Pricing and hedging variance swaps using stochastic volatility models

by

Kudakwashe Bopoto

(13110782)

Dissertation submitted in partial fulfillment of the requirements for the degree

Magister Scientiae

in Financial Engineering

in the Department of Mathematics and Applied Mathematics

in the Faculty of Natural and Agricultural Sciences

University of Pretoria

Pretoria

October 2019

Declaration

I, the undersigned, declare that the dissertation / thesis, which I hereby submit for the degree Magister Scientiae in Financial Engineering at the University of Pretoria, is my own independent work and has not previously been submitted by me for a degree at this or any other tertiary institution.

Signature: K Bopoto

Name: Kudakwashe Bopoto

Date: October 2019

Acknowledgements

I would like to thank the Almighty God for life and for giving me this opportunity to study. I would like to thank my wife, Rutendo, for her constant encouragement and my family for supporting me throughout the process. I would like to thank Prof. Eben Mare for his patience and sound advice which aided me in completing this dissertation.

Abstract

In this dissertation, the price of variance swaps under stochastic volatility
models based on the work done by Barndorff-Nielsen and Shepard (2001) and
Heston (1993) is discussed. The choice of these models is as a result of prop-
erties they possess which position them as an improvement to the traditional
Black-Scholes (1973) model. Furthermore, the popularity of these models in
literature makes them particularly attractive. A lot of work has been done
in the area of pricing variance swaps since their inception in the late 1990's.
The growth in the number of variance contracts written came as a result of
investors' increasing need to be hedged against exposure to future variance
fluctuations. The task at the core of this dissertation is to derive closed or
semi-closed form expressions of the fair price of variance swaps under the two
stochastic models. Although various researchers have shown that stochastic
models produce close to market results, it is more desirable to obtain the fair
price of variance derivatives using models under which no assumptions about
the dynamics of the underlying asset are made. This is the work of a use-
ful analytical formula derived by Demeterfi, Derman, Kamal and Zou (1999)
in which the price of variance swaps is hedged through a finite portfolio of
European call and put options of different strike prices. This scheme is practi-
cally explored in an example. Lastly, conclusions on pricing using each of the
methodologies are given.

Contents

IV Conclusions 158

13 Summary 158

14 Future research considerations 160

A Theoretical Concepts 161

B Stochastic Calculus 164

C MATLAB Code 174

List of Tables

List of Figures

List of Symbols and Abbreviations

- S_T: Stock price at time T

- $\sigma_R^2(T)$:Realised variance at time T

- K_{var} : Variance swap strike or fair variance swap strike

- $\mathbb{Q}$: Equivalent martingale measure

- $\mathbb{P}$: Real probability measure

- Ω: The whole space

- $\mathcal{F}_t$: Filtration containing information up to time t

- $\mathbb{E}_{\mathbb{Q}}$: Expectation under the equivalent martingale measure $\mathbb{Q}$

- $(X(t), 0 \leq t \leq T)$: Stochastic process X on interval $[0, T]$

- ϕ_X: Characteristic function of X

- ψ_X: Cumulant-generating function of X

- Γ : Gamma function or distribution

- $K_1(x)$: Modified Bessel function of the third kind

- $W(t)$: Wiener process or Brownian motion

- $L(\theta)$: Likelihood function for parameter set θ

- $\mathfrak{F}[f(t)]$: Fourier transform of $f(t)$

- B-NS: Barndorff-Nielsen and Shepard

- OU: Ornstein-Uhlenbeck

- BS: Black-Scholes

- RMSE: Root mean-square error

- MSE: Mean square-error

- MME: Method of moments estimator

- MLE: Maximum likelihood estimator

- NIG: Normal Inverse Gaussian distribution

- PDE: Partial Differential Equation

- SDE: Stochastic Differential Equation

- MGF: Moment Generating Function

1 Introduction

In this chapter, a background of variance swaps is introduced. Furthermore, a literature review on developments in the pricing of variance swaps is conducted. Lastly, an overview of the objectives of this dissertation is presented.

1.1 Background

Derivatives whose underlying is the variance of asset returns first became popular in the 1990s with an increasing need for investors to hedge themselves against future volatility fluctuations in turbulent times such as those which saw the collapse of the Long Term Capital Management (LTCM) in 1998 [49, 22]. Ideas of creating a volatility index on options have since been suggested from the 1970s emanating from the cornerstone option pricing ideas developed by Black and Scholes (1973) [12]. These ideas were refined by Fleming, Ostdiek and Whaley (1995) amongst other researchers to give the foundation which brought about the construction of the CBOE (Chicago Board Options Exchange) volatility index (VIX) which is used to predict expected future volatility by studying its short historic behaviour [32]. The VIX is commonly referred to as the 'fear index' because of its negative correlation with asset prices presents an idea of market perception in a given period. This index is particularly important to the establishment of derivatives on volatility which began trading OTC (over-the-counter) before being incorporated into the CBOE in 2006. Variance swaps, being derivatives on the realised variance of asset returns are closed tied to the activities of this index.

Throughout financial history, calm periods such as the boom before 2007 are followed are almost always followed turbulent periods (a phenomena known as volatility clustering) such as the 2007-2008 financial crisis in in which volatility soared to reach

unprecedented levels. These volatility movements though at different magnitudes are the same ones observed with political events such as the election of Donald Trump in November 2016 and sudden market movements such as the 5 February 2018 'stock market correction' which saw the Dow Jones Industrial Average (DJI) decreasing by 4.2% (over 1000 points) - one of the highest single-day drops in history, higher than the highest drop during the 2008 financial crisis and the 1987 black Monday. The CBOE Volatility Index (VIX) consequently increased by 14.5% over the same period [40].

With such events occurring unpredictably, the need for variance products is of paramount significance especially amongst traders dealing in over-the-counter (OTC) contracts. The need for trading variance products can be summarised by simple speculation of future variance levels thus presenting a money-making opportunity for the correctly positioned trades and the need to trade spreads between implied variance (variance which the market has implicitly used for valuing a benchmark option) and realised variance (actual historical variance over a fixed period) [11, 51].

1.1.1 Variance derivatives

Whilst the variance of asset returns has traditionally been seen as a measure of risk, over the past two decades variance itself has developed it into an asset class. Variance swaps are forward contracts or futures that provide direct exposure to future realised variance (square-root of realised volatility) of the returns of stocks rather than exposing the investors directly to the movements of the underlying stocks themselves [8, 49]. Their payoff is the product between the notional amount of the swap in dollars per annualised volatility point squared and the difference between realised variance of the asset over a predetermined period and the strike price of the variance agreed upon entry into the contract [8, 49]. This can be presented as:

$$N \left(\sigma_R^2(T) - K_{var} \right) \tag{1.1}$$

where N is the notional amount, $\sigma_R^2(T)$ is the realised variance at expiry and K_{var} the predetermined strike level.

The attractiveness of variance swaps to investors and traders lies in this direct exposure to the variance of the asset returns as well as the flexibility in pricing them. Carr and Madan (1998) showed how volatility and variance can be traded and priced through replicating a static portfolio of options. Furthermore, they showed how buying or selling contracts with a payoff explicitly containing realised variance as in the case of variance swaps and delta-hedging vanilla options is implemented [22]. However, the latter is less attractive because of the need to continuously re-balance and re-hedge the portfolio of options making contracts on variance more attractive. Replicating a static portfolio of options has a drawback in that sampling times of variance swaps are assumed to be continuous when they are discrete in real-life scenarios thereby pausing an estimation error as pointed out by Carr and Lee (2009) [20]. This leaves the contracts whose payoff contains variance more favourable. Since variance swaps are futures contracts, there is zero cost of entering into the contracts. This makes them even more appealing to some investors and traders.

Variance swaps were originally established on the underlying of indices and traded in over-the-counter (OTC) contracts. The payoff of such traded variance swaps is generally a function of realised or historical variance which is discretely sampled at predetermined times [51]. However, over the past decade variance swaps also saw themselves being traded in a second category as standardized contracts such as the CBOE S&P 500 variance futures. The payoff of such traded variance swaps is normally a function of implied variance [51, 8].

1.2 Stochastic volatility

In the pricing of many derivatives including variance swaps, the Black and Scholes (1973) pricing model is widely utilised. The BS model assumes that the volatility of the stocks is constant [12]. However, the constant volatility assumption fails to capture the market dynamics observed in reality which led to the development of stochastic models such as the Heston (1993) and the non-Gaussian Ornstein-Uhlenbeck (OU) model discussed in Barndorff-Nielsen and Shepard (B-NS) (2001) [5, 34].

Moreover, constant volatility models can not explain the volatility clustering under which there are periods in which prices abruptly move in a single direction in an observable random pattern. Over the decade preceding the 2007-2008 financial crisis, the volatility derivatives market received an increasing focus because of a growing need amongst market stakeholders to protect themselves against the risks from volatility movements and to be able to quantify and hedge these risks. Trading in instruments whose payoff has a component derived from stochastic volatility models such as the non-Gaussian OU model has enabled market participants to replicate the skewness and fat-tails observed in the returns of high-frequency assets markets thereby giving a more realistic implied volatility which shows aspects such as smiles experienced in reality [8].

Under the B-NS (2001) model, the volatility is given as a mean-reverting stationary stochastic process of the OU type driven by a subordinator (a Lévy process with no Gaussian component and positive increments) [5, 8, 44]. Another model that will be considered in this dissertation is the Heston (1993) model which allows the correlation between volatility and stock returns [34]. Zhu and Lian (2001) derived a closed-form solution for the price of discretely sampled variance swaps under the Heston (1993) model. The solution can be extended to continuous sampling as the

sampling frequency approaches infinity [51]. The Heston model is also implemented by Elliot, Siu and Chan (2006) to determine a martingale measure to price variance and volatility swaps after the adoption of a regime-switching Esscher transform [30]. However, contrary to their ability to represent market dynamics more accurately, stochastic volatility models are incomplete and thus most of the fundamental pricing theory becomes is violated.

1.3   Literature review

Owing to their ability to effectively provide direct exposure to variance, it is no surprise that there are various approaches to pricing discretely sampled variance swaps. Jia, Bi and Zhang (2015) have categorised these valuation techniques into i) numerical methods and ii) analytical methods [38]. Analytical methods of valuating variance swaps usually aim at deriving closed-form expressions for the fair-price of the variance swap under a risk-neutral measure. Examples of papers which analytically evaluate volatility and variance swaps are [8, 20, 28, 30, 44, 49]. Numerical methods on the other hand, approximate complex analytical methods to an acceptable degree and error.

An example of numerical method pricing is Little and Pant (2001) who approximate the price of a discretely sampled variance swaps using a finite-difference approach to solve a set of second-order parabolic partial differential equations [41]. The dimension reduction approach implemented in this numerical approach attains a high level of efficiency and accuracy for valuing the discretely sampled variance swaps [38]. Even though a high degree of accuracy is obtained using this numerical method, stochastic volatilities which are representative of actual market dynamics are not incorporated [38]. Dupire (1993), points out that the volatilities which are implied from the market prices of option prices valued under the BS framework exhibit a

degree of dependence between the implied volatility and the strike prices (smiles). However, attempting to prices the smiles into the BS model results in jumps and stochastic volatility as shown in Merton's 1976 model and Hull and White (1987) [26, 36, 42].

Carr and Lee (2009) navigate through the timeline of the development of variance derivatives and the earliest literature for the pricing of these derivatives [20]. Dupire (1993), derived the price of options with stochastic volatility which was extended by Carr and Madan (1998) to discuss ways in which variance can be traded. Carr and Madan (1998), described the trading of variance through replicating a portfolio of static positions in options, hedging of volatility exposure through dynamically trading the underlying and construction of contracts whose payoff is a function of variance [22, 20, 27]. Carr and Madan (1998), furthermore, derive the analytical formula without specifying the process of the underlying [23]. The dynamic replication strategy, in reality, is costly as there is a need for constant re-adjustment of the portfolio.

Literature categorises the pricing of variance swaps into those discretely sampled and ones which are continuously-sampled. Zhu and Lian (2011) extended the work done by Little and Pant (2001) to derive a closed-form exact solution for a partial differential equation system based on the Heston (1993) model with two-factors [51]. Under this framework, they look at the discretely sampled variance case. Most researchers aim to obtain a closed-form expression to effectively price instruments whose payoffs are functions of variance under a risk-neutral measure. However, this may not always be easily achievable thus quicker methods have to be implemented in some cases.

Furthermore, based on the Heston model, Elliot et al. (2006) [30] also developed a model for pricing volatility derivatives including variance swaps under a continuous-

time Markov-regulated version of the Heston (1993) stochastic model. They implemented a regime-switching Esscher transform to determine a martingale pricing measure for the valuation of volatility derivatives in an incomplete market [30]. After comparing this model to one without regime-switching, they showed that properties and valuation of volatility and variance swaps based on the regime-switching continuous-time Markov-regulated version of the Heston (1993) model were significantly higher than those without regime-switching [30].

Researchers have studied the impact of using continuously sampled variance swaps to estimated the discretely sampled variance swaps. Carr, Lee and Wu (2012), show that discretely sampled variance swaps have an increased price under an independence condition. This implies that the commonly quoted continuously-sampled versions often underestimate the price of variance swaps [21]. Broadie and Jain (2005) investigated the impact of jumps on the price of the variance swaps together with the impact of continuously sampling variance swaps compared to the price of actual discretely sampled variance swaps under the Heston model [13]. This was extended by Bernard and Cui (2013) who adopted a parametric approach that allowed the derivation of explicit closed-form expressions and asymptotic behavior with respect to key parameters such as the maturity of the contract, the risk-free rate, the sampling frequency, the volatility of the variance process and the correlation between the underlying stock and its volatility [10].

Carr, Madan and Yor (2005) suggest a method of pricing options on realised variance that are based on quadratic variation [18]. Benth et al. (2006) consider the non-Gaussian Ornstein-Uhlenbeck stochastic model suggested by Bandorff-Neilsen and Shepard (2001) to price variance swaps. They derived an analytical formula for the realized variance which enabled the price of the variance swap to be represented in terms of Laplace transforms. The fast Fourier transform method was used and

results compared to the approximation proposed by Brockhaus and Long [8]. Zhaoli et al. (2015) also derive a similar analytical formula to price variance swaps under the B-NS OU model and further compare the fair strike value based on the discrete model, continuous model, and Monte Carlo simulations [38].

1.4   Structure and objectives of the dissertation

The main objective of this dissertation is to derive closed/semi-closed expressions of the price of variance swaps under the stochastic volatility models by B-NS (2001) and Heston (1993). Furthermore, a secondary objective is to hedge variance swaps under a framework in which there are no assumptions made about the dynamics of the variance process. This is achieved through the static replication of the variance swap using a portfolio of vanilla options.

First, the instantaneous variance process is considered to have a non-Gaussian Ornstein-Uhlenbeck (OU) process which exhibits positive jumps only as presented by Barndoff-Nielsen and Shepard (2001) [5]. The analytical expression for the continuously sampled realised variance under the non-Gaussian OU model is derived using a key formula drawn from inverse Laplace transform and Fourier transform techniques. The work under Benth and Saltyte-Benth (2004) [9] and Benth et al. (2006) [8] is closely followed.

Next, the stochastic variance process is investigated under the model proposed by Heston (1993). The closed-form solution for pricing variance swaps with stochastic volatility is obtained from solving the model's partial differential equation (PDE) using Fourier transform techniques under the Feynman theorem. The valuation of the variance swap based on the discrete sampling and continuous sampling together with Monte Carlo simulations are presented in this chapter. Lastly, the Demeterfi

(1999) technique is implemented to a hypothetical portfolio of put and call options to hedge variance swaps. The rest of the dissertation is structured as follows:

In Chapter 2, the preliminary mathematical, finance and statistical concepts which form the building blocks of the concepts discussed in later chapters are introduced. Although all of the concepts could not be exhausted in this chapter, core concepts such as foundations of measure theory, stochastic calculus, stochastic processes, finance issues such as the imperfections of the Black-Scholes model, specific stochastic processes such as the Ornstein-Uhlenbeck processes and distributions considered under these processes such as the normal inverse-Gaussian (NIG) distribution are presented.

Chapter 3, presents the valuation of variance swaps under the B-NS non-Gaussian OU model. The chapter begins by giving the dynamics of the stock price and variance under this model. Next, the variance swap theory is re-introduced and then a solution of the model is presented. The superposition of the OU process is discussed then a transform approach is implemented to obtain the price of the variance swap from its exponent rather than the realised variance itself. The price of the variance swap is derived from inverse Laplace techniques. A semi-closed expression for the price of variance swaps is obtained through the Fast Fourier Transform (FFT). Lastly, simulations of the model and the variance swap valuations are presented.

Chapter 4, presents the Heston model. Here, the discretely sampled value of the variance swap is first considered. The terminal PDE of the model is considered under Feynman theory. An analysis of the payoff of the contingent claim changes the PDE into a two-part problem. The Feynman-Kac theorem is utilised in solving the two terminal-value problems. The closed-form of the variance swap is obtained as a combination of the two solutions. The continuously sampled case is also considered and lastly, the convergence of the discrete model to the continuous model considered

in a numerical example.

Chapter 5, discusses the hedging of variance swaps using a portfolio of European options. The chapter commences by showing that the fair price of a variance swap can be hedged by a log-price contract. However, since log-price contracts are not actively traded, the problem is shifted to replicating the payoff of the log-price contract. The proposition from Carr and Madan (1998) [23] is introduced and then the price of the variance swap is derived. However, the valuation requires a continuum of European options of all strikes which is not the case in reality where there are a finite number of options exist. A new approximation is derived for a finite number of options and then a numerical example is given to illustrate the price of variance swaps under this methodology.

Chapter 6, gives conclusions on the work considered in this dissertation together with the limitations and challenges.

Figure 1.1: Structure of the dissertation
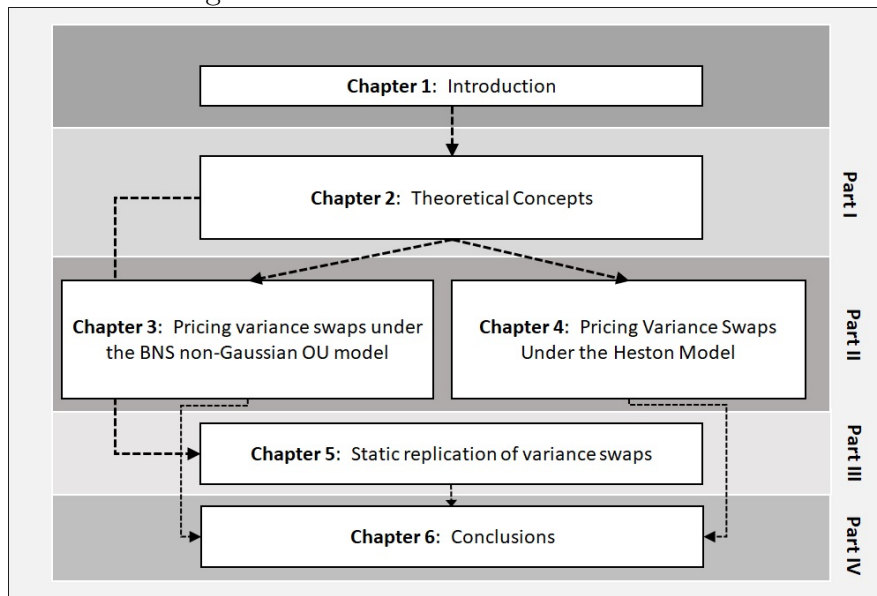
Part I

Theoretical Concepts

In this part, theoretical concepts that are directly or indirectly applied in this dissertation, are presented. These concepts cover a broad spectrum from variance swaps theory, measure theory, stochastic calculus, stochastic processes, examples of common stochastic processes such as the Ornstein-Uhlenbeck processes and the distributions considered under these processes such as the Normal Inverse Gaussian (NIG) distribution.

2   Probability and Measure theory fundamentals

The discussion begins with the definition of the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ under which the work in this dissertation will be considered. Where Ω is the whole space, $\mathcal{F}$ is a $\sigma - field$ (also called a $\sigma - algebra$) consisting of subsets of Ω and $\mathbb{P}$ is a probability measure such that $\mathbb{P} : \mathcal{F} \to \mathbb{R}$.

Definition 1. $(\sigma - field, Applebaum (2009)$ [1])

Let Ω be a non-empty set and $\mathcal{F}$ a collection of subsets of Ω. We call $\mathcal{F}$ a $\sigma - field$ if the following hold:

1. $\Omega \in \mathcal{F}$

2. If $A \in \mathcal{F}$, then $A^c \in \mathcal{F}$;

3. If $A_1, A_2, A_3, \dots$ is a sequence of elements of $\mathcal{F}$, then $\bigcup_{i=1}^{\infty} A_i \in \mathcal{F}$.

Definition 2. (Measure and Measurable Space, *Applebaum (2009)* [1])

The pair $(\Omega, \mathcal{F})$ is called a measurable space. A measure on $(\Omega, \mathcal{F})$ is a mapping $\mu : \mathcal{F} \to [0,\infty]$ that satisfies:

1. $\mu(\varnothing) = 0$;

2. $\mu \left(\bigcup\limits_{i=1}^{\infty} A_i \right) = \sum\limits_{i=1}^{\infty} \mu\left(A_i\right)$ for every sequence $A_1, A_2, A_3, ...$ of mutually disjoint sets in $\mathcal{F}$.

The triple $(\Omega, \mathcal{F}, \mu)$ is defined as measure space.

Definition 3. (Probability Measure and Probability Space, *Bain and Engelhardt (1992)* [2])

Suppose that Ω is a sample space and that $\mathcal{F}$ is a $\sigma - field$ of subsets of Ω, then the function $\mathbb{P} : \mathcal{F} \to \mathbb{R}$ is a probability measure function if:

1. $\mathbb{P}(A) \geq 0$, for all $A \in \mathcal{F}$;

2. $\mathbb{P}\left(\emptyset\right) = 0$ and $\mathbb{P}(\Omega) = 1$;

3. If $A_1, A_2, A_{3,...}$ is a sequence of mutually disjoint elements of $\mathcal{F}$, then $P \left(\bigcup\limits_{i=1}^{\infty} A_i \right) = \sum\limits_{i=1}^{\infty} P\left(A_i\right)$.

The triple $(\Omega, \mathcal{F}, P)$ is defined as probability space.

Definition 4. (Filtration, *Applebaum (2009)* [1])

Let $\mathcal{F}$ be a $\sigma - field$ of subsets of Ω, then a family of sub $\sigma - fields$, $\mathcal{F}_t := \{\mathcal{F}\}_{0 \leq t \leq T}$, is denoted a filtration if

$\mathcal{F}_s \subseteq \mathcal{F}_t$ whenever $s \leq t$.

The probability space $(\Omega, \mathcal{F}, \mathcal{F}_t, P)$ with the property defined above is called a *Filtered Probability Space.* In general, it is assumed that the filtered probability space satisfies the following 'usual' conditions [47]:

1. $\mathcal{F}$ is $\mathbb{P}$-complete;

2. $\mathcal{F}_0$ contains all the possible null sets of Ω;

3. $\{\mathcal{F}_t\}_{0 \leq t \leq T}$ is right continuous i.e. $\mathcal{F}_t = \bigcap_{s>t} \mathcal{F}_s$

Now that the fundamental concepts in probability theory have been presented, it is important to state results that are useful for measures. Particularly for absolutely continuous measures the *Radon-Nikodým Theorem* is presented. The theorem enables one to shift measures which critical in financial mathematics when moving from a probability measure in a risky world to that in a risk neutral world.

Definition 5. (Absolutely Continuous, *Barra (1981)* [7])

Let $\mathbb{P}$ and $\mathbb{Q}$ be measures on a measurable space $(\Omega, \mathcal{F})$. It is said that $\mathbb{Q}$ is *absolutely continuous* with respect to $\mathbb{P}$, denoted by $\mathbb{Q} \ll \mathbb{P}$, if $\mathbb{Q}(A) = 0$ for every $A \in \mathcal{F}$ for which $\mathbb{P}(A) = 0$ [7].

Two measures $\mathbb{P}$ and $\mathbb{Q}$ are said to be *equivalent* if they are mutually absolutely continuous i.e. $\mathbb{Q} \ll \mathbb{P}$ and $\mathbb{P} \ll \mathbb{Q}$ [1].

Theorem 1. *(Radon-Nikodým, Applebaum (2009) [1])*

Let $\mathbb{Q}$ and $\mathbb{P}$ be $\sigma - finite$ measures on a measurable space $(\Omega, \mathcal{F})$. Suppose that $\mathbb{Q} \ll \mathbb{P}$, then there exists a measurable function $g : \Omega \to \mathbb{R}^+$ such that, for all subsets $A \in \mathcal{F}$

$$\mathbb{Q}(A) = \int_A g d\mathbb{P}. \tag{2.1}$$

The function g is referred to as the Radon-Nikodým derivative of ν w.r.t μ. It is written $g = \frac{d\mathbb{Q}}{d\mathbb{P}}$ or $d\mathbb{Q} = g d\mathbb{P}$.

Proof. Refer to Theorem 5 in [7]. □

3 Stochastic Processes and Lévy Processes

In this section, Lévy processes are introduced and properties associated with them defined. The theoretical aspects discussed in the section are drawn mostly from texts by Schoutens [47] and Sato [46] for more advanced concepts. A time horizon, $T < \infty$, and a probability space, $(\Omega, \mathcal{F}, \mathbb{P})$ is assumed. When $T \subseteq \mathbb{N}$ the stochastic process X is referred to as a stochastic process being in *discrete time.* When $T \subseteq \mathbb{R}$ then the stochastic process X is referred to as being in *continuous time.*

Definition 6. (Stochastic Process, Schoutens (2003) [47])

A *Stochastic Process,* $X = (X(t), 0 \leq t \leq T)$ adapted to the filtration $\{\mathcal{F}_t\}_{t \leq 0 \leq T}$ (i.e $X(t)$ is $\mathcal{F}_t - measurable$ for every $t \in T$) is a collection of random variables on $\Omega \times [0, T]$.

The definition above implies that $X(t)$ is known at time t. Two stochastic processes $X(t)$ and $Y(t)$ are identical in law if and only if $X(t)$ converges in distribution to $Y(t)$. This is written as:

$$X(t) \overset{d}{=} Y(t). \tag{3.1}$$

Definition 7. (Continuous in Probability, Benth et. al. (2006) [8])

A stochastic processes is *stochastically continuous in probability* if for every $t \geq 0$ and $\varepsilon \geq 0$,

$$\lim_{s \to t} P\left[|X(s) - X(t)| > \varepsilon\right] = 0 \tag{3.2}$$

Next, a class of stochastic processes called *Lévy Processes* with independent and stationary increments is defined. This class is of particular interest because of the its properties and practicality. In the most basic form, random motions such as the Brownian Motion and random jump processes such as the Poison Process are typical examples of Lévy processes.

Definition 8. (Lévy Process, Schoutens (2003) [47])

An $\mathcal{F}_t - adapted$, càdlàg (i.e one with sample paths that are a.s. right-continuous and have limits from the left), real-valued stochastic process, $X = (X(t), 0 \leq t \leq T)$, is called a Lévy process if the following conditions are true:

1. For any choice of $n \geq 1$ and $0 \leq t_0 < t_1 < ... < t_n$, the random variables $X(t_0),\ X(t_1) - X(t_0),\ X(t_2) - X(t_1),......, X(t_n) - X(t_{n+1})$ are independent (*independent increments property*);

2. $X(0) = 0$ a.s;

3. The distribution of $X(t+s) - X(s)$ does not depend on s (*stationary increments property*) ;

4. X is *stochastically continuous* i.e. for every $t \geq 0$ and $\varepsilon \geq 0$, $\lim_{s \to t} P\left[|X(s) - X(t)| > \varepsilon\right] = 0$.

Definition 9. (Martingale, Applebaum (2009) [1])

A stochastic process, $X = (X(t), 0 \leq t \leq T)$ is called a martingale relative to $(\mathbb{P}, \mathcal{F})$ if

1. X is $\mathcal{F}$-adapted;

2. $\mathbb{E}[|X|] < \infty$ for $t \geq 0$;

3. $\mathbb{E}[X(t)|\mathcal{F}_s] = X(s)$, a.s. $(0 \leq s \leq t)$.

Definition 10. (Infinitely Divisible, Applebaum (2009) [1])

Let X be a random variable taking value in $\mathbb{R}$ with law μ_X. X is infinitely divisible, if for all $n \in \mathbb{N}$, there exists i.i.d random variables $X_1, ..., X_n$ such that

$$X \stackrel{d}{=} X_1^{(n)} + .. + X_n^{(n)}. \tag{3.3}$$

If $X_i^{(n)}$ has the law $\mu_{X^{(n)}}$ then $\mu_X = \mu_{X^{(n)}} * ... * \mu_{X^{(n)}}$ the convolution of $\mu_{X^{(n)}}$ n times.

The definition above can be illustrated in a simple example. Consider, $X \sim N(\mu, \sigma^2)$, and i.i.d random variables $Y_i \sim N(\frac{\mu}{n}, \sigma^2/n)$. Then from the properties of a Normal distribution $\sum_{i=1}^{n} Y_i \sim N(\mu, \sigma^2)$. Thus $X \stackrel{d}{=} \sum_{i=1}^{n} Y_i$.

Definition 11. (Characteristic Function, Applebaum (2009) [1])

Let X be a random variable defined on $(\Omega, \mathcal{F}, \mathbb{P})$ and taking values in $\mathbb{R}$ with the probability law p_X. Then its characteristic function $\phi_X(u) : \mathbb{R} \to \mathbb{C}$ is defined by

$$\phi_X(u) = \mathbb{E}\left(e^{iuX}\right) = \int_{\Omega} e^{i(u, X(\omega))} \mathbb{P}(d\omega) = \int_{\mathbb{R}} e^{iuy} p_X(dy) \tag{3.4}$$

for each $u \in \mathbb{R}$.

Some useful properties of a characteristic function are [1]:

1. $|\phi_X(u)| \leq 1$;

2. $\phi_X(-u) = \phi_X(u)$;

3. X is symmetric if and only if $\phi_X(u) \in \mathbb{R}$;

4. If $\mathbb{E}\left(\left|X_j^n\right|\right) < \infty$ for $n \in \mathbb{N}$ then

$$\mathbb{E}\left(X^n\right) = i^{-n} \frac{d^n}{du^n} \phi_X(u)\bigg|_{u=0}. \tag{3.5}$$

Some functions closely related to the characteristic function are:

- Cumulant function $k(u) = log\mathbb{E}\left(e^{-uX}\right) = log\phi_X(iu)$,

- Moment generating function $\vartheta(u) = \mathbb{E}\left(e^{uX}\right) = \phi_X(-iu)$ and

- Cumulant characteristic function $\psi(u) = log\mathbb{E}\left(e^{iuX}\right) = log\phi_X(u)$

Proposition 1. *(Sato (1999) [46])*

If X is a Lévy process, then for each $t \geq 0$, the random variable $X(t)$ is infinitely divisible. Conversely, if M is an infinitely divisible random variable, then there exists a Lévy process X such that $X(1) \overset{d}{=} M$.

Proof. For each $n \in \mathbb{N}$ we can write

$$X(t) = Y_1(t) + ... + Y_n(t)$$

where each

$$Y_k(t) = X\left(\left(\frac{kt}{n}\right)\right) - X\left(\frac{k(n-1)}{n}\right)$$

The $Y_k(t)$ are i.i.d.   □

The proof of the converse statement is not provided for purposes of this dissertation. The reader is referred to Sato (1999) [46].

Recalling the definition a cumulant characteristic function often referred to as the characteristic exponent , $\psi(u) = log\mathbb{E}\left(e^{iuX}\right) = log\phi_X(u)$ we are led to the Lévy-*Khintchine Formula/Representation.*

Theorem 2. (*Lévy*-Khintchine Formula, Schoutens (2003) [47])

Let X be a Lévy process with the cumulant characteristic function ψ. If $\gamma \in \mathbb{R}$, $\sigma^2 \geq 0$ and ν is a measure on $\mathbb{R} \setminus \{0\}$ such that

$$\int\limits_{-\infty}^{+\infty} inf\left\{1, x^2\right\}\nu(dx) = \int\limits_{-\infty}^{+\infty} \left(1 \wedge x^2\right)\nu\left(dx\right) < \infty$$

then from the given triplet (γ, σ^2, ν), for each $u \in \mathbb{R}$, we define

$$\psi(u) = i\gamma u - \frac{1}{2}\sigma^2 u^2 + \int\limits_{-\infty}^{+\infty} \left(e^{iux} - 1 - iux\mathbb{I}_{\{|x|\leq 1\}}\right)\nu\left(dx\right) \qquad (3.6)$$

Conversely, given the triplet (γ, σ^2, ν), there exists a Lévy process X with the cumulant characteristic function given by the expression above.

The proof of the Theorem is not given for purposes of this dissertation. The formula

above implies that the characteristic function is given by

$$\phi_X(u) = exp\left\{i\gamma u - \frac{1}{2}\sigma^2 u^2 + \int\limits_{-\infty}^{+\infty}\left(e^{iux} - 1 - iux\mathbb{I}_{\{|x|\leq 1\}}\right)\nu\left(dx\right)\right\} \qquad (3.7)$$

When a Lévy measure is of the form $\nu(dx) = u(x)dx$, the function $u(x)$ is referred to as the *Lévy density.* The Lévy density has the same properties as that of a probability density with the exception of the integrability condition and that it must be on $\mathbb{R}\backslash\{0\}$ [47].

From the Lévy-Khintchine formula, one can observe that a Lévy process is characterised by three parts which are independent i.e a linear deterministic part, a diffusion part and a pure jump part [47]. In particular $\gamma \in \mathbb{R}$ is the drift part, $\sigma^2 \in \mathbb{R}^+$ is the Gaussian or diffusion coefficient and $\nu(dx)$ the Lévy measure.

4   Common Lévy Processes in Finance

In this section, some of the popular Lévy processes in finance are introduced. Infinitely divisible distributions need to model the skewness and excess kurtosis present in the distribution of the log-returns of most financial assets [47]. Properties of these Lévy distributions such as their characteristic functions, density functions and Lévy triplets are discussed as in Schoutens (2003) [47].

An understanding of these properties will help understand the *Background Driving Lévy Processes* (BDLP) or subordinators of *Ornstein-Uhlenbeck (OU) Processes* which were introduced by Barndorff-Nielsen and Shephard (B-NS) (2001a,b, 2003b) [5] to model stochastic volatility. The *Gamma, Inverse Gaussian (IG)* and *Normal Inverse Gaussian (NIG)* processes are discussed in the following sections and will

pave the pathway for their use in providing positive BDLPs for the OU process described later on.

Definition 12. (Subordinator, Applebaum (2009) [1])

A subordinator, $T := \{T(t), t \geq 0\}$, is a one-dimensional Lévy process that is non-decreasing (a.s.). This means that

$$T(t) \geq 0$$

a.s.

and

$$T(t_1) \leq T(t_2)$$

a.s. whenever $t_1 \leq t_2$.

The transformation of one stochastic process to another by an increasing Lévy process (subordinator) independent of the original process under a random time change is called *Subordination*. It is a transformation of a Lévy process to another independent Lévy process [46].

Theorem 3. (Subordination of Lévy Processes, Sato (1999) [46])

Let $Z := Z(t) : t > 0$ be a subordinator (an increasing Lévy process on $\mathbb{R}$) with a Lévy measure ν, drift β_0, and $P_{Z_1} = \lambda$. Then it follows that

$$E\left(e^{-uZ_t}\right) = \int_{[0,\infty)} e^{-us}\lambda^t(ds) = e^{t\psi(-u)}, \qquad u \geq 0 \tag{4.1}$$

where for any complex ω with $\mathfrak{R}(\omega) \leq 0$,

$$\psi(\omega) = \beta_0\omega + \int_{[0,\infty)}$$

with $\beta_0 \geq 0$ and $\int_{[0,\infty)} (1 \wedge s)\nu(ds) < \infty$.

Definition 13. (Subordination of Lévy Processes, page 198 [46])

Let $X := \{X(t)\}$ be Lévy process on $\mathbb{R}$ with generating triplet (A, ν, γ) and let $\mu = P_{X_1}$. Suppose X and Z are independent, then the transformation of X to Y,

$$Y_t(\omega) = X_{Z_t(\omega)}, \quad t \geq 0 \tag{4.2}$$

is called the subordination by the subordinator $Z(t)$.

4.1   Gamma Process

Any positive real-valued stochastic process $L := \{L(t), t \geq 0\}$ is called a Gamma process with parameters $a > 0$ and $b > 0$ if $L \sim \Gamma(at, b)$ that is L has the distribution given by

$$f(x; a, b) = \frac{b^{at}}{\Gamma(a)} x^{at-1} e^{-bx}, \qquad x > 0 \tag{4.3}$$

[47].

The parameters $a, b > 0$ are responsible for the tail thickness and spread/scale of the distribution f.

