# Identification of currency jumps in the Rand to US Dollar exchange rate

**JS Labuschagne**

**orcid.org/0000-0003-3363-0634**

Dissertation accepted in partial fulfilment of the requirements for the degree *Master of Science in Applied Mathematics* at the North-West University

Supervisor: Dr GL Grobler

Co-supervisor: Prof IM Schoeman

Graduation May 2021

25069004

# Acknowledgments

This dissertation would not have been possible without the help and encouragement from many individuals.

To my supervisors Dr Gerrit Grobler and Prof Ilse Schoeman, I am truly grateful for the wisdom and guidance.

To my parents I am grateful for the unwavering support.

Finally I am grateful to God for His unending providence.

# Abstract

Modeling most financial assets, including exchange rates, proves to be a rather difficult task. One cause of this difficulty is the occasional unpredictable and large price movements that occur, known as jumps. Thus, including a jump component into a model may improve the accuracy of a model for financial assets. In order to do this jumps need to be identified. The question is whether this is possible or not?

Different tests have been developed to identify jumps, many of which are based upon the theory of bipower variation.

In this dissertation we will be studying the concepts of quadratic variation and bipower variation in order to study various tests based on bipower variation. We will additionally be applying one of the tests to identify the jumps in the Rand to United States Dollar exchange rate.

Having found the dates of jumps we assess potential causes and find that most jumps in the exchange rate are caused by political instability, most of which occurs in South Africa.


**Keywords:** Rand; US Dollar; currency jumps; bipower variation; quadratic variation; jump detection; jump diffusion model.

# Opsomming

Modellering van finansiële bates, insluitend wisselkoerse, is ’n taamlike moeilike taak. Een rede daarvoor is die onvoorspelbare en groot bewegings in die pryse wat plaasvind, ook bekend as spronge. Dus, kan die akkuraatheid van ’n model verbeter word wanneer, deur ’n sprongkomponent in ’n model ingesluit word. Om die taak uit te rig moet spronge geïdentifeer word. Die vraag is of dit moontlik is of nie?

Verskillende toetse is ontwikkel om spronge te identifiseer, waarvan ’n klomp gebaseer is op die teorie van dubbelmagvariasie.

In hierdie verhandeling bestudeer ons die konsepte van kwadratiese variasie en dubbelmagvariasie om sodoende verskeie toetse gebaseer op dubbelmagvariasie te bestudeer. Ons pas ook een van die toetse toe om spronge in die Rand tot Amerikaanse Dollar wisselkoers te identifiseer.

Deur die datums van spronge te identifiseer, word die moontlike oorsake van die spronge beoordeel. Ons bevind die dat meeste spronge as gevolg van politieke onstabiliteit is, waarvan die meeste van hierdie gebeurtenisse in Suid Afrika plaasgevind het.

**Sleutelwoorde:** Rand; Amerikaanse Dollar; wisselkoerssponge; dubbelmagvariasie;
kwadraatiese variasie; sprongopsporing; sprong-diffusie-model.

# Contents

# List of Figures

# List of Tables

CHAPTER 1

# Introduction

In this thesis we will be looking at testing for jumps. Specifically the testing for jumps in high-frequency data. A basic knowledge of stochastic processes may be required.

In recent years there has been an increase in the availability of data and data processing and as such the field of identifying jumps in high-frequency data has been increasing in popularity with many studies using intraday data, often at 5-minute intervals.

Baum and Zerilli (2016) modelled crude oil future prices (using high-frequency intraday data) with a stochastic volatility model, which included jumps of the prices. More specifically, the authors used bipower variation (a non-parametric method) to model the futures prices using a stochastic model with jumps by identifying the periods when these extreme events might have occurred. Furthermore, they identified extreme events, or jumps, by employing the standard Tukey formula to identify outliers. They also showed that accurately modelling jumps in a return series, leads to accuracy of volatility forecasts.

Laurent, Lecourt and Palm (2014) proposed a new semi-parametric test for additive jumps (similar to Lee and Mykland (2008)) in conditionally Gaussian ARMA-GARCH models with an additive jump component. Laurent, Lecourt and Palm (2014) applied their test to daily returns and found that less than 1% of movements in the data were jumps for the three exchange rates and between 1% and 3% of movements were jumps for all the stocks considered in the application (about 50 large capitalization stocks from the New York Stock Exchange or NYSE). The authors found that if they filtered their series (i.e. removed identified jumps) then they failed to reject the null hypothesis of conditional normality. Therefore, jumps may be the cause of deviation from normality after taking the effect of volatility clustering into account.

Patton and Sheppard (2015) assessed the effects of including the positive or negative signs of the high-frequency returns of stock prices (Standard and Poor 500 or S&P 500 index of 105 individual stocks) in forecasting volatility. The authors affirm that negative realized semivariance is much more important than positive semivariance for future volatility and extracting the effects of each of these leads to better forecasts, and also found that positive jumps lead to lower volatility while negative jumps lead to higher volatility.

When considering the subject of the identification of jumps and the effects macroeconomic news has on these jumps in financial data we found that there have been numerous studies over many years. Bajgrowicz and Scaillet (2011) proposed a new technique to identify spurious detection of jumps (using explicit thresholding on available test statistics). By application of this new technique, it is found that most of the jumps detected in the 30 Dow Jones stocks studied are spurious detections. Having removed the erroneously detected jumps Bajgrowicz and Scaillet studied the impact of news on jumps and found that a large portion of jumps occurs without any clear connection to news stories. However, Chatrath et al (2014) studied the impact of macroeconomic news on jumps in exchange rates. They looked at the exchange rate of the United States Dollar (USD) to four currencies and found that a disproportionate amount of jumps are associated with United States (US) news releases. Hence, jumps are a good proxy for news arrival in currency markets; there is a systematic reaction of currency prices to economic surprises and prices respond within 5 minutes of news releases. Fičura and Witzany (2015) estimated the stochastic volatility and jumps (in the Euro to USD exchange) using two different estimation methods, a non-parametric (power-variation) approach using 15-minute returns and Bayesian methods using daily returns. The results regarding the continuous stochastic volatility were very similar for the two methods. These two approaches, however, gave significantly different results for the jump estimation, the non-parametric approach resulted in significantly more jumps than the Bayesian method and the jumps identified with each method was on different days.

Now briefly reviewing the exchange rate models we see that Balaban (2004) in comparing the forecast performance of various models concluded that the exponential generalized autoregressive conditional heteroscedastic (EGARCH) model performs the best of the considered models. Ding and Vo (2012) who studied the volatility interactions between oil and exchange rate markets found that multivariate stochastic volatility models perform better than GARCH models when forecasting exchange rate volatility.

According to Hanousek et al (2011) researchers disagree about what causes jumps. Dewachter et al (2014) examined the impact of verbal announcements on the exchange rate markets, they made the claim that news releases significantly trigger large jumps or extreme events for up to an hour after the initial news release.

From the above studies, it follows that it is possible to identify jumps in economic data: Fičura and Witzany (2015) compared different methods of testing, Laurent, Lecourt and Palm (2014) recommended a new test for conditionally Gaussian ARMA-GARCH models and Bajgrowicz and Scaillet (2011) suggested a method of testing for spurious detections. The tests above, however, were based on bipower variation proposed by Barndorf-Nielsen and Shephard (2004) as well as the test by Lee and Mykland (2008) thus it becomes clear when studying the literature that the work by these authors is important in the field.

Barndorff-Nielsen and Shephard (2004) introduced a new method of assessing data in the form of the realised bipower variation and found that this new estimate can be used to estimate the instantaneous volatility of a process. This estimate is robust in the presence of jumps and thus can be used together with realised power variation to separate the quadratic variation into continuous and discontinuous, that is jump, components. Bearing in mind the importance of this work we will define these concepts in more detail later. Lee and Mykland (2008) then based on the results above of Barndorff-Nielsen and Shephard (2004) proposed a nonparametric test to detect jump times at the intraday level. So as previously stated since these two studies were important to the field they will be given special attention in our study.

Detection of jumps in financial data has some important applications. Some reasons pointed out by Lee and Mykland (2008) include that firstly when jumps are detected it helps us to assess what events may have lead to this jump occurring and thus testing for jumps may help in understanding financial market movements. Chatrath et al (2014) assessed jumps to gain an understanding of how news affects financial variables. Secondly testing for jumps helps in derivative hedging. Baum and Zerilli (2016) found that jumps in return series add to the accuracy of predicting the volatility.

This dissertation aims to:

(1) Assess some important tests to identify jumps and in practical applications we will be using daily data to apply one of these tests.
(2) Study how news affects jumps in the foreign exchange markets, specifically for the Rand-Dollar exchange rate. Since South Africa is a developing country rather than a developed country we believe this may produce new results.
(3) Discuss the work of Barndorff-Nielsen and Shephard (2004) and Lee and Mykland (2008) in full detail. This work forms the backbone of the detection of spurious jumps in this dissertation.
(4) Present a monograph that is easily accessible to students in the relevant research field. Hence, as much detail as required will be given in the discussions of proofs.

In the next chapter a literature study into tests for jumps based on power and bipower variation will be done. In Chapter 3 one of these tests will be applied to the Rand-Dollar exchange rate to identify news events that lead to the jumps identified. In Chapter 4 we conclude.

CHAPTER 2

# Literature Study

Let $(\Omega, \mathcal{F}, \{\mathcal{F}_{t\geq 0}\}, \mathcal{P})$ be a filtered probability space, where $\mathcal{F}_t : t \in [0,T]$ is a right continuous filtration on its $\sigma-$algebra $\mathcal{F}$. We define a stochastic process $S_t$ on this probability space as well as the process $X_t = \log S_t$, where $S_t$ is the asset price at time $t$ or in our application $S_t$ will be the Rand-Dollar exchange rate at time $t$. In this study we assume $X_t$ to be represented by a jump diffusion process given by

$$X_t = \int_0^t a(s)ds + \int_0^t \sigma(s)dW_s + \sum_{i=1}^{N_t} Z_i, \tag{2.1}$$

where $W_s$ is an $\mathcal{F}_s-$adapted standard Brownian motion, $a(s)$ and $\sigma(s)$ are $\mathcal{F}_t-$measurable functions, $N_t$ is a Poisson process with jump intensity $\lambda(t)$ and $(Z_i)_{i\geq 1}$ is a sequence of jumps independent of $N_t$ as well as $W_t$. This model is a nonparametric model in the sense that no specific structure is imposed on the functions $a(s)$, $\sigma(s)$ and $\lambda(t)$ or the jump distribution of $(Z_i)_{i\geq 1}$.

Suppose we have observations of $S_t$ at discrete times

$$0 = t_0 < t_1 < ... < t_n = T.$$

The objective of the test for jumps we are interested in is to test the hypothesis that no jump in $S_t$, or equivalently in $\log S_t$, occurred from time $t_{i-1}$ to time $t_i$ for any fixed $i = 1, 2, ..., n$. This is called testing for individual jumps. Therefore, under the null hypothesis we assume that from time $t_{i-1}$ to time $t_i$ our observations are a realisation of a diffusion process given by

$$X_{t_i} = X_{t_{i-1}} + \int_{t_{i-1}}^{t_i} a(s)ds + \int_{t_{i-1}}^{t_i} \sigma(s)dW_s. \tag{2.2}$$

We will apply the test by Lee and Mykland (2008), in which the instantaneous volatility $\sigma_{t_{i-1}}$, where

$$\sigma^2_{t_{i-1}} = \int_{t_{i-2}}^{t_{i-1}} \sigma^2(s)ds,$$

is estimated using a series of observations before the testing period of time $t_{i-1}$ to time $t_i$. If our observations up to time $t_{i-1}$ were a realisation of the pure diffusion process in (2.2) then a measure for the instantaneous volatility can be derived from an estimate of the quadratic variation of the process. However, in the presence of jumps, using an estimate of quadratic variation will lead to overestimation of $\sigma_{t_{i-1}}$. To obtain a consistent[1] and unbiased estimate for $\sigma_{t_{i-1}}$ in the presence of jumps, realised bipower variation, developed by Barndorff-Nielsen and Shephard (2004), will be used.

[1]In this dissertation we will use the term "consistent estimator" loosely. When $\sigma$ is constant or deterministic, then our estimator we refer to will be a consistent estimator. However when $\sigma$ is random, then our estimator is a consistent estimator of the expected value of instantaneous volatility.

In this chapter special attention will be given to the concepts of quadratic variation and bipower variation of a jump diffusion process. Therefore, in Section 2.1 jump diffusion processes will be discussed, with emphasis on compound Poisson processes. In Section 2.2 we will define various forms of variation, where we will also derive the bipower variation for some special cases of the more general jump diffusion process defined in (2.1). Several tests for jumps based on bipower variation will be considered in Section 2.3, where the test of Lee and Mykland (2008) will be discussed in detail.

## 2.1. Jump diffusion process

**2.1.1. Diffusion processes.** Our discussion on diffusion processes will begin with the definition of a Brownian motion $W_t$. Occasionally we will use the notation $W(t, \omega)$ when emphasizing that the process $W_t$ is a function of time $t$ and $\omega \in \Omega$.

DEFINITION 2.1. [Kuo (2006) page 7] A stochastic process $W(t, \omega)$ is called a Brownian motion when the following conditions are met:

- $P[\omega \in \Omega; W(0, \omega) = 0] = 1$, that is the process $W_t$ at time $t = 0$ almost surely takes on the value of 0.
- For any $0 \leq s < t$, the random variable $W_t - W_s$ is normally distributed with mean 0 and variance $(t - s)$, that is for any $a < b$
$$P[a \leq W_t - W_s \leq b] = \frac{1}{\sqrt{2\pi(t-s)}} \int_a^b e^{-\frac{x^2}{2(t-s)}} dx.$$
- $W_t$ has independent increments, i.e. for any $0 \leq t_1 < t_2 < ... < t_n$ the random variables
$$W_{t_1}, W_{t_2} - W_{t_1}, ..., W_{t_n} - W_{t_{n-1}}$$
are independent.
- Almost all sample paths of $W(t, \omega)$ are continuous, i.e.
$$P[\omega \in \Omega; W(\cdot, \omega) \text{ is continuous}] = 1.$$

Now having defined a Brownian motion process we can state some simple statistical properties. Since $W_t$ is normally distributed with mean zero and variance $t$, the following properties can easily be derived:

- the covariance of $W_t$ and $W_s$ is given by
$$\text{Cov}(W_t, W_s) = \min(s, t);$$
- the moment generating function is given by
$$M_X(x) = e^{\frac{tx^2}{2}},$$
in this case $X = W_t$.

Using Brownian motion we can define several types of stochastic integrals of the form

$$X(t, \omega) = \int_0^t f(s, \omega) dW(s, \omega). \tag{2.3}$$

These integrals are also stochastic processes, since they are functions of time $t$ as well as $\omega \in \Omega$. Therefore we are interested in the distributional properties of $X_t$. These properties will depend on the characteristics of the function $f$. For the case where $f$ is a function of time (s), and not $\omega$, the function $f$ is not random and is called a deterministic function. In this case the stochastic integral $X_t$ in (2.3) is called a Wiener integral.

Definition 2.2 (Kuo (2006) page 11). The Wiener integral is defined as the stochastic process

$$X_t = \int_0^t f(s)dW_s, \tag{2.4}$$

where $f$ is a deterministic function of time $t$.

One may see from this definition that this integral is a random variable for fixed $t$, thus we find it necessary to discuss statistical properties of the Wiener integral. The most important properties of which are as follows:

- the mean of $X_t$ is
$$\mathrm{E}(X_t) = 0;$$
- the variance of $X_t$ is given by
$$\mathrm{Var}(X_t) = \int_0^t f^2(s)ds;$$
- the covariance of two Wiener integrals $X_t$ and $X_s$ is given by
$$\mathrm{Cov}(X_t, X_s) = \int_0^{\min(t,s)} f^2(u)du.$$
This is a special case of the more general result that
$$\mathrm{Cov}(X_t, Y_s) = \int_0^{\min(t,s)} f(u)g(u)du, \tag{2.5}$$
where $X_t = \int_0^t f(u)dW_u$ and $Y_s = \int_0^s g(u)dW_u$.

It can also be proven that the Wiener integral $X_t$ is a Gaussian random variable; for this proof and the derivations of the properties above see Kuo (2006). Specifically in Kuo (2006) page 12 the covariance of $X_t$ and $Y_t$, given by

$$\mathrm{Cov}(X_t, Y_t) = \int_0^t f(u)g(u)du,$$

is derived. We will now extend this result to the result stated in (2.5).

Using $X_t$ and $Y_s$ as above for $s < t$ we find

$$\begin{aligned}\mathrm{Cov}(X_t, Y_s) &= \mathrm{Cov}\left(\int_0^t f(u)dW_u, \int_0^s g(u)dW_u\right) \\ &= \mathrm{Cov}\left(\int_0^s f(u)dW_u, \int_0^s g(u)dW_u\right) + \mathrm{Cov}\left(\int_s^t f(u)dW_u, \int_0^s g(u)dW_u\right) \\ &= \mathrm{Cov}(X_s, Y_s) + 0,\end{aligned}$$

having used independence of Brownian motion increments above. Using the result in Kuo (2006) for $\mathrm{Cov}(X_s, Y_s)$ we find that

$$\mathrm{Cov}(X_s, Y_t) = \int_0^s f(u)g(u)du$$

and that this result similarly holds for $t < s$ resulting in

$$\text{Cov}(X_s, Y_t) = \int_0^t f(u)g(u)du,$$

which proves the result in (2.5).

Since the Wiener integral is also a normally distributed random variable we can also easily realise that the moment generating function is simply a special case of the moment generating function of the normal distribution where the mean is 0 and the variance is $\text{Var}(X_t)$,

$$M_{X_t}(x) = e^{\frac{\text{Var}(X_t)x^2}{2}}.$$

In the special case when $f$ is a constant $\sigma$, then the Wiener integral $X_t$, given in equation (2.4), will become the more general Brownian motion. In this case $X_t$ will be normally distributed with mean 0 and variance $\sigma^2 t$.

For now we move on to another type of diffusion process we need to discuss. In the case of the Wiener integral we used a deterministic function in our definition and discussion. This is however not always the case hence the need to discuss the case when the integrand (in our case volatility) is a random variable.

