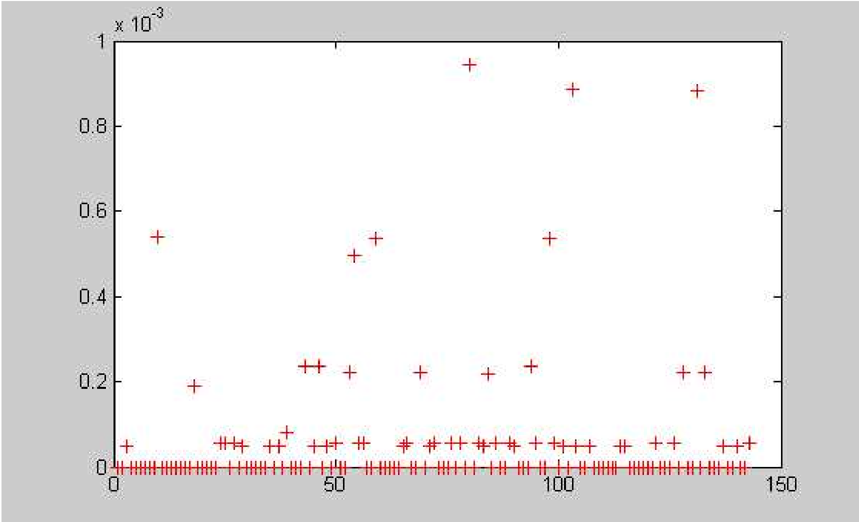

Figure 6.5: The Quadratic Variation of the jump process for the data.

From the economics aspects, by detecting jump part, we detect the risk that comes from the jump part. By this method, we can predict the links between the political or economical news and the jumps in the prices and this introduces new frameworks for the risk management.

CHAPTER 7

CONCLUSION

Jump detection by the realized variance and the (1,1)-order realized bipower variation process for the logarithmic price processes which are in the class of stochastic volatility semimartingale plus rare jump processes is the main objective of this study. The generalization of realized variance which is r-th order power variation is explained and we work on the bipower variation processes which are also the general version of power variations. The key property of the bipower variation process is the robustness to the jump part of the price process. Moreover, (1,1)-order bipower variation process estimates the integrated variance like realized variance. Robustness is the crucial point since when we add rare jumps to the price process which is in the class of stochastic volatility semimartingale, the probability limit of the bipower variation does not change. Thus, the difference of the realized variance and (1,1)order bipower variation estimates the quadratic variation of the jump part, i.e. jump detection is achieved.

By this method, the risk that comes from the discontinuous part of the logarithmic price processes is estimated. Jump detection have also some economic results. The link between some economical or political news and the size of the jumps can be predicted. In addition, the ability of separating the continuous and discontinuous components of the quadratic variation can be used to develop new volatility forecasting methods.

REFERENCES

[1] S. Figlewski. Forecasting volatility. *Financial Markets, Institutions and Instruments*, 6(1), 1997.

[2] L. H. Ederington and W. Guan. Forecasting volatility. *The Journal of Futures Markets*, 25(5):465–490, 2005.

[3] A. Gavolo N. Gospodinov and D. Jiang. Forecasting volatility. *Journal of Forecasting*, 25:381–400, 2006.

[4] S. Poon and C. W. J. Granger. Forecasting volatility in financial markets: A review. *Journal of Economic Literature*, 41:478–539, 2003.

[5] R. F. Engle and G. M. Gallo. A multiple indicators model for volatility using intra-daily data. *Journal of Applied Econometrics*, 131:3–27, 2006.

[6] F. Corsi G. Zumbach and A. Trapletti. Efficient estimation of volatility using high frequency data. *Journal of Applied Econometrics*, 2002. Research Institute for Applied Econometrics.

[7] J. Noh and T. Kim. Forecasting volatility of futures market: the sp 500 and ftse 100 futures using high frequency returns and implied volatility. *Applied Econometrics*, 38:395–413, 2006.

[8] J. R. Russell R. F. Engle. Analysis of high frequency financial data. 2004. Working Papers, New York University and University of California, University of Chicago.