Lemma 1. (Characteristic Function of the Gamma Process, Schoutens [47])

The characteristic function of L is given by

$$\phi(u) = \left(1 - \frac{iu}{b}\right)^{-at}. \tag{4.4}$$

Proof. From the definition of a characteristic function it follows that

$$
\begin{aligned}
\phi(u; a.b) &= E\left(e^{iuX}\right) \\
&= \int_0^\infty e^{iux} \frac{b^a}{\Gamma(a)} x^{a-1} e^{-bx} dx \\
&= \int_0^\infty \frac{b^a}{\Gamma(a)} x^{a-1} e^{iux-bx} dx \\
&= \int_0^\infty \frac{b^a}{\Gamma(a)} x^{a-1} e^{-b(x-\frac{iux}{b})} dx
\end{aligned}
$$

the if we set $y = x - \frac{iux}{b}$ it follows that

$$
\begin{aligned}
&= \int_0^\infty \frac{b^a}{\Gamma(a)} \left(\frac{by}{b-iu}\right)^{a-1} e^{-by} \left(\frac{b}{b-iu}\right) dy \\
&= \left(\frac{b}{b-iu}\right)^a \frac{b^a}{\Gamma(a)} y^{a-1} e^{-by} dy \\
&= \left(1 - \frac{iu}{b}\right)^{-a}.
\end{aligned}
$$

Then it follows that

$$
\phi(u; at, b) = \left(1 - \frac{iu}{b}\right)^{-at}.
$$

$\square$

If $X \sim Gamma(at, b)$ is a random variable with its characteristic function as shown in the Lemma above, a random variable $X_n^{(n)}$ with distribution, F_n, which is a $Gamma(\frac{at}{n}, b)$ implies that

$$
\left(\phi(u; (\frac{a}{n})t, b)\right)^n = \left(\left(1 - \frac{iu}{b}\right)^{-(\frac{a}{n})t}\right)^n = \phi(u; at, b). \tag{4.5}
$$

It follows that the Gamma Process is infinitely divisible. The Lévy triplet of a

Gamma(at, b) process is given by [47]

$$\left[\frac{at(1 - e^{-b})}{b}, 0, ate^{-bx}x^{-1}\mathbb{I}_{(x>0)}dx \right].$$

(4.6)

Its characteristic exponent can be written as

$$\psi(u) = i\left(\frac{at(1 - e^{-b})}{b} \right) u + \int_0^\infty \left(e^{iux} - 1 - iux\mathbb{I}_{\{|x| \le 1\}} \right) ate^{-bx}x^{-1}\mathbb{I}_{(x>0)}dx.$$

(4.7)

For the derivation of the triplet above refer to Corollary 8.9 in Sato (1999) [46]. In this case the distribution of *Gamma*$(\frac{at}{n}, b)$ is

$$F_n(dx) = \frac{b^{\frac{a}{n}t}}{\Gamma(\frac{at}{n})} x^{\frac{at}{n} - 1} e^{-bx} dx$$

(4.8)

so that for $x > 0$ if we set $\gamma_n = n \int_{|x|<1} x F_n(dx) := \gamma$ then the limit of as n tends to

infinity is

$$
\begin{aligned}
\lim_{n\to\infty} \gamma_n &= \lim_{n\to\infty} n \int_0^1 x F_n(dx) \\
&= \lim_{n\to\infty} n \int_0^1 x \frac{b^{\frac{a}{n}t}}{\Gamma(\frac{at}{n})} x^{\frac{at}{n}-1} e^{-bx} dx \\
&= \lim_{n\to\infty} n \int_0^1 \frac{b^{\frac{a}{n}t}}{\Gamma(\frac{at}{n})} x^{\frac{at}{n}} e^{-bx} dx \\
&= \lim_{n\to\infty} \int_0^1 \frac{atb^{\frac{a}{n}t}}{\Gamma(\frac{at}{n}+1)} x^{\frac{at}{n}} e^{-bx} dx \\
&= at \int_0^1 \lim_{n\to\infty} \left(\frac{b^{\frac{a}{n}t}}{\Gamma(\frac{at}{n}+1)} x^{\frac{at}{n}} e^{-bx} \right) dx \\
&= at \int_0^1 e^{-bx} dx \\
&= at \left(\frac{1-e^{-b}}{b} \right).
\end{aligned}
$$

Similarly for $x > 0$, if we set $G_n(a) = n \int_{-\infty}^a \frac{x^2}{1+x^2} F_n(dx)$, then taking the limit as n tends to infinity :

$$
\begin{aligned}
\lim_{n\to\infty} n F_n(x) &= \lim_{n\to\infty} n \frac{b^{\frac{a}{n}t}}{\Gamma(\frac{at}{n})} x^{\frac{at}{n}-1} e^{-bx} dx \\
&= at x^{-1} e^{-bx} dx.
\end{aligned}
$$

The jump which G makes at $x = 0$ is $\sigma^2 = \triangle G(0)$.

4.2   The Inverse Gamma Process

The probability density function of an Inverse Gaussian (IG) distribution with parameters $a > 0$ and $b > 0$ is given by [47]:

$$f(x; a, b) = \frac{a}{\sqrt{2\pi}} e^{ab} x^{-\frac{3}{2}} e^{\left(-\frac{1}{2}\left(a^2 x^{-1} + b^2 x\right)\right)} \quad x > 0 \tag{4.9}$$

If $L \sim IG(a, b)$, then for a positive constant c, $cL \sim IG\left(\sqrt{c}a, \frac{b}{\sqrt{c}}\right)$.

Lemma 2. *(Sato (1999) [46] and Schoutens (2003) [47])*

For $u \in \mathbb{R}$, the characteristic function of the $IG(a, b)$ is :

$$\phi(u; a.b) = E\left(e^{iuX}\right) = e^{\left(-a\left(\sqrt{-2iu + b^2} - b\right)\right)} \tag{4.10}$$

Proof. Consider the Laplace transform $F(s)$ that is:

$$
\begin{aligned}
F(s) &= \mathbb{E}(e^{-sX}) \\
&= \int_0^\infty e^{-sx} \frac{a}{\sqrt{2\pi}} e^{ab} x^{-\frac{3}{2}} e^{\left(-\frac{1}{2}\left(a^2 x^{-1} + b^2 x\right)\right)} dx \\
&= \int_0^\infty \frac{a}{\sqrt{2\pi}} e^{ab} x^{-\frac{3}{2}} e^{\left(-\frac{1}{2}\left(a^2 x^{-1} + b^2 x + 2sx\right)\right)} dx \\
&= \int_0^\infty \frac{a}{\sqrt{2\pi}} e^{\left(a\sqrt{2s+b^2} - a\sqrt{2s+b^2}\right) + ab} x^{-\frac{3}{2}} e^{\left(-\frac{1}{2}\left(a^2 x^{-1} + (2s+b^2)x\right)\right)} dx \\
&= \int_0^\infty \frac{a}{\sqrt{2\pi}} e^{a\sqrt{2s+b^2} - a\left(\sqrt{2s+b^2} - b\right)} x^{-\frac{3}{2}} e^{\left(-\frac{1}{2}\left(a^2 x^{-1} + (2s+b^2)x\right)\right)} dx \\
&= e^{-a\left(\sqrt{2s+b^2} - b\right)} \int_0^\infty \frac{a}{\sqrt{2\pi}} e^{a\sqrt{2s+b^2}} x^{-\frac{3}{2}} e^{\left(-\frac{1}{2}\left(a^2 x^{-1} + (2s+b^2)x\right)\right)} dx \\
&= e^{-a\left(\sqrt{2s+b^2} - b\right)} . 1 \\
&= e^{-a\left(\sqrt{2s+b^2} - b\right)} .
\end{aligned}
$$

In familiar notation, setting $u := s$, it follows that

$$\phi(u; a.b) = \mathbb{E}\left(e^{iuX}\right) = F(-iu) = e^{-a\left(\sqrt{-2iu+b^2}-b\right)}$$

$\square$

From the expression above, one can easily observe that $\phi(u; a.b) = \phi(u; \frac{a}{n}.b)^n$ hence , the Inverse Gaussian distribution is infinitely divisible.

Proposition 2. *(Sato (1999) [46])*

The Lévy triplet of an $IG(a, b)$ process is derived in a process as the one for a Gamma distribution as:

$$\left[\frac{a}{b}\left(2\Phi(b) - 1\right), 0, \frac{a}{\sqrt{2\pi x^3}}e^{\left(-\frac{1}{2}\left(b^2 x\right)\right)}\mathbb{I}_{(x>0)}dx\right]. \tag{4.11}$$

Proof. The distribution of $IG(\frac{a}{n}, b)$ is

$$F_n(dx) = \frac{a}{n\sqrt{2\pi}}e^{\frac{a}{n}b-1}x^{-\frac{3}{2}}e^{-\frac{1}{2}\left(a^2 x^{-1}+b^2 x\right)}dx \tag{4.12}$$

For $x > 0$

$$
\begin{aligned}
\lim_{n\to\infty} \gamma_n &= \lim_{n\to\infty} n \int_0^1 x F_n(dx) \\
&= \lim_{n\to\infty} n \int_0^1 x \frac{a}{n\sqrt{2\pi x^3}} e^{\frac{a}{n}b-1} e^{-\frac{1}{2}\left(a^2 x^{-1}+b^2 x\right)} dx \\
&= \frac{a}{\sqrt{2\pi}} \int_0^1 x \frac{1}{\sqrt{x^3}} e^{\left(-\frac{1}{2}\left(\frac{a/n}{\sqrt{x}}+b\sqrt{x}\right)^2\right)} \\
&= \frac{a}{\sqrt{2\pi}} \int_0^1 x^{-\frac{1}{2}} e^{\left(\frac{-b^2 x}{2}\right)} dx \\
&= \frac{a}{b} \frac{1}{\sqrt{2\pi}} \int_0^{b^2} y^{-\frac{1}{2}} e^{\left(\frac{-y}{2}\right)} dy \\
&= \frac{a}{b} \left(Y^2 \le b^2\right) \\
&= \frac{a}{b} \left(Y \le b\right) \\
&= \frac{a}{b} \left(2\Phi(b) - 1\right).
\end{aligned}
$$

Where $Y \sim N(0,1)$ and $Y^2 \sim \chi^2(1)$.

For $a > 0$

$$
\begin{aligned}
\lim_{n\to\infty} G_n(a) &= \lim_{n\to\infty} n \int_0^a \frac{x^2}{1+x^2} F_n(dx) \\
&= \lim_{n\to\infty} n \int_0^a \frac{x^2}{1+x^2} \frac{a}{n\sqrt{2\pi x^3}} e^{\frac{a}{n}b-1} e^{-\frac{1}{2}\left(a^2 x^{-1}+b^2 x\right)} dx \\
&= \lim_{n\to\infty} \int_0^a \frac{x^2}{1+x^2} \frac{a}{\sqrt{2\pi x^3}} e^{\left(-\frac{1}{2}\left(\frac{a/n}{\sqrt{x}}+b\sqrt{x}\right)^2\right)} dx \\
&= \int_0^a \frac{x^2}{1+x^2} \frac{a}{\sqrt{2\pi x^3}} e^{\left(-\frac{1}{2}\left(b^2 x\right)\right)} dx = G(a).
\end{aligned}
$$

For $a \ge 0$ it follows that $\lim_{n\to\infty} G_n(a) = G(a) \equiv 0$. Since G is continuous at $a = 0$, it implies that there is no Gaussian component.

Moreover for $x > 0$,

$$
\begin{aligned}
\lim_{n \to \infty} n F_n(x) &= \lim_{n \to \infty} n \frac{a/n}{\sqrt{2\pi x^3}} e^{\left(-\frac{1}{2}\left(\frac{a/n}{\sqrt{x}} + b\sqrt{x}\right)^2\right)} \\
&= \lim_{n \to \infty} \frac{a}{\sqrt{2\pi x^3}} e^{\left(-\frac{1}{2}\left(\frac{a/n}{\sqrt{x}} + b\sqrt{x}\right)^2\right)} \\
&= \frac{a}{\sqrt{2\pi x^3}} e^{\left(-\frac{1}{2}b^2 x\right)} := \nu(dx).
\end{aligned}
$$

The IG process $L = \{L_t : t \geq 0\}$ is a càdlàg process i.e one with independent and stationary increments and $L_t \sim IG(at, b)$. Following the discussion on page 53 of [47], if we let $L^{(a,b)}$ be the first time a standard Brownian motion with drift $b > 0$, $X(t) := \{B(t) + bt, t \geq 0\}$, reaches the positive level $a > 0$, then this random time follows the $IG(a, b)$ distribution. This first passage time defined on $(0, \infty)$ and distributed according to an $IG(a, b)$ distribution can be expressed as:

$$
L^{(a,b)} = \inf \{t > 0, X(t) = a\} \tag{4.13}
$$

$\square$

4.3   The Normal Inverse Gaussian Process

The Normal Inverse Gaussian (NIG) distribution (see [4]) has a probability density function with parameters $a, \delta > 0$ and $\beta < |a|$ given by:

$$
f(x; a, \beta, \delta) = \frac{a\delta}{\pi} e^{\left(\delta\sqrt{a^2 - \delta^2} + \beta x\right)} \frac{K_1\left(a\sqrt{\delta^2 + x^2}\right)}{\sqrt{\delta^2 + x^2}} \tag{4.14}
$$

where $K_1(x)$ is the modified Bessel function of the third kind.

Lemma 3. *(Barndorff-Nielsen and Shepard (2001) [6])*

For $u \in \mathbb{R}$, the characteristic function of the $NIG(a, \beta, \delta)$ is :

$$\phi(u) = E\left(e^{iuX}\right) = e^{\left(-\delta\left(\sqrt{a^2-(\beta+iu)^2}-\sqrt{a^2-\beta^2}\right)\right)}. \tag{4.15}$$

One can easily observe that $\phi(u; a, \beta, \delta) = \phi(u; a, \beta, \delta/n)^n$, which implies that the NIG distribution is infinitely divisible. The NIG process, X, can be defined by defined by:

$$X := \{X(t) : t \geq 0\}$$

with $X(t) \sim NIG(a, \beta, dt)$. The NIG process has no Gaussian component and its Lévy triplet is given by [47]:

$$\left[\frac{2a\delta}{\pi} \int\limits_0^1 sinh\left(\beta x\right) K_1\left(ax\right), 0, \frac{a\delta}{\pi} e^{\beta x} \frac{K_1\left(a\,|x|\right)}{|x|} dx\right]. \tag{4.16}$$

The NIG process can be related to an Inverse Gaussian time-changed Brownian motion . If $W = \{W(t) : t \geq 0\}$ is the standard Brownian motion and $I = \{I(t) : t \geq 0\}$ an $IG(a, b)$ where $a = 1$ and $b = \delta\sqrt{\alpha^2 - \beta^2}$ with $\alpha > 0$, $-\alpha < \beta < \alpha$ and $\delta > 0$. The stochastic process

$$X(t) = \beta\delta^2 I(t) + \delta W_{I(t)} \tag{4.17}$$

is a $NIG(\alpha, \beta, \delta)$. Furthermore, if $X \sim NIG(\alpha, \beta, \delta)$, then $-X \sim NIG(\alpha, -\beta, \delta)$. When $\beta = 0$, the distribution is symmetric [47].

4.4 The Generalised Hyperbolic Process

The Generalised Hyperbolic distribution $GH(\alpha, \beta, \delta, \nu)$ has a p.d.f defined by [4]:

$$f(x) = a(\alpha, \beta, \delta, \nu)\left(\delta^2 + x^2\right)^{\left(\nu - \frac{1}{2}\right)/2} K_{\nu - \frac{1}{2}}\left(\delta\sqrt{\alpha^2 - \beta^2}\right) e^{\beta x}, \tag{4.18}$$

where

$$a\left(\alpha, \beta, \delta, \nu\right) = \frac{\left(\alpha^2 - \beta^2\right)^{\nu/2}}{\sqrt{2\pi}\alpha^{\nu-\frac{1}{2}}\delta^\nu K_\nu\left(\delta\sqrt{\alpha^2 - \beta^2}\right)}$$

and

$\delta \geq 0, |\beta| < \alpha$ if $\nu > 0$;

$\delta > 0, |\beta| < \alpha$ if $\nu = 0$;

$\delta > 0, |\beta| \leq \alpha$ if $\nu < 0$

where $K_v(z)$ is the modified Bessel function of the third kind. The characteristic function of the Generalised Hyperbolic distribution $GH\left(\alpha, \beta, \delta, \nu\right)$ originally defined in Barndorff-Nielsen (1977) [3] is given by:

$$\phi\left(u\right) = \frac{\left(\alpha^2 - \beta^2\right)^{\nu/2}}{\left(\alpha^2 - (\beta + iu)^2\right)} \frac{K_\nu\left(\delta\sqrt{\alpha^2 - (\beta + iu)^2}\right)}{K_\nu\left(\delta\sqrt{\alpha^2 - \beta^2}\right)}. \tag{4.19}$$

The Generalised Hyperbolic distribution is proved to be infinitely divisible in Barndorff-Nielsen and Halgreen [3] . This implies that a Generalised Hyperbolic Lévy process $X(t)$ can be defined. The sophisticated Lévy measure of the GH process is defined as [47]:

$$\nu\left(dx\right) = \begin{cases} \frac{exp(\beta x)}{|x|}\left(\int_0^\infty \frac{exp\left(-|x|\sqrt{2y + \alpha^2}\right)}{\pi^2 y\left(J_\nu^2\left(\delta\sqrt{2y}\right) + N_\nu^2\left(\delta\sqrt{2y}\right)\right)}dy + \nu exp\left(-\alpha|x|\right)\right) & , \nu \geq 0 \\ \frac{exp(\beta x)}{|x|}\int_0^\infty \frac{exp\left(-|x|\sqrt{2y + \alpha^2}\right)}{\pi^2 y\left(J_{-\nu}^2\left(\delta\sqrt{2y}\right) + N_{-\nu}^2\left(\delta\sqrt{2y}\right)\right)}dy & , \nu < 0 \end{cases} \tag{4.20}$$

where $J_\nu(z)$ and $N_\nu(z)$ are Bessel functions of the first and second kind respectively. These are defined as:

$$J_\nu(z) = \left(\frac{z}{2}\right)^\nu \sum_{k=0}^{\infty} \frac{\left(-z^2/4\right)^k}{k!\Gamma\left(\nu + k + 1\right)} \tag{4.21}$$

and

$$N_v(z) = \frac{J_\nu(z)cos\left(\nu\pi\right) - J_{-\nu}(z)}{sin\left(\nu\pi\right)} \tag{4.22}$$

Special cases of the the Generalised Hyperbolic process are [47]:

1. The Variance Gamma process. This process can be derived from the GH by taking $\nu = \sigma^2/\nu^{VG}$, $\alpha = \sqrt{(2/\nu^{VG}) + (\theta^2/\sigma^4)}$, $\beta = \theta/\sigma^2$ and $\delta \to 0$.

2. The Normal Inverse Gaussian process. For $\nu = -\frac{1}{2}$ the NIG process is obtained. $GH\left(\alpha, \beta, \delta, -\frac{1}{2}\right) = NIG\left(\alpha, \beta, \delta\right)$.

4.5   Poisson Processes and Compensated Poisson Process

The *Poisson process* of mean/intensity $\lambda > 0$ is the stochastic process $N = \{N(t) : t \geq 0\}$ whose values lie in the set $\mathbb{N} \cup \{0\}$ such that each $N(t) \sim Poi\left(\lambda t\right)$ [1] that is:

$$P\left(N(t) = n\right) = \frac{e^{-\lambda t}\left(\lambda t\right)^n}{n!}N \tag{4.23}$$

for $n \in \{0, 1, 2, ...\}$. The characteristic function of the Poisson process is given by:

$$\phi(u) = e^{\left(\lambda t\left(e^{iu} - 1\right)\right)}. \tag{4.24}$$

From this, it can be easily observed that $\phi(u; \lambda) = \phi(u; \lambda/n)^n$. This implies that the Poisson process is infinitely divisible.

The *Compensated Poisson Process* is the stochastic process $\tilde{N} = \left\{ \tilde{N}(t) : t \geq 0 \right\}$ [47]. The Compensated Poisson process is defined as:

$$\tilde{N}(t) = N(t) - \lambda t \tag{4.25}$$

It is noted that $\mathbb{E}\left[\tilde{N}(t)\right] = 0$ and $\mathbb{E}\left[\left(\tilde{N}(t)\right)^2\right] = \lambda t$.

4.6   The Compound Poisson Process

Suppose that (Z_n) where $n \in \mathbb{N}$ is a sequence of i.i.d random variables with common law μ_Z and let a Poisson process $N(t) \sim Poi(\lambda t)$ independent of all Z_n. The Compound Poisson process Y_t is defined as [1]:

$$Y_t = Z_1 + Z_2 + ... + Z_{N(t)}$$

for all $t \geq 0$.

5   Ornstein-Uhlenbeck Lévy processes

In this section the properties of a family of Lévy processes called Ornstein-Uhlenbeck (OU) processes used to describe the dynamics of volatility in finance are discussed. The OU process was first suggested by Barndorff-Nielsen and Shepard (2001) [5] which in this dissertation is referred to as B-NS or in some instances B-NS. The

general form of the OU process $X(t) \geq 0$ is

$$dX(t) = -\lambda \left(X(t) - \alpha \right) dt + \sigma dZ(t), \qquad X(0) > 0 \qquad (5.1)$$

with

$$Z(0) = 0$$

where $\lambda > 0$, $\alpha \in \mathbb{R}$, $\sigma > 0$ and $Z(t)$ is a background driving Lévy process (BDLP) i.e. a Lévy process which has independent and stationary increments. B-NS (2001) focus on Lévy-driven processes known as non-Gaussian Ornstein-Uhlenbeck processes in which the BDLP, $Z(t)$, has positive increments and no Brownian part. These BDLPs are referred to as *subordinators* see (13). Since, the process $X(0) > 0$ and $Z(t)$ is an increasing process, it is clear that the process $X(t)$ is strictly positive [47]. If $Z(t)$ is a Brownian motion, then $X(t)$ is the usual Gaussian OU process.

5.1   D-OU and OU-D processes

The process $X := X(t) : t \geq 0$ is strictly stationary on the positive half-line that is there exists a law D, called the stationary law or the marginal law, such that $X(t)$ follows the law D for all t if the initial $X(0)$ is chosen according to D. Given a one-dimensional distribution D (not necessarily restricted to the positive half-line) , there exists a (stationary) OU process whose marginal law is D (i.e. a $D - OU$ process) if and only if D is self-decomposable [47]. The processes in the previous section can be used as D. According to Barndorff-Nielsen and Shephard's results (2001), it follows that if $\alpha = 0$ in Equation (5.1) then the solution is:

$$X(t) = e^{-\lambda t} X(0) + \int_0^t e^{-\lambda(t-s)} dZ_{\lambda s} \qquad (5.2)$$

$$= e^{-\lambda t} X(0) + e^{-\lambda t} \int_0^{\lambda t} e^s dZ_s \tag{5.3}$$

The process X is bounded below by $e^{-\lambda t} X(0)$ and is strictly stationary on the positive half-line. Let the Lévy measure of Z_1 be denoted by W. It is assumed that W has density $w(x)$. If the $u(x)$ is the Lévy density of D then $u(x)$ and $w(x)$ are related by [5]:

$$w(x) = -u(x) - xu'(x). \tag{5.4}$$

Let the tail mass function of $W(dx)$ be

$$W^+(x) = \int_x^\infty w(y)dy \tag{5.5}$$

then it follows from [6] that:

$$W^+(x) = xu(x) \tag{5.6}$$

and the the inverse of the function above is:

$$W^{-1}(x) = inf \left\{ y > 0 : W^+(y) \le x \right\} \tag{5.7}$$

Definition 14. (Self-decomposability, Schoutens (2002) [47])

Let ϕ_X be the characteristic function of a random variable X. Then X is self-decomposable if there exists a random variable Y_c independent of X such that:

$$\phi_X(u) = \phi_X(cu)\phi_{Y_c}(u) \tag{5.8}$$

for all $u \in \mathbb{R}$ and all $c \in (0,1)$ for some family of characteristic functions $\{\phi_c(u) : c \in (0,1)\}$.

Furthermore, the law of X belongs to the class of Lévys called L. A random variable with law in L is infinitely divisible.

5.1.1 Simulation via series representation

To simulate the OU process described in Equation (5.2), concentration is set on the Lévy integral. One approach could be to simulate the integrals directly, but this would be difficult due to the jumps present in the process. An estimate of the integral via infinite series representations is instead used [6]. If W is a Lévy measure for Z_1 and W^{-1} the inverse function for W^+, the series representation is given by:

$$\int_0^t g(s)dZ_s \overset{\mathcal{L}}{=} \sum_{i=1}^{\infty} W^{-1}\left(\frac{a_i}{t}\right) g\left(tu_i\right) \tag{5.9}$$

where $\{a_i\}$ and $\{u_i\}$ are two independent sequences of random variables with u_i independent replicas of a uniform $U(0,1)$ random variable and $a_1 < a_2 < ...$ the arrival times of a Poisson process with intensity 1 [47].

5.1.2 Simulating the IG process

To simulate the Inverse Gaussian process $IG(a,b)$ random variates can be used. The IG random variates are generated using the following algorithm [47]:

1. Generate a standard Normal random number v.

2. Set $y = \nu^2$ i.e $y \sim \chi^2(1)$.

3. Set $x = \left(\frac{a}{b}\right) + \frac{y}{2b^2} + \frac{\sqrt{4aby+y^2}}{2b^2}$.

4. Generate a uniform random number u.

5. If $u \leq \frac{a}{(a+xb)}$, then return the number x as the $IG(a,b)$ random number, else return $\frac{a^2}{b^2 x}$ as the $IG(a,b)$ random number.

Now to simulate the path of a process $I(t) = \{I(t) : t \geq 0\}$ an $IG(at,b)$ is simulated at time points $n \triangle t : n = 0, 1, ...$ as follows. First generate independent $IG(a \triangle t, b)$ random numbers $i_n; n \geq 1$, then

$$I(0) = 0$$

and

$$I(n\triangle t) = I((n-1)\triangle t) + i_n \tag{5.10}$$

for $n \geq 1$.

5.1.3   The NIG process

The NIG process can be simulated as time-changed Brownian motion. The $NIG\,(\alpha, \beta, \delta)$ process, $X(t) = \{X(t), t \geq 0\}$, can be represented as in Equation (4.17):

$$X(t) = \beta \delta^2 I(t) + \delta W_{I(t)} \tag{5.11}$$

where $I(t) \sim IG\left(t, \delta \sqrt{\alpha^2 - \beta^2}\right)$ for $\alpha, \delta > 0$, $|\beta| < \alpha$ and $W_{I(t)}$ is a Brownian motion.

5.1.4   The IG-OU process

The IG-OU process with $IG(a,b)$ marginals can be simulated via series representations in Equation (5.9). There is no explicit closed-form expression for the inverse W^{-1} of the tail mass function [47]. Thus the following approximation is made:

$$W^{-1}(x) \sim \frac{a^2}{2\pi x^2} \tag{5.12}$$

It should be noted that convergence can be slow using this approximation.

As a sum of two independent processes   The Lévy density of $IG(a,b)$, $u(x)$, is given by [4, 47]:

$$u(x) = \frac{a}{\sqrt{2\pi \left(\sqrt{x}\right)^3}} e^{\left(-\frac{1}{2}b^2 x\right)} \tag{5.13}$$

Then the Lévy density of the corresponding BDLP, $w(x)$, is given by [4]:

$$w(x) = \frac{a}{2\sqrt{2\pi}} \left(\frac{1}{x} + b^2\right) x^{-\frac{1}{2}} e^{\left(-\frac{1}{2}b^2 x\right)} \tag{5.14}$$

From [4], it is derived that the BDLP, $Z(t)$, in Equation (5.2) of the $IG(a,b) - OU$ process is a sum of two independent processes, $Z(t)^1$ and $Z(t)^2$ i.e.

$$Z(t) = Z(t)^1 + Z(t)^2, \tag{5.15}$$

where

$$Z(t)^1 \sim IG\left(\frac{a}{2}, b\right) \tag{5.16}$$

and $Z(t)^2$ is of the form:

$$Z(t)^2 = \frac{1}{b^2} \sum_{n=1}^{N_t} c_n^2, \tag{5.17}$$

where $N = \{N(t) : t \geq 0\}$ is a Poisson process with intensity $\frac{ab}{2}$ i.e $N(t) \sim Poi\left(\frac{abt}{2}\right)$ and $\{c_n, n = 1, 2, ...\}$ is a sequence of i.i.d $N(0, 1)$ random variables independent of the Poisson process N.

6   Finance

In this section, the finance theory in the valuation of contingent claims is presented. This theory forms the building blocks for developing the most basic and well-known continuous-time, continuous variable stochastic process for stock prices.

6.1   The Black-Scholes (BS) market model and risk-neutral pricing

Consider the probability space $(\Omega, \mathcal{F}, \mathbb{P})$. If investors are allowed to continuously trade up to some fixed time horizon $T > 0$, then let $\{W(t) : 0 \leq t \leq T\}$ be a Brownian motion with $\{\mathcal{F}_t : 0 \leq t \leq T\}$ a filtration for this Brownian. Then under the Black-Scholes model, the stock prices $\{S(t) : 0 \leq t \leq T\}$ evolve according to [11]:

$$dS(t) = \mu S(t)dt + \sigma S(t)dW(t), \qquad 0 \leq t \leq T \tag{6.1}$$

where μ is the drift parameter and σ constant volatility. The solution of this SDE is given by:

$$S(t) = S(0)e^{\left(\mu - \frac{1}{2}\sigma^2\right)(T-t) + \sigma(W(T) - W(t))} \tag{6.2}$$

The Black-Scholes model assumes that the market is complete and that there exist a risk-less asset $B(t)$ whose dynamics at $t \geq 0$ are:

$$dB(t) = rB(t) \tag{6.3}$$

with initial condition

$$B(0) = 1$$

where r is the continuously compounded risk-free rate [11]. Since the BS market model is complete, there exists only one equivalent martingale and the market can be completely replicated. This is the statement of Second Fundamental theorem of asset pricing.

Theorem 4. (First Fundamental Theorem of Asset Pricing, Bjork (2009) [11])

Assume that there exists a risk-free asset, and the corresponding risk-free interest rate by R. Then the market is arbitrage free if and only if there exists a probability measure $\mathbb{Q}$ such that:

$$S(0) = \frac{1}{1+R}\mathbb{E}^{\mathbb{Q}}[S(T)] \tag{6.4}$$

This is the risk-neutral pricing formula.

Theorem 5. (Second Fundamental Theorem of Asset Pricing, Bjork (2009) [11])

Consider a market model in which there exists a risk-neutral martingale measure, then the market is complete if and only if the martingale measure is unique.

The Black-Scholes market model is complete thus only one equivalent martingale measure $\mathbb{Q}$ exists. Under this risk-neutral measure the stock price follows the geometric Brownian motion. The risk-neutral stock price process, S_t, has constant volatility σ and the drift μ is the replaced by the continuously compounded risk-free rate r [47]. Thus the stock price process follows:

$$S(t) = S(0)e^{\left(\mu - \frac{1}{2}\sigma^2\right)(T-t) + \sigma(W(T) - W(t))} \tag{6.5}$$

7   Empirical Data

In this section, the empirical data from the FTSE/JSE Top Index 40 over a 10 year period from April 2009 to May 2019 is investigated. The fit of the Black-Scholes model to the data is investigated from a practical point of view. The BS model assumptions amongst others include a friction-less market with no taxes, transitional costs, no constraints in stock holding or short selling, normal log returns and a constant volatility over time. The point of this section is to show how these assumptions are not consistent with market data observations. This data will be utilised throughout this dissertation unless specified.

The log-returns of the FTSE/JSE Top 40 index are shown in the figure below:

Figure 7.1: FTSE/JSE Top 40 log-returns
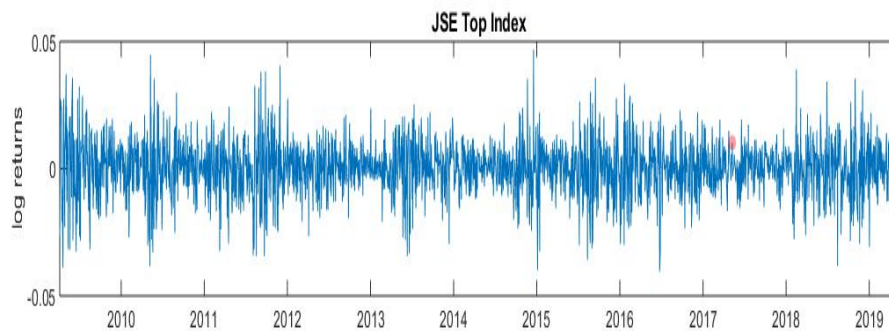
Next the closing prices of the index over the past decade are shown below:

Figure 7.2: FTSE/JSE Top 40 Price
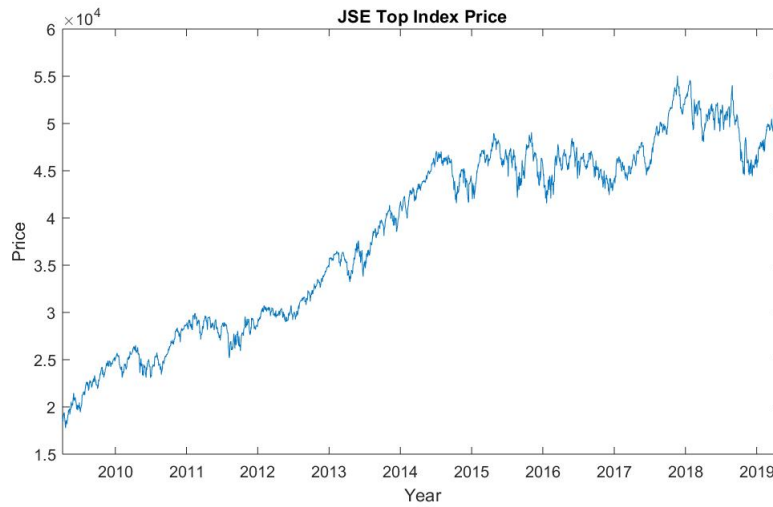
The figure 7 above shows the movement of the index price over the last decade. There are observable jumps in some periods.

7.0.1   Asymmetry and Excess Kurtosis

Table 1: Descriptive statistics of daily log-returns

No. of data points	Mean	Variance	SD	Skewness	Excess Kurtosis
2520	$3.8344{\times}10^{-4}$	$1.1660{\times}10^{-4}$	0.0108	-0.1348	4.2504
1000	$5.8122{\times}10^{-5}$	$1.1193{\times}10^{-4}$	0.0115	-0.1356	4.0184
500	$1.1964{\times}10^{-5}$	$9.9860{\times}10^{-5}$	0.0099	-0.1081	4.4812
250	$-7.2189{\times}10^{-5}$	$1.2810{\times}10^{-4}$	0.0113	-0.1272	3.7851

The data in the Table 1 above is observed from the empirical data. The moments of the daily log returns from 6 April 2009 to 10 May 2019 are compared over different time horizons.

Comments: The data is compared over different time horizons. The latest 1000 points, 500 points and 250 points by date are observed. The latest 250 days

are expected to provide more accurate estimates of the moments as they contain information relevant to the movements in that year.

Skewness measures the degree of a distribution's asymmetry and for a symmetric distribution like the $N(\mu = 0, \sigma^2 = 1)$, the skewness is 0 [47]. The data above is skewed although to a small degree. The market data for the JSE Top 40 Index shows that the excess kurtosis is greater than 3. This is more than the kurtosis observed for the Normal distribution.

The implication of modelling using a Normal distribution will be the inability to capture the tail distribution of the log returns. The data shows that the tails of the Normal distribution tend to zero quicker than the empirical data which in this case indicates a higher peaked distribution [1, 48, 44].

7.0.2 Empirical density function compared to the Normal density function

The study of the log returns of the JSE Top Index so far have shown that they do not fit a Normal distribution. However, further statistical analysis conducted to further support the preliminary observations. The Kolmogorov-Smirnov test is conducted as:

H_o: The index log returns follow a Normal distribution ;

H_1: The index log returns are not Normal distributed.

Using the MATLAB's *kstest* at 95% confidence level, we reject the null hypothesis and conclude that the log returns of the JSE Top 40 Index do not follow a Normal distribution. This is further illustrated graphically below using the Histogram and QQ-plot methods.

Figure 7.3: Histogram of the Log Returns Compared to the Normal Distribution
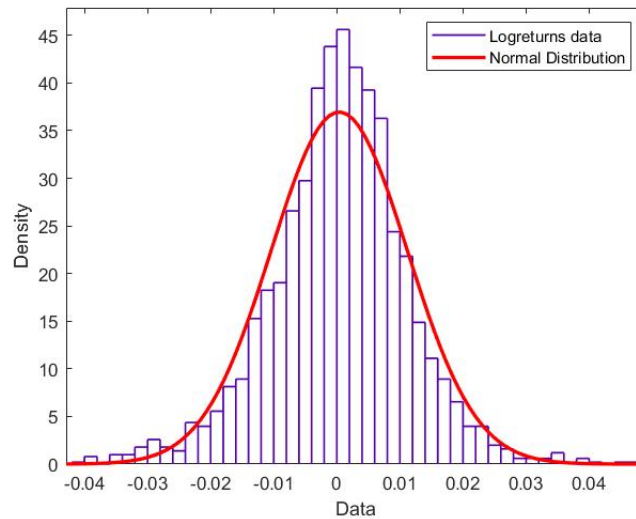
Figure 7.4: QQ Plot of the Empirical Data Compared to the Normal Distribution
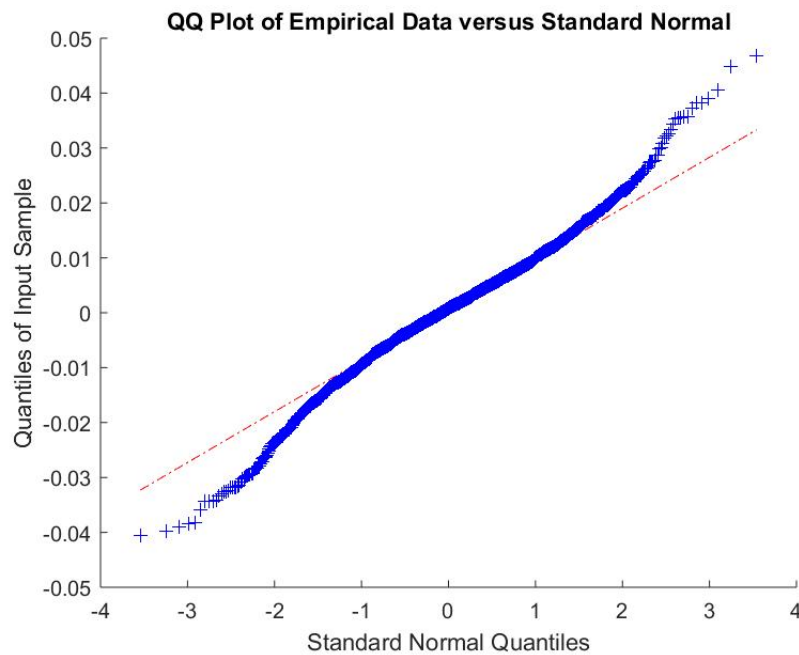
Comments: The figures above further support the Kolmogorov-Smirnov test. The QQ-plot shows that the log returns of the index do not fall perfectly on to the Normal distribution. The histogram shows that the Normal distribution has a lower peak

than the empirical distribution and it fails to capture the tail distribution as shown below:

Figure 7.5: The Normal Distribution underestimating the Tail Log Returns
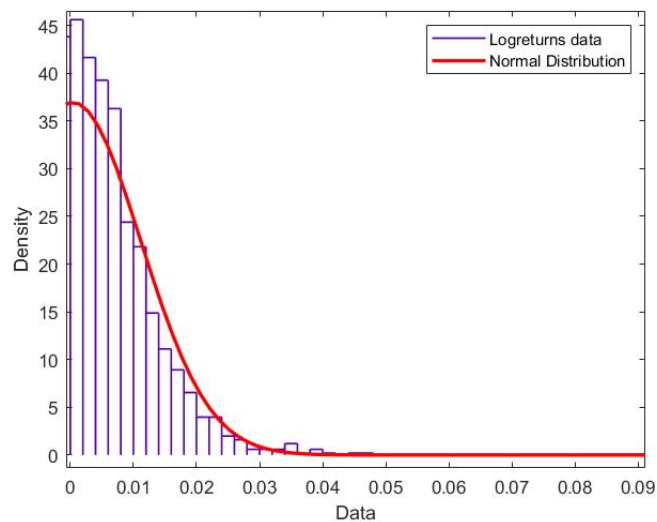
Comments: The Normal distribution does not capture the distribution of the peak and tail log returns accurately. In the figure above some tail log returns can be seen exceeding the Normal distribution density function. Benth et. al. (2006) [8] show that distribution of the tail and peak of log-returns can be better approximated by the Normal Inverse Gaussian (NIG) distribution. Using R's 'nigFit' function, a better fit is indicated. The QQ-plot below shows this:

Figure 7.6: QQ Plot of the Log Returns Compared to the NIG Distribution

Q-Q Plot of LogReturns
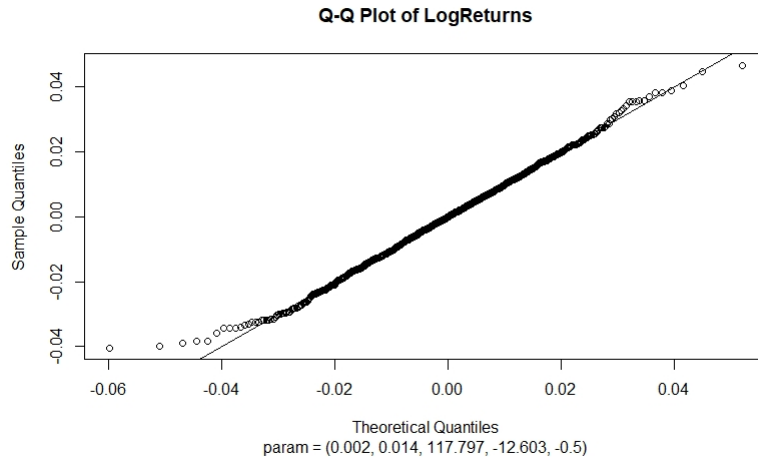
7.0.3   **Stochastic volatility**

Many researchers including [6, 15, 37, 38, 44] have shown that volatility is better approximated using models that show random change over time. The core of this dissertation is to use such models to price derivatives whose underlying is the square of volatility.