The reason why volatility should not necessarily be treated as fixed or a deterministic function is that Alizadeh et al (2001) found that the volatility of exchange rates is approximately Gaussian, that is it is a random variable and thus must be treated as such. Thus we continue by saying the volatility is a stochastic process in most practical examples, specifically in exchange rates.

Thus we consider the following integral, sometimes called an Itô integral

$$X_t = \int_0^t f(s)dW_s,$$

where $f \in L^2([0,t] \times \Omega)$. We use $L^2([0,t] \times \Omega)$ to refer to the space of all stochastic processes $f(s,\omega)$, $0 \le s \le t$, $\omega \in \Omega$ satisfying:

- $f(s,\omega)$ is $\mathcal{F}_s$ measurable, which implies that $f(s,\omega)$ can be a deterministic function or a random variable for each $s$ and is therefore a stochastic process.
- $\int_0^t \text{E}\left(|f(s)|^2\right) ds < \infty$.

Bearing in mind these conditions we can now state various important properties:

- the mean of $X_t$ is
$$\text{E}(X_t) = 0;$$
- the variance of $X_t$ is given by
$$\text{Var}(X_t) = \int_0^t \text{E}(|f(s)|^2)ds;$$
- for two Itô integrals $X_t = \int_0^t f(u)dW_u$ and $Y_s = \int_0^s g(u)dW_u$ the covariance is given by
$$\text{Cov}(X_t, Y_s) = \int_0^{\min(t,s)} \text{E}\left(f(u)g(u)\right) du.$$

For the derivations of these properties see Kuo (2006).

Since the function $f$ in this case is a random variable, this makes it more difficult to define the moment generating function of the integral without knowing the distribution of $f$. For our purposes, as previously stated, this function $f$ will be the volatility $\sigma \in L^2([0,t] \times \Omega)$.

In Barndorff-Nielsen and Shephard (2004) the stochastic volatility is defined in an even more general setting than above. In their setting $f \in \mathscr{L}\left(\Omega, L^2[0,t]\right)$, where the notation $\mathscr{L}\left(\Omega, L^2[0,t]\right)$ refers to the space of all stochastic processes $f$ such that

$$P\left(\omega \in \Omega : \int_0^t |f(s,\omega)| ds < \infty\right) = 1. \tag{2.6}$$

A key difference between this space of stochastic processes and the previous one is that $f$ in this case is not necessarily integrable, that is if $\mathrm{E}(f(t)) = \infty$, then $f \notin L^2([a,b] \times \Omega)$ but is included in the space $\mathscr{L}\left(\Omega, L^2[0,t]\right)$ if (2.6) is true, we can make a similar observation for $\mathrm{Var}\left(f(t)\right)$. We may wish to define the mean and variance for this stochastic process as in the previous cases. This is however difficult because we cannot say for certain that $\mathrm{E}[f(t)] < \infty$ is necessarily true. Thus since in the previous case we used $\mathrm{E}[W_t] = 0$ this will not help us for this larger class of stochastic processes $\mathscr{L}\left(\Omega, L^2[a,b]\right)$, since the mean in this case is heavily dependent on the function $f$.
Now having done a basic overview of various types of diffusion processes which we will use to model the volatility, we can move on to briefly assessing the types of processes used to model the jumps.

**2.1.2. Compound Poisson processes.** In this study we will be using the compound Poisson process to model the jump process, that is the discontinuous part of the jump diffusion process as defined in equation (2.1). In order to understand and discuss compound Poisson processes we need to first define and discuss Poisson processes.

Definition 2.3 (Kuo (2006) page 76). A Poisson process with parameter $\lambda > 0$ is a stochastic process $N(t,\omega)$ satisfying the following:

- $P\{\omega \in \Omega; N(0,\omega) = 0\} = 1$.
- For any $0 \leq s < t$ the random variable $N_t - N_s$ is a Poisson random variable with parameter $\lambda(t-s)$, that is

$$P\left[N_t - N_s = k\right] = e^{-\lambda(t-s)} \frac{\left(\lambda(t-s)\right)^k}{k!}, k = 0, 1, 2, ..., n, ...$$

- $N_t$ has independent increments, that is for any $0 \leq t_1 < t_2 < ... < t_n$, the random variables

$$N_{t_1}, N_{t_2} - N_{t_1}, ..., N_{t_n} - N_{t_{n-1}}$$

  are independent.
- Almost all sample paths are right continuous with left-hand limits, or càdlàg, that is

$$P\{\omega \in \Omega; N(\cdot,\omega) \text{ is right continuous with left-hand limits}\} = 1.$$

Now using this definition we can discuss similar statistical properties as with the diffusion processes. We start by stating the expected value and variance for this type of process:

- the mean of $N_t$ is given by

$$\mathrm{E}(N_t) = \lambda t;$$

- the variance of $N_t$ is given by

$$\mathrm{Var}(N_t) = \lambda t;$$

- the covariance of $N_t$ and $N_s$ is given by

$$\operatorname{Cov}(N_t, N_s) = \lambda \min(s, t).$$

We shall now derive the covariance stated above. We start by letting $N_t$ and $N_s$ both be Poisson random processes at time $t$ and $s$ respectively. If we assume $s < t$ then,

$$\begin{aligned}
\operatorname{Cov}(N_t, N_s) &= \mathrm{E}(N_t N_s) - \mathrm{E}(N_t)\,\mathrm{E}(N_s) \\
&= \mathrm{E}(N_t N_s - N_s^2 + N_s^2) - \lambda t \lambda s \\
&= \mathrm{E}\left(N_s\,(N_t - N_s) + N_s^2\right) - \lambda t \lambda s \\
&= \mathrm{E}(N_s)\,\mathrm{E}(N_t - N_s) + \mathrm{E}(N_s^2) - \lambda t \lambda s \qquad (2.7) \\
&= \lambda s \lambda (t - s) + \lambda s + \lambda^2 s^2 - \lambda t \lambda s \\
&= \lambda s \lambda t - \lambda^2 s^2 + \lambda s + \lambda^2 s^2 - \lambda s \lambda t \\
&= \lambda s.
\end{aligned}$$

Similarly for $t < s$ we will obtain

$$\operatorname{Cov}(N_t, N_s) = \lambda t.$$

Thus we obtain the result,

$$\operatorname{Cov}(N_t, N_s) = \lambda \min(t, s).$$

If we want an easier way to calculate the higher moments we can use the moment generating function. The moment generating function for a Poisson distribution with parameter $\lambda t$ is as follows:

$$M_{N_t}(x) = e^{(\lambda t)(e^x - 1)}.$$

One can use the Poisson distribution to define another important distribution which is essential in modeling jumps, the compound Poisson distribution.

When assessing jumps one finds that the jumps can occur at any time and can be of any size. For this reason we cannot use simply the Poisson distribution. For a Poisson distribution the time and size of the jumps will be integers, and for the jump size we would want, a continuous random variable.

Definition 2.4 (Ait-Sahalia and Jacod (2014) page 21). A compound Poisson process is a stochastic process that can be written in the form

$$Y_t = \sum_{i=1}^{N_t} X_i,$$

where $N_t$ is a Poisson process and each $X_i$ for $1 \leq i \leq N_t$ are independent and identically distributed random variables.

This distribution is particularly noteworthy since we can use $N_t$ to denote the jump times and $X_i$ the jump sizes. Using this definition for the compound Poisson distribution we can state the expected value and variance of such a stochastic process:

- the mean of $Y_t$ is given by

$$\mathrm{E}(Y_t) = \lambda t\, \mathrm{E}(X_1);$$

- the variance of $Y_t$ is given by

$$\mathrm{Var}(Y_t) = \lambda t \, \mathrm{E}(X_1^2);$$

- the covariance of $Y_t$ and $Y_s$ is given by

$$\mathrm{Cov}(Y_t, Y_s) = \lambda \min(t, s) \, \mathrm{E}(X_1^2).$$

Before we proceed to derive the covariance result we will state Wald's equation since this will be used in this derivation. This is

$$\mathrm{E}(X_1 + ... + X_N) = E(N)E(X_1) \tag{2.8}$$

as found in Khoshnevisan (2007) where the $X_i$'s are independent and identically distributed. Now for two compound Poisson processes $Y_t$ and $Y_s$, assuming $s < t$ we get,

$$\begin{aligned}
\mathrm{Cov}(Y_t, Y_s) &= \mathrm{E}(Y_t Y_s) - \mathrm{E}(Y_t)\,\mathrm{E}(Y_s) \\
&= \mathrm{E}\left(\sum_{i=1}^{N_t} X_i \sum_{j=1}^{N_s} X_j\right) - \mathrm{E}\left(\sum_{i=1}^{N_t} X_i\right)\mathrm{E}\left(\sum_{j=1}^{N_s} X_j\right) \\
&= \mathrm{E}\left(\left[\sum_{i=1}^{N_s} X_i + \sum_{i=N_s+1}^{N_t} X_i\right]\sum_{j=1}^{N_s} X_j\right) - \mathrm{E}\left(\sum_{i=1}^{N_t} X_i\right)\mathrm{E}\left(\sum_{j=1}^{N_s} X_j\right) \\
&= \mathrm{E}\left(\left(\sum_{i=1}^{N_s} X_i\right)^2 + \sum_{i=1}^{N_s} X_i \sum_{j=N_s+1}^{N_t} X_j\right) - \lambda^2 ts E(X_1)^2 \\
&= \mathrm{E}\left(\sum_{i=1}^{N_s} X_i^2 + \sum_{\substack{j=1 \\ j\neq k}}^{N_s}\sum_{k=1}^{N_s} X_j X_k + \sum_{i=1}^{N_s} X_i \sum_{j=N_s+1}^{N_t} X_j\right) - \lambda^2 ts E(X_1)^2 \\
&= \mathrm{E}(N_s)\,\mathrm{E}(X_1^2) + \mathrm{E}(N_s^2 - N_s)\,\mathrm{E}(X_1)^2 + \mathrm{E}(N_s(N_t - N_s))\,\mathrm{E}(X_1)^2 - \lambda^2 ts E(X_1)^2 \\
&= \lambda s \, \mathrm{E}(X_1^2) + (\mathrm{E}(N_s N_t) - \mathrm{E}(N_s))\,\mathrm{E}(X_1)^2 - \lambda^2 ts E(X_1)^2 \\
&= \lambda s \, \mathrm{E}(X_1^2) + \mathrm{E}(N_t N_s - N_s)\,\mathrm{E}(X_1)^2 - \lambda^2 ts E(X_1)^2 \\
&= \lambda s \, \mathrm{E}(X_1^2) + (\lambda s + \lambda^2 ts - \lambda s)\,\mathrm{E}(X_1)^2 - \lambda^2 ts E(X_1)^2 \\
&= \lambda s \, \mathrm{E}(X_1^2),
\end{aligned}$$

where we have used Wald's equation as in (2.8) and our previous result in (2.7). We similarly find for $t < s$ that

$$\mathrm{Cov}(Y_t, Y_s) = \lambda t \, \mathrm{E}(X_1^2).$$

Thus the result follows.

For an easier way to determine the higher moments we can use the moment generating function. If we don't assume any specific distribution for the random variables $X_i$ then we cannot know the exact moment generating function but we can derive the following expression:

$$M_Y(x) = e^{(M_X(x)-1)\lambda t},$$

where $M_X(x)$ represents the moment generating function of the random variables $X_i$.

A special case that we will be interested in is when we let the random variables $X_i$ be given by $Z_i$ which are normally distributed with mean 0 and some variance $\sigma_Z^2$. This allows us to determine the mean, variance and moment generating function more precisely than above, given by

$$\mathrm{E}\left(\sum_{i=1}^{N_t} Z_i\right) = 0$$

$$\mathrm{Var}\left(\sum_{i=1}^{N_t} Z_i\right) = \sigma_Z^2 \lambda t$$

$$M_Y(x) = e^{\left(\exp\left(\frac{\sigma_Z^2 x}{2}\right) - 1\right)\lambda t}$$

Another result which we will find useful later on is that of the absolute value of the normal distribution which is sometimes called the half-normal distribution. We find that the expected value of this random variable is as follows:

Proposition 2.5. *For some normally distributed random variable $X$ with mean $0$ and variance $\sigma_X^2$, the expected value of $|X|$ is found to be* $\mathrm{E}\,|X| = \sigma_X\sqrt{\frac{2}{\pi}}$

Proof. We use the definition of the expected value and double the positive part since we are looking at the absolute value. This gives

$$\begin{aligned}
\mathrm{E}\,|X| &= 2\int_0^\infty \left| \frac{x}{\sqrt{2\sigma_X^2\pi}} e^{-\frac{x^2}{2\sigma_X^2}} \right| dx \\
&= \sqrt{\frac{2}{\sigma_X^2\pi}} \left[ -\sigma_X^2 e^{-\frac{x^2}{2\sigma_X^2}} \right|_{x=0}^{\infty} \\
&= \sigma_X\sqrt{\frac{2}{\pi}}.
\end{aligned}$$

□

In our discussion of the jump process we will be assuming that the jumps are normal with mean 0.

Now having assessed the types of random processes that we will use to model the jumps and volatility, we can discuss some basic properties of the jump diffusion process as in (2.1), which combines these two. Since the drift component is deterministic, which is not random, there is not much we can say about the statistical properties thereof. Using $X_t$ as in (2.1) we can state some of the basic properties of the jump diffusion model:

- the mean of $X_t$ is given by

$$\mathrm{E}(X_t) = \int_0^t a(s)ds + \lambda t\,\mathrm{E}(Z_1);$$

- the variance of $X_t$ is given by

$$\mathrm{Var}(X_t) = \int_0^t \mathrm{E}(|\sigma(s)|^2)ds + \lambda t E(Z_1^2); \tag{2.9}$$

- the covariance of $X_t$ and some $X_s$ is given by

$$\mathrm{Cov}(X_t, X_s) = \int_0^{\min(t,s)} \mathrm{E}(|\sigma(s)|^2)ds + \lambda \min(t,s) E(Z_1^2),$$

where we assume $\lambda$ is constant, the results however can be extended for $\lambda(t)$.

The results above are true if $\sigma \in L^2$. However if $\sigma \in \mathscr{L}$, then the variance of $X_t$ may not exist. In this case, we will have to define the quadratic variation in order to define a quantity that will describe the deviation of the sample paths of $X_t$. This will be the topic of the next section as well as the concepts of bipower and quadratic variation.

## 2.2. Quadratic and bipower variation

In testing for jumps it is common to use the instantaneous volatility, that is the volatility $\sigma$ at some time $t$. This is the method we will be employing in our practical application. This may sound simple but there is significant theory to discuss before we can do it, since as previously noted for $\sigma \in \mathscr{L}$ the variance of the jump-diffusion process $X_t$ may not exist.

In order to resolve this issue we require a statistic that will describe the variation of $X_t$ that will still be defined in the general case when $\sigma \in \mathscr{L}$. For this we define the quadratic variation. As will be seen later in this section we find that in the presence of jumps the quadratic variation will include both the continuous and jump components. This will not be the end of our discussion however since our goal is to isolate the instantaneous volatility.

Another quantity which will help us achieve this goal is the bipower variation, which will also help to describe the deviation of sample paths except that it is unaffected by jumps, as we will see shortly. These two quantities of quadratic variation and bipower variation can then be used to estimate the volatility at each point. Thus we will proceed to discuss these concepts in more detail, starting with more formal definitions.

For the following definition, as well as in the sequel, we let $\Delta_n = \{0 = t_0 < t_1 < ... < t_n = t\}$ be a partition of $[0, t]$, we use $\|\Delta_n\| = \max_{1 \leq i \leq n}(t_i - t_{i-1})$ and finally note that we will use the notation $X_n \xrightarrow{p} X$ or $X = p \lim_{n \to \infty} X_n$ for convergence in probability.

Definition 2.6 (Barndorff-Nielsen and Shephard (2004) page 5). The quadratic variation for a process $X_t$ is defined to be:

$$[X]_t = p \lim_{\|\Delta_n\| \to 0} \sum_{j=1}^{n} [X_{t_j} - X_{t_{j-1}}]^2.$$

For a compound Poisson process this can be stated more simply without the limit as

$$[X]_t = \sum_{0 \leq s \leq t} (\Delta X_s)^2 \,, \tag{2.10}$$

where $\Delta X_s$ are the jump sizes defined by $\Delta X_s = X_s - \lim_{t \to s^-} X_t$.

Quadratic variation is a special case of the more general power variation defined on page 8 of Barndorff-Nielsen and Shephard (2004).

Using this definition for quadratic variation we can now determine the quadratic variation for each of the types of processes discussed in the previous section. We will start with the diffusion process.

THEOREM 2.7. *Suppose $\sigma \in \mathscr{L}$. The quadratic variation of a stochastic volatility process, of the form $Y_t = \int_0^t \sigma(s)dW_s$ where $\sigma$ is the volatility process and $W_t$ is a Brownian motion, is*

$$[Y]_t = \int_0^t \sigma^2(s)ds$$

A method that can be used to derive the quadratic variation of a process is to form a sequence of random variables

$$X_n = \sum_{i=1}^{n}(X_{t_i} - X_{t_{i-1}})^2,$$

and to show that $X_n$ converges to $[X]_t$ in $L^2$, denoted $X_n \xrightarrow{L^2} [X]_t$. However, if $X_t$ is an Itô integral with $\sigma \in \mathscr{L}$ then this approach is difficult to apply. In this dissertation we will supply proofs for some special cases of the function $\sigma$. For the case where $\sigma$ is constant the proof for the quadratic variation can be found on page 38 of Kuo (2006). For the more general proof see Barndorff-Nielsen and Shepard (2004). We will supply the proof for the special case where $\sigma$ is a deterministic function of time $t$ (see Appendix A.1).

We now move on to the results for the compound Poisson process, using normally distributed $Z_i$. The quadratic variation is as follows:

THEOREM 2.8. *The quadratic variation for a compound Poisson process $Y$ is*

$$[Y]_t = \sum_{i=1}^{N_t} Z_i^2.$$

PROOF. Using the alternate definition given by equation (2.10), we see that $\Delta X_s$ in that equation will be $\Delta Y_s$ here and for $0 \le s \le t$ each jump size will be given by $Z_{N_s}$ thus leading us to find,

$$\begin{aligned}
\Delta Y_s &= Y_s - \lim_{t \to s^-} Y_t \\
&= Z_{N_s} 1_{\{\Delta N_s = 1\}} \\
\therefore (\Delta Y_s)^2 &= Z_{N_s}^2 1_{\{\Delta N_s = 1\}} \\
\therefore \sum_{0 \le s \le t} (\Delta Y_s)^2 &= \sum_{0 \le s \le t} Z_{N_s}^2 1_{\{\Delta N_s = 1\}} \\
&= \sum_{i=1}^{N_t} Z_i^2
\end{aligned}$$

A more rigorous proof can be seen in Appendix A.2 □

Finally we can look at a jump diffusion process $X_t$ as in (2.1).

