Incorporating market attention in option pricing with applications to Bitcoin derivatives

Álvaro Guinea Juliá

Doctor of Philosophy
University of York
Mathematics

March 2022

Declaration of Authorship

I, Álvaro Guinea Juliá, declare that this thesis titled, “Incorporating market attention in option pricing with applications to Bitcoin derivatives ” and the work presented in it are my own. I confirm that:

- This work was done wholly or mainly while in candidature for a research degree at this University.
- Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated.
- Where I have consulted the published work of others, this is always clearly attributed.
- Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work.
- I have acknowledged all main sources of help.
- Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself.

Signed: Álvaro Guinea Juliá

Date: 27/03/2022

"*All models are wrong, but some are useful.*"

George E. P. Box

UNIVERSITY OF YORK

Abstract

University of York
Mathematics

Doctor of Philosophy

Incorporating market attention in option pricing with applications to Bitcoin derivatives

by Álvaro Guinea Juliá

The attention that media or investors pay to the market affects the prices of stocks and assets. This attention is usually called market attention or market interest. It has been shown that this attention affects stocks and indexes. In recent years, due to the development of cryptocurrencies, there is an increasing literature that analyzes the relation between cryptocurrency prices and market attention. Because the value of cryptocurrencies has increased during the last years, new exchanges have appeared that offer European options on Bitcoin.

In this thesis we develop six different models that incorporate market attention into the modelling of Bitcoin option prices. Firstly, we construct two continuous time models that incorporate market attention into the volatility structure, building on existing work by Cretarola, Figà-Talamanca, and Patacca (2020). For these two models we show how we can estimate the parameters and give a closed formula for pricing European options. Then we construct two continuous time models that contain jumps in the price structure to take into account that the distribution of Bitcoin returns has fat tails. Again, for these models, we estimate the parameters and develop a closed formula for pricing Bitcoin options. Lastly, we construct two discrete time models in which the volatility is explained by the market attention but also by an unobserved process. The estimation of these models is quite complex. Because of that, we will use sequential Monte Carlo methods for the estimation of these models.

Acknowledgements

First of all, I would like to express my gratitude to my supervisor Dr. Alet Roux for her patience and for teaching me how to think critically. I would also like to thank to my TAP members Dr. Andrea Meireles Rodrigues and Dr. Alexei Daletskii for their questions and for creating a good atmosphere during the TAP meetings. Many thanks to the Department of Mathematics at the University of York for their financial support in the form of a graduate teaching studentship.

In addition I would like to thank to my friends of the 66 Ambrose Gang: Chiara, Laura, Martín, Ángela and Mahy for their support and for making this journey easier. I would also like to thank to Peiyun and the people of the Department of Mathematics for their help.

Lastly, I would like to thank to my parents and to my brother for their help and their endless support during these four years.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

The term market attention is defined as the attention that investors or media pay to a particular stock or asset. This term is also called investor attention or market interest. Different proxies can be selected to measure market attention. Traditionally, trading volume, news coverage, or extreme past returns have been used. These proxies have been shown to affect stock prices (see K. Hou, Xiong, and Peng (2009) and J. Chen, Tang, Yao, and Zhou (2022)). Due to the increase in the use of the Internet in the last two decades, new proxies for market attention have appeared, such as the number of Google searches or the Twitter volume. Da, Engelberg, and Gao (2011) show that the number of Google searches affects the prices of stocks in the Russel 3000 index. It has also been shown that Twitter sentiment affects the prices of the stocks in the Dow Jones Industrial Average index (Ranco et al., 2015). In addition, Twitter volume also affects the prices of options (Wei, Mao, & Wang, 2016).

In this thesis, we are interested in developing models for pricing Bitcoin options that incorporate market attention. Bitcoin was presented by Nakamoto (2008) in 2008 and since then the number of people buying and selling Bitcoin has rapidly increased. In recent years, due to the increase in the value of Bitcoin, new exchanges have emerged that offer European Bitcoin options. Some of these exchanges are Deribit (https://www.deribit.com/), LedgerX (https://derivs.ftx.us/) and Bit (https://www.bit.com/).

There is a constantly increasing literature that studies the relation between cryptocurrencies and market attention. For example, Smales (2022) builds a panel regression model to show that the number of Google searches generates greater returns and greater volatility for the most important cryptocurrencies. Eom, Kaizoji, Kang, and Pichl (2019) construct an autoregressive model for volatility that includes the number of Google searches and they show that the number of Google searches improves the predictability of Bitcoin volatility. Suardi, Rasel, and Liu (2022) use VADER (Valence Aware Dictionary for Sentiment Reasoning) to extract information from a collection of tweets that contain the hashtag Bitcoin. They construct two factors, that they call sentiment dispersion and investor attention. Using several econometric models, they show that an increase in sentiment dispersion increases volatility, and investor attention predicts trading volume. Also, Philippas, Rjiba, Guesmi, and Goutte (2019) construct a dual process diffusion model to express that media attention (Google searches and Twitter volume) partially affects the prices of Bitcoin. Furthermore, López-Cabarcos, Pérez-Pico, Piñeiro-Chousa, and Šević (2021) construct a sentiment index for Bitcoin using the software Stanford Core NLP and the web page StockTwits.com. The authors then build GARCH type models and show that the sentiment index affects the volatility of Bitcoin returns. Aalborg, Molnár, and de Vries (2019) construct several linear regression models to show that trading volume affects Bitcoin volatility. In the work of Al Guindy (2021), the author uses a VAR model to prove that an increase in investor attention, in this case it is the Twitter volume of the last five days, produces an increase in Bitcoin volatility. In addition, using a GARCH-MIDAS model, Liang, Zhang, Li, and Ma (2022) show that the number of Google searches has an impact on predicting the volatility of Bitcoin. Kristoufek (2015) uses wavelet analysis to indicate that an increase in interest (number of Google searches and Wikipedia views) produces an increase in prices during bubble formation and a decrease in prices during bubble burst. Evidence of this phenomenon can be found in the work of Zhang, Lu, Tao, and Wang (2021), where the authors show using the Granger causality test that an increase in the number of Google searches contributes to bubble formation in Bitcoin prices. Ciaian, Rajcaniova, and Kancs (2016) show using time series models that the number of views on Wikipedia also has an impact on the price of Bitcoin in the short term. Also, Aslanidis, Bariviera, and López (2021) build a Google trends cryptocurrency index and show that there is a short term bidirectional relation between the price of Bitcoin and the Google trends index. Figà-Talamanca and Patacca (2019) using ARMA-GARCH models show that trading volume affects the mean and the volatility of Bitcoin returns, while the number of Google searches only affects the volatility of Bitcoin returns. In other paper (Figà-Talamanca & Patacca, 2020), the same authors using a VAR-EGARCH model show that neither the trading volume nor the number of Google searches affect the mean of Bitcoin returns. However, both attention proxies affect the volatility of Bitcoin returns.

We observe that the results in the literature seem to disagree in some aspects. Some authors claim that market attention influences the mean and volatility of Bitcoin returns (Ciaian, Rajcaniova, & Kancs, 2016; Figà-Talamanca & Patacca, 2019; Kristoufek, 2015; Smales, 2022). While other sources have found that market attention affects only the volatility of Bitcoin returns (Aalborg, Molnár, & de Vries, 2019; Al Guindy, 2021; Figà-Talamanca & Patacca, 2020; Suardi, Rasel, & Liu, 2022). In this thesis, we will assume that market attention affects only the volatility of Bitcoin returns; see Chapter 3.

In most of the econometric models that appear in the previous literature review, market attention is included in those models with a delay. That is, market attention does not act instantaneously with respect to the response variable (usually the return of Bitcoin or the volatility of Bitcoin returns). Because of that, in all models included in this thesis, market attention acts on the return of Bitcoin with a constant delay.

There is previous work in the literature that analyses the problem of Bitcoin option pricing. Because Bitcoin prices have high volatility, several models include jumps in the price structure (K. Chen & Huang, 2021; Olivares, 2020; Shirvani, Mittnik, Lindquist, & Rachev, 2021). Cao and Celik (2021) propose an equilibrium model to price Bitcoin options. Cretarola, Figà-Talamanca, and Patacca (2020) construct a stochastic volatility model that incorporates market attention into the option pricing. We base the models presented in Chapter 3 and Chapter 4 on this model. All previous references build models that are continuous in time, but discrete time models have also been used for pricing Bitcoin options (Siu & Elliott, 2021; Venter, Mare, & Pindza, 2020).

In this thesis, we propose different models to price Bitcoin options that incorporate market attention. In Chapter 2, we model the interest as a stochastic process and show how we can perform estimation and validation. In Chapter 3, we construct a first simple model for pricing Bitcoin options. In this model the volatility of the log-returns is proportional to the market attention. Based on the model in Chapter 3, we construct a new model in Chapter 4 that allows us to include jumps in the volatility structure. In Chapter 5, we build a time changed model, in which an increase in interest produces an increase in the probability of a price jump. Chapter 6 is somewhat distinct from the rest of the thesis. In this chapter, we build models in which the volatility is explained by an unobserved process and by market attention. To estimate these types of models, we need to use a set of techniques called sequential Monte Carlo methods.

Before introducing the models of this thesis, we need to introduce some mathematical concepts and some tools that we will use through the thesis.

1.1 Lévy process

In this section, we will provide a brief presentation of Lévy processes. To this end, let us take $(\Omega, \mathcal{F}, \mathbb{P}, (\mathcal{F}_t))_{t\geq 0}$ to be a filtered probability space. Here, we are interested in knowing when an adapted stochastic process X is a Lévy process and what its main properties are. We will assume the filtration $(\mathcal{F}_t))_{t\geq 0}$ to be the filtration generated by the process X.

Definition 1.1.1. *(Pascucci, 2011, Definition 13.10) An adapted stochastic process $X = (X(t))_{t\geq 0}$ that takes values in $\mathbb{R}$, is a Lévy process if $X(0) = 0$ almost surely and*

1. *X has independent increments, that is, $X(t) - X(s)$ is independent of $\mathcal{F}_s$ for all $0 \leq s < t$.*

2. *X has stationary increments, that is, $X(t) - X(s)$ has the same distribution as $X(t - s)$, for all $0 \leq s < t$.*

3. *X is stochastically continuous that is, for all $\epsilon > 0$ and for all $t \leq 0$, we have*

$$\lim_{h\to 0} \mathbb{P}\left(|X(t+h) - X(t)| > \epsilon\right) = 0.$$

An interesting property of Lévy processes is that they have almost surely cádlág paths.

Theorem 1.1.1. *(Applebaum, 2009, Theorem 2.1.7) Every Lévy process has a cádlág modification that is itself a Lévy process.*

In this thesis, when we work with a Lévy process, we always work with its cádlág modification. Because Lévy processes have almost surely cádlág paths, the number of jumps larger than any $\epsilon > 0$ is finite. However, the number of jumps smaller than ϵ is countable and could be infinite (Pascucci, 2011, p. 438). That is, in the case that a Lévy process has jumps, it has a finite number of large jumps, but it could have an infinite number of small jumps.

Example 1.1.1. *(Pascucci, 2011, Definition 13.8)*

One example of a Lévy process is the compound Poisson process. This process X is formed by a Poisson process N with parameter $\lambda > 0$ and a sequence $\{Z_n\}_{n=1}^{\infty}$ of independent and identically distributed random variables, which are independent of the Poisson process N. The compound Poisson process X is defined as

$$X(t) = \sum_{n=1}^{N(t)} Z_n \textit{ for } t \geq 0.$$

1.1.1 Infinitely divisible distributions

Lévy processes are related to the concept of infinitely divisible distributions.

Definition 1.1.2. *(Pascucci, 2011, Definition 13.16) A random variable Y is said to be infinitely divisible if, for any $n \geq 2$, there exist independent and identically distributed random variables $Y_1, \ldots, Y_n$ such that:*

$$Y \stackrel{d}{=} Y_1 + \ldots + Y_n,$$

where $\stackrel{d}{=}$ means equal in distribution.

Examples of infinitely divisible distributions are the normal distribution, the Poisson distribution, the inverse Gaussian distribution, or the gamma distribution. The following result shows how Lévy processes and infinitely divisible distributions relate.

Proposition 1.1.1. *(Cont & Tankov, 2004, Proposition 3.1) Let X be a Lévy process. Then for every $t \geq 0$, $X(t)$ has an infinitely divisible distribution. Conversely, if F is an infinitely divisible distribution then there exists a Lévy process X such that the distribution of $X(1)$ is given by F.*

We are now interested in the structure of the characteristic function of Lévy processes.

Theorem 1.1.2. *(Pascucci, 2011, Theorem 13.15) If X is a Lévy process, then there exists a unique continuous function $\Psi^X : \mathbb{R} \to \mathbb{C}$ such that $\Psi(0) = 0$ and*

$$\Phi^{X(t)}(u) = E\left[e^{iuX(t)}\right] = e^{t\Psi^X(u)}, \quad t \geq 0, \; u \in \mathbb{R}.$$

The function Ψ^X is called the characteristic exponent of X.

Theorem 1.1.2 tells us that the distribution of a Lévy process X is determined by the characteristic exponent or what is the same by the distribution of $X(1)$. If X is a Lévy process then we have from Theorem 1.1.2 that

$$\Phi^{X(t)}(u) = \left(\Phi^{X(\frac{t}{n})}(u)\right)^n, \text{ for } u \in \mathbb{R}, \; n \in \mathbb{N}, \, t \geq 0.$$

It is possible to obtain an analytical formula for the characteristic function of $X(t)$ where X is a Lévy process. But first we have to introduce the concepts of Lévy measure and Jump measure.

1.1.2 Jump measures and Lévy–Itô decomposition

Let us now study the jumps of Lévy processes. If X is a Lévy process, the size of the jump at time $t \geq 0$ is defined as $\Delta X(t) = X(t) - X(t^-)$. Since X has almost surely cádlág paths we have that $\Delta X(t) = 0$ almost surely for a fixed $t \geq 0$ (Applebaum, 2009, Lemma 2.3.2). It is possible to define a random measure that counts the number of jumps in a given time period $[0, t]$. We define the random measure J as

$$\begin{aligned} J(t, H) &= \#\{s \in [0, t] : \Delta X(s) \in H\} \\ &= \sum_{0 \leq s \leq t} \mathbb{1}_H(\Delta X(s)) \text{ for } H \in \mathcal{B}(\mathbb{R} \setminus \{0\}) \text{ and } t \geq 0, \end{aligned}$$

where

$$\mathbb{1}_H(x) = \begin{cases} 1 & \text{if } x \in H, \\ 0 & \text{otherwise.} \end{cases}$$

The random measure J is usually called the jump measure of the process X.

Using the jump measure J, it is possible to define a new measure v on $\mathcal{B}(\mathbb{R}\setminus\{0\})$ as

$$v(H) = E\left[J(1,H)\right], \text{ for } H \in \mathcal{B}(\mathbb{R}\setminus\{0\}).$$

The measure v is called the Lévy measure of the process X. This measure is not a probability measure and is not necessarily finite. The measure v counts the expected number of jumps per unit of time (Applebaum, 2009, p. 87).

We have seen that a Lévy process can have an infinite number of jumps so the random measure J can take infinite values. Remember that a Lévy process can only take a finite number of large jumps. Because of that we define $H \in \mathcal{B}(\mathbb{R}\setminus\{0\})$ to be bounded below if $0 \notin \bar{H}$, where $\bar{H}$ represents the closure of H. If H is bounded below, then the random measure $J(.,H)$ takes only finite values (Applebaum, 2009, Lemma 2.3.4). When the set H is bounded below the jump measure, J satisfies certain interesting properties.

Theorem 1.1.3. *(Pascucci, 2011, Lemma 13.33)*
Let X be a Lévy process with jump measure J and Lévy measure v. Then

1. *If $H \in \mathcal{B}(\mathbb{R})$ is bounded below, then the process $(J(t,H))_{t\geq 0}$ defined as*

$$J(t,H) = \#\left\{s \in [0,t] : \Delta X(s) \in H\right\} \text{ for } t \geq 0,$$

is a Poisson process with intensity $v(H)$ and the compensated process $(\tilde{J}(t,H))_{t\geq 0}$ defined as

$$\tilde{J}(t,H) = J(t,H) - tv(H) \text{ for } t \geq 0,$$

is a martingale.

2. *If $H \in \mathcal{B}(\mathbb{R})$ is bounded below and f is a measurable function then the process $(J_t(f,H))_{t\geq 0}$ defined as*

$$J_t(f,H) = \int_0^t \int_H f(s.x)J(ds,dx) = \sum_{0\leq s\leq t} f(s,\Delta X(s))\mathbb{1}_H(\Delta X(s)) \text{ for } t \geq 0,$$

is a compound Poisson process.

It is possible to show that Lévy processes can be expressed as the sum of a drift term, a Brownian motion, a compound Poisson process and a sequence of compensated compound Poisson processes.

Theorem 1.1.4. *(Cont & Tankov, 2004, Proposition 3.7) Lévy-Itô decomposition*
Let X be a Lévy process with jump measure J and Lévy measure v. Then the Lévy measure v satisfies

$$\begin{aligned}\int_{|x|\geq 1} v(dx) &< \infty,\\ \int_{|x|<1} |x|^2 v(dx) &< \infty.\end{aligned}$$

Moreover there exists a Brownian motion B and two constants $\sigma \geq 0$ and $\gamma \in \mathbb{R}$ such that:

$$X(t) = \gamma t + \sigma B(t) + X^l(t) + \lim_{\epsilon\downarrow 0} \tilde{X}^\epsilon(t), \tag{1.1.1}$$

where

$$\begin{aligned} X^l(t) &= \int_0^t \int_{|x|\geq 1} xJ(ds,dx) \text{ is a compound Poisson process,} \\ \tilde{X}^\epsilon(t) &= \int_0^t \int_{\epsilon\leq|x|<1} x\tilde{J}(ds,dx) \\ &= \int_0^t \int_{\epsilon\leq|x|<1} x\left(J(ds,dx) - v(dx)ds\right) \\ &\quad \text{is a compensated compound Poisson process.} \end{aligned}$$

The terms in (1.1.1) are independent and the convergence of the last term is almost sure and uniform in t on $[0,T]$.

Definition 1.1.3. *(Cont & Tankov, 2004, p. 80) The triplet (γ,σ^2,v) in Theorem 1.1.4 is called the characteristic triplet or the Lévy triplet of the Lévy process X.*

1.1.3 Lévy-Khintchine representation

From the Lévy triplet of a Lévy process it is possible to obtain an analytical formula for its characteristic function. This formula is called the Lévy-Khintchine representation.

Theorem 1.1.5. *(Sato, 1999, p. 37) Lévy-Khintchine representation*
The characteristic exponent of a Lévy process X with Lévy triplet (γ,σ^2,v) is

$$\Psi^X(u) = i\gamma u - \frac{1}{2}\sigma^2u^2 + \int_{\mathbb{R}} \left(e^{iux} - 1 - iux\mathbb{1}_{\{|x|\leq 1\}}\right) v(dx),$$

where $\gamma \in \mathbb{R}$, $\sigma > 0$ and v is a Lévy measure that satisfies $\int_{\mathbb{R}}(x^2 \wedge 1)v(dx) < \infty$.

As we have seen before, the Lévy process is determined by its Lévy triplet. Now let us check two examples that are used in Chapter 4 and in Chapter 5.

Example 1.1.2. *Let us consider X to be a compound Poisson process, then X has characteristic triplet $(\gamma,0,v)$ with*

$$\begin{aligned} v(dx) &= \lambda F(dx), \\ \gamma &= \int_{|x|\leq 1} xv(dx), \end{aligned}$$

where $\lambda > 0$ represents the intensity of the Poisson process and F is the distribution of the jump. The characteristic exponent of X can be written as

$$\Psi^X(u) = \lambda \int_{-\infty}^{\infty} \left(e^{iux} - 1\right) F(dx) \text{ for } u \in \mathbb{R}$$

(Cont & Tankov, 2004, Proposition 3.4, Proposition 3.8).

Example 1.1.3. *Let X be a Lévy process of finite variation with Lévy triplet given by $(\gamma,0,v)$. By Cont and Tankov (2004, Corollary 3.1) we have that X can be expressed as the sum of its jumps between 0 and t and a linear drift term:*

$$X(t) = bt + \int_0^t \int_{\mathbb{R}} xJ(dx,dx) = bt + \sum_{0<s\leq t}^{\Delta X(s)\neq 0} \Delta X(s)$$

and its characteristic function can be written as

$$E\left[e^{iuX(t)}\right] = \exp\left\{t\left(ibu + \int_{\mathbb{R}} \left(e^{iux} - 1\right) v(dx)\right)\right\},$$

where $b = \gamma - \int_{|x|\leq 1} xv(dx)$.

Example 1.1.4. *(Schoutens, 2003, p. 53)*

A well known Lévy process is the inverse Gaussian Lévy process. This Lévy process is related to the inverse Gaussian distribution with parameters $a, b > 0$. *The characteristic function of the inverse Gaussian distribution is:*

$$\Phi_{IG}(u) = \exp\left\{-a\left(\sqrt{-2iu + b^2} - b\right)\right\} \text{ for } u \in \mathbb{R},$$

and the density function is

$$f_{IG}(x) = \frac{a}{\sqrt{2\pi}} \exp\{ab\} x^{-3/2} \exp\left\{-\frac{1}{2}\left(a^2 x^{-1} + b^2 x\right)\right\} \mathbb{1}_{(0,\infty)}(x) \text{ for } x \in \mathbb{R}.$$

The inverse Gaussian distribution is an infinitely divisible distribution, hence by Proposition 1.1.1 there exists a Levy process X *such that* $X(1)$ *follows an inverse Gaussian distribution. So we have that the characteristic function of* $X(t)$ *is:*

$$\Phi^{X(t)}(u) = E\left[e^{iuX(t)}\right] = \exp\left\{-at\left(\sqrt{-2iu + b^2} - b\right)\right\} \text{ for } u \in \mathbb{R} \text{ and } t \geq 0.$$

The inverse Gaussian Lévy process X *has Lévy triplet* $(\gamma, 0, v)$ *with:*

$$\begin{aligned} \gamma &= \frac{a}{b}(2F_N(b|0,1) - 1), \\ v(dx) &= (2\pi)^{-1/2} a x^{-3/2} \exp\left\{\frac{-1}{2} b^2 x\right\} \mathbb{1}_{(0,\infty)}(x) dx, \end{aligned}$$

where $F_N(.|0,1)$ *is the cumulative distribution function of a standard normal random variable.*

1.1.4 Martingale properties

This section is devoted to analysing when a Lévy process or an exponential Lévy process is a martingale. But first let us check when the exponential moment of a Lévy process is finite.

Proposition 1.1.2. *(Pascucci, 2011, Proposition 13.49) Let* X *be a Lévy process on* $\mathbb{R}$ *with characteristic triplet* (γ, σ^2, v). *The exponential moment* $E\left[e^{uX(t)}\right]$, $u \in \mathbb{R}$ *is finite iff*

$$\int_{|x|\geq 1} e^{ux} v(dx) < \infty.$$

In this case

$$E\left[e^{uX(t)}\right] = e^{t\Psi^X(-iu)},$$

where Ψ^X *is the characteristic exponent of* X.

Processes that are martingales play a central role in finance because options are priced when the discounted stock price is a martingale. The next result tells us when a Lévy process is a martingale, but also when an exponential Lévy process is a martingale. This result is important because it allows us to verify if we are in an equivalent risk-neutral measure when the stock price is modelled by an exponential Lévy process.

Theorem 1.1.6. *(Pascucci, 2011, Theorem 13.50) Let X be a real valued Lévy process with characteristic triplet (μ, σ, v), we have:*

1. *If $E[|X(1)|] < \infty$ then $(X(t) - E[X(t)])_{t\geq 0}$ is a martingale.*
2. *If $E\left[e^{uX(t)}\right] < \infty$ for some $u \in \mathbb{R}$ then*
$$\left(\frac{e^{uX(t)}}{E[e^{uX(t)}]}\right)_{t\geq 0}$$
is a martingale.

1.1.5 Change of measure for Lévy process

If we consider a Lévy process X living in a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, we will be interested in probability measures $\mathbb{Q}$ such that they are equivalent to the probability measure $\mathbb{P}$ and that X under $\mathbb{Q}$ is still a Lévy process but with a different characteristic triplet.

Proposition 1.1.3. *(Cont & Tankov, 2004, Proposition 9.8)(Pascucci, 2011, Theorem 13.50)*

Let X be a Lévy process with triplet (γ, σ^2, v) under $\mathbb{P}$. Then the following two conditions are equivalent:

1. *There is a probability measure $\mathbb{Q}$, equivalent to $\mathbb{P}$, such that $X = (X(t))_{t\in[0,T]}$ is a Lévy process with triplet $(\tilde{\gamma}, \tilde{\sigma}, \tilde{v})$ under $\mathbb{Q}$;*
2. *all of the following conditions hold:*
   (a) *$\tilde{v}(dx) = H(x)v(dx)$ for some Borel function $H : \mathbb{R} \to (0, \infty)$;*
   (b) *$\tilde{\sigma} = \sigma$;*
   (c) *$\tilde{\gamma} = \gamma + \int_{|x|\leq 1} x(H(x) - 1)v(dx) + \sigma^2\eta$ for some $\eta \in \mathbb{R}$;*
   (d) *$\int_{\mathbb{R}} \left(1 - \sqrt{H(x)}\right)^2 v(dx) < \infty$.*

In addition, if the measures $\mathbb{P}$ and $\mathbb{Q}$ are equivalent then we have that $\mathbb{Q}$ can be expressed as
$$\mathbb{Q}(A) = \int_A Z(T)d\mathbb{P} \text{ for } A \in \mathcal{F},$$
where the process Z is defined for $t \geq 0$ as
$$\begin{aligned} Z(t) &= \exp\left\{\eta X^c(t) - \frac{\eta^2\sigma^2 t}{2} - \eta\gamma t\right\} \\ &\quad \exp\left\{\lim_{\epsilon\downarrow 0}\left(\sum_{\substack{s\leq t \\ |\Delta X(s)|>\epsilon}} \log H(\Delta X(s)) - t\int_{|x|>\epsilon}(H(x)-1)v(dx)\right)\right\}, \end{aligned}$$
with the process X^c being the continuous part of the process X.

Notice that if the Lévy process has a diffusion part we can freely change the drift, but the volatility σ has to remain the same. Let us now see an example of how this change of measure works.

Example 1.1.5. *Let X be a compound Poisson process with Lévy triplet $(\gamma, 0, v)$ where*

$$\begin{aligned} v(dx) &= \lambda f(x)\mathbb{1}_D(x)dx, \\ \gamma &= \int_{|x|\leq 1} xv(dx), \end{aligned}$$

with $\lambda > 0$ and f is a density function with support D. We would like to show that there exists an equivalent measure $\mathbb{Q}$ under which the process X is a Lévy process with triplet $(\tilde{\gamma}, 0, \tilde{v})$ where

$$\begin{aligned} \tilde{v}(dx) &= \tilde{\lambda}\tilde{f}(x)\mathbb{1}_D(x)dx, \\ \tilde{\gamma} &= \int_{|x|\leq 1} x\tilde{v}(dx), \end{aligned}$$

with $\tilde{\lambda} > 0$ and $\tilde{f}$ is a density function with the same support D.

To this end, let us define the function H as

$$H(x) = \frac{\tilde{\lambda}\tilde{f}(x)}{\lambda f(x)}\mathbb{1}_D(x).$$

Notice that with this selection of H we have that

$$\tilde{v}(dx) = H(x)v(dx).$$

For applying Proposition 1.1.3 we need to show that $\int_{\mathbb{R}} \left(1 - \sqrt{H(x)}\right)^2 v(dx) < \infty$. Using the fact that $(a-b)^2 \leq a^2 + b^2$ for $a, b \geq 0$ we can write:

$$\begin{aligned} \int_{\mathbb{R}} \left(1 - \sqrt{H(x)}\right)^2 v(dx) &\leq \int_{\mathbb{R}} (1 + H(x))v(dx) \\ &= \int_D \lambda f(x)dx + \int_D \tilde{\lambda}\tilde{f}(x)dx \\ &= \lambda + \tilde{\lambda} < \infty. \end{aligned}$$

Lastly, notice that:

$$\begin{aligned} \tilde{\gamma} &= \int_{|x|<1} x(H(x) - 1)v(dx) + \gamma \\ &= \int_{|x|\leq 1} x\tilde{v}(dx) - \int_{|x|\leq 1} xv(dx) + \int_{|x|\leq 1} xv(dx) \\ &= \int_{|x|\leq 1} x\tilde{v}(dx). \end{aligned}$$

So we can apply Proposition 1.1.3 and obtain the desired equivalent probability measure.

1.2 Change of Brownian motion

Let $\left(\Omega, \mathcal{F}, \mathbb{P}, (\mathcal{F}_t)_{t\geq 0}\right)$ be a filtered probability space, and let W be a $\mathcal{F}_t$-Brownian motion. This section is concerned with the computation of the following Itô integral

$$\int_\tau^T \theta(s-\tau)dW(s) \text{ for } T > \tau > 0,$$

where $(\theta(t-\tau))_{t\geq 0}$ is a left-continuous adapted process with $\theta(t) = \phi(t)$ when $t \in [-\tau, 0]$, and where $\phi : [-\tau, 0] \to \mathbb{R}$ is a left-continuous deterministic function.

Theorem 1.2.1. *Let $T, \tau \in (0, \infty)$ with $T > \tau$ and consider the left-continuous and adapted process $(\theta(t-\tau))_{t\geq 0}$ with $\theta(t) = \phi(t)$ when $t \in [-\tau, 0]$, and where $\phi : [-\tau, 0] \to \mathbb{R}$ is a left-continuous deterministic function.*

If θ satisfies

$$\mathbb{P}\left(\left\{\int_0^T \theta^2(s-\tau)ds < \infty\right\}\right) = 1, \tag{1.2.1}$$

then

$$\int_\tau^T \theta(s-\tau)dW(s) = \int_0^{T-\tau} \theta(s)dB(s) \text{ almost surely,}$$

where B is the Brownian motion defined as

$$B(t) = W(t+\tau) - W(\tau) \text{ for } t \geq 0.$$

Proof. Define $X(T)$ as

$$\begin{aligned} X(T) &= \int_\tau^T \theta(s-\tau)dW(s) \\ &= \int_\tau^T \theta(s \vee \tau - \tau)dW(s) \\ &= \int_0^T \mathbb{1}_{(\tau,T]}(s)\theta(s \vee \tau - \tau)dW(s), \end{aligned}$$

where $x \vee y = \max\{x, y\}$ and the last equality comes from Klebaner (2012, Theorem 4.3.2). Also, let us define $Y(T)$ as

$$Y(T) = \int_0^{T-\tau} \theta(s)dB(s).$$

Let us take sequences of partitions $0 = t_0^{(n)} < t_1^{(n)} < ... < t_n^{(n)} = T - \tau$ of $[0, T-\tau]$ such that $\max_{1\leq k\leq n}\left\{t_k^{(n)} - t_{k-1}^{(n)}\right\} \to 0$ as $n \to \infty$. Because θ is left-continuous we can approximate θ by:

$$\theta^{(n)}(s) = \sum_{k=1}^n \theta\left(t_{k-1}^{(n)}\right) \mathbb{1}_{\left(t_{k-1}^{(n)}, t_k^{(n)}\right]}(s),$$

(Applebaum, 2009, Lemma 4.3.1). The Itô integral of $\theta^{(n)}$ is

$$Y^{(n)}(T) = \sum_{k=1}^n \theta\left(t_{k-1}^{(n)}\right)\left(B\left(t_k^{(n)}\right) - B\left(t_{k-1}^{(n)}\right)\right).$$

Because θ satisfies (1.2.1) then $Y^{(n)}(T)$ converges to $Y(T)$ in probability (Applebaum, 2009, p. 206).