Historical volatility for the JSE Top 40 Index can be estimated by the standard deviation of the daily log returns over one year preceding the day. The historic volatility is calculated as an annualised value of the daily standard deviation by multiplying the standard deviation with the square root of 250 days.

Figure 7.7: JSE Top 40 Index Historical Volatility
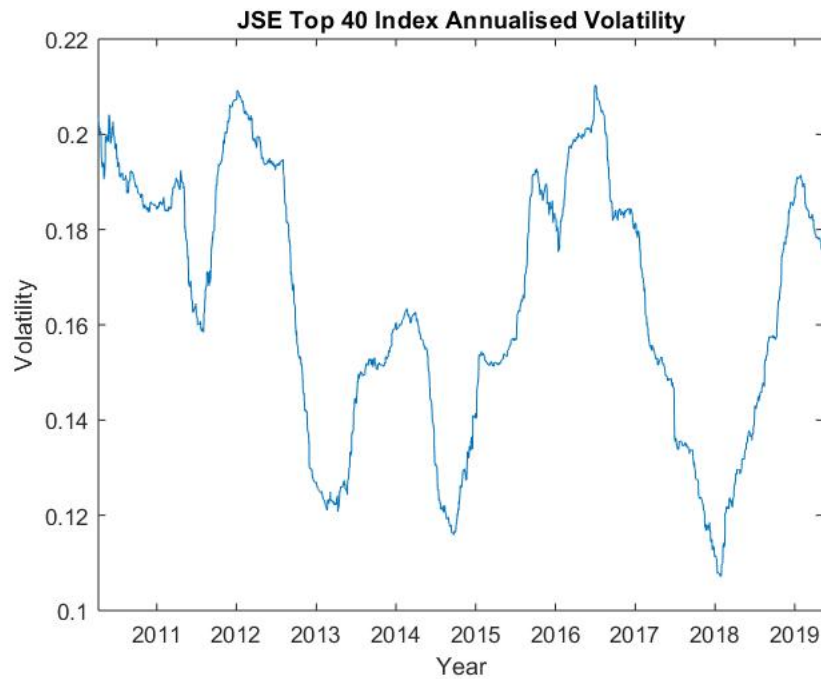
Comments: The figure above shows that the volatility of the index is stochastic over time. Furthermore, a mean-reversion effect is observed. This will be explained in the chapter in which the Heston (1993) stochastic volatility model is discussed.

7.0.4   Volatility clustering

Figure below shows that there is evidence for volatility clustering. There are groupings of periods with high returns and groupings of periods with low returns. This further supports the need for models that consider the stochastic pattern of volatility rather than assuming that it is constant over time.

Figure 7.8: JSE Top 40 absolute log returns
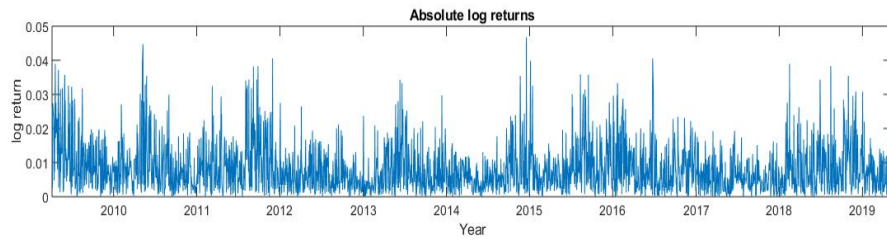
7.0.5 Market price error

If the model parameters are estimated by minimizing the root-mean-square error (RMSE) between market prices and BS model prices, a difference, is observed. This is known as the calibration error. The figure below shows the difference between the market call option prices and BS model prices for $T = 0.19047619$ (2-month data using a year with 250 days). The MATLAB's *lsqnonlin* function is used to calibrate the BS model to the market prices. The calibration techniques are shown in detail in a later chapter.

Figure 7.9: BS versus market European call option prices
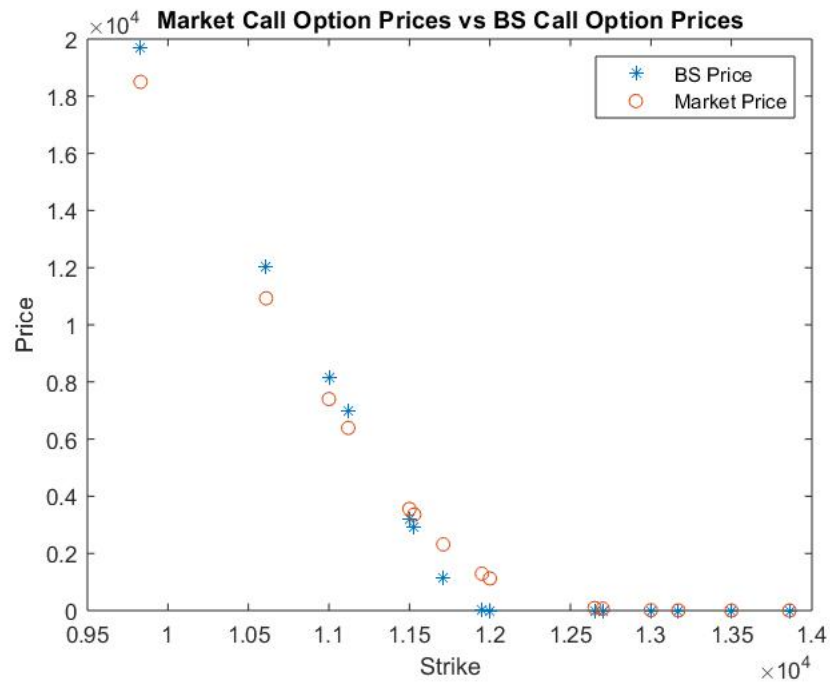
8   Statistical Considerations

In this section, methods for fitting Lévy distributions to the distribution of index price log returns are shown. The NIG distribution is used in this dissertation under the Barndorff-Nielsen and Shepard (2001) Ornstein-Uhlenbeck model. The parameters of this model are estimated using the method of moments estimation which is described in this section. The density function of the Lévy distribution is denoted $f(x;\theta)$ where θ represents the set of unknown parameters to be estimated. The log returns are then used to determine an acceptable set of these parameters under an assumption that they are i.i.d over non-overlapping periods in a Lévy setting [47].

8.0.1 Method of Moments Estimation

The method of moments estimation (MME) is a popular point estimation methodology which relies on the law of large numbers. Suppose $X_1, X_2, ...X_n$ are i.i.d random variables, then the kth population moment is given by [47]:

$$\mu_k' = \mathbb{E}\left(Y^k\right) = \int_0^\infty y^k f\left(y; \theta\right) dy \tag{8.1}$$

and the corresponding k^{th} sample moment is given by:

$$\hat{\mu}_k = \frac{1}{n} \sum_{i=1}^n Y_i^k. \tag{8.2}$$

It then follows that $\hat{\mu}_k$ in an unbiased estimator of μ_k' since $\mathbb{E}\left(\hat{\mu}_k\right) = \mathbb{E}\left(\frac{1}{n} \sum_{i=1}^n Y_i^k\right) = \frac{1}{n} \sum_{i=1}^n \mathbb{E}(Y_i^k) = \mu_k$. Then, the *Method of Moments estimator* for the unknown set of parameters which solves the equation:

$$\hat{\mu}_k = \mu_k'. \tag{8.3}$$

The kth moment can also be calculated from the characteristic function as [47]:

$$\mathbb{E}\left(X^k\right) = i^{-k} \frac{d^k}{du^k} \phi(u) \bigg|_{u=0} .$$

However, it may be a tedious task to find the derivatives of most characteristic functions.

8.0.2 Maximum Likelihood Estimation

The maximum-likelihood estimator (MLE) is a set of parameters θ that maximises the likelihood function as follows [2]:

$$L(\theta) = \prod_{i=1}^{n} f(x_i, \theta) \tag{8.4}$$

Since the log-function is monotonically increasing, maximising the function is equivalent to maximising the logarithm of the function. The log-likelihood is:

$$logL(\theta) = \sum_{i=1}^{n} log f(x_i, \theta) \tag{8.5}$$

Example 1. MLE for the Normal Inverse Gaussian distribution

The $NIG(\alpha, \beta, \delta, \mu)$ is investigated in a later chapter in this dissertation. Its pdf is known explicitly as [4, 44]:

$$f(x; \alpha, \beta, \delta, \mu) = \frac{\alpha\delta K_1\left(\alpha\sqrt{\delta^2 + (x-\mu)^2}\right)}{\pi\sqrt{\delta^2 + (x-\mu)^2}} e^{\delta\sqrt{\alpha^2 - \beta^2} + \beta(x-\mu)} \tag{8.6}$$

Then the likelihood function is given by:

$$L(\alpha, \beta, \delta, \mu) = \prod_{i=1}^{n} \frac{\alpha\delta K_1\left(\alpha\sqrt{\delta^2 + (x_i-\mu)^2}\right)}{\pi\sqrt{\delta^2 + (x_i-\mu)^2}} e^{\delta\sqrt{\alpha^2 - \beta^2} + \beta(x_i-\mu)} \tag{8.7}$$

$$= \frac{(\alpha\delta)^n \sum_{i=1}^{n} K_1\left(\alpha\sqrt{\delta^2 + (x_i-\mu)^2}\right)}{\pi^n \prod_{i=1}^{n} \sqrt{\delta^2 + (x_i-\mu)^2}} e^{n\delta\sqrt{\alpha^2 - \beta^2} + \beta \sum_{i=1}^{n}(x_i-\mu)}$$

which implies that the log-likelihood function is given by:

$$
\begin{aligned}
log\left(L(\alpha,\beta,\delta,\mu)\right) \;=\;& log\left((\alpha\delta)^n \sum_{i=1}^{n} K_1\left(\alpha\sqrt{\delta^2+(x_i-\mu)^2}\right)\right) \\
& -nlog\left(\pi\right) - \frac{1}{2}\sum_{i=1}^{n} log\left(\delta^2+(x_i-\mu)^2\right) + n\delta\sqrt{\alpha^2-\beta^2} \quad (8.8)\\
& +\beta\sum_{i=1}^{n}(x_i-\mu) \\
\;=\;& -nlog\left(\pi\right) + nlog\left(\alpha\delta\right) + \sum_{i=1}^{n} log\left(K_1\left(\alpha\sqrt{\delta^2+(x_i-\mu)^2}\right)\right)\\
& -\frac{1}{2}\sum_{i=1}^{n} log\left(\delta^2+(x_i-\mu)^2\right) + n\delta\sqrt{\alpha^2-\beta^2} + \beta\sum_{i=1}^{n}(x_i-\mu) \quad (8.9)
\end{aligned}
$$

Obtaining the partial derivatives of the log-likelihood function can be a cumbersome task since it contains the Bessel function. However, practically, optimisation of the NIG can be done easily using software.

9   Pricing Variance Swaps

In this section, the foundation of the theory behind the mechanism of a variance swap is laid out for future chapters. The variance swap is defined and various approaches to how variance swaps can be priced are discussed. In particular, the discretely-sampled and continuously-sampled variance swaps are defined.

Definition 15. (Variance Swap Payoff, Zhaoli et. al. (2015) [38])

A variance swap is a forward contract to exchange an agreed variance swap strike, K_{var}, for a future annualised realised variance, $\sigma_R^2(T)$, such that the payoff of at expiry is given by:

$$N(\sigma_R^2(T) - K_{var}) \tag{9.1}$$

for a given notional amount, N.

The definition above means that at maturity the holder of the variance swap contract receives an amount N for every unit by which the realised variance exceeds the variance swap price, K_{var}. Since variance swaps are forward contracts, there is no cost of entering the contract. Hence, from Asset Pricing Theory and the definition of a forward contract, it follows that the risk-neutral price of the variance swap (V_t) at $t = 0$ is given by [38]:

$$V_0 = 0 = e^{-rT}\mathbb{E}_{\mathbb{Q}}\left(N(\sigma_R^2(T) - K_{var})|\mathcal{F}_t\right) \tag{9.2}$$

for an equivalent martingale measure $\mathbb{Q}$.

Thus,

$$K_{var} = \mathbb{E}_{\mathbb{Q}}\left((\sigma_R^2(T)|\mathcal{F}_t\right) \tag{9.3}$$

where the notional amount has been set to $N = 1$.

This implies that to price a variance swap in the risk-neutral world, the initial problem can be reduced to calculating the expected future level of realised variance in a risk-neutral world. The expected realised variance can be obtained discretely or continuously. This gives rise to the discretely-sampled and continuously-sampled versions of variance swaps.

Definition 16. (Discretely-Sampled Expected Realised Variance, Little and Pant (2001) [41])

The expectation of the discretely-sampled level of realised variance of an underlying asset, S, at maturity is given by:

$$\mathbb{E}_{\mathbb{Q}}\left[(\sigma_R^2(T)\right] = \frac{1}{N\Delta t}\sum_{i=1}^{N}\mathbb{E}_{\mathbb{Q}}\left[\left(\frac{S(t_i) - S(t_{i-1})}{S(t_{i-1})}\right)^2\right].100^2 \qquad (9.4)$$

where $S(t_i)$ is the price of the underlying asset at time t_i.

Its continuously-sampled counterpart is defined as:

Definition 17. (Continuously-Sampled Expected Realised Variance, Benth. et al. (2006) [8])

The expectation of the continuously-sampled level of realised variance of an underlying asset, S, at maturity is given by:

$$\mathbb{E}_{\mathbb{Q}}\left[(\sigma_R^2(T)\right] = \frac{1}{T}\mathbb{E}_{\mathbb{Q}}\left[\int_0^T \sigma^2(s)ds\right].100^2. \qquad (9.5)$$

Now the fair price of variance swaps can be deduced. This leads to the next part of this dissertation.

Part II

Pricing variance swaps using stochastic volatility models

In this part, the pricing of variance swaps under the Bandorff-Neilsen and Shepard (2001) and Heston (1993) stochastic models is studied. The aim of this part is to derive semi-closed and closed-form expressions of pricing variance swaps under these models.

10 Variance swaps under the B-NS model

In this chapter, the semi-closed form expression for the price of continuously-sampled realised variance swaps under a stochastic model is derived. The B-NS non-Gaussian Ornstein-Uhlenbeck model where the log-returns are constructed using a mean-reverting stationary process with a background-driving Lévy process is implemented.

10.1 Introduction

Volumes of traded variance swaps have rapidly expanded since the first variance swaps were traded. This is an anticipated development as investors seek to be directly exposed to realised variance which gives them an idea of the market's perspective and hedges them against future variance fluctuations. Stochastic volatility models were born out of the deficiencies of the classical Black-Scholes (1973) model

including the constant volatility assumption, the volatility smile effect, and the kurtosis and skewness effects. Researchers in this light, are constantly trying to develop pricing techniques that accurately capture real market conditions. This chapter presents the movement of the price and variance processes as random independent processes. The instantaneous volatility in this chapter is considered to have a process as that under the non-Gaussian Ornstein-Uhlenbeck (OU) model with positive jumps only as presented in Barndoff-Nielsen and Shepard (2001). The analytical expression for the price of continuously-sampled realised variance swaps under the non-Gaussian OU model is derived. The work under Benth and Saltyte-Benth (2004) and Benth et al. (2006) is closely followed.

10.2   The B-NS non-Gaussian OU model

In 2001, Barndorff-Nielsen and Shepard constructed a stochastic model for the continuously-sampled variance of stock returns. The variance is assumed to be a superposition of positive non-Gaussian processes of the Ornstein-Uhlenbeck type.

Definition 18. (Ornstein-Uhlenbeck Process, [5])

An OU process is a stochastic process with dynamics given by:

$$dX(t) = -\lambda \left(X(t) - \alpha \right) dt + \sigma dZ(t) \tag{10.1}$$

$$Z(0) = 0$$

where $\lambda > 0$, $\alpha \in \mathbb{R}$ and $\sigma > 0$. $Z(t)$ is a background driving Lévy process (BDLP) that is a Lévy process which has independent and stationary increments. In the definition above, the parameters are defined as follows:

- λ : how "strongly" the system reacts to perturbations (the "decay-rate" or "growth-rate")

- σ^2 : the variation or the size of the noise.

- α : the asymptotic mean

In this chapter, the focus will be on a specific case where $Z = \{Z(t)\}_{t \geq 0}$ has no Gaussian component and has positive increments. In similar fashion to the Gaussian Ornstein-Uhlenbeck process, $Z = \{Z(t)\}_{t \geq 0}$ is also mean-reverting.

Definition 19. (The non-Gaussian OU model, Barndorff-Nielsen and Shepard (2001) [6])

Consider a completely filtered probability space $(\Omega, \mathcal{F}, \mathcal{F}_t, P)$ with $\{\mathcal{F}_t\}_{t \geq 0}$ the completion of $\sigma(W_s, L_{\lambda S}; s \leq t)$ and assume the existence of a risk-neutral probability measure (equivalent martingale measure (EMM)),$\mathbb{Q}$, then the price, $S(t)$, and the variance, $\sigma^2(t)$, have dynamics given by:

$$
\begin{cases}
dS(t) = \left(\mu + \beta\sigma^2(t)\right) S(t)dt + \sqrt{\sigma^2(t)}S(t)dW(t) & , S(0) = s > 0 \\
d\sigma^2(t) = -\lambda\sigma^2(t)dt + dZ(\lambda t) & , \sigma^2(0) = \sigma > 0
\end{cases}
\tag{10.2}
$$

$Z(0) = 0$, $\lambda > 0$, $\mu, \beta \in \mathbb{R}$ and $W = \{W(t)\}_{t \geq 0}$, a Wiener process, independent from the subordinator $Z(t)$.

The variance process is said to be a non-Gaussian Ornstein-Uhlenbeck process [5]. It is a mean reverting process like its Gaussian counterpart. However, the variance process has positive jumps and revert downwards which results in $\sigma^2(t)$ being positive.

The unusual timing $Z(\lambda t)$ has been chosen so that the marginal distribution of $\sigma^2(t)$ does not change with the value of λ. The Lévy measure, is represented by

$z(dl) \in \mathbb{R}^+$ since Z is a subordinator (BDLP) [4, 6]. In the generalised version of
(10.2), a leverage term $\rho dZ(\lambda t)$ is added to the right-hand side of the stock process.
However, this case is not covered for purposes of this dissertation.

10.3   Stochastic volatility processes

The continuous-time models derived from Brownian motion such as the Samuelson
or Black-Scholes which model the log-price of the assets by a process whose solution
is of the general case of the first equation in (10.2) are important in derivative pricing
[6]. In B-NS (2001), the asset's aggregate log-returns, $\{y_n\}$ are defined as:

$$y_n = \int\limits_{(n-1)\triangle}^{n\triangle} dS\,(t) = S(n\triangle) - S\,((n-1)\triangle) \tag{10.3}$$

where $\triangle > 0$ is the interval length are scaled mixtures of Normal distributions since:

$$y_n | \sigma_n^2 \sim N\left(\mu\triangle + \beta\sigma^2\triangle, \triangle\sigma^2\right) \tag{10.4}$$

that is the aggregate log-returns conditional on variance follow a Normal distribution
[6]. BNS (2001), implement this to implicitly model the marginal distribution of the
log-returns under a specified stationary distribution for the variance, $\sigma^2(t)$. Under
this assertion, there exists a subordinator Z such that the variance is the solution of
the OU second equation of (10.2) [8].

Proposition 3. *(Stationary OU process, Barndorff-Nielsen and Shepard (2001) [6])*

The stationary process, $\sigma^2(t)$, is of Ornstein-Uhlenbeck type satisfying the second

equation of (10.2) is a stationary OU process if it can be represented as:

$$\sigma^2(t) = \int_{-\infty}^{0} e^s dZ(\lambda t + s), \tag{10.5}$$

which can also be expressed (via the Itô formula for semi-martingales) as

$$\sigma^2(t) = e^{-\lambda t}\sigma^2(0) + \int_{0}^{t} e^{-\lambda(t-s)}dZ(\lambda s), \tag{10.6}$$

where $\lambda > 0$, and $Z_{\lambda s}$ is the BDLP.

Proof. If $F(t, v(t)) \in C^2$, then by the Taylor series expansion we have that:

$$dF = \frac{\partial F}{\partial t}dt + \frac{\partial F}{\partial v}dv + \frac{1}{2}\frac{\partial^2 F}{\partial t^2}(dt)^2 + \frac{1}{2}\frac{\partial^2 F}{\partial v^2}(dv)^2 + \frac{\partial^2 F}{\partial v \partial t}\partial v \partial t.$$

Let

$$F(t, v(t)) = v(t)e^{\lambda t}$$

If we define $F(0,0) = v(0)$ and $v(t) := \sigma^2(t)$, then

$$\begin{cases} \frac{\partial F}{\partial t} &= \lambda v(t)e^{\lambda t} \\[2mm] \frac{\partial F}{\partial v} &= e^{\lambda t} \\[2mm] \frac{\partial^2 F}{\partial v^2} &= 0. \end{cases}$$

Then by Itô's Lemma we have that,

$$dF = \frac{\partial F}{\partial t}dt + \frac{\partial F}{\partial v}dv = \lambda v(t)e^{\lambda t}dt + e^{\lambda t}dv = \lambda v(t)e^{\lambda t}dt + e^{\lambda t}(-\lambda v(t)dt + dZ_{\lambda t}) = e^{\lambda t}dZ_{\lambda t}.$$

Then it follows from the definition of F that,

$$F(t, v(t)) = v(0) + \int_0^t e^{\lambda s} dZ_{\lambda s.}$$

It then implies that,

$$v(t) = e^{-\lambda t} v(0) + e^{-\lambda t} \int_0^t e^{\lambda s} dZ_{\lambda s.}$$

which gives (10.6) above. This does not contain the deterministic component, dt, thus non-Gaussian. $\qquad\square$

10.4   Superposition of non-Gaussian OU-process

B-NS [6], suggested that a superposition of Ornstein-Uhlenbeck processes $Y_j(t)$, with different mean-reversion speeds, λ_j , should be used to obtain a more general correlation pattern in the volatility structure. The proposition below suggests the dynamics of the weighted sum variance.

Proposition 4. *(General non-Gaussian OU process, Benth et. al. (2006) [8])*

Consider positive weights w_j, for $j = 1, 2, ..., m$ such that $\sum_{j=1}^{m} w_j = 1$. Then define the general structure of the stationary process, $\sigma^2(t)$, by the superposition of m different non- Gaussian Ornstein-Uhlenbeck processes. Following the notation in Benth et al.(2006), the stationary process is defined as:

$$\sigma^2(t) = \sum_{j=1}^{m} w_j Y_j(t) \tag{10.7}$$

where the stochastic volatility processes $Y_j(t)$ are defined as

$$dY_j(t) = -\lambda_j Y_j(t) dt + dL_j(\lambda_j t), \tag{10.8}$$

for $j = 1, 2, ..., m$ and m independent BDLPs,L_j. Then Lévy measures, $l(dz)$, which correspond to the BDLPs, L_j are assumed positive since they are supported by these BDLPs (subordinators). The dynamics of $Y_j(t)$ are then assumed to be as follows[8]:

$$Y_j(t) = e^{-\lambda t} Y_j(0) + \int_0^t e^{-\lambda(t-s)} dL_j(\lambda_j s) \qquad (10.9)$$

for $0 \leq t$.

Proof. Let $v(t) := \sigma^2(t) = \sum_{j=1}^m w_j Y_j(t)$, and $F(t, v(t)) = v(t)e^{\lambda t} \in C^2$. Then applying Itô's Lemma for semi-martingales (as above) and applying summation rules to the definition of the stochastic integral we obtain (10.9). This completes the proof. $\square$

Recalling the definition, continuously-sampled realised variance over a period $[0, T]$ is defined as,

$$\sigma_R^2(T) = \frac{1}{T} \int_0^T \sigma^2(s) ds \qquad (10.10)$$

In the real world, the variance is sampled at discrete times and thus (10.10) above becomes:

$$\sigma_n^2(T) = \frac{1}{n} \sum_{i=1}^n \sigma^2(s_i) \qquad (10.11)$$

where sampling is done at times $s_i \in [0, T]$. In this chapter, however, all derivations will be based on the continuously-sampled variance. As defined in Part I, a variance swap pays the holder of the contract $N(\sigma_R^2(T) - K_{var})$, where N is the notional amount chosen to be $N = 1$ in order to simplify the expressions. K_{var}, is the fixed level of variance for the variance swap or its fair strike. It is expected that $\lim_{n \to \infty} \sigma_n^2(T) = \sigma_R^2(T)$. The quadratic variation of the log-price is connected to the realised variance as follows [8]:

For any sequence of partitions $t_0 = 0 < t_1 < ... < t_m$ with $sup(t_{i+1} - t_i) \to 0$ as $n \to \infty$

$$\lim_{n \to \infty} \frac{n}{(n-1)t} \sum_{i=1}^{n} ln \left(\frac{S(t_{i+1})}{S(t_i)} \right)^2 = \int_0^t \sigma^2(s)ds \qquad (10.12)$$

The fixed level of variance, K_{var}, is chosen so that the variance swap has a risk-neutral value equal to zero. This the implies that at time $0 \leq t \leq T$, the fixed level is given by conditional expectation of realised variance under a risk-neutral equivalent martingale measure (EMM), $\mathbb{Q}$.

$$K_{var}(t, T) = E_{\mathbb{Q}}[\sigma_R^2(T)|\mathcal{F}_t] \qquad (10.13)$$

This is a forward contract written on realised variance. It follows that:

$$K_{var}(0, T) = E_{\mathbb{Q}}[\sigma_R^2(T)] \qquad (10.14)$$

$$K_{var}(T, T) = \sigma_R^2(T) \qquad (10.15)$$

For a volatility swap we have,

$$K_{vol}(t, T) = E_{\mathbb{Q}}[\sigma_R(T)|\mathcal{F}_t] \qquad (10.16)$$

where $\sigma_R(T)$ is realised volatility. Hence in general we obtain,

$$K(t, T) = E_{\mathbb{Q}}[\sigma_R^{2\gamma}(T)|\mathcal{F}_t]$$

for $\gamma > -1$.

10.5   Using the Esscher transform

Benth and Saltyte-Benth (2004) pointed out that the B-NS non-Gaussian OU model gives rise to an incomplete market [9] because of its positive jump feature which is a result of the background driving Lévy process. However, in an incomplete market setting, several martingale measures can be used to price the variance swaps in a way consistent with no arbitrage [11]. If the price of the contingent claim, in this case, the variance swap, is attainable, then all choices of $\mathbb{Q}$ will produce the same price process. However, to specify one price, one choice of the equivalent martingale measure must be utilised. This is the work of the Esscher transform. The Esscher transform is used to develop a sub-class of the equivalent martingale measures which results in a sub-class of variances which can be calibrated with market data [8]. The equivalent martingale measure,$\mathbb{Q}$, is the probability measure equal to the real probability measure, $\mathbb{P}$, such that all continuously tradeable securities are martingales after discounting. This EMM minimises the relative entropy.

The negation of the Second Fundamental Theorem of asset pricing implies that for an incomplete market the requirement of no arbitrage is no longer sufficient to determine a unique price for the derivative. There are several martingale measures, all of which can be used to price derivatives in a way consistent with no arbitrage [11]. This explains the need for the Esscher transform. Following the work in Benth and Saltyte-Benth, an exponential integrability condition on the Lévy measure L is imposed to ensure the existence of moments of the stock price process [9]. The pricing of contingent claims such as options and variance swaps requires the first moment of the underlying i.e. $E_{\mathbb{Q}}[\sigma_R^2(T)|\mathcal{F}_t]$ in our case.

Definition 20. (Esscher transform, Schoutens (2003) [47])

Consider the stochastic market model such that $S(t) = S(0)e^{X(t)}$ where $X =$

$\{X(t) : t \geq 0\}$ is a Lévy process. Then let $f_t(x)$ be the probability density of $X(t)$. For $\theta \in \mathbb{R}$ such that $\int\limits_{-\infty}^{\infty} e^{\theta x} f_t(x) dx < \infty$, a new probability measure $f_t(x; \theta)$ can be defined as [47]:

$$f_t(x; \theta) = \frac{e^{\theta x} f_t(x)}{\int\limits_{-\infty}^{\infty} e^{\theta x} f_t(x) dx} \tag{10.17}$$

Now, θ, is selected such that the discounted price process is a martingale i.e $S(0) = e^{-rT} \mathbb{E}_\theta [S(t)]$ with the expectation taken with respect to the law with the new density function $f_t(x; \theta)$ [47]. Then for a stochastic process $X(t)$ defined on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$, the Esscher transform is a change of probability measure $\mathbb{P}$ by the process and a constant θ to the EMM, $\mathbb{Q}$,such that [50]:

$$\left.\frac{d\mathbb{Q}}{d\mathbb{P}}\right| \mathcal{F}_t = \frac{e^{\theta X(t)}}{\mathbb{E}\left[e^{\theta X(t)}\right]}. \tag{10.18}$$

Now, the EMMs, $\mathbb{Q}$, are constructed using the Esscher transform, following the techniques in [9]. Assume that $\theta_k(t) : [0, T] \to \mathbb{R}$,$j = 1, 2, ..., m$ are real-valued measurable and bounded functions. As in Benth and Saltyte-Benth (2004), consider the stochastic exponent defined as:

$$Z^\theta(t) = exp\left(\left(\sum_{j=1}^{m}\left(\int_0^t \theta_j(s) \, dL_j(\lambda_j s) - \int_0^t \lambda_j \psi_j(\theta_j(s)) \, ds\right)\right)\right) \tag{10.19}$$

which can also be represented as:

$$Z^\theta(t) = \prod_{j=1}^{m}\left(exp\left(\left(\int_0^t \theta_j(s) \, dL_j(\lambda_j s) - \int_0^t \lambda_j \psi_j(\theta_j(s)) \, ds\right)\right)\right) \tag{10.20}$$

where $\psi_j(x)$ are the log-moment generating functions of $L_j(t)$ that is $\psi_j(x) =$

$lnE\left(e^{xL_j(1)}\right)$ [9, 8]. To ensure the existence of the moments of the Lévy measure, an exponential integrability is imposed by the following condition:

Remark 1. Integrability Condition (**L**): There exists a constant $\kappa > 0$ such that the Lévy measure l_j satisfies the integrability condition

$$\int_0^t e^{z\kappa} l_j\,(dz) < \infty. \tag{10.21}$$

The constant $\kappa > 0$ determines the finite order of moments for the stochastic processes $Z^\theta(t)$. This follows from the following Lemma [9]:

Lemma 4. *(Key formula, Eberlein and Raible 1999 [29]))*

*If $g\colon [0,T] \to \mathbb{R}$ is a bounded and measurable function and the condition **(L)** holds for $\kappa = sup_{s\epsilon[0,T]}|g(s)|$, then*

$$E\left[exp\left(\int_0^t g(s)dL_s\right)\right] = exp\left(\int_0^t \psi(g(s))ds\right) \tag{10.22}$$

Proof. For any partition $0 = t_0 < .... < t_{N+1} = t$ of the interval $[0,t]$ we get by the independence and stationarity of the increments of the Lévy process, L_s, that :

$$E\left[exp\left(\sum_{j=o}^N g(t_j)\left(L_{t_{j+1}} - L_{t_j}\right)\right)\right] = \prod_{j=o}^N E\left[exp\left(g\left(t_j\right)\left(L_{t_{j+1}} - L_{t_j}\right)\right)\right]$$

$$= \prod_{j=o}^N exp\left(\psi\left(g(t_j)\right)(t_{j+1} - t_j)\right)$$

$$= exp\left(\sum_{j=o}^N \psi\left(g(t_j)\right)(t_{j+1} - t_j)\right)$$

Then, if with let the partition $(t_{j+1} - t_j) \to 0$, the right-hand side of the equations

above converges to $exp\left(\int\limits_{o}^{t}\psi\left(g(s)\right)ds\right)$ and the left-hand side of the equations above

converges in measure to $\int\limits_{o}^{t}g(s)dL_s$. Hence:

$$exp\left(\sum_{j=o}^{N}g(t_j)\left(L_{t_{j+1}}-L_{t_j}\right)\right)\rightarrow exp\left(\int\limits_{o}^{t}g(s)dL_s\right) \tag{10.23}$$

in measure. $\qquad\square$

Thus If we take $g(s):=\sigma(s,T)$, the immediate result from the Lemma above is that:

$$E\left[exp\left(\int\limits_{0}^{t}\sigma(s,T)dL_s\right)\right]=exp\left(\left(\int\limits_{0}^{t}\psi(\sigma(s,T))ds\right)\right) \tag{10.24}$$

Remark 2. The stochastic processes $Z^{\theta}(t)$, as defined in (10.19), are well-defined under the natural exponential integrability conditions on the Lévy measures $l_k\left(dz\right)$ which were assumed to hold. This means that, the processes $Z^{\theta}(t)$ are well defined for $t\in[0,T]$ if the condition **(L)** holds for $\kappa=sup_{j=1,..,m,s\epsilon[0,T]}|\theta_j(s)|$. The probability measure,

$$Q^{\theta}\left(A\right)=E[I_A Z^{\theta}(\tau_{max})] \tag{10.25}$$

where I_A is the indicator function and τ_{max} is a fixed time horizon including all trading times. The measure Q^{θ} is equivalent to $\mathbb{P}$ thus an EMM. The expectation due to probability measure Q^{θ} is denoted by $\mathbb{E}_{\theta}[.]$. The Lemma below ensures a sufficient condition for the existence of price dynamics of the variance swaps under EMM, Q^{θ}.

Lemma 5. *Suppose that condition **(L)** holds for* $\kappa:=1+sup_{j=1,..,m,s\epsilon[0,T]}|\theta_j(s)|$, *then*

$$\mathbb{E}_\theta[f(\sigma^2(T))] < \infty. \tag{10.26}$$

Proof. Let f be a real-valued measurable function with at most linear growth. Then from the assumption in the Lemma, it follows that $Z^\theta(T)$ is well defined. Since $\mathbb{Q}^\theta$ is equivalent to $\mathbb{P}$ we find that

$$\mathbb{E}_\theta\left[|f(\sigma^2(T))|\right] = E\left[|f(\sigma^2(T))|Z^\theta(T)\right] \tag{10.27}$$

and by linear growth of f it follows that $\mathbb{E}\left[exp\left(\int\limits_0^T \theta\left(s\right) + e^{-\lambda(T-s)}dL_s\right)\right]$ must be finite for the Lemma to hold, and this is ensured by the assumption. $\square$

10.6   Formula for deriving variance swap prices using Laplace transforms

In this section the formula for deriving the price of variance swaps in terms of Laplace transforms is summarised by the Lemma below. This is simply the Laplace transform of the conditional distribution of realised variance up to time $t \leq T$ [8].