THEOREM 2.9. *The quadratic variation of a jump diffusion process of the form*

$$X_t = \int_0^t a(s)ds + \int_0^t \sigma(s)dW_s + \sum_{i=1}^{N_t} Z_i,$$

*is given by*

$$[X]_t = \int_0^t \sigma^2(s)ds + \sum_{i=1}^{N_t} Z_i^2.$$

This result is a consequence of the independence of $N_t$ and $W_s$ as well as the fact that, for instance $\int_{t_{i-1}}^{t_i} a(s)ds \to a(s)ds$. One can then use the table provided on page 103 of Kuo (2006) to show that $\left[\int_{t_{i-1}}^{t_i} a(s)ds\right]^2 \to 0$, $\int_{t_{i-1}}^{t_i} a(s)ds \int_{t_{i-1}}^{t_i} \sigma(s)dW_s^2 \to 0$, and finally that $\int_{t_{i-1}}^{t_i} a(s)ds \sum_{i=1}^{N_t - N_{t-1}} Z_i \to 0$, as $\Delta t \to 0$. A noteworthy detail is that the $\mathrm{Var}(X_t)$, as in (2.9), will be equal to $\mathrm{E}([X]_t)$ as above if $\sigma \in L^2$. For now though having calculated the quadratic variation we can move on to discuss the bipower variation.

Definition 2.10 (Barndorff-Nielsen and Shephard (2004) page 9). The bipower variation for a process $X_t$ is defined to be:

$$[X]_t^* = p \lim_{\|\Delta_n\| \to 0} \sum_{j=1}^{n-1} |X_{t_j} - X_{t_{j-1}}||X_{t_{j+1}} - X_{t_j}|$$

This definition is a special case which can be extended to a more general form which can be seen in Barndorff-Nielsen and Shephard (2004). Similar to above we will look at the bipower variation for the diffusion process first.

Theorem 2.11. *The bipower variation of a stochastic volatility process, of the form* $Y_t = \int_0^t \sigma(s)dW_s$ *where* $\sigma$ *is the volatility process and* $W_t$ *is a Brownian motion, is*

$$[Y]_t^* = \frac{2}{\pi} \int_0^t \sigma^2(s)ds.$$

Proof. We will assess the case where $\sigma$ is a fixed constant. Let

$$X_i = |\Delta Y_{t_i}||\Delta Y_{t_{i+1}}| - \frac{2}{\pi}\sigma^2 \Delta t_i$$

where $\Delta Y_{t_i} = \int_{t_{i-1}}^{t_i} \sigma(s)dW_s$ and $\Delta t_i = t_i - t_{i-1}$ , since this will allow us to prove that

$$\sum_{i=1}^{n-1} |\Delta Y_{t_i}||\Delta Y_{t_{i+1}}| \xrightarrow{P} \frac{2}{\pi}\sigma^2 t$$

by showing that $\sum_{i=1}^{n-1} X_i \xrightarrow{P} 0$. To show convergence in probability we will show that $\mathrm{E}\left(\sum_{i=1}^{n-1} X_i\right)^2 \to 0$. We will be demonstrating that $\mathrm{E}\left(\sum_{i=1}^{n-1} X_i^2\right) \to 0$ and $\mathrm{E}\left(\sum_{j=1}^{n-1} \sum_{k=1}^{n-1}{}_{j \neq k} X_i X_j\right) \to 0$. To do this we proceed by calculating $\mathrm{E}(X_i^2)$, $\mathrm{E}(X_i X_{i+1})$ and $\mathrm{E}(X_i X_j)$ for $j \notin \{i-1, i, i+1\}$. Since $\mathrm{E}(|\Delta Y_{t_i}|) = \sigma\sqrt{\frac{2\Delta t_i}{\pi}}$, we have that

$$\begin{aligned}
\mathrm{E}\left(X_i^2\right) &= \mathrm{E}\left[(\Delta Y_{t_i})^2 \left(\Delta Y_{t_{i+1}}\right)^2 - \frac{4}{\pi}\sigma^2 \Delta t_i |\Delta Y_{t_i}||\Delta Y_{t_{i-1}}| + \frac{4}{\pi^2}\sigma^4 \Delta t_i^2\right] \\
&= \sigma^4 \Delta t_i \Delta t_{i+1} - \frac{4}{\pi}\sigma^4 \sqrt{\frac{2\Delta t_i}{\pi}} \sqrt{\frac{2\Delta t_{i+1}}{\pi}} \Delta t_i + \frac{4}{\pi^2}\sigma^4 \Delta t_i^2 \\
&= \sigma^4 \Delta t_i \Delta t_{i+1} - \frac{8}{\pi^2}\sigma^4 \Delta t_i \sqrt{\Delta t_i \Delta t_{i+1}} + \frac{4}{\pi^2}\sigma^4 \Delta t_i^2,
\end{aligned}$$

and that

$$\begin{aligned}
\mathrm{E}\left(X_i X_{i+1}\right) &= \mathrm{E}\bigg[\left(|\Delta Y_{t_i}||\Delta Y_{t_{i+1}}| - \frac{2}{\pi}\sigma^2\Delta t_i\right)\left(|\Delta Y_{t_{i+1}}||\Delta Y_{t_{i+2}}| - \frac{2}{\pi}\sigma^2\Delta t_{i+1}\right)\bigg] \\
&= \mathrm{E}\bigg[|\Delta Y_{t_i}|(\Delta Y_{t_{i+1}})^2|\Delta Y_{t_{i+2}}| - \frac{2}{\pi}\sigma^2\Delta t_i|\Delta Y_{t_{i+1}}||\Delta Y_{t_{i+2}}| \\
&\quad - \frac{2}{\pi}\sigma^2\Delta t_{i+1}|\Delta Y_{t_i}||\Delta Y_{t_{i+1}}| + \frac{4}{\pi^2}\sigma^4\Delta t_i\Delta t_{i+1}\bigg] \\
&= \frac{2}{\pi}\sigma^4\Delta t_{i+1}\sqrt{\Delta t_i\Delta t_{i+2}} - \frac{4}{\pi^2}\sigma^4\Delta t_i\sqrt{\Delta t_{i+1}\Delta t_{i+2}} \\
&\quad - \frac{4}{\pi^2}\sigma^4\Delta t_{i+1}\sqrt{\Delta t_i\Delta t_{i+1}} + \frac{4}{\pi^2}\sigma^4\Delta t_i\Delta t_{i+1}.
\end{aligned}$$

Note that $X_i$ and $X_j$ are independent for $j \neq i-1$ and $j \neq i$ since $\Delta Y_i$, $\Delta Y_{i+1}$, $\Delta Y_j$, $\Delta Y_{j+1}$ are independent for $i < j-1$ and $i > j+1$. Thus we have that

$$\begin{aligned}
\mathrm{E}(X_i X_j) &= \mathrm{E}(X_i)\,\mathrm{E}(X_j) \\
&= 0,
\end{aligned}$$

since $\mathrm{E}(X_i) = 0$ and $\mathrm{E}(X_j) = 0$.

Finally we have that

$$\begin{aligned}
\mathrm{E}\left(\sum_{i=1}^{n-1} X_i\right)^2 &= \sum_{i=1}^{n-1}\mathrm{E}\left(X_i^2\right) + 2\sum_{i=1}^{n-2}\mathrm{E}\left(X_i X_{i+1}\right) + \mathrm{E}\left(\sum_{\substack{j=1 \\ j\neq k \;\; j\neq k+1 \;\; j\neq k-1}}^{n-1}\sum_{k=1}^{n-1} X_j X_k\right) \\
&\leq \sum_{i=1}^{n-1}\left(\sigma^4\Delta t_i\Delta t_{i+1} + \frac{4}{\pi^2}\sigma^4\Delta t_i^2\right) \\
&\quad + 2\sum_{i=1}^{n-2}\left(\frac{2}{\pi}\sigma^4\Delta t_{i+1}\sqrt{\Delta t_i\Delta t_{i+2}} + \frac{4}{\pi^2}\sigma^4\Delta t_i\Delta t_{i+1}\right) \\
&\leq \left(1 + \frac{4}{\pi^2}\right)\sigma^4 t\|\Delta_n\| + 2\left(\frac{2}{\pi}\sigma^4\|\Delta_n\|t + \frac{4}{\pi^2}\sigma^4\|\Delta_n\|t\right) \\
&\to 0
\end{aligned}$$

as $\|\Delta_n\| \to 0$.

As mentioned above this is for the case when the volatility is a fixed constant, which will not usually be the case. Hence we look at more general cases in Appendix A.3. □

We now assess the bipower variation for the compound Poisson process given as follows;

Theorem 2.12. *The bipower variation for a compound Poisson process is 0.*

Proof. Let $X_i = |Y_{t_i} - Y_{t_{i-1}}||Y_{t_{i+1}} - Y_{t_i}|$. In this proof we assess $\mathrm{E}\left(\sum_{i=1}^{n-1} X_i\right)^2$ in order to prove convergence in probability. Let $\Delta Y_i = Y_{t_i} - Y_{t_{i-1}}$, $\Delta t_i = t_i - t_{i-1}$ and $\Delta N_i = N_{t_i} - N_{t_{i-1}}$. We now turn our attention to $\mathrm{E}\,|\Delta Y_i|$. Using the triangle inequality and Wald's equation we have the following:

$$
\begin{aligned}
\mathrm{E}\left|\Delta Y_i\right| &\leq \mathrm{E}\left(\sum_{j=N_{t_{i-1}}}^{N_{t_i}} |Z_j|\right) \\
&= \mathrm{E}\left(\Delta N_i\right)\mathrm{E}\left|Z_j\right| \\
&= \lambda \Delta t_i \sigma_Z \sqrt{\frac{2}{\pi}}.
\end{aligned}
$$

This makes estimating $\mathrm{E}\left(X_i\right)$ easier, for which we also use the fact that a compound Poisson process has independent increments:

$$
\begin{aligned}
\mathrm{E}\left(X_i\right) &= \mathrm{E}\left(|\Delta Y_i||\Delta Y_{i+1}|\right) \\
&\leq \lambda^2 \Delta t_i \Delta t_{i+1} \sigma_Z^2 \frac{2}{\pi}.
\end{aligned}
$$

We now calculate $\mathrm{E}\left(X_i^2\right)$ using the fact that $|\Delta Y_i|^2 = (\Delta Y_i)^2$:

$$
\begin{aligned}
\mathrm{E}\left(X_i^2\right) &= \mathrm{E}\left(|\Delta Y_i||\Delta Y_{i+1}||\Delta Y_i||\Delta Y_{i+1}|\right) \\
&= \mathrm{E}\left[\left(\Delta Y_i\right)^2\left(\Delta Y_{i+1}\right)^2\right] \\
&= \sigma_Z^4 \lambda^2 \Delta t_i \Delta t_{i+1}.
\end{aligned}
$$

Finally we require the following two results before we can proceed. We firstly have that

$$
\begin{aligned}
\mathrm{E}\left(X_i X_{i+1}\right) &= \mathrm{E}\left(|\Delta Y_i||\Delta Y_{i+1}||\Delta Y_{i+1}||\Delta Y_{i+2}|\right) \\
&= \mathrm{E}\left(|\Delta Y_i||\Delta Y_{i+1}|^2|\Delta Y_{i+2}|\right) \\
&\leq \left(\lambda \Delta t_i \sigma_Z \sqrt{\frac{2}{\pi}}\right)\left(\sigma_Z^2 \lambda \Delta t_{i+1}\right)\left(\lambda \Delta t_{i+2} \sigma_Z \sqrt{\frac{2}{\pi}}\right) \\
&= \lambda^3 \sigma_Z^4 \frac{2}{\pi} \Delta t_i \Delta t_{i+1} \Delta t_{i+2}.
\end{aligned}
$$

We calculate the following for $i \neq j$, $i \neq j+1$ and $i \neq j-1$:

$$
\begin{aligned}
\mathrm{E}\left(X_i X_j\right) &= \mathrm{E}\left(|\Delta Y_i||\Delta Y_{i+1}||\Delta Y_j||\Delta Y_{j+1}|\right) \\
&= \mathrm{E}\left|\Delta Y_i\right|\mathrm{E}\left|\Delta Y_{i+1}\right|\mathrm{E}\left|\Delta Y_j\right|\mathrm{E}\left|\Delta Y_{j+1}\right| \\
&\leq \lambda^4 \sigma_Z^4 \frac{4}{\pi^2} \Delta t_i \Delta t_{i+1} \Delta t_j \Delta t_{j+1}.
\end{aligned}
$$

Let $\|\Delta_n\| = \max_{1\le i\le n}(\Delta t_i)$ be as above. We then have that

$$\begin{aligned}
\mathrm{E}\left(\sum_{i=1}^{n-1} X_i\right)^2 &= \mathrm{E}\left(\sum_{i=1}^{n-1} X_i^2\right) + 2\,\mathrm{E}\left(\sum_{i=1}^{n-2} X_i X_{i+1}\right) + \mathrm{E}\left(\sum_{\substack{j=1\\ j\neq k}}^{n-1}\sum_{\substack{k=1\\ j\neq k+1\\ j\neq k-1}}^{n-1} X_j X_k\right)\\
&= \sum_{i=1}^{n-1} \mathrm{E}\left(X_i^2\right) + 2\sum_{i=1}^{n-2}\mathrm{E}\left(X_i X_{i+1}\right) + \sum_{\substack{j=1\\ j\neq k}}^{n-1}\sum_{\substack{k=1\\ j\neq k+1\\ j\neq k-1}}^{n-1} \mathrm{E}\left(X_j X_k\right)\\
&\le \sum_{i=1}^{n-1}\left(\sigma_Z^4\lambda^2\Delta t_i\Delta t_{i+1}\right) + 2\sum_{i=1}^{n-2}\left(\lambda^3\sigma_Z^4\frac{2}{\pi}\Delta t_i\Delta t_{i+1}\Delta t_{i+2}\right)\\
&\quad + \sum_{\substack{j=1\\ j\neq k}}^{n-1}\sum_{\substack{k=1\\ j\neq k+1\\ j\neq k-1}}^{n-1}\left(\lambda^4\sigma_Z^4\frac{4}{\pi^2}\Delta t_i\Delta t_{i+1}\Delta t_j\Delta t_{j+1}\right)\\
&\le \sigma_Z^4\lambda^2\|\Delta_n\|\sum_{i=1}^{n-1}\Delta t_{i+1} + 2\lambda^3\sigma_Z^4\frac{2}{\pi}\|\Delta_n\|^2\sum_{i=1}^{n-2}\Delta t_i\\
&\quad + \lambda^4\sigma_Z^4\frac{4}{\pi^2}\|\Delta_n\|^2\sum_{i=1}^{n}\sum_{j=1}^{n}\Delta t_i\Delta t_j\\
&\le \sigma_Z^4\lambda^2\|\Delta_n\|t + 2\lambda^3\sigma_Z^4\frac{2}{\pi}\|\Delta_n\|^2 t\\
&\quad + \lambda^4\sigma_Z^4\frac{4}{\pi^2}\|\Delta_n\|^2 t^2\\
&\to 0
\end{aligned}$$

as $\|\Delta_n\| \to 0$. □

Finally we look at the bipower variation of the jump diffusion process $X_t$ as in (2.1).

THEOREM 2.13. *The bipower variation for a jump-diffusion process of the form*

$$X_t = \int_0^t a(s)ds + \int_0^t \sigma(s)dW_s + \sum_{i=1}^{N_t} Z_i,$$

*is given by*

$$[X]_t^* = \frac{2}{\pi}\int_0^t \sigma(s)^2 ds.$$

This result follows similarly to Theorem 2.9.

We require these results to eventually attempt to estimate the variance at a specific point in time, that is the instantaneous variance.

DEFINITION 2.14. The instantaneous variance for a volatility process $\sigma$ at a time $t_i$ is given by

$$\sigma_{t_i}^2 = \frac{1}{t_i - t_{i-1}}\int_{t_{i-1}}^{t_i}\sigma^2(s)ds.$$

The instantaneous volatility is the square root of the instantaneous variance $\sigma_{t_i}^2$. The objective is to find a measure of instantaneous volatility in the presence of jumps from our data. When testing for jumps the daily exchange rate movement will be scaled with the reciprocal of this measure of instantaneous volatility.

Please note that since $\sigma$ can be random the instantaneous volatility can also be random. Thus we will need to state the expected value as we have done in the previous section. If the process $\sigma(s)$ is assumed to be stationary then Barndorff-Nielsen and Shephard (2001; 2002) state the expected value as

$$\begin{aligned} \mathrm{E}(\sigma_{t_i}^2) &= \int_{t_{i-1}}^{t_i} \mathrm{E}(\sigma^2(s))ds \\ &= \xi(t_i - t_{i-1}), \end{aligned}$$

where $\xi$ refers to the constant mean of $\sigma^2(s)$. Stationarity, which ensures that the mean is constant, will however not always hold. Therefore measuring the instantaneous variance will not be as easy as finding a constant $\xi$. We require a means of measuring the variance and as such we will assess the quadratic variance and bipower variation.

We start by assessing an estimate of the quadratic variation, the estimate is referred to as the realised variance.

Definition 2.15 (Barndorff-Nielsen and Shephard (2004) page 4). The realised variance is defined to be

$$\widetilde{[X]}_t = \sum_{i=1}^{n} |X_{t_i} - X_{t_{i-1}}|^2.$$

We can restate the above to help us define instantaneous realised variance rather than simply realised variance as follows,

$$\begin{aligned} \widetilde{[X]}_{t_{i-K+2},t_i} &= \widetilde{[X]}_{t_i} - \widetilde{[X]}_{t_{i-K+2}} \\ &= \sum_{j=1}^{i} |X_{t_j} - X_{t_{j-1}}|^2 - \sum_{j=1}^{i-K+2} |X_{t_j} - X_{t_{j-1}}|^2 \\ &= \sum_{j=i-K+3}^{i} |X_{t_j} - X_{t_{j-1}}|^2, \end{aligned} \tag{2.11}$$

where $K$ will refer to the window size that is being assessed in the bipower case (i.e. the number of data points from time $t_{i-K+1}$ to time $t_i$), for now we will use $K-1$ data points. As with the quadratic variation this is again a special case of the more general realised power variation found in Barndorff-Niesen and Shephard (2004).