The random variable $Y^{(n)}$ can be written as

$$\begin{aligned} Y^{(n)}(T) &= \sum_{k=1}^{n} \theta\left(t_{k-1}^{(n)}\right)\left(B\left(t_k^{(n)}\right) - B\left(t_{k-1}^{(n)}\right)\right) \\ &= \sum_{k=1}^{n} \theta\left(t_{k-1}^{(n)}\right)\left(W\left(t_k^{(n)} + \tau\right) - W\left(t_{k-1}^{(n)} + \tau\right)\right). \end{aligned}$$

Define $s_i^{(n)} = t_k^{(n)} + \tau$, $\forall k = 0, \ldots, n$ then we have $\tau = s_0^{(n)} < s_1^{(n)} < \ldots < s_n^{(n)} = T$. Notice that $\left(s_i^{(n)}\right)_{i=0}^{n}$ is a sequence of partitions of the interval $[\tau, T]$ such that $\max_{1\le k\le n}\left\{s_k^{(n)} - s_{k-1}^{(n)}\right\} \to 0$ as $n \to \infty$. Then we can write

$$\begin{aligned} Y^{(n)}(T) &= \sum_{k=1}^{n} \theta\left(t_{k-1}^{(n)}\right)\left(W\left(t_k^{(n)} + \tau\right) - W\left(t_{k-1}^{(n)} + \tau\right)\right) \\ &= \sum_{k=1}^{n} \theta\left(s_{k-1}^{(n)} - \tau\right)\left(W\left(s_k^{(n)}\right) - W\left(s_{k-1}^{(n)}\right)\right) \\ &= \sum_{k=1}^{n} \mathbb{1}_{(\tau,T]}\left(s_{k-1}^{(n)}\right)\theta\left(s_{k-1}^{(n)} \vee \tau - \tau\right)\left(W\left(s_k^{(n)}\right) - W\left(s_{k-1}^{(n)}\right)\right) \\ &\quad + \sum_{l=1}^{n} \mathbb{1}_{(\tau,T]}\left(r_{l-1}^{(n)}\right)\theta\left(r_{l-1}^{(n)} \vee \tau - \tau\right)\left(W\left(r_l^{(n)}\right) - W\left(r_{l-1}^{(n)}\right)\right), \end{aligned}$$

where $0 = r_0^{(n)} < r_1^{(n)} < \ldots < r_n^{(n)} = \tau$ is a sequence of partitions of $[0, \tau]$ such that $\max_{1\le l\le n}\left\{r_l^{(n)} - r_{l-1}^{(n)}\right\} \to 0$ as $n \to \infty$. Now, let us define

$$\tilde{s}_k^{(2n)} = \begin{cases} r_k^{(n)} & \text{if } 0 \le k \le n \\ s_{k-n}^{(n)} & \text{if } n+1 \le k \le 2n. \end{cases}$$

for $k = 0, 1 \ldots, 2n$. Notice that the sequence $\left(\tilde{s}_k^{(2n)}\right)_{k=0}^{2n}$ is a sequence of partitions of $[0, T]$ such that $\max_{1\le k\le 2n}\left\{\tilde{s}_k^{(n)} - \tilde{s}_{k-1}^{(n)}\right\} \to 0$ as $n \to \infty$. If we define $X_n(T)$ as

$$X_n(T) = \sum_{k=0}^{2n} \mathbb{1}_{(\tau,T]}\left(\tilde{s}_{k-1}^{(2n)}\right)\theta\left(\tilde{s}_{k-1}^{(2n)} \vee \tau - \tau\right)\left(W\left(\tilde{s}_k^{(2n)}\right) - W\left(\tilde{s}_{k-1}^{(2n)}\right)\right) = Y_n(T).$$

We know that $X_n(T)$ converges in probability to $X(T)$ (Applebaum, 2009, p.206). Because $X_n(T)$ converges in probability to $X(T)$ and also converges in probability to $Y(T)$, then $X(T) = Y(T)$ almost surely (Sokol & Rønn-Nielsen, 2013, Lemma 1.2.5). □

1.3 Conditional characteristic functions

In this subsection we present the concept of the conditional characteristic function and prove a crucial result that will allow us to do the change of measure and price options.

Here we consider a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and we would like to compute the conditional characteristic function of random variables given certain sub-σ-algebras of $\mathcal{F}$.

Definition 1.3.1. *(Yuan & Lei, 2016, p. 3707) Let X be a random variable and let $\mathcal{G} \subset \mathcal{F}$ be a sub-σ-algebra. The conditional characteristic function of X given $\mathcal{G}$ ($\mathcal{G}$-characteristic function) is:*

$$\Phi^{X|\mathcal{G}}(\lambda) = E\left[e^{i\lambda X}|\mathcal{G}\right] = \int_{\mathbb{R}} e^{i\lambda x} dF^{X|\mathcal{G}}(x|\mathcal{G}) \text{ for all } \lambda \in \mathbb{R},$$

where $F^{X|\mathcal{G}}(x|\mathcal{G})$ defined on $\mathbb{R} \times \Omega$ is the conditional distribution function of X given $\mathcal{G}$, that is $F^{X|\mathcal{G}}(x|\mathcal{G}) = \mathbb{P}(X \leq x|\mathcal{G})$.

The conditional characteristic function identifies the distribution of a random variable given a sub-σ-algebra $\mathcal{G} \subset \mathcal{F}$.

Theorem 1.3.1. *(Yuan & Lei, 2016, p. 3717)*
Let X and Y be two random variables with respective $\mathcal{G}$-characteristic functions $\Phi^{X|\mathcal{G}}(\lambda)$ and $\Phi^{Y|\mathcal{G}}(\lambda)$. Then $\Phi^{X|\mathcal{G}}(\lambda) = \Phi^{Y|\mathcal{G}}(\lambda)$ almost surely if and only if X and Y are $\mathcal{G}$-identically distributed.

Assume that our probability space is equipped with a filtration $(\mathcal{F}_t)_{t\geq 0}$. Let $(W(t))_{t\geq 0}$ be a one-dimensional Brownian motion, and let $\left(\mathcal{F}_t^W\right)_{t\geq 0}$ be the filtration generated by this Brownian motion and $\mathcal{F}_t^W \subset \mathcal{F}_t$ for all $t \geq 0$. Define the filtration $(\mathcal{G}_t)_{t\geq 0}$ that is independent of $\left(\mathcal{F}_t^W\right)_{t\geq 0}$ and $\mathcal{G}_t \subset \mathcal{F}_t$ for all $t \geq 0$. Using the definition of conditional characteristic function (Definition 1.3.1) and Theorem 1.3.1, we will show that the Itô integral with respect to W of a left-continuous and $(\mathcal{G}_t)_{t\geq 0}$-adapted process follows a Normal distribution under the sub-σ-algebra $\mathcal{G}_t$. First let us prove the following Lemma.

Lemma 1.3.1. *Let $(\theta(t))_{t\geq 0}$ be a random step process such that it is independent of $\left(\mathcal{F}_t^W\right)_{t\geq 0}$ and it is adapted with respect to the filtration $(\mathcal{G}_t)_{t\geq 0}$. Define $X_t := \int_0^t \theta(s)dW(s)$ hence we have that*

$$\Phi^{X_t|\mathcal{G}_t}(\lambda) = \exp\left\{-\lambda^2 \int_0^t \frac{1}{2}\theta^2(s)ds\right\} \ \forall \lambda \in \mathbb{R}.$$

Proof. Because θ is a random step process then exists a partition $0 = t_0 < t_1 < t_2 < \ldots < t_l = t$ of $[0, t]$ such that:

$$\theta(s) = \sum_{k=1}^{l} \mathbb{1}_{(t_{k-1},t_k]}(s)\theta_k,$$

where θ_k is $\mathcal{G}_{t_{k-1}}$-measurable for all $k = 1, \ldots, l$. Define $\Delta W(k) = W(t_k) - W(t_{k-1})$ and $\Delta t_k = t_k - t_{k-1}$, then the Itô integral of this random step process is

$$\int_0^t \theta(s)dW(s) = \sum_{k=1}^{l} \theta_k \Delta W(k).$$

For $k = 1, 2, \ldots, l-1$ define $\mathcal{G}_t \vee \mathcal{F}_{t_k}^W$ to be the smallest σ-algebra containing $\mathcal{G}_t$ and $\mathcal{F}_{t_k}^W$. The conditional characteristic function of $\sum_{k=1}^{l} \theta_k \Delta W(k)$ given $\mathcal{G}_t$ can be expressed as:

$$
\begin{aligned}
&E\left[\exp\left\{i\lambda\sum_{k=1}^{l}\theta_k\Delta W(k)\right\}\middle|\,\mathcal{G}_t\right]\\
&\qquad\qquad = E\left[\prod_{k=1}^{l}\exp\left\{i\lambda\theta_k\Delta W(k)\right\}\middle|\,\mathcal{G}_t\right]\\
&\qquad\qquad = E\left[E\left[\prod_{k=1}^{l}\exp\{i\lambda\theta_k\Delta W(k)\}\middle|\,\mathcal{G}_t\vee\mathcal{F}^W_{t_1}\right]\middle|\,\mathcal{G}_t\right]\\
&\qquad\qquad = E\left[exp\{i\lambda\theta_1\Delta W(1)\}E\left[\prod_{k=2}^{l}\exp\{i\lambda\theta_k\Delta W(k)\}\middle|\,\mathcal{G}_t\vee\mathcal{F}^W_{t_1}\right]\middle|\,\mathcal{G}_t\right].
\end{aligned}
$$

Repeating $l-2$ times, we obtain the following.

$$
\begin{aligned}
&E\left[\exp\{i\lambda\sum_{k=1}^{l}\theta_k\Delta W(k)\}\middle|\,\mathcal{G}_t\right]=\\
&E\left[\exp\{i\lambda\theta_1\Delta W(1)\}E\left[\ \dots\ E\left[exp\{i\lambda\theta_l\Delta W(l)\}|\,\mathcal{G}_t\vee\mathcal{F}^W_{t_{l-1}}\right]\ \dots\ \middle|\,\mathcal{G}_t\vee\mathcal{F}^W_{t_1}\right]\middle|\,\mathcal{G}_t\right].
\end{aligned}
$$

Because θ_l is $\mathcal{G}_t\vee\mathcal{F}^W_{t_{l-1}}$-measurable and $\Delta W(l)$ is independent of $\mathcal{G}_t\vee\mathcal{F}^W_{t_{l-1}}$, we have that

$$
E\left[\exp\left\{i\lambda\theta_l\Delta W(l)\right\}|\,\mathcal{G}_t\vee\mathcal{F}^W_{l-1}\right] \quad = \quad \exp\left\{\frac{-1}{2}\lambda^2\theta_l^2\Delta t_l\right\},
$$

(Kopp, Malczak, & Zastawniak, 2014, Theorem 4.27). Notice that $\exp\left\{\frac{-1}{2}\lambda\theta_l^2\Delta t_l\right\}$ is $\mathcal{G}_t$-measurable, so it can be extracted outside the expectation.

Repeating this process $l-1$ times:

$$
\begin{aligned}
E\left[\exp\left\{i\lambda\sum_{k=1}^{l}\theta_k\Delta W(k)\right\}\middle|\,\mathcal{G}_t\right] \quad &= \quad \prod_{k=1}^{l}\exp\left\{-\lambda^2\frac{1}{2}\theta_k^2\Delta t_k\right\}\\
&= \quad \exp\left\{\sum_{k=1}^{l}-\lambda^2\frac{1}{2}\theta_k^2\Delta t_k\right\}.
\end{aligned}
$$

□

Using previous Lemma, we can generalise the result to left-continuous processes.

Theorem 1.3.2. *Let $(\theta(t))_{t\geq 0}$ be a left-continuous process such that*

$$
\mathbb{P}\left(\int_0^t\theta^2(s)ds<\infty\right)=1.
$$

This process is independent of $\left(\mathcal{F}^W_t\right)_{t\geq 0}$ and it is adapted with respect to the filtration $(\mathcal{G}_t)_{t\geq 0}$. Define the random variable $X_t:=\int_0^t\theta(s)dW(s)$ hence we have that

$$
\Phi^{X_t|\mathcal{G}_t}(\lambda)=\exp\left\{-\lambda^2\int_0^t\frac{1}{2}\theta^2(s)ds\right\},\ \forall\lambda\in\mathbb{R}.
$$

Proof. The process θ can be approximated by a sequence of random step processes

$$\theta^{(l)}(s) = \sum_{k=1}^{l} \mathbb{1}_{\left(t_{k-1}^{(l)}, t_k^{(l)}\right]}(s)\theta\left(t_{k-1}^{(l)}\right),$$

where $l = 1, 2, \ldots$ and $0 = t_0^{(l)} < t_1^{(1)} < \ldots < t_l^{(l)} = t$ is a sequence of partitions of $[0, t]$ such that $\max_i \left\{t_i^{(l)} - t_{i-1}^{(l)}\right\} \to 0$ as $l \to \infty$ (Applebaum, 2009, Lemma 4.3.1). Define $\Delta W\left(k^{(l)}\right) = W\left(t_k^{(l)}\right) - W\left(t_{k-1}^{(l)}\right)$ and $\Delta t_k^{(l)} = t_k^{(l)} - t_{k-1}^{(l)}$. We know that

$$\sum_{k=0}^{l-1} \theta\left(t_{k-1}^{(l)}\right) \Delta W(k^{(l)}) \to \int_0^t \theta(s) dW(s) \text{ in probability as } l \to \infty,$$

(Applebaum, 2009, p.206). Convergence in probability implies the existence of a subsequence that converges almost surely (Klebaner, 2012, p. 38). Define the following sequence $(X_l)_{l=1}^{\infty}$ where $X_l = \sum_{k=1}^{l} \theta\left(t_{k-1}^{(l)}\right) \Delta W\left(k^{(l)}\right)$. Now let us consider the subsequence $(Y_h)_{h=1}^{\infty}$ where $Y_h = X_{l_h}$ and $l_1 < l_2 < \ldots$ is an increasing sequence of indices, such that:

$$Y_h = X_{l_h} = \sum_{k=1}^{l_h} \theta\left(t_{k-1}^{(l_h)}\right) \Delta W\left(k^{(l_h)}\right) \to \int_0^t \theta(s) dW(s) \text{ almost surely as } h \to \infty\ .$$

Because the convergence is almost surely, and by continuity of $x \to e^{i\lambda x}$, then we know that:

$$\exp\left\{i\lambda \sum_{k=1}^{l_h} \theta\left(t_{k-1}^{(l_h)}\right) \Delta W\left(k^{(l_h)}\right)\right\} \to \exp\left\{i\lambda \int_0^t \theta(s) dW(s)\right\} \text{ as } h \to \infty$$

almost surely. Applying the Conditional Dominated Convergence Theorem (Williams, 1991, Theorem 9.7), we have that

$$\lim_{h\to\infty} E\left[\exp\left\{i\lambda \sum_{k=1}^{l_h} \theta\left(t_{k-1}^{(l_h)}\right) \Delta W\left(k^{(l_h)}\right)\right\} \middle| \mathcal{G}_t\right] = E\left[\exp\left\{i\lambda \int_0^t \theta(s) dW(s)\right\} \middle| \mathcal{G}_t\right]$$

almost surely. By Lemma 1.3.1 we obtain that

$$E\left[\exp\left\{i\lambda \sum_{k=1}^{l_h} \theta\left(t_{k-1}^{(l_h)}\right) \Delta W\left(k^{(l_h)}\right)\right\} \middle| \mathcal{G}_t\right] = \exp\left\{-\lambda^2 \sum_{k=1}^{l_h} \frac{1}{2}\theta^2\left(t_{k-1}^{(l_h)}\right) \Delta t_k^{(l_h)}\right\}.$$

At the end we have:

$$\lim_{h\to\infty} E\left[\exp\left\{i\lambda\sum_{k=1}^{l_h}\theta\left(t_{k-1}^{(l_h)}\right)\Delta W\left(k^{(l_h)}\right)\right\}\middle|\mathcal{G}_t\right]$$
$$= \lim_{h\to\infty}\exp\left\{-\lambda^2\sum_{k=1}^{l_h}\frac{1}{2}\theta^2\left(t_{k-1}^{(l_h)}\right)\Delta t_k^{(l_h)}\right\}$$
$$= \exp\left(\int_0^t -\lambda^2\frac{1}{2}\theta^2(s)ds\right) \text{ almost surely.}$$

The last equality comes from the left-continuous paths of θ and from the definition of the Lebesgue integral. □

1.4 Martingale property

This section is devoted to summarising some results that will help us to prove when a stochastic exponential is a martingale. In this case, let $\left(\Omega, \mathcal{F}, \mathbb{P}, (\mathcal{F}_t)_{t\geq 0}\right)$ be a filtered probability space, let W be a $\mathcal{F}_t$-Brownian motion, define $\left(\mathcal{F}_t^W\right)_{t\geq 0}$ to be the filtration generated by W and assume that $\mathcal{F}_t^W \subset \mathcal{F}_t$ for all $t \geq 0$. Also, consider the filtration $(\mathcal{G}_t)_{t\geq 0}$ such that $\mathcal{G}_t \subset \mathcal{F}_t$ for all $t \geq 0$ and it is independent of $\left(\mathcal{F}_t^W\right)_{t\geq 0}$. The result of Proposition 1.4.1 will allow us to compute the conditional distribution of an Itô integral with respect to the σ-algebra $\mathcal{G}_t$ under some restrictions.

Proposition 1.4.1. *Let $(\theta(t))_{t\geq 0}$ be a left-continuous process such that*

$$\mathbb{P}\left(\int_0^T \theta^2(s)ds < \infty\right) = 1$$

for all $T > 0$. If the process θ is independent of $\left(\mathcal{F}_t^W\right)_{t\geq 0}$ and it is adapted with respect to the filtration $(\mathcal{G}_t)_{t\geq 0}$, then

$$\int_0^T \theta(s)dW(s)\Big|\mathcal{G}_T \sim N\left(0, \int_0^T \theta^2(s)ds\right) \quad \text{where } 0 < T < \infty.$$

Proof. By Theorem 1.3.2 we know that the conditional characteristic function of $\int_0^t \theta(s)dW(s)$ with respect to $\mathcal{G}_t$ is the same as the conditional characteristic function of a normal random variable with mean zero and variance $\int_0^t \theta^2(s)ds$. From Theorem 1.3.1, we have the desired result. □

Proposition 1.4.2. *Let $(\theta(t))_{t\geq 0}$ be a left-continuous process such that*

$$\mathbb{P}\left(\int_0^T \theta^2(s)ds < \infty\right) = 1$$

for all $T > 0$. If this process θ is independent of $\left(\mathcal{F}_t^W\right)_{t\geq 0}$ and it is adapted with respect to the filtration $(\mathcal{G}_t)_{t\geq 0}$, then

$$\begin{aligned}\int_u^T \theta(s)dW(s)\Big|\mathcal{G}_T &\sim N\left(0, \int_u^T \theta^2(s)ds\right),\\ \int_u^T \theta(s)dW(s)\Big|\mathcal{G}_T \vee \mathcal{F}_u^W &\sim N\left(0, \int_u^T \theta^2(s)ds\right),\end{aligned}$$

where $0 < u < T < \infty$.

Proof. From Theorem 1.2.1 we have that

$$\int_u^T \theta(s)dW(s) = \int_0^{T-u} \theta(s+u)dB(s),$$

where

$$B(s) = W(s+u) - W(u) \text{ for } s \geq 0.$$

Notice that $\theta(s+u)$ is measurable with respect to $\mathcal{G}_T$ when $s \in [0, T-u]$. By application of Proposition 1.4.1 we have

$$\int_0^{T-u} \theta(s+u)dB(s)\Big| \mathcal{G}_T \sim N\left(0, \int_0^{T-u} \theta^2(s+u)ds\right).$$

In addition, by the independence of increments of Brownian motion, B is independent of the σ-algebra $\mathcal{F}_u^W$. Then by Proposition 1.4.1 we have

$$\int_0^{T-u} \theta(s+u)dB(s)\Big| \mathcal{G}_T \vee \mathcal{F}_u^W \sim N\left(0, \int_0^{T-u} \theta^2(s+u)ds\right).$$

Finally, notice that

$$\int_0^{T-u} \theta^2(s+u)ds = \int_u^T \theta^2(s)ds.$$

□

The next result comes from the work of Mijatović and Urusov (2012). But first let us introduce the concept of local integrability.

Definition 1.4.1. *(Cherny & Engelbert, 2005, Definition 2.1) A measurable function $f : \mathbb{R} \to \mathbb{R}$ is locally integrable in a set $D \subseteq \mathbb{R}$, if $\forall x \in D$ $\exists \epsilon > 0$ such that*

$$\int_{x-\epsilon}^{x+\epsilon} |f(y)|dy \quad < \quad \infty.$$

Let us consider the state space $J = (l, r)$, where $-\infty \leq l < r \leq \infty$. Let Y be an J-valued diffusion governed by the stochastic differential equation

$$dY(t) = \mu(Y(t))dt + \sigma(Y(t))dW(t), \; Y(0) = y \in J,$$

where $\mu, \sigma : J \to \mathbb{R}$ are Borel functions satisfying the conditions

$$\sigma(x) \neq 0 \; \forall x \in J, \tag{1.4.1}$$

$$\frac{1}{\sigma^2}, \frac{\mu}{\sigma^2} \text{ are locally integrable functions in } J. \tag{1.4.2}$$

Define the stopping time ξ to be the exit time of Y from J. In our case, we are interested only in the case when $\mathbb{P}(\xi = \infty) = 1$. Consider the stochastic exponential

$$Z(t) = \exp\left\{\int_0^{t\wedge\xi} b(Y(u))dW(u) - \frac{1}{2}\int_0^{t\wedge\xi} b^2(Y(u))du\right\} \; t \in [0, \infty),$$

where $b : J \to \mathbb{R}$ and it satisfies the condition:

$$\frac{b^2}{\sigma^2} \text{ is locally integrable in } J. \tag{1.4.3}$$

The next result shows when the stochastic exponential Z is a martingale.

Theorem 1.4.1. *(Mijatović & Urusov, 2012, Corollary 2.2)*
Assume that Y does not exit J with probability 1 and the conditions (1.4.1),(1.4.2) and (1.4.3) are satisfied. Then Z is a martingale if and only if $\tilde{Y}$ does not exit J with probability 1, where $\tilde{Y}$ is the auxiliary process governed by the stochastic differential equation:

$$d\tilde{Y}(t) = \left(\mu(\tilde{Y}(t)) + b(\tilde{Y}(t))\sigma(\tilde{Y}(t))\right) dt + \sigma(\tilde{Y}(t))dW(t) \text{ with } \tilde{Y}(0) = y. \tag{1.4.4}$$

Now we would like to show how Theorem 1.4.1 can be applied to a particular example. Let us consider the stochastic differential equation

$$dX(t) = a(b - X(t))dt + \gamma\sqrt{X(t)}dW(t) \text{ with } X(0) = x \in (0, \infty), \tag{1.4.5}$$

where $b \in \mathbb{R}$, $a, \gamma > 0$ and $\frac{2ab}{\gamma^2} \geq 1$. We know that the equation (1.4.5) has a strong solution and X remains in the interval $(0, \infty)$ almost surely (Gulisashvili, 2012, p. 44, Theorem 2.27). Let us define the process Z' as follows.

$$\begin{aligned} Z'(t) &= \exp\left\{\int_0^t \left(-\frac{\lambda_1}{\gamma\sqrt{X(u)}} - \frac{\lambda_2}{\gamma}\sqrt{X(u)}\right) dW(u)\right\} \\ &\quad \exp\left\{-\frac{1}{2}\int_0^t \left(\frac{\lambda_1}{\gamma\sqrt{X(u)}} + \frac{\lambda_2}{\gamma}\sqrt{X(u)}\right)^2 du\right\} \text{ for } t \geq 0, \end{aligned} \tag{1.4.6}$$

where $\lambda_1, \lambda_2 \in \mathbb{R}$. Now we would like to prove that Z' is a martingale.

Proposition 1.4.3. *Consider a process X that satisfies the equation (1.4.5) and a process Z' that is defined in equation (1.4.6). Let us define $\tilde{a} = a + \lambda_2$ and $\tilde{b} = \frac{ab-\lambda_1}{a+\lambda_2}$. If the conditions*

$$\frac{2\tilde{a}\tilde{b}}{\gamma^2} \geq 1 and \tag{1.4.7}$$

$$\tilde{a} > 0 \tag{1.4.8}$$

are satisfied, then Z' is a martingale.

Proof. First, we are going to check that conditions (1.4.1),(1.4.2) and (1.4.3) are satisfied. Notice that the condition (1.4.1) is satisfied because $\gamma\sqrt{x} > 0$, $\forall x \in (0, \infty)$. For proving conditions (1.4.2) and (1.4.3) let us define the functions f, g and h as

$$f(x) = \frac{1}{\gamma^2 x}, \tag{1.4.9}$$

$$g(x) = \frac{ab - ax}{\gamma^2 x}, \tag{1.4.10}$$

$$h(x) = \frac{\left(\frac{\lambda_1}{\gamma\sqrt{x}} + \frac{\lambda_2}{\gamma}\sqrt{x}\right)^2}{\gamma^2 x} = \frac{\lambda_1^2}{\gamma^4 x^2} + \frac{\lambda_2^2}{\gamma^4} + \frac{2\lambda_1\lambda_2}{\gamma^4 x}, \tag{1.4.11}$$

for $x \in (0,\infty)$. Now we will show that the functions f,g and h are locally integrable in $(0,\infty)$. To that end, let us pick an $x \in (0,\infty)$ and choose an $\epsilon = \frac{|x|}{2}$. For the function f we have

$$\begin{aligned}\int_{x-\epsilon}^{x+\epsilon} \frac{1}{\gamma^2 y} dy &= \frac{1}{\gamma^2} \log(y)|_{x-\epsilon}^{x+\epsilon} \\ &= \frac{1}{\gamma^2} \log\left(\frac{x+\epsilon}{x-\epsilon}\right) < \infty.\end{aligned}$$

In the case of the function g we obtain

$$\begin{aligned}\int_{x-\epsilon}^{x+\epsilon} \left|\frac{ab-ay}{\gamma^2 y}\right| dy &\leq \frac{1}{\gamma^2}\left(ab\int_{x-\epsilon}^{x+\epsilon} \frac{1}{y}dy + a\int_{x-\epsilon}^{x+\epsilon} dy\right) \\ &= \frac{ab}{\gamma^2}\log\left(\frac{x+\epsilon}{x-\epsilon}\right) + \frac{2a}{\gamma^2}\epsilon < \infty.\end{aligned}$$

Lastly, for the function h, we can show that

$$\begin{aligned}&\int_{x-\epsilon}^{x+\epsilon} \left|\frac{\lambda_1^2}{\gamma^4 y^2} + \frac{\lambda_2^2}{\gamma^4} + \frac{2\lambda_1\lambda_2}{\gamma^4 y}\right| dy \\ &\leq \frac{1}{\gamma^4}\left(\lambda_1^2 \int_{x-\epsilon}^{x+\epsilon} \frac{1}{y^2}dy + \lambda_2^2\int_{x-\epsilon}^{x+\epsilon} dy + 2\lambda_1\lambda_2\int_{x-\epsilon}^{x+\epsilon}\frac{1}{y}dy\right) \\ &= -\left.\frac{\lambda_1^2}{\gamma^4 y}\right|_{x-\epsilon}^{x+\epsilon} + \frac{2\lambda_2^2}{\gamma^4}\epsilon + \frac{2\lambda_1\lambda_2}{\gamma^4}\log\left(\frac{x+\epsilon}{x-\epsilon}\right) \\ &= -\frac{\lambda_1^2}{\gamma^4(x+\epsilon)} + \frac{\lambda_1^2}{\gamma^4(x-\epsilon)} + \frac{2\lambda_2^2}{\gamma^4}\epsilon + \frac{2\lambda_1\lambda_2}{\gamma^4}\log\left(\frac{x+\epsilon}{x-\epsilon}\right) < \infty.\end{aligned}$$

We have just shown that the functions f, g and h are locally integrable in $(0,\infty)$. Notice that the auxiliary process $\tilde{X}$ satisfies the stochastic differential equation

$$d\tilde{X}(t) = \left(ab - a\tilde{X}(t) - \lambda_1 - \lambda_2\tilde{X}(t)\right)dt + \gamma\sqrt{\tilde{X}(t)}dW(t) \text{ with } \tilde{X}(0) = x \in (0,\infty). \tag{1.4.12}$$

Equation (1.4.12) can be rewritten as

$$d\tilde{X}(t) = \tilde{a}\left(\tilde{b} - \tilde{X}(t)\right)dt + \gamma\sqrt{\tilde{X}(t)}dW(t) \text{ with } \tilde{X}(0) = x \in (0,\infty), \tag{1.4.13}$$

where $\tilde{a} = a + \lambda_2$ and $\tilde{b} = \frac{ab-\lambda_1}{a+\lambda_2}$. Because conditions (1.4.7) and (1.4.8) are satisfied, we know that equation (1.4.13) has a strong solution and it does not exit the interval $(0,\infty)$ almost surely (Gulisashvili, 2012, p. 44, Theorem 2.27). So by application of Theorem 1.4.1 we know that Z' is a martingale. □

1.5 Maximum likelihood estimator

Suppose that we have a random variable Y and we have the observation y of this random variable. Assume that the density function f^Y of Y is known. The density function f^Y depends on the data y and on a vector of unknown parameters θ. Because of that, we express the density function as $f^Y(y|\theta)$. The parameter θ takes values in a parameter space Θ. The objective is to estimate the parameter θ based on the observation y.