[9] R. F. Engle and J. R. Russell. Autoregressive conditional duration: A new model for irregularly spaced transaction data. *Econometrica*, 66(5):1127–1162, 1998.

[10] H. Luckock F. Mainardi M. Mantelli E. Scalas, R. Gorenflo and M. Raberto. Anomalous waiting times in high-frequency financial data. *Quantitative Finance*, 4:695–702, 2004.

[11] M. M. Dacorogna, R. Gençay, U. A. Müller, R. B. Olsen, and O. V. Pictet. *An Introduction to High-frequency Finance*. Springer Finance Series.

[12] R. C. Merton. On estimating the expected return on the market: An exploratory investigation. *Journal of Financial Economics*, 8:323–361, 1980.

[13] T. G. Andersen and T. Bollerslev. Answering the skeptics: Yes, standard volatility models do provide accurate forecasts. *International Economic Review*, 39:885–905, 1998.

[14] O. E. Barndorff-Nielsen and N. Shephard. Estimating quadratic variation using realized variance. *Journal of Applied Econometrics*, 17:457–477, 2002.

[15] O. E. Barndorff-Nielsen and N. Shephard. Econometric analysis of realized volatility and its use in estimating stochastic volatility models. *Journal of the Royal Statistical Society*, 64(2):253–280, 2002.

[16] T. G. Andersen, T. Bollerslev, F. X. Diebold, and P. Labys. Modelling and forecasting realized volatility. *Econometrica*, 71(2):579–625, 2003.

[17] T. G. Andersen, T. Bollerslev, and N. Meddahi. Correcting the errors: Volatility forecast evaluation using high-frequency data and realized volatilities. *Econometrica*, 73(1):279–296, 2005.

[18] F. M. Bandi and J. R. Russell. Microstructure noise, realized volatility, and optimal sampling. 2003. CIRANO conference "realized volatility".

[19] J. M. Maheu. Modelling persistent time series data with application to realized volatility. 2004. working papers.

[20] T. G. Andersen and T. Bollerslev. Realized return volatility, asset pricing, and risk management. 2006. NBER reporter.

[21] O. E. Barndorff-Nielsen, N. Shephard, S. E. Graversen, J. Jacod, and M. Podolskij. A central limit theorem for realized power and bipower variations of continuous semimartingales.

[22] K. Back. Asset pricing for general processes. *Journal of Mathematical Economics*, 20:371–395, 1991.

[23] P. E. Protter. *Stochastic Integration and Differential Equations.* Applications of Mathematics Series. Second edition.

[24] O. E. Barndorff-Nielsen and N. Shephard. Variation, jumps, market frictions and high frequency data in financial econometrics. 2005. Economics papers: 2005-W16, University of Oxford, Nuffield College.

[25] O. E. Barndorff-Nielsen, N. Shephard, and S. E. Graversen. Power variation and stochastic volatility: a review and some new results. 2003. Economics papers: 2003-W19, University of Oxford, Nuffield College.

[26] O. E. Barndorff-Nielsen and N. Shephard. Power and bipower variation with stochastic volatility and jumps (with discussion). *Journal of Financial Econometrics*, 2:1–48, 2004.

[27] O. E. Barndorff-Nielsen and N. Shephard. Econometrics of testing for jumps in financial economics using bipower variation. 2003. Economics papers: 2003-W2, University of Oxford, Nuffield College.

[28] H. Körezlioğlu and A. Hayfavi. *Elements of Probability Theory.*

[29] S. E. Shreve. *Stochastic Calculus for Finance Continuous-Time Models.* Springer Finance Series.

[30] D. Lamberton and B. Lapeyre. *Introduction to Stochastic Calculus Applied to Finance.* First edition.

[31] R. Durrett. *Stochastic Calculus: A Practical Introduction.* Probability and Stochastic Series. Second edition.

[32] D. L. Snyder. *Random Point Processes.*

[33] M. Beine, J. Lahaye, S. Laurent, C. J. Neely, and F. C. Palm. Central bank intervention and exchange rate volatility, its continuous and jump components. *Federal Reserve Bank of St. Louis Working Paper Series 2006-031C*, 2006.