Lemma 6. *(Laplace transform of realised variance, Benth and Saltyte-Benth (2004) [9, 8])*

Consider real and measurable functions $\theta_j(t) : [0,T] \rightarrow \mathbb{R}, j = 1, 2, ..., m$ and complex number z. Suppose that condition ($\boldsymbol{L}$) is satisfied and well-defined for $|Re(z)| < \left(\frac{1}{\lambda_j T}\left(1 - e^{-\lambda_j(T-s)}\right)\right)^{-1} \kappa, \forall j$ where $\kappa = sup_{j=1,..,m,s\epsilon[0,T]}|\theta_j(s)|$. Then,

$$E^\theta\left[e^{z\sigma_R^2(T)}|\mathcal{F}_t\right] = exp\left[\sum_{j=1}^m \lambda_j \left(\int_t^T \psi_j\left(\frac{zw_j}{\lambda_j T}\left(1 - e^{-\lambda_j(T-s)}\right) + \theta_j(s)\right) - \psi_j\left(\theta_j(s)\right)ds\right)\right]$$

$$* \exp\left[\frac{z}{T}\left(t\sigma_R^2(t) + \sum_{j=1}^{m}\frac{1}{\lambda_j}\left(1 - e^{-\lambda_j(T-t)}\right)w_jY_j(t)\right)\right] \qquad (10.28)$$

Proof. From the definition of continuously-sampled realised variance we obtain,

$$E^\theta\left[e^{z\sigma_R^2(T)}|\mathcal{F}_t\right] = E^\theta\left[e^{z\left(\frac{1}{T}\int_0^T\sigma^2(s)ds\right)}|\mathcal{F}_t\right]$$

It then follows from the superposition of m different non- Gaussian Ornstein-Uhlenbeck processes that,

$$E^\theta\left[e^{z\sigma_R^2(T)}|\mathcal{F}_t\right] = E^\theta\left[e^{z\left(\frac{1}{T}\int_0^T\left(\sum_{j=1}^{m}w_jY_j(s)\right)ds\right)}|\mathcal{F}_t\right]$$

then,

$$E^\theta\left[e^{z\sigma_R^2(T)}|\mathcal{F}_t\right] = E^\theta\left[e^{z\left(\frac{1}{T}\sum_{j=1}^{m}w_j\int_0^T Y_j(s)ds\right)}|\mathcal{F}_t\right]$$

from Bayes' formula we have that ,

$$E^\theta\left[e^{z\sigma_R^2(T)}|\mathcal{F}_t\right] = E\left[e^{z\left(\sum_{j=1}^{m}\frac{w_j}{T}\int_0^T Y_j(s)ds\right)}\frac{Z^\theta(T)}{Z^\theta(t)}|\mathcal{F}_t\right]$$

substituting $Z^\theta(t)$ we have that ,

$$E^\theta\left[e^{z\sigma_R^2(T)}|\mathcal{F}_t\right] = E\left[e^{\left(\sum_{j=1}^{m}\left(\frac{zw_j}{T}\int_0^T Y_j(s)ds + \int_t^T\theta_j(s)dL_j(\lambda_j s)\right)\right)}|\mathcal{F}_t\right].e^{\sum_{j=1}^{m}-\lambda_j\int_t^T\psi_j(\theta_j(s))ds}$$

now since $\sigma_R^2(s)$ is $\mathcal{F}_s$-adapted, and from the linearity of integrals we have that,

$$E^\theta\left[e^{z\sigma_R^2(T)}|\mathcal{F}_t\right] = E\left[e^{\left(\sum_{j=1}^{m}\left(\frac{zw_j}{T}\left(\int_0^t Y_j(s)ds + \int_t^T Y_j(s)ds\right) + \int_t^T\theta_j(s)dL_j(\lambda_j s)\right)\right)}|\mathcal{F}_t\right].e^{\sum_{j=1}^{m}-\lambda_j\int_t^T\psi_j(\theta_j(s))ds}$$

then,

$$
E^\theta\left[e^{z\sigma_R^2(T)}|\mathcal{F}_t\right] = E\left[e^{\left(\sum\limits_{j=1}^{m}\left(\frac{zw_j}{T}\int\limits_t^T Y_j(s)ds+\int\limits_t^T \theta_j(s)dL_j(\lambda_j s)\right)\right)}|\mathcal{F}_t\right].e^{\sum\limits_{j=1}^{m}\left(\frac{zw_j}{T}\int\limits_0^t Y_j(s)ds-\lambda_j\int\limits_t^T \psi_j(\theta_j(s))ds\right)}
$$

from the dynamics of the stochastic volatility processes $Y_j(t)$ in Equation (10.8) we have that (integral form),

$$
Y_j(T) = Y_j(t) - \int\limits_t^T \lambda_j Y_j(s)ds + \int\limits_t^T dL_j(\lambda_j s),
$$

then,

$$
\lambda_j\int\limits_t^T Y_j(s)ds = Y_j(t) - Y_j(T) + \int\limits_t^T dL_j(\lambda_j s),
$$

from the explicit expression of $Y_j(t) = e^{-\lambda t}Y_j(0) + \int\limits_0^t e^{-\lambda(t-s)}dL_j(\lambda_j s)$ [44] it follows that,

$$
\int\limits_t^T Y_j(s)ds = \frac{1}{\lambda_j}Y_j(t)\left(1 - e^{-\lambda_j(T-s)}\right) + \frac{1}{\lambda_j}\int\limits_t^T \left(1 - e^{-\lambda(T-s)}\right)dL_j(\lambda_j s)
$$

then,

$$
E^\theta\left[e^{z\sigma_R^2(T)}|\mathcal{F}_t\right] = E\left[e^{\left(\sum\limits_{j=1}^{m}\left(\int\limits_t^T \frac{zw_j}{T\lambda_j}\left(1-e^{-\lambda(T-s)}\right)+\theta_j(s)dL_j(\lambda_j s)\right)\right)}|\mathcal{F}_t\right]
$$

$$
*e^{\left(\frac{z}{T}t\sigma_R^2(t)+\sum\limits_{j=1}^{m}\left(\frac{zw_j}{T\lambda_j}\left(1-e^{-\lambda(T-t)}\right)Y_j(t)-\lambda_j\int\limits_t^T \psi_j(\theta_j(s))ds\right)\right)}
$$

implementing the independent increment property of the subordinator we have that [8],

$$E^{\theta}\left[e^{z\sigma_R^2(T)}|\mathcal{F}_t\right] = e^{\left(\sum_{j=1}^{m}\lambda_j\left(\int_t^T\psi_j\left(\frac{zw_j}{T\lambda_j}\left(1-e^{-\lambda(T-s)}\right)+\theta_j(s)\right)-\psi_j(\theta_j(s))ds\right)\right)}$$

$$*e^{\left(\frac{z}{T}\left(t\sigma_R^2(t)+\sum_{j=1}^{m}\left(\frac{1}{\lambda_j}\left(1-e^{-\lambda(T-t)}\right)w_jY_j(t)\right)\right)\right)}$$

which completes the proof. $\square$

Now the Lemma above can be used to derive the Laplace transform of variance price process.

Proposition 5. *(Laplace transform of B-NS realised variance, Benth et. al (2006)[8])For every $\gamma > -1$ and $a > 0$ satisfying $a < \left(\frac{\lambda_j}{T}\left(1-e^{-\lambda(T-s)}\right)\right)^{-1}\kappa$, for all j such that $\kappa = sup_{j=1,..,m,s\epsilon[0,T]}|\theta_j(s)|$, it holds that*

$$\sigma^{2\gamma}(t,T) = \frac{\Gamma(\gamma+1)}{2\pi i}\int_{a-i\infty}^{a+i\infty}z^{-(\gamma+1)}\eta(t,T,z)$$

$$*e^{\left(\frac{z}{T}\left(t\sigma_R^2(t)+\sum_{j=1}^{m}\left(\frac{1}{\lambda_j}\left(1-e^{-\lambda_j(T-t)}\right)w_jY_j(t)\right)\right)\right)}dz \qquad (10.29)$$

where

$$\eta(t,T,z) = e^{\left(\sum_{j=1}^{m}\lambda_j\left(\int_t^T\psi_j\left(\frac{zw_j}{T\lambda_j}\left(1-e^{-\lambda_j(T-s)}\right)+\theta_j(s)\right)-\psi_j(\theta_j(s))ds\right)\right)}$$

Proof. From the definition of the inverse Laplace transforms that for any $a > 0$ and $\gamma > -1$:

$$x^{\gamma} = \frac{\Gamma(\gamma+1)}{2\pi i}\int_{a-i\infty}^{a+i\infty}z^{-(\gamma+1)}e^{zx}dz$$

substituting the conditional realised variance under a well-defined moment ,

$$\sigma^{2\gamma}(t,T) \quad = \quad \frac{\Gamma(\gamma+1)}{2\pi i} \int\limits_{a-i\infty}^{a+i\infty} z^{-(\gamma+1)} E^{\theta}\left[e^{z\sigma_R^2(T)}|\mathcal{F}_t\right] dz$$

substituting $E^{\theta}\left[e^{z\sigma_R^2(T)}|\mathcal{F}_t\right]$ from Lemma 6 the result follows. $\square$

The proposition above gives the generalised formula for the dynamics of variance swaps in terms of the Laplace transform. Benth et al. (2007) mention that the generalised formula for the price of variance swaps derived above is applicable for numerical calculations via the fast Fourier transform (FFT) and fast numerical inversion techniques for the Laplace transform [8].

Proposition 6. *(Variance swap price, Benth and Saltyte-Benth (2004) [9, 8])*

The price of the variance swap is given by the following expression:

$$\sigma^2(t,T) \quad = \quad \frac{1}{T}\left(t\sigma_R^2(t) + \sum_{j=1}^{m}\left(\frac{1}{\lambda_j}\left(1 - e^{-\lambda_j(T-t)}\right) w_j Y_j(t)\right)\right) \qquad (10.30)$$
$$+ \sum_{j=1}^{m}\left(\frac{w_j}{T}\int_{t}^{T} \psi_j'(\theta_j(s))\left(1 - e^{-\lambda_j(T-s)}\right) ds\right)$$

Proof. Consider $z \in \mathbb{R}$, differentiating the expression for $E^{\theta}\left[e^{z\sigma_R^2(t,T)}\right]$ in Lemma 6, it follows that:

$$\frac{d}{dz}E^{\theta}\left[e^{z\sigma_R^2(t,T)}\right] \quad = \quad \sigma_R^2(t,T).e^{z\sigma_R^2(t,T)} \quad , for\, t \leq T$$

then letting $z = 0$ one obtains $\sigma_R^2(T)$, the realised variance at time T. Implementing this to the expression in Lemma 6:

$$
\frac{d}{dz} E^\theta \left[e^{z\sigma_R^2(t,T)} \right] = e^{\left(\sum_{j=1}^{m} \lambda_j \left(\int_t^T \psi_j \left(\frac{zw_j}{T\lambda_j} \left(1 - e^{-\lambda_j(T-s)} \right) + \theta_j(s) \right) - \psi_j(\theta_j(s)) ds \right) \right)}
$$
$$
* \sum_{j=1}^{m} \left(\frac{w_j}{T} \int_t^T \psi_j^{'}(\theta_j(s)) \left(1 - e^{-\lambda_j(T-t)} \right) ds \right)
$$
$$
* e^{\left(\frac{z}{T} \left(t\sigma_R^2(t) + \sum_{j=1}^{m} \left(\frac{1}{\lambda_j} \left(1 - e^{-\lambda_j(T-t)} \right) w_j Y_j(t) \right) \right) \right)}
$$
$$
* \frac{1}{T} \left(t\sigma_R^2(t) + \sum_{j=1}^{m} \left(\frac{1}{\lambda_j} \left(1 - e^{-\lambda_j(T-t)} \right) w_j Y_j(t) \right) \right)
$$

letting $z = 0$ it follows that:

$$
\sigma^2(t,T) = \frac{1}{T} \left(t\sigma_R^2(t) + \sum_{j=1}^{m} \left(\frac{1}{\lambda_j} \left(1 - e^{-\lambda_j(T-t)} \right) w_j Y_j(t) \right) \right)
$$
$$
+ \sum_{j=1}^{m} \left(\frac{w_j}{T} \int_t^T \psi_j^{'}(\theta_j(s)) \left(1 - e^{-\lambda_j(T-t)} \right) ds \right).
$$

It is of importance to observe that the explicit expression for the price of the variance swaps obtained from the Laplace transforms is dependent on both $\sigma^2(t)$ and $\sigma_R^2(T)$ (the current level of variance and the realised level of variance). The relationship between the price of the variance swap and the variance level will be demonstrated later in the chapter. □

10.7   Theoretical Fourier Transform (FFT) Evaluation

In this section, the Fast Fourier Transform (FFT) is theoretically applied to obtain the numerical solution of the expressions for realised variance in terms of the Laplace

transforms obtained in the previous section. The representation of the price dynamics of the variance swaps in terms of the Laplace transform as in Proposition 5 is well suited to numerical evaluations based on the FFT and other inversion techniques for the Laplace transform.

The FFT is a reliable and computationally efficient method of evaluating the discrete Fourier Transform [24]. This method has been implemented in option pricing since over 2 decades ago by Carr and Madan (1999) who assume that the characteristic function of the risk-neutral density is known analytically then develop a simple analytic expression for the Fourier Transform of the option value. The FFT is then used to numerically solve for the option price or its time value [24].

10.7.1   Theoretical FFT in Pricing Variance Swaps

As mentioned above, the Fast Fourier Transform is a computationally efficient way of calculating the sum (discrete Fourier transform) [24]:

$$w(k) = \sum_{j=1}^{N} e^{-i\frac{2\pi}{N}(j-1)(k-1)} x(j) \quad for \, k = 1, 2, ..., N \tag{10.31}$$

in which N complex numbers are fed as inputs to obtain another sequence of N complex numbers. The implementation of the sum above reduces the number of multiplications required in the N summations from $\mathcal{O}(N^2)$ to $\mathcal{O}(Nln_2(N))$ which is a large reduction [24]. Although this method proposed by Carr and Madan (1999) [24] is fast and reliable, one has to have in mind that the analytical form of the characteristic function has to be known. The expression in Proposition 5, given its form, can be implemented using the FFT. To achieve this, there is need to approximate the integral with a finite sum thus z and $\tilde{\sigma}^2 := \frac{t}{T}\sigma_R^2$ needs to be

discretised. Let

$$\tilde{\sigma}^2(k) = \triangle\tilde{\sigma}^2.(k-1) \tag{10.32}$$

and

$$z(j) = a + i\triangle z(j-1) \tag{10.33}$$

then the expression in Proposition 5 can be re-written in standard FFT form as:

$$
\begin{aligned}
\sigma^{2\gamma}(t,T) &= \frac{\Gamma(\gamma+1)}{2\pi i}\int_{a-i\infty}^{a+i\infty} z^{-(\gamma+1)} E^\theta\left[e^{z\sigma_R^2(T)}|\mathcal{F}_t\right]dz \\[2mm]
&\approx \frac{\Gamma(\gamma+1)}{2\pi i}\sum_{j=1}^{N} z(j)^{-(\gamma+1)} E^\theta\left[e^{z\sigma_R^2(T)}|\mathcal{F}_t\right]i\triangle z \\[2mm]
&= \frac{\Gamma(\gamma+1)}{2\pi i}\sum_{j=1}^{N} z(j)^{-(\gamma+1)} e^{\left(\sum_{j=1}^{m}\lambda_j\left(\int_t^T \psi_j\left(\frac{z(j)w_j}{T\lambda_j}\left(1-e^{-\lambda(T-s)}\right)+\theta_j(s)\right)-\psi_j(\theta_j(s))ds\right)\right)} \\[1mm]
&\quad .e^{\left(z(j)\left(\frac{t}{T}\sigma_R^2(t)\right)\right)} e^{\frac{z(j)}{T}\sum_{j=1}^{m}\left(\frac{1}{\lambda_j}\left(1-e^{-\lambda(T-t)}\right)w_jY_j(t)\right)}i\triangle z \\[2mm]
&= \frac{\Gamma(\gamma+1)}{2\pi i}\sum_{j=1}^{N} (a+i\triangle z(j-1))^{-(\gamma+1)} e^{\left(\sum_{j=1}^{m}\lambda_j\left(\int_t^T \psi_j\left(\frac{a+i\triangle z(j-1)(w_j)}{T\lambda_j}\left(1-e^{-\lambda(T-s)}\right)+\theta_j(s)\right)-\psi_j(\theta_j(s))ds\right)\right)} \\[1mm]
&\quad .e^{\left((a+i\triangle z(j-1))\left(\triangle\tilde{\sigma}^2.(t-1)\right)\right)} .e^{\sum_{j=1}^{m}\frac{a+i\triangle z(j-1)}{T}\left(\frac{1}{\lambda_j}\left(1-e^{-\lambda(T-t)}\right)w_jY_j(t)\right)}i\triangle z \\[2mm]
&= \frac{\Gamma(\gamma+1)}{2\pi i}\sum_{j=1}^{N} (a+i\triangle z(j-1))^{-(\gamma+1)} e^{\left(\sum_{j=1}^{m}\lambda_j\left(\int_t^T \psi_j\left(\frac{a+i\triangle z(j-1)(w_j)}{T\lambda_j}\left(1-e^{-\lambda(T-s)}\right)+\theta_j(s)\right)-\psi_j(\theta_j(s))ds\right)\right)} \\[1mm]
&\quad .e^{i\frac{2\pi}{N}(j-1)(t-1)} .e^{a\left(\triangle\tilde{\sigma}^2.(t-1)\right)} .e^{\sum_{j=1}^{m}\frac{a+i\triangle z(j-1)}{T}\left(\frac{1}{\lambda_j}\left(1-e^{-\lambda(T-t)}\right)w_jY_j(t)\right)}i\triangle z
\end{aligned}
$$

if we let $\triangle z = \frac{2\pi}{N\triangle\tilde{\sigma}^2}$ then a sum of the form (10.31) is obtained as:

$$
\begin{aligned}
\sigma^{2\gamma}(t,T) &\approx \frac{\Gamma(\gamma+1)}{N\triangle\tilde{\sigma}^2}\sum_{j=1}^{N} (a+i\triangle z(j-1))^{-(\gamma+1)} e^{\left(\sum_{j=1}^{m}\lambda_j\left(\int_t^T \psi_j\left(\frac{a+i\triangle z(j-1)(w_j)}{T\lambda_j}\left(1-e^{-\lambda(T-s)}\right)+\theta_j(s)\right)-\psi_j(\theta_j(s))ds\right)\right)} \\[1mm]
&\quad .e^{i\frac{2\pi}{N}(j-1)(t-1)} .e^{a\left(\triangle\tilde{\sigma}^2.(t-1)\right)} .e^{\sum_{j=1}^{m}\frac{a+i\triangle z(j-1)}{T}\left(\frac{1}{\lambda_j}\left(1-e^{-\lambda(T-t)}\right)w_jY_j(t)\right)}
\end{aligned}
$$

for $t = 1, 2, .., N$.

Carr and Madan (1999) pointed out that Fourier transforms are expressed in terms of the characteristic function of the log price. The characteristic function is known analytically in many scenarios. Carr and Madan (1999) point out that where the dynamics of the log-price are infinitely divisible processes, the characteristic function is obtained naturally from the representation of these processes [24]. Moreover, to obtain the log-moment generating functions explicitly, there is a need to specify the BDLP Lévy processes, L_j. A common approach is to specify a stationary distribution of the OU process and then derive the cumulant-generating (log-moment generating) function for the Lévy process from the distribution [8]. Some of the distribution classes used to specify the characteristic function are independent stable increments, the Variance-Gamma process [24] and the Inverse Gaussian (IG) distribution [6].

10.7.2   The Inverse Gaussian distribution

In this section, the generalised Inverse Gaussian (IG) is specified for the OU process and the cumulant-generating function of the Lévy obtained from the IG distribution. If $x \sim IG(v, \delta, \gamma)$ (note that the IG distribution is stated in this way since the standard $IG(\lambda, \gamma, \delta)$ is unavailable due to the notation used in this chapter) then it has the density function [6]:

$$\frac{(\gamma/\delta)^\nu}{2K_\nu(\delta\gamma)} x^{\nu-1} exp \left\{ -\frac{1}{2} \left(\delta^2 x^{-1} + \gamma^2 x \right) \right\}, \qquad x > 0 \qquad (10.34)$$

where K_ν is a Bessel function of the third kind with index ν, $\gamma > 0$, and $\delta > 0$. Special cases are [6]:

1. the Inverse Gaussian Law, where $\nu = -\frac{1}{2}$;

2. the positive Hyperbolic Law, where $\nu = 1$;

3. the inverse $\chi^2 - Law$ with df degrees of freedom where $\nu = -df/2$, $\delta = \sqrt{df}$ and $\gamma = 0$;

4. Γ,where $\delta = 0$ and $\nu > 0$.

The law for $IG(-\frac{1}{2}, \delta, \gamma)$, the Inverse Gaussian is given by:

$$\frac{\delta}{\sqrt{2\pi}}exp(\delta\gamma)x^{-\frac{3}{2}}exp\left\{-\frac{1}{2}\left(\delta^2x^{-1} + \gamma^2x\right)\right\}, \qquad x > 0 \qquad\qquad (10.35)$$

with $\gamma > 0$ and $\delta > 0$. The impact of the parameters δ (shape parameter) and impact of γ (scale parameter) is shown below in Figure 10.7.2 and Figure 10.7.2:

Figure 10.1: The impact of $\delta \in (2, 4, 6, 8)$ for $\gamma = 2$
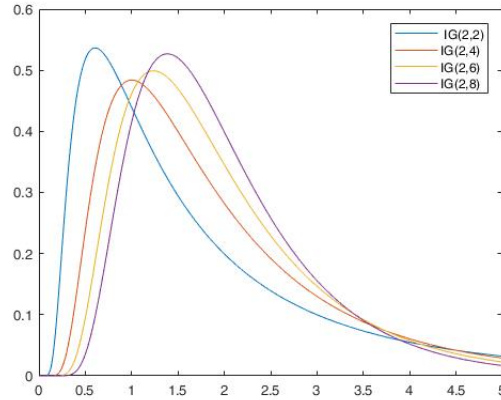
Figure 10.2: The impact of $\gamma \in (2, 4, 6, 8)$ for $\delta = 2$
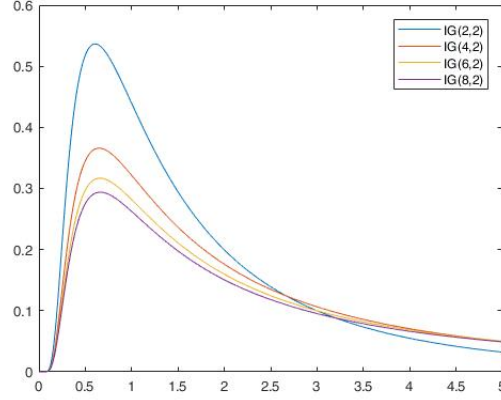
Nicola to and Vernados (2003), state that it is possible to state the variance process as the Inverse Gaussian Ornstein-Uhlenbeck process (IG-OU) with the stationary distribution of the BDLP being given by an $IG(\delta, \gamma)$ law. The log-moment generating function (cumulant-generating function $\psi_j(x) = lnE\left(e^{xL_j(1)}\right)$ of $IG(\delta, \gamma)$ is given by [44] :

$$\psi^{IG}(\theta) = \delta\gamma - \delta(\gamma^2 - 2\theta)^{\frac{1}{2}} \qquad (10.36)$$

The cumulant generating function of, X, $\psi^X(\theta) = lnE[e^{\theta X}]$, is related to the cumulant function, $\psi(\theta)$, of the BDLP by the formula:

$$\psi(\theta) = \theta\frac{d\psi^X(\theta)}{d\theta} \qquad (10.37)$$

Thus the cumulant-generating function of the $IG(\delta, \gamma)$ law is:

$$\psi(\theta) = \theta\delta(\gamma^2 - 2\theta)^{-\frac{1}{2}} \qquad (10.38)$$

If the stationary distribution of the OU process is inverse Gaussian i.e. IG-OU as above, then the log-returns of the stock will be approximately Normal Inverse Gaus-

sian (NIG) [8]. This four-parameter distribution fits well to the distribution of log-returns of stock prices [44]. The probability density function of the $NIG(\alpha, \beta, \delta, \mu)$ is given explicitly as follows [9]:

$$f\left(x; \alpha, \beta, \delta, \mu\right) = \frac{\alpha\delta K_1\left(\alpha\sqrt{\delta^2 + (x - \mu)^2}\right)}{\pi\sqrt{\delta^2 + (x - \mu)^2}} e^{\delta\sqrt{\alpha^2 - \beta^2} + \beta(x - \mu)} \tag{10.39}$$

where $\mu \in \mathbb{R}$, $\delta > 0$ and $0 \leq |\beta| \leq \alpha$

The method of moments estimation (MME) method can be implemented as an initial point of numerical estimation of parametric models. In this chapter, the MME estimation is applied to the NIG distribution's first and second cumulant, the skewness and the excess kurtosis to construct a system of non-linear equations for the four parameters in the $X \sim NIG(\alpha, \beta, \delta, \mu)$ distribution. The central moments for $X \sim NIG(\alpha, \beta, \delta, \mu)$ are [39]: $\mathbb{E}\left(X\right) = \mu + \delta\frac{\alpha^2}{\gamma}$, $Var\left(X\right) = \delta\frac{\alpha^2}{\gamma^3}$, $Skew\left(X\right) = \delta\frac{3\beta}{\alpha\sqrt{\delta\gamma}}$ and $Kurt\left(X\right) = 3 + 3\left(1 + 4\left(\frac{\beta}{\alpha}\right)^2\right)\frac{1}{\delta\gamma}$. The solution of the system of equations yields the following parameter estimates (see Theorem 2 [31]):

$$\hat{\beta} = \frac{3}{s(\rho - 1)\gamma_1} \tag{10.40}$$

$$\hat{\delta} = \frac{3s\sqrt{(\rho - 1)}}{\rho|\gamma_1|} \tag{10.41}$$

$$\hat{\mu} = \bar{x} - \frac{3s}{\rho\gamma_1} \tag{10.42}$$

$$\hat{\alpha} = \frac{3\sqrt{\rho}}{(\rho - 1)s|\gamma_1|} \tag{10.43}$$

where

$$\rho = 3\left(\frac{\gamma_2}{\gamma_1{}^2}\right) - 4 > 1 \qquad (10.44)$$

for s^2 the sample variance of log returns, $\bar{x}$ the sample mean for the log returns series , γ_1 the sample skewness and γ_2 the sample kurtosis.

10.7.3   Numerical examples

In this section, numerical examples to evaluate variance swap prices are presented. The parameters of the NIG distribution are derived using estimates from the method of moments estimation (MME). These estimates are then used in the simulation of the NIG process. Furthermore, the analytical formula derived in Proposition 6 used to estimate the price of the variance swap in Equation (12.25) where the OU process of the variance of log-returns is driven by the Normal Inverse Gaussian process.

10.7.4   Market Data

The data-set considered used this analysis consists of a set of 2520 FTSE/JSE Top 40 index daily closing prices over the period 6 April 2009 to 10 May 2019. This is the empirical data introduced earlier in Part I. The summary statistics of the daily log-returns on the JSE Top 40 index are stated again below as:

Table 2: Descriptive statistics for the JSE Top 40 daily log-returns

Variable	Value
Mean return	3.8344×10^{-4}
Standard deviation	0.0108
Variance	1.1660×10^{-4}
Minimum	-0.0405
Maximum	0.0468
Kurtosis	4.2504
Skewness	-0.1348

Parameter estimation

The parameters of the $NIG(\alpha, \beta, \delta, \mu)$ for the log-returns of the FTSE/JSE Top 40 index are estimated via the MME and MLE. To estimate the parameters via the MME, the system of equations resulting from substituting central moments with the observed sample moments of the log-returns empirical data is solved. In the MLE method, the log-likelihood function in Equation (8.8) is optimised to a local minimum in R.

Table 3: Estimates of the NIG Parameters

	α	β	δ	μ
MME	78.1947	-2.9565	0.009098	7.2828×10^{-4}
MLE	117.82900	-12.59909	0.01363	1.85×10^{-3}

The MME parameters are optimised to a local minimum using MATLAB's fminunc and the MLE are obtained using 'nigFit' in R. The fit of the MLE parameters against the empirical data shown below.

Figure 10.3: Parameter Fit to Empirical Data
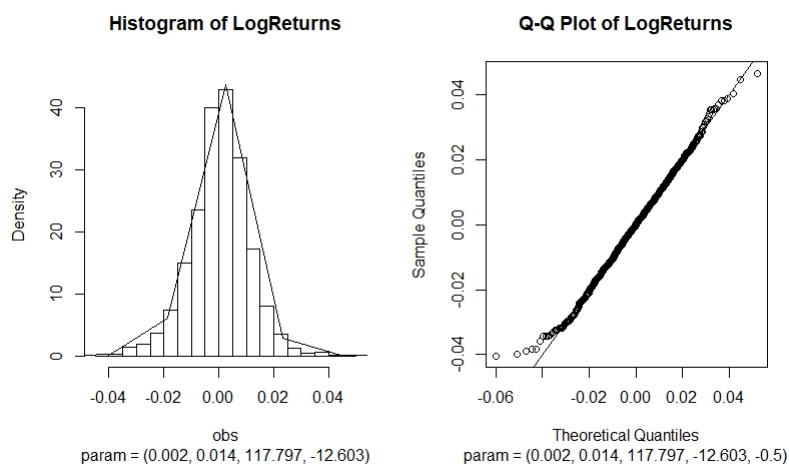
The NIG captures the peak and tail distribution of the index log returns.

10.7.5   Path simulation

Using the techniques shown a previous chapter and the parameters in the table above, the sample IG and NIG processes are shown below:

Figure 10.4: Inverse Gaussian process
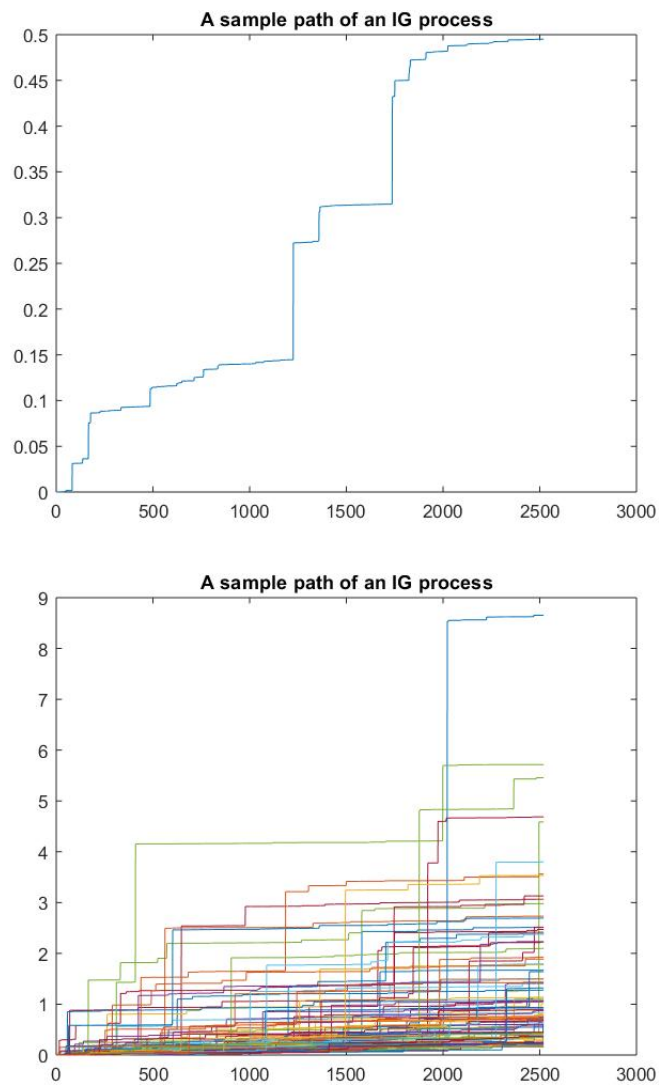
Positive jumps with mean-reversion are observed in the simulation.

Figure 10.5: Normal Inverse Gaussian Process
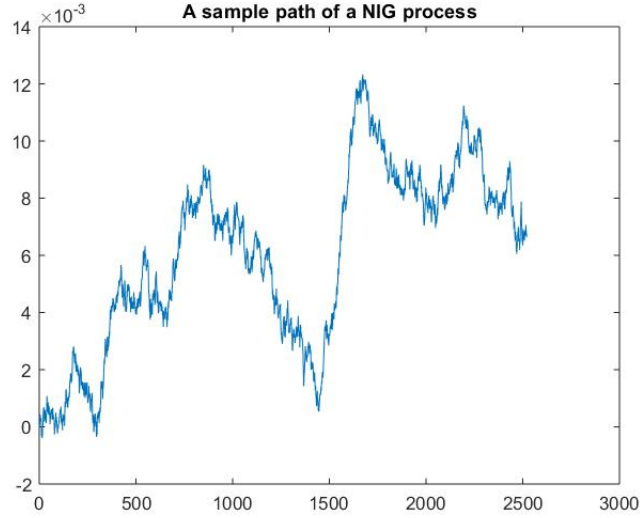
The stock process can be described by the path above.

10.7.6   Exploring the explicit formula for the price of the variance swap

Consider the price of the variance swap derived in Equation (10.30). Then since the Inverse Gaussian distribution is self-decompasable, the law of the BDLP of the variance process in Equation (10.2) can be specified as the $IG\,(a,b)$ distribution. Moreover, the cumulant generating function of the $IG(a,b)$ distribution is given by [44]:

$$\psi^{IG}(\theta) = log\mathbb{E}\left(e^{-\theta X}\right) = ab - a\left(b^2 - 2\theta\right)^{\frac{1}{2}}. \qquad (10.45)$$

Thus to obtain the *IG–OU* process, the BDLP must have a cumulant function given by Equation (10.38), given by $\psi(\theta) = \theta\left(\frac{d}{d\theta}\psi^{IG}(\theta)\right)$. This gives $\psi(\theta) = \theta a\left(b^2 - 2\theta\right)^{-\frac{1}{2}}$. This case is of particular interest because when $\rho = 0$, the IG distributed stationary OU process results in NIG distributed log returns of the stock. This class of distributions has been shown to provide a good fit to log returns of

stock prices [44].

In the explicit formula for the price of a variance swap in Equation (10.30), m, is assumed to be 2 to demonstrate the superposition of two different non-Gaussian Ornstein–Uhlenbeck processes. The formula requires estimates of the weights w_j, $Y_j(t)$ the variance processes and $\theta_j(s)$ the decay rates. Using the MME parameters of the $NIG(\alpha, \beta, \delta, \mu)$ we obtain the variance estimate as $\alpha^2 \delta \left(\alpha^2 - \beta^2\right)^{-3/2} = 1.3163 \times 10^{-4}$ [47]. Assuming equal weights i.e. $w_1 = w_2 = 0.5$, then it follows that $Y_1(t) = Y_2(t) = 6.5817 \times 10^{-5}$. The decay rates are assumed to be $\theta_1(s) = 0.9$ and $\theta_2(s) = 0.03$ for the two OU processes respectively. For various levels of realised variance, the value of the variance swap is given the following curve:

Figure 10.6: Explicit price of the variance swap
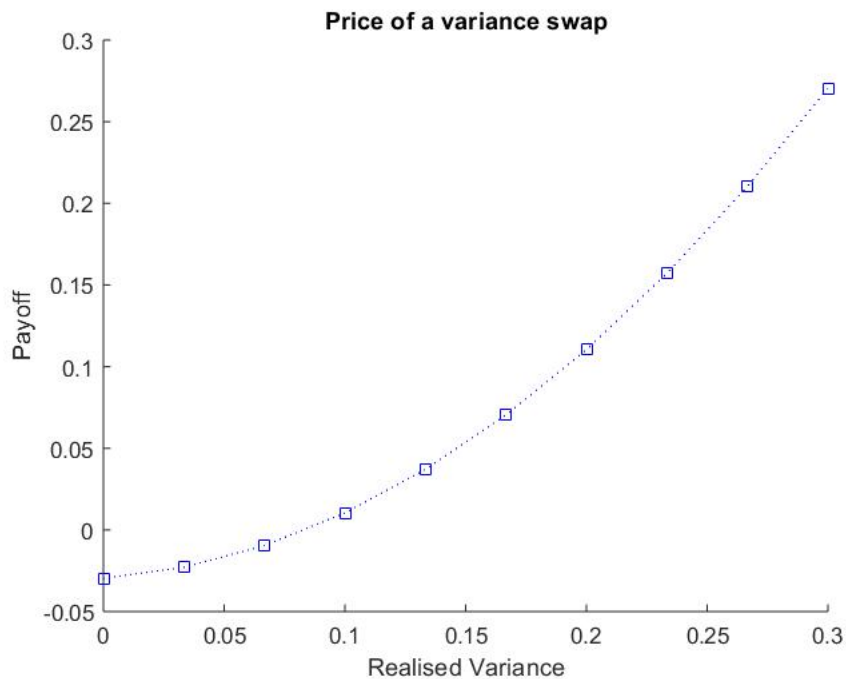
The price was obtained in MATLAB from the explicit price derived in (10.30).

The impact of the decay parameters $\theta_1(s)$ and $\theta_2(s)$ on the price of the variance swaps is shown below:

Figure 10.7: The impact of decay parameters on the price of variance swaps
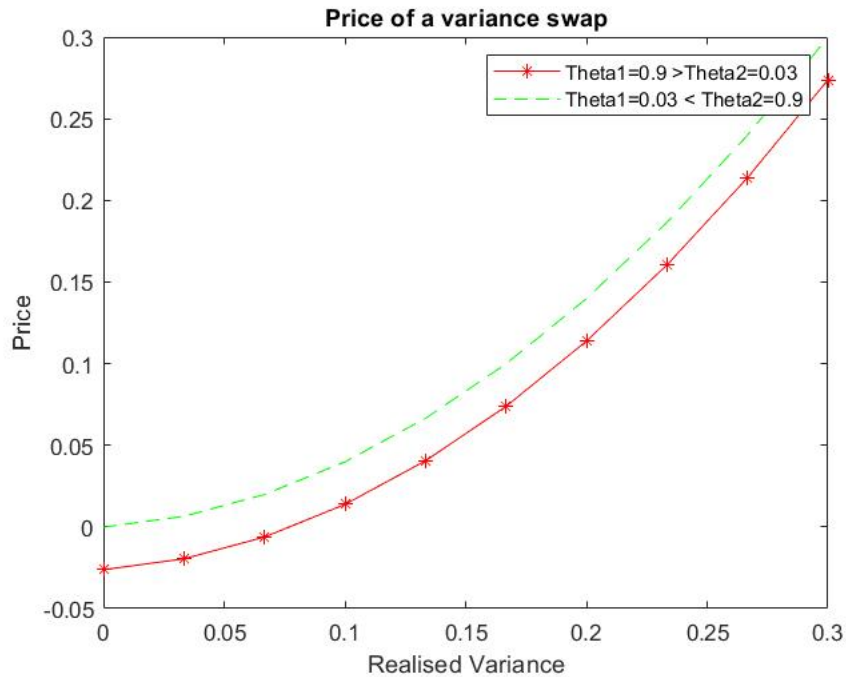
Although the variance of the $m = 2$ OU processes is equally weighted, the decay parameters have an impact on the price.

10.8   Conclusions

The model developed by Barndoff-Nielsen and Shepard (2001) provides attractive features to the volatility process such as positive jumps that are observable in market data. The analytical formula for continuously-sampled variance swaps was obtained using Laplace transform under an integrability condition for Levy processes. Although elementary results such as the Laplace transforms were implemented to obtain the semi closed-form analytical formula for the price of variance swaps, the formula itself is still practically complex. The estimates of the NIG were obtained from the method of moments estimation (MME). An implementation of the formula obtained showed that the price of continuously-sampled variance swaps is convex in realised variance (smile effect). The this consistent with the derived formula for the

price of variance swaps in which the realised variance is present.