The likelihood of θ based on y is defined as

$$L^Y(\theta) = f^Y(y|\theta) \text{ for } \theta \in \Theta$$

(Davison, 2003, p. 94). We can consider the likelihood as a function of θ for a fixed y. Usually, it is preferable to work with the log-likelihood that is defined as

$$l^Y(\theta) = \log(L^Y(\theta)).$$

The maximum likelihood estimator of θ, that we defined as $\hat{\theta}$, is the value of θ that maximizes the likelihood function, that is

$$\hat{\theta} = \arg\max_{\theta \in \Theta} L^Y(\theta).$$

Because the logarithm is a strictly increasing function, we have that

$$\hat{\theta} = \arg\max_{\theta \in \Theta} l^Y(\theta).$$

1.5.1 Conditional likelihood

In this case, the parameter θ that appears in the density function f^Y of the random variable Y can be expressed as

$$\theta = (\varphi, \lambda),$$

where φ is a $p \times 1$ vector of parameters of interest and λ is a $q \times 1$ vector of nuisance parameters. If we can find a sufficient statistic R_λ for the vector of nuisance parameters λ that is not a function of φ, then the conditional log-likelihood for the parameter φ is defined as

$$l^{Y|R_\lambda}(\varphi) = \log(f^{Y|R_\lambda}(y|R_\lambda)),$$

where $f^{Y|R_\lambda}(.|R_\lambda)$ is the conditional density function of the random variable Y given R_λ (Kalbfleisch & Sprott, 1970, p. 181). The vector of parameters φ can be estimated by maximising the function $l^{Y|R_\lambda}$,

$$\hat{\varphi} = \arg\max_{\varphi} l^{Y|R_\lambda}(\varphi).$$

Properties of conditional maximum likelihood estimates can be found in Andersen (1970).

1.6 Data

In this thesis, we will use some data to estimate and calibrate the parameters of the models presented. In this section, we will introduce the data that will be used in the next chapters.

As we shall see in Chapter 2 the data that are going to be used as proxies for the market attention are the number of Wikipedia views of the keyword 'Altcoin" and the number of unique active addresses. The data related to the number of views on Wikipedia are taken from https://pageviews.toolforge.org. The number of unique active addresses is taken from https://charts.coinmetrics.io/network-data/. The number of unique active addresses represents the number of active addresses that participate in successful transactions on the blockchain network, either as a sender or as a receiver. The frequency of market attention data is daily. The periods for which the data are taken are the following:

1. In Chapter 2 and Chapter 6 market attention data were taken from 2018-09-20 to 2019-09-20.

2. Market attention data in Chapter 3, Chapter 4 and Chapter 5 correspond to the period from 2018-04-01 to 2019-09-01.

For estimating the parameters of our models, we will also need the historical prices of Bitcoin in United States dollars. These are taken from https://charts.coinmetrics.io/network-data/ and consist of the closing price of Bitcoin. The Bitcoin price offered by Coinmetrics is taken when the New York market exchange closes (16:00 EST). Due to the fact that Bitcoin is traded on different exchanges, the price given by Coinmetrics is calculated using their own methodology using the most important exchanges [1]. The data again have daily frequency, and the periods for which the data are taken are the following:

1. In Chapter 6 market attention data were taken from 2018-09-19 to 2019-09-20.

2. Market attention data in Chapter 3, Chapter 4 and Chapter 5 correspond to the period from 2018-04-01 to 2019-09-01.

The Bitcoin historical data and the historical data of the market attention, is going to be used for the estimation of the parameters of the models presented in this thesis.

Apart from the historical data used to estimate the parameters of the proposed models, we will require option market data to compare the results given by our models with real market data. The option market data is taken from Tardis (https://docs.tardis.dev/historical-data-details/deribit). Tardis allows us to download the option market data of the exchange Deribit for the first day of each month for free. We take the option market data for the first day of each month, from April 2019 to September 2021. Deribit also offers perpetual options, but in our case, we are only interested in plain vanilla European call and put options. The underlying asset of the options offered by Deribit is a Bitcoin index, which is built using the prices of Bitcoin in United States dollars from different exchanges. [2] Tardis also offers the price in United States dollars of this index. We will use the value of this index as the initial value of the underlying for pricing options.

The ask price and bid price of the options are given in number of Bitcoins, not in United States dollars. Because of that, we need to convert the option prices from number of Bitcoins to United States dollars. We use the previously mentioned Deribit Bitcoin index to convert from the price in Bitcoins to the price in United States dollars. However, the value of the strikes are given in United States dollars. In this thesis, we assume that the value of the options are given in United States dollars. This approach has been followed by the literature previously cited. But Alexander and Imeraj (2021) proposed a method to price options in number of Bitcoins instead of pricing options in number of United States dollars.

Remark 1.6.1. *The data provided by Tardis are intraday data, but the bid and ask prices vary through the day. Because of that, for each day, we take the first ask and*

[1] For more information, see https://coinmetrics.io/reference-rates/ and https://docs.coinmetrics.io/methodologies/reference-rates/real-time-reference-rates-methodology.

[2] For a detailed explanation, see https://legacy.deribit.com/pages/docs/general and https://legacy.deribit.com/pages/docs/options.

bid prices of the option that appear in the order book starting at 00:00 UTC. For each option, we take the last value of the Deribit Bitcoin index that appears on the order book, before the date in which the option is taken.

Remark 1.6.2. *The bid-ask spread in Bitcoin options given in United States dollars can be quite high. Since we are using the mid price to evaluate our model, we only take the options that satisfy the following:*

$$\frac{AskPrice - BidPrice}{AskPrice} < 0.1.$$

In that way, the bid and ask prices are not too widely spread and it makes sense to compare the prices given by our model with the mid price.

Chapter 2

Modelling market attention

2.1 Introduction

As we have seen in Chapter 1, market attention seems to affect Bitcoin volatility. In this chapter, we are interested in modelling market attention as a stochastic process. Market attention or interest is always a non-negative quantity, and because of that we will use only non-negative stochastic processes. We will use continuous stochastic processes for modelling, since we would like to apply market attention to option pricing. Because of that we will use stochastic differential equations for modelling the interest.

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and let process $I = (I(t))_{t\geq 0}$ represents market attention. We assume that I is the strong solution of the following stochastic differential equation:

$$dI(t) = \alpha_I(I(t))dt + \beta_I(I(t))dX_I(t), \tag{2.1.1}$$

where X_I is a Lévy process and the functions $\alpha_I : \mathbb{R} \to \mathbb{R}$, $\beta_I : \mathbb{R} \to (0, \infty)$ are such that the equation (2.1.1) has a strong solution.

There are different possibilities for equation (2.1.1), but we impose the following requirements:

1. The strong solution of (2.1.1) has to be positive with probability 1.
2. The integrated interest process has to be analytically tractable. That is, it should be possible to compute the characteristic function of the random variable $\int_0^t I(s)ds$ for $t \geq 0$.

With these requirements in mind, we propose that the process I could be a Cox-Ingersoll-Ross process or a positive Ornstein-Uhlenbeck process.

The stochastic differential equation (2.1.1) depends on an unknown vector of parameters θ_I. In this chapter, we are concerned with the estimation of θ_I. We will assume that we have a discrete observed sample $\{y_j\}_{j=0}^N$ of $N+1$ observations of market attention. These observations are equally spaced in time with a time step $\Delta > 0$. Each observation y_j is assumed to be a realisation of the random variable $I(t_j)$ for $j = 0, , 1, \ldots, N$. We take the initial time t_0 to be $t_0 = 0$ and since the observations are equally spaced in time, we can write $t_j = j\Delta$ for $j = 0, 1, \ldots, N$. Using these observations , we will estimate the parameter θ_I.

Once we know how to estimate the vector of parameters θ_I, we would like to check that the observed sequence $\{y_j\}_{j=0}^N$ has been generated by the proposed process. That is, we would like to validate our model.

2.2 Cox-Ingersoll-Ross process

This process has been used in a financial context, to model short-term interest rates (Cox, Ingersoll, & Ross, 1985) and the volatility of stock prices (Heston, 1993). The Cox-Ingersoll-Ross process satisfies the following stochastic differential equation:

$$dI(t) = a_I(b_I - I(t))dt + \sigma_I \sqrt{I(t)} dW_I(t) \quad \text{with } I(0) = y_0, \tag{2.2.1}$$

where $W_I = (W_I(t))_{t \geq 0}$ is a Brownian motion, $b_I \in \mathbb{R}$, $\sigma_I, a_I, y_0 > 0$ and we assume that

$$\frac{2a_I b_I}{\sigma_I^2} \geq 1. \tag{2.2.2}$$

It has been proved that equation (2.2.1) has a unique strong solution (Gulisashvili, 2012, p. 44). Furthermore, because we have the assumption (2.2.2) we know that $I(t) > 0$ for all $t \geq 0$ almost surely (Gulisashvili, 2012, Theorem 2.27).

2.2.1 Estimation

Now we are concerned with the estimation of the parameters a_I, b_I and σ_I that appear in equation (2.2.1).

As we have said, we have a finite number of $N+1$ of observations $\{y_j\}_{j=0}^N$, equally spaced in time with time step Δ. For the estimation of the parameters of the Cox-Ingersoll-Ross process, we will use the maximum likelihood estimation method.

This method of estimation for the Cox-Ingersoll-Ross process is also used by Iacus (2009, p. 112) and by Kladívko (2007).

Because the Cox-Ingersoll-Ross process I is a strong solution of a stochastic differential equation, it has the Markov property (Karatzas & Shreve, 1998, Theorem 5.4.20). We assume $I(0)$ to be a constant random variable, hence by the Markov property we can write the likelihood function as

$$L^I(\theta_I) = f^I_{1:N}(y_1, \ldots, y_N | y_0, \theta_I) = \prod_{j=0}^{N-1} f^I_{j+1|j}(y_{j+1}|y_j, \theta_I), \tag{2.2.3}$$

where $f^I_{1:N}$ is the joint density function of the random variables $\{I(t_j)\}_{j=1}^N$ given $I(0)$ and the function $f^I_{j+1|j}$ is the conditional density function of $I(t_{j+1})$ given $I(t_j)$. In this case, the vector of parameters θ_I satisfies

$$\theta_I = (a_I, b_I, \sigma_I).$$

The conditional density $f^I_{j+1|j}$ is also called the transition density function. Due to the fact that I is a Cox-Ingersoll-Ross process, the transition density function $f^I_{j+1|j}$ is known (Jondeau, Poon, & Rockinger, 2007, p. 421). Let us define:

$$\begin{aligned} c &= \frac{2a_I}{\sigma_I^2(1 - \exp(-a_I\Delta))}, \\ q &= \frac{2a_I b_I}{\sigma_I^2} - 1, \\ u_j &= c y_j e^{-a_I \Delta} \text{ and} \\ p_{j+1} &= 2c y_{j+1}, \end{aligned}$$

for $j = 0, 1 \ldots, N-1$. Then the transition density function satisfies

$$f^I_{j+1|j}(y_{j+1}|y_j, \theta_I) = 2cf_{\chi^2}(p_{j+1}|2q+2, 2u_j) \quad \text{for } j = 0, \ldots, N-1, \tag{2.2.4}$$

where $f_{\chi^2}(.|2q+2, 2u_j)$ is the density function of a non central chi-square with $2q+2$ degrees freedom and non-centrality parameter $2u_j$. Using the transition density (2.2.4) and equation (2.2.3) we can construct the log-likelihood function as

$$\begin{aligned} l^I(\theta_I) &= \log(L^I(\theta_I)) \\ &= \sum_{j=0}^{N-1} \log\left(f^I_{j+1|j}(y_{j+1}|y_j, \theta_I)\right) \\ &= \sum_{j=0}^{N-1} \log\left(2cf_{\chi^2}(p_{j+1}|2q+2, 2u_j)\right). \end{aligned} \tag{2.2.5}$$

The maximum likelihood estimation method consists of maximising the likelihood L^I to estimate the vector of parameters θ_I. Because the logarithmic function is an increasing function, maximising L_I is equivalent to maximising the log-likelihood function l^I.

2.2.2 Validation

Now that we know how to estimate the parameters of the Cox-Ingersoll-Ross process, we will check that the observed data follows the proposed process. To this end, we will use the method of the generalized Gaussian residuals (Lindström, 2004, p.64).

This method requires us to determine the conditional distribution function of $I(t_{j+1})$ given $I(t_j)$, which is defined as

$$\begin{aligned} F^I_{j+1|j}(y_{j+1}) &= \mathbb{P}\left(I(t_{j+1}) \leq y_{j+1}|I(t_j)\right) \\ &= \int_0^{y_{j+1}} f^I_{j+1|j}(x|y_j, \theta_I)dx \quad \text{for } j = 0, 1, \ldots, N-1. \end{aligned}$$

We define the sequence $\{U_j\}_{j=1}^N$ as

$$U_j = F^I_{j|j-1}(I(t_j)) \quad \text{for } j = 1, 2, \ldots, N. \tag{2.2.6}$$

This is a sequence of independent and identically distributed standard uniform random variables (Lindström, 2004, Lemma 2.1).

Now let us define the sequence $\{M_j\}_{j=1}^N$ as

$$M_j = F_N^{-1}(U_j|0, 1) \quad \text{for } j = 1, \ldots N, \tag{2.2.7}$$

where $F_N(.|0, 1)$ is the cumulative distribution function of a normal random variable with mean 0 and variance 1. The sequence $\{M_j\}_{j=1}^N$ is a sequence of independent and identically distributed standard normal random variables (Lindström, 2004, p.68). Because of this, the $M_j's$ are called the generalised Gaussian residuals.

Going back to the observed sample $\{y_j\}_{j=0}^N$ we can compute a realisation of the Gaussian residuals $\{M_j\}_{j=1}^N$, that we call $\{m_j\}_{j=1}^N$. Hence for $j = 1, \ldots, N$ we have that:

$$m_j = F_N^{-1}\left(F^I_{j|j-1}(y_j)\right).$$

Under the hypotheses that the observations $\{y_j\}_{j=0}^N$ have been generated by a Cox-Ingersoll-Ross process with parameter θ_I, we have that $\{m_j\}_{j=1}^N$ has been generated by a sequence of independent and identically distributed standard normal random variables. It is possible to validate the model by checking the assumptions of normality and independence. To check the normality assumption, we can use the Kolmogorov–Smirnov test (Massey, 1951). For independence, we can use the Ljung-Box test (Ljung & Box, 1978) and check if there is any significant auto-correlation in the Gaussian residuals and in the squared Gaussian residuals.

2.2.3 Numerical experiments

For testing the methods explained in Section 2.2.1 and in Section 2.2.2 we will develop different experiments using simulated data. See Appendix D.1 to see how we generate the synthetic data.

In the first experiment, we generate 1000 simulations of $N = 1000$ steps of a Cox-Ingersoll-Ross process with parameters

$$a_I = 0.5, \quad b_I = 0.06, \quad \sigma_I = 0.15,$$

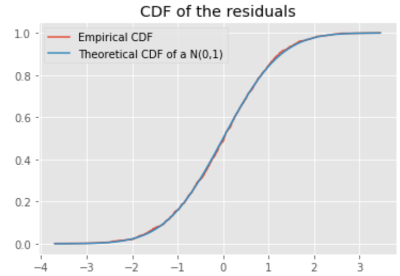
an initial value of $I(0) = 0.07$ and a time step $\Delta = 0.1$. For each of the simulations we estimate the parameters of the Cox-Ingersoll-Ross process using the method explained in Section 2.2.1. Once we estimate the parameters of all of the simulations, we compute the mean of the estimated parameters and its standard deviation. The results obtained in this experiment are shown in Table 2.1. As we can observe the method explained in Section 2.2.1 gives good estimates for the parameters. For obtaining the estimates of the parameters, we need to maximise the function l^I defined in (2.2.5). To maximise the function l^I we use the Python function “minimize” with the method “Sequential Least Squares Programming”, a function that is included in the Python package “SciPy”. This optimization algorithm needs an initial estimator of the parameters to start working. We can generate initial estimators of the parameters, using the technique explained in Appendix E.1.

Parameter	True value	Mean	Standard deviation
a_I	0.5	0.544686	0.11257
b_I	0.06	0.060179	0.007503
σ_I	0.15	0.150045	0.003391

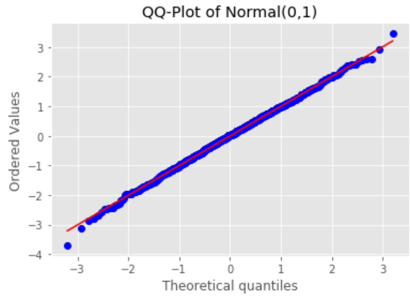
TABLE 2.1: Mean and standard deviation of the estimated parameters computed using the maximum likelihood estimation method for 1000 realizations.

For the second experiment, we generate one realisation of $N = 1000$ steps of the Cox-Ingersoll-Ross process with the same parameters, initial value, and time step as in the previous experiment. Using this realization, we first estimate the parameters and utilising the estimated parameters we compute the Gaussian residuals. We use the Kolmogorov–Smirnov test and the Ljung-Box test to check the normality and the dependence of the residuals, respectively. For the Kolmogorov–Smirnov test we obtain a p-value of 0.932116, so the normal assumption cannot be rejected. In the case of the Ljung-Box test, we obtain a p-value of 0.537033 for the realization of the residuals and a p-value of 0.729391 for the realization of the squared residuals. In both cases we cannot reject the assumption that the auto-correlations of the first 20 lags are zero.

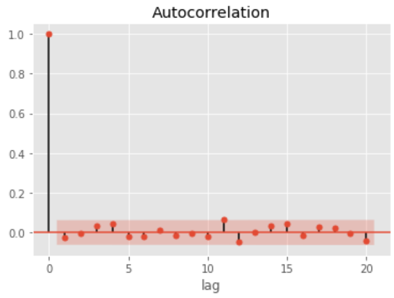
The results obtained by the statistical tests are in line with the results that appear in Figure 2.1.



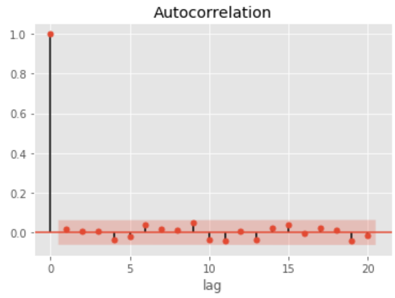
(A) Empirical cumulative distribution function of the residuals and theoretical distribution function of a standard normal.



(B) Standard normal quantile-quantile-plot.



(C) Sample auto-correlation function of the residuals.



(D) Sample auto-correlation function of the squared residuals.

FIGURE 2.1: Empirical CDF, QQ-plot and sample auto-correlation of one realisation of the Gaussian residuals defined in (2.2.7).

2.3 Ornstein-Uhlenbeck processes

Laying aside the Cox-Ingersoll-Ross process introduced in Section 2.2, we must turn to other processes as well. In this case, we want to estimate and validate Ornstein-Uhlenbeck processes. Since market attention is always positive, we focus ourselves only on positive Ornstein-Uhlenbeck processes. Positive Ornstein-Uhlenbeck processes have been used for modelling energy commodities (Benth, Kallsen, & Meyer-Brandis, 2007) and the volatility of stock prices (Barndorff-Nielsen & Shephard, 2003).

2.3.1 Introduction to Ornstein-Uhlenbeck processes

The overview of Ornstein-Uhlenbeck processes included here is based on the material presented by Valdivieso (2005) and by Valdivieso, Schoutens, and Tuerlinckx (2009). This includes Definition 2.3.1 and Propositions 2.3.1–2.3.4.

Let us consider a one dimensional Lévy process $Z_I = (Z_I(t))_{t\geq 0}$ with Lévy characteristic $(\gamma_0, \sigma_0, v_0)$. A process $I = (I(t))_{t\geq 0}$ is said to be an Ornstein-Uhlenbeck process if it has cádlág paths and it satisfies the following stochastic differential equation

$$dI(t) = -\lambda_I I(t)dt + dZ_I(\lambda_I t) \quad \text{with } I(0) = y_0, \tag{2.3.1}$$

where $\lambda_I, y_0 > 0$. The process Z_I is usually called the background driving Lévy process. The process I is also named a process of Ornstein-Uhlenbeck type generated by $(\gamma_0, \sigma_0, v_0, \lambda_I)$.

For any $t \geq 0$ the stochastic differential equation (2.3.1) has a strong solution that satisfies:

$$I(t) = e^{-\lambda_I t}\left(y_0 + \int_0^t e^{\lambda_I s} dZ_I(\lambda_I s)\right). \tag{2.3.2}$$

Sometimes it is more convenient to express the solution in equation (2.3.1) recursively. Let $\Delta > 0$ represents a time step. The strong solution of equation (2.3.1) at time $t + \Delta$ can be written in terms of the value at time t as

$$I(t+\Delta) = e^{-\lambda_I \Delta}\left(I(t) + e^{-\lambda_I t}\int_t^{t+\Delta} e^{\lambda_I s} dZ_I(\lambda_I s)\right), \tag{2.3.3}$$

this tells us that the process I is a Markov process.

Now let us focus on the random variable $Z_I^*(\Delta)$ defined by

$$Z_I^*(\Delta) = \int_0^{\lambda_I \Delta} e^s dZ_I(s). \tag{2.3.4}$$

The random variable $Z_I^*(\Delta)$ is quite important because it will allow us to compute the conditional characteristic function of the random variable $I(t+\Delta)$ given $I(t)$, to that end we will need the following result.

Proposition 2.3.1. *For any $t \geq 0$ and $\Delta \geq 0$*

$$Z_I^*(\Delta) \stackrel{d}{=} \int_0^{\Delta} e^{\lambda_I s} dZ_I(\lambda_I s) \stackrel{d}{=} e^{-\lambda_I t}\int_t^{t+\Delta} e^{\lambda_I s} dZ_I(\lambda_I s), \tag{2.3.5}$$

where $\stackrel{d}{=}$ means equal in distribution.

Using Proposition 2.3.1 and equation (2.3.3) we have that

$$e^{\lambda_I \Delta} I(t+\Delta) - I(t) = e^{-\lambda_I t}\int_t^{t+\Delta} e^{\lambda_I s} dZ_I(\lambda_I s) \stackrel{d}{=} Z_I^*(\Delta). \tag{2.3.6}$$

Equation (2.3.6) will play an important role when computing the maximum likelihood estimator.

Remark 2.3.1. *If μ is a distribution on $\mathbb{R}$ we will use the notation Φ_μ to indicate the characteristic function associated with μ and C_μ to represent the cumulant function, which are defined as*

$$\begin{aligned} \Phi_\mu(u) &= \int_{\mathbb{R}} e^{iux}\mu(dx) \text{ for all } u \in \mathbb{R}, \\ C_\mu(u) &= \log(\Phi_\mu(u)) \quad \text{for all } u \in \mathbb{R}. \end{aligned}$$

If we have a random variable X with law μ then we denote the characteristic function of X as Φ^X and the cumulant function of X as C^X. Both of them are defined as

$$\begin{aligned} \Phi^X(u) &= E\left[e^{iuX}\right] = \int_{\mathbb{R}} e^{iux}\mu(dx) \text{ for all } u \in \mathbb{R}, \\ C^X(u) &= \log\left(\Phi^X(u)\right) \text{ for all } u \in \mathbb{R}. \end{aligned}$$

Ornstein-Uhlenbeck processes are related to a special type of distribution called self-decomposable distribution, which is defined below.

Definition 2.3.1. *A distribution μ is said to be* self-decomposable *if for any $0 < a < 1$, there exists a distribution v_a such that:*

$$\Phi_\mu(u) = \Phi_\mu(au)\Phi_{v_a}(u) \text{ for all } u \in \mathbb{R}.$$

Alternatively, a random variable X is said to have a self-decomposable *distribution if for any $0 < a < 1$, there exists a random variable Y_a, independent of X, such that:*

$$X \stackrel{d}{=} aX + Y_a.$$

The proposition below tells us that if we can find a distribution D that is self-decomposable, then we can construct an Ornstein-Uhlenbeck process that satisfies equation (2.3.1) and that has a stationary distribution D.

Proposition 2.3.2. *If I is an Ornstein-Uhlenbeck process generated by a background driving Lévy process Z_I with Lévy triplet $(\gamma_0, \sigma_0, v_0)$ and a parameter $\lambda_I > 0$ such that*

$$\int_{|x|>2} \log(|x|) dv_0(x) < \infty, \tag{2.3.7}$$

then I has a unique self-decomposable stationary distribution μ.

Conversely, for any $\lambda_I > 0$ and any self-decomposable distribution D, there exists a unique Lévy triplet $(\gamma_0, \sigma_0, v_0)$ satisfying (2.3.7) and a process of Ornstein-Uhlenbeck type generated by $(\gamma_0, \sigma_0, v_0, \lambda_I)$ such that D is the stationary distribution of I.

A process I satisfying the converse result of Proposition 2.3.2 is called a D-Ornstein-Uhlenbeck process. In this thesis we are only interested in stationary D-Ornstein-Uhlenbeck processes.

Proposition 2.3.2 says that for every self-decomposable distribution D, we can construct an Ornstein-Uhlenbeck process with stationary distribution D. However, it does not say anything related to the distribution of the background driving Lévy process Z_I. The next result tells us how the distributions of Z_I and D relate.

Proposition 2.3.3. *For any $u \in \mathbb{R}$:*

$$C^{Z_I(1)}(u) = \log\left(E\left[e^{iuZ_I(1)}\right]\right) = u\frac{\partial C_D(u)}{\partial u}.$$

We have seen that the distribution of the random variable $Z_I^*(\Delta)$ is important in determining the transition distribution of the Ornstein-Uhlenbeck process. Proposition 2.3.4 allows us to compute the characteristic function of $Z_I^*(\Delta)$ using the distribution of the background driving Lévy process Z_I.

Proposition 2.3.4. *For any $\Delta > 0$ and $u \in \mathbb{R}$:*

$$\Phi^{Z_I^*(\Delta)}(u) = E\left[e^{iuZ_I^*(\Delta)}\right] = e^{\lambda \int_0^\Delta C^{Z_I(1)}(ue^{\lambda_I s})ds}.$$

Using Proposition 2.3.3 and Proposition 2.3.4 we can compute the characteristic function of the random variable $Z_I^*(\Delta)$ from a self-decomposable distribution D. From equation (2.3.6) we have that $e^{-\lambda_I\Delta}I(t+\Delta) - I(t)$ has the same distribution as $Z_I^*(\Delta)$. So the characteristic distribution of $e^{-\lambda_I\Delta}I(t+\Delta) - I(t)$ can also be computed from the self-decomposable distribution D.

2.3.2 Inverse Gaussian Ornstein-Uhlenbeck process

This type of stationary Ornstein-Uhlenbeck process has as stationary distribution an inverse Gaussian distribution with parameters $a_I > 0$ and $b_I > 0$, which we denote as $IG(a_I, b_I)$. The density function of an inverse Gaussian random variable with parameters a_I and b_I is

$$f_{IG}(x|a_I, b_I) = \frac{a_I e^{a_I b_I}}{\sqrt{2\pi x^3}} e^{\frac{-1}{2}\left(\frac{a_I^2}{x}+b_I^2 x\right)} \quad \text{for all } x > 0,$$

and its characteristic function is:

$$\Phi_{IG}(u|a_I, b_I) = e^{a_I(b_I - \sqrt{b^2 - 2iu})} \quad \text{for all } u \in \mathbb{R}.$$

The inverse Gaussian distribution is self-decomposable. By application of Proposition 2.3.2, there exists an $IG(a_I, b_I)$-Ornstein-Uhlenbeck process, since the inverse Gaussian distribution is self-decomposable. Using Proposition 2.3.3 we can compute the cumulant function of the background driving Lévy process Z_I of the $IG(a_I, b_I)$-Ornstein-Uhlenbeck process:

$$\begin{aligned} C^{Z_I(1)}(u) &= u\frac{\partial C_{IG}(u|a_I, b_I)}{\partial u} \\ &= \frac{a_I u i}{\sqrt{b_I^2 - 2iu}} \quad \text{for all } u \in \mathbb{R}. \end{aligned} \tag{2.3.8}$$

From equation (2.3.8) and by application of Proposition 2.3.4 we have that

$$\begin{aligned} \Phi^{Z_I^*(\Delta)}(u) &= e^{\lambda_I \int_0^\Delta C^{Z_I(1)}(ue^{\lambda_I s})ds} \\ &= e^{\lambda_I \int_0^\Delta \frac{a_I iue^{\lambda_I s}}{\sqrt{b_I^2 - 2 - iue^{\lambda s}}} ds} \\ &= e^{a_I\left(\sqrt{b_I^2 - 2iu} - \sqrt{b_I^2 - 2iue^{\lambda_I \Delta}}\right)} \quad \text{for all } u \in \mathbb{R}. \end{aligned} \tag{2.3.9}$$

We have just computed the characteristic function of the random variable $Z_I^*(\Delta)$ of an $IG(a_I, b_I)$-Ornstein-Uhlenbeck process. The result obtained in equation (2.3.9) will allow us to approximate the transition density function of an $IG(a_I, b_I)$-Ornstein-Uhlenbeck process. The results presented in this section come from the work of Valdivieso (2005).

2.3.3 Maximum likelihood estimation for the inverse Gaussian Ornstein Uhlenbeck process

We turn our attention to estimate the parameters of an $IG(a_I, b_I)$-Ornstein-Uhlenbeck process. Let I be an $IG(a_I, b_I)$-Ornstein-Uhlenbeck process, with an unknown vector of parameters $\theta_I = (a_I, b_I, \lambda_I)$.

As in Section 2.2.1 we assume a collection of $N+1$ observations $\{y_j\}_{j=0}^N$, which are equally spaced in time with a time step $\Delta > 0$. Each y_j is a realisation of the random variable $I(t_j)$ for all $j = 0, 1, \ldots, N$. Again, we consider $I(0)$ to be a constant random variable.

By the Markov property, the likelihood function is:

$$L^I(\theta_I) = f^I_{1:N}(y_1, \ldots, y_N|y_0, \theta_I) = \prod_{j=0}^{N-1} f^I_{j+1|j}(y_{j+1}|y_j, \theta_I), \tag{2.3.10}$$

where $f^I_{1:N}$ is the joint density function of the random variables $\{I(t_j)\}_{j=1}^N$ given $I(0)$ and the function $f^I_{j+1|j}$ is the conditional density function of $I(t_{j+1})$ given $I(t_j)$.

Because the observations are equally spaced in time, we can express $t_{j+1} = t_j + \Delta$, for $j = 0, 1, \ldots, N-1$. From equation (2.3.6) we have that

$$e^{\lambda_I \Delta} I(t_{j+1}) - I(t_j) = e^{\lambda_I \Delta} I(t_j + \Delta) - I(t_j) \stackrel{\mathrm{d}}{=} Z_I^*(\Delta) \tag{2.3.11}$$

for all $j = 0, 2, \ldots, N-1$. Hence, we can write

$$\mathbb{P}\left(I(t_{j+1}) \leq y_{j+1}|I(t_j) = y_j\right) = \mathbb{P}\left(Z_I^*(\Delta) \leq e^{\lambda_I \Delta} y_{j+1} - y_j\right)$$

(Valdivieso, Schoutens, & Tuerlinckx, 2009, p. 11), and the conditional density $f^I_{j+1|j}$ can be written as

$$f^I_{j+1|j}(y_{j+1}|y_j, \theta_I) = e^{\lambda_I \Delta} f^{Z_I^*(\Delta)}\left(e^{\lambda_I \Delta} y_{j+1} - y_j|\theta_I\right) \tag{2.3.12}$$

(Valdivieso, 2005, p. 99).