The window size $K$ will refer to the number of data points used in the bipower variation later and not the number of intervals assessed. The example in table (2.1) helps us see that since for the realised variance $K-1$ data points are used this means that only $K-2$ intervals can be used for the sum. This is the reason we use $t_{i-K+2}$ rather than simply $t_{i-K}$ in equation (2.11) and why we will use $\frac{1}{K-2}$ in equation (2.13) when we attempt to define a measure of the instantaneous variance.

By the definition of quadratic variation (2.6) we can easily see that the realised variance will converge in probability to the quadratic variation, that is,

TABLE 2.1. Window size example

| | |
|---|---|
| **Example $K$ value** | 5 |
| **Example Data points:** | $\{X_{t_1}, X_{t_2}, X_{t_3}, X_{t_4}, X_{t_5}\}$ |
| **Realised variance:** | $(X_{t_3} - X_{t_2})^2 + (X_{t_4} - X_{t_3})^2 + (X_{t_5} - X_{t_4})^2$ |
| **Realised bipower variation:** | $\lvert X_{t_2} - X_{t_1}\rvert\lvert X_{t_3} - X_{t_2}\rvert + \lvert X_{t_3} - X_{t_2}\rvert\lvert X_{t_4} - X_{t_3}\rvert + \lvert X_{t_4} - X_{t_3}\rvert\lvert X_{t_5} - X_{t_4}\rvert$ |

- for the case where there are no jumps we have,

$$\begin{aligned}\widetilde{[X]}_{t_{i-K+2},t_i} &\xrightarrow{P} [X]_{t_i} - [X]_{t_{i-K+2}} \\ &= \int_0^{t_i} \sigma^2(s)ds - \int_0^{t_{i-K+2}} \sigma^2(s)ds \\ &= \int_{t_{i-K+2}}^{t_i} \sigma^2(s)ds;\end{aligned}$$

- for the case where there are jumps we have,

$$\begin{aligned}\widetilde{[X]}_{t_{i-K+2},t_i} &\xrightarrow{P} [X]_{t_i} - [X]_{t_{i-K+2}} \\ &= \int_{t_{i-K+2}}^{t_i} \sigma^2(s)ds + \sum_{j=N_{t_{i-K+2}}}^{N_{t_i}} Z_j^2.\end{aligned}$$

Since the drift process will have a quadratic variation of 0 we ignore it. Now having stated the limit in probability of the realised variance we also need to assess the expected value of the realised variance. We can calculate the expected value as follows:

$$\begin{aligned}\mathrm{E}\left(\widetilde{[X]}_t\right) &= \mathrm{E}\left(\sum_{i=1}^{n} |X_{t_i} - X_{t_{i-1}}|^2\right) \\ &= \sum_{i=1}^{n} \mathrm{E}\left(|X_{t_i} - X_{t_{i-1}}|^2\right) \\ &= \sum_{i=1}^{n} \mathrm{E}\left(|\int_{t_{i-1}}^{t_i} \sigma(s)dW_s + \sum_{j=N_{t_{i-1}}}^{N_{t_i}} Z_j|^2\right) \\ &= \sum_{i=1}^{n}\left(\mathrm{E}\left(\int_{t_{i-1}}^{t_i} \sigma(s)^2 ds\right) + \mathrm{E}\left(\sum_{j=N_{t_{i-1}}}^{N_{t_i}} Z_j^2\right) + 2\,\mathrm{E}\left(\int_{t_{i-1}}^{t_i} \sigma(s)dW_s \sum_{j=N_{t_{i-1}}}^{N_{t_i}} Z_j\right)\right) \\ &= \sum_{i=1}^{n}\left(\int_{t_{i-1}}^{t_i} \mathrm{E}\left(\sigma(s)^2\right) ds + \mathrm{E}\left(N_{t_{i-1}} - N_{t_i}\right) \mathrm{E}\left(Z_1^2\right)\right) \\ &= \int_0^t \mathrm{E}\left(\sigma(s)^2\right) ds + \lambda t \sigma_Z^2.\end{aligned}$$

Thus if we state the expected value for the two cases we find,

- for the case when there are no jumps we have,

$$\mathrm{E}\left(\widetilde{[X]}_t\right) = \int_0^t \mathrm{E}\left(\sigma(s)^2\right) ds;$$

- for the case when there are jumps we have,

$$\mathrm{E}\left(\widetilde{[X]}_t\right) = \int_0^t \mathrm{E}\left(\sigma(s)^2\right) ds + \lambda t \sigma_Z^2. \tag{2.12}$$

We can now try to use the realised variance to define a measure for the instantaneous variance. We can use the form given in equation (2.11) and let,

$$\begin{aligned} \frac{1}{\Delta t}\tilde{\sigma}_{t_i}^2 &= \frac{1}{\Delta t(K-2)}\widetilde{[X]}_{t_{i-K+2},t_i} \\ &\xrightarrow{P} \sigma \end{aligned} \tag{2.13}$$

We will first assess this measure to ensure that it is unbiased.

Proposition 2.16. *The measure based on the realised variance*

$$\widetilde{\sigma_{t_i}^2} = \frac{1}{K-2}\widetilde{[X]}_{t_{i-K+2},t_i},$$

*is an unbiased measure of the instantaneous variance if no jumps occur.*

Proof. Observe that

$$\begin{aligned} \mathrm{E}\left(\frac{\widetilde{\sigma_{t_i}^2}}{\Delta t}\right) &= \mathrm{E}\left(\frac{1}{\Delta t(K-2)}\widetilde{[X]}_{t_{i-K+2},t_i}\right) \\ &= \frac{1}{\Delta t(K-2)}\int_{t_{i-K+2}}^{t_i} \sigma^2(s) ds. \end{aligned}$$

If we work with the case when $\sigma$ is a constant and when each $\Delta t_i$ is constant we can further simplify the above expression as follows:

$$\begin{aligned} \frac{1}{\Delta t(K-2)}\int_{t_{i-K+2}}^{t_i} \sigma^2 ds &= \frac{1}{\Delta t(K-2)}\sigma^2(t_i - t_{i-K+2}) \\ &= \frac{\sigma^2 \Delta t(K-2)}{\Delta t(K-2)} \\ &= \sigma^2. \end{aligned}$$

To see that this will not be the case when jumps are involved we note the result in equation (2.12). □

Additionally, we will need this to be a consistent measure for the instantaneous variance, thus for that we check the limit in probability as $\|\Delta_n\| \to 0$ where $\Delta_n = \max_{1\le j\le i}(t_j - t_{j-1})$. Therefore we have the following results:

- for the case where there are no jumps,

$$\begin{aligned} \widetilde{\sigma_{t_i}^2} &= \frac{1}{K-2}\widetilde{[X]}_{t_{i-K+2},t_i} \\ &= \frac{1}{K-2}\sum_{j=1}^{i} |X_{t_j} - X_{t_{j-1}}|^2 - \sum_{j=1}^{i-K+2} |X_{t_j} - X_{t_{j-1}}|^2 \\ &\xrightarrow{P} \int_0^{t_i} \sigma(s)^2 ds - \int_0^{t_{i-K+2}} \sigma(s)^2 ds \\ &= \frac{1}{K-2}\int_{t_{i-K+2}}^{t_i} \sigma(s)^2 ds; \end{aligned}$$

- for the case where there are jumps,

$$
\begin{aligned}
\widetilde{\sigma^2_{t_i}} &= \frac{1}{K-2}\widetilde{[X]}_{t_{i-K+2},t_i} \\
&\xrightarrow{P} \frac{1}{K-2}\int_{t_{i-K+2}}^{t_i} \sigma(s)^2 ds + \frac{t_i - t_{i-K+2}}{K-2}\lambda\sigma_Z^2.
\end{aligned}
$$

For certain special cases when $\sigma$ is stationary we could simplify the above expressions. However this is unnecessary since it is already clear from the above that we cannot estimate instantaneous variance in this way since in the case where there are jumps, the discontinuous jump component will not be zero. For this reason we will turn our attention to the bipower variation.

When estimating the bipower variation for a stochastic process we use what is known as the realised bipower variation.

Definition 2.17. The realised bipower variation is given by

$$\widetilde{[X]}^*_t = \sum_{i=1}^{n-1} |X_{t_i} - X_{t_{i-1}}||X_{t_{i+1}} - X_{t_i}|.$$

As with the realised variance we can restate the realised bipower variation in terms of a window size $K$ in order to help estimate the instantaneous variance. This yields

$$
\begin{aligned}
\widetilde{[X]}^*_{t_{i-K+2},t_i} &= \widetilde{[X]}^*_{t_i} - \widetilde{[X]}^*_{t_{i-K+2}} \\
&= \sum_{j=1}^{i-1} |X_{t_j} - X_{t_{j-1}}||X_{t_{j+1}} - X_{t_j}| - \sum_{j=1}^{i-K+1} |X_{t_j} - X_{t_{j-1}}||X_{t_{j+1}} - X_{t_j}| \\
&= \sum_{j=i-K+2}^{i-1} |X_{t_j} - X_{t_{j-1}}||X_{t_{j+1}} - X_{t_j}|
\end{aligned}
\tag{2.14}
$$

The above is again a special case, and thus can be extended to a more general form, of realised bipower variation, found in Barndorff-Nielsen and Shephard (2004).

Once again we need to iron out some of the details of what exactly $K$ refers to. As previously stated $K$ refers to the number of data points and not the number of intervals, and so to see what this means for bipower variation we refer to the example in Table 2.1 on page 19. In the example we see that only using $K$ data points means we only have $K-2$ elements in the sum in equation (2.14). Again this is the reason why we use $t_{i-K+2}, t_i$ and why we will use $\frac{1}{K-2}$ as part of equation (2.17). Furthermore we will want the $K$th data value we use to be equal to $t_i$, meaning the first one will need to be $t_{i-K+1}$ and for that we need the summation in equation (2.14) to be from $j = i-K+2$ up to $j = i-1$.

We can see by the definition of bipower variation that the realised bipower variation will converge in probability to the bipower variation as $\|\Delta_n\| \to 0$. We can state this more clearly for the two cases as follows:

- for the case when there are no jumps we have,

$$
\begin{aligned}
\widetilde{[X]}^*_t &\xrightarrow{P} [X]^*_t \\
&= \frac{2}{\pi}\int_0^t \sigma(s)^2 ds;
\end{aligned}
$$

- for the case when there are jumps we have,

$$\begin{aligned} \widetilde{[X]}_t^* &\xrightarrow{P} [X]_t^* \\ &= \frac{2}{\pi}\int_0^t \sigma(s)^2 ds. \end{aligned} \tag{2.15}$$

Since the above two cases are the same we immediately begin to suspect it may prove to be a superior estimate for the required instantaneous variance than the estimate based on realised variance. We move on to assess the expected value of the realised bipower variation, and see that

$$\begin{aligned} \mathrm{E}\left(\widetilde{[X]}_t^*\right) &\xrightarrow{P} \mathrm{E}\left([X]_t^*\right) \\ &= \frac{2}{\pi}\int_0^t \mathrm{E}\left(\sigma(s)^2\right) ds. \end{aligned} \tag{2.16}$$

The above follows by using equation (2.15) and then using Theorem 2.13.

Consequently if we attempt to define a measure of the instantaneous variance we may use the following,

$$\begin{aligned} s_{t_i}^{*2} &= \frac{1}{K-2}\widetilde{[X]}^*_{t_{i-K+2},t_i} \\ &= \frac{1}{K-2}\sum_{j=i-K+2}^{i-1} |X_{t_j} - X_{t_{j-1}}||X_{t_{j+1}} - X_{t_j}|. \end{aligned} \tag{2.17}$$

We now need to make certain this estimate will be consistent where the previous one was not.

THEOREM 2.18. *The following estimator is a consistent estimator for $\sigma$ in the presence of jumps*

$$\frac{\pi}{2\Delta t}s_{t_i}^{*2} \xrightarrow{P} \frac{1}{(K-2)\Delta t}\int_{t_{i-K+2}}^{t_i} \sigma(s)^2 ds \tag{2.18}$$

PROOF. For the following we take the limit in probability as $\|\Delta_n\| \to 0$. Observe that

$$\begin{aligned} \frac{s_{t_i}^{*2}}{\Delta t} &= \frac{1}{\Delta t(K-2)}[X]^*_{t_{i-K+2},t_i} \\ &= \frac{1}{\Delta t(K-2)}\left(\sum_{j=1}^{i-1}|X_{t_j} - X_{t_{j-1}}||X_{t_{j+1}} - X_{t_j}| - \sum_{j=1}^{i-K+1}|X_{t_j} - X_{t_{j-1}}||X_{t_{j+1}} - X_{t_j}|\right) \\ &\xrightarrow{P} \frac{2}{\Delta t(K-2)\pi}\left(\int_{t_0}^{t_i}\sigma^2(s)ds - \int_{t_0}^{t_{i-K+2}}\sigma^2(s)ds\right) \\ &= \frac{2}{\Delta t(K-2)\pi}\int_{t_{i-K+2}}^{t_i}\sigma^2(s)ds. \end{aligned}$$

From this we can already see that only the continuous component will affect this estimate and thus it will be consistent given the appropriate constant coefficient. However we use the simple case where $\sigma$ is constant to show this more precisely,

$$
\begin{aligned}
\frac{s_{t_i}^{*2}}{\Delta t} &\xrightarrow{P} \frac{2}{\Delta t(K-2)\pi}\int_{t_{i-K+2}}^{t_i}\sigma^2 ds \\
&= \frac{2}{\Delta t(K-2)\pi}\sigma^2(t_i - t_{i-K+2}) \\
&= \frac{2}{\Delta t(K-2)\pi}\sigma^2\Delta t(K-2) \\
&= \frac{2}{\pi}\sigma^2,
\end{aligned}
$$

□

The coefficient $\frac{2}{\pi}$ above will be used as part of the variance of the test that we will define in the next section. Note that $\frac{\pi}{2} \times s_{t_i}^{*2}$ as in equation (2.17) will converge to the same value as $\widetilde{\sigma}_{t_i}^2$ as in equation (2.13) when there are no jumps. Having found a satisfactory measure for instantaneous variance we end this section with a table summarising the theoretical and estimated values above given in Table 2.2.

TABLE 2.2. Summary of estimated vs theoretical parameters

| **Theoretical value** | **Estimation** |
|---|---|
| **Quadratic variation:** $[X]_t = p\lim_{M\to\infty}\sum_{i=1}^{M}[X_{t_j} - X_{t_{j-1}}]^2$ | **Realised variance:** $\widetilde{[X]}_t = \sum_{i=1}^{n}\left(X_{t_i} - X_{t_{i-1}}\right)^2$ <br> $\widetilde{\sigma_{t_i}^2} = \frac{1}{K-2}\widetilde{[X]}_{t_{i-K+2},t_i}$ |
| **Bipower Variation:** $[X]_t^* = p\lim_{\delta\to 0}\sum_{j=1}^{\lfloor\frac{t}{\delta}\rfloor-1}\lvert X_{t_j} - X_{t_{j-1}}\rvert\lvert X_{t_{j+1}} - X_{t_j}\rvert$ | **Realised bipower Variation:** $\widetilde{[X]}_t^* = \sum_{i=1}^{n-1}\lvert X_{t_i} - X_{t_{i-1}}\rvert\lvert X_{t_{i+1}} - X_{t_i}\rvert$ <br> ${s^{*2}}_{t_i} = \frac{1}{K-2}\widetilde{[X]}^*_{t_{i-K+2},t_i}$ |

In the following section we discuss more specifically how we might use the above estimates to test for jumps in various ways.

## 2.3. Tests for jumps

**2.3.1. Tests based on bipower variation.** As can be seen in the previous sections for a jump process $Y_t$ the quadratic variation, $[Y]_t = \sum_{i=1}^{N_t} Z_i^2$, does not reduce to zero, but the bipower variation, $[Y]_t^* \to 0$, reduces to 0. In contrast to this, for the stochastic volatility process without jumps we find that for quadratic variation and bipower variation that $[Y]_t^* = \frac{2}{\pi}[Y]_t$. We can thus use the difference between the quadratic variation and bipower variation to test for jumps.

Barndorff-Nielsen and Shephard (2006) used these results regarding the quadratic variation and bipower variation in two ways; in terms of the linear difference between the quadratic variation and bipower variation or in terms of the ratio of the quadratic variation and bipower variation. They define test statistics in terms of the difference as,

$$G = \frac{\delta^{-\frac{1}{2}}(\mu_1^{-2}[Y_\delta]_t^* - [Y_\delta]_t)}{\sqrt{\int_0^t (\frac{\pi^2}{4} + \pi - 5)\sigma(u)^4 du}} \to N(0,1)$$

where $\delta$ is the length of the time intervals used, $Y$ is a stochastic volatility plus drift model and $\mu_1 = \sqrt{\frac{2}{\pi}}$. The second test statistic in terms of the ratio is defined by

$$H = \frac{\delta^{-\frac{1}{2}}\left(\frac{\mu_1^{-2}[Y_\delta]_t^*}{[Y_\delta]_t} - 1\right)}{\sqrt{\left(\frac{\pi^2}{4} + \pi - 5\right)\frac{\int_0^t \sigma(s)^4 ds}{\left(\int_0^t \sigma(s)^2 ds\right)^2}}} \to N(0,1)$$

as $\delta \to 0$. Corsi et al (2010) built on the above idea by introducing threshold bipower estimation, that is combining the ideas of bipower variation estimation and threshold estimation.

Baum and Zerilli (2016) in their research into crude oil simply used the realized variance and realized bipower variation as estimates for the quadratic variation and integrated volatility and used the standard Tukey formula to identify outliers. Fičura and Witzany (2015) similarly used the realized variance and bipower variation to estimate the jump component, but there is often noise in the data and thus only the significant jumps should be considered using this method.