11 Pricing Variance Swaps Under the Heston Model

The Heston Model(1993) [34] is amongst the most popular models to describe the pathway of the volatility of an underlying asset. As in the B-NS model described earlier, the dynamics of the asset's volatility are described by a mean-reverting stochastic process. In Zhu and Lian (2011) [51], an analytical closed-form solution for pricing variance swaps under stochastic volatility with an OU process is derived. The closed-form solution for pricing discretely sampled variance swaps with stochastic volatility is obtained from solving the model's two-part partial differential equation (PDE) using Fourier transform techniques under Feynman theory. The asymptotic form of the analytical formula for the value of the variance swaps based on the discrete sampling is studied. In particular, convergence of the discretely-sampled case to the continuously-sampled case under the Heston model is investigated.

11.1 The Heston Model

If $(\Omega, \mathcal{F}, \mathcal{F}(t), \mathbb{P})$ is a probability space adapted to the filtration $\{\mathcal{F}(t) : 0 \leq t \leq T\}$, then Heston (1993) suggests the following dynamics for the stock price and variance processes:

$$
\begin{cases}
dS(t) = \mu S(t)dt + \sqrt{v(t)}S(t)dW(t)^1 \\
dv(t) = \kappa \left(\theta - v(t) \right) dt + \sigma_\nu \sqrt{v(t)}dW(t)^2 \\
dW(t)^1 dW(t)^2 = \rho dt
\end{cases}
\tag{11.1}
$$

with initial conditions:

$$\begin{cases} S(0) \geq 0 \\ \nu(0) \geq 0 \end{cases}$$

where

- μ is the expected return of the underlying asset, which is r in a risk neutral setting

- θ is the long-term mean of variance, which the variance is assumed to revert to regardless of the initial level

- κ is a mean-reverting speed parameter of the variance towards θ

- σ_v is the 'volatility of volatility ' which contributes to the variation in volatility together with the kurtosis of the underlying asset's distribution.

The notation for the volatility, $\nu(t) := \sigma^2(t)$ in this section is changed to not be confused with the volatility of volatility σ_ν. The two Wiener processes $dW(t)^1$ and $dW(t)^2$ describe the random noise in the underlying stock and its volatility respectively. These two are assumed to be correlated with a constant correlation coefficient, ρ [51]. The variance should be positive and this is ensured by fixing $2\kappa\theta \geq \sigma^2$ [51]. This is known as the *Feller condition*. For purposes of this discussion, the expected return of the underlying is assumed to be constant. However, this assumption does not significantly affect any of the results presented in this section and the representation can be easily generalized to incorporate the case of deterministic interest rates [38].

11.2 Risk neutral pricing

It should be observed that the stock and variance processes described in Equation (11.1) are under the real-world probability $\mathbb{P}$. The existence theorem of equivalent martingale measure (EMM), enables us to change the real probability measure to a risk-neutral probability measure under which contingent claims can be priced. This is particularly achieved by the Girsanov Theorem (see [11], page 161). Define

$$d\widetilde{W(t)^1} = dW(t)^1 + \varphi(t)dt \tag{11.2}$$

and

$$d\widetilde{W(t)^2} = dW(t)^2 + \lambda(S, \nu, t)dt \tag{11.3}$$

then let process $L(t)$ be:

$$L(t) = e^{\left\{\int_0^t \varphi(s)dW(t)^1 + \int_0^t \varphi(s)dW(t)^2 - \frac{1}{2}\int_0^t \varphi(s)^2 + \lambda^2(S,\nu,t)ds\right\}} \tag{11.4}$$

where $\varphi(t) = \frac{\mu-r}{\sqrt{\nu(t)}}$, $\widetilde{W(t)}$ is the $\mathbb{Q}$-Wiener process and $W(t)$ is the $\mathbb{P}$-Wiener process. The new probability measure $\mathbb{Q}$ on $\mathcal{F}(t)$ is given by:

$$\frac{d\mathbb{Q}}{d\mathbb{P}} = L(t).$$

The Heston Model (11.1) can now be redefined under $\mathbb{Q}$ as:

$$\begin{cases} dS(t) & = \mu S(t)dt + \sqrt{v(t)}S(t)dW(t)^1 \\ dv(t) & = \kappa^* \left(\theta^* - v(t)\right)dt + \sigma_\nu \sqrt{v(t)}dW(t)^2 \end{cases} \tag{11.5}$$

with $\{S(t)\}_{t\geq 0}$ and $\{\nu(t)\}_{t\geq 0}$ with

$$\kappa^* = \kappa + \lambda \tag{11.6}$$

$$\theta^* = \frac{\kappa\theta}{\kappa + \lambda} \tag{11.7}$$

The function $\lambda(S, \nu, t)$ is the *market price of volatility risk.* As explained in Heston (1993), the function yields a premium proportional to the variance so that $\lambda(S, \nu, t) = \lambda\nu(t)$ [35, 45].

Recalling the value of a variance swap at time t we have that:

$$V(t) = e^{-r(T-t)}\mathbb{E}_{\mathbb{Q}}\left(N(\sigma_R^2(T) - K_{var})\right) \tag{11.8}$$

and since the swap must be entered into with no cost at time $t = 0$ it implies that:

$$K_{var} = E_{\mathbb{Q}}[\sigma_R^2(T)] \tag{11.9}$$

11.3   Approach for Pricing Discretely-Sampled Variance Swaps

In this section, the Fourier transform approach is used to obtain a closed-form analytical solution for the fair delivery price of a variance swap. The associated PDE is analytically solved and an explicit closed-form solution is obtained as in Heston [34]. The expected value of realised variance under the risk-neutral measure $\mathbb{Q}$ can be expressed as [38]:

$$\mathbb{E}_{\mathbb{Q}}\left[(\sigma_R^2(T))\right] = \frac{1}{N\Delta t}\sum_{i=1}^{N}\mathbb{E}_{\mathbb{Q}}\left[\left(\frac{S(t_i)-S(t_{i-1})}{S(t_{i-1})}\right)^2\right].100^2 \qquad (11.10)$$

Looking at the expression above, the problem of valuing a variance swap can be reduced to:

$$\mathbb{E}_{\mathbb{Q}}\left[\left(\frac{S(t_i)-S(t_{i-1})}{S(t_{i-1})}\right)^2\right] \qquad (11.11)$$

To calculate the expectation two cases need to be considered:

1. i $= 1$ and,

2. i > 1.

This is due to the difference in the calculation procedures. The i is fixed as a constant, hence both t_i and t_{i-1} are regarded as known constants [38].

Case $i > 1$ In this case the time $t_{i-1} > 0$ and $S(t_{i-1}) > 0$ are unknown at the current time $t = 0$. This means that the payoff function depends on two unknown values that is $S(t_{i-1})$ and $S(t_i)$ (the future underlying price value). The two-dimensional problem becomes difficult to solve, thus a new function $I(t)$ is introduced to solve the two-dimensional problem as in Little and Pant (2001) [41]:

$$I(t) = \int_0^t \delta\left(t_{i-1}-\tau\right)S(\tau)d\tau \qquad (11.12)$$

where $\delta(x)$ is the Dirac delta function (see Appendix). Then

$$I(t) = \begin{cases} 0, & ;0 \le t < t_{i-1} \\ S(t_{i-1}), & ;t \ge t_{i-1} \end{cases} \qquad (11.13)$$

The expected value of an expression involving $S(t)$ and $I(t)$ have to be evaluated . A contingent claim whose payoff at expiry t_i depends on $S(t_i)$ and $I(t_i)$ is considered. Under the assumptions made regarding the dynamics of the underlying asset, the set of historical prices are independent variables [51]. Following the notation in Heston(1993), the value of a contingent claim, U, can be written as a function of three independent variables $U_i = U_i(S(t), \nu(t), I(t), t)$. If we let the payoff of this contingent claim at expiry be $\left(\frac{S(t_i)}{I(t_i)} - 1\right)^2$ then under standard no-arbitrage arguments we obtain the following PDE (subscripts removed for the simplicity of the expression) for the value of the contingent claim [51]:

$$
\begin{cases}
\frac{\partial U_i}{\partial t} + \frac{1}{2}\nu S^2 \frac{\partial U_i^2}{\partial S^2} + \rho\sigma_V\nu S \frac{\partial U_i}{\partial S \partial \nu} + \frac{1}{2}\sigma_V^2\nu\frac{\partial U_i^2}{\partial \nu^2} + rS\frac{\partial U_i}{\partial S} + \left[\kappa^*(\theta^* - \nu)\right]\frac{\partial U_i}{\partial \nu} - rU_i + \delta\left(t_{i-1} - t\right)\frac{\partial U_i}{\partial I} = 0 \\
U_i(S, \nu, I, t) = \left(\frac{S}{I} - 1\right)^2
\end{cases}
$$

$$(11.14)$$

From *Feynman Kac Theorem* (see Appendix), the solution of this terminal PDE can be written as [51]:

$$
E_{\mathbb{Q}}\left[\left(\frac{S(t_i)}{I(t_i)} - 1\right)^2\right] = e^{rt_i}U_i\left(S(0), \nu(0), I(0), t_0\right). \qquad (11.15)
$$

This is the expression that needs to be solved to obtain the price of the variance swap. Therefore, it is sufficient to solve the PDE and terminal condition stated above to obtain the expected value. If properties of the Dirac delta function and the indicator variable I_t are implemented for $t \neq t_{i-1}$, the terminal PDE is to reduced to:

$$\begin{cases} \frac{\partial U_i}{\partial t} + \frac{1}{2}\nu S^2 \frac{\partial U_i^2}{\partial S^2} + \rho\sigma_V\nu S \frac{\partial U_i}{\partial S\partial\nu} + \frac{1}{2}\sigma_V^2\nu\frac{\partial U_i^2}{\partial\nu^2} + rS\frac{\partial U_i}{\partial S} + [\kappa^*(\theta^* - \nu)]\frac{\partial U_i}{\partial\nu} - rU_i = 0 \\ U_i(S,\nu,I,t) = \left(\frac{S}{I} - 1\right)^2 \end{cases}$$

$$(11.16)$$

This implies that the indicator variable has been eliminated from the equation except at $t = t_{i-1}$. However, as seen above the variable, I_t, is still present in the terminal condition. By definition, I_t experiences a jump in value at $t = t_{i-1}$. Under no-arbitrage theory, this variable should be continuous so that the option remains continuous across time. Therefore, the original PDE system is split into two systems across times $[0, t_{i-1}]$ and $[t_{i-1}, t_i]$ as jump conditions [38, 51]:

$$\begin{cases} \frac{\partial U_i}{\partial t} + \frac{1}{2}\nu S^2 \frac{\partial U_i^2}{\partial S^2} + \rho\sigma_V\nu S \frac{\partial U_i}{\partial S\partial\nu} + \frac{1}{2}\sigma_V^2\nu\frac{\partial U_i^2}{\partial\nu^2} + rS\frac{\partial U_i}{\partial S} + [\kappa^*(\theta^* - \nu)]\frac{\partial U_i}{\partial\nu} - rU_i = 0 \\ \lim_{t\uparrow t_{i-1}} U_i(S,\nu,I,t) = \lim_{t\downarrow t_{i-1}} U_i(S,\nu,I,t) \qquad\qquad\qquad\qquad , 0 \le t \le t_{i-1} \end{cases}$$

$$(11.17)$$

in which $I(t) = 0$ and

$$\begin{cases} \frac{\partial U_i}{\partial t} + \frac{1}{2}\nu S^2 \frac{\partial U_i^2}{\partial S^2} + \rho\sigma_V\nu S \frac{\partial U_i}{\partial S\partial\nu} + \frac{1}{2}\sigma_V^2\nu\frac{\partial U_i^2}{\partial\nu^2} + rS\frac{\partial U_i}{\partial S} + [\kappa^*(\theta^* - \nu)]\frac{\partial U_i}{\partial\nu} - rU_i = 0 \\ U_i(S,\nu,I,t) = \left(\frac{S}{I} - 1\right)^2 \qquad\qquad\qquad\qquad\qquad\qquad\qquad , t_{i-1} \le t \le t_i \end{cases}$$

$$(11.18)$$

in which $I(t) = S(t_{i-1})$.

The latter system is solved analytically using the generalised Fourier transform method.

Definition 21. (Fourier Transform, Zhaoli et. al. (2015) [38])

If $f(t)$ satisfies *Dirichlet* conditions, and absolutely integrable, then the Fourier transform $F(\omega)$ of $f(t)$ is defined as:

$$\mathfrak{F}\left[f(t)\right] = F(\omega) = \int\limits_{-\infty}^{\infty} f(t)e^{-i\omega t}dt \tag{11.19}$$

where $i^2 = -1$.

Proposition 7. (*Zhu and Lian (2011) [51]*)

By using Fourier transform algorithms, the PDE system of a derivative which has a payoff function of the form $U(S,\nu,t) = H(S)$ at expiry T and whose underlying asset follows the dynamics of the Heston stochastic volatility model (11.1) is:

$$\begin{cases} \frac{\partial U}{\partial t} + \frac{1}{2}\nu S^2 \frac{\partial U^2}{\partial S^2} + \rho\sigma_V \nu S \frac{\partial U}{\partial S\partial\nu} + \frac{1}{2}\sigma_V^2 \nu \frac{\partial U^2}{\partial \nu^2} + rS\frac{\partial U}{\partial S} + \left[\kappa^*(\theta^* - \nu)\right]\frac{\partial U}{\partial \nu} - rU = 0 \\ U\left(S,\nu,T\right) = H\left(S\right) \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad , t \leq T \end{cases} \tag{11.20}$$

and its solution can be expressed as:

$$U\left(x,\nu,t\right) = \mathfrak{F}^{-1}\left[e^{C(\omega,T-t)+D(\omega,T-t)\nu}\mathfrak{F}\left[H\left(e^x\right)\right]\right] \tag{11.21}$$

where $x = ln(S)$ and ω is the Fourier transform variable, and

$$
\begin{cases}
C(\omega, \tau) = r(\omega i - 1)\tau + \frac{\kappa^* \theta^*}{\sigma_V^2}\left[(a+b)\tau - 2ln\left(\frac{1-ge^{b\tau}}{1-g}\right)\right] \\[2mm]
D(\omega, \tau) = \frac{a+b}{\sigma_V^2}\left(\frac{1-e^{b\tau}}{1-ge^{b\tau}}\right) \\[2mm]
a = \kappa^* - \rho\sigma_V\omega i \\[2mm]
b = \sqrt{a^2 + \sigma_V^2(\omega^2 + \omega i)} \\[2mm]
g = \frac{a+b}{a-b}
\end{cases}
\tag{11.22}
$$

Proof. The outline of the proof is given below.

Considering PDE (11.20) above, the following transforms are used:

$$
\begin{cases}
\tau = T - t \\[2mm]
x = lnS
\end{cases}
\tag{11.23}
$$

The PDE then becomes:

$$
\begin{cases}
\frac{\partial U}{\partial \tau} = \frac{1}{2}\nu\frac{\partial U^2}{\partial x^2} + \rho\sigma_V\nu\frac{\partial U}{\partial x\partial \nu} + \frac{1}{2}\sigma_V^2\nu\frac{\partial U^2}{\partial \nu^2} + \left(r - \frac{1}{2}\nu\right)\frac{\partial U}{\partial x} + \left[\kappa^*(\theta^* - \nu)\right]\frac{\partial U}{\partial \nu} - rU \\[2mm]
U(x, \nu, 0) = H(e^x)
\end{cases}
\tag{11.24}
$$

The solution of this PDE is obtained from generalised Fourier transform with respect to x. Then utilising a Lemma on generalised Fourier transforms (see Lemma 8 in the Chapter Appendix), it follows that if the transform is applied to the PDE w.r.t x, the following terminal value problem for $\tilde{U}(\omega, \nu, \tau) = \mathfrak{F}[U(\omega, \nu, \tau)]$ is obtained:

$$
\begin{cases}
\frac{\partial \tilde{U}}{\partial \tau} = \frac{1}{2}\sigma_V^2 \nu \frac{\partial \tilde{U}^2}{\partial \nu^2} + [\kappa^*\theta^* - (\rho\sigma_\nu\omega i - \kappa)\nu] \frac{\partial \tilde{U}}{\partial \nu} - \left[(r\omega i - r) - \frac{1}{2}(\omega i + \omega^2)\nu\right]\tilde{U} \\
\tilde{U}(\omega, \nu, 0) = \mathfrak{F}[H(e^x)]
\end{cases}
$$

$$(11.25)$$

Following the discussion in Heston (1993) , the solution of the PDE can be assumed to have a form:

$$
\tilde{U}(\omega, \nu, \tau) = e^{C(\omega,\tau)+D(\omega,\tau)\nu}\tilde{U}(\omega, \nu, 0) \tag{11.26}
$$

Substituting the function above into the PDE ((11.25)) , $C(\omega, \tau)$ and $D(\omega, \tau)$ will satisfy ordinary differential equations (ODEs):

$$
\begin{cases}
\frac{dD}{d\tau} = \frac{1}{2}\sigma_V^2 D^2 + (\rho\sigma_\nu\omega i - \kappa^*)D - \frac{1}{2}(\omega i + \omega^2) \\
\frac{dC}{d\tau} = \kappa^*\theta^* D + r(\omega i - 1)
\end{cases}
\tag{11.27}
$$

where $C(\omega, 0) = 0$ and $D(\omega, 0) = 0$ are initial conditions. Considering the first equation in the system of ODEs above, the solution for D can be obtained analytically *(see the Chapter Appendix)*. The MATLAB function *ode45* can be used to obtained the solution for C via numerical integration. The solutions as in [51] are:

$$
\begin{cases}
C(\omega, \tau) = r(\omega i - 1)\tau + \frac{\kappa^*\theta^*}{\sigma_V^2}\left[(a+b)\tau - 2ln\left(\frac{1-ge^{b\tau}}{1-g}\right)\right] \\
D(\omega, \tau) = \frac{a+b}{\sigma_V^2}\left(\frac{1-e^{b\tau}}{1-ge^{b\tau}}\right)
\end{cases}
\tag{11.28}
$$

where

$$
\begin{cases}
a = \kappa^* - \rho\sigma_V\omega i \\[2mm]
b = \sqrt{a^2 + \sigma_V^2\left(\omega^2 + \omega i\right)} \\[2mm]
g = \frac{a+b}{a-b}
\end{cases}
\tag{11.29}
$$

Since the Fourier transform parameter, ω, still appears in $C\left(\omega,\tau\right)$ and $D\left(\omega,\tau\right)$, the solution of the original PDE can be obtained via the inverse Fourier transform as [38]:

$$
U\left(x,\nu,\tau\right) = \mathfrak{F}^{-1}\left[\tilde{U}\left(\omega,\nu,\tau\right)\right] = \mathfrak{F}^{-1}\left[e^{C(\omega,T-t)+D(\omega,T-t)\nu}\mathfrak{F}\left[H\left(e^x\right)\right]\right]
\tag{11.30}
$$

$\square$

In order to solve PDE ((11.18)) using the Proposition above, the payoff function $H(S)$ is substituted by $\left(\frac{S}{I}-1\right)^2$. Considering $x = lnS$ and taking into account that I is a constant, the generalized Fourier transform performed to the terminal payoff function $H(e^x) = \left(\frac{e^x}{I}-1\right)^2$ as:

$$
\mathfrak{F}\left[\left(\frac{e^x}{I}-1\right)^2\right] = 2\pi\left[\frac{\delta_{-2i}(\omega)}{I^2} - 2\frac{\delta_{-i}(\omega)}{I} + \delta_0(\omega)\right]
\tag{11.31}
$$

Now using Proposition (7), the solution of PDE ((11.18)) is:

$$
\begin{aligned}
U_i\left(S, \nu, I, t\right) &= \mathfrak{F}^{-1}\left[e^{C(\omega, t_i-t)+D(\omega, t_i-t)\nu}2\pi\left[\frac{\delta_{-2i}(\omega)}{I^2} - 2\frac{\delta_{-i}(\omega)}{I} + \delta_0(\omega)\right]\right] \\
&= \int_{-\infty}^{\infty}\frac{1}{2\pi}e^{C(\omega, t_i-t)+D(\omega, t_i-t)\nu}2\pi\left[\frac{\delta_{-2i}(\omega)}{I^2} - 2\frac{\delta_{-i}(\omega)}{I} + \delta_0(\omega)\right]e^{i\omega x}d\omega \\
&= \left[\frac{1}{I^2}e^{C(\omega, t_i-t)+D(\omega, t_i-t)\nu+i\omega x}\right]_{\omega=-2i} - \left[\frac{2}{I}e^{C(\omega, t_i-t)+D(\omega, t_i-t)\nu+i\omega x}\right]_{\omega=-i} \\
&\quad + \left[e^{C(\omega, t_i-t)+D(\omega, t_i-t)\nu+i\omega x}\right]_{\omega=0} \\
&= \frac{e^{2x}}{I^2}e^{\tilde{C}(t_i-t)+\tilde{D}(t_i-t)\nu} - \frac{2e^x}{I} + e^{-r(t_i-t)}
\end{aligned}
\tag{11.32}
$$

where $\tilde{C}(t) := C(-2i, t)$, $\tilde{D}(t) := D(-2i, t)$, and $x = lnS$ for $t_{i-1} \le t \le t_i$. These can be expressed as:

$$
\begin{cases}
\tilde{C}\left(\tau\right) = r\tau + \frac{\kappa^*\theta^*}{\sigma_V^2}\left[(\tilde{a}+\tilde{b})\tau - 2ln\left(\frac{1-\tilde{g}e^{\tilde{b}\tau}}{1-\tilde{g}}\right)\right] \\
\tilde{D}\left(\tau\right) = \frac{\tilde{a}+\tilde{b}}{\sigma_V^2}\left(\frac{1-e^{\tilde{b}\tau}}{1-\tilde{g}e^{\tilde{b}\tau}}\right)
\end{cases}
\tag{11.33}
$$

where

$$
\begin{cases}
\tilde{a} = \kappa^* - \rho\sigma_V\omega i \\
\tilde{b} = \sqrt{\tilde{a}^2 + 2\sigma_V^2} \\
\tilde{g} = \left(\frac{\tilde{a}}{\sigma_V}\right)^2 - 1 + \left(\frac{\tilde{a}}{\sigma_V}\right)\sqrt{\left(\frac{\tilde{a}}{\sigma_V}\right)^2 - 2}
\end{cases}
\tag{11.34}
$$

This forms the explicit form solution of the PDE ((11.18)) and completes the first stage of the two PDEs for calculating the expectation in Equation ((11.11)).

Next we solve PDE ((11.17)) to conclude the calculation of the expectation in Equation ((11.11)). By definition of $I := I(t)$ in Equation ((11.13)), it can be

deduced that $\lim\limits_{t\downarrow t_{i-1}} ln(S(t)) = lnI$. Then it follows from Equation ((11.32)) that:

$$
\begin{aligned}
\lim\limits_{t\downarrow t_{i-1}} U_i\left(S,\nu,I,t\right) &= \lim\limits_{t\downarrow t_{i-1}} \left[\frac{e^{2x}}{I(t)^2} e^{\tilde{C}(\Delta t)+\tilde{D}(\Delta t)\nu} - \frac{2e^x}{I(t)} + e^{-r(\Delta t)} \right] \\
&= \lim\limits_{t\downarrow t_{i-1}} \left[\frac{e^{2lnS_t}}{I(t)^2} e^{\tilde{C}(\Delta t)+\tilde{D}(\Delta t)\nu} - \frac{2e^{lnS_t}}{I(t)} + e^{-r(\Delta t)} \right] \\
&= \frac{e^{2lnI_t}}{I(t)^2} e^{\tilde{C}(\Delta t)+\tilde{D}(\Delta t)\nu} - \frac{2e^{lnI(t)}}{I(t)} + e^{-r(\Delta t)} \\
&= e^{\tilde{C}(\Delta t)+\tilde{D}(\Delta t)\nu} - 2 + e^{-r(\Delta t)}
\end{aligned}
\tag{11.35}
$$

To simplify the notation we can let the r.h.s of Equation ((11.35)) be denoted by $f(\nu)$ as:

$$
f\left(\nu\right) = e^{\tilde{C}(\Delta t)+\tilde{D}(\Delta t)\nu} + e^{-r(\Delta t)} - 2.
\tag{11.36}
$$

which is the terminal condition of PDE ((11.17)) in the period $0 \le t \le t_{i-1}$ according to the jump conditions $\lim\limits_{t\uparrow t_{i-1}} U_i\left(S,\nu,I,t\right) = \lim\limits_{t\downarrow t_{i-1}} U_i\left(S,\nu,I,t\right)$. As seen in Equation ((11.36)), the terminal condition for PDE ((11.17)) in period $0 \le t \le t_{i-1}$ contains a single independent variable ν. The advantage of this in solving the PDE ((11.17)) can be exploited using the following Proposition from Zhu and Lian (2011) [51] which is stated without proof for purposes of this dissertation.

Proposition 8. *(Zhu and Lian(2011) [51])*

The solution of the PDE

$$
\begin{cases}
\frac{\partial U}{\partial t} + \frac{1}{2}\nu S^2 \frac{\partial U^2}{\partial S^2} + \rho\sigma_V\nu S \frac{\partial U}{\partial S\partial\nu} + \frac{1}{2}\sigma_V^2\nu \frac{\partial U^2}{\partial\nu^2} + rS\frac{\partial U}{\partial S} + \left[\kappa^*(\theta^* - \nu)\right]\frac{\partial U}{\partial\nu} - rU = 0 \\
U\left(S,\nu,T\right) = f\left(\nu\right)
\end{cases}
\tag{11.37}
$$

is obtained analytically in the form,

$$U\left(S, \nu, t\right) = \int_0^\infty e^{-r(T-t)} f\left(\nu(T)\right) \int_0^\infty e^{-r(T-t)} f\left(\nu(T)\right) p\left(\nu(T)|\nu(t)\right) d\nu(T) d\nu(T)$$

(11.38)

where

$$p\left(\nu(T)|\nu(t)\right) = ce^{-W-V} \left(\frac{V}{W}\right)^{q/2} K_q\left(2\sqrt{WV}\right)$$

(11.39)

with

$$c = \frac{2\kappa^*}{\sigma_V^2\left(1-e^{-\kappa^*(T-t)}\right)}$$

$$W = c\nu_t e^{-\kappa^*(T-t)}$$

$$V = c\nu_T$$

$$q = \frac{2\kappa^*\theta^*}{\sigma_V^2} - 1$$

(11.40)

and $K_q(\cdot)$ is the Bessel function of the first kind of order q.

Then it follows from Proposition (8) that the solution of the PDE ((11.17)) can be expressed in the form:

$$U_i\left(S, \nu, I, t\right) = \int_0^\infty e^{-r(t_{i-1}-t)} f\left(\nu(t_{i-1})\right) p\left(\nu(t_{i-1})|\nu(t)\right) d\nu(t_{i-1})$$

(11.41)

where $f(\nu(t_{i-1}))$ and $p\left(\nu(t_{i-1})|\nu(t)\right)$ is transition probability density function of OU process. The expectation in Equation ((11.11)) has been obtained for the Case $i > 1$ by solving PDE ((11.18)) and PDE ((11.17)).

$$\mathbb{E}_0^{\mathbb{Q}}\left[\left(\frac{S(t_i) - S(t_{i-1})}{S(t_{i-1})}\right)^2\right] = e^{rt_i} U_i\left(S(0), \nu(0), I(0), 0\right)$$

(11.42)

$$= \int_0^\infty e^{r(\Delta t)} f\left(\nu(t_{i-1})\right) p\left(\nu(t_{i-1})|\nu(0)\right) d\nu(t_{i-1}) \tag{11.43}$$

The form of $f(\nu)$ used in deriving a solution for PDE ((11.18)) can now be implemented again to obtain a closed form solution. Zhu and Lain (2011), set $\chi_t^2 = 2c\nu_t$. This implies that the stochastic variable is subject to a non-central chi-squared distribution, $\chi(t)^2 \sim \chi^2(2V; 2q+2, 2W)$ with a pdf denoted $p_{\chi(t)^2}(x)$. Then it follows that $p\left(\nu(T)|\nu(t)\right) = 2cp_{\chi(T-t)^2}(2c\nu(t))$. The parameters are defined as in Proposition (8) . Then, using the MGF of a non-central chi-squared distribution it follows that:

$$
\begin{aligned}
\mathbb{E}_0^{\mathbb{Q}}\left[\left(\frac{S(t_i) - S(t_{i-1})}{S(t_{i-1})}\right)^2\right] &= \int_0^\infty e^{r(\Delta t)} f\left(\nu(t_{i-1})\right) p\left(\nu(t_{i-1})|\nu(0)\right) d\nu(t_{i-1}) \\
&= e^{r(\Delta t)}\mathbb{E}_0^{\mathbb{Q}}\left[e^{\tilde{C}(\Delta t)+\tilde{D}(\Delta t)\nu(t_{i-1})} + e^{-r(\Delta t)} - 2\right] \\
&= e^{r(\Delta t)}\left(e^{\tilde{C}(\Delta t)}\mathbb{E}_0^{\mathbb{Q}}\left[e^{\tilde{D}(\Delta t)\nu(t_{i-1})}\right] + e^{-r(\Delta t)} - 2\right) \\
&= e^{r(\Delta t)}\left(e^{\tilde{C}(\Delta t)}\mathbb{E}_0^{\mathbb{Q}}\left[e^{\tilde{D}(\Delta t)\frac{\chi_{t_i}^2}{2c}}\right] + e^{-r(\Delta t)} - 2\right) \tag{11.44} \\
&= e^{r(\Delta t)}\left(e^{\tilde{C}(\Delta t)}\left[(1 - 2\xi)^{-(q+1)} e^{\frac{2W\xi}{1-2\xi}}\right]_{\xi=\frac{\tilde{D}(\Delta t)}{2c}} + e^{-r(\Delta t)} - 2\right) \\
&= e^{r(\Delta t)}\left(e^{\tilde{C}(\Delta t)+\frac{W\tilde{D}(\Delta t)}{c-\tilde{D}(\Delta t)}}\left[\left(\frac{c}{c - \tilde{D}(\Delta t)}\right)^{\frac{2\kappa^*\theta^*}{\sigma_V^2}} e^{\frac{2W\Phi}{1-2\Phi}}\right]_{\Phi=\frac{\tilde{D}(\Delta t)}{2c}}\right) \\
&\quad + e^{r(\Delta t)}\left(e^{-r(\Delta t)} - 2\right) \tag{11.45}
\end{aligned}
$$

The parameters c and W were determined by replacing time $t = 0$ and $T = t_{i-1}$ in Equation ((11.40)). Thus,

$$f_i\left(\nu(0)\right) = e^{\tilde{C}(\Delta t)+\frac{c_i e^{-\kappa^* t_{i-1}}}{c_i - \tilde{D}(\Delta t)}\tilde{D}(\Delta t)\nu_0}\left(\frac{c_i}{c_i - \tilde{D}(\Delta t)}\right)^{\frac{2\kappa^*\theta^*}{\sigma_V^2}} + e^{-r(\Delta t)} - 2 \tag{11.46}$$

where,

$$c_i = \frac{2\kappa^*}{\sigma_V^2 \left(1 - e^{-\kappa^* t_{i-1}}\right)}$$

then the solution of the expectation is obtained as:

$$\mathbb{E}_0^{\mathbb{Q}} \left[\left(\frac{S(t_i) - S(t_{i-1})}{S(t_{i-1})} \right)^2 \right] = e^{r(\Delta t)} f_i \left(\nu(0) \right). \qquad (11.47)$$

Then the sum in Equation ((11.10)) for the calculation of the variance swap case be determined for all $i > 1$.

Case $i = 1$ For the case $i = 1$, $t_{i-1} = 0$ and $S(t_{i-1}) = S(0)$ which is the current price of the underlying, and therefore known. It is important to note that for the case $i > 1$, the $S(t_i)$ were unknown values. The expectation in Equation ((11.11)) then reduces to a form:

$$\mathbb{E}_0^{\mathbb{Q}} \left[\left(\frac{S(t_i)}{S(0)} - 1 \right)^2 \right] \qquad (11.48)$$

The expectation can be obtained as:

$$\mathbb{E}_0^{\mathbb{Q}} \left[\left(\frac{S(t_i)}{S(0)} - 1 \right)^2 \right] = e^{r(\Delta t)} f \left(\nu(0) \right). \qquad (11.49)$$

Therefore, the case $i > 1$ and case $i = 1$ can be combined to deduce the fair value of a discretely sampled variance swap as:

$$K_{var} = \mathbb{E}_0^{\mathbb{Q}}[\sigma_R^2(T)] = 100^2 \cdot \frac{e^{r(\Delta t)}}{T} \left[f \left(\nu(0) \right) + \sum_{i=2}^{N} f_i \left(\nu(0) \right) \right] \qquad (11.50)$$

where N is the finite number of sampling times of the swap contract.

11.4   The Continuously-sampled Variance Swap

The continuously-sampled Heston (1993) model case has been proposed by many researchers such as Elliot and Siu (2007) [30] and Swishchuk (2011) [49] to price variance swaps for stochastic volatility. In the continuously-sampled case the expected realised variance is given by:

$$K_{var} = \mathbb{E}_{\mathbb{Q}}[\sigma_R^2(T)] = 100^2.\mathbb{E}_{\mathbb{Q}}\left[\frac{1}{T}\int_0^T \nu(t)dt\right] \tag{11.51}$$

Swishchuk (2004) [48], defines a transformation of the variance process by:

$$h(t) := e^{\kappa t}(\nu(t) - \theta) \tag{11.52}$$

and then applying Itô's formula we obtain the stochastic equation for $h(t)$ as :

$$dh(t) = \sigma_\nu e^{\kappa t}\sqrt{e^{-\kappa t}h(t) + \theta}dW(t) \tag{11.53}$$

then the solution for Equation (11.53) is [48]:

$$h(t) = \nu(0) - \theta + \tilde{W}\left(\phi(t)^{-1}\right) \tag{11.54}$$

then from Equation (11.52) :

$$\nu(t) = e^{-\kappa t}\left(\nu(0) - \theta + \tilde{W}\left(\phi(t)^{-1}\right)\right) + \theta \tag{11.55}$$

where $\tilde{W}(\cdot)$ is a $\mathcal{F}(t)$-measurable Wiener process and $\phi(t)$ is defined as:

$$\phi(t) = \sigma_{\nu}^{-2} \int_{0}^{t} \left\{ e^{\kappa\phi(s)} \left(\nu(0) - \theta + \tilde{W}\left(\phi(t)^{-1}\right) \right) + \theta e^{2\kappa\phi(s)} \right\}^{-1} ds.$$

Then it follows that:

$$\mathbb{E}\left(\nu(t)\right) = e^{-\kappa t}\left(\nu(0) - \theta\right) + \theta. \tag{11.56}$$

The expectation $\mathbb{E}_{\mathbb{Q}}[\sigma_R^2(T)]$ and thus the delivery price of the variance swap is

$$\begin{aligned}
\mathbb{E}_{\mathbb{Q}}\left(\sigma_R^2(T)\right) &= \frac{1}{T}\int_{0}^{T} \mathbb{E}_{\mathbb{Q}}\left(\nu(t)\right) dt \\
&= \frac{1}{T}\int_{0}^{T} e^{-\kappa t}\left(\nu(0) - \theta\right) + \theta dt \\
&= \frac{1}{-\kappa T}\left[e^{-\kappa t}\left(\nu(0) - \theta\right)\right]_0^T + \theta \\
&= \frac{1 - e^{-\kappa T}}{\kappa T}\left(\nu(0) - \theta\right) + \theta
\end{aligned}$$

Which can be re-written in familiar notation as :

$$\mathbb{E}_{\mathbb{Q}}\left(\sigma_R^2(T)\right) = \nu(0)\frac{1 - e^{-\kappa^* T}}{\kappa^* T} + \theta\left(1 - \frac{1 - e^{-\kappa^* T}}{\kappa^* T}\right) \tag{11.57}$$

The expression above is also found in Brockhaus and Long (2000) [16] and is translated as the weighted average spot variance, $\nu(0)$, and the mean of the variance over a long term, θ [51]. Theoretically, the discrete model should converge to the continuous model as the sampling times increases that is $\Delta t \to 0$. This will be explored practically in the next section.

Proposition 9. *(Convergence, Zhaoli et. al. (2015) [51, 38])*

If we consider the stochastic volatility in Equation (11.1) , then the risk-neutral price of a variance swap obtained from a finite number of discrete sampling times can be approximated by a continuum of sampling times as the sampling times increase that is:

$$\lim_{\Delta t \to 0} \frac{e^{r(\Delta t)}}{T} \left[f\left(\nu(0)\right) + \sum_{i=2}^{N} f_i\left(\nu(0)\right) \right] = \nu(0) \left(\frac{1 - e^{-\kappa^* T}}{\kappa^* T} \right) + \theta^* \left(1 - \frac{1 - e^{-\kappa^* T}}{\kappa^* T} \right)$$

(11.58)

where the parameters are defined as in Equation (11.40) and Equation (11.46).

Proof. Using the L'Hopital's rule, it can be deduced that:

$$\lim_{\Delta t \to 0} \tilde{C}\left(\Delta t\right) = \lim_{\Delta t \to 0} \left(r\Delta t + \frac{\kappa^* \theta^*}{\sigma_V^2} \left[(\tilde{a} + \tilde{b})\Delta t - 2ln\left(\frac{1 - \tilde{g}e^{\tilde{b}\Delta t}}{1 - \tilde{g}} \right) \right] \right) = 0 \quad (11.59)$$

$$\lim_{\Delta t \to 0} \tilde{D}\left(\Delta t\right) = \lim_{\Delta t \to 0} \left(\frac{\tilde{a} + \tilde{b}}{\sigma_V^2} \left(\frac{1 - e^{\tilde{b}\Delta t}}{1 - \tilde{g}e^{\tilde{b}\Delta t}} \right) \right) = 0 \quad (11.60)$$

then,

$$\lim_{\Delta t \to 0} f\left(\nu(0)\right) = \lim_{\Delta t \to 0} \left(e^{\tilde{C}(\Delta t) + \tilde{D}(\Delta t)\nu(0)} + e^{-r(\Delta t)} - 2 \right) = 0 \quad (11.61)$$

$$\lim_{\Delta t \to 0} \frac{f\left(\nu(0)\right)}{\Delta t} = \lim_{\Delta t \to 0} \left(\frac{e^{\tilde{C}(\Delta t) + \tilde{D}(\Delta t)\nu(0)} + e^{-r(\Delta t)} - 2}{\Delta t} \right) = \nu_0 \quad (11.62)$$

and

$$\lim_{\Delta t \to 0} \frac{f_i\left(\nu(0)\right)}{\Delta t} = \nu(0)e^{-\kappa^*(i-1)\Delta t} + \theta^*\left(1 - e^{-\kappa^*(i-1)\Delta t}\right) \tag{11.63}$$

thus we can derive the limit as:

$$
\begin{aligned}
\lim_{\Delta t \to 0} \frac{e^{r(\Delta t)}}{T}\left[f\left(\nu(0)\right) + \sum_{i=2}^{N} f_i\left(\nu(0)\right)\right] &= \frac{1}{T}\lim_{\Delta t \to 0}\sum_{i=2}^{N}\Delta t\left(\nu_0 + \frac{f_i\left(\nu(0)\right)}{\Delta t}\right) \\
&= 0 + \frac{1}{T}\lim_{\Delta t \to 0}\Delta t\sum_{i=2}^{N}\left(\nu(0)e^{-\kappa^*(i-1)\Delta t} + \theta\left(1 - e^{-\kappa^*(i-1)\Delta t}\right)\right) \\
&= \frac{1}{T}\int_{0}^{T}\left(\nu(0)e^{-\kappa^* t} + \theta\left(1 - e^{-\kappa^* t}\right)\right) dt \\
&= \nu(0)\left(\frac{1 - e^{-\kappa^* T}}{\kappa^* T}\right) + \theta^*\left(1 - \frac{1 - e^{-\kappa^* T}}{\kappa^* T}\right).
\end{aligned}
$$

which concludes the proof. $\square$

11.5   Numerical examples

The results which were derived are practically implemented in this section. The stock and price processes as in the Heston Model (11.1) are simulated using Monte Carlo methods. To be able to simulate the model, calibration is done to minimise the mean square-error between the market and model prices of European style calls. The closed-form exact solution for the discretely and continuously sampled variance swaps was extensively studied. The price of variance swaps is compared in these two scenarios of the Heston model. The price of variance swaps with increase in years to maturity of the variance swap contract is also studied in a practical setting.