Now we can write the likelihood (2.3.10) as

$$\begin{aligned} L^I(\theta_I) &= e^{N\lambda_I \Delta} \prod_{j=0}^{N-1} f^{Z_I^*(\Delta)}\left(e^{\lambda_I \Delta} y_{j+1} - y_j|\theta_I\right) \\ &= e^{N\lambda_I \Delta} \prod_{j=0}^{N-1} f^{Z_I^*(\Delta)}\left(z^*_{j+1}|\theta_I\right), \end{aligned} \tag{2.3.13}$$

where $z^*_{j+1} = e^{\lambda_I \Delta} y_{j+1} - y_j$ for $j = 0, 1, \ldots, N-1$. Since we have an analytical formula for the characteristic function of $Z_I^*(\Delta)$, the density function of $Z_I^*(\Delta)$ can be computed as

$$f^{Z_I^*(\Delta)}(z_j^*|\theta_I) = \frac{1}{\pi} \int_0^\infty \Re\left[e^{-iuz_j^*} \Phi^{Z_I^*(\Delta)}(u|\theta_I)\right] du$$

(Jondeau, Poon, & Rockinger, 2007, p. 483). From equation (2.3.13) we have that the log-likelihood can be expressed as

$$l^I(\theta_I) = N\lambda_I \Delta + \sum_{j=0}^{N-1} \log\left(f^{Z_I^*(\Delta)}\left(z^*_{j+1}|\theta_I\right)\right). \tag{2.3.14}$$

We maximise the function l^I to estimate the parameter θ_I.

2.3.4 Validation

Once we know how to estimate the parameters of an $IG(a_I, b_I)$-Ornstein-Uhlenbeck process, we would like to verify that the observed sample $\{y_j\}_{k=0}^N$ truly follows an inverse Gaussian Ornstein-Uhlenbeck process with vector of parameters $\theta_I = (a_I, b_I, \lambda_I)$.

To that end we define the residuals $\{M_j\}_{j=1}^N$ as

$$M_{j+1} = e^{\lambda_I \Delta} I(t_{j+1}) - I(t_j) = e^{-\lambda_I t_j} \int_{t_j}^{t_j+\Delta} e^{\lambda_I s} dZ_I(\lambda_I s) \quad \text{for } j = 0, \dots, N-1. \tag{2.3.15}$$

Notice that $\{M_j\}_{j=1}^N$ is a sequence of independent random variables, since the sequence of stochastic integrals $\{\int_{t_j}^{t_j+\Delta} e^{\lambda_I s} dZ_I(\lambda_I s)\}_{j=0}^{N-1}$ consists of independent random variables. By Proposition 2.3.1 we have that

$$M_j \overset{\mathrm{d}}{=} Z_I^*(\Delta) \quad \text{for } j = 1, 2, \dots, N.$$

If we assume that the data have been generated by an inverse Gaussian Ornstein-Uhlenbeck process with parameter θ_I, then the sequence $\{z_j^*\}_{j=1}^N$, where $z_j^* = e^{\lambda_I \Delta} y_j - y_{j-1}$ for $j = 1, \dots, N$, has been generated by the sequence of independent and identically distributed random variables $\{M_j\}_{j=1}^N$.

So, under the null hypotheses we have that $\{z_j^*\}_{j=1}^N$ has been generated by a sequence of independent and identically distributed random variables, whose distribution is equal to the distribution of $Z_I^*(\Delta)$. Because $Z_I^*(\Delta)$ is a continuous random variable, we can apply the one-sided Kolmogorov–Smirnov test (Massey, 1951). This test allows us to check if a sample has been generated by a known distribution. This test is based in comparing the theoretical cumulative distribution function with the empirical cumulative distribution function.

For applying the one-sided Kolmogorov–Smirnov test, we need to be able to compute the cumulative distribution function of the random variable $Z_I^*(\Delta)$. We do not have an analytical form for the cumulative distribution function of $Z_I^*(\Delta)$. However, as it is continuous, the cumulative distribution function can be computed form its characteristic function using the inversion theorem that is,

$$F^{Z_I^*(\Delta)}(z_j^*|\theta_I) = \frac{1}{2} - \frac{1}{\pi} \int_0^\infty \frac{1}{u} \Im \left[e^{-iuz_j^*} \Phi^{Z_I^*(\Delta)}(u|\theta_I) \right] du$$

(Jondeau, Poon, & Rockinger, 2007, Proposition 15.2). In addition to checking that the sequence $\{z_j^*\}_{j=1}^N$ has been generated by a sequence of random variables with the same distribution as $Z_I^*(\Delta)$, we also need to check the independence assumption. To this end we will use the Ljung-Box test (Ljung & Box, 1978) for checking the significance of the auto-correlations of the residuals and the squared residuals.

2.3.5 Numerical experiments

As we did with the Cox-Ingersoll-Ross process, we would like to test the methods explained in Section 2.3.3 and in Section 2.3.4. To that end, we will generate simulated data (see Appendix D.2).

We first generate a realization of $N = 800$ steps of an inverse Gaussian Ornstein-Uhlenbeck process with parameters

$$a_I = 4, \quad b_I = 10, \quad \lambda_I = 5,$$

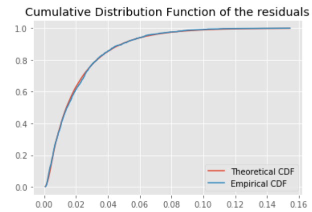
an initial value $I(0) = 0.4$ and a time step $\Delta = 0.0125$. For this realisation, we estimate the parameters of the process using the method explained in Section 2.3.3. As we can observe in Table 2.2, the method gives good estimates for the parameters.

Recall that to obtain the parameters we need to maximise the function l^I defined in (2.3.14). To maximize the function l^I we use the Python function “minimize” with the method “Sequential Least Squares Programming”; this function is included in the Python package “SciPy”. This optimisation algorithm needs an initial estimator of the parameters to start working. For this reason, we generate initial estimators of the parameters, using the technique explained in Appendix E.2.

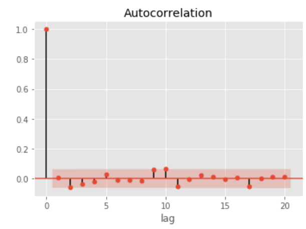
Parameter	True value	Estimated value
a_I	4	3.981078
b_I	10	9.744054
λ_I	5	5.072365

TABLE 2.2: Estimated parameters using the maximum likelihood estimation method.

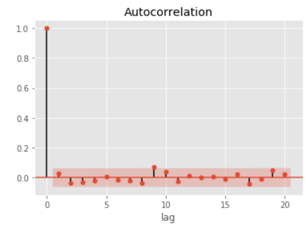
Once we have estimated the parameters of the inverse Gaussian Ornstein-Uhlenbeck process, we compute the residuals introduced in Section 2.3.4, using the estimated parameters. The Kolmogorov–Smirnov test gives us a p-value of 0.447723, so we cannot reject the assumption that the residuals satisfy the same distribution as $Z_I^*(\Delta)$. In the case of the Ljung-Box test, we obtain a p-value of 0.380174 for the realisation of the residuals and a p-value of 0.586722 for the realisation of the squared residuals. In both cases, we cannot reject the assumption that the auto-correlations of the first 20 lags are zero. These results are in line with the graphs shown in Figure 2.2.



(A) Empirical cumulative distribution function of the residuals and the theoretical distribution function of the random variable $Z_I^*(\Delta)$.



(B) Sample auto-correlation function of the residuals.



(C) Sample auto-correlation function of the squared residuals.

FIGURE 2.2: Empirical CDF and sample auto-correlation of a realization of the residuals $\{M_j\}_{j=1}^N$, defined in Section 2.3.4.

2.4 Proxies for the market attention

With the tools presented in previous sections, we would like to verify if the selected proxies for the interest follow a Cox-Ingersoll-Ross process or an inverse Gaussian Ornstein-Uhlenbeck process. Some examples of attention proxies that can be found in the literature are: volume (Figà-Talamanca & Patacca, 2019, 2020), the number of Google searches (Figà-Talamanca & Patacca, 2019, 2020) , the number of Wikipedia views (Kristoufek, 2015) or number of tweets (Al Guindy, 2021).

2.4.1 Cox-Ingersoll-Ross process

We are focused on modelling market attention proxies using a Cox-Ingersoll-Ross process. Firstly, we analyse the number of daily Wikipedia views for certain keywords.

All selected keywords are related to Bitcoin. The keywords that we examine are: "Bitcoin", "Blockchain", "Cryptocurrency", "Bitcoin Network", "Bitcoin Wallet", "Binance", "Hodl", "Sathoshi Nakamoto" and "Altcoin". The only keyword that seems to follow a Cox-Ingersoll-Ross process is "Altcoin". The results are presented in Table 2.3. For each of the different temporal windows, we estimate the parameters and compute the Gaussian residuals presented in Section 2.2.2. As we can observe, all p-values of the Kolmogorov-Smirnov test are above 0.05, meaning that we cannot reject the assumption of normality for the residuals. Hence, in all temporal windows, there is statistical evidence that the number of views follows a Cox-Ingersoll-Ross process. Notice that the estimated parameters vary depending on which temporal window are estimated. This could indicate that the Wikipedia views follow a regime switching process, but these types of processes are outside of the scope of this thesis.

Temporal window	a_I	b_I	σ_I	p-value
2018-09-20 to 2019-09-20	114.083863	21.76599	35.05684	0.611167
2019-09-20 to 2020-09-20	388.902639	11.288746	36.771933	0.300911
2020-09-20 to 2021-09-20	38.984017	39.869911	33.281242	0.884317
2018-09-20 to 2020-09-20	123.01943	16.502741	33.380029	0.27324
2019-09-20 to 2021-09-20	43.16364	25.446899	31.667167	0.306409
2018-09-20 to 2021-09-20	54.944194	24.149982	32.541913	0.174079

TABLE 2.3: Estimated parameters of the Cox-Ingersoll-Ross process when using the number of Wikipedia views of the keyword "Altcoin" as a proxy for the interest. The p-value of the Kolmogorov-Smirnov test applied to the Gaussian residuals is also shown.

Another proxy studied is the daily number of unique active addresses. The results are shown in Table 2.4. The p-values of the Kolmogorov-Smirnov test are above 0.05, except for the temporal window from 2018-09-20 to 2021-09-20, which has a value of 0.009635. However, we assume that we have enough evidence to assume that the number of unique active addresses satisfies a Cox-Ingersoll-Ross process.

Temporal window	a_I	b_I	σ_I	p-value
2018-09-20 to 2019-09-20	157.196897	696049.730926	2323.112482	0.498187
2019-09-20 to 2020-09-20	106.85478	811354.646069	2178.042426	0.599642
2020-09-20 to 2021-09-20	100.873074	1005752.522156	2267.874974	0.958168
2018-09-20 to 2020-09-20	102.281197	753978.695629	2207.131253	0.087094
2019-09-20 to 2021-09-20	69.263311	909024.347716	2174.369875	0.225312
2018-09-20 to 2021-09-20	58.811268	838703.355729	2165.110738	0.009635

TABLE 2.4: Estimated parameters of the Cox-Ingersoll-Ross process when using the number of unique active addresses as a proxy for the interest. The p-value of the Kolmogorov-Smirnov test applied to the Gaussian residuals is also shown.

2.4.2 Inverse Gaussian Ornstein-Uhlenbeck process

In the case of the Inverse Gaussian Ornstein-Uhlenbeck process, we analyse the logarithm of the daily number of unique active addresses. Since the values of the number of addresses are on the order of 10^5, we do not have to worry about getting negative values. So, the time series can still be used to fit a positive Ornstein-Uhlenbeck process. We fit the inverse Gaussian Ornstein-Uhlenbeck process to different temporal windows of the time series. For each of the temporal windows, we estimate the parameters and compute the residuals introduced in Section 2.3.4. The results are shown in Table 2.5. Notice that the p-values of the Kolmogorov-Smirnov test that appear in Table 2.5 are all above 0.1. Hence, there is sufficient statistical evidence to assume that the logarithm of the number of unique active addresses follows an inverse Gaussian Ornstein-Uhlenbeck process.

Temporal window	a_I	b_I	λ_I	p-value
2018-09-20 to 2019-09-20	310.839434	23.126722	162.028074	0.414629
2019-09-20 to 2020-09-20	299.362691	22.023851	112.464686	0.736125
2020-09-20 to 2021-09-20	317.616805	23.000978	107.376863	0.957407
2018-09-20 to 2020-09-20	276.400641	20.448327	106.875749	0.128057
2019-09-20 to 2021-09-20	258.05562	18.83415	73.417364	0.672384
2018-09-20 to 2021-09-20	226.78821	16.657239	62.177681	0.136611

TABLE 2.5: Estimated parameters of the inverse Gaussian Ornstein-Uhlenbeck process when using the logarithmic number of unique active addresses as a proxy for the interest. The p-value of the Kolmogorov-Smirnov test applied to the residuals is also shown.

We also analyse the number of Wikipedia views and the logarithm of Wikipedia views, but in both cases we cannot find any evidence that the data follow an inverse Gaussian Ornstein-Uhlenbeck process.

Remark 2.4.1. *We have attempted to test the autocorrelation of the residuals and squared residuals using the Ljung-Box test, but the results show that there is evidence of autocorrelation in the residuals and squared residuals, for all cases studied in Tables 2.3, 2.4 and 2.5. This could indicate that there is some kind of memory in the interest process and because of that Markov processes are not appropriate. One could use processes that satisfy fractional stochastic differential equations, like the fractional Cox-Ingersoll-Ross process (Mishura & Yurchenko-Tytarenko, 2018) or the geometric fractional Brownian motion (Pipiras & Taqqu, 2017, p. 412). These types of processes are outside of the scope of this thesis. A good result was obtained for the Kolmogorov-Smirnov test, so we assume that the Cox-Ingersoll-Ross process and the inverse Gaussian Ornstein-Uhlenbeck process are valid processes for modelling market attention.*

2.5 Conclusion and future work

In this chapter, we discussed the estimation and validation of the Cox-Ingersoll-Ross process and the inverse Gaussian Ornstein-Uhlenbeck process. In addition, we showed how these processes can be used to model market attention. In the following chapters, we will use this knowledge to build models to price Bitcoin options.

Here, we focused only on two processes, but other processes can be used. For example, Cretarola, Figà-Talamanca, and Patacca (2020) use a geometric Brownian motion to model the volume of Bitcoin. Apart from the inverse Gaussian Ornstein-Uhlenbeck process, other types of positive Ornstein-Uhlenbeck processes can be used, like the gamma Ornstein-Uhlenbeck process or the tempered stable Ornstein-Uhlenbeck process (Schoutens, 2003, p.68–70). Other processes that could be used to model the interest are processes used to model population growth, like the stochastic logistic model (Skiadas, 2010, p. 266).

Chapter 3

A first simple model

3.1 Introduction

In Chapter 2 we introduced the concept of market attention. In this chapter, we will be concerned with using market attention to model Bitcoin prices. Here, we proposed a simple model that relates the prices of Bitcoin to market interest. For this model we will estimate its parameters and price vanilla options.

How the returns of Bitcoin relates to market attention has previously been studied in the literature. But the findings of previous literature have been inconsistent. For example, some sources claim to find that the interest affects both the mean and the volatility of Bitcoin returns (Figà-Talamanca & Patacca, 2019; Kristoufek, 2015). In other cases, the authors say that the market attention seems to affect only the volatility of Bitcoin returns (Aalborg, Molnár, & de Vries, 2019; Figà-Talamanca & Patacca, 2020). In our case, we assume that market attention influences only the volatility of Bitcoin returns. From our point of view this seems logical because an increase in the number of news stories or in the number of searches means that investors are paying attention to Bitcoin and it is likely to produce a change in the price. However, it is not possible to infer directly from the interest the direction of this change. In addition, the change in the price is not produced instantaneously, but rather with a delay (Kristoufek, 2015). Because of that, we propose that the logarithm of Bitcoin prices satisfies the following stochastic differential equation:

$$dX(t) = \mu dt + g(I(t-\tau))dW_P(t), \tag{3.1.1}$$

where X and I represent the logarithm of the price and the interest processes respectively, W_P is a Brownian motion, $\mu \in \mathbb{R}$, $\tau \geq 0$ and g is a non-decreasing continuous function. By application of the Itô formula we have that the price process $P(t) = e^{X(t)}$ for $t \geq 0$ will satisfy the following stochastic differential equation

$$dP(t) = \left(\mu + \frac{1}{2}g^2(I(t-\tau))\right) P(t)dt + g(I(t-\tau))P(t)dW_P(t). \tag{3.1.2}$$

The model that appears in equation (3.1.2) is partially based on the model proposed by Cretarola, Figà-Talamanca, and Patacca (2020). Inspired by equation (3.1.1) we propose a model in which the interest follows an affine process and $g(x) = c\sqrt{x}$ for some $c > 0$. In that way, we can have a model that is analytically tractable.

Due to the fact that the model presented in this chapter is similar to the model proposed by Cretarola, Figà-Talamanca, and Patacca (2020), we briefly analyse the main characteristics of this model. In this article, the authors proposed the following model for the Bitcoin price:

$$\begin{aligned} dP(t) &= \mu_P I(t-\tau)P(t)dt + \sigma_P\sqrt{I(t-\tau)}P(t)dW_P(t) \text{ with } P(0) = p \in \mathbb{R} \\ dI(t) &= \mu_I I(t)dt + \sigma_I I(t)dW_I(t) \text{ for } t > 0 \\ \text{if} \quad & I(t) = \phi^I(t) \text{ when } t \in [-L, 0], \end{aligned}$$

where $\mu_P, \mu_I \in \mathbb{R}$, $\sigma_P, \sigma_I, L > 0$ and $\tau \in [0, L]$ and W_P, W_I are two independent Brownian motions. They use as proxies for market attention the volume of transactions and Google searches of the word "Bitcoin". Notice that in this case the interest process follows a geometric Brownian motion. The model presented by Cretarola, Figà-Talamanca, and Patacca (2020) is similar to the stochastic volatility model of Hull and White (1987), with the difference in the presence of the delay parameter. Notice that in the model proposed by Cretarola, Figà-Talamanca, and Patacca (2020), the market attention affects the mean and volatility of the log returns. In the model that we propose, the market attention only affects the volatility of log returns. We identify some space for improvement,

1. The model presented by Cretarola, Figà-Talamanca, and Patacca (2020), does not have a closed formula for pricing plain vanilla European options. We will show in this chapter that it is possible to construct a stochastic volatility model that contains a delay parameter.

2. In addition, the empirical density function of the Bitcoin returns has fat tails (Chan, Chu, Nadarajah, & Osterrieder, 2017). Due to this fact, it is convenient to include jumps in the price modelling of Bitcoin. We will show that it is possible to construct models that have jumps in Chapter 5, which have a closed formula for pricing European options.

3.2 Model specification

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space that contains two independent Brownian motions W_P and W_I. Assume that the price P and the interest I satisfy the following system of stochastic differential equations:

$$\begin{aligned} dP(t) &= \left(\mu + \frac{\sigma_P^2}{2} I(t-\tau)\right) P(t)dt + \sigma_P\sqrt{I(t-\tau)}P(t)dW_P(t) \text{ with} && (3.2.1) \\ P(0) &= p \in \mathbb{R}_+ = (0, \infty), \\ dI(t) &= a_I(b_I - I(t))dt + \sigma_I\sqrt{I(t)}dW_I(t) \text{ when } t > 0 \text{ and} && (3.2.2) \\ I(t) &= \phi^I(t),\ t \in [-L, 0], \end{aligned}$$

where $\mu, b_I \in \mathbb{R}$, $\sigma_P, \sigma_I, a_I, L > 0$, $\tau \in [0, L]$ and we impose the condition

$$\frac{2a_I b_I}{\sigma_I^2} \geq 1. \tag{3.2.3}$$

Here τ is a fixed lag parameter and we have the following continuous and deterministic initial function

$$\phi^I : [-L, 0] \to (0, \infty).$$

Notice that the function ϕ^I is always positive because the interest is required to be greater than zero.

In Section 3.3 we will show that stochastic differential equations (3.2.1)–(3.2.2) have strong solutions. So equations (3.2.1) and (3.2.2) define the price and the market attention processes respectively.

In addition, we assume that there is a bond or a market account $(B(t))_{t\geq 0}$ with known interest rate $r \geq 0$, that satisfies:

$$B(t) = B(0)e^{rt}.$$

Define $(\mathcal{F}_t^{W_P})_{t\geq 0}$ and $(\mathcal{F}_t^{W_I})_{t\geq 0}$ to be the filtrations generated by Brownian motions W_P and W_I respectively. Let us consider the general filtration $(\mathcal{F}_t)_{t\geq 0}$, which is defined as

$$\mathcal{F}_t = \mathcal{F}_t^{W_P} \vee \mathcal{F}_t^{W_I} \text{ for } t \geq 0,$$

where $\mathcal{F}_t^{W_P} \vee \mathcal{F}_t^{W_I}$ is the smallest σ-field containing $\mathcal{F}_t^{W_P}$ and $\mathcal{F}_t^{W_I}$. In addition, we define the delayed filtration $(\tilde{\mathcal{F}}_t)_{t\geq 0}$ as

$$\tilde{\mathcal{F}}_t = \begin{cases} \mathcal{F}_t^{W_P} \text{ if } t \leq \tau, \\ \mathcal{F}_t^{W_P} \vee \mathcal{F}_{t-\tau}^{W_I} \text{ if } t > \tau. \end{cases}$$

Notice that

$$\tilde{\mathcal{F}}_t \subseteq \mathcal{F}_t \;\forall t \geq 0.$$

We assume that both filtrations satisfy the usual hypotheses. These usual hypotheses are:

1. Completeness: $\mathcal{F}_0$ and $\tilde{\mathcal{F}}_0$ contain all sets of $\mathbb{P}$-measure zero.

2. Right continuity:

$$\mathcal{F}_t = \mathcal{F}_{t^+} \qquad \text{where } \mathcal{F}_{t^+} = \bigcap_{\epsilon>0} \mathcal{F}_{t+\epsilon},$$
$$\tilde{\mathcal{F}}_t = \tilde{\mathcal{F}}_{t^+} \qquad \text{where } \tilde{\mathcal{F}}_{t^+} = \bigcap_{\epsilon>0} \tilde{\mathcal{F}}_{t+\epsilon},$$

(Applebaum, 2009, p. 72). In addition, we also assume that $\mathcal{F}_t \subseteq \mathcal{F}$ for all $t \geq 0$.

It is clear that the processes $(P(t))_{t\geq 0}$ and $(I(t))_{t\geq 0}$ are adapted with respect to the filtration $(\mathcal{F}_t)_{t\geq 0}$. Because the deterministic function ϕ^I is adapted to any filtration and $\tau \geq 0$, the process $(I(t-\tau))_{t\geq 0}$ is also adapted to $(\mathcal{F}_t)_{t\geq 0}$. The processes $(P(t))_{t\geq 0}$ and $(I(t-\tau))_{t\geq 0}$ are also adapted with respect to the delayed filtration $(\tilde{\mathcal{F}}_t)_{t\geq 0}$.

3.3 Strong solutions

In this section, we will show that the stochastic differential equations (3.2.1)–(3.2.2) have strong solutions.

First, let us show that equation (3.2.2) has a strong solution. Notice that when $t \in [-L, 0]$ we have that $I(t) = \phi^I(t)$. In the case when $t > 0$, we know that the interest satisfies the stochastic differential equation (3.2.2). From Section 2.2 we know that there is a unique strong solution for equation (3.2.2) and because of condition (3.2.3) we have that $I(t) > 0$ for all $t > 0$ almost surely.

Now we will show that a strong solution exists for the stochastic differential equation (3.2.1).

Proposition 3.3.1. *The stochastic differential equation (3.2.1) has a strong solution and it has the following form:*

$$P(t) = e^{X(t)} \text{ for } t \geq 0. \tag{3.3.1}$$

where

$$X(t) = X(0) + \mu t + \int_0^t \sigma_P \sqrt{I(u-\tau)} dW_P(u) \text{ for } t \geq 0, \tag{3.3.2}$$

with $X(0) = \log P(0)$.

Proof. The first thing to notice is that the integral involved in the proposed solution (3.3.1) is well defined, since $(I(t-\tau))_{t\geq 0}$ is a continuous and $(\mathcal{F}_t)_{t\geq 0}$-adapted process. Now we will show that the proposed solution (3.3.1) satisfies equation (3.2.1). Using the Itô formula to compute $e^{X(t)}$, we obtain that

$$\begin{aligned} dP(t) &= \left(e^{X(t)}\mu + \frac{1}{2} e^{X(t)} \sigma_P^2 I(t-\tau) \right) dt \\ &\quad + e^{X(t)} \sigma_P \sqrt{I(t-\tau)} dW_P(t) \\ &= \left(\mu + \frac{\sigma_P^2}{2} I(t-\tau) \right) P(t) dt + \sigma_P \sqrt{I(t-\tau)} P(t) dW_P(t), \end{aligned}$$

as required. □

In general, it is more convenient to work with the log price than with the price.

3.4 Markov property

In this section, we are concerned with the Markov property of the process Y defined as

$$Y(t) = (X(t), I(t-\tau)) \text{ for } t \geq 0. \tag{3.4.1}$$

Proposition 3.4.1. *The process Y defined in (3.4.1) is a Markov process with respect to the filtration $(\tilde{\mathcal{F}}_t)_{t\geq 0}$.*

Proof. When $t \leq \tau$ we have

$$\begin{aligned} X(t) &= X(0) + \mu t + \int_0^t \sigma_P \sqrt{\phi^I(u-\tau)} dW_P(u), \\ I(t-\tau) &= \phi^I(t). \end{aligned}$$

Observe that when $t \leq \tau$ the process I is deterministic and X is the strong solution of a stochastic differential equation. This means that when $t \leq \tau$ we have that Y is a Markov process (Karatzas & Shreve, 1998, Theorem 5.4.20). When $t > \tau$ the process X can be written as

$$\begin{aligned} X(t) &= X(0) + \int_0^\tau \mu du + \int_0^\tau \sigma_P \sqrt{\phi^I(u-\tau)} dW_P(u) \\ &\quad + \int_0^{t-\tau} \mu du + \int_\tau^t \sigma_P \sqrt{I(u-\tau)} dW_P(u) \\ &= X(\tau) + \int_0^{t-\tau} \mu du + \int_\tau^t \sigma_P \sqrt{I(u-\tau)} dW_P(u). \end{aligned}$$

From Theorem 1.2.1 we have that

$$\int_\tau^t \sigma_P \sqrt{I(u-\tau)}dW_P(u) = \int_0^{t-\tau} \sigma_P \sqrt{I(u)}dB_P(u) \text{ almost surely,} \tag{3.4.2}$$

where $B_P(u) = W_P(u+\tau) - W_P(\tau)$ for $u \geq 0$. Then we can write

$$X(t) \;=\; X(\tau) + \int_0^{t-\tau} \mu du + \int_0^{t-\tau} \sigma_P \sqrt{I(u)}dB_P(u).$$

Define a new process Z as

$$Z(t) = Z(0) + \int_0^t \mu du + \int_0^t \sigma_P \sqrt{I(u)}dB_P(u) \text{ with } Z(0) = X(\tau). \tag{3.4.3}$$

Notice that the process Z is adapted with respect to the filtration $(\mathcal{F}_t^*)_{t\geq 0}$ defined as

$$\mathcal{F}_t^* = \mathcal{F}_\tau^{W_P} \vee \mathcal{F}_t^{B_P} \vee \mathcal{F}_t^{W_I} \text{ for } t \geq 0,$$

where $\left(\mathcal{F}_t^{B_P}\right)_{t\geq 0}$ is the filtration generated by the Brownian motion B_P. Let us consider $s > \tau$. Now we would like to prove that $\mathcal{F}_s^{W_P} = \mathcal{F}_\tau^{W_P} \vee \mathcal{F}_{s-\tau}^{B_P}$. Since $B_P(s-\tau) = W_P(s) - W_P(\tau)$ we have that $B_P(s-\tau)$ is $\mathcal{F}_s^{W_P}$ measurable and

$$\mathcal{F}_{s-\tau}^{B_P} \subseteq \mathcal{F}_s^{W_P}.$$

Also we have that $(\mathcal{F}_s^{W_P})_{s\geq 0}$ is a filtration so

$$\mathcal{F}_{s-\tau}^{B_P} \vee \mathcal{F}_\tau^{W_P} \subseteq \mathcal{F}_s^{W_P}.$$

Now notice that $W_P(s) = B_P(s-\tau) + W_P(\tau)$ and that $B_P(s-\tau)$ is independent of $W_P(\tau)$ then

$$\mathcal{F}_s^{W_P} \subseteq \mathcal{F}_{s-\tau}^{B_P} \vee \mathcal{F}_\tau^{W_P}.$$

We have just shown that $\mathcal{F}_s^{W_P} = \mathcal{F}_\tau^{W_P} \vee \mathcal{F}_{s-\tau}^{B_P}$ when $s > \tau$ and hence $\tilde{\mathcal{F}}_s = \mathcal{F}_{s-\tau}^*$.

From equation (3.4.3) we have that the random variable $X(t) = Z(t-\tau)$ when $t > \tau$. Let us define the process Y^* as

$$Y^*(t) = (Z(t), I(t)) \text{ for } t \geq 0.$$

The process Y^* is $(\mathcal{F}_t^*)_{t\geq 0}$-adapted and it is the strong solution of the following system of stochastic differential equations:

$$dZ(t) \;=\; \mu dt + \sigma_P \sqrt{I(t)}dB_P(t) \text{ with } Z(0) = X(\tau), \tag{3.4.4}$$
$$dI(t) \;=\; a_I(b_I - I(t))dt + \sigma_I \sqrt{I(t)}dW_I(t) \text{ with } I(0) = \phi^I(0). \tag{3.4.5}$$

Since the process Y^* is the strong solution of a system of stochastic differential equations, it is a Markov process with respect to the filtration $(\mathcal{F}_t^*)_{t\geq 0}$ (Karatzas & Shreve, 1998, Theorem 5.4.20). We have that $Y(t) = Y^*(t-\tau) = (X(t), I(t-\tau))$ when $t > \tau$.

We have seen that $\mathcal{F}_t^{W_P} = \mathcal{F}_\tau^{W_P} \vee \mathcal{F}_{t-\tau}^{B_P}$ when $t > \tau$, hence we have that $\tilde{\mathcal{F}}_t = \mathcal{F}_{t-\tau}^*$.

We have just shown that $(Y^*(t-\tau))_{t\geq\tau}$ has the Markov property under the filtration $(\tilde{\mathcal{F}}_t)_{t\geq\tau}$. So $Y(t)$ has the Markov property under $(\tilde{\mathcal{F}}_t)_{t\geq\tau}$. □

3.5 Conditional independence of the logarithmic returns

In some instances, we will be interested in computing the distribution of the logarithmic returns given the interest process. Since the market attention is observed this conditional distribution will help us to estimate the parameters. Given the time $H > 0$ we are interested in seeing if the logarithmic return $R(s,t) = X(t) - X(s)$ with $H \geq t > s \geq 0$ when conditioned to the σ-algebra $\mathcal{F}_H^{W_I}$, is independent of $\mathcal{F}_s$.