Huang and Tauchen (2005) rather than just testing for jumps proposed a measure for how much a jump has affected returns, the relative jump measure, namely

$$RJ_{t_{i-K+1},t_i} = \frac{\widetilde{[X]}_{t_{i-K+1},t_i} - \frac{2K}{\pi(K-1)}\widetilde{[X]}^*_{t_{i-K+1},t_i}}{\widetilde{[X]}_{t_{i-K+1},t_i}}.$$

Even while using these tests above, jumps are inherently difficult to identify since only discrete data will be available. Hence it is usually necessary to take the data at as high a frequency as possible; usually 10 or 15 minute interval data since any more than this and the data will be prone to noise.

We can additionally use the bipower variation to define various estimators for the volatility, as we have done at the end of the previous section. One such estimator based on the bipower variation which will be useful to us later is $s^{*2}$ as defined in equation (2.17).

**2.3.2. A test for individual jumps.** The tests discussed above are all very useful in testing for jumps, but in general the difference between quadratic and bipower variation, and by extension the tests that are based thereupon, usually identify whether or not jumps have occurred within a period of time. It is therefore more difficult to use the above tests to identify individual jumps. For this we turn to the test of Lee and Mykland (2008). This test is shown to outperform the test by Barndorff-Nielsen and Shephard (2006) in terms of size and power.

Before we move on to discuss the test as it is in Lee and Mykland's (2008) original derivation we may first want to get a feel for the logic and intuition behind the test. We start by looking at a simplified test statistic T,

$$T = \frac{\log S_{t_i} - \log S_{t_{i-1}}}{\sqrt{\Delta t}\sigma},$$

where we start with the case where $\sigma$ is constant and known. In this rather unrealistic case when $\sigma$ is a known constant we have that the null distribution for this test statistic T will be the standard normal distribution, under the null hypothesis of no jumps.

Using the estimator defined by equation (2.17) we can generalize the test statistic $T$.

THEOREM 2.19. *The test statistic $T^*$,*

$$T^* = \frac{\log S_{t_i} - \log S_{t_{i-1}}}{\sqrt{\Delta t} s^*_{t_{i-1}}},$$

*converges in distribution to the standard normal distribution, $N(0,1)$ under the null hypothesis of no jumps.*

PROOF. Using the simple case when $\sigma$ is constant we find that

$$\begin{aligned} T^* &= \frac{\log S_{t_i} - \log S_{t_{i-1}}}{s^*_{t_i}} \\ &= \frac{\log S_{t_i} - \log S_{t_{i-1}}}{\sqrt{\Delta t}\sigma} \frac{\sqrt{\Delta t}\sigma}{s^*_{t_i}} \\ &= T\frac{\sqrt{\Delta t}\sigma}{s^*_{t_i}}. \end{aligned}$$

We also have

$$\frac{\pi}{2}\frac{1}{\Delta t} s^{*2}_{t_i} \xrightarrow{P} \sigma^2$$

from equation (2.18) on page 22 and as such by the continuous mapping theorem we can say

$$\sqrt{\frac{\pi}{2}}\frac{1}{\sqrt{\Delta t}} s^*_{t_i} \xrightarrow{P} \sigma.$$

Therefore, $\frac{\sqrt{\Delta t}\sigma}{s^*_{t_i}} \xrightarrow{P} \sqrt{\frac{\pi}{2}}$ and since $T \to N(0,1)$, we have according to Slutsky's theorem that

$$T^* \xrightarrow{d} N\left(0, \frac{\pi}{2}\right).$$

□

Lee and Mykland (2008) do not use the test statistic $T^*$ defined above, but this simplification helps to understand the intuition and logic behind how the test works. We will now discuss Lee and Mykland's form of the test in more detail.

A potentially important feature of Lee and Mykland's (2008) test is that it makes no assumption as to whether or not jumps have occured before or after the time being tested. The logic behind this test is to estimate what Lee and Mykland refer to as "instantaneous volatility" and compare that to the asset return, or in our case log return of the exchange rate. This instananeous volatility is commonly calculated using the quadratic variation, that is the sum of squared returns or log returns $p\lim_{n\to\infty}\sum_{i=2}^{n}(\log S_{t_i} - \log S_{t_{i-1}})^2$.

The quadratic variation, however, as previously discussed, is not robust with respect to jumps and as such the bipower variation will need to be used, that is the product of consecutive absolute log-returns,

$$p\lim_{n\to\infty}\sum_{i=2}^{n-1}|\log S_{t_i} - \log S_{t_{i-1}}||\log S_{t_{i+1}} - \log S_{t_i}|.$$

The reason for this is that as previously discussed and proven by Barndorff-Nielsen and Shephard (2004) the bipower variation estimator is consistent in the presence of jumps.

For the test by Lee and Mykland (2008) we use a fixed value for $\Delta t = \frac{T}{n}$, where $T$ is the time horizon and $n$ is the number of observations that occur within $[0, T]$. We assess a movement of the process within a window size of $K$, that is we use the previous $K$ returns when testing a given observation for jumps. The instantaneous volatility is then estimated using these previous $K$ observations. Finally the test statistic, denoted $\mathcal{L}$, is defined in terms of this estimation of bipower variation and the next return. Written more formally the test statistic $\mathcal{L}$ is defined as follows:

Definition 2.20 (Lee and Mykland (2008) page 8). The test statistic $\mathcal{L}_i$ used to test at time $t_i$ if a jump occurred from $t_{i-1}$ to $t_i$ is defined as

$$\mathcal{L}_i \equiv \frac{\log S(t_i) - \log S(t_{i-1})}{s^*_{t_i}} \sim N\left(0, \frac{\pi}{2}\right),$$

under the null hypothesis, where we have

$$s^{*2}_{t_i} = \frac{1}{K-2} \sum_{j=i-K+1}^{i-2} |\log S_{t_i} - \log S_{t_{i-1}}||\log S_{t_{i+1}} - \log S_{t_i}|.$$

Since the test looks at high-frequency data, the effect of the drift is negligible. Note that in this definition the indices in $s^{*2}$ are different from our original definition in the previous section.

The window size $K$ is chosen in such a way as to eliminate the effects of jumps on the volatility estimation, that is it must be large enough so that the effect of jumps disappears but must clearly be smaller than the total number of observations $n$. It is clear that the choice of frequency $\Delta t$ will affect the necessary choice of $K$. A rule of thumb for choosing the window size $K$ is that the choice of $K$ should be between $\sqrt{252 \times n_{day}}$ and $252 \times n_{day}$ where $n_{day}$ refers to the number of observations in a day. Since higher values for $K$ within range do not add much to the accuracy of results, it is better to use smaller values for $K$ that are within the range.

When using the above test statistic for the null hypothesis of there being no jumps when selecting a rejection region, we need to bear the following in mind:

$$\frac{\max|\mathcal{L}_i| - C_n}{S_n} \to \xi \tag{2.19}$$

as $\Delta t \to 0$ when there are no jumps and $\xi$ has a cumulative distribution function $P(\xi \le x) = e^{-e^{-x}}$, where

$$C_n = \frac{\sqrt{2\log n}}{\sqrt{\frac{2}{\pi}}} - \frac{\log \pi + \log\log n}{2\sqrt{\frac{2}{\pi}(2\log n)}},$$

$$S_n = \frac{1}{\sqrt{\frac{2}{\pi}(2\log n)}},$$

with $n$ the number of observations. Thus this value $n$ will usually be the same as $K$. Now we look at choosing a rejection region, or put more simply choosing a significance level $\alpha$. For $\frac{|\mathcal{L}_i| - C_n}{S_n}$ let the threshold be $\beta^*$. Then since $P(\xi > \beta^*) = e^{-e^{-\beta^*}} = \alpha$ we can say the critical value will then be $-\log(-\log(1-\alpha))$. So for $\alpha = 5\%$ we will have $CV = -\log(-\log(0.95)) = 2.9702$ and so if $\frac{|\mathcal{L}_i| - C_n}{S_n} > 2.9702$, we can reject the null hypothesis that there are no jumps.

Note that in equation (2.19) above we use the standard Gumbel distribution to choose a rejection region. One may begin to wonder why this is the case and why we have used $\max \mathcal{L}_i$ rather than simply $\mathcal{L}_i$, for that we note the following,

$$\begin{aligned} T_{i,n}^{(1)} &= \frac{|\mathcal{L}_i| - C_n}{S_n} \\ &\leq \frac{\max|\mathcal{L}_i| - C_n}{S_n} \\ &\to \xi. \end{aligned}$$

If we use $n = K = 20$ we find $S_n = 0.5120$ and $C_n = 2.4938$. As a result using the Gumbel distribution will have an effect on the rejection region since we will expect the 95th percentile of the Gumbel distribution to be greater than the 95th percentile of $T_{i,n}^{(1)}$ as we have defined it here.

We can calculate the theoretical expected value of $T_{i,n}^{(1)}$, since we know that $|\mathcal{L}_i|$ is half normally distributed, thus

$$\begin{aligned} \mathrm{E}(T_{i,n}^{(1)}) &= \frac{\mathrm{E}(|\mathcal{L}_i|) - C_n}{S_n} \\ &= \frac{1 - C_n}{S_n} \\ &= -2.9176. \end{aligned}$$

We find the theoretical expected value of the standard Gumbel distribution to be approximately $\mathrm{E}(\xi) = -0.5772$. These results give us an indication that the two distribution will drastically differ.

However, since we are more interested in the 95th percentiles of the two distributions we can use Monte Carlo simulation to approximate the 95th percentile of $T_{i,n}^{(1)}$ as follows:

- begin by simulating a value $Z_i \sim N(0, 1)$;
- using $Z_i$ we calculate $|\mathcal{L}_i|$ using
$$|\mathcal{L}_i| = \sqrt{\frac{\pi}{2}}|Z_i|;$$
- determine $T_{i,n}^{(1)}$ using
$$T_{i,n}^{(1)} = \frac{|\mathcal{L}_i| - C_n}{S_n},$$
where $C_n = 2.4938$ and $S_n = 0.5120$;
- repeat the above $MC = 100\,000$ times to obtain $T_{i,n}^{(1)}$ for $i = 1, 2, 3, ..., MC$;
- calculate the order statistics $T_{(i),n}^{(1)}$, where $T_{(1),n}^{(1)} < T_{(2),n}^{(1)} < ... < T_{(MC),n}^{(1)}$;
- the 95th percentile is then given by $T_{(0.95MC),n}^{(1)}$.

The result obtained this way is that the critical value for $\alpha = 5\%$ is $CV = -0.0538$. Since this critical value is lower than the original value, $CV = 2.9702$, it is reasonable to say that using $T_{i,n}^{(1)}$ would increase the likelihood of spurious detection of jumps. Thus the test proposed by Lee and Mykland (2008) is a conservative test. The question now arises of which method of choosing a rejection region better serves our purposes. For this we first need to discuss the likelihood of misclassification of jumps in Lee and Mykland's (2008) original test.

In their development of the above test, Lee and Mykland (2008), noted that the likelihood of misclassification becomes negligible when using high-frequency data. For our purposes, however,

we will be using this test with daily data which may impact this likelihood of misclassification.
We assess the sensitivity and specificity of the test by assessing various types of misclassification errors. According to Lee and Mykland (2008), there are four types of misclassification errors, which are failure to detect an individual jump, global failure to detect jumps, spurious detection of an individual jump and global spurious detection of jumps. The failure to detect an individual jump is when a jump occurs within the time interval $(t_{i-1}; t_i]$ and the test fails to detect it. Spurious detection of an individual jump is when a jump does not occur within the time interval $(t_{i-1}; t_i]$ but the test detects a jump. The global variations of the types of misclassification are simply, the failure to detect any jump in the whole time horizon $[0; T]$, or similarly the spurious detection of one or more jumps in the whole time horizon.
In their development of this test Lee and Mykland (2008) performed a Monte Carlo study for which they have results on the probability of spurious detections and failure to detect jumps, for individual jumps. These results will be very useful for us. Table 2.3 summarises the probability of the various errors using the original critical value of $-log(-log(\alpha))$.

TABLE 2.3. Summary of the probability of various errors for stochastic volatility

| **Frequency:** | 24 **hours** | 15 **minutes** |
|---|---|---|
| **Probability of individual spurious detections:** | 3.9110e−03 | 8.9449e−05 |
| **Probability of failure to detect individual jumps of size:** | | |
| $3\sigma$ | 0.053 | 0.003 |
| $2\sigma$ | 0.067 | 0.001 |
| $\sigma$ | 0.146 | 0.002 |
| $0.5\sigma$ | 0.428 | 0.008 |
| $0.25\sigma$ | 0.750 | 0.039 |
| $0.1\sigma$ | 0.968 | 0.190 |

Source: Lee and Mykland (2008)

As seen in Table 2.3 the probability of spurious detection of jumps is low, even for 24 hour data.
The results for the failure to detect jumps depend on the size of the jump. These results are for the stochastic volatility model case. For jumps three times the size of the volatility the probability of detection at a frequency of 24 hours is 0.9470 and for smaller jumps this value decreases. For jumps twice the volatility it is 0.9330, while for jumps the same size as the volatility 0.8540 and for jumps half the size of the volatility it is 0.5720. This value gets significantly smaller for smaller jumps. The probabilities of detecting jumps increases as the frequency increases.
Bearing these probabilities in mind we see that as far as spurious detection is concerned using a frequency of 24 hours is good enough. For a failure to detect jumps it is simply not feasible to use a frequency of 24 hours to try to detect jumps about half the size of the volatility or smaller. Our goal is however not to detect small jumps, since we will be focusing on extreme events in the data.

Thus we can with reasonable certainty say that for our purposes using daily data is good enough. Similarly we can also conclude that it would not be better to use the standard normal distribution to choose a rejection region. That is we will be using Lee and Mykland's (2008) method of choosing a rejection region, the original critical value of $-log(-log(\alpha))$.

**2.3.3. Other tests.** The previous two tests discussed, the one by Barndorff-Nielsen and Shephard (2004) and the one by Lee and Mykland (2008), are by no means the only tests in the literature. The purpose of this section is to give a brief overview and comparison of various other tests in the literature.

The first test we will be discussing is that of Jiang and Oomen (2008). They proposed a new test based on what they refer to as swap variance. Various types of tests are built upon this newly developed swap variance, that is a difference test, a logarithm test and a ratio test. The logic behind this type of test is similar to the bipower variation test except that bipower variation is robust to the presence of jumps and swap variance is sensitive to the presence of jumps. In calculating the test statistics defined by Jiang and Oomen (2008) the bipower variation is used as an estimate for the variance and as such this test is based on bipower variation. This test performs better with a larger sample size and thus it is more difficult to test for individual jumps. Rather the whole sample is tested for jumps. This test performs well using high-frequency data as it can be modified to correct for market microstructure noise. This is not a requirement however and the test can likely be used with daily data.

The test suggested by Corsi et al (2010) is one based on the bipower variation. This test is built by using threshold bipower (or multipower) variation, rather than simply using bipower (or multipower) variation. The test is used to examine a fixed time period for jumps and as such it is in all likelihood not useful in detecting individual jumps within the same time period, such as on the same day. This test can be used with daily data with a sufficiently large time window, such as a month.

The test by Bajgrowicz and Scaillet (2011) does not actually test for jumps. It is however worth mentioning as it helps to test for spurious detections of jumps. This test is used on daily test statistics and uses a thresholding technique. The test is built around thresholding daily test statistics, and as such this test is not an individual test but rather tests for any spurious detections within each day. This also means that this test more than likely cannot be used in its current form on daily data and needs to be altered in order to be used on daily data. Using the proposed thresholding technique it is seen that testing for jumps produces significantly less spurious detection of jumps. However a drawback is that for smaller jumps there are also significantly fewer true jumps detected. This technique can therefore be used to test for spurious detections among extreme events.

The next test we will be discussing is that of Laurent, Lecourt and Palm (2014). This test is similar to that of Lee and Mykland (2008). This test uses a test statistic defined in terms of robust estimates for the mean and variance, in other words this test is not based on bipower variation. This test is not significantly weaker when many jumps are present and as such it can be useful in not just testing whether there are jumps in the whole sample but also in testing for specific jumps, that is it is an individual jump test. In the example given, Laurent, Lecourt and Palm (2014) used daily exchange rate data, and thus the test clearly works for daily data. As can be expected, the proportion of correct jump detections increases as the size of the jumps increase.

Having given a brief overview of these tests we can see that each has its strengths and weaknesses. Now in the following chapter we will look at an application of the test by Lee and Mykland (2008) to test for jumps in the Rand to US Dollar exchange rate.

CHAPTER 3

# Application: Rand-Dollar exchange rate

## 3.1. Data

We will be assessing the data for the Rand-Dollar exchange rate for the period from January 2012 to March 2020, since this period is between two recent major financial crises. Figure 3.1 contains the raw data and Figure 3.2 contains the daily log returns of the Rand-Dollar exchange rate.

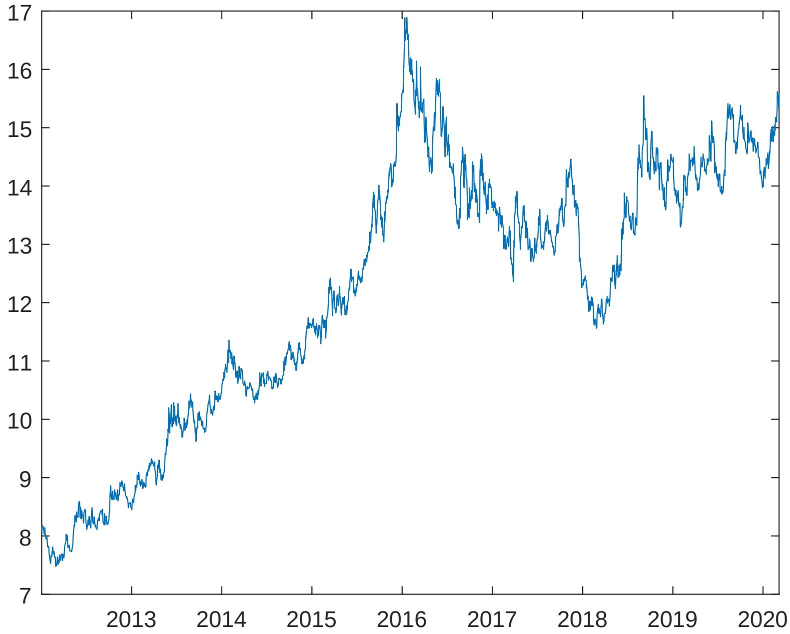


FIGURE 3.1. Data for the Rand-Dollar exchange rate from January 2012 to December 2020

Source: Data obtained from the South African Reserve Bank

Upon inspection of the data we have reason to suspect that there may be a break in the data which would be noteworthy to investigate. We will specifically take a closer look at January 2013 to December 2017. It seems that the log returns for the period after the 9th of December 2015 may have a higher variance than those before. So now we look at some more detailed statistics and basic tests, using the skewness and kurtosis, with the null hypothesis being that the data is normally distributed before and after 9th December 2015.