11.5.1 Model Calibration

To approximate the Heston model parameters, a popular approach is loss function estimation (that is mean square-error (MSE), sum of square-error (SSE) etc). In this methodology the model parameters are chosen so that market option prices are as close as possible to the model option prices [45]. Let the Heston parameters we wish to approximate be the vector $\Theta := (\nu(0), \kappa^*, \theta^*, \sigma_\nu, \rho)$. If there are a set of N_T maturities $\tau_i(t = 1, 2, ..., N_T)$ and a set of N_K strikes $K_j(j = 1, 2, ...N_K)$, then each strike-maturity combination has a market call option price $C^{market}(\tau_i, K_j) := C_{ij}^{market}$ and corresponding model price $C^{\Theta}(\tau_i, K_j) := C_{ij}^{\Theta}$. The calibration problem in this dissertation is set as

$$min \frac{1}{N} \sum_{i,j} \left(C_{ij}^{market} - C_{ij}^{\Theta} \right)^2. \tag{11.64}$$

This is the minimum MSE for each combination of market and model prices. The parameters obtained from the MATLAB calibration are:

Table 4: Parameters estimates for the Heston Model

$\nu(0)$	κ^*	θ^*	σ_ν	ρ	μ
$6.6602{\times}10^{-7}$	0.01238	0.00735	0.003446	-0.7576	$3.8344{\times}10^{-4}$

μ is set mean of the FTSE/JSE Top 40 Index data in Chapter 3. The maturity is chosen as one year, i.e. $T = 1$ and the initial price of the underlying set at earliest price in the data set, $S(0) = 263.68$. In the calibration, the Heston Model's vanilla call option values were compared with the market values (assumed to be Black-Scholes model prices). The calibration results using the parameters above are shown below:

Figure 11.1: Calibration of the Heston model to the FTSE/JSE Top 40 Index
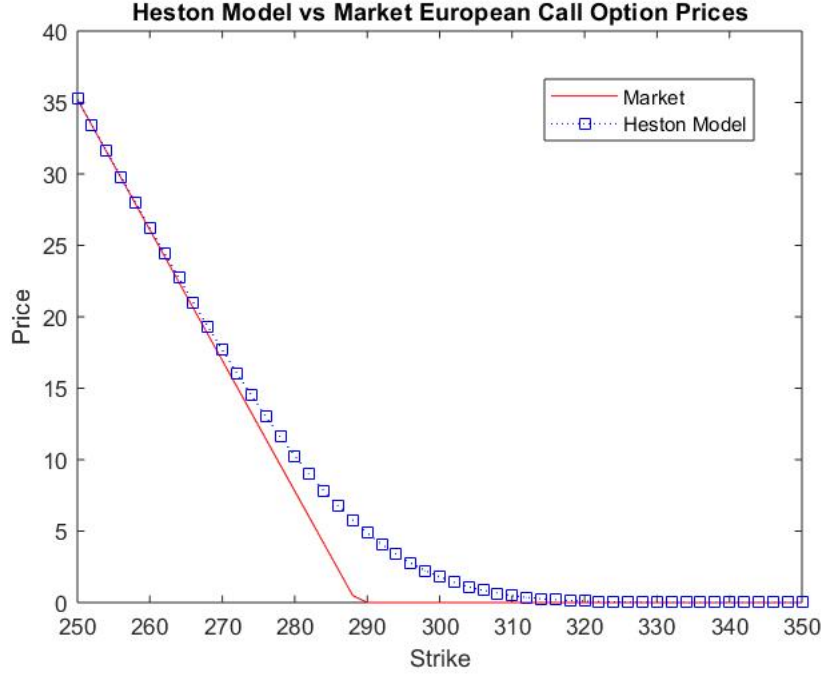
11.5.2   Monte Carlo simulation of the Heston Stochastic volatility

Monte Carlo simulation are a number of techniques to 'artificially' generate the movement of the stock and variance in the Heston Model (11.1) over given period. There are many Monte Carlo techniques which can be implemented to this effect. Rouah (2013), in his text mentions standard approaches such as the Euler and the Milstein, or the implicit Milstein scheme. The advantage of these schemes is that they are easy to understand, and their convergence properties are well-known [45]. Recalling the dynamics of the variance process in the Heston Model (11.1), a simple Euler discretization of the variance process could be:

$$\nu(t+1) - \nu(t) = \kappa^*(\theta^* - \nu(t))\Delta t + \sigma_\nu \sqrt{\nu(t)}\sqrt{\Delta t}Z \qquad (11.65)$$

where $Z \sim N(0,1)$. However, the problem with this simple discretization is that the

variance process could result in negative values even if the Feller condition, $2\kappa\theta \geq \sigma_\nu^2$, is met [45]. To make the variance process, $\nu(t)$, non-negative the following schemes can be used:

1. *Full trunctation* - the negative values for $\nu(t)$ are floored at 0. This means that $\nu(t)$ is replaced by $\nu(t)^+ = max\,(0, \nu(t))$.

2. *Reflection* - negative values for $\nu(t)$ are reflected with $-\nu(t)$. This means that $\nu(t)$ is replaced by $|\nu(t)|$.

The flaw in the Reflection approach is that large negative variances are reflected to large positive variances which means that realisations of low volatility are transformed into high volatility. Various approaches to simulate values of the variance process, $\nu(t)$, which do not produce negative values are investigated in the CIR variance process [45]. A clever approach could be to simulate the log process or the square-root process of $\nu(t)$ then exponentiate and square the results respectively. Simulating the stock process, $S(t)$, then becomes straight forward. However, the log price process $x(t) = lnS(t)$ can instead be simulated and the result exponentiated. From Itô's Lemma, the Heston Model (11.1) log stock price process becomes :

$$
\begin{cases}
dx(t) &= \left(\mu - \frac{1}{2}\nu(t)\right) dt + \sqrt{\nu(t)}dW(t)^1 \\
d\nu(t) &= \kappa^* \left(\theta^* - \nu(t)\right) dt + \sigma_\nu \sqrt{\nu(t)}dW(t)^2
\end{cases}
\tag{11.66}
$$

Simulating the Heston model follows the following algorithm [45]:

1. Set the $S(0)$ to the current spot price (or $x(0)$ to the current log spot price), and set $\nu(0)$ to the current variance.

2. Generate two independent standard normal random variables Z_1 and Z_2. Define $Z_\nu = Z_1$ and $Z_S = \rho Z_\nu + \sqrt{1 - \rho^2} Z_2$. Generate the Wiener process by $dW(t)^1 = \sqrt{\Delta t} Z_S$ and $dW(t)^2 = \sqrt{\Delta t} Z_\nu$.

3. Obtain the updated variance $\nu(t + 1)$.

4. Given $\nu(t+1)$, obtain $S(t+1)$ (or $x(t+1)$ in the log process case) then return to (1).

Note that $E[Z_S] = E[Z_\nu] = 0$ and then $E[Z_\nu Z_S] = \rho E[Z_\nu^2] + \sqrt{1 - \rho^2} E[Z_1 Z_2] = \rho$.

Figure 11.2: Example: Variance Process under the Heston Model
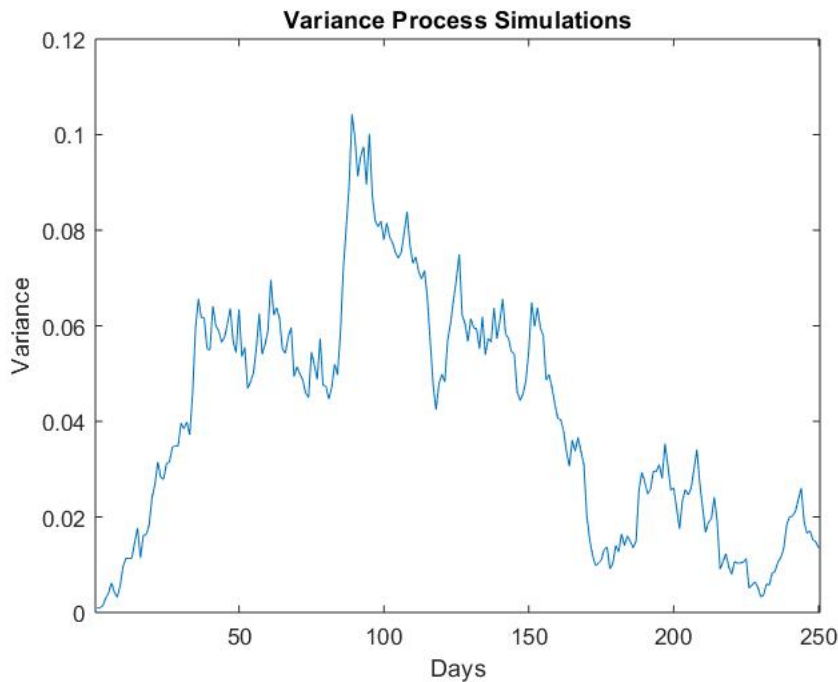
The variance process above can have multiple simulations for the same number of days to produce the following:

Figure 11.3: Simulations of the Variance Process
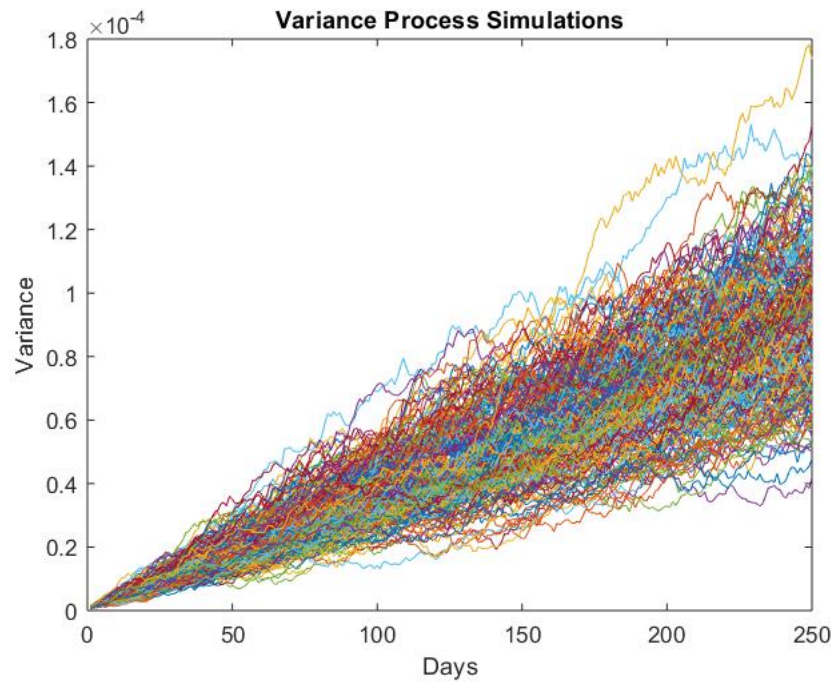
It should be noted that the variance process is non-negative in the Figure above. The reflection scheme was applied to this effect. This is also reproduced in a more appealing graph as:

Figure 11.4: Variance Process representation

Figure 11.5: Simulations of the Stock Price under the Heston Model

The price of variance swaps based on the parameters of the Heston model calibrated

to the JSE Top 40 Index market data are presented below. The variance swap strikes for the discretely-sampled and continuously-sampled are compared for increasing sampling times.

Table 5: Variance Swap Strike for the Discrete and Continuous Models

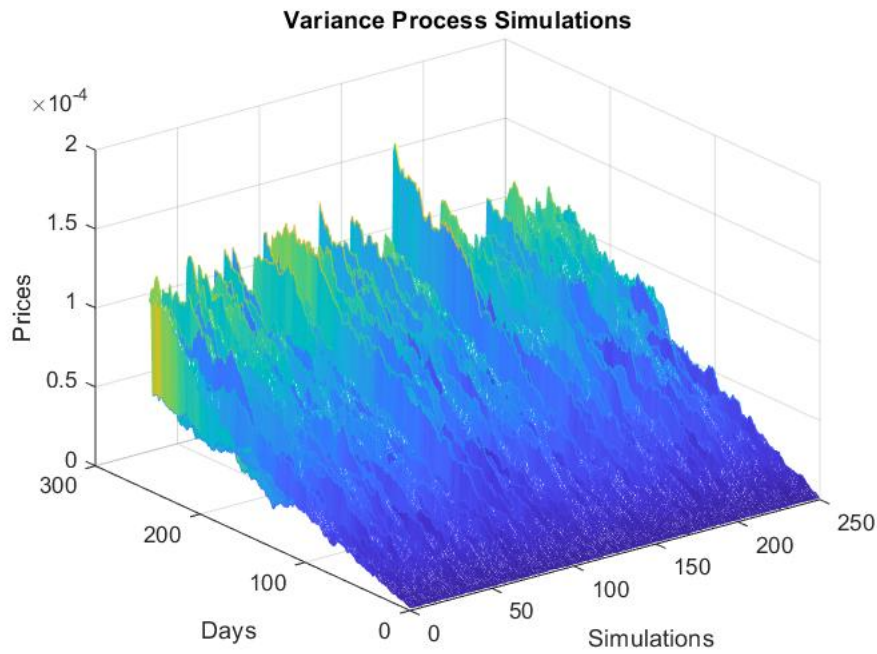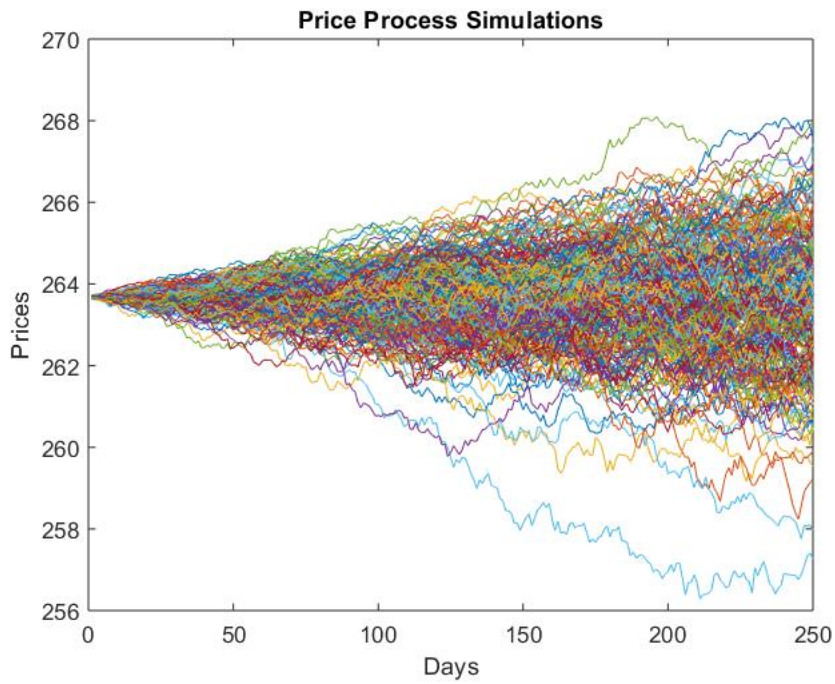| Sampling Frequency	Discrete model K_{var}	Continuous model K_{var}
N=4	4.5990036×10^{-5}	4.5971609×10^{-5}
N=12	4.5977199×10^{-5}	4.5971609×10^{-5}
N=52	4.5972850×10^{-5}	4.5971609×10^{-5}
N=252	4.5971863×10^{-5}	4.5971609×10^{-5}
N=504	4.5971735×10^{-5}	4.5971609×10^{-5}
N=1512	4.5971680×10^{-5}	4.5971609×10^{-5}
N=5 000	4.5971650×10^{-5}	4.5971609×10^{-5}
N=10 000	4.5971620×10^{-5}	4.5971609×10^{-5}
N=100 000	4.5971631×10^{-5}	4.5971609×10^{-5}

Thus the discrete model converges to the continuous model as the frequency of sampling increases that is $\Delta t \to 0$. However, in this case converge was slow with sampling required for up to 1512 times a day to reach the continuous model result. The information in Table 5 above is plotted in Excel as:

Figure 11.6: Convergence of the discrete approximation on to the continuous model
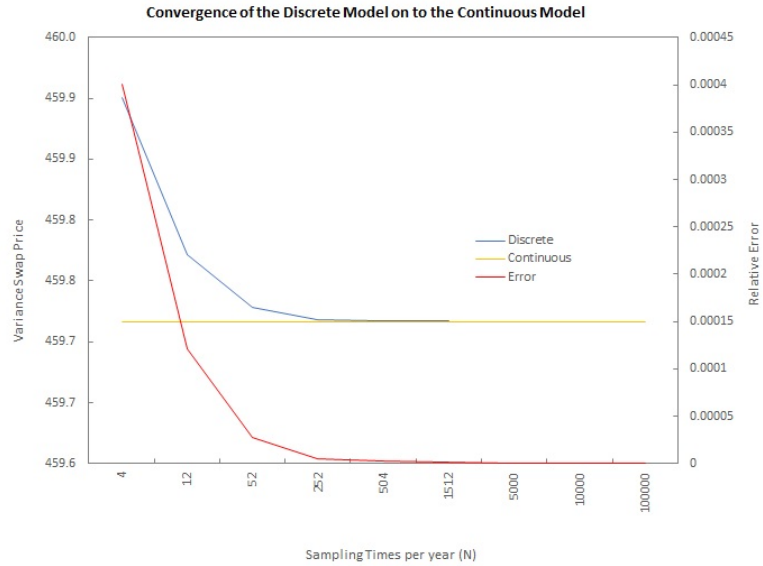
The variance swap prices have been scaled by a factor of 10^7.

Comments: Figure 11.6 above shows that for a few sampling times the relative error between the discrete and the continuous model is large. The relative error however decreases with increasing sampling times and becomes negligible as $N \to \infty$ that is as $\Delta t \to 0$. This depicts the assertions in Proposition 9.

Next, the continuously-sampled price for variance swaps relationship with the maturity is investigated.

Figure 11.7: Variance swap price against maturity (years)
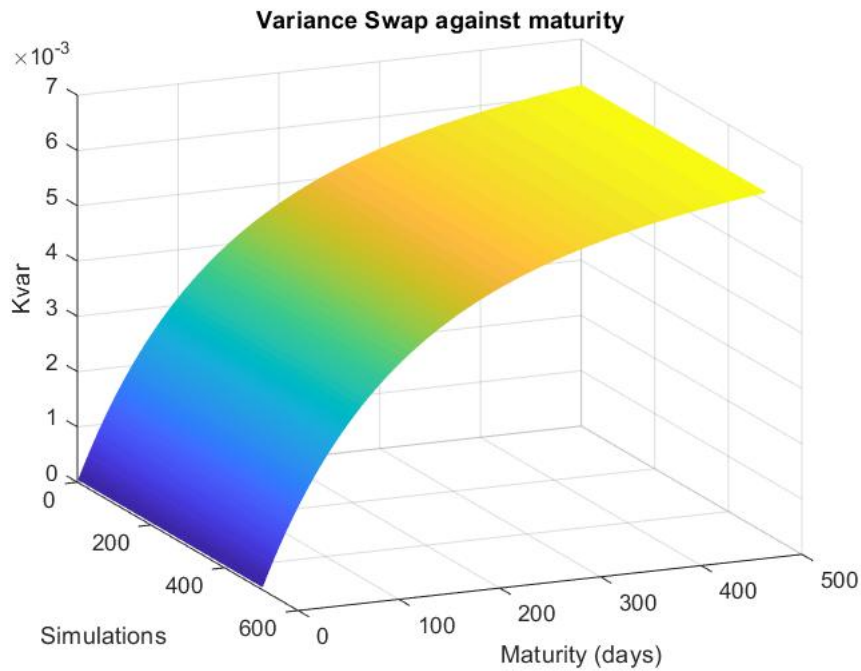
In this simulation $\nu_0 < \theta$.

Comments: With the parameters given in Table 4, the rate of change of the price for variance swaps with maturity is high for short-dated maturities but however, gradually becomes low for long-dated maturities. This is observed from the steepness of the curve.

The calibration of the model to market data resulted in a specific set of parameters and the prices as shown in the previous figure. The general sensitivity of the variance swap prices to changes in parameters is studied below:

Figure 11.8: Sensitivity of variance prices to parameter changes
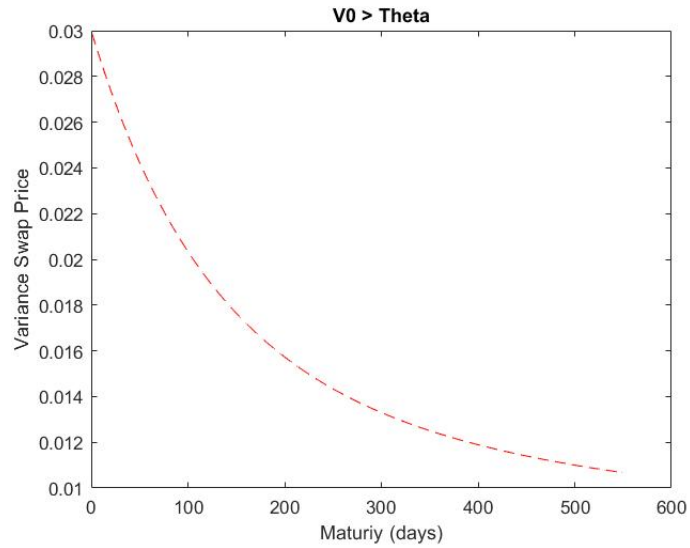
$\kappa* = 0.01238, \theta* = 0.000735,$
$\sigma_v = 0.0034458, \nu_0 = 0.03$
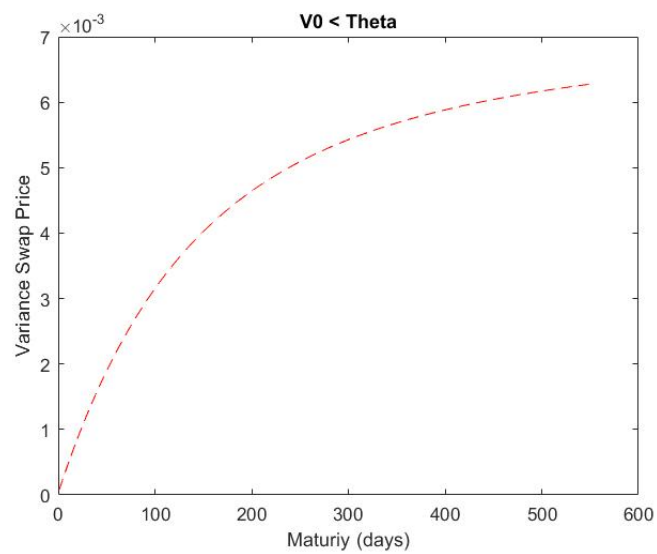
$\kappa* = 0.01238, \theta* = 0.000735,$
$\sigma_v = 0.0034458, \nu_0 = 0.00003$

Figure 11.9: Sensitivity of variance prices to kappa
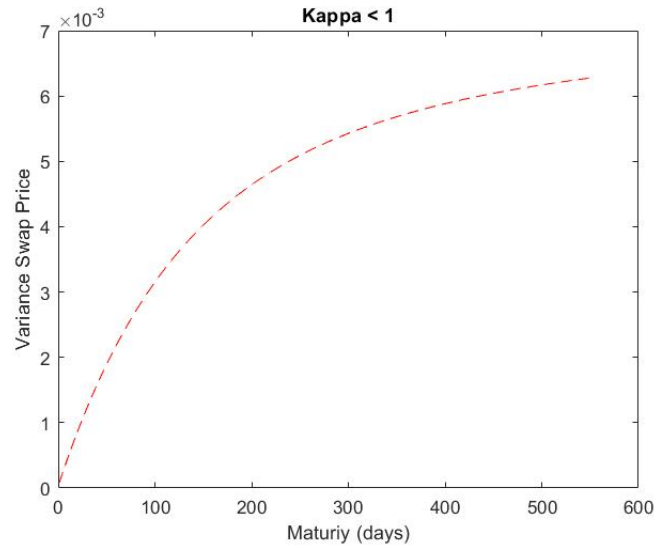
$\kappa* = 0.01238, \theta* = 0.000735,$
$\sigma_v = 0.0034458, \nu_0 = 0.00003$
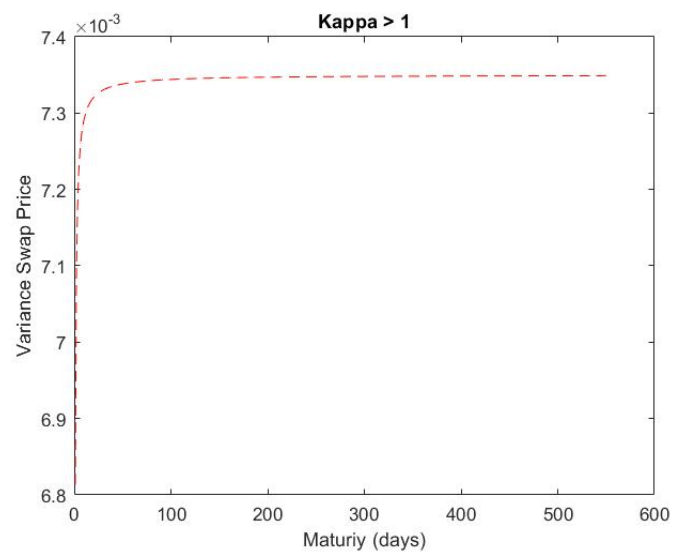
$\kappa* = 12.38, \theta* = 0.000735,$
$\sigma_v = 0.0034458, \nu_0 = 0.00003$

11.6   Conclusions

The Heston (1993) model provides an alternative to the Black-Scholes model for pricing derivatives. The Heston Model (11.1) provides a more realistic description of the paths of the stock price and volatility processes. However, the analytical formula for the discretely-sampled variance swaps was not obtained as easy as in the continuously-sampled case. The analytical formula for the discretely-sampled variance swaps was obtained from solving two-part PDE using Fourier transform techniques under Feynman-Kac theory. In a practical setting, to obtain the parameters of the Heston model in a risk-neutral setting, the parameters had to be estimated using historical data. The calibration problem was set in to minimise the MSE (mean squared error) between the market European call prices and the European call prices under the Heston Model. In the approximation of the parameters lies a challenge of the long computational time required to produce the parameter estimates. This makes the Heston Model particularly unattractive.

Furthermore, it was shown that the continuously-sampled model overstates the price of variance swaps when compared to the discretely-sampled model for a few sampling times. However, the discrete model was both shown analytically and practically to converge to the continuous model at the sampling times tend to infinity. The sensitivity of the variance swap prices to changes in parameters was investigated. Furthermore, the rate of change of the price of continuously-sampled variance swaps with maturity was shown to be higher for shorter maturities than for longer-maturities.

Chapter Appendix

Definition 22. (Dirac Delta Function)

The Dirac delta $\delta(x)$ for $x \in \mathbb{R}$ is a function defined as:

$$\delta(x) = \begin{cases} +\infty, & x = 0 \\ 0 & x \neq 0 \end{cases}$$

and

$$\int\limits_{-\infty}^{\infty} \delta(x)dx = 1$$

Some useful function of the Dirac Delta function is that $\delta(x-a) = 0$ for $x \neq a$.

The Fourier Transform is defined as:

Definition 23. (Fourier Transform, [38])

If $f(t)$ satisfies *Dirichlet* conditions, and absolutely integrable, then the Fourier transform $F(\omega)$ of $f(t)$ is defined as:

$$\mathfrak{F}[f(t)] = F(\omega) = \int\limits_{-\infty}^{\infty} f(t)e^{-i\omega t}dt \tag{11.67}$$

where $i^2 = -1$.

The Fourier Transform is then coupled with the following theorems to be used in the valuation of the expression of the expected value of realised variance under the risk neutral measure.

Lemma 7. *(Zhaoli et. al (2015), [38])*

If f(t) → 0 for |t| → ∞, then

$$\mathfrak{F}\left[\frac{df(t)}{dt}\right] = i\omega\mathfrak{F}\left[f(t)\right]. \tag{11.68}$$

In general,

$$\mathfrak{F}[a_n f^{(n)}(t) + a_{n-1} f^{(n-1)}(t) + \cdots + a_1 f'(t) + a_0 f]$$

$$= \left[a_n(i\omega)^n + a_{n-1}(i\omega)^{n-1} + \cdots + a_1(i\omega) + a_0\right]\mathfrak{F}[f(t)]. \tag{11.69}$$

Lemma 8. *(Zhaoli et.al (2015), [38])*

Let $\delta_s(\omega)$ be the Dirac delta function, then $\mathfrak{F}[e^{ist}] = 2\pi\delta_s(\omega)$. If $\phi(t)$ is continuous function, then

$$\int\limits_{-\infty}^{\infty} \delta_s(t)\phi(t)dt = \phi(s).$$

The derivation of the *Heston PDE* In this section the Heston PDE is derived. Unlike the Black-Scholes case where the source of stochasticity is the underlying stock, in the Heston case the random changes in volatility also need to be hedged in order to form a riskless portfolio [33]. As in the text by Gatheral (2011) a portfolio, $\Pi := \Pi_t$, which is made up of the option being priced $V = V(S, v, t)$, Δ units of the stock S, and ϕ units of another option $U = U(S, v, t)$ that is used to hedge the volatility is considered. The value of the portfolio is then:

$$\Pi = V + \Delta S + \phi U \tag{11.70}$$

Assuming that this portfolio is self-financing then

$$d\Pi = dV + \Delta dS + \phi dU \tag{11.71}$$

Then Itô's Lemma is then applied to dV and differentiation with respect to variables S, t and ν is conducted. We then obtain:

$$dV = \frac{\partial V}{\partial t}dt + \frac{\partial V}{\partial S}dS + \frac{\partial V}{\partial \nu}d\nu + \frac{1}{2}\nu S^2\frac{\partial^2 V}{\partial S^2}dt + \frac{1}{2}\nu\sigma^2\frac{\partial^2 V}{\partial \nu^2}dt + \rho\sigma\nu S\frac{\partial^2 V}{\partial S\partial \nu}dt$$

It is easy to show that applying Itô's Lemma to dU results in a similar PDE with terms in V replaced by U [33]. By grouping terms in dt, dS and $d\nu$, the change in the value of the portfolio in a time dt is given by:

$$d\Pi = \left\{\frac{\partial V}{\partial t} + \frac{1}{2}\nu S^2\frac{\partial^2 V}{\partial S^2} + \frac{1}{2}\nu\sigma^2\frac{\partial^2 V}{\partial \nu^2} + \rho\sigma\nu S\frac{\partial^2 V}{\partial S\partial \nu}\right\}dt \tag{11.72}$$

$$+\phi\left\{\frac{\partial U}{\partial t} + \frac{1}{2}\nu S^2\frac{\partial^2 U}{\partial S^2} + \frac{1}{2}\nu\sigma^2\frac{\partial^2 U}{\partial \nu^2} + \rho\sigma\nu S\frac{\partial^2 U}{\partial S\partial \nu}\right\}dt$$

$$+\left\{\frac{\partial V}{\partial S} + \phi\frac{\partial U}{\partial S} + \Delta\right\}dS + \left\{\frac{\partial V}{\partial \nu} + \phi\frac{\partial U}{\partial \nu}\right\}d\nu$$

For the portfolio to be instantaneously risk-free (that is hedges against stock and volatility in this case) the terms in dS and $d\nu$ must be equal to zero [33]. Then it follows that:

$$\frac{\partial V}{\partial S} + \phi\frac{\partial U}{\partial S} + \Delta = 0 \tag{11.73}$$

and

$$\frac{\partial V}{\partial \nu} + \phi\frac{\partial U}{\partial \nu} = 0 \tag{11.74}$$

then it implies that the hedging parameters are given by:

$$\phi = -\frac{\partial V}{\partial \nu} \Big/ \frac{\partial U}{\partial \nu} \tag{11.75}$$

and

$$\Delta = -\phi \frac{\partial U}{\partial S} - \frac{\partial V}{\partial S} \tag{11.76}$$

The change in the value of the portfolio is then left as:

$$d\Pi = \left\{ \frac{\partial V}{\partial t} + \frac{1}{2}\nu S^2 \frac{\partial^2 V}{\partial S^2} + \frac{1}{2}\nu\sigma^2 \frac{\partial^2 V}{\partial \nu^2} + \rho\sigma\nu S \frac{\partial^2 V}{\partial S \partial \nu} \right\} dt \tag{11.77}$$

$$+\phi \left\{ \frac{\partial U}{\partial t} + \frac{1}{2}\nu S^2 \frac{\partial^2 U}{\partial S^2} + \frac{1}{2}\nu\sigma^2 \frac{\partial^2 U}{\partial \nu^2} + \rho\sigma\nu S \frac{\partial^2 U}{\partial S \partial \nu} \right\} dt$$

which can be re-written as:

$$d\Pi = (A + \phi B)dt \tag{11.78}$$

Using the fact that the return on a risk-free portfolio must equal the risk-free rate r, which is assumed to be deterministic for purposes of this discussion, it follows $d\Pi = r\Pi dt$. Then,

$$(A + \phi B)dt = r\Pi dt \tag{11.79}$$

$$= r\left(V + \Delta S + \phi U\right) dt$$

Then it follows that,

$$A + \phi B = r\left(V + \Delta S + \phi U\right)$$

Substituting $\phi = -\frac{\partial V}{\partial \nu} / \frac{\partial U}{\partial \nu}$ and rearranging produces:

$$\frac{A - rV + rS\frac{\partial V}{\partial S}}{\frac{\partial V}{\partial \nu}} = \frac{B - rU + rS\frac{\partial U}{\partial S}}{\frac{\partial U}{\partial \nu}}. \tag{11.80}$$

The right-hand sided of the equation above is in terms of U alone and the left-hand side in terms of V only. Thus, a function $f(S, \nu, t) = -\kappa(\theta - \nu) + \lambda(S, \nu, t)$ of S, ν and t can be written for both sides of the equation as in Heston (1993). The term $\lambda(S, \nu, t)$ is the market price of volatility risk [35]. Now if, $f(S, \nu, t)$ is substituted for the left-hand side of the equation above, it follows that:

$$-\kappa(\theta - \nu) + \lambda(S, \nu, t) = \frac{B - rU + rS\frac{\partial U}{\partial S}}{\frac{\partial U}{\partial \nu}} \tag{11.81}$$

$$-\left[\kappa(\theta - \nu) - \lambda(S, \nu, t)\right]\frac{\partial U}{\partial \nu} = B - rU + rS\frac{\partial U}{\partial S}$$

$$B - rU + rS\frac{\partial U}{\partial S} + \left[\kappa(\theta - \nu) - \lambda(S, \nu, t)\right]\frac{\partial U}{\partial \nu} = 0$$

now substituting for B, it follows that:

$$\frac{\partial U}{\partial t} + \frac{1}{2}\nu S^2\frac{\partial^2 U}{\partial S^2} + \frac{1}{2}\nu\sigma^2\frac{\partial^2 U}{\partial \nu^2} + \rho\sigma\nu S\frac{\partial^2 U}{\partial S\partial \nu} \tag{11.82}$$

$$-rU + rS\frac{\partial U}{\partial S} + \left[\kappa(\theta - \nu) - \lambda(S, \nu, t)\right]\frac{\partial U}{\partial \nu} = 0$$

which is the Heston(1993) PDE.

Theorem 6. *(Feynman-Kac Theorem)*

Consider the terminal PDE:

$$
\begin{cases}
\frac{\partial U}{\partial t} + a(t,x)\frac{\partial U}{\partial x} + \frac{1}{2}b^2(t,x)\frac{\partial^2 U}{\partial x^2} - G(t,x)u(t,x) + f(t,x) = 0 \\
\\
u(t,x) = \Phi(x)
\end{cases}
$$

for all $x \in \mathbb{R}$ and $t \in [0,T]$ where a,b,Φ and G are known functions and $u : \mathbb{R} \to [0,T]$

is the unknown function. Then the solution of the PDE can be written as:

$$
u(x,t) = \mathbb{E}_{\mathbb{Q}} \left[\int_t^T e^{-\int_t^s G(X_u)du} f(X_s)\, ds + e^{-\int_t^T G(X_s)ds} \Phi(X_T) \,|\, X_t = x \right] \tag{11.83}
$$

under the probability measure $\mathbb{Q}$ such that X is an Itô process driven by:

$$
dX_t = a(X_t)\, dt + b(X_t)\, dW_t
$$

where W_t is a Wiener process and the initial condition is $X_t = x$.