Proposition 3.5.1. *Let $\lambda \in \mathbb{R}$ and let $H \geq t > s \geq 0$. Then:*

$$E\left[e^{i\lambda R(s,t)} \middle| \mathcal{F}_s \vee \mathcal{F}_H^{W_I}\right] = E\left[e^{i\lambda R(s,t)} \middle| \mathcal{F}_H^{W_I}\right]. \tag{3.5.1}$$

Proof. Notice that $R(s,t)$ can be written as

$$R(s,t) = \mu(t-s) + \int_s^t \sigma_P \sqrt{I(u-\tau)} dW^P(u).$$

By application of Proposition 1.4.2, the left hand side of equation (3.5.1) can be written as

$$\begin{aligned} E\left[e^{i\lambda R(s,t)} \middle| \mathcal{F}_s \vee \mathcal{F}_H^{W_I}\right] &= e^{i\lambda\mu(t-s)} E\left[e^{i\lambda \int_s^t \sigma_P \sqrt{I(u-\tau)} dW^P(u)} \middle| \mathcal{F}_s \vee \mathcal{F}_H^{W_I}\right] \\ &= e^{i\lambda\mu(t-s)} e^{-\lambda^2 \frac{\sigma_P^2}{2} \int_s^t I(u-\tau) du}. \end{aligned}$$

In the case of the right hand side of equation (3.5.1) we have from Proposition 1.4.2:

$$\begin{aligned} E\left[e^{i\lambda R(s,t)} \middle| \mathcal{F}_H^{W_I}\right] &= e^{i\lambda\mu(t-s)} E\left[e^{i\lambda \int_s^t \sigma_P \sqrt{I(u-\tau)} dW^P(u)} \middle| \mathcal{F}_H^{W_I}\right] \\ &= e^{i\lambda\mu(t-s)} e^{-\lambda^2 \frac{\sigma_P^2}{2} \int_s^t I(u-\tau) du}. \end{aligned}$$

□

Remark 3.5.1. *From Proposition 3.5.1 it is clear that the logarithmic returns are independent given the σ-algebra $\mathcal{F}_H^{W_I}$.*

3.6 Estimation

Because it is not possible to obtain continuous observations in the real world, we will have discrete observations of the processes X and I. Using these observations, we would like to estimate the parameters that appear in the system (3.2.1)–(3.2.2).

We will assume that there are $N+1 \in \mathbb{N}$ observations equally spaced in time, with a time step $\Delta = \frac{H}{N}$ and a time horizon H. The lag parameter τ can only take a finite number of non-negative values, that is $\tau \in \{0, \Delta, 2\Delta, \ldots, M\Delta\}$ where $M \in \mathbb{N}$ with $N \gg M$ and we assume that there exists a $k \in \{0, 1, 2, \ldots, M\}$ such that $\tau = k\Delta$. We also define $L = M\Delta$.

Suppose that we have the following discrete and equally spaced in time observations $\{x_j\}_{j=0}^N$, $\{y_j\}_{j=-M}^N$for the processes X and I respectively, where x_j is the observation of the variable $X(t_j)$ and y_j is the observation of $I(t_j)$ with $t_j = j\Delta$. The horizon H can be expressed as $H = N\Delta = t_N$.

Remember that when $t \in [-L, 0]$ we have that $I(t) = \phi^I(t)$. For the deterministic function ϕ^I we approximate it by

$$\hat{\phi}^I(t) = y_j + (t - t_j)\frac{y_{j+1} - y_j}{t_{j+1} - t_j} \text{ when } t \in [t_j, t_{j+1}] \text{ and } j = -M, \dots, -1. \quad (3.6.1)$$

The approximation function $\hat{\phi}^I$ is just linear interpolations of the observations of the process I when $t \in [-L, 0]$. We use the approximation defined in equation (3.6.1) to approximate the integrals:

$$\int_{t_j}^{t_{j+1}} \phi^I(s - \tau)ds \approx \int_{t_j}^{t_{j+1}} \hat{\phi}^I(s - \tau)ds \text{ when } t \in [t_j, t_{j+1}] \text{ and } j = -M, \dots, -1.$$

For the estimation of the parameters of the system (3.2.1)–(3.2.2), we first estimate the parameters of the interest process and then we estimate the parameters that appear in equation (3.2.1). For the estimation of the parameters related to the market attention process, we will use the maximum likelihood estimation method explained in Chapter 2. In the case of the parameters that appear in equation (3.2.1) we will use the conditional likelihood estimation method.

It is possible to estimate all the parameters that appear on the system of equations (3.2.1)–(3.2.2) together, using the maximum likelihood estimation method. Since we have an affine model, it is possible to compute the joint characteristic function of the random variables $X(t)$ and $I(t-\tau)$. Once the characteristic function is computed, it is possible to recover the density function. The problem is that we need to approximate a double integral.

3.6.1 Conditional likelihood estimator

In this method, we will estimate the parameters of the system (3.2.1)–(3.2.2) in a two-step approach. Since we have the observations of the interest process, we can use the methods explained in Chapter 2 to estimate the parameters that are related to the market attention process. For the parameters that appear in the price equation (3.2.1) we will use the conditional likelihood estimation method (see Section 1.5.1).

From now on we will focus on the estimation of the parameters (μ, σ_P). We will assume that the lag parameter τ is given. The parameter τ will be estimated in Section 3.6.2. In this case, we will use the conditional likelihood method for estimating the parameters μ and σ_P. This method is also used by Cretarola and Figà-Talamanca (2021).

To perform this estimation, we will use the logarithmic returns of the price. Because of that, we define the vector of the logarithmic returns R as

$$R = (R(t_0, t_1), R(t_1, t_2), \dots R(t_{N-1}, t_N)),$$

where

$$R(t_j, t_{j+1}) = X(t_{j+1}) - X(t_j) \text{ for } j = 0, \dots, N-1.$$

Let us define

$$r_{j+1} = x_{j+1} - x_j \text{ for } j = 0, 1, \dots, N-1$$

as the realizations of the logarithmic returns.

The variable $R(t_j, t_{j+1})$ can be expressed as:

$$R(t_j, t_{j+1}) = \int_{t_j}^{t_{j+1}} \mu du + \int_{t_j}^{t_{j+1}} \sigma_P \sqrt{I(u - \tau)} dW_P(u),$$

for $j = 0, \dots, N-1$. In light of Proposition 1.4.2 we know that

$$f^R_{j+1}\left(r_{j+1}|\mathcal{F}^{W_I}_H\right) = f_N(r_{j+1}|\mu\Delta, v^{2,j}_R) \text{ for } j = 0, \dots, N-1, \tag{3.6.2}$$

where $f_N(.|\mu\Delta, v^{2,j}_R)$ is the density function of a normal random variable with mean $\mu\Delta$ and variance $v^{2,j}_R$ that is defined by

$$v^{2,j}_R \quad = \quad \sigma^2_P \int_{t_j}^{t_{j+1}} I(u-\tau)du.$$

Using Remark 3.5.1 we can express the conditional density of $\{R(t_j, t_{j+1})\}_{j=0}^{N-1}$ given $\mathcal{F}^{W_I}_H$ as

$$f^R_{1:N}\left(r_{1:N}|\mathcal{F}^{W_I}_H\right) = \prod_{j=0}^{N-1} f^R_{j+1}\left(r_{j+1}|\mathcal{F}^{W_I}_H\right). \tag{3.6.3}$$

Let us define the random vector J_I as

$$J_I = \left(J^{t_0}_I, J^{t_1}_I, \dots, J^{t_{N-1}}_I\right),$$

where

$$J^{t_j}_I \quad = \quad \int_{t_j}^{t_{j+1}} I(u-\tau)du \text{ for } j = 0, 1 \dots N-1.$$

Proposition 3.6.1. *Let $u_j \in \mathbb{R}$ for $j = 0, \dots, N-1$ and let $\sigma(J_I)$ be the σ-algebra generated by the vector of random variables J_I. Then:*

$$\begin{aligned} E\left[e^{\sum_{j=0}^{N-1} iu_j R(t_j,t_{j+1})}\middle|\mathcal{F}^{W_I}_H\right] &= E\left[e^{\sum_{j=0}^{N-1} iu_j R(t_j,t_{j+1})}\middle|\sigma(J_I)\right] \\ &= \prod_{j=0}^{N-1} e^{iu_j\mu\Delta - \frac{1}{2}u_j^2 v^{2,j}_R}. \end{aligned} \tag{3.6.4}$$

A particular case of equation (3.6.4) is

$$\begin{aligned} E\left[e^{iu_j R(t_j,t_{j+1})}\middle|\mathcal{F}^{W_I}_H\right] &= E\left[e^{iu_j R(t_j,t_{j+1})}\middle|\sigma(J_I)\right] \\ &= e^{iu_j\mu\Delta - \frac{1}{2}u_j^2 v^{2,j}_R} \ \textit{for } j = 0, 1, \dots, N-1. \end{aligned}$$

Proof. Notice that from equation (3.6.2) we have that

$$E\left[e^{iu_j R(t_j,t_{j+1})}\middle|\mathcal{F}^{W_I}_H\right] = e^{iu_j\mu\Delta - \frac{1}{2}u_j^2 v^{2,j}_R} \text{ for } j = 0, 1, \dots, N-1. \tag{3.6.5}$$

Notice that the random variable J_I is measurable with respect to the σ-algebra $\mathcal{F}^{W_I}_H$, hence $\sigma(J_I) \subseteq \mathcal{F}^{W_I}_H$. By the tower property we have that:

$$E\left[e^{\sum_{j=0}^{N-1} iu_j R(t_j,t_{j+1})}\middle|\sigma(J_I)\right] = E\left[E\left[e^{\sum_{j=0}^{N-1} iu_j R(t_j,t_{j+1})}\middle|\mathcal{F}^{W_I}_H\right]\middle|\sigma(J_I)\right]. \tag{3.6.6}$$

Using the conditional independence of logarithmic returns (Remark 3.5.1) we get

$$E\left[e^{\sum_{j=0}^{N-1} iu_j R(t_j,t_{j+1})} \middle| \mathcal{F}_H^{W_I} \right] = \prod_{j=0}^{N-1} E\left[e^{iu_j R(t_j,t_{j+1})} \middle| \mathcal{F}_H^{W_I} \right]. \tag{3.6.7}$$

From the result in (3.6.5) we obtain:

$$E\left[e^{\sum_{j=0}^{N-1} iu_j R(t_j,t_{j+1})} \middle| \mathcal{F}_H^{W_I} \right] = \prod_{j=0}^{N-1} e^{iu_j\mu\Delta - \frac{1}{2}u_j^2 v_R^{2,j}}. \tag{3.6.8}$$

Substituting equation (3.6.8) into equation (3.6.6), we have that

$$\begin{aligned} E\left[e^{\sum_{j=0}^{N-1} iu_j R(t_j,t_{j+1})} \middle| \sigma(J_I) \right] &= E\left[\prod_{j=0}^{N-1} e^{iu_j\mu\Delta - \frac{1}{2}u_j^2 v_R^{2,j}} \middle| \sigma(J_I) \right] \\ &= \prod_{j=0}^{N-1} e^{iu_j\mu\Delta - \frac{1}{2}u_j^2 v_R^{2,j}}, \end{aligned}$$

where the last equality comes from the fact that $v_R^{2,j}$ is measurable with respect to the σ-algebra $\sigma(J_I)$. □

From Proposition 3.6.1 we have that

$$f_{j+1}^R\left(r_{j+1}|\mathcal{F}_H^{W_I}\right) = f_{j+1}^R\left(r_{j+1}|\sigma\left(J_I\right)\right) \text{ for } j = 0, 1, \ldots, N-1,$$

and we also have that

$$f_{1:N}^R\left(r_{1:N}|\mathcal{F}_H^{W_I}\right) = f_{1:N}^R\left(r_{1:N}\,|\sigma\left(J_I\right)\right).$$

The random vector R depends on the parameter $\theta_R = (\lambda_R, \xi_R)$, where

$$\lambda_R = (\mu, \sigma_P), \qquad \xi_I = (a_I, b_I, \sigma_I).$$

The parameter of interest is the vector λ_R and ξ_I is the vector of nuisance parameters. We will assume the vector J_I to be a sufficient statistic for the parameter ξ_I.

The conditional log-likelihood is

$$\begin{aligned} l^R\left(\lambda_R\,|\sigma\left(J_I\right),\tau\right) &= \log\left(f_{1:N}^R\left(r_{1:N}\,|\sigma\left(J_I\right)\right)\right) \\ &= \sum_{j=0}^{N-1} \log f_{j+1}^R\left(r_{j+1}|\sigma\left(J_I\right)\right) \\ &= \sum_{j=0}^{N-1} \log f_N(r_{j+1}|\mu\Delta, v_R^{2,j}). \end{aligned} \tag{3.6.9}$$

We will maximize the function $l^R\left(.\,|\sigma\left(J_I\right),\tau\right)$ to estimate the parameter λ_R.

For computing the function $l^R\left(.\,|\sigma\left(J_I\right),\tau\right)$ we need to know the value of the Lebesgue integral that appears in $v_R^{2,j}$. This would not be a problem if the process I was continuously observed, but because the process I is discretely observed we need to estimate those integrals. The integrals are estimated by the trapezoid method:

$$
\begin{aligned}
J_I^{t_j} = \int_{t_j}^{t_{j+1}} I(u-\tau)du &\approx \int_{t_j}^{t_{j+1}} \left(y_{j-k} + (u-t_j)\frac{y_{j+1-k}-y_{j-k}}{t_{j+1}-t_j} \right) du \\
&= \frac{\Delta}{2}(y_{j+1-k}+y_{j-k}).
\end{aligned}
$$

The approximation is based in integrating the linear interpolation between the observations at times t_j and t_{j+1}.

Due to how we define the approximation function $\hat{\phi}^I$ in (3.6.1). We do not need to differentiate between the returns at times before τ and after τ, because the conditional distribution of the returns satisfies equation (3.6.2) independently if $t_{j+1} \leq t_k = \tau$ or $t_{j+1} > t_k = \tau$.

It is possible to obtain an analytical expression for the maximum likelihood estimators of μ and σ_P.

Proposition 3.6.2. *The maximum likelihood estimates of μ and σ_P are*

$$
\begin{aligned}
\hat{\mu} &= \frac{\sum_{j=0}^{N-1} \frac{r_j}{J_I^{t_j}}}{\Delta \sum_{j=0}^{N-1} \frac{1}{J_I^{t_j}}}, \\
\hat{\sigma}_P &= \sqrt{\frac{1}{N}\sum_{j=0}^{N-1}\left(\frac{r_j - \hat{\mu}\Delta}{\sqrt{J_I^{t_j}}}\right)^2}.
\end{aligned}
$$

Proof. From equation (3.6.9) we have that

$$
\begin{aligned}
l^R\left(\lambda_R \left| \sigma\left(J_I\right), \tau\right.\right) &= \sum_{j=0}^{N-1}\left(-\log\left(\sigma_P\sqrt{J_I^{t_j}2\pi}\right) - \frac{1}{2}\left(\frac{r_j-\mu\Delta}{\sigma_P\sqrt{J_I^{t_j}}}\right)^2 \right) \\
&= -\sum_{j=0}^{N-1}\log\sigma_P - \sum_{j=0}^{N-1}\log\sqrt{J_I^{t_j}2\pi} \\
&\quad -\frac{1}{2}\sum_{j=0}^{N-1}\left(\frac{r_j^2+\mu^2\Delta^2-2r_j\mu\Delta}{\sigma_P^2 J_I^{t_j}}\right) \\
&= -N\log\sigma_P - \sum_{j=0}^{N-1}\log\sqrt{J_I^{t_j}2\pi} \\
&\quad -\frac{1}{2\sigma_P^2}\sum_{j=0}^{N-1}\frac{r_j^2}{J_I^{t_j}} - \frac{\mu^2\Delta^2}{2\sigma_P^2}\sum_{j=0}^{N-1}\frac{1}{J_I^{t_j}} + \frac{\mu\Delta}{\sigma_P^2}\sum_{j=0}^{N-1}\frac{r_j}{J_I^{t_j}}.
\end{aligned}
$$

Let us define $z = \sigma_P^2$, so the conditional log-likelihood can be rewritten as

$$
\begin{aligned}
l^R\left(\lambda_R \left| \sigma\left(J_I\right), \tau\right.\right) &= -\frac{N}{2}\log z - \sum_{j=0}^{N-1}\log\sqrt{J_I^{t_j}2\pi} \\
&\quad -\frac{1}{2z}\sum_{j=0}^{N-1}\frac{r_j^2}{J_I^{t_j}} - \frac{\mu^2\Delta^2}{2z}\sum_{j=0}^{N-1}\frac{1}{J_I^{t_j}} + \frac{\mu\Delta}{z}\sum_{j=0}^{N-1}\frac{r_j}{J_I^{t_j}}.
\end{aligned}
$$

Let us now take the partial derivatives with respect to μ and z:

$$
\begin{aligned}
\frac{\partial l^R\left(\lambda_R \left|\sigma\left(J_I\right), \tau\right.\right)}{\partial \mu} &= -\frac{\mu\Delta^2}{z}\sum_{j=0}^{N-1}\frac{1}{J_I^{t_j}} + \frac{\Delta}{z}\sum_{j=0}^{N-1}\frac{r_j}{J_I^{t_j}} \\
&= \frac{1}{z}\left(-\mu\Delta^2\sum_{j=0}^{N-1}\frac{1}{J_I^{t_j}} + \Delta\sum_{j=0}^{N-1}\frac{r_j}{J_I^{t_j}}\right), \\
\frac{\partial l^R\left(\lambda_R \left|\sigma\left(J_I\right), \tau\right.\right)}{\partial z} &= -\frac{N}{2z} + \frac{1}{2z^2}\sum_{j=0}^{N-1}\frac{r_j^2}{J_I^{t_j}} + \frac{\mu^2\Delta^2}{2z^2}\sum_{j=0}^{N-1}\frac{1}{J_I^{t_j}} - \frac{\mu\Delta}{z^2}\sum_{j=0}^{N-1}\frac{r_j}{J_I^{t_j}} \\
&= \frac{1}{z^2}\left(-\frac{N}{2}z + \frac{1}{2}\sum_{j=0}^{N-1}\frac{r_j^2}{J_I^{t_j}} + \frac{\mu^2\Delta^2}{2}\sum_{j=0}^{N-1}\frac{1}{J_I^{t_j}} - \mu\Delta\sum_{j=0}^{N-1}\frac{r_j}{J_I^{t_j}}\right).
\end{aligned}
$$

To get the critical values we have to equate the partial derivatives to zero. Since $z > 0$ it is sufficient to solve the following equations

$$
-\mu\Delta^2\sum_{j=0}^{N-1}\frac{1}{J_I^{t_j}} + \Delta\sum_{j=0}^{N-1}\frac{r_j}{J_I^{t_j}} = 0, \tag{3.6.10}
$$

$$
-\frac{N}{2}z + \frac{1}{2}\sum_{j=0}^{N-1}\frac{r_j^2}{J_I^{t_j}} + \frac{\mu^2\Delta^2}{2}\sum_{j=0}^{N-1}\frac{1}{J_I^{t_j}} - \mu\Delta\sum_{j=0}^{N-1}\frac{r_j}{J_I^{t_j}} = 0. \tag{3.6.11}
$$

For equation (3.6.10) we have that a possible estimate for μ is

$$
\hat{\mu} = \frac{\sum_{j=0}^{N-1}\frac{r_j}{J_I^{t_j}}}{\Delta\sum_{j=0}^{N-1}\frac{1}{J_I^{t_j}}}. \tag{3.6.12}
$$

Also from equation (3.6.11) we get that a possible estimate for z is

$$
\begin{aligned}
\hat{z} &= \frac{2}{N}\left(\frac{1}{2}\sum_{j=0}^{N-1}\frac{r_j^2}{J_I^{t_j}} + \frac{\hat{\mu}^2\Delta^2}{2}\sum_{j=0}^{N-1}\frac{1}{J_I^{t_j}} - \hat{\mu}\Delta\sum_{j=0}^{N-1}\frac{r_j}{J_I^{t_j}}\right) \\
&= \frac{1}{N}\sum_{j=0}^{N-1}\left(\frac{r_j - \hat{\mu}\Delta}{\sqrt{J_I^{t_j}}}\right)^2.
\end{aligned} \tag{3.6.13}
$$

So, a possible estimate for σ_P is

$$
\hat{\sigma}_P = \sqrt{\frac{1}{N}\sum_{j=0}^{N-1}\left(\frac{r_j - \hat{\mu}\Delta}{\sqrt{J_I^{t_j}}}\right)^2}. \tag{3.6.14}
$$

The Hessian matrix of the conditional log-likelihood is the following

$$\begin{pmatrix} \frac{\partial^2 l^R(\lambda_R|\sigma(J_I),\tau)}{\partial\mu^2} & \frac{\partial^2 l^R(\lambda_R|\sigma(J_I),\tau)}{\partial\mu\partial z} \\ \frac{\partial^2 l^R(\lambda_R|\sigma(J_I),\tau)}{\partial z\partial\mu} & \frac{\partial^2 l^R(\lambda_R|\sigma(J_I),\tau)}{\partial z^2} \end{pmatrix}$$
$$= \begin{pmatrix} -\frac{\Delta^2}{z}\sum_{j=1}^{N-1}\frac{1}{J_I^{t_j}} & \frac{1}{z^2}\left(\mu\Delta^2\sum_{j=0}^{N-1}\frac{1}{J_I^{t_j}} - \Delta\sum_{j=0}^{N-1}\frac{r_j}{J_I^{t_j}}\right) \\ \frac{1}{z^2}\left(\mu\Delta^2\sum_{j=0}^{N-1}\frac{1}{J_I^{t_j}} - \Delta\sum_{j=0}^{N-1}\frac{r_j}{J_I^{t_j}}\right) & \frac{1}{z^3}\left(-\frac{1}{2}\sum_{j=1}^{N-1}\left(\frac{r_j-\mu\Delta}{\sqrt{J_I^{t_j}}}\right)^2\right) \end{pmatrix}.$$

Substituting the possible estimates for μ and z defined in equation (3.6.12) and (3.6.13) respectively into the Hessian matrix, we obtain the following

$$\begin{pmatrix} -\frac{\Delta^2}{\hat{z}}\sum_{j=1}^{N-1}\frac{1}{J_I^{t_j}} & 0 \\ 0 & \frac{1}{\hat{z}^3}\left(-\frac{1}{2}\sum_{j=1}^{N-1}\left(\frac{r_j-\hat{\mu}\Delta}{\sqrt{J_I^{t_j}}}\right)^2\right) \end{pmatrix}. \tag{3.6.15}$$

The eigenvalues of the matrix (3.6.15) are

$$\begin{aligned} \lambda_1 &= -\frac{\Delta^2}{\hat{z}}\sum_{j=1}^{N-1}\frac{1}{J_I^{t_j}}, \\ \lambda_2 &= -\frac{1}{2\hat{z}^3}\sum_{j=1}^{N-1}\left(\frac{r_j-\hat{\mu}\Delta}{\sqrt{J_I^{t_j}}}\right)^2. \end{aligned}$$

Since both eigenvalues are negative we have that $\hat{\mu}$ and $\hat{z}$ maximize the conditional log-likelihood function. □

3.6.2 Model Selection

Since τ can only take a finite number of values, remember that $\tau \in \{0, \Delta, 2\Delta, \ldots, M\Delta = L\}$. The problem of the estimation of τ can be reduced to a model selection problem. That is, we would select among the following different models for the price data:

$$X(t) = X(0) + \int_0^t \mu du + \int_0^t \sigma_P\sqrt{I(u-\Delta r)}dW_P(u),$$

for $r = 0, 1, \ldots, M$. Since the process $(I(t))_{t\geq 0}$ is given, we can use the conditional likelihood function defined in Section 3.6.1. According to the literature (deLeeuw, 1992, p. 605), we should select the model that minimizes the Akaike information criterion, which is defined as

$$AIK_r = 2q - 2l^R\left(\hat{\lambda}_R\,|\sigma\left(J_I\right), \tau = r\Delta\right) \text{ for } r = 0, 1, \ldots M, \tag{3.6.16}$$

where q is the number of parameters of the model and $\hat{\lambda}_R = (\hat{\mu}, \hat{\sigma}_P)$ is defined as in Proposition 3.6.2. Another method consists of minimizing the Bayesian information criterion (Neath & Cavanaugh, 2012), defined as

$$BIC_r = q\log(N) - 2l^R\left(\hat{\lambda}_R\,|\sigma\left(J_I\right), \tau = r\Delta\right) \text{ for } r = 0, 1, \dots M. \tag{3.6.17}$$

Since the number of parameters is $q = 2$ for all $r = 0, 1, \dots M$ then, minimizing the Akaike information criterion and the Bayesian information criterion is equivalent to maximizing the following expression

$$l^R\left(\hat{\lambda}_R\,|\sigma\left(J_I\right), \tau = r\Delta\right) \text{ for } r = 0, 1, \dots M.$$

That is, we select the model that maximizes the conditional log-likelihood.

3.6.3 Numerical experiments

For the evaluation of the techniques presented in Section 3.6.1 and in Section 3.6.2, we perform the following experiment. We generate 1000 realizations of $N = 1000$ steps of the processes defined in equations (3.2.1)–(3.2.2) with the parameters

$$a_I = 0.5, \quad b_I = 0.06, \quad \sigma_I = 0.15, \quad \mu = 0.2, \quad \sigma_P = 2, \quad \tau = 10\Delta,$$

with time step $\Delta = 0.1$, $L = 20\Delta$ and initial values $P(0) = 100$ and

$$\phi^I(t) = 0.05 + 0.1\cos^2\left(4t + \frac{\pi}{2}\right) \text{ for } t \in [-L, 0].$$

For each of the realizations we estimate the parameters of the model using the techniques explained in Section 3.6. Once we estimate the parameters of all of the simulations, we compute the mean of the estimated parameters and its standard deviation. The results are shown in Table 3.1. The method gives good estimates for the parameters, especially for those that appear in the price equation (3.2.1).

Parameter	True value	Mean	Standard deviation
a_I	0.5	0.540466	0.110240
b_I	0.06	0.060265	0.007219
σ_I	0.15	0.150133	0.003414
μ	0.2	0.199620	0.03945
σ_P	2	2.000897	0.045047
τ	10 Δ	10.013 Δ	0.304682 Δ

Table 3.1: Mean and the standard deviation of the estimated parameters using the estimation method explained in Section 3.6 for 1000 realizations.

3.7 Option pricing

In this section, we will be interested in pricing plain vanilla options with strike price K and expiration date T. That is, we are interested in pricing European call options with payoff $(P(T)-K)^+$ and European put options with payoff $(K-P(T))^+$. Because the expiration date of the options is T we assume that $t \in [0, T]$ and $\mathcal{F} = \mathcal{F}_T$.

3.7.1 Change of measure

For pricing options, we would like to use a risk-neutral measure. To obtain a risk-neutral measure, let us define the following processes

$$\begin{aligned} W_P^*(t) &= W_P(t) + \int_0^t \theta_P(s)ds \text{ for } t \in [0,T], \\ W_I^*(t) &= W_I(t) + \int_0^t \theta_I(s)ds \text{ for } t \in [0,T] \end{aligned}$$

and

$$\begin{aligned} Z(t) &= \exp\left\{-\int_0^t \theta_P(s)dW_P(s) - \int_0^t \theta_I(s)dW_I(s)\right\} \\ &\quad \exp\left\{-\frac{1}{2}\int_0^t \theta_P^2(s)ds - \frac{1}{2}\int_0^t \theta_I^2(s)ds\right\} \text{ for } t \in [0,T], \end{aligned}$$

where $(\theta_P(t))_{t\in[0,T]}$ and $(\theta_I(t))_{t\in[0,T]}$ are two adapted processes with respect to $(\mathcal{F}_t)_{t\in[0,T]}$. If the process $(Z(t))_{t\in[0,T]}$ is a martingale with respect to $(\mathcal{F}_t)_{t\in[0,T]}$ then we can apply the Girsanov theorem. If Z is a martingale then by the Girsanov theorem $(W_P^*(t), W_I^*(t))_{t\in[0,T]}$ is a two dimensional Brownian motion under the measure $\mathbb{Q}$, where

$$\mathbb{Q}(A) = \int_A Z(T)d\mathbb{P} \ \forall A \in \mathcal{F}.$$

We will show that under the appropriate choice of θ_P and θ_I the process Z is a martingale.

To have a risk-neutral measure, we need to ensure that the discounted price process under the new measure $\mathbb{Q}$ is a martingale. If the price process under $\mathbb{Q}$ satisfies the following stochastic differential equation

$$dP(t) = rP(t)dt + \sigma_P P(t)\sqrt{I(t-\tau)}dW_P^*(t) \tag{3.7.1}$$

then it is possible to show that the discounted price process is a martingale.

To obtain equation (3.7.1), different values for the adapted processes θ_P and θ_I can be taken, implying that there are infinitely many risk-neutral measures. In this case, we choose the following values for the processes θ_P and θ_I:

$$\theta_P(t) = \frac{\mu + \frac{\sigma_P^2}{2}I(t-\tau) - r}{\sigma_P\sqrt{I(t-\tau)}}, \tag{3.7.2}$$

$$\theta_I(t) = \frac{\lambda_1}{\sigma_I\sqrt{I(t)}} + \frac{\lambda_2}{\sigma_I}\sqrt{I(t)}, \tag{3.7.3}$$

where $\lambda_1, \lambda_2 \in \mathbb{R}$.

We chose θ_P and θ_I as in equations (3.7.2) and (3.7.3) respectively, for the following reasons:

1. It is possible to obtain the risk-neutral price equation (3.7.1).

2. The interest is still a Cox-Ingersoll-Ross process under the measure $\mathbb{Q}$.

To obtain equations (3.7.4) and (3.7.5), we need to prove that $(Z(t))_{t\in[0,T]}$ is a martingale with respect to $(\mathcal{F}_t)_{t\in[0,T]}$.

If we assume for a moment that $(Z(t))_{t\in[0,T]}$ is a martingale, then we can apply the Girsanov theorem, and under the measure $\mathbb{Q}$ we have the following system of stochastic differential equations

$$\begin{aligned} dP(t) &= rP(t)dt + \sigma_P P(t)\sqrt{I(t-\tau)}dW_P^*(t) \text{ with } P(0) = p \in \mathbb{R}_+, && (3.7.4)\\ dI(t) &= (a_I b_I - \lambda_1 - (a_I + \lambda_2)I(t))dt + \sigma_I\sqrt{I(t)}dW_I^*(t) && (3.7.5)\\ \text{if} \quad & t \in [0,T] \text{ and } I(t) = \phi^I(t) \text{ when } t \in [-L,0]. \end{aligned}$$

Under the probability measure $\mathbb{Q}$ equation (3.7.5) can be rewritten as

$$\begin{aligned} dI(t) &= \tilde{a}_I(\tilde{b}_I - I(t))dt + \sigma_I\sqrt{I(t)}dW_I^*(t) \text{ when } t \in [0,T] && (3.7.6)\\ \text{and} \quad & I(t) = \phi^I(t) \text{ when } t \in [-L,0], \end{aligned}$$

where $\tilde{a}_I = a_I + \lambda_2$ and $\tilde{b}_I = \frac{a_I b_I - \lambda_1}{a_I + \lambda_2}$. If we impose the conditions

$$\begin{aligned} \frac{2\tilde{a}_I\tilde{b}_I}{\sigma_I^2} &\geq 1 && (3.7.7)\\ \tilde{a}_I &> 0, && (3.7.8) \end{aligned}$$

then equation (3.7.6) has a strong solution (Gulisashvili, 2012, p. 44) and it is greater than zero almost surely (Gulisashvili, 2012, Theorem 2.27).