The test we will use in Table 3.1 is the Jarque-Bera (JB) test found on page 10 of Tsay (2010). This test involves two other test statistics which can each be used to test for the normal distribution in their own right.

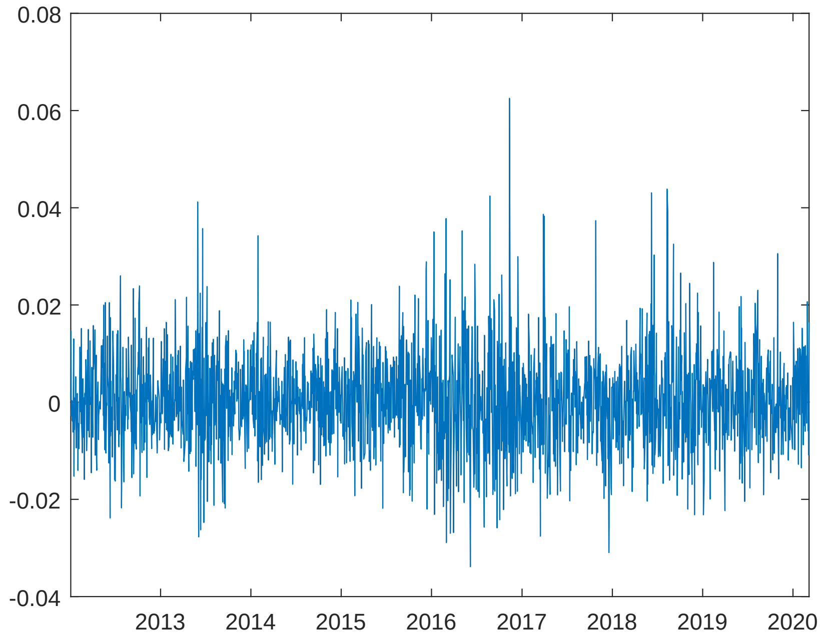


FIGURE 3.2. Log returns for the exchange rate data
Source: Own construction from exchange rate data

The first test which we will refer to as $t_1$ uses the skewness of the data to calculate a test statistic

$$t_1 = \frac{\hat{S}}{\sqrt{6/n}},$$

where $\hat{S}$ is the skewness of the sample data and $n$ is the number of observations within the sample. This test statistic has a standard normal distribution, under the null hypothesis of no skewness and normaility.

The second test statistic which we will refer to as $t_2$ uses the excess kurtosis of the sample data to calculate the test statistic

$$t_2 = \frac{\hat{K} - 3}{\sqrt{24/n}},$$

where $\hat{K}$ refers to the kurtosis and as such we subtract 3 in order to calculate the excess kurtosis and $n$ is again the number of observations in the sample. This test statistic has a standard normal distribution as well, under the null hypothesis of no excess kurtosis and normaility.

Finally the combination of these two test statistics is the Jarque-Bera test,

$$\begin{aligned} JB &= (t_1^2 + t_2^2) \\ &= \frac{\hat{S}^2}{6/n} + \frac{(\hat{K} - 3)^2}{24/n}. \end{aligned}$$

This test statistic has a chi-squared distribution with two degrees of freedom, under the null hypothesis of normality.

We begin our assessment of Table 3.1 by assessing the data before the 9th of December 2015. When using the $t_1$ test statistic we find that for a significance level of 1% we would get the result

TABLE 3.1. Statistics showing a possible break in the data on the 9th of December 2015

| **Log Returns** | **Before** | **After** | **Entire Set** |
|---|---|---|---|
| **Number of observations ($n$):** | 735 | 513 | 1248 |
| **Mean:** | 7.3261e-04 | -3.2654e-04 | 2.9724e-04 |
| **Skewness ($\hat{S}$):** | 0.1789 | 0.5950 | 0.4110 |
| $t_1 = \frac{\hat{S}}{\sqrt{6/n}}$: | 1.98 | 5.50 | 5.93 |
| **p-value for $\|t_1\|$** | 0.0477 | 3.7979e-07 | 3.0293-09 |
| **Kurtosis ($\hat{K}$):** | 4.0162 | 5.1101 | 5.1656 |
| $t_2 = \frac{\hat{K}-3}{\sqrt{24/n}}$: | 5.62 | 9.76 | 15.62 |
| **p-value for $\|t_2\|$ :** | 1.9096e-08 | Less than 2.2e-16 | Less than 2.2e-16 |
| **JB test ($t_1^2 + t_2^2$):** | 497.90 | 588.44 | 1422.67 |
| **p-value for JB-test:** | Less than 2.2e-16 | Less than 2.2e-16 | Less than 2.2e-16 |
| **Standard deviation:** | 0.0087 | 0.0118 | 0.0101 |
| **Max:** | 0.0412 | 0.0625 | 0.0625 |
| **99th percentile:** | 0.0216 | 0.0375 | 0.0289 |
| **95th percentile:** | 0.0141 | 0.0177 | 0.0157 |
| **90th percentile:** | 0.0114 | 0.0133 | 0.0121 |
| **10th percentile:** | -0.0099 | -0.0146 | -0.0118 |
| **5th percentile:** | -0.0125 | -0.0190 | -0.0161 |
| **1st percentile:** | -0.0206 | -0.0269 | -0.0242 |
| **Min:** | -0.0277 | -0.0338 | -0.0338 |

that the data is symmetric. The $t_2$ test statistic on the other hand is extremely significant meaning the data is unlikely to have no excess kurtosis. When combining these two tests in the JB-test we find that the p-value is extremely small making it unlikely that the data is normally distributed. Bearing in mind that the kurtosis test gave us much more significant results we may conclude that these heavier tails may be caused by jumps.

We move on to assess the data after the 9th of December 2015. In this case we find that in all three test statistics, that is in $t_1$, $t_2$ and the JB-test, the p-values are small enough that the data is skew, the kurtosis differs from the normal distribution and therefore the data is not normally distributed. We can very confidently reject the null hypotheses of all three tests. It is noteworthy that the test using skewness had much less significant results for the period before when compared to after, we may conclude then that the events of the 9th of December 2015 may have had a great effect on the data's skewness.

Finally we assess the entire data set including both before and after. We find that once again all three of the test statistics produce such minuscule p-values that we can again reject the null hypotheses with extreme confidence.

We may begin to wonder now whether these results may be affected by jumps that we have not been taken into account. To counter this we now move forward in the next section to analyse the power and bipower variation so that we can test for jumps.

## 3.2. Analysis of bipower and quadratic variation

We can test the data using the test of Lee and Mykland (2008) which allows us to test the data using a window size $K$ for which we use 20. That is we are testing for jumps by assessing the previous 20 observations. Before we test for jumps we first need to look at the quadratic variation and bipower variation, since these will be our indicators of whether there is a jump or not.

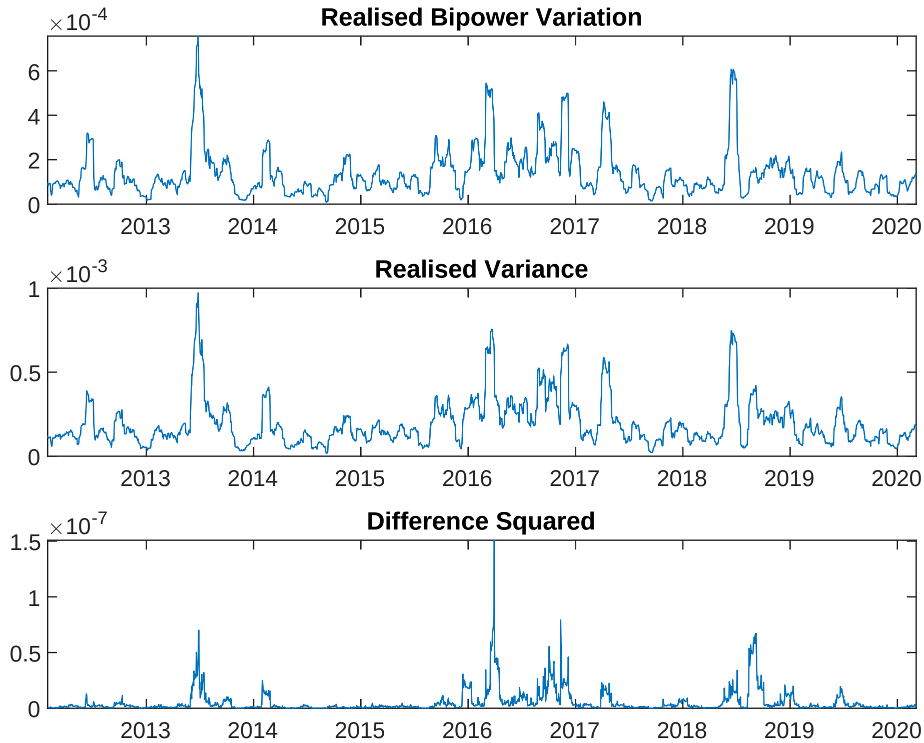


FIGURE 3.3. Bipower variation and realized variance and the difference of these two values squared

Source: Own construction from exchange rate data

We see in Figure 3.3 the difference squared has distinct spikes in value on certain dates and it is these dates that we are interested in.

## 3.3. Detection of jumps

Starting with a confidence level of 90% and a window size of $K = 20$ we finally test for jumps and obtain results as in Figure 3.4. The exact dates of the detected jumps are 31 May 2013, 30

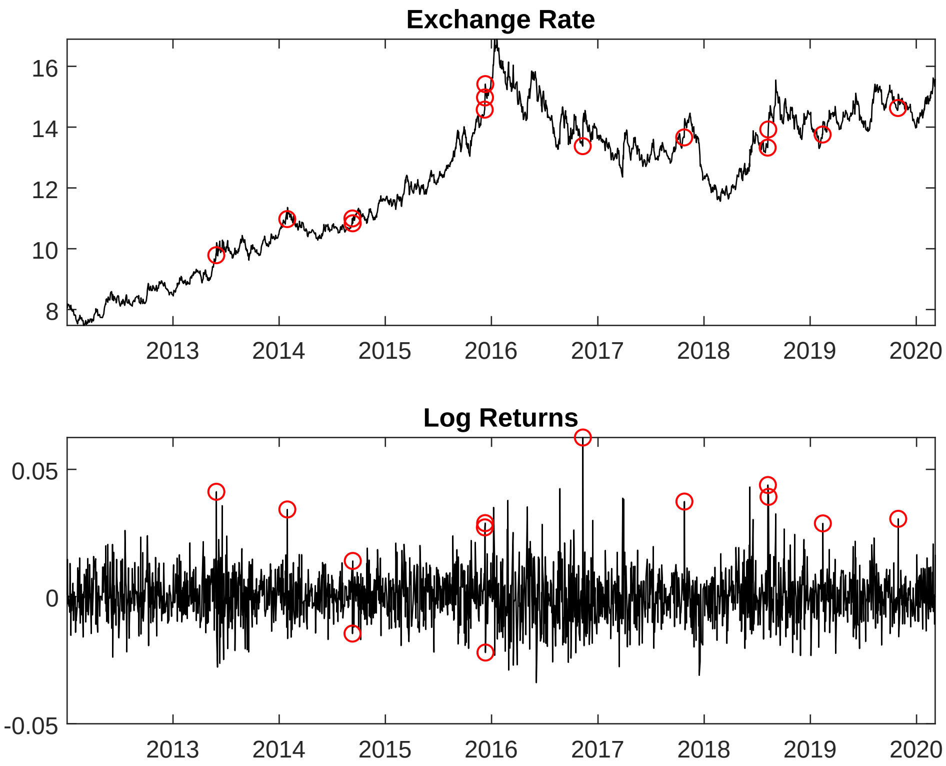


FIGURE 3.4. Exchange rate data with jumps for a confidence level of 90%
Source: Own construction from exchange rate data

January 2014, 11 September 2014, 12 September 2014, 10 December 2015, 11 December 2015, 14 December 2015, 11 November 2016, 26 October 2017, 10 August 2018, 13 August 2018, 14 February 2019 and 31 October 2019. In Figure 3.5 we test for jumps using a confidence level of 95%.

Now having moved on to a higher confidence level we have fewer jumps than before. The exact dates for the detected jumps are 31 May 2013, 11 September 2014, 10 December 2015, 11 December 2015, 14 December 2015, 11 November 2016, 26 October 2017, 10 August 2018, 13 August 2018 and 31 October 2019. Finally in Figure 3.6 we will test the data for jumps at a confidence level of 99%.

We see that for this higher confidence level jumps are only detected on the dates, 10 December 2015, 11 December 2015, 26 October 2017, 10 August 2018 and 13 August 2018. Table 3.2 compiles the above jump dates and possible reasons for the jump. The significance level is indicated by $*$, that is 10% is denoted by $*$, 5% by $**$ and 1% by $***$, where we use $S_n = 0.5120$ and $C_n = 2.4938$.

Table 3.2 uses the price of a US Dollar in rands. Thus a jump up refers to a relative strengthening of the Dollar and weakening of the Rand and a jump down refers to a relative strengthening of the Rand and weakening of the Dollar. Some of the jumps have more obvious economic causes than others. The majority of the jumps however seem to be caused by events within South Africa. It is worth noting that the maximum movement we had in Table 3.1 which corresponds to the jump on 11 November 2016 does not have the highest value of the test statistic $\frac{|\mathcal{L}_i|-C_n}{S_n}$. The reason for this

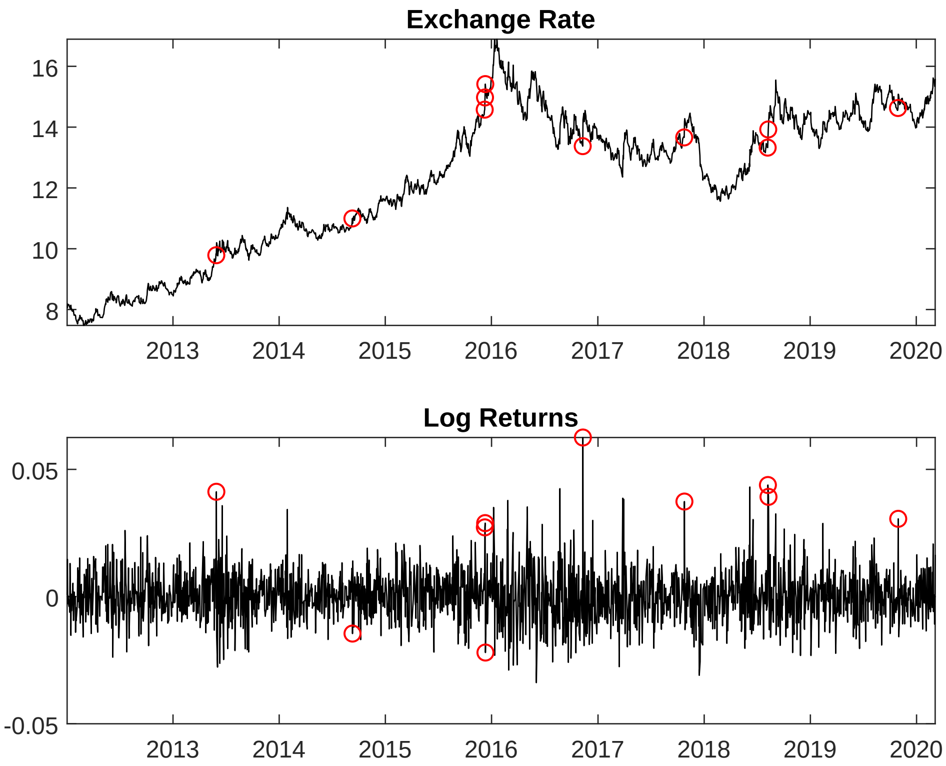


FIGURE 3.5. Exchange rate data with jumps for a confidence level of 95%
Source: Own construction from exchange rate data

is that although the movement on this date was the largest in the data set the movements before this jump caused the test statistic to be lower than one might expect for the largest movement in the data set since the instantaneous variance would have been higher than in other cases.

When assessing Table 3.2 one can immediately see that most of the jumps are upward movements, that is a relative weakening of the Rand and strengthening of the Dollar. To begin to understand the reason behind this we need to assess what types of events most likely caused each jump.

We see that jumps occurred on the 10th, the 11th and 14th of December, all three of these likely stemmed from the same event, that is the removal of the South African finance minister. It is worth noting that the initial jump on the 10th of December was the second largest jump.

In order to make any reasonable conclusions however we need to discuss more than just one event. We find that most of the jumps are caused by political instability, most of which is in South Africa. The next prominent factor is related to interest rates in South Africa and the US. Finally various other economic factors also caused several jumps. Overall one may thus find that regardless of the specific types of factors that caused a jump, that is political instability, interest rates or other economic factors, it is clear that the majority of the jumps identified can be linked to occurrences in South Africa rather than in the US.