Solution of the ODE in Equation (11.27)

Considering the ODE

$$
\frac{dD}{d\tau} = \frac{1}{2}\sigma_V^2 D^2 + (\rho\sigma_\nu\omega i - \kappa^*)D - \frac{1}{2}(\omega i + \omega^2) \tag{11.84}
$$

for

To simplify notation let

$$
\begin{cases}
\alpha := \kappa^* - \rho\sigma_\nu\omega i \\
\\
\beta := -\frac{1}{2}(\omega i + \omega^2) \\
\\
\gamma := \frac{1}{2}\sigma_V^2
\end{cases} \tag{11.85}
$$

then it follows that 11.84 can be rewritten as:

$$\frac{dD}{d\tau} = \gamma D(\tau)^2 - \alpha D(\tau) + \beta \tag{11.86}$$

where $D(\omega, 0) := D(0) = 0$ is the initial condition. This is the form of a *Riccati Equation.*

The general solution of the *Riccati Equation* can be obtained via substitution as follows:

Assume that $\beta, \gamma \neq 0$ then define

$$m = D\gamma$$

and assume that that $m, \gamma \in C^1$. The it follows, from the chain rule that

$$
\begin{aligned}
m' = (D\gamma)' &= D'\gamma + \gamma'D \\
&= \left(\gamma D^2 - \alpha D + \beta\right)\gamma + \gamma'D \\
&= \gamma^2 D^2 - \alpha D\gamma + \beta\gamma + \gamma'D \\
&= \gamma^2 \left(\frac{m^2}{\gamma^2}\right) - \alpha m + \beta\gamma + \gamma'D \\
&= m^2 - \alpha m + \beta\gamma + m\left(\frac{\gamma'}{\gamma}\right) \\
&= m^2 + m\left(\frac{\gamma'}{\gamma} + \alpha\right) + \beta\gamma
\end{aligned} \tag{11.87}
$$

This is the form of a linear second order differential equation. Then 11.87 can be re-written in the following form:

$$m' = m^2 + mP(\tau) + Q(\tau) \tag{11.88}$$

then if we use the substitution $m = \frac{-u'}{u}$ where $u \in C^2$, it follows that

$$m' = \frac{-u''}{u} + \left(\frac{u'}{u}\right)^2 \tag{11.89}$$

$\Longleftrightarrow$

$$m' = \frac{-u''}{u} + m^2$$
$$\frac{-u''}{u} = m' - m^2 = mP + Q \tag{11.90}$$

$\Longleftrightarrow$

$$\frac{-u''}{u} = \left(\frac{-u'}{u}\right)P + Q$$
$$u'' = u'P - Qu \tag{11.91}$$
$$u'' - u'P + Qu = 0$$

Equation 11.91 is a homogeneous second linear differential equation. Substituting back $P := \frac{\gamma'}{\gamma} + \alpha = -\alpha$ and $Q := \beta\gamma$ we obtain:

$$u'' + \alpha u' + \beta\gamma u = 0 \tag{11.92}$$

This homogeneous ODE has characteristic equation $r^2 + \alpha r + \beta\gamma = 0$. This implies that $r_1 = \frac{-\alpha + \sqrt{\alpha^2 - 4\beta\gamma}}{2}$ and $r_2 = \frac{-\alpha - \sqrt{\alpha^2 - 4\beta\gamma}}{2}$ thus $u_1 = e^{r_1\tau}$ and $u_2 = e^{r_2\tau}$ are solutions.

Checking the solution $u_1 = e^{r_1\tau}$, it follows that $u_1' = r_1 e^{r_1\tau}$ and $u_1'' = r_1^2 e^{r_1\tau}$. Then:

$$\begin{aligned}
0 = u'' + \alpha u' + \beta\gamma u &= r_1^2 e^{r_1\tau} + \alpha r_1 e^{r_1\tau} + \beta\gamma e^{r_1\tau} \\
&= \frac{1}{4}\left(\alpha^2 - 2\alpha\sqrt{\alpha^2 - 4\beta\gamma} + \alpha^2 - 4\beta\gamma\right)e^{r_1\tau} \\
&\quad + \alpha\left(\frac{-\alpha + \sqrt{\alpha^2 - 4\beta\gamma}}{2}\right)e^{r_1\tau} + \beta\gamma e^{r_1\tau}
\end{aligned} \tag{11.93}$$

This is a solution if for $c_1, c_2 \in \mathbb{C}$, it follows that:

$$u = c_1 u_1 + c_2 u_2 = c_1 e^{r_1 \tau} + c_2 e^{r_2 \tau} \qquad (11.94)$$

but we recall that $m = \frac{-u'}{u}$ then it follows that $D = -\frac{-u'}{u\gamma}$. Therefore,

$$D(\tau) = \frac{-\left(c_1 r_1 e^{r_1 \tau} + c_2 r_2 e^{r_2 \tau}\right)}{\gamma \left(c_1 e^{r_1 \tau} + c_2 e^{r_2 \tau}\right)} \qquad (11.95)$$

and since $D(0) = 0$ it implies that

$$\frac{-\left(c_1 r_1 + c_2 r_2\right)}{\gamma \left(c_1 + c_2\right)} = 0 \qquad (11.96)$$

$\Rightarrow$

$$c_1 = -c_2 \left(\frac{r_2}{r_1}\right) \qquad (11.97)$$

Thus

$$
\begin{aligned}
D(\tau) &= \frac{-\left(-c_2 \left(\frac{r_2}{r_1}\right) r_1 e^{r_1 \tau} + c_2 r_2 e^{r_2 \tau}\right)}{\gamma \left(-c_2 \left(\frac{r_2}{r_1}\right) e^{r_1 \tau} + c_2 e^{r_2 \tau}\right)} \\
&= \frac{e^{r_1 \tau} + e^{r_2 \tau}}{\left(\frac{\gamma}{r_1}\right) \left(\frac{r_2}{r_1} e^{r_1 \tau} - e^{r_2 \tau}\right)} \\
&= \frac{1 - e^{r_2 \tau - r_1 \tau}}{\frac{-\gamma}{r_2} \left(1 - \frac{r_2}{r_1} e^{r_2 \tau - r_1 \tau}\right)} \\
&= \frac{1 - e^{-d\tau}}{\frac{-\gamma}{r_2} \left(1 - g e^{-d\tau}\right)}
\end{aligned}
\qquad (11.98)
$$

where $d = r_2 - r_1 = \sqrt{\alpha^2 - 4\beta\gamma}$ and $g = \frac{r_2}{r_1} = \frac{-\alpha - \sqrt{\alpha^2 - 4\beta\gamma}}{-\alpha + \sqrt{\alpha^2 - 4\beta\gamma}}$

Part III

Hedging variance swaps

12 Static replication of variance swaps

Demeterfi, Derman, Kamal and Zou (1999) mention that the volatility of an asset is the simplest measure of risk and uncertainty of the asset [25]. As mentioned earlier, the need to hedge oneself from volatility risk saw the increase in the trading of volatility and variance-based derivatives over the past two decades. The payoff of a variance swap is a function of realised variance (the variance of the underlying asset's return over the lifetime of the variance swap). In this section, a methodology discussed in Demeterfi et al. (1999) and Carr and Lee (2009) to replicate the payoff of variance swaps is discussed.

In a non-parametric setting, the underlying asset and vanilla options are traded in quantities which are expressed in terms of options (vanilla options) without specifying parameters of a particular distribution [19, 25]. This is in contrast to the price of variance swaps developed under stochastic volatility models as indicated earlier. Neuberger (1994) showed that the hedging error obtained from delta-hedging a log contract accumulates the difference between realised variance and the fixed variance implemented in the hedge [43].

Carr and Lee (2009) postulate that under an independence condition, the risk-neutral distribution of path-dependent realised variance can be used to determine the value of a stock option [19]. This relationship is inverted in a generalised setting to given market prices of options at a given expiry and invert a Hull-White-type [36]

relationship to infer the entire risk-neutral distribution of the stochastic realised volatility [19].

Thus, a profile of option prices with expiry T, will therefore, in a non-parametric way, infer the non-arbitrage prices of claims whose payoffs are a function of realised variance (variance derivatives). Furthermore, the option prices will allow the replication of variance derivatives by dynamic trading in standard options and the underlying stock [19, 22]. In this section, the call prices on S for maturity T are assumed to be known for all strike prices. The realised variance can be replicated by hedging a log contract whose payoff at time T is $-2log\left(\frac{S_T}{S_0}\right)$ [24, 28, 43]. As mentioned in Demeterfi et al. (1999) [25], the risk-neutral price of a variance swap in this context is thus based on:

1. The capability of replicating a log contract through a portfolio of options which has a continuous range of strike prices,

2. Adherence to the fundamental options valuation theory, under which the stock price is assumed to evolve under continuous dynamics.

In reality, however, the range of strikes are limited to a finite range [25].

Neuberger [43] showed that by delta-hedging a contract paying the log of the price, the hedging error accumulates to the difference between the realized variance and the fixed variance used in the delta-hedge. The contract paying the log of the price can be created with a static position in options.

12.1 Assumptions

Fixing an arbitrary time horizon $T > 0$, it is assumed that there exists a risk-neutral world with a constant risk-free rate r. Furthermore, assume that markets

are frictionless. Then on a filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}(t)\}, \mathbb{P})$ and under an equivalent probability measure, $\mathbb{Q}$, the underlying stock's dynamics are as follows:

$$dS(t) = rdt + \sigma(t, ..)S(t)dW(t) \tag{12.1}$$

for some $(\mathcal{F}(t), \mathbb{Q})$ -Wiener process $W(t)$ and a measurable process $\mathcal{F}(t) - adapted$ process $\sigma(t, ..)$ which satisfies

$$\int_0^T \sigma^2(t, ..) \, dt < m \in \mathbb{R} \tag{12.2}$$

with σ and W being assumed independent.

12.2 Theoretical replication of the fair strike of a variance swap using an options portfolio

As in Demeterfi et. al. (1999), the only assumption made about the dynamics of the underlying asset in this section is that the price of the underlying is continuous [25].

This implies that the underlying stock's dynamics can be represented by:

$$dS(t) = \mu(t)S(t)dt + \sigma(t)S(t)dW(t) \tag{12.3}$$

where W_t represents the Wiener process, with the drift, μ, and the continuously-sampled volatility, σ, assumed to be arbitrary functions of time and other parameters. It is assumed that the underlying stock does not pay dividends.

The definition of realised variance presented earlier can now be represented theoretically as :

$$\sigma_R^2(T) = \frac{1}{T} \int\limits_0^T \sigma^2(t)\,dt \tag{12.4}$$

As shown earlier, since variance swaps are forward contracts on realised variance, the fair future fixed level of variance, K_{var}, is given by:

$$K_{var} = E\left[\sigma_R^2(T)\right] \tag{12.5}$$

Which can be rewritten as:

$$K_{var} = \frac{1}{T} E\left[\int\limits_0^T \sigma^2(t)\,dt\right] \tag{12.6}$$

The expression for K_{var} shown above does not give insight into the replication scheme using options since the future value for variance is unknown. The general idea behind a replicating scheme is to derive a position which over the next instance of time generates a payoff which is proportional to the incremental variance of the stock [25].

From Equation (12.3), the approach is shifted by implementing Itô's Lemma to $ln\left(S(t)\right)$ as follows:

$$dln\left(S(t)\right) = \left(\mu(t) - \frac{1}{2}\sigma^2(t)\right)dt + \sigma(t)dW_t \tag{12.7}$$

Then subtracting Equation (12.7) from Equation (12.3) it follows that,

$$\frac{dS(t)}{S(t)} - d\ln\left(S(t)\right) = \frac{1}{2}\sigma^2(t) \tag{12.8}$$

Then integrating both sides from 0 to T,

$$\int_0^T \frac{dS(t)}{S(t)}dt - \ln\left(\frac{S(T)}{S(0)}\right) = \int_0^T \frac{1}{2}\sigma^2(t)dt. \tag{12.9}$$

We then obtain the continuously-sampled realised variance as ,

$$\frac{2}{T}\left(\int_0^T \frac{dS(t)}{S(t)}dt - \ln\left(\frac{S(T)}{S(0)}\right)\right) = \frac{1}{T}\int_0^T \sigma^2(t)dt = \sigma_R^2(T). \tag{12.10}$$

From Equation (12.10), it thus can be obtained that realised variance can be repli-
cated by a static short position in a contract which at expiration pays the logarithm
of the total return and a dynamic long position of $\frac{1}{S(T)}$ shares of stock until expiry
[14, 25]. The first term in on the left-hand side of Equation (12.10) can be thought
of as the result from continuously re-balancing a stock position so that it always
pays \$1 and the second term represents a short position in a contract which pays
the log of total return at expiry [25].

Broadie and Jain (2008) specify that the result in Equation (12.9) holds for both
stochastic models such as the Heston model and the the classic Black-Scholes model
[14].

Equation (12.10) thus provides an alternative method for deducing the fair strike
price of a variance swap contract. Now, rather than taking the average of future
expected realised variance as in Equation (12.6), one can can take the expected value
of the L.H.S of Equation (12.10) under the risk-neutral measure to obtain the cost
of replication directly as [25]:

$$K_{var} = \frac{2}{T} E \left(\int_0^T \frac{dS(t)}{S(t)} dt - ln \left(\frac{S(T)}{S(0)} \right) \right) \qquad (12.11)$$

In a risk-neutral world with a constant risk-free rate r, the dynamics of the underlying price can be represented by:

$$\frac{dS(t)}{S(t)} = rdt + \sigma(t)dW_t \qquad (12.12)$$

Now if one integrates Equation (12.12) and takes the risk-neutral expected value on both sides one obtains:

$$E \left(\int_0^T \frac{dS(t)}{S(t)} dt \right) = rT \qquad (12.13)$$

This then implies that the fair value of the variance swap strike is:

$$K_{var} = \frac{2}{T} \left(rT - E \left[ln \left(\frac{S(T)}{S(0)} \right) \right] \right) \qquad (12.14)$$

Now since log-contracts are not actively traded the second term on the R.H.S of Equation (12.14) can be duplicated by decomposing the shape of the log payoff into linear and curved components [25]. To this regard the following should be noted:

1. The linear component can be replicated with a forward contract on the underlying stock with delivery time T

2. The curved component can be replicated using vanilla options with all strike levels and same expiry T.

Demeterfi et al. (1999) [25], suggest that for practical reasons there is need to replicate the log payoff with liquid options. The liquid options in this regard are a mixture of out-of-the-money calls for high stock values and out-of-the-money puts for low stock values. A new parameter, $S*$, following the notation in [25] is introduced to separate puts and calls. With this new parameter, the log payoff can now be rewritten as:

$$ln\left(\frac{S(T)}{S(0)}\right) = ln\left(\frac{S(T)}{S*}\right) + ln\left(\frac{S*}{S(0)}\right) \qquad (12.15)$$

The second term on the R.H.S in Equation (12.15) is a constant thus there is need to only replicate the first term on the R.H.S. To replicate this term a proposition is implemented as in Demeterfi et al. (1999) and proof from Carr and Madan (1998) [25, 22] :

Proposition 10. *(Carr and Madan (1998), [22])*

Any twice-differentiable function $f(S(T))$ can be re-written as:

$$
\begin{aligned}
f\left(S(T)\right) &= f\left(K\right) + f'\left(K\right)\left[(S(T) - K)^+ - (K - S(T))^+\right] + \int_0^K f''(K)\left(K - S(T)\right)^+ dK \\
&\quad + \int_K^\infty f''(K)\left(S(T) - K\right)^+ dK \qquad (12.16)
\end{aligned}
$$

Proof. For any payoff $f(S(T))$ the Dirac delta sifting property implies that :

$$
\begin{aligned}
f\left(S(T)\right) &= \int_0^\infty f(K)\delta\left(S(T)-K\right)dK, \\
&= \int_0^\kappa f(K)\delta\left(S(T)-K\right)dK + \int_\kappa^\infty f(K)\delta\left(S(T)-K\right)dK,
\end{aligned}
$$

for any $\kappa \in \mathbb{R}^+$. Then integrating each term by parts, it follows that,

$$
\begin{aligned}
f\left(S(T)\right) &= \left[f\left(K\right)1_{\{S(T)<K\}}\right]_0^\kappa - \int_0^\kappa f'(K)1_{\{S(T)<K\}}dK \\
&\quad + \left[f\left(K\right)1_{\{S(T)\geq K\}}\right]_0^\infty + \int_\kappa^\infty f'(K)1_{\{S(T)\geq K\}}dK
\end{aligned}
$$

Integrating by parts for the second time, it follows that:

$$
\begin{aligned}
f\left(S(T)\right) &= f(\kappa)1_{\{S(T)<K\}} - \left[f'(K)\left(K-S(T)\right)^+\right]_0^\kappa + \int_0^\kappa f''(K)\left(K-S(T)\right)^+ dK \\
&\quad + f(\kappa)1_{\{S(T)\geq K\}} - \left[f'(K)\left(S(T)-K\right)^+\right]_\kappa^\infty + \int_\kappa^\infty f''(K)\left(S(T)-K\right)^+ dK \\
&= f(\kappa) + f'(\kappa)\left[\left(S(T)-\kappa\right)^+ - \left(\kappa-S(T)\right)^+\right] \\
&\quad + \int_0^\kappa f''(K)\left(K-S(T)\right)^+ dK + \int_\kappa^\infty f''(K)\left(S(T)-K\right)^+ dK
\end{aligned}
$$

This concludes the proof. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\quad$ $\square$

Proposition 11. *The payoff of the log-contract, $\ln\left(\frac{S(T)}{S*}\right)$, can be decomposed as :*

$$ln\left(\frac{S(T)}{S*}\right) = \frac{S(T) - S*}{S*} - \int_0^{S*} \frac{1}{K^2}\left(K - S(T)\right)^+ dK - \int_{S*}^{\infty} \frac{1}{K^2}\left(S(T) - K\right)^+ dK$$

$$(12.17)$$

for $S*, K \in \mathbb{R}^+$

Proof. Since the logarithm function is twice differentiable, if we let $\kappa = S*$ and $f(X) = ln(X)$, the result follows directly from the expression in the statement of Proposition 10. $\square$

Equation (12.17) means that the payoff of a short log contract can be represented by:

- long $\frac{1}{S*}$ forward contract struck at $S*$

- short in $\frac{1}{K^2}$ puts struck at K, for all strike prices less than $S*$

- and short in $\frac{1}{K^2}$ calls struck at K, for all strike prices greater than $S*$

- $(K - S(T))^+$ is the payoff of a European Put Option and $(S(T) - K)^+$ the payoff of a European Call Option

all expiring at time T.

In the absence of arbitrage, the decomposition in Proposition 10 must exist [22]. Thus using the initial values of the Proposition 11 above implies that the fair strike in Equation (12.14) can be re-written as:

$$
\begin{aligned}
K_{var} &= \frac{2}{T}\left(rT - E\left[ln\left(\frac{S(T)}{S(0)}\right)\right]\right) \\
&= \frac{2}{T}\left(rT - E\left[ln\left(\frac{S(T)}{S*}\right) + ln\left(\frac{S*}{S(0)}\right)\right]\right) \\
&= \frac{2}{T}\left(rT - E\left[ln\left(\frac{S(T)}{S*}\right)\right] - ln\left(\frac{S*}{S(0)}\right)\right) \\
&= \frac{2}{T}\left(rT - E\left[\frac{S(T)-S*}{S*} - \int_0^{S*}\frac{1}{K^2}(K-S_T)^+ dK - \int_{S*}^{\infty}\frac{1}{K^2}(S(T)-K)^+ dK\right]\right) \\
&\quad -\frac{2}{T}ln\left(\frac{S*}{S(0)}\right) & (12.18) \\
&= \frac{2}{T}\left(rT - \left(\frac{S(0)}{S*}e^{rT} - 1\right) + e^{rT}\int_0^{S*}\frac{1}{K^2}(K-S(0))^+ dK\right) & (12.19) \\
&\quad +\frac{2}{T}\left(+e^{rT}\int_{S*}^{\infty}\frac{1}{K^2}(S(0)-K)^+ dK + ln\left(\frac{S*}{S(0)}\right)\right)
\end{aligned}
$$

If one denotes $P_0(K) = (K - S(0))^+$ and $C_0(K) = (S(0) - K)^+$ as initial values of a European put and call with strike K, respectively, then it follows that,

$$
\begin{aligned}
K_{var} &= \frac{2}{T}\left(rT - \left(\frac{S(0)}{S*}e^{rT} - 1\right) + e^{rT}\int_0^{S*}\frac{1}{K^2}P(K)\,dK\right) & (12.20) \\
&\quad +\frac{2}{T}\left(+e^{rT}\int_{S*}^{\infty}\frac{1}{K^2}C(K)\,dK - ln\left(\frac{S*}{S(0)}\right)\right)
\end{aligned}
$$

The integrals in the equation above sum an infinite number of vanilla options and calls in a continuum of strike prices. The same underlying asset as the one for the variance swap which is being replicated is used to write the options. The above

expression can be be simplifies by setting $S* = S(0)$. Then one obtains:

$$K_{var} = \frac{2}{T} \left(rT - (e^{rT} - 1) + e^{rT} \int_0^{S(0)} \frac{1}{K^2} P(K) \, dK + e^{rT} \int_{S(0)}^{\infty} \frac{1}{K^2} C(K) \, dK \right)$$

(12.21)

12.3   Replication in a practical setting

In the previous section it was shown that the log price can be re-written as a sum $ln\left(\frac{S(T)}{S*}\right) + ln\left(\frac{S*}{S(0)}\right)$ with $S*$ the boundary between European calls and puts. Thus the recalling the expression for the variance swap strike, K_{var} it can be written as:

$$
\begin{aligned}
K_{var} &= \frac{2}{T}\left(rT - E\left[ln\left(\frac{S(T)}{S(0)}\right)\right]\right) \\
&= \frac{2}{T}\left(rT - E\left[ln\left(\frac{S(T)}{S*}\right) + ln\left(\frac{S*}{S(0)}\right)\right]\right) \\
&= \frac{2}{T}\left(rT - E\left[ln\left(\frac{S(T)}{S*}\right) - \frac{S(T) - S*}{S*} + \frac{S(T) - S*}{S*} + ln\left(\frac{S*}{S(0)}\right)\right]\right)
\end{aligned}
$$

(12.22)

after taking the expectation then,

$$K_{var} = \frac{2}{T}\left[rT - \left(\frac{S(0)}{S*}e^{rT} - 1\right) - ln\left(\frac{S*}{S(0)}\right)\right] + \frac{2}{T}E\left[\frac{S(T) - S*}{S*} - ln\left(\frac{S(T)}{S*}\right)\right].$$

(12.23)

The problem is now centered on solving the second term. The portfolio has the payoff at expiry denoted by

$$f(S(T)) = \frac{2}{T}\left[\frac{S(T) - S*}{S*} - ln\left(\frac{S(T)}{S*}\right)\right]$$

(12.24)

If V is the present value of the portfolio, then strike of the variance swap can be re-written as:

$$K_{var} = \frac{2}{T}\left[rT - \left(\frac{S(0)}{S*}e^{rT} - 1\right) - ln\left(\frac{S*}{S(0)}\right)\right] + e^{rT}V. \tag{12.25}$$

The objective, as shown before is to estimate $f(S(T))$. If a portfolio of European call options with strikes $S(0) = S* < K_0^c < K_1^c < K_2^c < ...$ and European put options with strikes $S(0) = S* > K_0^p > K_1^p > K_2^p > ...$. To replicate $f(S(T))$, a piece-wise function of these individually weighted options can be used as:

$$V = \sum_{i=0} w_i^c (K_i^c) C (K_i^c) + \sum_{i=0} w_i^p (K_i^p) P (K_i^p) \tag{12.26}$$

where $w_i^c (K_i^c)$ and $w_i^p (K_i^p)$ are weights for calls and puts respectively. The weights are derived for the curve approximating $f(S_T)$, as follows:

Figure 12.1: Approximation of $f(S(T))$
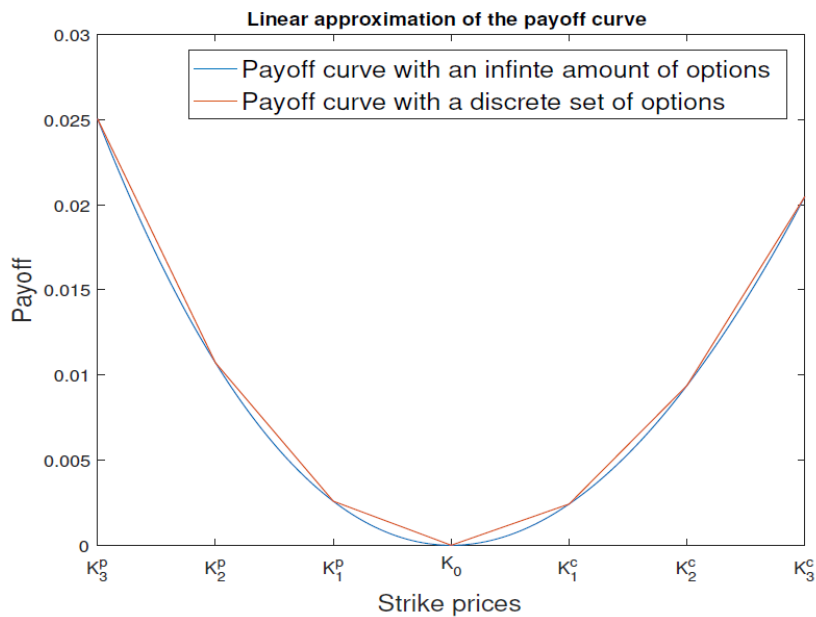
Source: Demeterfi et. al. (1999) [25]

From the Figure above K_0^c to K_1^c can be approximated by a payoff of $w\left(K_0^c\right)$ European call options with strike K_0^c. Then it follows that:

$$w\left(K_0^c\right) = \frac{f(K_1^c) - f(K_0^c)}{K_1^c - K_0^c} \tag{12.27}$$

Similarly for $w\left(K_1^c\right)$, the part from K_1^c to K_2^c is derived as a combination of European calls options with strikes K_0^c and K_1^c. Assuming we already hold $w\left(K_0^c\right)$ European call options, then:

$$w\left(K_1^c\right)\left(S - K_1^c\right) + w\left(K_0^c\right)\left(S - K_0^c\right) = f(S) \tag{12.28}$$

for $S = K_0^c, K_1^c$. Then solving the previous equation, by substitution results in:

$$w\left(K_1^c\right) = \frac{f(K_1^c) - f(K_0^c)}{K_1^c - K_0^c} - w\left(K_0^c\right)$$

which in general can be written for European call options as:

$$w\left(K_{n+1}^c\right) = \frac{f(K_{n+1}^c) - f(K_n^c)}{K_{n+1}^c - K_n^c} - \sum_{i=0}^{n-1} w\left(K_i^c\right) \tag{12.29}$$

and for European put options as:

$$w\left(K_{n+1}^p\right) = -\frac{f(K_{n+1}^p) - f(K_n^p)}{K_{n+1}^p - K_n^p} - \sum_{i=0}^{n-1} w\left(K_i^p\right) \tag{12.30}$$

12.4   Numerical Example

In this section, linear programming techniques are implemented in the hedging of the log contract which entails the hedging of variance swaps using vanilla European

options. The problems that arise in this implementation are discussed in numerical examples. Particularly, it is difficult to select the range of strike prices to be used and also the number of vanilla options to be included in the replicating options portfolio. Consider the following parameters:

Table 6: Hypothetical Parameters

S(T)	volatility	r	T	dividend rate
292.373	0.0108	0.00038344	1	0%

Now considering the parameters above. The European call options and put options are calculated for twenty strike prices ranging from 10 points below the stock price to 10 points above the stock price. The formula for calculating weights results in heavier weights for out-of-the-money (OTM) options. The resulting variance swap prices for standard maturities is:

Table 7: Price of variance swaps

Maturity	K_{var}
T=0.0833	0.0019
T=0.25	0.0026
T=0.5	0.0035
T=1	0.052
T=2	0.081

The payoff approximation as in Figure 12.3 is replicated in MATLAB as:

Figure 12.2: Payoff using the Demeterfi approximation
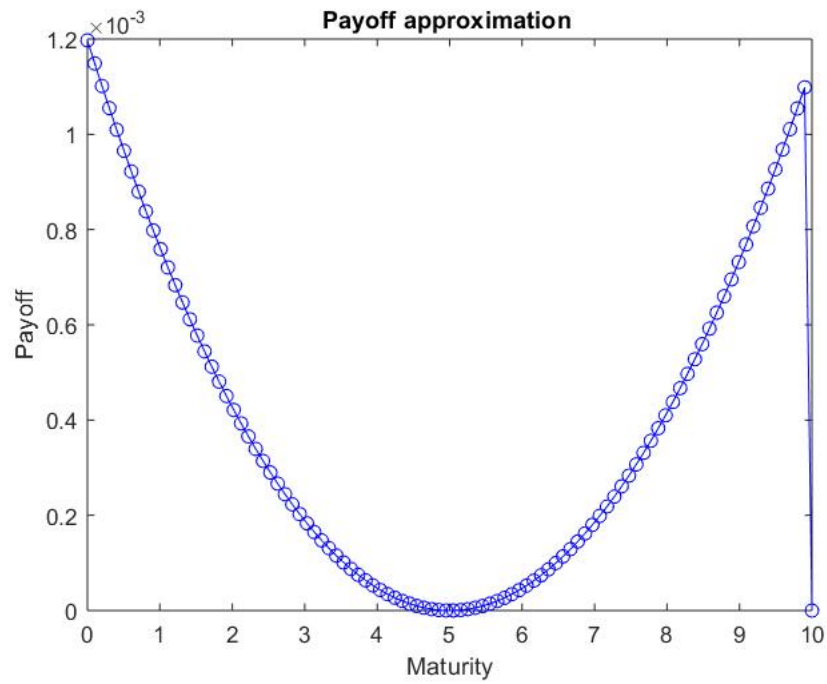
The price of variance swaps at various maturities is also investigated.

Figure 12.3: Variance Swap Price Term Dependence
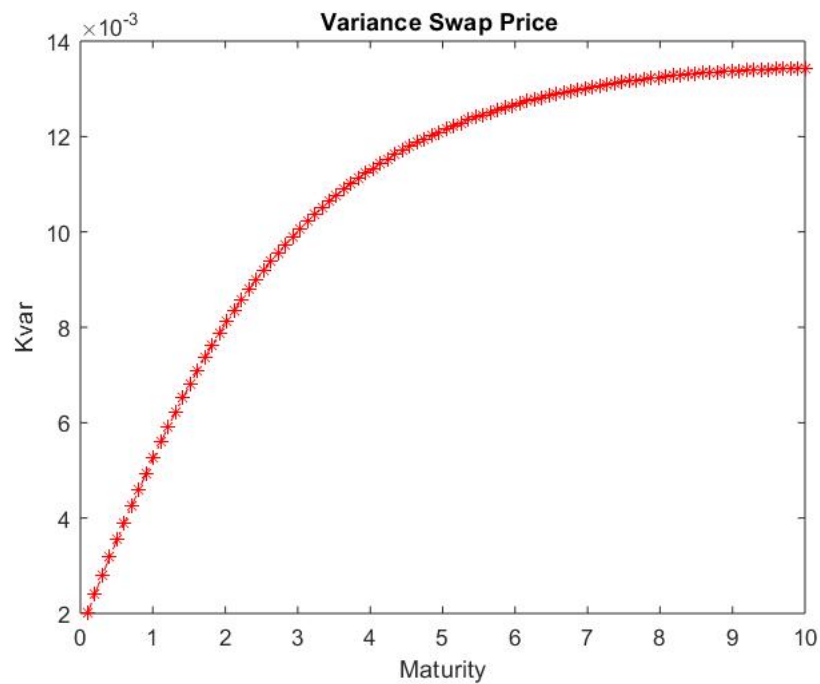
Comments: The Figure 12.4 above shows that the price of variance swaps relies on the maturity. Although the price monotonically increases with maturity, the rate of change of the price with maturity is higher for shorter maturities than for longer maturities.

12.5   Empirical evidence

The implementation in the previous example is re-performed using market data of 10 European call options and 45 European put options on the JSE Top 40 index as at 3 May 2019. The expiry date of the contracts chosen is on 20 June 2019 ($T = 0.1315068$). The price of variance swaps as in Equation (12.25) is derived using the Demeterfi et. al (1999) [25] non-linear payoff approximation. The table below shows how V in Equation (12.25) is derived from the market options data.

Table 8: Static Replication of the Price of Variance Swaps

Contracts Expiry:			20 June 2020	
Spot Price (ZAR)			533.43	
Puts	Strike (ZAR)	Price (ZAR)	Weight	Contribution
	400.00	0.10	9.112038×10^{-3}	0.000911204
	408.50	0.17	8.651058×10^{-3}	0.00147068
	410.00	0.19	8.137902×10^{-3}	0.001546201
	420.00	0.33	7.275507×10^{-3}	0.002400917
	430.00	0.58	6.653711×10^{-3}	0.003859152
	...	...	...	...
	523	7.93	1.36668×10^{-5}	
Calls	535	89.3	7.92073×10^{-5}	0.0014184807
	537	80.36	7.87176×10^{-5}	0.007444697
	538	44.22	3.09882×10^{-4}	0.007029486
	540	21.23	5.02838×10^{-4}	0.024902153
	550	4.55	6.26138×10^{-4}	0.022235494
	...	...	...	...
	628.50	0.01	2.41311×10^{-2}	0.026544310
Total Portfolio Cost (V)				1.22199
Variance Swap Price (K_{var})				1.22205

Comments: In the Table above, the weights are calculated using Equation (12.30) and Equation (12.29). More weight is given to more OTM options when calculating the total portfolio cost, V.

The MATLAB code for the generating the weights above is as follows:

MATLAB Code:

```matlab
    S0=Spot(1);%S0=S* the boundary between puts and calls;

    r=0.00038344;

    T= Term(1);

    K=Strike; % Conversion to familiar notation

    m=sum(Contract_Type=="Put"); %position of last Put value

    n=length(K);

    f=@(x)((2/T)*((x-S0)/S0 -log(x/S0))); %Demeterfi
Approximation

    Df =zeros(n,1); weight=zeros(n,1); Contrib=zeros(n,1);

    %Puts

    for i=1:m

    Df(i)= ((f(K(i))- f(K(i+1)))/ (K(i+1)-K(i)));

    if i==m weight(i)=Df(i);

    elseif i< m weight(i)=Df(i)-Df(i+1);

    end

    Contrib(i)=Price(i)*weight(i);

    end

    %Calls

    for i=m+1:n-1

    Df(i)= ((f(K(i+1))- f(K(i)))/ (K(i+1)-K(i)));

    if i==m+1 weight(i)=Df(i); elseif i>m
weight(i)=Df(i)-Df(i-1);

    end

    Contrib(i)=Price(i)*weight(i);

    end

    V= sum(Contrib);

    Kvar = (2/T)*(r*T-(exp(r*T)-1))+(exp(r*T)*V); %Price of
Variance Swap
```

12.6   Comparison of the Pricing Methods

So far, three methods for pricing variance swaps have been considered. The stochastic volatility approach have been useful in closely estimating the real-world dynamics of the stocks and their variance processes. The non-parametric replication approach has also been successful in pricing the variance swaps. Each of the models has had flaws in the pricing process. In particular, the stochastic models have been not straight forward to implement and moreover required the calibration of models to market data which weighs more on computational time. However, although the replication scheme does not have this flaw, it is difficult to determine the number of options to be considered in the hedging of variance swaps. Furthermore, in the real world there are challenges with obtaining market prices for a continuum of strikes. The price of variance swaps under each model is compared below:

Table 9: Comparison of variance swap pricing schemes

Models	Maturity (Years)			
	0.083	0.25	0.5	1
Heston	1.2032×10^{-4}	1.2778×10^{-4}	1.3894×10^{-4}	1.6119×10^{-4}
B-NS	1.0543×10^{-4}	1.0553×10^{-4}	1.0569×10^{-4}	1.0605×10^{-4}
Demeterfi Replication	3.7101×10^{-5}	1.7863×10^{-4}	4.0796×10^{-4}	8.7716×10^{-4}

The pricing of variance swaps under the three schemes produces inconsistent results. However, there seems to be cohesion between the stochastic models. The initial variance is low which means that the Replication scheme is overstating the prices for longer maturities.

12.7   Conclusion

The Demeterfi et al. (1999) methodology provides an alternative approximation method of obtaining the price of variance swaps without specifying the dynamics of

the stock and variance processes. The value of a variance swap was obtained from replicating a static short position in a contract that pays the logarithm of the total return at expiry and a dynamic long position of $\frac{1}{S(T)}$ shares of stock until expiry in Equation (12.11). In a risk-neutral setting, this value of the variance swap is a function of the expectation of the log-contract. Thus the problem of obtaining the value of the variance swap can be translated into a problem of estimating the log contract's payoff.

Since log-contracts are not actively traded, the payoff is replicated by a continuum of out-of-the-money (OTM) vanilla call and put options of all strike as in Proposition 11 from an approximation obtained in Carr and Madan (1998) [23]. In practice, however, the options have a finite number of strikes. This is the key aspect of the Demeterfi et al. (1999) methodology, which computes the price of variance swaps via the estimate of the value of the portfolio of OTM European call and put options with a finite number of strike prices using a linear approximation of the payoff of the log-contract. The numerical example shows that the price of the variance swaps is dependent on the term, differing significantly for shorter maturities and differing less for longer maturities. However, in this case, the price of the variance swaps peaks at a maturity of 3 years and slowly declines for maturities higher than 3 years. The shape of the curve is consistent with the one obtained in the stochastic volatility models shown in earlier sections. The price of variance swaps is hedged using market prices of European call and put options. The replication approximation gives more weight to options which are further out-of-the-money. Furthermore, a comparison of the pricing methods shows inconsistencies between the different schemes.

Part IV

Conclusions

13   Summary

In this research, the closed and semi-closed form expressions for the price of variance swaps whose underlying processes were assumed to have the B-NS (2001) non-Gaussian OU and Heston (1993) stochastic models were derived. Furthermore, using the Demeterfi et. al. (1999) closed-form expression it was shown that variance swaps can be hedged by a portfolio of out-of-the-money vanilla call and put options with a finite number of strike prices.

In Chapter 2, the mathematical, finance and statistical theory used in this dissertation was discussed. Although it was shown that the Black-Scholes model is not entirely perfect in determining the dynamics of the underlying asset, it is a more tractable model whose constructs can be easily understood and have a wide arsenal of computational tools available. Deriving the price of the variance swaps under the more accurate Heston (1993) and B-NS (2001) models involved enhanced techniques that were mostly made difficult by the complexity of the expressions.

In Chapter 3, the model developed by Barndoff-Nielsen and Shepard (2001) provided attractive features to the volatility process such as positive jumps which are observable in market data. The analytical formula for continuously-sampled variance swaps was obtained using Laplace transform and Fourier transform theory under an integrability condition for Levy processes. Although elementary results such as the Laplace transform and Fourier transform were implemented to obtain the analytical

formula, it was practically challenging to obtain the price of variance swap using the formula compared to the traditional results as those under the Black-Scholes model. The initial estimates of the NIG were obtained from the method of moments estimation (MME) but however more accurate estimates were obtained from maximum likelihood estimation (MLE). Theoretical assumptions about the individual OU processes had to be made to derive the price of the variance swap in a practical setting. The solutions obtained thus relied on the accuracy of the approximations. Lastly, it was shown that the price of continuously-sampled variance swaps is convex in realised variance.