Theorem 3.7.1. *If conditions (3.7.7) and (3.7.8) are satisfied, then $(Z(t))_{t\in[0,T]}$ is a martingale with respect to $(\mathcal{F}_t)_{t\in[0,T]}$.*

Proof. We know that $(Z(t))_{t\in[0,T]}$ is a non-negative local martingale because Z is the stochastic exponential of a local martingale (Klebaner, 2012, p. 227). Hence it is a supermartingale (Klebaner, 2012, Theorem 7.23). Because of that it is enough to prove that $E[Z(t)] = Z(0) = 1$. Conditioning with respect to $\mathcal{F}_t^{W_I}$ and using Proposition 1.4.1, we have

$$\begin{aligned} & E[Z(t)] \\ =& E\left[\exp\left\{-\int_0^t \theta_P(s)dW_P(s) - \frac{1}{2}\int_0^t \theta_P^2(s)ds - \int_0^t \theta_I(s)dW_I(s) - \frac{1}{2}\int_0^t \theta_I^2(s)ds\right\}\right] \\ =& E\left[\exp\left\{-\int_0^t \theta_I(s)dW_I(s) - \frac{1}{2}\int_0^t \theta_I^2(s)ds - \frac{1}{2}\int_0^t \theta_P^2(s)ds\right\}\right. \\ & \quad \left. E\left[\exp\left\{-\int_0^t \theta_P(s)dW_P(s)\right\}\middle| \mathcal{F}_t^{W_I}\right]\right] \\ =& E\left[\exp\left\{-\int_0^t \theta_I(s)dW_I(s) - \frac{1}{2}\int_0^t \theta_I^2(s)ds\right\}\right]. \end{aligned}$$

Now define the process Z' as

$$Z'(t) = \exp\left\{\int_0^t \tilde{\theta}_I(s)dW_I(s) - \frac{1}{2}\int_0^t \tilde{\theta}_I^2(s)ds\right\} \text{ for } t \in [0,T],$$

where $\tilde{\theta}_I = -\theta_I$. Because conditions (3.7.7) and (3.7.8) are satisfied, then by Proposition 1.4.3 we know that Z' is a martingale. So, we have just shown that

$$E[Z(t)] = E[Z'(t)] = Z'(0) = 1.$$

□

3.7.2 Pricing call options

In Section 3.7.1 we proved the existence of a risk-neutral probability $\mathbb{Q}$ under which we have the following dynamics for the Bitcoin price:

$$\begin{aligned} dP(t) &= rP(t)dt + \sigma_P P(t)\sqrt{I(t-\tau)}dW_P^*(t) &\quad (3.7.9)\\ &\quad \text{with } P(0) = p \in \mathbb{R}_+ = (0,\infty), \\ dI(t) &= \tilde{a}_I(\tilde{b}_I - I(t))dt + \sigma_I\sqrt{I(t)}dW_I^*(t) \text{ when } t \in [0,T] &\quad (3.7.10)\\ &\quad \text{and } I(t) = \phi^I(t)\ t \in [-L,0], \end{aligned}$$

where $\tilde{b}_I \in \mathbb{R}$, $\tau \in [0,L]$, $r \geq 0$ and $L, \sigma_P, \sigma_I, \tilde{a}_I > 0$ with the condition $\frac{2\tilde{a}_I\tilde{b}_I}{\sigma_I^2} \geq 1$. It is possible to show that the discounted process $\tilde{P}$, that is defined as $\tilde{P}(t) = e^{-rt}P(t)$ for $t \in [0,T]$, is a martingale with respect to the filtration $(\mathcal{F}_t)_{t\in[0,T]}$ under the risk-neutral measure $\mathbb{Q}$. That is

$$E_{\mathbb{Q}}\left[\tilde{P}(t)\,|\mathcal{F}_s\right] = \tilde{P}(s) \text{ for } T \geq t \geq s,$$

where $E_{\mathbb{Q}}$ symbolizes the expected value under the measure $\mathbb{Q}$.

Proposition 3.7.1. *The discounted price process $\tilde{P}$ is a martingale with respect to the filtration $(\mathcal{F}_t)_{t\in[0,T]}$ on the probability space $(\Omega, \mathcal{F}, \mathbb{Q})$.*

Proof. Applying the Itô formula to $\left(e^{-rt}P(t)\right)_{t\in[0,T]}$ we have

$$d\tilde{P}(t) = \sigma_P\tilde{P}(t)\sqrt{I(t-\tau)}dW_P^*(t) \text{ with } \tilde{P}(0) = p > 0. \quad (3.7.11)$$

From equation (3.7.11) we know that $\tilde{P}$ is a local martingale (Klebaner, 2012, Remark 4.6). Using the Itô formula it is possible to show that

$$\tilde{P}(t) = \tilde{P}(0)\exp\left\{-\frac{\sigma_P^2}{2}\int_0^t I(u-\tau)du + \sigma_P\int_0^t \sqrt{I(u-\tau)}dW_P^*(u)\right\}.$$

Because the process $\tilde{P}$ is a positive local martingale, we know that $\tilde{P}$ is a supermartingale (Klebaner, 2012, Theorem 7.23). So it is enough to show that

$$E_{\mathbb{Q}}[\tilde{P}(t)] = \tilde{P}(0).$$

Conditioning with respect to $\mathcal{F}_t^{W_I^*}$ and using Proposition 1.4.1, we have:

$$\begin{aligned} E_{\mathbb{Q}}[\tilde{P}(t)] &= E_{\mathbb{Q}}\left[\tilde{P}(0)\exp\left\{-\frac{\sigma_P^2}{2}\int_0^t I(u-\tau)du + \sigma_P\int_0^t\sqrt{I(u-\tau)}dW_P^*(u)\right\}\right] \\ &= \tilde{P}(0)E_{\mathbb{Q}}\left[\exp\left\{-\frac{\sigma_P^2}{2}\int_0^t I(u-\tau)du\right\}\right. \\ &\qquad \left. E_{\mathbb{Q}}\left[\exp\left\{\sigma_P\int_0^t\sqrt{I(u-\tau)}dW_P^*(u)\right\}\middle|\,\mathcal{F}_t^{W_I^*}\right]\right] \\ &= \tilde{P}(0). \end{aligned}$$

□

It is possible to show that the discounted price process $\tilde{P}$ is also a martingale with respect to the delayed filtration $\left(\tilde{\mathcal{F}}_t\right)_{t\in[0,T]}$.

Corollary 3.7.1. *The discounted price process $\tilde{P}$ is a martingale with respect to the filtration $(\tilde{\mathcal{F}}_t)_{t\in[0,T]}$ on the probability space $(\Omega, \mathcal{F}, \mathbb{Q})$.*

Proof. Since $\tilde{\mathcal{F}}_t \subset \mathcal{F}_t$ for all $t \in [0,T]$, we obtain by application of Theorem 3.7.1:

$$\begin{aligned} E_{\mathbb{Q}}\left[\tilde{P}(t)\middle|\tilde{\mathcal{F}}_s\right] &= E_{\mathbb{Q}}\left[E_{\mathbb{Q}}\left[\tilde{P}(t)\middle|\mathcal{F}_s\right]\middle|\tilde{\mathcal{F}}_s\right] \\ &= E_{\mathbb{Q}}\left[\tilde{P}(s)\middle|\tilde{\mathcal{F}}_s\right] \\ &= \tilde{P}(s), \end{aligned}$$

where $T \geq t \geq s \geq 0$ and the last equality comes from the fact that the process $\tilde{P}$ is adapted with respect to the filtration $(\tilde{\mathcal{F}}_t)_{t\in[0,T]}$. □

In this section, we would like to price a European call option with expiration date T and strike K. Because the discounted price process $\tilde{P}$ is a martingale, the price of the call option under $\mathbb{Q}$ is

$$C(0) = E_{\mathbb{Q}}\left[e^{-rT}(P(T)-K)^+\right].$$

Before pricing a call option, notice that:

$$\int_0^T I(u-\tau)du = \int_{-\tau}^{T-\tau} I(u)du = \begin{cases} \int_{-\tau}^0 \phi^I(u)du + \int_0^{T-\tau} I(u)du & \text{if } T > \tau \\ \int_{-\tau}^{T-\tau} \phi^I(u)du & \text{if } T \leq \tau. \end{cases} \tag{3.7.12}$$

From equation (3.7.12) we need to differentiate two cases for pricing options, depending on whether $\tau < T$ or $\tau \geq T$.

It is not difficult to see that the price process satisfies the following equation

$$P(t) = P(0)\exp\left\{rT - \frac{\sigma_P^2}{2}\int_0^t I(u-\tau)du + \sigma_P\int_0^t \sqrt{I(u-\tau)}dW_P^*(u)\right\}. \tag{3.7.13}$$

From equation (3.7.13) the logarithm of the price process X, with $X(t) = \log P(t)$ for $t \in [0,T]$ can be expressed as

$$X(t) = x + rt - \frac{1}{2}\int_0^t \sigma_P^2 I(u-\tau)du + \int_0^t \sigma_P\sqrt{I(u-\tau)}dW_P^*(u) \tag{3.7.14}$$

with $X(0) = x = \log p$.

3.7.3 Pricing a call option when $\tau \geq T$

From equation (3.7.14) and because $\tau \geq T$ we have that

$$X(T) = x + rT - \frac{1}{2}\int_0^T \sigma_P^2\phi^I(u-\tau)du + \int_0^T \sigma_P\sqrt{\phi^I(u-\tau)}dW_P^*(u) \text{ with } X(0) = x. \tag{3.7.15}$$

From the properties of the Itô integral (Klebaner, 2012, Theorem 4.11) we have that

$$X(T) \sim N\left(\mu_X, \sigma_X^2\right),$$

where

$$\mu_X = x + rT - \frac{1}{2}\int_{-\tau}^{T-\tau} \sigma_P^2 \phi^I(u)du, \tag{3.7.16}$$

$$\sigma_X^2 = \int_{-\tau}^{T-\tau} \sigma_P^2 \phi^I(u)du. \tag{3.7.17}$$

The price of a call option when the expiration time $T \leq \tau$ is just a generalization of the Black-Scholes-Merton formula, so the price of a European call option with expiration date $T \leq \tau$ and strike K is

$$C(0) = P(0)F_N(d_1|0,1) - Ke^{-rT}F_N(d_2|0,1),$$

where

$$\begin{aligned} d_1 &= \frac{\log\left(\frac{P(0)}{K}\right) + rT + \frac{\sigma_P^2}{2}\int_{-\tau}^{T-\tau}\phi^I(u)du}{\sqrt{\sigma_P^2\int_{-\tau}^{T-\tau}\phi^I(u)du}} \\ d_2 &= d_1 - \sqrt{\sigma_P^2\int_{-\tau}^{T-\tau}\phi^I(u)du}, \end{aligned}$$

and $F_N(.|0,1)$ is the cumulative distribution function of a standard normal random variable (Wilmott, 2006, p. 148).

3.7.4 Pricing a call option when $\tau < T$

For computing the price of the call option, we will first compute the characteristic function of the random variable $X(T)$. The characteristic function of $X(T)$ is defined as

$$\Phi^{X(T)}(\lambda) = E_{\mathbb{Q}}\left[e^{i\lambda X(T)}\right] \text{ for all } \lambda \in \mathbb{R}.$$

For computing the characteristic function of $X(T)$ with $\tau < T$, it convenient to differentiate between what happens before τ and after τ. To that end, let us define the process Z as

$$\begin{aligned} Z(t) &= X(t+\tau) - X(\tau) + x - rt \\ &= x - \frac{1}{2}\int_0^t \sigma_P^2 I(u)du + \int_0^t \sigma_P\sqrt{I(u)}dB_P^*(u) \text{ for } t \geq 0, \end{aligned} \tag{3.7.18}$$

where the last equality comes from the application of Theorem 1.2.1 and $B_P^*(t) = W_P^*(t+\tau) - W_P^*(\tau)$ for $t \geq 0$. Notice that Z satisfies the stochastic differential equation

$$dZ(t) = -\frac{1}{2}\sigma_P^2 I(t)dt + \sigma_P\sqrt{I(t)}dB_P^*(t) \text{ with } Z(0) = x.$$

The random variable $X(T)$ can be expressed in terms of the random variable $Z(T-\tau)$ as

$$\begin{aligned} X(T) &= X(\tau) - x + r(T-\tau) + Z(T-\tau) \\ &= rT - \frac{1}{2}\int_0^\tau \sigma_P^2 \phi^I(u-\tau)du + \int_0^\tau \sigma_P \sqrt{\phi^I(u-\tau)}dW_P^*(u) + Z(T-\tau). \end{aligned}$$

Because $B_P^*(t)$ is independent of $\mathcal{F}_\tau^{W_P^*}$ for all $t \geq 0$ then we have that $Z(t)$ is independent of $\mathcal{F}_\tau^{W_P^*}$ for all $t \geq 0$.

Conditioning with respect to $\mathcal{F}_\tau^{W_P^*}$, we obtain

$$\begin{aligned} \Phi^{X(T)}(\lambda) &= E\left[e^{i\lambda(X(\tau)-x+r(T-\tau)+Z(T-\tau))}\right] \\ &= E_{\mathbb{Q}}\left[E_{\mathbb{Q}}\left[e^{i\lambda(X(\tau)-x+r(T-\tau)+Z(T-\tau))}\middle|\mathcal{F}_\tau^{W_P^*}\right]\right] \\ &= E_{\mathbb{Q}}\left[e^{i\lambda(X(\tau)-x+r(T-\tau))}E_{\mathbb{Q}}\left[e^{i\lambda Z(T-\tau)}\middle|\mathcal{F}_\tau^{W_P^*}\right]\right] \\ &= E_{\mathbb{Q}}\left[e^{i\lambda(X(\tau)-x+r(T-\tau))}\right]E_{\mathbb{Q}}\left[e^{i\lambda Z(T-\tau)}\right]. \end{aligned} \tag{3.7.19}$$

The first expectation of equation (3.7.19) can be computed as

$$\begin{aligned} E_{\mathbb{Q}}\left[e^{i\lambda(X(\tau)-x+r(T-\tau))}\right] &= E_{\mathbb{Q}}\left[e^{i\lambda rT - \frac{i\lambda}{2}\int_0^\tau \sigma_P^2\phi^I(u-\tau)du + i\lambda\int_0^\tau \sigma_P\sqrt{\phi^I(u-\tau)}dW_P^*(u)}\right] \\ &= e^{i\lambda rT - \frac{\sigma_P^2}{2}\epsilon^0(i\lambda+\lambda^2)}, \end{aligned}$$

where the last equality comes from the properties of the Itô integral (Klebaner, 2012, Theorem 4.11) and $\epsilon^0 = \int_{-\tau}^0 \phi^I(u)du$.

Now, for computing the characteristic function of $X(T)$ we first need to compute the characteristic function of $Z(T-\tau)$. If we define the process V as $V(t) = \sigma_P^2 I(t)$ for $t \geq 0$, then $Z(T-\tau)$ can be written as

$$Z(T-\tau) = x - \frac{1}{2}\int_0^{T-\tau} V(u)du + \int_0^{T-\tau} \sqrt{V(u)}dB_P^*(u).$$

By application of the Itô formula, we have that V satisfies the stochastic differential equation

$$dV(t) = a_V(b_V - V(t))dt + \sigma_V\sqrt{V(t)}dW_I^*(t) \text{ with } V(0) = \sigma_P^2\phi^I(0),$$

where $a_V = \tilde{a}_I$, $b_V = \sigma_P^2\tilde{b}_I$ and $\sigma_V = \sigma_P\sigma_I$. From condition (3.7.7) we have that

$$\frac{2a_V b_V}{\sigma_V^2} = \frac{2\tilde{a}_I\tilde{b}_I}{\sigma_I^2} \geq 1.$$

The Feller condition is still satisfied, so the process V is greater than zero with probability one. The process (Z, V) satisfies the Heston model (Heston, 1993). The characteristic function of the log price in the Heston model is known, so the characteristic function of $Z(t)$ with $t \geq 0$ can be written as

$$\begin{aligned}\Phi^{Z(t)}(\lambda) &= E_{\mathbb{Q}}\left[e^{i\lambda Z(t)}\right] \\ &= \exp\left\{i\lambda x + b_V a_V \sigma_V^{-2}\left[(a_V - d)t - 2\log\left(\frac{1-ge^{-dt}}{1-g}\right)\right]\right\} \\ &\quad \exp\left\{V(0)\sigma_V^{-2}\left[\frac{(a_V-d)(1-e^{-dt})}{1-ge^{-dt}}\right]\right\},\end{aligned}$$

with

$$\begin{aligned} d &= \left(a_V^2 - \sigma_V^2(-i\lambda - \lambda^2)\right)^{1/2} \\ g &= \frac{a_V - d}{a_V + d}\end{aligned}$$

(Madan, Reyners, & Schoutens, 2019). We have just shown that

$$\Phi^{X(T)}(\lambda) = e^{i\lambda rT - \frac{\sigma_P^2}{2}\epsilon^0(i\lambda+\lambda^2)}\Phi^{Z(T-\tau)}(\lambda).$$

The benefit of having a closed formula for the characteristic function of the log price is that we can price European options with it.

If we assume that $E_{\mathbb{Q}}[e^{(\delta+1)X(T)}]$ is finite for some $\delta > 0$, then the price of the call option with strike K and maturity T can be computed as

$$C(0) = E_{\mathbb{Q}}\left[e^{-rT}(P(T)-K)^+\right] = \frac{e^{-\delta\log K}}{\pi}\int_0^\infty e^{-i\lambda\log K}\varphi(\lambda)d\lambda, \tag{3.7.20}$$

where

$$\varphi(\lambda) = \frac{e^{-rT}\Phi^{X(T)}\left(\lambda - (\delta+1)i\right)}{\delta^2 + \delta - \lambda^2 + i(2\delta+1)\lambda}.$$

(Carr & Madan, 1999, p.64). This result allows us to compute the price of European call options using the characteristic function of $X(T)$. But first we need to show for which values of δ the expectation $E_{\mathbb{Q}}[e^{(\delta+1)X(T)}]$ is finite.

Proposition 3.7.2. *Let $\tilde{\delta} \in \mathbb{R}$ and $D = \tilde{a}_I^2 + \sigma_P^2\sigma_I^2(\tilde{\delta} - \tilde{\delta}^2)$, then*

1. *If $D \geq 0$ then $E_{\mathbb{Q}}[e^{\tilde{\delta}X(T)}] < \infty$.*
2. *If $D < 0$ then $\lim_{T\to T^*} E_{\mathbb{Q}}[e^{\tilde{\delta}X(T)}] = \infty$, where*

$$T^* = \frac{2}{\sqrt{-D}}\left(\pi + \arctan\left(\frac{\sqrt{-D}}{-a_I}\right)\right) + \tau.$$

Proof. Using similar techniques as the ones used to obtain equation (3.7.19), we can show that $E_{\mathbb{Q}}[e^{\tilde{\delta}X(T)}]$ can be written as

$$\begin{aligned} E_{\mathbb{Q}}\left[e^{\tilde{\delta}X(T)}\right] &= E_{\mathbb{Q}}\left[e^{\tilde{\delta}(X(\tau)-x+r(T-\tau))}\right]E_{\mathbb{Q}}\left[e^{\tilde{\delta}Z(T-\tau)}\right] \\ &= E_{\mathbb{Q}}\left[e^{\tilde{\delta}rT - \frac{\tilde{\delta}}{2}\int_0^\tau \sigma_P^2\phi^I(u-\tau)du + \tilde{\delta}\int_0^\tau \sigma_P\sqrt{\phi^I(u-\tau)}dW_P^*(u)}\right] \\ &\quad E_{\mathbb{Q}}\left[e^{\tilde{\delta}Z(T-\tau)}\right]. \end{aligned} \tag{3.7.21}$$

The first expected value that appears on equation (3.7.21) is the moment generating function of a normal random variable, hence this expectation is finite. So $E_{\mathbb{Q}}\left[e^{\tilde{\delta}X(T)}\right]$ will be finite if and only if $E_{\mathbb{Q}}\left[e^{\tilde{\delta}Z(T-\tau)}\right]$ is finite. We have seen that Z is the log-price of the Heston model (Heston, 1993). The values for which the expected value $E_{\mathbb{Q}}\left[e^{\tilde{\delta}Z(t)}\right]$ is finite have been studied previously. If we define $D = a_V^2 - \sigma_V^2(\tilde{\delta}^2 - \tilde{\delta})$, then from the results given by Friz and Keller-Ressel (2010) we have that

1. If $D \geq 0$ then $E_{\mathbb{Q}}[e^{\tilde{\delta}Z(t)}] < \infty$ for all $t \geq 0$.

2. If $D < 0$ then $\lim_{t\to t^*} E_{\mathbb{Q}}[e^{\tilde{\delta}Z(t)}] = \infty$, where

$$t^* = \frac{2}{\sqrt{-D}}\left(\pi + \arctan\left(\frac{\sqrt{-D}}{-a_V}\right)\right). \tag{3.7.22}$$

Notice that the value D can be written as $D = a_I^2 - \sigma_P^2\sigma_I^2(\tilde{\delta}^2 - \tilde{\delta})$. From equation (3.7.22) we have that when $D < 0$: $\lim_{T\to T^*} E_{\mathbb{Q}}[e^{\tilde{\delta}Z(T-\tau)}] = \infty$, where

$$T^* = \frac{2}{\sqrt{-D}}\left(\pi + \arctan\left(\frac{\sqrt{-D}}{-a_I}\right)\right) + \tau.$$

□

Remark 3.7.1. *In this chapter, we assume that the Brownian motions that model the price and interest processes are independent. A discussion about how we could create a model with correlated Brownian motions is presented in Appendix A.*

3.8 Market option prices

In this subsection, we price European options using the model introduced in this chapter. The parameters, of the model are estimated from historical market data. With these values for the parameters we price different options and compare the results obtained with their market prices. Firstly we select a risk-neutral measure with $\lambda_1 = 0$ and $\lambda_2 = 0$ (defined in equation (3.7.3)). That is, the interest process is the same under the physical measure $\mathbb{P}$ and under the risk-neutral measure $\mathbb{Q}$. For pricing the options, we assume that $r = 0$. To check the performance of our model, we compute the root mean square error:

$$rmse = \sqrt{\frac{1}{n}\sum_{j=1}^{n}\left(MidPrice_j - ModelPrice_j\right)^2}$$

and the relative root mean square error:

$$rrmse = \sqrt{\frac{\frac{1}{n}\sum_{j=1}^{n}\left(MidPrice_j - ModelPrice_j\right)^2}{\sum_{j=1}^{n} MidPrice_j^2}},$$

where n is the number of options and $MidPrice_j = \frac{AskPrice_j + BidPrice_j}{2}$.

As we have seen in Section 3.7.1 we can pick any values for λ_1 and λ_2 as long as conditions (3.7.7) and (3.7.8) are satisfied. This gives a certain flexibility to our model. We would like to choose a risk-neutral measure in which the values for λ_1 and

λ_2 are optimal. Because of that the values for the parameters λ_1 and λ_2 are selected by minimizing the relative root mean square error. In this case, the drift term of the interest process will be different under the measure $\mathbb{P}$ and under the probability $\mathbb{Q}$. However, the diffusion part will be the same in both measures.

The option market data used in this section is explained in Section 1.6. For estimating the parameters of the model, we use a temporal window of one year of historical data. That is, for pricing the options taken at the date 2019-04-01, the parameters of the model are estimated using the historical data from 2018-04-01 to 2019-04-01. In the case of the options taken at 2019-05-01, we use historical data from 2018-05-01 to 2019-05-01 for estimating the model parameters. We proceed in a similar manner for the rest of the options taken at different dates.

As we have seen in Section 2.4 we identify two valid proxies for the interest, when it is modelled by a Cox-Ingersoll-Ross process. One of these proxies is the number of Wikipedia views of the keyword "Altcoin". The other one is the unique number of active addresses. Because of that we separate the analysis into two parts. In the first part, we do the pricing of options using Wikipedia views as proxy for the interest (Section 3.8.1). In Section 3.8.2 we price the options using the number of addresses as the proxy for the interest. All historical data have a daily frequency that is, $\Delta = \frac{1}{365}$ and τ can take values in $\{0, 1\Delta, 2\Delta, \dots, 20\Delta\}$. An explanation of the market attention data used in this section can be found in Section 1.6.

3.8.1 Wikipedia views

In the case we select the number of Wikipedia views of the keyword "Altcoin" as a proxy for the interest, the values for the estimated parameters are shown in Figure 3.1 and in Figure 3.2. Figure 3.1 shows the estimated parameters of the Cox-Ingersoll-Ross process for each of the temporal windows. We can observe that the values of the estimated parameters of the interest process vary with time, this phenomenon was also observed in Chapter 2. The estimated parameters related to the price equation (3.2.1) are shown in Figure 3.2. For each temporal window, we estimate the parameters using the techniques explained in Section 3.6. The results show that in all windows, the estimated delay is always greater than zero.

Each temporal window has a duration of one year and, for example, when the date 2019-09-01 appears on the x-axis of Figures 3.1 and 3.2, this means that the data used to estimate the model parameter are taken from 2018-09-01 to 2019-09-01.

Remark 3.8.1. *The first temporal window shown in Figures 3.1 and 3.2 should be from 2018-04-01 to 2019-04-01. However, during the temporal frame 2017-09-01 to 2018-09-01, the of number of Wikipedia views takes the value zero several times. Because of that, we start estimating the parameters of the model from 2018-09-01 to 2019-09-01.*

With the values of the parameters shown in Figures 3.1 and 3.2 we compute the prices of the options taken from 2019-09-01 to 2021-09-01 and we compare them with their market values. The root mean square error and the relative root mean square error are shown in Figure 3.3 and in Figure 3.4 respectively. The values of these options are computed under the risk-neutral probability $\mathbb{Q}$, where $\lambda_1 = \lambda_2 = 0$. From Section 3.7.1 we have seen that λ_1 and λ_2 can have different values as long as conditions (3.7.7) and (3.7.8) are satisfied. We select the values of λ_1 and λ_2 that minimize the relative root mean squared error. To that end we use the use the Python function "minimize" with the algorithm "Sequential Least Squares Programming", this function is included in the Python package "SciPy". For these new values of λ_1 and λ_2 we obtain a new root mean square error and a new relative root mean square error. These are also shown in Figure 3.3 and in Figure 3.4.

In addition, for these new values of λ_1 and λ_2 we obtain new values for the parameters $\tilde{a}_I$ and $\tilde{b}_I$, which are shown in Figure 3.5. As we can observe in Figure 3.5 the values of the calibrated $\tilde{a}_I$ and $\tilde{b}_I$ vary more through time than the estimated ones.

We also compute the prices given by the Black-Scholes-Merton model. The parameters of the model are estimated by the maximum likelihood estimator method using historical data. The temporal windows used for this model are the same as the ones defined above. The root mean square error and the relative root mean square error obtained by the use of Black-Scholes-Merton model, are shown in Figure 3.3 and in Figure 3.4 respectively. The results obtained by the Black-Scholes-Merton model are in some instances better and in some instances worse than the proposed model. However, when we calibrate the proposed model, the model defined in Section 3.2 obtains better results. In addition, we divide the relative root mean square error obtained by the proposed model by the relative root mean square error obtained by the Black-Scholes-Merton model. The results are shown in Figure 3.6. When the value is less than one, it means that the proposed model gives better results than the Black-Scholes-Merton model. We can observe in Figure 3.6, that the model with calibration is better that the Black-Scholes-Merton model. When the model is not calibrated, the proposed model is not consistent in achieving better results than the Black-Scholes-Merton model.

3.8.2 Unique addresses

Another proxy that we consider is the daily number of unique active addresses. The results of the parameters related to the interest process are shown in Figure 3.7. The estimated parameters that appear in the price equation (3.2.1) are shown in Figure 3.8. In general, the delay parameter is always greater than zero, except for some periods in which it has the value 0.

As in Section 3.8.1 we price the options taken for the dates 2019-04-01 to 2021-09-01. We compute the root mean square error and the relative root mean square error, using the values of the parameters shown in Figures 3.7 and 3.8. The root mean square error and the relative root mean square error that we obtain are shown in Figure 3.9 and in Figure 3.10 respectively.

As we did before, we would like to select the value for λ_1 and λ_2 optimally. For these new values of λ_1 and λ_2 we obtain a new root mean squared error and a new relative root mean square error, that are also shown in Figures 3.9–3.10.

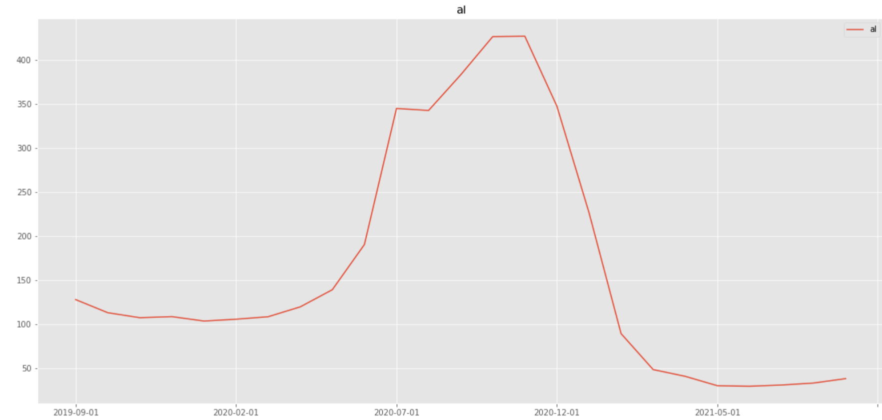
As in Section 3.8.1 for these new values for λ_1 and λ_2, we have new values for the parameters $\tilde{a}_I$ and $\tilde{b}_I$. These are shown in Figure 3.11. In this case, we can see that in general, the calibrated parameters are greater than the estimated parameters.

We also show the values for the root mean square error and the relative root mean square root in Figures 3.9–3.10 obtained using the Black-Scholes-Merton model. As we can observe, the Black-Scholes-Merton model gives similar results to the proposed model. But when we calibrate the proposed model, we obtain better results than the Black-Scholes-Merton model. We again compute the ratio between the relative root mean error given by our model and the ratio between the relative root mean error obtained by the Black-Scholes-Merton model. These results are shown in Figure 3.12. When the model is calibrated, the results given by the proposed model are better than the results given by the Black-Scholes-Merton model. However, this is not the case when the model is not calibrated.