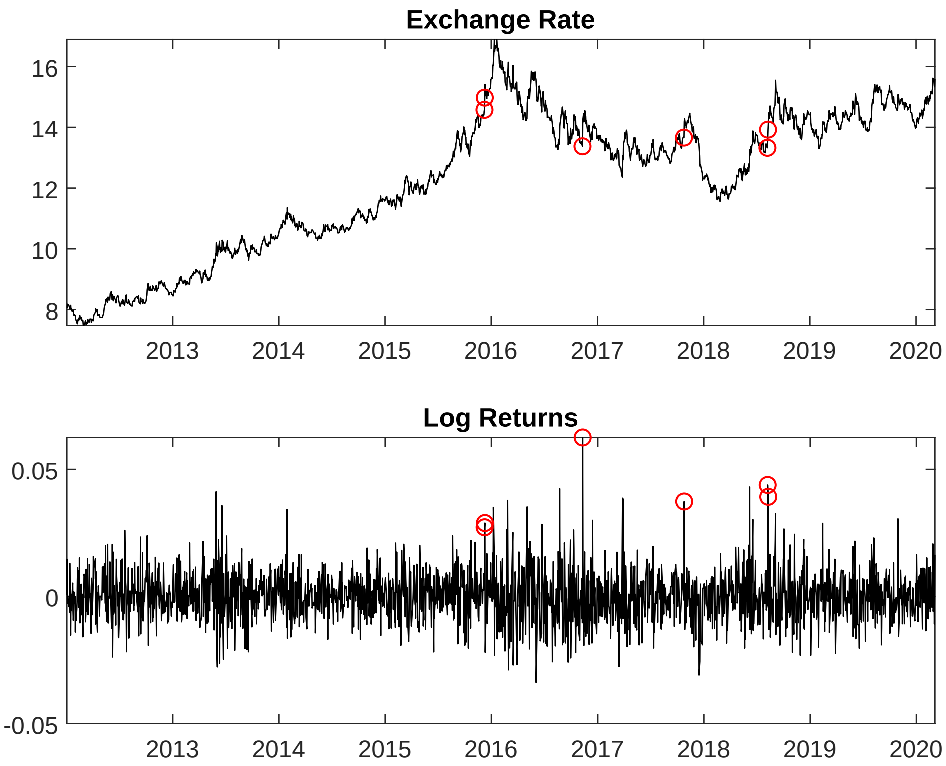


FIGURE 3.6. Exchange rate data with jumps for a confidence level of 99%

Source: Own construction from exchange rate data

Table 3.2. Jump dates and possible causes

| **Jump Date** | **Up or Down** | **P-Values** | **Test statistic** $\frac{\lvert\mathcal{L}_i\rvert - C_n}{S_n}$ | **Possible causes** |
|---|---|---|---|---|
| 31 May 2013 | Up | 0.0454 | $**3.0698$ | Overall economic factors, mining labour issues and a speech by the president of South Africa (SAPA, 2013) |
| 30 January 2014 | Up | 0.0707 | $*2.6128$ | South African surprise interest rate hike (Zacks, 2014) |
| 11 September 2014 | Down | 0.0472 | $**3.0294$ | US interest rate climb (Anon, 2014) |
| 12 September 2014 | Up | 0.0823 | $*2.4541$ | Peugeot choose not to manufacture in South Africa (Cokayne, 2014) |
| 10 December 2015 | Up | 0.0008 | $***7.1220$ | Removal of South African finance minister (Hirsch, 2015) |
| 11 December 2015 | Up | 0.0020 | $***6.2214$ | Address by the new South African finance minister (SA News, 2015) |
| 14 December 2015 | Down | 0.0304 | $**3.4771$ | Another new South African finance minister appointed (Areff, 2015) |
| 11 November 2016 | Up | 0.0049 | $***5.3157$ | Overall South African political struggles (Anon, 2016). Dow Jones at record high following Trump's victory (Farrell & Wearden, 2016). |
| 26 October 2017 | Up | 0.0020 | $***6.2260$ | Gupta deals (Anon, 2017) and South Africa's mini-budget (Muller, 2017) |
| 10 August 2018 | Up | 0.0003 | $***8.0515$ | Various economic factors (Anon, 2018a), including Turkey rout (Anon, 2018b) |
| 13 August 2018 | Up | 0.0018 | $***6.3155$ | Dow Jones jumps up as Turkish Lira rebounds (Imbert & Gibbs, 2018) |
| 14 February 2019 | Up | 0.0790 | $*2.4971$ | Eskom split (Toyana & Roelf, 2019) |
| 31 October 2019 | Up | 0.0244 | $**3.7008$ | South African mini-budget (Muller, 2019). No US rate hike (Imbert, 2019). |

CHAPTER 4

# Conclusion

In order to test for jumps one must assess certain estimators of the instantaneous volatility. The more well known estimators are realised quadratic variation and realised bipower variation, first introduced by Barndorff-Nielsen and Shephard (2004). An interesting and very useful property of these two estimators is that realised bipower variation is robust to the presence of jumps, but realised quadratic variation is not. Thus this allows us to use these estimations to determine the jump component of a diffusion process.

This property of the quadratic variation and bipower variation can be used to test for the presence of jumps. In the literature there are many papers on tests for jumps based on the bipower variation but two that were important in this study were the test by Barndorff-Nielsen and Shephard (2006), which is heavily based on the work from Barndorff-Nielsen and Shephard (2004), and the test by Lee and Mykland (2008). Although the test by Lee and Mykland (2008) incorporates bipower variation it can be used to test for individual jumps rather than testing for the presence of jumps within a certain time interval.

We were able to present the techincal work of these authors in an understandable manner for any interested researcher in the field of mathematical statistics without neglecting any crucial concepts.

As an application of the test by Lee and Mykland we tested for large jumps using daily Rand to US Dollar exchange rate data. Once identified we attempted to find possible events that may have caused the jumps. In their study Chatrath et al (2014) make the claim that when studying the exchange rate of the US Dollar to British pound, Euro, Japanese yen and Swiss franc, most of the jumps are caused by events in the US. Having assessed this for the Rand to US Dollar exchange rate we see that we cannot extend their claim to hold true for this additional case. Many of the jumps identified were as a result of political instability, which might be the reason for this contradicting result.

However, further research using higher frequency data and more rigorous assessment of events in the world may be needed to understand why the claim that most jumps are caused by US events does not hold true for the Rand to US Dollar exchange rate. Additionally such research would benefit from including more currencies from developing countries or countries with unstable economies. I suspect the outcome would be that jumps are caused by unexpected events and as such countries with unstable economies, with more unexpected events, will cause more jumps.

# Bibliography


[1] Aït-Sahalia, Y. and Jacod, J. 2014. *High-Frequency Financial Econometrics.* Princeton University Press.

[2] Alizadeh, S., Brandt, M.W. and Diebold, F.X. 2001. *High-and low-frequency exchange rate volatility dynamics: range-based estimation of stochastic volatility models* (No. w8162). National Bureau of Economic Research.

[3] Anon. 2014. Stock prices fall as interest rates climb. *CBS News*, 12 Sep. https://www.cbsnews.com/news/stock-prices-fall-as-interest-rates-climb/ Date of access: 2 Aug. 2020

[4] Anon. 2016. South Africa's political struggle hits economy. *BBC News*, 11 Nov. https://www.bbc.com/news/av/business-37951250/south-africa-s-political-struggle-hits-economy Date of access: 2 Jul. 2020

[5] Anon. 2017. SAP comes clean on R100 million Gupta 'kickbacks', *Business Tech*, 26 Oct. https://businesstech.co.za/news/business/207607/sap-comes-clean-on-r100-million-gupta-kickbacks/ Date of access: 2 Jul. 2020

[6] Anon. 2018a. 6 reasons why the Rand is getting hammered. *Business Tech*, 13 Aug. https://businesstech.co.za/news/finance/264313/6-reasons-why-the-rand-is-getting-hammered/ Date of access: 2 Jul. 2020

[7] Anon. 2018b. South Africa's Rand falls to 2-Year low as Turkey rout spreads. *VOA*, 13 Aug. https://www.voanews.com/africa/south-africas-rand-falls-2-year-low-turkey-rout-spreads Date of access: 2 Jul. 2020

[8] Areff, A. 2015. Timeline: how South Africa got three finance ministers in four days. *News24*, 14 Dec. https://www.news24.com/news24/southafrica/news/timeline-how-south-africa-got-three-finance-ministers-in-four-days-20151214 Date of access: 2 Jul. 2020

[9] Bajgrowicz, P., Scaillet, O. and Treccani, A. 2016. Jumps in high-frequency data: Spurious detections, dynamics, and news. *Management Science,* 62(8), 2198-2217.

[10] Bajgrowicz, P. and Scaillet, O. 2011. We propose a technique to avoid spurious detections of jumps in high-frequency data via an explicit thresholding on available test statistics *Swiss Finance Institute Research Paper Series*, 11-36.

[11] Balaban, E., 2004. Comparative forecasting performance of symmetric and asymmetric conditional volatility models of an exchange rate. *Economics Letters*, 83(1), 99-105.

[12] Barndorff-Nielsen, O.E. and Shephard, N., 2001. Non-Gaussian Ornstein-Uhlenbeck-based models and some of their uses in financial economics. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 63(2), 167-241.

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

[14] Barndorff-Nielsen, O.E. and Shephard, N. 2004. Power and bipower variation with stochastic volatility and jumps. *Journal of Financial Econometrics,* 2(1), 1-37.

[15] Barndorff-Nielsen, O.E. and Shephard, N. 2006. Econometrics of testing for jumps in financial economics using bipower variation. *Journal of Financial Econometrics*, 4(1), 1-30.

[16] Baum, C.F. and Zerilli, P. 2016. Jumps and stochastic volatility in crude oil futures prices using conditional moments of integrated volatility. *Energy Economics,* 53, 175-181.

[17] Chatrath, A., Miao, H., Ramchander, S. and Villupuram, S. 2014. Currency jumps, cojumps and the role of macro news. *Journal of International Money and Finance,* 40, 42-62.

[18] Cokayne, R. 2014. Peugeot chooses Nigeria over SA. *IOL*, 12 Sep. https://www.iol.co.za/business-report/economy/peugeot-chooses-nigeria-over-sa-1749744 Date of access: 2 Jul. 2020

[19] Corsi, F., Pirino, D. and Renò, R., 2010. Threshold bipower variation and the impact of jumps on volatility forecasting. *Journal of Econometrics*, 159(2), 276-288.

[20] Dewachter, H., Erdemlioglu, D., Gnabo, J.Y. and Lecourt, C., 2014. The intra-day impact of communication on euro-dollar volatility and jumps. *Journal of International Money and Finance*, 43, 131-154.

[21] Ding, L. and Vo, M., 2012. Exchange rates and oil prices: A multivariate stochastic volatility analysis. *The Quarterly Review of Economics and Finance*, 52(1), 15-37.

[22] Farrell, S. and Wearden, G. 2016. Why has Dow Jones hit a record high following Trump's victory? *The Guardian*, 10 Nov. https://www.theguardian.com/business/2016/nov/10/why-has-dow-jones-hit-a-record-high-following-trumps-victory Date of access: 2 Aug. 2020

[23] Fičura, M. and Witzany, J. 2015. Estimating stochastic volatility and jumps using high-frequency data and Bayesian methods. *Available at SSRN 2551807.*

[24] Hanousek, J., Kocenda, E. and Novotný, J., 2012. The identification of price jumps. *Monte Carlo Methods and Applications*, 18(1), 53-77.

[25] Hirsch, A. 2015. The removal of South Africa's finance minister is bad news for the country. *The Conversation*, 10 Dec. https://theconversation.com/the-removal-of-south-africas-finance-minister-is-bad-news-for-the-country-52170 Date of access: 2 Jul. 2020

[26] Huang, X. and Tauchen, G. 2005. The relative contribution of jumps to total price variance. *Journal of financial econometrics*, 3(4), 456-499.

[27] Imbert, F. and Gibbs, A. 2018. Dow jumps more than 100 points, snaps 4-day slide as Turkish lira rebounds from record low. *CNBC*, 14 Aug. https://www.cnbc.com/2018/08/14/us-markets-investors-shake-off-turkeys-economic-crisis.html Date of access: 2 Aug. 2020

[28] Imbert, F. 2019. S&P 500 closes at a record as Fed signals no rate hikes until inflation rises 'significantly'. *CNBC*, 30 Oct. https://www.cnbc.com/2019/10/30/stock-market-fed-rate-decision-and-earnings-in-focus-on-wall-street.html Date of access: 2 Aug. 2020

[29] Jiang, G.J. and Oomen, R.C., 2008. Testing for jumps when asset prices are observed with noise-a "swap variance" approach. *Journal of Econometrics*, 144(2), 352-370.

[30] Khoshnevisan, D. *Probability* American Mathematical Society, Providence, RI, 2007.

[31] Kuo, H.H. *Introduction to Stochastic Integration.* Springer, 2006.

[32] Laurent, S., Lecourt, C. and Palm, F.C. 2014. Testing for jumps in conditionally Gaussian ARMA-GARCH models, a robust approach. *Computational Statistics & Data Analysis,* 100, 383-400.

[33] Lee, S.S. and Mykland, P.A. 2008. Jumps in financial markets: A new nonparametric test and jump dynamics. *The Review of Financial Studies,* 21(6), 2535-2563.

[34] Muller, S.M. 2017. Latest budget underscores desperate state of South Africa's finances. *The Conversation*, 25 Oct. https://theconversation.com/latest-budget-underscores-desperate-state-of-south-africas-finances-86362 Date of access: 2 Jul. 2020

[35] Muller, S.M. 2019. Mini-budget underscores bad state of South Africa's economy. *The Conversation*, 30 Oct. https://theconversation.com/mini-budget-underscores-bad-state-of-south-africas-economy-126137 Date of access: 2 Jul. 2020

[36] Patton, A.J. and Sheppard, K. 2015. Good volatility, bad volatility: Signed jumps and the persistence of volatility. *Review of Economics and Statistics*, 97(3), 683-697.

[37] SA News. 2015. Watch: first address from South Africa's new finance minister. *Business Tech*, 11 Dec. https://businesstech.co.za/news/government/106877/watch-first-address-from-south-africas-new-finance-minister/ Date of access: 2 Jul. 2020

[38] SAPA. 2013. Rand weakness has positives. *IOL*, 31 May. https://www.iol.co.za/business-report/markets/currencies/rand-weakness-has-positives-1525374 Date of access: 2 Jul. 2020

[39] South African Reserve Bank. *Selected Historical Rates* https://www.resbank.co.za/en/home/what-we-do/statistics/key-statistics/selected-historical-rates Date of access: 3 Jun. 2020

[40] Tsay, R.S. *Analysis of financial time series. 3rd ed.* John Wiley and sons, Hoboken, NJ, 2010.

[41] Toyana, M. and Roelf, W. 2019. Eskom split to minimize risks to South Africa, Ramaphosa says. *Reuters*, 14 Feb. https://www.reuters.com/article/us-safrica-eskom/eskom-split-to-minimize-risks-to-south-africa-ramaphosa-says-idUSKCN1Q32BL Date of access: 2 Jul. 2020

[42] Zacks. 2014. Stock market news for January 30, 2014 - market news. *Nasdaq*, 30 Jan. https://www.nasdaq.com/articles/stock-market-news-for-january-30-2014-market-news-2014-01-30 Date of access: 2 Jul. 2020

APPENDIX A

# Omitted proofs

## A.1. Quadratic variation of diffusion processes

In this section we intend to prove that for the stochastic volatility process

$$Y_t = \int_0^t \sigma(s) dW_s,$$

the quadratic variation is given by

$$[Y]_t = \int_0^t \sigma^2(s) ds,$$

where $\sigma$ is the volatility process and $W_t$ is a Brownian motion.

Now as previously mentioned we will look at the case when $\sigma$ is a continuous deterministic function of time. Let $Y_t = \int_0^t \sigma(s) dW_s$ and let $\Delta Y_{t_i} = \int_{t_{i-1}}^{t_i} \sigma(s) dW_s$. Immediately we can say by Kuo (2006) Theorem 2.3.4 that $\int_0^t \sigma(s) dW_s \sim N(0, \mathrm{Var}(\sigma))$ where $\mathrm{Var}\,\sigma = \int_0^t \sigma^2(s) ds$. Let $X_i = \Delta Y_{t_i}^2 - \int_{t_{i-1}}^{t_i} \sigma(s)^2 ds$. We start by calculating $\mathrm{E}(X_i)$ and $\mathrm{E}(X_i X_j)$:

$$\begin{aligned}
\mathrm{E}\left(X_i^2\right) &= \mathrm{E}\Bigg[\left(\int_{t_{i-1}}^{t_i} \sigma(s) dW_s\right)^4 - 2\left(\int_{t_{i-1}}^{t_i} \sigma(s)^2 ds\right)\left(\int_{t_{i-1}}^{t_i} \sigma(s) dW_s\right)^2 \\
&\qquad + \left(\int_{t_{i-1}}^{t_i} \sigma(s)^2 ds\right)^2\Bigg] \\
&= 3\left(\int_{t_{i-1}}^{t_i} \sigma(s)^2 ds\right)^2 - 2\left(\int_{t_{i-1}}^{t_i} \sigma(s)^2 ds\right)^2 + \left(\int_{t_{i-1}}^{t_i} \sigma(s)^2 ds\right)^2 \\
&= 2\left(\int_{t_{i-1}}^{t_i} \sigma(s)^2 ds\right)^2
\end{aligned}$$

We calculate the following for $i \neq j$:

$$\begin{aligned}
\mathrm{E}(X_i X_j) &= \mathrm{E}\Bigg[\Delta Y_{t_i}^2 \Delta Y_{t_j}^2 - \Delta Y_{t_i}^2\left(\int_{t_{j-1}}^{t_j} \sigma(s)^2 ds\right) - \Delta Y_{t_j}^2\left(\int_{t_{i-1}}^{t_i} \sigma(s)^2 ds\right) \\
&\qquad + \left(\int_{t_{i-1}}^{t_i} \sigma(s)^2 ds\right)\left(\int_{t_{j-1}}^{t_j} \sigma(s)^2 ds\right)\Bigg] \\
&= \left(\int_{t_{i-1}}^{t_i} \sigma(s)^2 ds\right)\left(\int_{t_{j-1}}^{t_j} \sigma(s)^2 ds\right) - \left(\int_{t_{i-1}}^{t_i} \sigma(s)^2 ds\right)\left(\int_{t_{j-1}}^{t_j} \sigma(s)^2 ds\right) \\
&\qquad - \left(\int_{t_{i-1}}^{t_i} \sigma(s)^2 ds\right)\left(\int_{t_{j-1}}^{t_j} \sigma(s)^2 ds\right) + \left(\int_{t_{i-1}}^{t_i} \sigma(s)^2 ds\right)\left(\int_{t_{j-1}}^{t_j} \sigma(s)^2 ds\right) \\
&= 0.
\end{aligned}$$

Let $f(t) = \int_0^t \sigma(s)^2 ds$ thus by our assumptions regarding $\sigma$ in this case we find that this function $f$ is continuous with finite variation and thus the quadratic variation of $f$ will be 0.

$$\begin{aligned}
\mathrm{E}\left(\sum_{i=1}^{n} X_i\right)^2 &= \sum_{i=1}^{n} \mathrm{E}\left(X_i^2\right) + \sum_{\substack{i=1 \\ i \neq j}}^{n} \sum_{j=1}^{n} \mathrm{E}\left(X_i X_j\right) \\
&= 2\sum_{i=1}^{n} \left(\int_{t_{i-1}}^{t_i} \sigma(s)^2 ds\right)^2 \\
&= \sum_{i=1}^{n} \left|f(t_i) - f(t_{i-1})\right|^2 \to 0
\end{aligned}$$

## A.2. Compound Poisson quadratic variation

In this section we intend to prove that for a compound Poisson process

$$Y_t = \sum_{i=1}^{N_t} Z_i,$$

the quadratic variation is given by

$$[Y]_t = \sum_{i=1}^{N_t} Z_i^2,$$

where we assume that $Z_i \sim N(0, \sigma_Z^2)$.