In Chapter 4, the Heston (1993) model provided a more realistic description of the paths of the stock price and volatility processes. However, the analytical formula for the discretely-sampled variance swaps was not obtained as easy as in the continuously-sampled case. The analytical formula for the discretely-sampled variance swaps was obtained from solving two-part PDEs using Fourier transform techniques under Feynman-Kac theory. Parameter calibration was then conducted in a way that minimised the mean squared error between the market European call prices and the European call prices under the Heston Model. Long computational time was required to produce the parameter estimates in MATLAB making the Heston model particularly unattractive in this regard. Furthermore, the continuously-sampled case overstated the price of variance swaps when compared to the discretely-sampled case where fewer sampling times were considered. However, the price of variance swaps under the discretely sampled case model was shown to converge to the price under the continuous model as the sampling times were increased. The convergence was however slow with the number of sampling times required reaching more than 1512 times per annum (about 6 times per day). The price of continuously-sampled variance swaps was shown to differ significantly for shorter maturities and became more consistent for longer-maturities.

In Chapter 5, it was shown that variance swaps can be theoretically hedged using a continuum of vanillas options of all strike prices. However, in practice, there are a finite number of vanilla options. Thus an expression for the fair price of the variance swap for a finite number of European call and put options for a finite number of strikes was derived. Demeterfi et al. (1999) provide a methodology which computes the price of variance swaps via the estimate of the value of the portfolio of OTM European call and put options with a finite number of strike prices using a linear approximation of the payoff of the log-contract. A numerical example showed that the price of the variance swaps is dependent on the term, increasing significantly for shorter maturities and slowing down for longer maturities. The shape of the curve was consistent with the ones obtained under stochastic assumptions for the variance process models as in the Heston (1993) and B-NS (2001). Furthermore, a comparison of the three pricing approaches studied did not produce consistent results.

14   Future research considerations

Since the stochastic volatility models considered are path dependent, it would be of interest to study the models under a delayed time-process. There would be need to prove whether the pricing principles remain the same under a delayed-time setting. Furthermore, jumps can also be considered in an adjusted Heston model.

Appendix

A   Theoretical Concepts

In this Appendix, some theoretical comments which were applied in this dissertation are shown. Furthermore, some MATLAB function implemented are presented.

Definition 24. (Borel sets, Sato (1999) [46])

Let Ω be a non-empty set. The $\sigma - field$ generated on Ω by subsets A of the non-empty set Ω,denoted by $\sigma(A)$ is defined as:

$\sigma(A) := \bigcap \{\mathcal{F} : A \subseteq \mathcal{F}\}$ and $\mathcal{F}$ is a σ-field on Ω.

This implies that the $\sigma - field$ generated by A is the smallest $\sigma - field$ which contains A. This leads us to the definition of *Borel sets*.

The *Borel* $\sigma - algebra$ of $\mathbb{R}$, denoted $\mathcal{B}$, is the $\sigma - algebra$ generated by the family of all open sets on the real line. That is, if $\mathscr{O}$ denotes the collection of all open subsets of $\mathbb{R}$, then $\mathcal{B} = \sigma(\mathscr{O})$. A real-valued function $f(x)$ on $\mathbb{R}$ is called measurable, if it is $\mathcal{B}(\mathbb{R})$-measurable that is $\{x : f(x) \in B\}$ is in $\mathscr{O}$ for each $B \in \mathcal{B}(\mathbb{R})$ [46].

Now a random variable can be defined.

Definition 25. (Random Variable, Applebaum (2009) [1])

Given a probability space $(\Omega, \mathcal{F}, P)$, a random variable, X, is the mapping $X : \Omega \to \mathbb{R}$.

The measurable mapping $Z = X + Yi$ from Ω to $\mathbb{C}$ is called a complex random variable. If X is a random variable, its law (or distribution) is the Borel probability

measure $\mathbb{P}_X$ on $\mathbb{R}$. One can write, $\mathbb{P}[\omega \in \Omega : X(\omega) \in B] = \mathbb{P}[X \in B]$, the mapping of B which is a probability measure on $\mathbb{R}$. This probability measure can be denoted $\mathbb{P}_X(B)$ and is called the distribution (or law) of X [46]. Now the fundamental concepts of a the *Expectation* and *Variance* of a random variable can be introduced as they are particularly important in probability theory.

Definition 26. (Expectation, *Bain and Engelhardt (1992)* [2])

If X is a real-valued ($\mathbb{R}-valued$) random variable and if the integral $\int_\Omega X(\omega)\,\mathbb{P}(d\omega)$ exists then it is called the *Expectation* of X, denoted by $\mathbb{E}[X]$. If X is a random variable on $\mathbb{R}$ and if $f(X)$ is a bounded and measurable function on $\mathbb{R}$, then

$$\mathbb{E}_\mathbb{P}[f(X)] = \int_\Omega f(x)\,\mathbb{P}_X(dx). \tag{A.1}$$

The variance of $f(X)$ is given by,

$$Var[f(X)] = \int_\Omega (f(x) - \mathbb{E}_\mathbb{P}[f(x)])^2\,\mathbb{P}_X(dx) \tag{A.2}$$

A random variable X is said to have a Property A *almost surely (a. s.)* or *with probability 1*, if there is a $\Omega_0 \in \mathcal{F}$ with $\mathbb{P}[\Omega_0] = 1$ such that $X(\omega)$ has Property A for every $\omega \in \Omega_0$.

Next, the concept of a *Conditional Expectation* is introduced. This is the usual expectation is conditioned on a $\sigma - algebra$, set or random variable. Now we let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and $\mathcal{G}$ a sub-$\sigma - algebra$ of $\mathcal{F}$ so that:

1. $\mathcal{G}$ is a $\sigma - algebra$;

2. $\mathcal{G} \subseteq \mathcal{F}$.

Definition 27. (Conditional Expectation, *Bain and Engelhardt (1992)* [2])

Let $f(X)$ be an integrable function on $\mathbb{R}$ and let $\mathcal{H} \subseteq \mathcal{F}$ be a sub-$\sigma-field$, then given the usual probability space $(\Omega, \mathcal{F}, P)$, the conditional expectation of $f(X)$ given $\mathcal{H}$ is the random variable, $E[f(X)|\mathcal{H}]$ such that

1. $E[f(X)|\mathcal{G}]$ is $\mathcal{H}-measurable$;

2. For any $A \in \mathcal{H}$,
$$\int_A E^P[f(X)|\mathcal{H}]\,dP = \int_A X\,dP.$$

The conditional expectation has the some useful properties which are implemented in this dissertation. The proposition below highlights some of these properties.

Proposition 12. *(Brzeźniak and Zastawniak (2000) [17])*

Let $\mathcal{H}$ be a $sub-\sigma-algebra$ of $\mathcal{F}$, then the conditional expectation has the following properties:

1. $E(\alpha X + \beta Y|\mathcal{H}) = \alpha E(X) + \beta E(Y)$ for $\alpha, \beta \in \mathbb{R}$;

2. $E(E(X|\mathcal{H})) = E(X)$;

3. $E(XY|\mathcal{H}) = XE(Y|\mathcal{H})$ if X is $\mathcal{H}-measurable$;

4. $E(X|\mathcal{H}) = E(YX)$ if X is independent of $\mathcal{H}$;

5. $E(E(X|\mathcal{G})|\mathcal{H}) = E(X|\mathcal{H})$ if $\mathcal{H} \subset \mathcal{G}$;

6. If $X \geq 0$ then $E(X|\mathcal{H}) \geq 0$

Proof. See [17], Proposition 2.4 on pages 29-31. $\qquad\square$

Definition 28. (Càdlàg Function, Schoutens [47])

A function, $f : E \to \mathbb{R}$, is called a càdlàg if for every $t \in E$ if the left limit $f(t-) = lim_{s\uparrow t}f(s)$ exists and the right limit $f(t+) = lim_{s\downarrow t}f(s)$ exists and are equal to $f(t)$. This is often abbreviated RCLL (right continuous left limit).

Definition 29. (Lévy Measure, Sato (1999) [46])

If X is a Lévy process on $\mathbb{R}$,then its Lévy measure ν on $\mathbb{R}$ is defined as

$$\nu(A) = \frac{1}{t}\mathbb{E}\left(\sum_{0 \leq s \leq t} \mathbb{I}_{\{(X_s - X_{s-}) \in A\}}\right), A \in \mathcal{B}\left(\mathbb{R}\right) \tag{A.3}$$

The jump at time s is defined as $\triangle X(s) = X(s) - X(s-)$ with $X(s-) = lim_{s\uparrow t}X(s)$. The measure is an average number of number of jumps whose size is an element of A for a given period, t. This measure is defined on $\mathbb{R}\backslash\{0\}$ and

$$\int\limits_{-\infty}^{+\infty} inf\left\{1, x^2\right\}\nu(dx) = \int\limits_{-\infty}^{+\infty}\left(1 \wedge x^2\right)\nu\left(dx\right) < \infty \tag{A.4}$$

Lévy measures are characterised with no mass at the origin but infinitely many jumps called *singularities*.

B   Stochastic Calculus

In this section, some fundamentals about stochastic calculus are presented. These preliminaries form a core part of asset pricing theory in mathematical finance which recognises the fact that asset prices behave randomly under uncertainties . This section is mostly drawn from the [11].

Definition 30. (Wiener Process,[11])

A process $W = (W(t), 0 \leq t \leq T)$ is called a Wiener process if the following conditions hold:

1. $W(0) = 0$

2. W has independent increments. For any choice of $n \geq 1$ and $0 \leq t_0 < t_1 < ... < t_n$, the random variables $W(t_0)$, $W(t_1) - W(t_0)$, $W(t_2) - W(t_1)$,....,$W(t_n) - W(t_{n-1})$ are independent stochastic random variables.

3. For $s < t$ distribution of $X(t) - X(s)$ has the normal distribution $N\left(0, \sqrt{t-s}\right)$.

4. X is *stochastically continuous* i.e. for every $t \geq 0$ and $\varepsilon \geq 0$, $\lim\limits_{s \to t} P\left[|W(s) - W(t)| > \varepsilon\right] = 0$

B.1   Construction of the Stochastic Integral

Consider the Wiener process, W, and another stochastic process Y .The class $\mathcal{L}^2$ ensures the existence of the stochastic integral under certain integrability conditions.

Definition 31. (Class $\cdot^2$, page 40[11])

The stochastic process Y belongs to class $E^2[a, b]$ if the following conditions are satisfied:

1. $\int_a^b E\left[Y^2(s)\right] ds < \infty$

2. The process Y is adapted to the filtration $\mathcal{F}_t^W$ (i.e Y can be completely determined given the observations of the path $W = (W_t, 0 \leq t \leq T)$).

The stochastic integral, Y, belong to the class $\mathcal{L}^2$ if $Y \in \mathcal{L}^2[0, t]$ for $t > 0$. The task now is to define a stochastic integral $\int_a^b Y_s dW_s$ for the process $Y \in E^2[a, b]$. This can be carried out in two steps:

Suppose $Y \in \mathcal{L}^2[a, b]$ is simple, that is the interval $[a, b]$ can be partitioned as $a = t_0 < t_1 < ... < t_n = b$ such that Y is constant on each sub-interval. This means that $Y(s) = Y(t_j)$ for each $s \in [t_j, t_{j+1}]$. Then the stochastic integral can be defined as:

$$\int_a^b Y(s)dW_s = \sum_{j=0}^{n-1} Y(t_j) \left(W(t_{j+1}) - W(t_j) \right). \tag{B.1}$$

For a general $Y \in E^2[a, b]$, to define the stochastic integral, Y is approximated by a sequence of simple stochastic processes Y_n as [11]:

1. $\int_a^b \mathbb{E}\left\{ (Y_n(s) - Y(s))^2 \right\} ds \to 0.$

2. For each integral $Z_n := \int_a^b Y_n(s)dW_s$ there exists $Z \in \mathcal{L}^2$ such that $Z_n \to Z$ as $n \to \infty$.

3. The stochastic integral is then defined as: $\int_a^b Y(s)dW_s = \lim_{n \to \infty} \int_a^b Y_n(s)dW_s$.

The most important properties of the stochastic integral can be summarised by the following proposition:

Proposition 13. *(Properties of the stochastic integral, [11])*

Let Y be an $\mathcal{F}_t^W -$ adapted stochastic process that satisfies the following condition:

$$\int_0^t \mathbb{E}\left[Y^2(s) \right] ds < \infty \tag{B.2}$$

then the following hold:

$$\mathbb{E}\left[\int_0^t Y(s)dW_s \right] = 0 \tag{B.3}$$

and

$$\mathbb{E}\left[\left(\int_0^t Y(s)dW_s\right)^2\right] = \mathbb{E}\left[\int_0^t Y(s)^2 ds\right] \tag{B.4}$$

Proof. Let the interval $[0, t]$ be partitioned as $0 = t_0 < t_1 < ... < t_n = t$ and let $\Delta W_j = W_{t_{j+1}} - W_{t_j}$.

Consider a random variable $Z_n = \sum_{j=0}^{n-1} Y(t_j)\Delta W_j$. If we let $n \to \infty$ as $\Delta W_j \to 0$, then the sequence of random variables (Z_n) converges to $\int_0^t Y(s)dW_s$ (mean-square). Then if we consider

$$\mathbb{E}\left[\sum_{j=0}^{n-1} Y(t_j)\Delta W_j\right] = \sum_{j=0}^{n-1} \mathbb{E}\left[Y(t_j)\right]\mathbb{E}\left[\Delta W_j\right]. \tag{B.5}$$

Looking at the random variables, since ΔW_j is independent to W_{t_j} so are $Y(t_j)$ and ΔW_j. Thus the product on the right-hand side is 0. This implies that $\mathbb{E}\left[Z_n\right] = 0$ from which the first result in Equation (B.3) follows.

The proof of (B.4) follows as:

Since $Z_n \to \int_0^t Y(s)dW_s$ (mean square) then it follows that

$$\mathbb{E}\left[Z_n^2\right] \to \mathbb{E}\left[\left(\int_0^t Y(s)dW_s\right)^2\right]. \tag{B.6}$$

This means that the we only look at $\mathbb{E}\left[Z_n^2\right]$. Using the property of independent increments of the Wiener process it follows that $\mathbb{E}\left[\Delta W_j\right] = 0$. We also utilise the fact that $\mathbb{E}\left[(\Delta W_j)^2\right] = \Delta_j s$. then,

$$
\begin{aligned}
Z_n^2 &= \left(\sum_{j=0}^{n-1} Y(t_j) \Delta W_j \right)^2 \\
&= \sum_{j=0}^{n-1} \sum_{k=0}^{n-1} Y(t_j) Y(t_k) \Delta W_j \Delta W_k \\
&= \sum_{j=0}^{n-1} (Y(t_j))^2 (\Delta W_j)^2 + 2 \sum_{k<j}^{n-1} Y(t_j) Y(t_k) \Delta W_j \Delta W_k \quad\text{(B.7)}
\end{aligned}
$$

Since $Y(t_j)$ and ΔW_j are independent we have that :

$$
\mathbb{E}\left((Y(t_j))^2 (\Delta W_j)^2 \right) = \mathbb{E}\left((Y(t_j))^2 \right) \Delta_j s \quad\text{(B.8)}
$$

and for $k < j$

$$
\mathbb{E}\left(Y(t_j) Y(t_k) \Delta W_j \Delta W_k \right) = \mathbb{E}\left(Y(t_j) Y(t_k) \Delta W_j \right) \mathbb{E}\left(\Delta W_k \right) = 0 \quad\text{(B.9)}
$$

then from Equations (B.7), (B.8) and (B.9), it follows that:

$$
\mathbb{E}\left[Z_n^2 \right] = \sum_{j=0}^{n-1} \mathbb{E}\left((Y(t_j))^2 \right) \Delta_j s \quad\text{(B.10)}
$$

Looking at Equation (B.10) above, the right-hand side is the Riemann sum which converges to $\int_0^t Y_s ds$. Then it follows that:

$$\mathbb{E}\left(\left|\int_0^t Y(s)dW_s\right|^2\right) = \lim_{n\to\infty}\mathbb{E}\left[Z_n^2\right]$$

$$= \lim_{n\to\infty}\sum_{j=0}^{n-1}\mathbb{E}\left((Y(t_j))^2\right)\Delta_j s$$

$$= \int_0^t \mathbb{E}\left(Y(s)^2\right)ds$$

$$= \mathbb{E}\left[\int_0^t Y(s)^2 ds\right]$$

The last step is obtained from the fact that the definition of an expected value is in essence an integral , thus the order of integration can be reverse to give the second result. This concludes the proof. □

B.2 The Itô formula

Definition 32. (Itô Process)

A stochastic process $X = (X(t), t \geq 0)$ is an Itô process is it is defined by the equation:

$$X(t) = X(0) + \int_0^t Y(s)ds + \int_0^t Z(s)dW_S \tag{B.11}$$

where $Y(s), Z(s)$ are driven by the Brownian motion and satisfy:

$$\int_0^t E\left[Y(s)^2\right]ds < \infty \tag{B.12}$$

and,

$$\int_0^t E\left[Z(s)^2\right]ds < \infty \tag{B.13}$$

The Itô process can be re-written in stochastic differential form as:

$$dX(t) = Y(t)dt + Z(t)dW_t. \tag{B.14}$$

The following theorem presents the Itô formula without proof. It's heuristic proof can be given from second Taylor expansion where higher order terms are considered negligible.

Theorem 7. *(Itô formula, [11])*

Let the stochastic process X have a stochastic differential equation defined by:

$$dX(t) = \mu(t)dt + \sigma(t)dW_t$$

where $\mu(t)$ and $\sigma(t)$ are adapted process and let $H \in \mathcal{C}^{1,2}$ be a real-valued function. If we define a stochastic process $Z_t = H(t, X(t))$, then Z has a stochastic differential equation given by:

$$dH(t, X(t)) = H_1(t, X(t))dt + H_2(t, X(t)) \, dX_t + \frac{1}{2}H_{22}(t, X(t)) \, (dX(t))^2 \tag{B.15}$$

where we define the notation as:

$$
\begin{aligned}
H_1(t, x) &= \frac{\partial}{\partial t}H(t, x) \\
H_2(t, x) &= \frac{\partial}{\partial x}H(t, x) \\
H_{12}(t, x) &= \frac{\partial^2}{\partial t \partial x}H(t, x) \\
H_{22}(t, x) &= \frac{\partial^2}{\partial x^2}H(t, x)
\end{aligned}
$$

and $(dX(t))^2$ is calculated form the Itô multiplication table:

	dt	dW_t
dt	0	0
dW_t	0	dt

Table 10:   Itô multiplication table

B.3   Stochastic Calculus for Lévy processes

In this section stochastic integrals whose integrators are Lévy processes are examined. The Itô's formula for Lévy-type stochastic integrals together with the conditions for existence and uniqueness is presented.

Definition 33. (One-dimensional Lévy-type stochastic process, [1])

Let $W = \{W(t) : 0 \leq t \leq T\}$ be Wiener process and $N(dt, dx)$ a Poisson random measure on $\mathbb{R}^+$ with a corresponding compensated Poisson random measure $\tilde{N}(dt, dx)$. Then the Lévy-type stochastic process $Y = \{Y(t) : 0 \leq t \leq T\}$ on $(\Omega, \mathcal{F}, \mathbb{P})$ has a stochastic integral of the form:

$$Y_t = Y_0 + \int_0^t g(s)ds + \int_0^t f(s)dW_s + \int_0^t \int_{\mathbb{R}} h(s,x)\, \tilde{N}(ds, dx) \tag{B.16}$$

$$+ \int_0^t \int_{\mathbb{R}} k(s,x)\, N(ds, dx)$$

where $g : [0,T] \to \mathbb{R}, f : [0,T] \to \mathbb{R}, h : [0,T] \times \mathbb{R} \to \mathbb{R}$ and $k : [0,T] \times \mathbb{R} \to \mathbb{R}$ are real-valued functions satisfying the following conditions:

$\int_0^t |g(s)|ds < \infty$; $\int_0^t f^2(s)ds < \infty$; $\int_0^t \int_{\mathbb{R}} h^2(s,x)\, \nu(dx)ds < \infty$ and $\int_0^t \int_{\mathbb{R}} k^2(s,x)\, \nu(dx)ds < \infty$.

The stochastic integral in (B.16) above can be re-written in the form:

$$
\begin{aligned}
dY_t &= g(t)dt + f(t)dW_t + \int_{\mathbb{R}} h(t,x)\tilde{N}(dt,dx) \qquad (B.17)\\
&\quad + \int_{\mathbb{R}} k(t,x)N(dt,dx)
\end{aligned}
$$

Theorem 8. (Itô's Theorem 2, [1])

Let Y be a Lévy-type stochastic integral given by:

$$
\begin{aligned}
dY_t &= g(t)dt + f(t)dW_t + \int_{\mathbb{R}} h(t,x)\tilde{N}(dt,dx)\\
&\quad + \int_{\mathbb{R}} k(t,x)N(dt,dx)
\end{aligned}
$$

then for each real-valued function $H \in \mathcal{C}^2$, a.s we have that:

$$
\begin{aligned}
H(Y_t) - H(Y_0) &= \int_0^t H_1\left(g(t)dt + f(t)dW_t\right) + \frac{1}{2}\int_0^t H_{12}\left(g(t)dt + f(t)dW_t\right)^2\\
&\quad + \int_0^t \int_{\mathbb{R}} \left[H(Y_{s^-}) + k(s,x) - H(Y_{s^-})\right]N(ds,dx)\\
&\quad + \int_0^t \int_{\mathbb{R}} \left[H(Y_{s^-}) + h(s,x) - H(Y_{s^-})\right]N(ds,dx)\\
&\quad + \int_0^t \int_{\mathbb{R}} \left[H(Y_{s^-}) + h(s,x) - H(Y_{s^-})\right.\\
&\quad \left. - h(s,x)H_1\right]\nu(dx)ds
\end{aligned}
$$

where $H_1 = \frac{\partial}{\partial t}(H(Y_{s^-})); H_2 = \frac{\partial}{\partial x}(H(Y_{s^-}))$ and $H_{12} = \frac{\partial^2}{\partial t \partial x}(H(Y_{s^-}))$.

Proof. See [1] page 226. □

Theorem 9. (Existence and Uniqueness of Lévy-type SDE, [1])

Consider the modified SDE

$$dX(t) = b(X(t))dt + \sigma(X(t))dW_t + \int_{\mathbb{R}} F(X(t), x)\tilde{N}(dt, dz) \qquad (B.18)$$

with initial condition

$$X(0) = X_0.$$

where $b : [0, T] \times \Omega \to \mathbb{R}^{N \times M}; \sigma : [0, T] \times \Omega \to \mathbb{R}^{N \times M}$ and $F : [0, T] \times \mathbb{R}^l \times \Omega \to \mathbb{R}^{N \times l}$.

If we impose the following conditions:

- **(C1) Lipschitz Condition** . There exists a $K_1 > 0$ such that for all $y_1, y_2 \in \mathbb{R}^N$

$$|b(y_1) - b(y_2)|^2 + \|\sigma(y_1) - \sigma(y_2)\|^2$$

$$+ \int_{\mathbb{R}} |F(y_1, x) - F(y_2, x)|^2 \nu(dx) \le K_1 |y_1 - y_2|^2$$

- **(C2) Growth Condition.** There exists a $K_2 > 0$ such that for all $y \in \mathbb{R}^N$

$$|b(y)|^2 + \|\sigma(y)\|^2 + \int_{\mathbb{R}} |F(y, x)|^2 \nu(dx) \le K_2 \left(1 + |y|^2\right)$$

Then the process $X = \{X_t : t \ge 0\}$ has a unique solution.

Proof. See [1] page 305. $\qquad \square$

C MATLAB Code

```
%This code demonstrates the Demeterfi et al.  (1999) hedging of variance %swaps
function y= Replication2(S0,Val,T,r,sigma)
%Replication2(292.373,10,T,0.00038344,0.0108)
K=linspace(S0-Val,S0+Val)';
% length of the Strike price column vector /array n=length(K);
f=zeros(n,1); P=zeros(n,1); C=zeros(n,1); Df=zeros(n,1);
weight=zeros(n,1); Contrib=zeros(n,1);
%Puts for i=1:(length(K)/2)-1 %Demeterfi (1999) approximation
f(i,1) = (2/T)*(((K(i,1)-S0)/S0)-log(K(i,1)/S0)); % Df(i,1)=
(f(i,1)-f(i+1,1))/(K(i+1,1)-K(i,1)); if i==(length(K)/2)-1 weight(i,1)= Df(i,1);
end P(i,1)= bsput(S0,K(i,1),r,T,sigma); if i<(length(K)/2)-1 weight(i,1)=
Df(i,1)-Df(i+1,1); end Contrib(i,1)=P(i,1)*weight(i,1); end
%Calls for i=(length(K)/2):(length(K))-1 %Demeterfi (1999) approximation
f(i,1) = (2/T)*(((K(i,1)-S0)/S0)-log(K(i,1)/S0)); % Df(i,1)=
(f(i+1,1)-f(i,1))/(K(i+1,1)-K(i,1)); if i==(length(K)/2) weight(i,1)= Df(i,1);
end C(i,1)= bscall(S0,K(i,1),r,T,sigma); if i>(length(K)/2) weight(i,1)=
Df(i,1)-Df(i-1,1); end Contrib(i,1)=C(i,1)*weight(i,1); end V=sum(Contrib);
Kvar = (2/T)*(r*T-(((S0/S0)*exp(r*T))-1))+(exp(r*T)*V);
y= Kvar;
end
```

```
%Estimate of Price of a variance swap under the B-NS Model
%Parameters for NIG(alpha,beta,delta,mu=0)
alpha= 78.1947; %The tail-heaviness parameter of the NIG distribution
beta=21.9762; %The asymmetry parameter of NIG distribution delta=0.009098;

%The scale parameter of the NIG lambda1=0.9; %Decay rate estimate for the first

OU lambda2=0.03; %Decay rate estimate for the second OU Y1=0.000065817;

%Variance 1 (of the 2 superpositioned variance processes) Y2=0.000065817;

%Variance 2 w1=0.5; %Weight for variance 1 w2=0.5; %Weight for variance 2

Vstart=0;

%The initial realised variance Vend=0.3; %The realised variance at T NumOfVs=10;

%Number of Spaces between Vstart and Vend nSims=10; %Number of Variance prices

to be calculated

gamma=sqrt(alpha^2-beta^2); %The gamma from NIG
Price=zeros(nSims,1); Price1=zeros(nSims,1); Price2=zeros(nSims,1);
%realised variances at t vt=0.041304663; t=1; %Current time or time zero T=10;
%maturity (in years)
%integral 1 in the variance price

syms s f1= (delta*((gamma^2)- s)/((gamma^2)-2*s)^(3/2))*(1-exp(-lambda1*(T-t)));

m1=vpaintegral(f1,s,[t T]);

%integral 2 in the variance price

syms s f2= (delta*((gamma^2)- s)/((gamma^2)-2*s)^(3/2))*(1-exp(-lambda2*(T-t)));

m2=vpaintegral(f2,s,[t T]);

%The price for various realised variances for i=1:T Price=
(t/T)*(vt) + ((w1/(T*lambda1))*(1-exp(lambda1*(T-t)))*Y1)+
((w2/(T*lambda2))*(1-exp(lambda2*(T-t)))*Y2);

Price2(i,1)= ((w1/T)*m1)+ ((w2/T)*m2); Price(i,1)=(Price1(i,1)+Price2(i,1)); end

figure hold on plot(x0,Price,':bs') title('Price of a variance swap')

xlabel('Realised Variance') ylabel('Price')
```

```matlab
%This script impliments the Disrete model for pricing variance swaps which was %
obtained using PDEs of the Heston model.
%Heston Parameters:  N = 100000;% Number of sampling times
kappa=0.0123800596307453; theta=0.00735002387651258; mu=0.00038344;
sigmaV=0.00344581787186608; S0=263.68; V0=6.66019106816959e-07;
rho=-0.757595353042044;
T=1; dt=T/N; %parameter for calculating f(V0) a = kappa-2*rho*sigmaV; b =
sqrt(a^2-(2*sigmaV^2));
%g = ((a/sigmaV)^2)-1+(a/sigmaV)*sqrt(((a/sigmaV)^2)-2); g= (a+b)/(a-b); C_dt =
mu*dt+ ((kappa*theta)/sigmaV^2)*((a+b)*dt-2*log((1-g*exp(b*dt))/(1-g)));
D_dt =((a+b)/sigmaV^2)*((1-exp(b*dt))/(1-(g*exp(b*dt))));
%Formula for f(V0) f_V0=exp(C_dt+(D_dt*V0))+exp(-mu*dt)-2;
%Calculating sum fi(V0) for N sampling times f=zeros(N,1);
for i=2:N f(1,1)=f_V0; c_i=(2*kappa)/((sigmaV^2)*(1-exp(-kappa*((i-1)/N))));
f(i,1)= exp(C_dt + (c_i* (exp(-kappa*((i-1)/N))/(c_i-D_dt))*D_dt*V0))*...
((c_i/(c_i-D_dt))^((2*kappa*theta)/sigmaV^2))+exp(-mu*dt)-2; end
sum_f=sum(f);
Kvar=((exp(mu*dt)/T)*(sum_f)); Kvar1=(V0*((1-exp(-kappa*T))/(kappa*T)) +
theta*(1- ( (1-exp(-kappa*T))/(kappa*T))));
```

```
% This Script defines the pdf of the Normal Inverse Gaussian distribution % fx =
NIGPDF(X, ALPHA, BETA, MU, DELTA) returns the pdf of the NIG i.e
% fx = (((alpha*delta)/pi)*bessek(1, (alpha*z)))/(pi*z)*exp(delta*sqrt(alpha^2 -
beta^2)) % *exp(beta(x-mu)) % function y=NIG_distribution(x,alpha,beta,mu,delta)
%%Constraints for the parameters if alpha <= 0 error('ALPHA must be positive.');
end if delta <= 0 error('DELTA must be positive.'); end
if abs(beta)>=alpha error('|BETA| must be smaller than ALPHA'); end % transform
input into column vector [nx, mx] = size(x); x = reshape(x, nx * mx, 1);
% calculate leading factor q = (delta * alpha / pi) * exp(delta * sqrt(alpha^2 -
beta^2)); xbar = x - mu; z = sqrt(delta^2 + xbar .* xbar);
% modified Bessel function of the third kind K = besselk(1, (alpha * z)); y = q
./ z .* K .* exp(beta * (x - mu));
% transform input vector back to input format y = reshape(y, nx, mx);
y(isinf(x)) = 0;
```

References

[1] David Applebaum. *Lévy processes and stochastic calculus*. Cambridge university press, 2009.

[2] Lee Bain and Max Engelhardt. *Introduction to Probability and Mathematical Statistics*, volume 2. Duxbury Classic Series, 1992.

[3] Ole Barndorff-Nielsen and Christian Halgreen. Infinite divisibility of the hyperbolic and generalized inverse gaussian distributions. *Probability Theory and Related Fields*, 38(4):309–311, 1977.

[4] Ole E Barndorff-Nielsen. Processes of normal inverse gaussian type. *Finance and stochastics*, 2(1):41–68, 1997.

[5] Ole E Barndorff-Nielsen and Neil Shephard. 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, 2001.

[6] Ole E Barndorff-Nielsen and Neil Shephard. 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, 2001.

[7] G de Barra. Measure theory and integration. *John Willy and sons*, 1981.

[8] Fred Espen Benth, Martin Groth, and Rodwell Kufakunesu. Valuing volatility and variance swaps for a non-gaussian ornstein–uhlenbeck stochastic volatility model. *Applied Mathematical Finance*, 14(4):347–363, 2007.

[9] Fred Espen Benth and Jūratė Šaltytė-Benth. The normal inverse gaussian distribution and spot price modelling in energy markets. *International journal of theoretical and applied finance*, 7(02):177–192, 2004.

[10] Carole Bernard and Zhenyu Cui. Prices and asymptotics for discrete variance swaps. *Applied Mathematical Finance*, 21(2):140–173, 2014.

[11] Tomas Björk. *Arbitrage theory in continuous time*. Oxford university press, 2009.

[12] Fischer Black and Myron Scholes. The pricing of options and corporate liabilities. *Journal of political economy*, 81(3):637–654, 1973.

[13] Mark Broadie and Ashish Jain. The effect of jumps and discrete sampling on volatility and variance swaps. *International Journal of Theoretical and Applied Finance*, 11(08):761–797, 2008.

[14] Mark Broadie and Ashish Jain. Pricing and hedging volatility derivatives. *Journal of Derivatives*, 15(3):7, 2008.

[15] Mark Broadie and Özgür Kaya. Exact simulation of stochastic volatility and other affine jump diffusion processes. *Operations research*, 54(2):217–231, 2006.

[16] Oliver Brockhaus and Douglas Long. Volatility swaps made simple. *RISK-LONDON-RISK MAGAZINE LIMITED-*, 13(1):92–95, 2000.

[17] Zdzislaw Brzezniak and Tomasz Zastawniak. *Basic stochastic processes: a course through exercises*. Springer Science & Business Media, 2000.

[18] Peter Carr, Hélyette Geman, Dilip B Madan, and Marc Yor. Pricing options on realized variance. *Finance and Stochastics*, 9(4):453–475, 2005.

[19] Peter Carr and Roger Lee. Robust replication of volatility derivatives. In *Prmia award for best paper in derivatives, mfa 2008 annual meeting*. Citeseer, 2008.

[20] Peter Carr and Roger Lee. Volatility derivatives. *Annu. Rev. Financ. Econ.*, 1(1):319–339, 2009.

[21] Peter Carr, Roger Lee, and Liuren Wu. Variance swaps on time-changed lévy processes. *Finance and Stochastics*, 16(2):335–355, 2012.

[22] Peter Carr and Dilip Madan. Towards a theory of volatility trading. *Volatility: New estimation techniques for pricing derivatives*, 1998.

[23] Peter Carr and Dilip Madan. Towards a theory of volatility trading. *Volatility: New estimation techniques for pricing derivatives*, 1998.

[24] Peter Carr and Dilip Madan. Option valuation using the fast fourier transform. *Journal of computational finance*, 2(4):61–73, 1999.

[25] Kresimir Demeterfi, Emanuel Derman, Michael Kamal, and Joseph Zou. More than you ever wanted to know about volatility swaps. *Goldman Sachs quantitative strategies research notes*, 41:1–56, 1999.

[26] Bruno Dupire. Pricing with a smile. risk 7, 18-20. *EBRI Issue Brief*, 1993.

[27] Bruno Dupire. A unified theory of volatility. *Derivatives pricing: The classic collection*, pages 185–196, 1996.

[28] Bruno Dupire. Pricing and hedging with smiles. *Mathematics of derivative securities*, 1(1):103–111, 1997.

[29] Ernst Eberlein and Sebastian Raible. Term structure models driven by general lévy processes. *Mathematical Finance*, 9(1):31–53, 1999.

[30] Robert J Elliott, Tak Kuen Siu, and Leunglung Chan. Pricing volatility swaps under heston's stochastic volatility model with regime switching. *Applied Mathematical Finance*, 14(1):41–62, 2007.

[31] Anders Eriksson, Eric Ghysels, and Fangfang Wang. The normal inverse gaussian distribution and the pricing of derivatives. *Journal of Derivatives*, 16(3):23, 2009.

[32] Jeff Fleming, Barbara Ostdiek, and Robert E Whaley. Predicting stock market volatility: A new measure. *Journal of Futures Markets*, 15(3):265–302, 1995.

[33] Jim Gatheral. *The volatility surface: a practitioner's guide*, volume 357. John Wiley & Sons, 2011.

[34] Steven L Heston. A closed-form solution for options with stochastic volatility with applications to bond and currency options. *The review of financial studies*, 6(2):327–343, 1993.

[35] Steven L Heston. A closed-form solution for options with stochastic volatility with applications to bond and currency options. *The review of financial studies*, 6(2):327–343, 1993.

[36] John Hull and Alan White. The pricing of options on assets with stochastic volatilities. *The journal of finance*, 42(2):281–300, 1987.

[37] John Hull and Alan White. The pricing of options on assets with stochastic volatilities. *The journal of finance*, 42(2):281–300, 1987.

[38] Zhaoli Jia, Xiuchun Bi, and Shuguang Zhang. Pricing variance swaps under stochastic volatility with an ornstein-uhlenbeck process. *Journal of Systems Science and Complexity*, 28(6):1412–1425, 2015.

[39] Anna Kalemanova, Bernd Schmid, and Ralf Werner. The normal inverse gaussian distribution for synthetic cdo pricing. *The journal of derivatives*, 14(3):80–94, 2007.

[40] Andy Kiersz. Here are the biggest one-day point drops in the dow's history, February 2018.

[41] Thomas Little and Vijay Pant. A finite difference method for the valuation of variance swaps. In *Quantitative Analysis In Financial Markets: Collected Papers of the New York University Mathematical Finance Seminar (Volume III)*, pages 275–295. World Scientific, 2001.

[42] Robert C Merton. Option pricing when underlying stock returns are discontinuous. *Journal of financial economics*, 3(1-2):125–144, 1976.

[43] Anthony Neuberger. The log contract. *Journal of portfolio management*, 20:74–74, 1994.

[44] Elisa Nicolato and Emmanouil Venardos. Option pricing in stochastic volatility models of the ornstein-uhlenbeck type. *Mathematical finance*, 13(4):445–466, 2003.

[45] Fabrice D Rouah. *The Heston Model and Its Extensions in Matlab and C*. John Wiley & Sons, 2013.

[46] Ken-iti Sato and Sato Ken-Iti. *Lévy processes and infinitely divisible distributions*. Cambridge university press, 1999.

[47] Wim Schoutens. *Lévy processes in finance*. Wiley, 2003.

[48] Anatoliy Swishchuk. Modeling of variance and volatility swaps for financial markets with stochastic volatilities. *WILMOTT magazine*, 2:64–72, 2004.

[49] Anatoliy Swishchuk and Li Xu. Pricing variance swaps for stochastic volatilities with delay and jumps. *International Journal of Stochastic Analysis*, 2011, 2011.

[50] Luogen Yao, Gang Yang, and Xiangqun Yang. A note on the mean correcting martingale measure for geometric lévy processes. *Applied Mathematics Letters*, 24(5):593–597, 2011.

[51] Song-Ping Zhu and Guang-Hua Lian. A closed-form exact solution for pricing variance swaps with stochastic volatility. *Mathematical Finance*, 21(2):233–256, 2011.