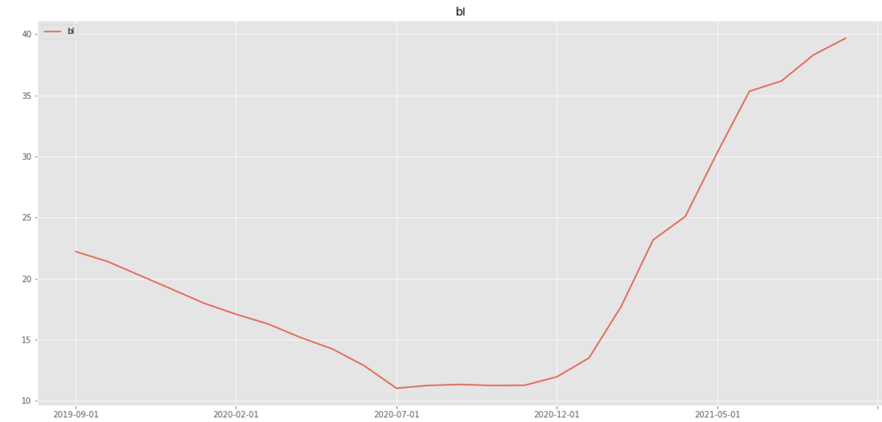
3.9 Conclusion and future work

In this chapter, we proposed a model for pricing Bitcoin options in which the volatility is proportional to the market attention. In this case, we assume that the interest follows a Cox-Ingersoll-Ross process. In Chapter 4 we will use an inverse Gaussian Ornstein-Uhlenbeck process, and this will allow us to incorporate jumps into the volatility process. In addition, we showed how the change of measure could be performed and we derived a semi-closed formula for pricing plain European options. We compared the prices given by our model with the real market data and showed that our model gives better results than the Black-Scholes-Merton model.

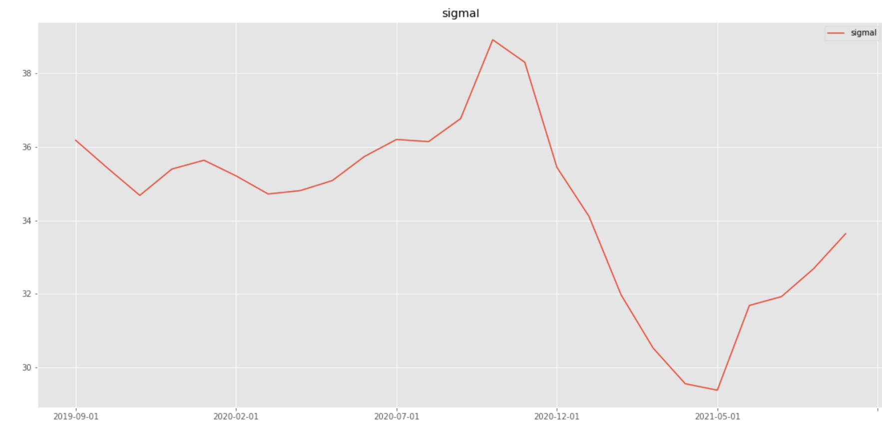
In the model presented in this chapter, the Brownian motions that appears in equations (3.2.1) and (3.2.2) are uncorrelated. In Appendix A we introduced a model with correlation and we derive a semi-closed formula for pricing European options. In future work, it would be interesting to study this correlated model. In addition, as has been proposed before, another interesting topic for study would be the construction of models in which the volatility is affected by several proxies of the interest (Cretarola, Figà-Talamanca, & Patacca, 2020). In this case we will have a model that is similar to the double Heston model (Christoffersen, Heston, & Jacobs, 2009, p. 8).



(A) Estimation of the parameter a_I.

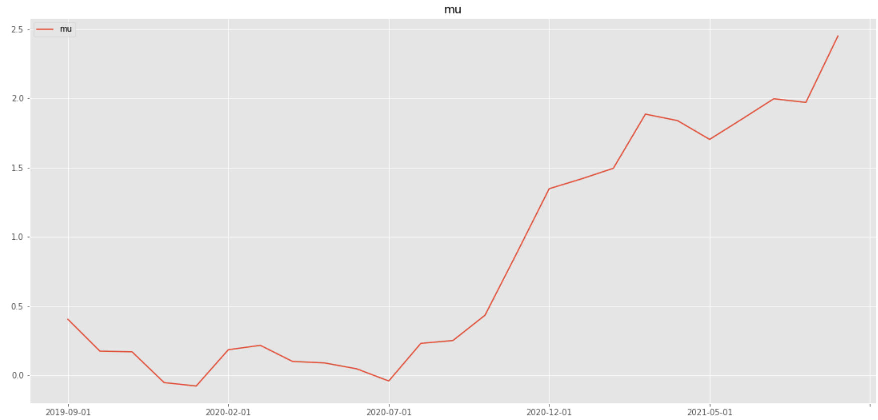


(B) Estimation of the parameter b_I.

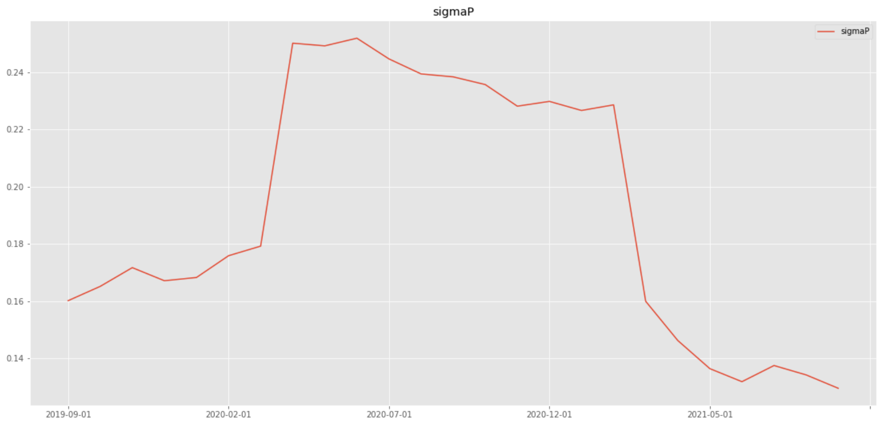


(C) Estimation of the parameter σ_I.

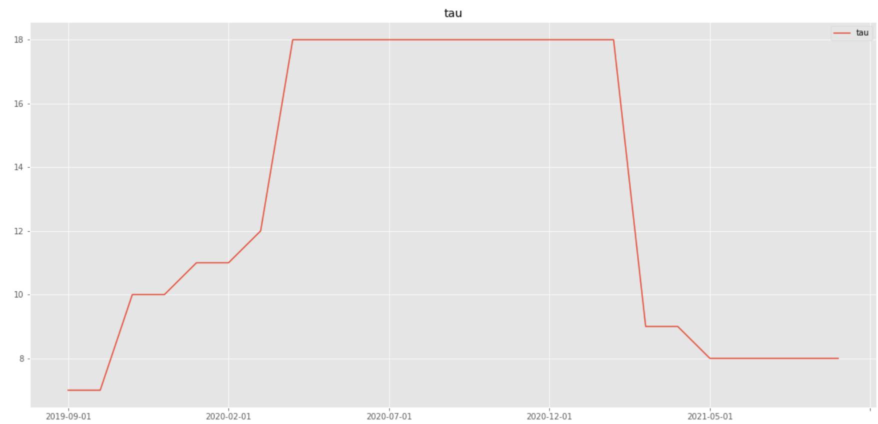
FIGURE 3.1: Estimated values for the parameters of the model defined in Section 3.2 related to market attention process, when the proxy of the market interest is the number of views on Wikipedia of the keyword "Altcoin".



(A) Estimation of the parameter μ.



(B) Estimation of the parameter σ_P.



(C) Estimation of the parameter τ.

FIGURE 3.2: Estimated values for the parameters of the model defined in Section 3.2 related to the price equation, when the proxy of the market interest is the number of views on Wikipedia of the keyword "Altcoin".

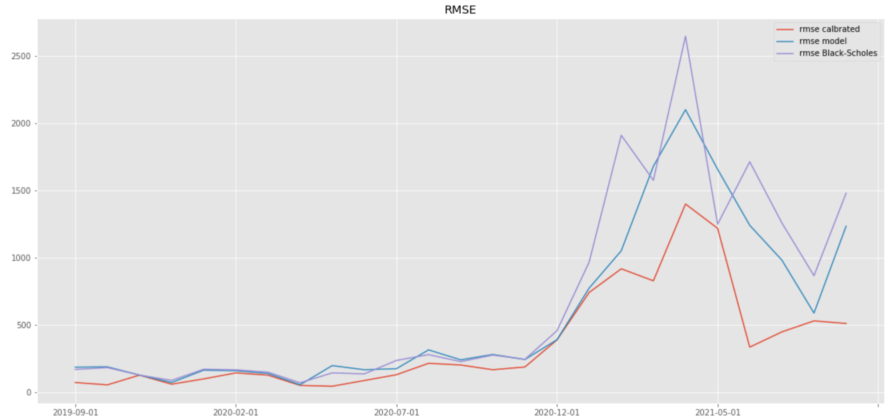


FIGURE 3.3: Root mean square error when the proxy of market attention is the number of Wikipedia views of the word "Altcoin".

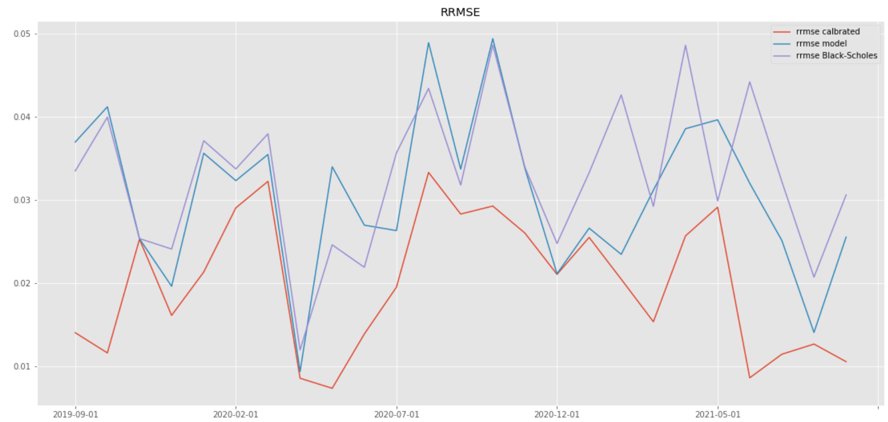
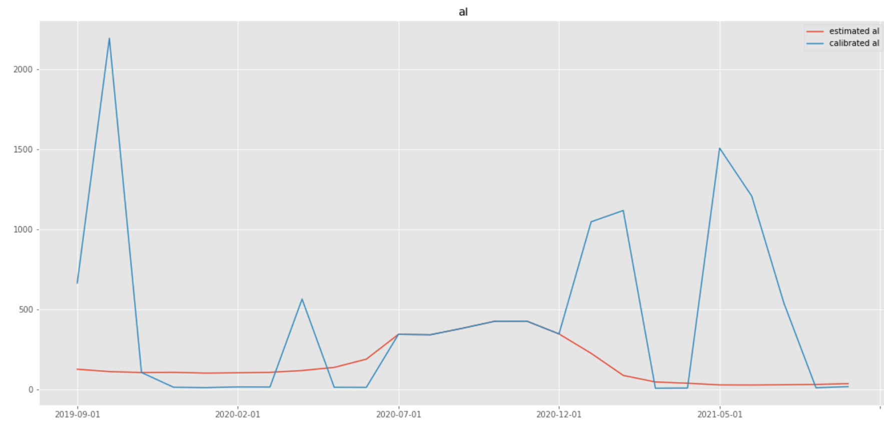
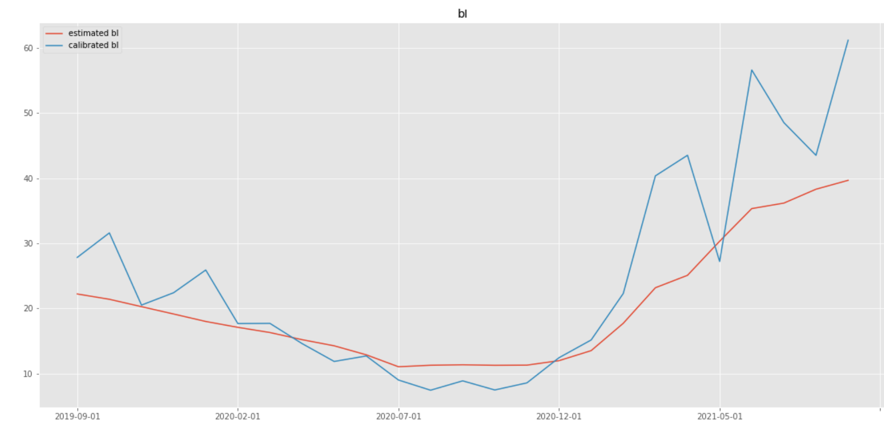


FIGURE 3.4: Relative root mean square error when the proxy for market attention is the number of Wikipedia views of the word "Altcoin".



(A) Estimated and calibrated values for a_I.



(B) Estimated and calibrated values for b_I.

FIGURE 3.5: Calibrated values for the parameters a_I and b_I when the proxy of the market interest is the number of views on Wikipedia of the keyword "Altcoin".

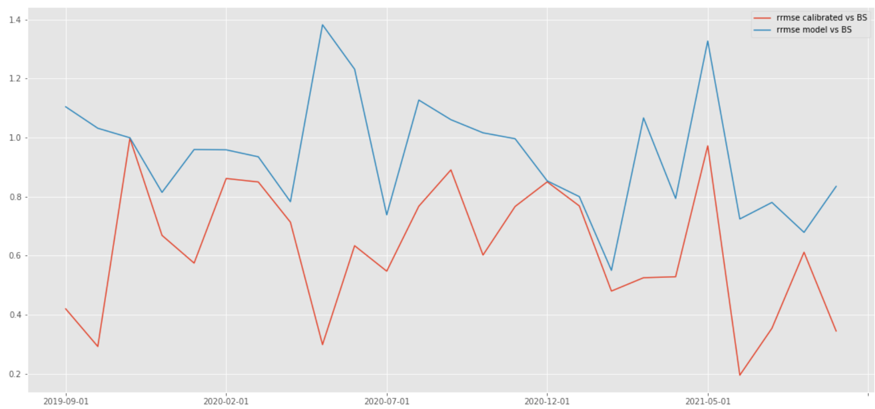
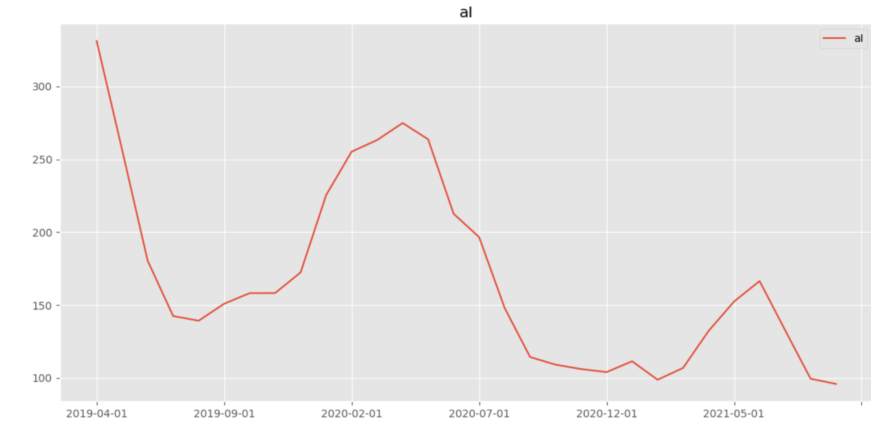
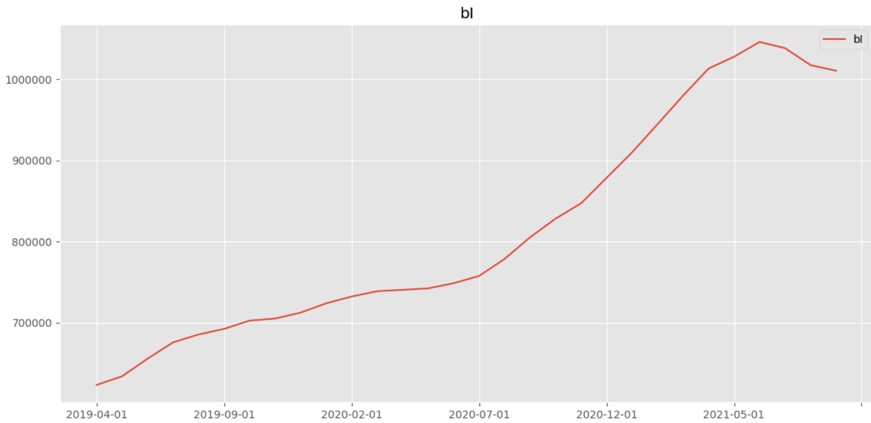


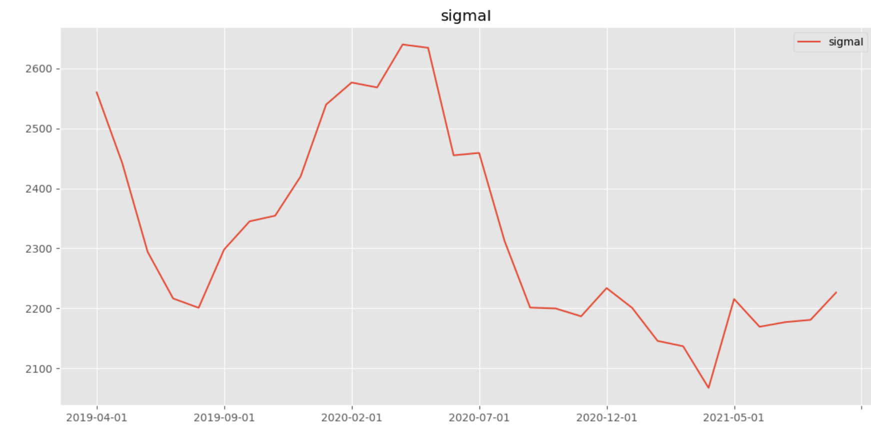
FIGURE 3.6: Relative root mean square error given by our model divided by the relative root mean square error obtained by the Black-Scholes-Merton model, when the proxy for market attention is the number of Wikipedia views of the word "Altcoin".



(A) Estimation of the parameter a_I.

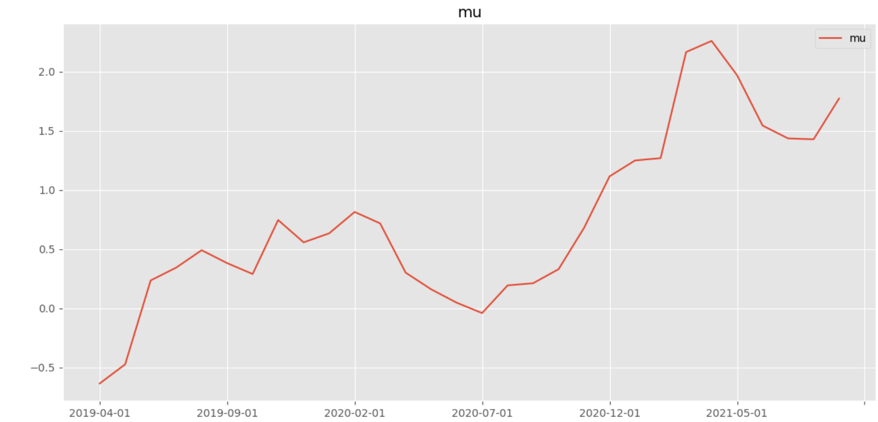


(B) Estimation of the parameter b_I.

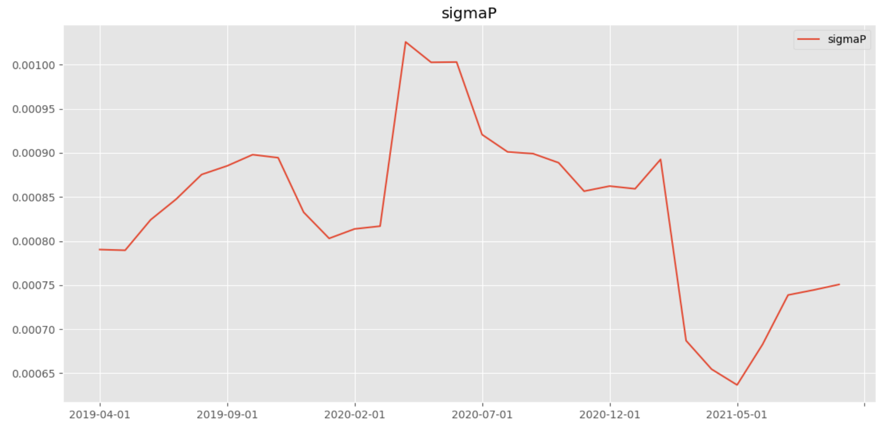


(C) Estimation of the parameter σ_I.

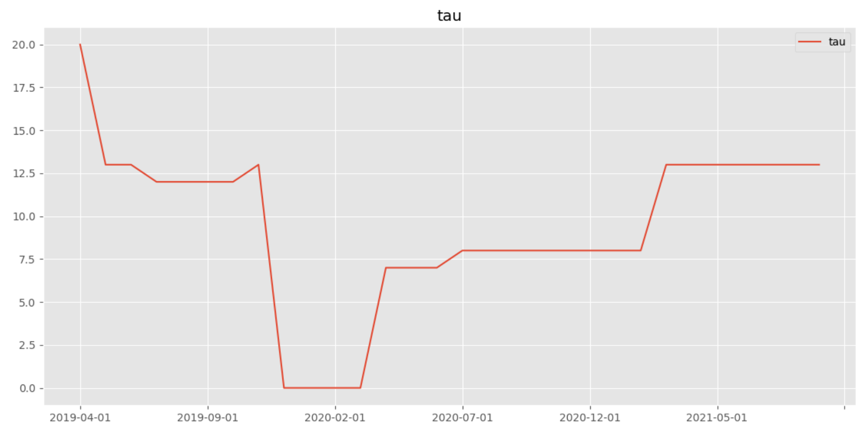
FIGURE 3.7: Estimated values for the parameters of the model defined in Section 3.2 related to market attention process, when the proxy of the market interest is the number of unique active addresses.



(A) Estimation of the parameter μ.



(B) Estimation of the parameter σ_P.



(C) Estimation of the parameter τ.

FIGURE 3.8: Estimated values for the parameters of the model defined in Section 3.2 related to the price equation, when the proxy of the market interest is the number of unique active addresses.

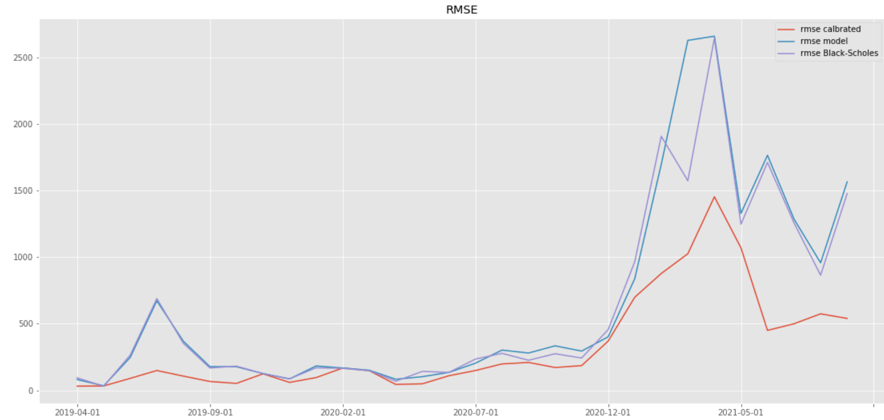


FIGURE 3.9: Root mean square error when the proxy of the market interest is the number of unique active addresses.

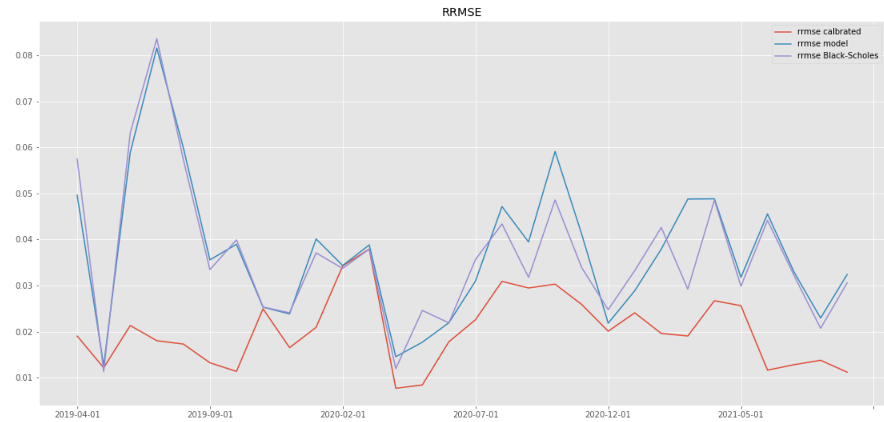
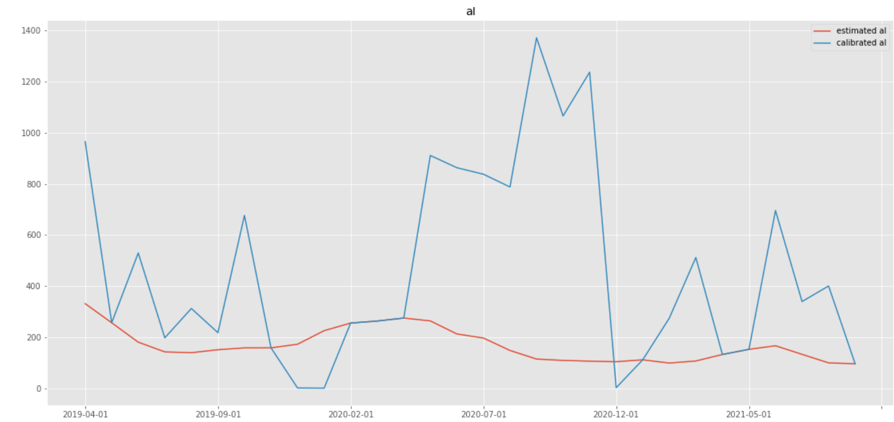
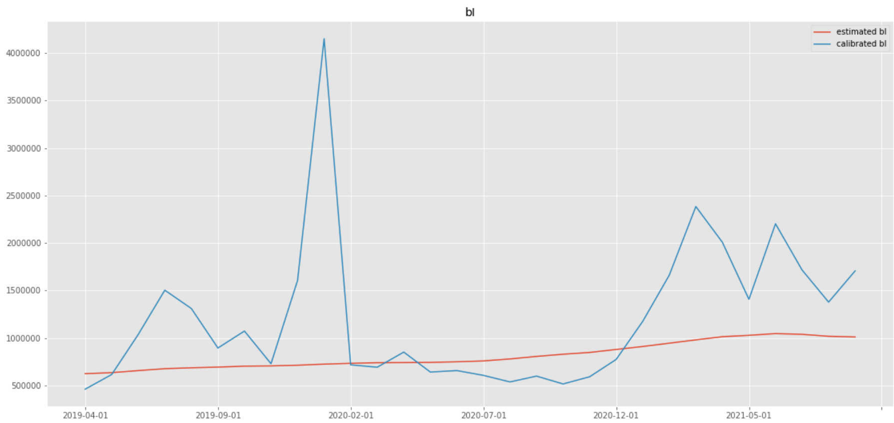


FIGURE 3.10: Relative root mean square error when the proxy of the market interest is the number of unique active addresses.



(A) Estimated and calibrated values for a_I.



(B) Estimated and calibrated values for b_I.

FIGURE 3.11: Calibrated values for the parameters a_I and b_I when the proxy of the market interest is the number of unique active addresses.

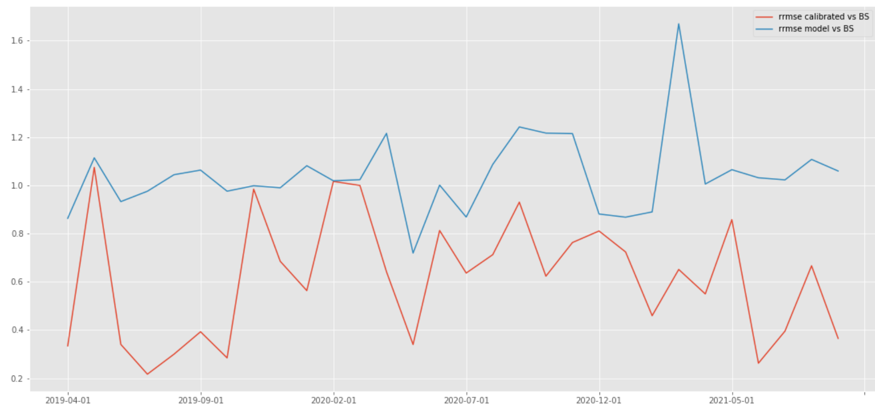


FIGURE 3.12: Relative root mean square error given by our model divided by the relative root mean square error obtained by the Black-Scholes-Merton model, when the proxy of the market interest is the number of unique active addresses.

Chapter 4

A Simple model with an Ornstein-Uhlenbeck interest process

4.1 Introduction

In Chapter 3 we constructed a simple model for Bitcoin option prices. In Chapter 3 we assumed that the market attention follows a Cox-Ingersoll-Ross process. As we saw in Chapter 2 we could in some cases use an Ornstein-Uhlenbeck process to model the interest. In this chapter we will introduce a model that has the same price structure as the model introduced in Chapter 3, meaning that only the volatility of the log-price is affected by the market attention. In this case, we will assume that the interest process follows an inverse Gaussian Ornstein-Uhlenbeck process, so in that way we include jumps when modelling the interest process. The reason for extending the model presented in Chapter 3, adding jumps in the interest structure comes from A. Hou, Wang, Chen, and Härdle (2020). In this article, the authors fit to the historical data of Bitcoin several stochastic volatility models. They show that the model that gives the best fit to the data is the model that includes jumps in the volatility structure.

We will show that some results obtained in Chapter 3 can also be applied to this particular case. The idea of using an Ornstein-Uhlenbeck process for modelling the volatility of the price was introduced by Barndorff-Nielsen and Shephard (2001) and these types of models are usually called Barndorff-Nielsen and Shephard models.

Formally, we introduce the following model. Let us assume $(\Omega, \mathcal{F}, \mathbb{P})$ to be a probability space, that contains a Brownian motion W_P and a Lévy process Z_I. Furthermore, assume that W_P and Z_I are independent. The process Z_I is the background driving Lévy process of an $IG(a_I, b_I)$-Ornstein-Uhlenbeck process with parameter $\lambda_I > 0$. We assume that the interest process I is this inverse Gaussian Ornstein-Uhlenbeck process with background driving Lévy process Z_I.