Let $Y_t$ be a compound Poisson process with intensity $\lambda$ and normally distributed jump sizes. Using the moment generating function we can derive the first four moments of the compound Poisson distribution which we will need later. These moments can be shown, for $s < t$, to be

$$\begin{aligned} \mathrm{E}\,(Y_t - Y_s) &= 0 \\ \mathrm{E}\,(Y_t - Y_s)^2 &= \sigma_Z^2 \lambda\,(t-s) \\ \mathrm{E}\,(Y_t - Y_s)^3 &= 0 \\ \mathrm{E}\,(Y_t - Y_s)^4 &= 3\sigma_Z^4 \left(\lambda^2\,(t-s)^2 + \lambda\,(t-s)\right). \end{aligned} \tag{A.1}$$

Let $X_i = \left(Y_{t_i} - Y_{t_{i-1}}\right)^2 - \left([Y]_{t_i} - [Y]_{t_{i-1}}\right)$, where $[Y]_{t_i} = \sum_{j=1}^{N_{t_i}} Z_j^2$. In order to prove the required result we will need to show that $\sum_{i=1}^n X_i$ converges to 0 in probability and to do this it is sufficient to show $\mathrm{E}\left(\sum_{i=1}^n X_i\right)^2$ converges to 0. Observe that

$$\mathrm{E}\left(\sum_{i=1}^n X_i\right)^2 = \sum_{i=1}^n \mathrm{E}\left(X_i^2\right) + \sum_{\substack{j=1 \\ j\neq k}}^n \sum_{k=1}^n \mathrm{E}\,(X_j X_k)$$

In order to calculate the above we need to know $\mathrm{E}\left(X_i^2\right)$ and $\mathrm{E}\,(X_j X_k)$. Thus we continue by first calculating these values for some $1 \le i \le n$.

Let $\Delta Y_i = Y_{t_i} - Y_{t_{i-1}} = \sum_{j=N_{t_{i-1}}}^{N_{t_i}} Z_j$ and $\Delta N_i = N_{t_i} - N_{t_{i-1}}$. Then

$$\begin{aligned}
\mathrm{E}\left(X_i^2\right) &= \mathrm{E}\left(\Delta Y_i^4 - 2\Delta Y_i^2 \sum_{j=N_{t_{i-1}}}^{N_{t_i}} Z_j^2 + \left(\sum_{j=N_{t_{i-1}}}^{N_{t_i}} Z_j^2\right)^2\right) \\
&= \mathrm{E}\left(\Delta Y_i^4 - 2\left(\sum_{a=N_{t_{i-1}}}^{N_{t_i}} Z_a^2 + \sum_{\substack{b=N_{t_{i-1}} \\ b\neq c}}^{N_{t_i}} \sum_{c=N_{t_{i-1}}}^{N_{t_i}} Z_b Z_c\right) \sum_{j=N_{t_{i-1}}}^{N_{t_i}} Z_j^2\right. \\
&\qquad \left. + \sum_{j=N_{t_{i-1}}}^{N_{t_i}} Z_j^4 + \sum_{\substack{d=N_{t_{i-1}} \\ d\neq e}}^{N_{t_i}} \sum_{e=N_{t_{i-1}}}^{N_{t_i}} Z_d^2 Z_e^2\right) \\
&= 3\sigma_Z^4\left(\lambda^2\Delta t_i^2 + \lambda\Delta t_i\right) - 2\,\mathrm{E}\left[\sum_{a=N_{t_{i-1}}}^{N_{t_i}} Z_a^4 + \sum_{\substack{a=N_{t_{i-1}} \\ a\neq j}}^{N_{t_i}} \sum_{j=N_{t_{i-1}}}^{N_{t_i}} Z_a^2 Z_j^2\right. \\
&\qquad \left. + \sum_{\substack{b=N_{t_{i-1}} \\ b\neq c}}^{N_{t_i}} \sum_{c=N_{t_{i-1}}}^{N_{t_i}} Z_b Z_c \left(\sum_{j=N_{t_{i-1}}}^{N_{t_i}} Z_j^2\right)\right] + \mathrm{E}\left(\Delta N_i\right)\mathrm{E}\left(Z_j^4\right) \\
&\qquad + \mathrm{E}\left(\sum_{\substack{d=N_{t_{i-1}} \\ d\neq e}}^{N_{t_i}} \sum_{e=N_{t_{i-1}}}^{N_{t_i}} Z_d^2 Z_e^2\right) \\
&= 3\sigma_Z^4\left(\lambda^2\Delta t_i^2 + \lambda\Delta t_i\right) - 2\,\mathrm{E}\left(\Delta N_i\right)\mathrm{E}\left(Z_a^4\right) - 2\sigma_Z^4\lambda^2\Delta t_i^2 \\
&\qquad + 0 + 3\sigma_Z^4\lambda\Delta t_i + \sigma_Z^4\lambda^2\Delta t_i^2 \\
&= 2\lambda^2\Delta t_i^2\sigma_Z^4.
\end{aligned}$$

In the above we used the following two results: firstly

$$\begin{aligned}
\mathrm{E}\left(\sum_{\substack{k=N_{t_{i-1}}\\k\neq l}}^{N_{t_i}}\sum_{l=N_{t_{i-1}}}^{N_{t_i}} Z_k^2 Z_l^2\right) &= \mathrm{E}\left(\mathrm{E}\left(\sum_{\substack{k=N_{t_{i-1}}\\k\neq l}}^{N_{t_i}}\sum_{l=N_{t_{i-1}}}^{N_{t_i}} Z_k^2 Z_l^2 | \Delta N_i\right)\right)\\
&= \sum_{n=0}^{\infty}\mathrm{E}\left(\sum_{\substack{k=1\\k\neq l}}^{n}\sum_{l=1}^{n} Z_k^2 Z_l^2\right)P\left(\Delta N_i = n\right)\\
&= \sum_{n=0}^{\infty}\sum_{\substack{k=1\\k\neq l}}^{n}\sum_{l=1}^{n}\mathrm{E}\left(Z_k^2 Z_l^2\right)P\left(\Delta N_i = n\right)\\
&= \sum_{n=0}^{\infty}\sum_{\substack{k=1\\k\neq l}}^{n}\sum_{l=1}^{n}\sigma_Z^2\sigma_Z^2 P\left(\Delta N_i = n\right)\\
&= \sum_{n=0}^{\infty} n\left(n-1\right)\sigma_Z^4 P\left(\Delta N_i = n\right)\\
&= \sigma_Z^4\left[\mathrm{E}\left(\Delta N_i^2\right) - \mathrm{E}\left(\Delta N_i\right)\right]\\
&= \sigma_Z^4\left(\lambda^2\Delta t_i^2\right),
\end{aligned}$$

which follows using the tower property and the independence of $N_t$ and $Z_i$, and secondly

$$\mathrm{E}\left(\sum_{\substack{b=N_{t_{i-1}}\\b\neq c}}^{N_{t_i}}\sum_{c=N_{t_{i-1}}}^{N_{t_i}} Z_b Z_c\left(\sum_{j=N_{t_{i-1}}}^{N_{t_1}} Z_j^2\right)\right) = 0$$

which follows since $b \neq c$. We see that for any given $b$ and $c$ one or both of $Z_b$ and $Z_c$ will be independent from the other random variables and since $\mathrm{E}\left(Z_b\right) = 0$ for any valid $b$ the sum reduces to 0 as above. We now need the value $\mathrm{E}\left(X_i X_j\right)$ for $i \neq j$. It will be useful to calculate $\mathrm{E}\left(X_i\right)$ first:

$$\begin{aligned}
\mathrm{E}\left(X_i\right) &= \mathrm{E}\left(\Delta Y_i^2 - \sum_{j=N_{t_{i-1}}}^{N_{t_i}} Z_j^2\right)\\
&= \sigma_Z^2\lambda\Delta t_i - \mathrm{E}\left(\Delta N_i\right)\mathrm{E}\left(Z_j^2\right)\\
&= \sigma_Z^2\lambda\Delta t_i - \sigma_Z^2\lambda\Delta t_i = 0.
\end{aligned}$$

Now we look at $\mathrm{E}\left(X_i X_j\right)$:

$$\begin{aligned}
\mathrm{E}\left(X_i X_j\right) &= \mathrm{E}\left(\mathrm{E}\left(X_i X_j | \mathcal{F}_{t_{j-1}}\right)\right)\\
&= \mathrm{E}\left(X_i\,\mathrm{E}\left(X_j | \mathcal{F}_{t_{j-1}}\right)\right)\\
&= \mathrm{E}\left(X_i\,\mathrm{E}\left(X_j\right)\right)\\
&= \mathrm{E}\left(0\right) = 0
\end{aligned}$$

where we have used the fact that $X_j$ is independent of $\mathcal{F}_{t_{j-1}}$ since $Y_t$ has independent increments. Now we can finally look at $\mathrm{E}\left(\sum_{i=1}^n X_i\right)^2$. Let $\|\Delta_n\| = \max_{1\leq i\leq n}(\Delta t_i)$. Then

$$\begin{aligned} E\left(\sum_{i=1}^n X_i\right)^2 &= \sum_{i=1}^n E\left(X_i^2\right) + \sum_{\substack{j=1 \\ j\neq k}}^n \sum_{k=1}^n E\left(X_j X_k\right) \\ &= \sum_{i=1}^n 2\lambda^2 \Delta t_i^2 \sigma_Z^4 \\ &\leq 2\lambda^2 \sigma_Z^4 n \|\Delta_n\|^2 \to 0 \end{aligned} \tag{A.2}$$

as $\|\Delta_n\| \to 0$. This proves the result.

## A.3. Diffusion processes bipower variation

In this section we intend to prove that for the stochastic volatility process

$$Y_t = \int_0^t \sigma(s)dW_s,$$

the bipower variation is given by

$$[Y]_t^* = \frac{2}{\pi}\int_0^t \sigma^2(s)ds,$$

where $\sigma$ is the volatility process and $W_t$ is a Brownian motion.

We turn our attention to the case where $\sigma$ is a continuous deterministic function, this will allow us to use theorem 2.3.4 on page 11 of Kuo (2006). As before we proceed by estimating $\mathrm{E}\left(X_i^2\right)$, $\mathrm{E}\left(X_iX_{i+1}\right)$ and $\mathrm{E}\left(X_iX_j\right)$. An important difference though is that we will define $X_i$ slightly differently than in the proof of Theorem 2.11, first we let $f(t_i) = \int_{t_{i-1}}^{t_i}\sigma(s)^2ds$, now we let

$$X_i = |\Delta Y_{t_i}||\Delta Y_{t_{i+1}}| - \frac{2}{\pi}\sqrt{f(t_i)}\sqrt{f(t_{i+1})}.$$

Here we have $\Delta Y_{t_i} = \int_{t_{i-1}}^{t_i}\sigma(s)dW_s$ and we know that $|\Delta Y_{t_i}|$ will follow a half-normal distribution. Bearing all this in mind we find that,

$$\begin{aligned}
\mathrm{E}\left(X_i^2\right) &= \mathrm{E}\left[\Delta Y_{t_i}^2\Delta Y_{t_{i+1}}^2 - 2|\Delta Y_{t_i}||\Delta Y_{t_{i+1}}|\left(\frac{2}{\pi}\sqrt{f(t_i)}\sqrt{f(t_{i+1})}\right) + \frac{4}{\pi^2}\left(\sqrt{f(t_i)}\sqrt{f(t_{i+1})}\right)^2\right] \\
&\le \left(1+\frac{4}{\pi^2}\right) f(t_i)f(t_{i+1}),
\end{aligned}$$

and that

$$\begin{aligned}
\mathrm{E}\left(X_iX_{i+1}\right) &= \mathrm{E}\Big[|\Delta Y_{t_i}|\left(\Delta Y_{t_{i+1}}\right)^2|\Delta Y_{t_{i+2}}| - \frac{2}{\pi}|\Delta Y_{t_i}||\Delta Y_{t_{i+1}}|\left(\sqrt{f(t_{i+1})f(t_{i+2})}\right) \\
&\qquad - \frac{2}{\pi}|\Delta Y_{t_{i+1}}||\Delta Y_{t_{i+2}}|\left(\sqrt{f(t_i)f(t_{i+1}))}\right) + \frac{4}{\pi^2}f(t_{i+1})\left(\sqrt{f(t_i)f(t_{i+2})}\right)\Big] \\
&\le \frac{2}{\pi}\left(\int_{t_{i-1}}^{t_i}\sigma(s)^2ds\int_{t_{i+1}}^{t_{i+2}}\sigma(s)^2ds\right)^{\frac{1}{2}}\left(\int_{t_{i-1}}^{t_i}\sigma(s)^2ds\right) \\
&\qquad + \frac{4}{\pi^2}f(t_{i+1})\left(\sqrt{f(t_i)f(t_{i+2})}\right) \\
&= \left(\frac{2}{\pi}+\frac{4}{\pi^2}\right) f(t_{i+1})\left(\sqrt{f(t_i)f(t_{i+2})}\right).
\end{aligned}$$

We find that $X_i$ and $X_j$ are independent for $j \ne i-1$ and $j \ne i$ since $\Delta Y_i$, $\Delta Y_{i+1}$, $\Delta Y_j$, $\Delta Y_{j+1}$ are independent for $i < j-1$ and $i > j+1$. Thus we have that

$$\begin{aligned}
\mathrm{E}\left(X_i\right) &= \mathrm{E}\left(|\Delta Y_{t_i}||\Delta Y_{t_{i+1}}| - \frac{2}{\pi}\sqrt{f(t_i)}\sqrt{f(t_{i+1})}\right) \\
&= \mathrm{E}\left(|\Delta Y_{t_i}|\right)\mathrm{E}\left(|\Delta Y_{t_{i+1}}|\right) - \frac{2}{\pi}\sqrt{f(t_i)}\sqrt{f(t_{i+1})} \\
&= \frac{2}{\pi}\sqrt{\int_{t_{i-1}}^{t_i}\sigma(s)^2ds}\sqrt{\int_{t_i}^{t_{i+1}}\sigma(s)^2ds} - \frac{2}{\pi}\sqrt{\int_{t_{i-1}}^{t_i}\sigma(s)^2ds}\sqrt{\int_{t_i}^{t_{i+1}}\sigma(s)^2ds} \\
&= 0,
\end{aligned}$$

this follows since $|\Delta Y_{t_i}|$ follows the half-normal distribution, thus we now easily see the following:

$$\begin{aligned}
\mathrm{E}\left(X_iX_j\right) &= \mathrm{E}(X_i)\,\mathrm{E}(X_j) \\
&= 0.
\end{aligned}$$

Finally we observe that

$$
\begin{aligned}
\mathrm{E}\left(\sum_{i=1}^{n-1} X_i\right)^2 &= \sum_{i=1}^{n-1} \mathrm{E}\left(X_i^2\right) + 2\sum_{i=1}^{n-2} \mathrm{E}\left(X_i X_{i+1}\right) + \mathrm{E}\left(\sum_{\substack{j=1 \\ j\neq k}}^{n-1} \sum_{\substack{k=1 \\ j\neq k+1 \\ j\neq k-1}}^{n-1} X_j X_k\right) \\
&\leq \sum_{i=1}^{n-1}\left(1+\frac{4}{\pi^2}\right) f(t_i) f(t_{i+1}) + 2\sum_{i=1}^{n-2}\left(\frac{2}{\pi}+\frac{4}{\pi^2}\right) f(t_{i+1})\left(\sqrt{f(t_i) f(t_{i+2})}\right) + 0 \\
&\leq \left(1+\frac{4}{\pi^2}\right) \max_{1\leq i\leq n} f(t_i) \int_0^t \sigma^2(s)ds + 2\left(\frac{2}{\pi}+\frac{4}{\pi^2}\right)\sqrt{\max_{1\leq i\leq n} f(t_i)^2} \int_0^t \sigma^2(s)ds \\
&\to 0
\end{aligned}
$$

as $\|\Delta_n\| = \max\left(t_i - t_{i-1}\right) \to 0$ and the above follows since $\max_{1\leq i\leq n} f(t_i) \to 0$. Thus our desired final result will follow since as $\|\Delta_n\| = \max\left(t_i - t_{i-1}\right) \to 0$ we have

$$
\begin{aligned}
\frac{2}{\pi}\sum_{i=1}^{n-1}\sqrt{\int_{t_{i-1}}^{t_i} \sigma^2(s)ds}\sqrt{\int_{t_i}^{t_{i+1}} \sigma^2(s)ds} &\simeq \frac{2}{\pi}\sum_{i=1}^{n-1}\int_{t_{i-1}}^{t_i} \sigma^2(s)ds \\
&= \int_0^t \sigma^2(s)ds.
\end{aligned}
$$

In the above case we have heavily used the fact that a deterministic function $\sigma(s)$ will allow us to treat the stochastic integral $\int_0^t \sigma(s)dW_s$ as a Gaussian random variable. Unfortunately for the case when $\sigma(s)$ is itself a random variable as well this fact does not hold anymore. As such our method of proof will not work thus for the proof of the most general case, that is when $\sigma(s)$ is random, see Barndorff-Nielsen and Shephard (2004).

# Notation Index

# Subject Index