Due to the fact that Z_I is the background driving Lévy process of an $IG(a_I, b_I)$-Ornstein-Uhlenbeck process, we have that Z_I is a positive Lévy process. Hence Z_I has finite variation on a bounded time interval (Cont & Tankov, 2004, Proposition 3.10). In addition, we have that Z_I satisfies:

$$E\left[e^{-uZ_I(1)}\right] = \exp\left\{-\int_0^\infty \left(1 - e^{-ux}\right) v_I(dx)\right\} \text{ for } u \in \mathbb{R},$$

where v_I is a Lévy measure (Barndorff-Nielsen & Shephard, 2001, p. 6). So the Lévy triplet of Z_I can be written as $(\gamma_I, 0, v_I)$ where

$$\gamma_I = \int_{|x|\leq 1} x v_I(dx).$$

In addition, the Lévy measure v_I can be written in terms of a Lévy density w_I as $v_I(dx) = w_I(x)dx$ where

$$w_I(x) = \frac{a_I}{2\sqrt{2\pi}} x^{-3/2}(1 + b_I^2 x)e^{-\frac{1}{2}b_I^2 x}\mathbb{1}_{(0,\infty)}(x).$$

Also we have that the process Z_I jumps infinitely often because $\int_0^\infty w_I(x)dx = \infty$ (Nicolato & Venardos, 2003, p. 449).

Since the process I is an $IG(a_I, b_I)$-Ornstein-Uhlenbeck process with parameter $\lambda_I > 0$, it satisfies the following stochastic differential equation

$$dI(t) = -\lambda_I I(t)dt + dZ_I(\lambda_I t) \text{ when } t > 0.$$

When $t \in [-L, 0]$ we assume that $I(t) = \phi^I(t)$, where $\phi^I : [-L, 0] \to (0, \infty)$ is a continuous and deterministic function and $L > 0$.

Some of the results used in previous chapters are valid for left-continuous processes. We know that the process I has almost surely càdlàg paths (Valdivieso, Schoutens, & Tuerlinckx, 2009, p. 4). Because of that we define the process $I^- = (I^-(t))_{t\geq 0}$ as

$$I^-(t) = \lim_{s\to t^-} I(s) \text{ almost surely.} \tag{4.1.1}$$

We know that the limit (4.1.1) converges almost surely, because the process I has almost surely càdlàg paths. Notice that since the process I is càdlàg, we have that the process I^- has almost surely càglàd paths.

Finally, for the price process P we assume that it satisfies the following stochastic differential equation

$$dP(t) = \left(\mu + \frac{\sigma_P^2}{2} I^-(t-\tau)\right) P(t)dt + \sigma_P\sqrt{I^-(t-\tau)}P(t)dW_P(t),$$

where $\mu \in \mathbb{R}$, $\sigma_P > 0$, $\tau \in [0, L]$ and $P(0) = p > 0$.

In addition, we assume that there is a bond or a market account $(B(t))_{t\geq 0}$ with known interest rate $r \geq 0$, that satisfies:

$$B(t) = B(0)e^{rt}.$$

So, at the end, we have assumed that the price process P and the interest process I satisfy the following system of stochastic differential equations:

$$\begin{aligned} dP(t) &= \left(\mu + \frac{\sigma_P^2}{2} I^-(t-\tau)\right) P(t)dt + \sigma_P\sqrt{I^-(t-\tau)}P(t)dW_P(t) && (4.1.2)\\ &\quad \text{with } P(0) = p \in \mathbb{R}_+, \\ dI(t) &= -\lambda_I I(t)dt + dZ_I(\lambda_I t) \text{ when } t > 0 && (4.1.3)\\ &\quad \text{with } I(t) = \phi^I(t) \text{ when } t \in [-L, 0]. \end{aligned}$$

Let us define $\left(\mathcal{F}_t^{W_P}\right)_{t\geq 0}$ and $\left(\mathcal{F}_t^{Z_I}\right)_{t\geq 0}$ to be the filtrations generated by the

Brownian motion W_P and the Lévy process Z_I respectively, and we define the general filtration $(\mathcal{F}_t)_{t\geq 0}$ to be the filtration defined as

$$\mathcal{F}_t = \mathcal{F}_t^{W_P} \vee \mathcal{F}_t^{Z_I} \text{ for } t \geq 0,$$

where $\mathcal{F}_t^{W_P} \vee \mathcal{F}_t^{Z_I}$ is the smallest σ-field containing $\mathcal{F}_t^{W_P}$ and $\mathcal{F}_t^{Z_I}$. As we did in Chapter 3 we also define the delayed filtration $(\tilde{\mathcal{F}}_t)_{t\geq 0}$ as

$$\tilde{\mathcal{F}}_t = \begin{cases} \mathcal{F}_t^{W_P} \text{ if } t \leq \tau, \\ \mathcal{F}_t^{W_P} \vee \mathcal{F}_{t-\tau}^{W_I} \text{ if } t > \tau. \end{cases}$$

We have seen in Chapter 2 that equation (4.1.3) has a strong solution, so the market attention follows an Inverse Gaussian Ornstein-Uhlenbeck processes. Now we would like to show that the price equation (4.1.2) has a strong solution. By application of the Itô formula we can show that equation (4.1.2) has the following strong solution

$$P(t) = P(0) \exp\left\{\mu t + \int_0^t \sigma_P \sqrt{I^-(u-\tau)} dW_P(u)\right\} \text{ for } t \geq 0. \tag{4.1.4}$$

Notice that the Itô integral that appears in equation (4.1.4) is well defined since the process I^- has almost surely càglàd paths.

In most cases, it is more convenient to work with the log-price. Because of that we define the process X as

$$X(t) = \log(P(t)) = x + \mu t + \int_0^t \sigma_P \sqrt{I^-(u-\tau)} dW_P(u) \quad \text{for } t \geq 0,$$

where $x = \log(P(0)) = X(0)$.

4.2 Estimation

For the estimation procedure, we assume that we have discrete observations of the processes X and I. We would like to estimate the parameters that appear in the system (4.1.2)–(4.1.3) using these discrete observations.

Here we will assume that there are $N+1 \in \mathbb{N}$ discrete observations with a time horizon H. All of these observations are assumed to be equispaced in time with a time step $\Delta = \frac{H}{N}$. The lag parameter τ can only take a finite number of non-negative values, that is $\tau \in \{0, \Delta, 2\Delta, \ldots, M\Delta\}$ where $M \in \mathbb{N}$ with $N \gg M$ and we assume that there exists a $k \in \{0, 1, 2, \ldots, M\}$ such that $\tau = k\Delta$. We also define $L = M\Delta$.

For the estimation of the parameters, we assume that we have the following observations $\{x_j\}_{j=0}^N$, $\{y_j\}_{j=-M}^N$ for the processes X and I respectively, where x_j is the observation of the variable $X(t_j)$ and y_j is the observation of $I(t_j)$ with $t_j = j\Delta$. The horizon H can be expressed as $H = N\Delta = t_N$.

Remember that when $t \in [-L, 0]$ we have $I(t) = \phi^I(t)$. For the deterministic function ϕ^I we approximate it by linear interpolation, i.e.:

$$\hat{\phi}^I(t) = y_j + (t - t_j)\frac{y_{j+1} - y_j}{t_{j+1} - t_j} \text{ when } t \in [t_j, t_{j+1}] \text{ and } j = -M, \ldots, -1. \tag{4.2.1}$$

The approximation function $\hat{\phi}^I$ is just linear interpolations of the observations of the process I when $t \in [-L, 0]$. We use the approximation defined in equation (4.2.1) to approximate the integrals:

$$\int_{t_j}^{t_{j+1}} \phi^I(s-\tau)ds \approx \int_{t_j}^{t_{j+1}} \hat{\phi}^I(s-\tau)ds \text{ when } t \in [t_j, t_{j+1}] \text{ and } j = -M, \ldots, -1.$$

For the estimation of the parameters of the system (4.1.2)–(4.1.3), we first estimate the parameters of the interest process and then we estimate the parameters that appear in equation (4.1.2). For the estimation of the parameters related to the market attention process, we will use the maximum likelihood estimation method explained in Chapter 2. In the case of the parameters that appear in equation (4.1.2) we will use the conditional likelihood estimation method.

4.2.1 Conditional likelihood estimator

As we have said, we estimate the parameters (a_I, b_I, λ_I) that appear in equation (4.1.3) using the methods explained in Chapter 2. In the case of the parameters (μ, σ_P) that appear on equation (4.1.2), we will use the conditional likelihood estimation method.

We are interested in the estimation of the parameters (μ, σ_P). We will assume that the lag parameter τ is given. For the estimation of the parameters μ and σ_P we will use the conditional likelihood method. Instead of using the log-price for the estimation, we will use the logarithmic returns of the price. Because of this we define the vector of the logarithmic returns R as

$$R = (R(t_0, t_1), R(t_1, t_2), \ldots R(t_{N-1}, t_N)),$$

where

$$R(t_j, t_{j+1}) = X(t_{j+1}) - X(t_j) \text{ for } j = 0, \ldots, N-1.$$

Let us define

$$r_{j+1} = x_{j+1} - x_j \text{ for } j = 0, 1, \ldots, N-1$$

as the realizations of the logarithmic returns. Let us define the random vector J_I as

$$J_I = \left(J_I^{t_0}, J_I^{t_1}, \ldots, J_I^{t_{N-1}}\right),$$

where

$$J_I^{t_j} = \int_{t_j}^{t_{j+1}} I^-(u-\tau)du \text{ for } j = 0, 1 \ldots N-1.$$

Remark 4.2.1. *Due to the fact that I is a process with almost surely càdlàg paths we have that*

$$\int_{t_j}^{t_{j+1}} I^-(u-\tau)du = \int_{t_j}^{t_{j+1}} I(u-\tau)du \text{ almost surely for } j = 0, 1 \ldots N-1.$$

Using similar techniques as in Chapter 3, it is possible to show that the logarithmic returns are independent given the σ-algebra $\sigma(J_I)$ generated by the random variable J_I. In addition, it is possible to show that

$$f_{j+1}^R(r_{j+1}|\sigma(J_I)) = f_N(r_{j+1}|\mu\Delta, v_R^{2,j}) \text{ for } j = 0, \ldots, N-1, \tag{4.2.2}$$

where $f_N(.|\mu\Delta, v_R^{2,j})$ is the density function of a normal random variable with mean $\mu\Delta$ and variance $v_R^{2,j}$ that is defined as

$$v_R^{2,j} = \sigma_P^2 \int_{t_j}^{t_{j+1}} I(u-\tau)du.$$

Notice that the random vector R depends on the parameter $\theta_R = (\lambda_R, \xi_R)$, where

$$\lambda_R = (\mu, \sigma_P), \qquad \xi_R = (a_I, b_I, \lambda_I).$$

The parameter ξ_R is estimated using the methods of Chapter 2. So, in this case the parameter of interest is the vector λ_R and ξ_R is the vector of nuisance parameters. The conditional log-likelihood given J_I is

$$\begin{aligned} l^R\left(\lambda_R \left|\sigma\left(J_I\right), \tau\right.\right) &= \log f_{1:N}^R\left(r_{1:N} \left|\sigma\left(J_I\right)\right.\right) \\ &= \sum_{j=0}^{N-1} \log f_{j+1}^R\left(r_{j+1}|\sigma\left(J_I\right)\right) \\ &= \sum_{j=0}^{N-1} \log f_N(r_{j+1}|\mu\Delta, v_R^{2,j}). \end{aligned} \tag{4.2.3}$$

We will maximize the function $l^R\left(.\left|\sigma\left(J_I\right), \tau\right.\right)$ to estimate the parameter λ_R. It is possible to obtain an analytical expression for the maximum likelihood estimators of μ and σ_P.

Proposition 4.2.1. *The maximum likelihood estimators of μ and σ_P are*

$$\begin{aligned} \hat{\mu} &= \frac{\sum_{j=0}^{N-1} \frac{r_j}{J_I^{t_j}}}{\Delta \sum_{j=0}^{N-1} \frac{1}{J_I^{t_j}}}, \\ \hat{\sigma}_P &= \sqrt{\frac{1}{N} \sum_{j=0}^{N-1} \left(\frac{r_j - \hat{\mu}\Delta}{\sqrt{J_I^{t_j}}}\right)^2}. \end{aligned}$$

Proof. See proof of Proposition 3.6.2 in Chapter 3. □

For computing the function $l^R\left(.\left|\sigma\left(J_I\right), \tau\right.\right)$ we need to know the value of the Lebesgue integral that appears in $v_R^{2,j}$. As we did in Chapter 3 these integrals are estimated by the Trapezoid method:

$$\begin{aligned} J_I^{t_j} = \int_{t_j}^{t_{j+1}} I(u-\tau)du &\approx \int_{t_j}^{t_{j+1}} \left(y_{j-k} + (u-t_j)\frac{y_{j+1-k} - y_{j-k}}{t_{j+1} - t_j}\right) du \\ &= \frac{\Delta}{2}(y_{j+1-k} + y_{j-k}). \end{aligned}$$

The approximation is based on integrating the linear interpolation between the observations at times t_j and t_{j+1}. Due to how we define the approximation function $\hat{\phi}^I$ in (4.2.1), we do not need to differentiate between the returns at times before τ and after τ, because the conditional distribution of the returns satisfies equation (4.2.2) independently of whether $t_{j+1} \leq t_k = \tau$ or $t_{j+1} > t_k = \tau$.

Remark 4.2.2. *As we did in Section 3.6.2 we reduce the problem of estimating the parameter τ to a problem of model selection. We select the value of τ that maximizes the expression:*

$$l^R\left(\hat{\lambda}_R\,|\sigma\left(J_I\right),\tau=r\Delta\right)\ \textit{for}\ r=0,1,\ldots M,$$

where $\lambda_R=(\hat{\mu},\hat{\sigma}_P)$ is defined as in Proposition 4.2.1.

4.2.2 Numerical experiments

Now we would like to test the techniques presented in Section 4.2.1. To that end, we produce 20 realizations of $N=800$ steps of the processes defined in equations (4.1.2)–(4.1.3) with the following values for the parameters

$$a_I=4,\quad b_I=10,\quad \lambda_I=5,\quad \mu=0.1,\quad \sigma_P=0.3,\quad \tau=10\Delta,$$

with time step $\Delta=0.0125$, $L=20\Delta$ and initial values $P(0)=100$ and

$$\phi^I(t)=0.4+0.2\cos^2\left(30t+\frac{\pi}{2}\right)\ \text{for}\ t\in[-L,0].$$

For each of the realizations, we estimate the parameters of the model using the techniques explained in Section 4.2. Once we estimate the parameters of all of the simulations, we compute the mean of the estimated parameters and its standard deviation. The results are shown in Table 4.1. We can observe that the methods explained in Section 4.2.1 give estimates that are near to the true value of the parameters.

Parameter	True value	Mean	Standard deviation
a_I	4	3.935182	0.1068459
b_I	10	9.7396181	0.284834
λ_I	5	5.111530	0.064071
μ	0.1	0.084563	0.038136
σ_P	0.3	0.299311	0.0075191
τ	10 Δ	10.3 Δ	2.304343 Δ

TABLE 4.1: Mean and standard deviation of the estimated parameters using the estimation method explained in Section 4.2 for 20 realizations.

4.3 Pricing vanilla options

Now we are interested in pricing European call and put options with strike price K and expiration date T. Because the expiration date of the options is T we assume that the time $t\in[0,T]$ and the σ-algebra $\mathcal{F}=\mathcal{F}_T$.

4.3.1 Change of measure

For pricing options, it is required to be under an equivalent probability measure called a risk-neutral measure. Under a risk-neutral measure, the discounted stock price is a martingale. As shown by Nicolato and Venardos (2003) in Barndorff-Nielsen and Shephard models, it is possible to obtain several risk-neutral measures and under these measures the volatility process is still an Ornstein-Uhlenbeck process. We cannot apply the results of Nicolato and Venardos (2003) directly due to the presence of the delay parameter τ.

Our objective is to construct a risk-neutral measure $\mathbb{Q}$ under which the discounted stock price is a martingale, and the interest process is still an inverse Gaussian Ornstein-Uhlenbeck process, possibly with different parameters. To do that, we perform the change of measure in two steps. First, we change to a measure under which the market attention process I is an inverse Gaussian Ornstein-Uhlenbeck process with different parameters a_I and b_I. Finally, from this intermediate measure we change to the desired risk-neutral measure.

First, we would like to find an equivalent $\mathbb{Q}^*$ such that the Lévy process Z_I has the Lévy triplet $(\tilde{\gamma}_I, 0, \tilde{v}_I)$ where

$$\begin{aligned} \tilde{v}_I(dx) &= \tilde{w}_I(x)dx = \frac{\tilde{a}_I}{2\sqrt{2\pi}} x^{-3/2}(1+\tilde{b}_I^2 x)e^{-\frac{1}{2}\tilde{b}_I^2 x}\mathbb{1}_{(0,\infty)}(x)dx, && (4.3.1)\\ \tilde{\gamma}_I &= \int_{|x|\le 1} x\tilde{v}_I(dx), && (4.3.2)\end{aligned}$$

with $\tilde{a}_I, \tilde{b}_I > 0$. We choose the Lévy triplet $(\tilde{\gamma}_I, 0, \tilde{v}_I)$ as shown in equations (4.3.1)–(4.3.2), because in that way the interest process I is still an $IG(\tilde{a}_I, \tilde{b}_I)$-Ornstein-Uhlenbeck process. That is, under this new proposed probability measure $\mathbb{Q}^*$ the interest process has different parameters $\tilde{a}_I$ and $\tilde{b}_I$, but it is still an inverse Gaussian Ornstein-Uhlenbeck process with parameter $\lambda_I > 0$.

Proposition 4.3.1. *If $\tilde{a}_I = a_I$ then there exists an equivalent probability measure $\mathbb{Q}^*$ with respect to the measure $\mathbb{P}$ such that the Lévy process Z_I has Lévy triplet $(\tilde{\gamma}_I, 0, \tilde{v}_I)$ defined as in equations (4.3.1)–(4.3.2).*

The probability measure $\mathbb{Q}^*$ can be expressed as

$$\mathbb{Q}^*(A) = \int_A Z^*(T)d\mathbb{P} \text{ for } A \in \mathcal{F}, \tag{4.3.3}$$

where the process Z^* is defined as

$$Z^*(t) = \exp\left\{ \lim_{\epsilon\downarrow 0}\left(\sum_{\substack{s\le t\\ |\Delta Z_I(s)|>\epsilon}} \log(H(\Delta Z_I(s))) - t\int_{|x|>\epsilon}(H(x)-1)v_I(dx)\right)\right\} \tag{4.3.4}$$

for $t \in [0,T]$, with

$$H(x) = \frac{1+\tilde{b}_I^2 x}{1+b_I^2 x} e^{-\frac{x}{2}(\tilde{b}_I^2 - b_I^2)}\mathbb{1}_{(0,\infty)}(x).$$

Proof. Let us consider the function H to be the function defined as:

$$H(x) = \frac{\tilde{a}_I}{a_I}\frac{1+\tilde{b}_I^2 x}{1+b_I^2 x} e^{-\frac{x}{2}(\tilde{b}_I^2 - b_I^2)}\mathbb{1}_{(0,\infty)}(x).$$

Notice that v_I and $\tilde{v}_I$ satisfy the following relation:

$$\tilde{v}_I(dx) = H(x)v_I(dx).$$

Nicolato and Venardos (2003, p. 455) showed that if $\tilde{a}_I = a_I$ then we have that

$$\int_{\mathbb{R}} \left(1 - \sqrt{H(x)}\right)^2 v_I(dx) < \infty.$$

Finally notice that

$$\begin{aligned}
\tilde{\gamma}_I &= \int_{|x|\leq 1} x(H(x)-1)v_I(dx) + \gamma_I \\
&= \int_{|x|\leq 1} x\tilde{v}_I(dx) - \int_{|x|\leq 1} xv_I(dx) + \int_{|x|\leq 1} xv_I(dx) \\
&= \int_{|x|\leq 1} x\tilde{v}_I(dx).
\end{aligned}$$

So, from Proposition 1.1.3 there is an equivalent measure $\mathbb{Q}^*$ under which Z_I is a Lévy process with Lévy triplet $(\tilde{\gamma}_I, 0, \tilde{v}_I)$. In addition, $\mathbb{Q}^*$ is defined as in equation (4.3.3) and the process Z^* is defined as in equation (4.3.4). □

In Proposition 4.3.1 we showed that Z_I is the background driving Lévy process of an $IG(a_I, \tilde{b}_I)$-Ornstein-Uhlenbeck process with parameter $\lambda_I > 0$ under the measure $\mathbb{Q}^*$. Notice that the market attention process I is still an inverse Gaussian process and its parameters are the same as in the physical measure except for the parameter b_I; that has changed to the parameter $\tilde{b}_I$.

Under the probability measure $\mathbb{Q}^*$ the process Z_I is still a Lévy process. But what happens with the process W_P, it is still a Brownian motion? It is possible to show that under the measure $\mathbb{Q}^*$ the process W_P is a Brownian motion and the processes W_P and Z_I are independent (see Section B.1).

Now we would like to find an equivalent measure $\mathbb{Q}$ with respect to the measure $\mathbb{Q}^*$ under which the discounted stock price is a martingale. To that end, let us define the process Z as

$$Z(t) = \exp\left\{-\int_0^t \theta_P(s)dW_P(s) - \frac{1}{2}\int_0^t \theta_P^2(s)ds\right\} \text{ for } t \in [0,T], \quad (4.3.5)$$

where θ_P is an adapted process that is defined as

$$\theta_P(t) = \frac{\mu + \frac{\sigma_P^2}{2}I^-(t-\tau) - r}{\sigma_P\sqrt{I^-(t-\tau)}} \text{ for } t \in [0,T]. \quad (4.3.6)$$

We show in Theorem 4.3.1 that Z is a martingale with respect to $(\mathcal{F}_t)_{t\in[0,T]}$ then by the Girsanov theorem the process W_P^* defined as

$$W_P^*(t) = W_P(t) + \int_0^t \theta_P(s)ds \text{ for } t \in [0,T], \quad (4.3.7)$$

is a Brownian motion under the measure $\mathbb{Q}$, where

$$\mathbb{Q}(A) = \int_A Z(T)d\mathbb{Q}^* \; \forall A \in \mathcal{F}.$$

Furthermore, the price process P satisfies the stochastic differential equation

$$dP(t) = rP(t)dt + \sigma_P P(t)\sqrt{I^-(t-\tau)}dW_P^*(t). \quad (4.3.8)$$

It is possible to show that when the price process P satisfies the equation (4.3.8), then the discounted stock price is a martingale with respect to $(\mathcal{F}_t)_{t\in[0,T]}$.

Theorem 4.3.1. *The process $(Z(t))_{t\in[0,T]}$ defined in (4.3.5) is a martingale with respect to $(\mathcal{F}_t)_{t\in[0,T]}$.*

Proof. We know that Z is a non-negative local martingale because Z is the stochastic exponential of a local martingale (Klebaner, 2012, p. 227). Hence it is a supermartingale (Klebaner, 2012, Theorem 7.23). Because of that, it is enough to prove that $E[Z(t)] = Z(0) = 1$. Conditioning with respect to $\mathcal{F}_t^{Z_I}$ and using Proposition 1.4.1, we have that

$$\begin{aligned}
E[Z(t)] &= E\left[\exp\left\{-\int_0^t \theta_P(s)dW_P(s) - \frac{1}{2}\int_0^t \theta_P^2(s)ds\right\}\right] \\
&= E\left[\exp\left\{-\frac{1}{2}\int_0^t \theta_P^2(s)ds\right\} E\left[\exp\left\{-\int_0^t \theta_P(s)dW_P(s)\right\}\middle| \mathcal{F}_t^{Z_I}\right]\right] \\
&= E\left[\exp\left\{-\frac{1}{2}\int_0^t \theta_P^2(s)ds\right\}\exp\left\{\frac{1}{2}\int_0^t \theta_P^2(s)ds\right\}\right] \\
&= 1 \text{ for } t \in [0,T].
\end{aligned}$$

□

In this case under the measure $\mathbb{Q}$ the process Z_I is still a Lévy process with triplet $(\tilde{\gamma}_I, 0, \tilde{v}_I)$ and the processes W_P^* and Z_I are independent (see SectionB.2). In addition, because $\mathbb{P}$ is equivalent to $\mathbb{Q}^*$ and $\mathbb{Q}^*$ is equivalent to $\mathbb{Q}$, we have that $\mathbb{P}$ is equivalent to $\mathbb{Q}$. We have just proved that there exists an equivalent probability measure $\mathbb{Q}$ with respect to $\mathbb{P}$ such that the price process P satisfies the equation (4.3.8) and the interest process is an $IG(a_I, \tilde{b}_I)$-Ornstein-Uhlenbeck process with parameter $\lambda_I > 0$.

We still have to show that the discounted stock price is a martingale with respect to the filtration $(\mathcal{F}_t)_{t\in[0,T]}$. Using the Itô formula it is easy to see that under $\mathbb{Q}$ the price P can be written as

$$P(t) = P(0)\exp\left\{rt - \frac{\sigma_P^2}{2}\int_0^t I^-(u-\tau)du + \sigma_P\int_0^t \sqrt{I^-(u-\tau)}dW_P^*(u)\right\} \quad (4.3.9)$$

for $t \in [0,T]$. Now we would like to show that

$$E_{\mathbb{Q}}\left[\tilde{P}(t)\,|\mathcal{F}_s\right] = \tilde{P}(s) \text{ for } t,s \in [0,T] \text{ such that } t \geq s,$$

where $E_{\mathbb{Q}}$ symbolizes the expected value under the measure $\mathbb{Q}$ and $\tilde{P}$ is the discounted stock price, that is defined as

$$\tilde{P}(t) = e^{-rt}P(t) \text{ for } t \in [0,T].$$

Proposition 4.3.2. *The discounted price process $\tilde{P}$ is a martingale with respect to the filtration $(\mathcal{F}_t)_{t\in[0,T]}$ on the probability space $(\Omega, \mathcal{F}, \mathbb{Q})$.*

Proof. From equation (4.3.9) we know that $\tilde{P}$ satisfies the following equation:

$$\tilde{P}(t) = \tilde{P}(0)\exp\left\{-\frac{\sigma_P^2}{2}\int_0^t I^-(u-\tau)du + \sigma_P\int_0^t \sqrt{I^-(u-\tau)}dW_P^*(u)\right\}.$$

Notice that the process $\tilde{P}$ is the stochastic exponential of a local martingale, hence it is a non-negative local martingale (Klebaner, 2012, p. 227) and hence it is a supermartingale (Klebaner, 2012, Theorem 7.23). So it is enough to show that

$$E_{\mathbb{Q}}[\tilde{P}(t)] = \tilde{P}(0).$$

Conditioning with respect to $\mathcal{F}_t^{Z_I}$ and using Proposition 1.4.1, we have that

$$\begin{aligned}
E_{\mathbb{Q}}[\tilde{P}(t)] &= E_{\mathbb{Q}}\left[\tilde{P}(0)\exp\left\{-\frac{\sigma_P^2}{2}\int_0^t I^-(u-\tau)du + \sigma_P\int_0^t \sqrt{I^-(u-\tau)}dW_P^*(u)\right\}\right] \\
&= \tilde{P}(0)E_{\mathbb{Q}}\left[\exp\left\{-\frac{\sigma_P^2}{2}\int_0^t I^-(u-\tau)du\right\}\right. \\
&\quad \left. E_{\mathbb{Q}}\left[\exp\left\{\sigma_P\int_0^t \sqrt{I^-(u-\tau)}dW_P^*(u)\right\}\middle| \mathcal{F}_t^{Z_I}\right]\right] \\
&= \tilde{P}(0).
\end{aligned}$$

In addition, the discounted price process $\tilde{P}$ is also a martingale with respect to the delayed filtration $\left(\tilde{\mathcal{F}}_t\right)_{t\in[0,T]}$. □

Corollary 4.3.1. *The discounted price process $\tilde{P}$ is a martingale with respect to the filtration $(\tilde{\mathcal{F}}_t)_{t\in[0,T]}$ on the probability space $(\Omega, \mathcal{F}, \mathbb{Q})$.*

Proof. Since $\tilde{\mathcal{F}}_t \subset \mathcal{F}_t$ for all $t \in [0,T]$, we obtain by application of Proposition 4.3.2:

$$\begin{aligned}
E_{\mathbb{Q}}\left[\tilde{P}(t)\middle|\tilde{\mathcal{F}}_s\right] &= E_{\mathbb{Q}}\left[E_{\mathbb{Q}}\left[\tilde{P}(t)\middle|\mathcal{F}_s\right]\middle|\tilde{\mathcal{F}}_s\right] \\
&= E_{\mathbb{Q}}\left[\tilde{P}(s)\middle|\tilde{\mathcal{F}}_s\right] \\
&= \tilde{P}(s),
\end{aligned}$$

where $T \geq t \geq s \geq 0$ and the last equality comes from the fact that the process $\tilde{P}$ is adapted with respect to the filtration $(\tilde{\mathcal{F}}_t)_{t\in[0,T]}$. □

4.4 Pricing call options

We showed that under the risk-neutral probability, the price process P satisfies the following equation:

$$dP(t) = rP(t)dt + \sigma_P\sqrt{I^-(t-\tau)}dW_P^* \text{ with } P(0) = p \in \mathbb{R}_+, \tag{4.4.1}$$

where r is a known interest rate, W_P^* is a Brownian motion, and $\sigma_P > 0$. In addition we have that the interest process satisfies an $IG(a_I, \tilde{b}_I)$-Ornstein-Uhlenbeck process with parameter $\lambda_I > 0$, so the interest I satisfies the stochastic differential equation:

$$dI(t) = -\lambda_I I(t)dt + dZ_I(\lambda_I t) \text{ with } I(0) = y_0, \tag{4.4.2}$$

where Z_I is a Lévy process with Lévy triplet $(\tilde{\gamma}_I, 0, \tilde{v}_I)$ defined as in (4.3.1)–(4.3.2). In Chapter 2 we have seen that equation (4.4.2) has the following strong solution

$$I(t) = e^{-\lambda_I t}\left(I(0) + \int_0^t e^{\lambda_I s}dZ_I(\lambda_I s)\right) \text{ for } t \in [0,T].$$

Define the process Z_I^* as

$$Z_I^*(t) = \int_0^t e^{\lambda_I s} dZ_I(\lambda_I s) \text{ for } t \in [0, T].$$

So, we can express I as

$$I(t) = e^{-\lambda_I t} \left(I(0) + Z_I^*(t) \right) \text{ for } t \in [0, T]. \tag{4.4.3}$$

From the right hand side of equation (4.4.3) we have that the random variable $Z_I^*(t)$ identifies the distribution of $I(t)$. From Chapter 2 we have that the characteristic function of the random variable $Z_I^*(t)$ is

$$\Phi^{Z_I^*(t)}(u) = e^{a_I\left(\sqrt{\tilde{b}_I^2 - 2iu} - \sqrt{\tilde{b}_I^2 - 2iue^{\lambda_I t}}\right)} \text{ for all } u \in \mathbb{R}. \tag{4.4.4}$$

In some cases, we will be interested in the process I multiplied by some positive constant $C > 0$. In this case, we will be interested in the distribution of the random variable $CI(t)$.

Proposition 4.4.1. *If V is the process defined as $V(t) = CI(t)$ for $t \in [0, T]$ and $C > 0$ then V is an $IG(a_V, b_V)$-Ornstein-Uhlenbeck process with parameter $\lambda_V > 0$, where*

$$\begin{aligned} a_V &= \sqrt{C} a_I, & (4.4.5)\\ b_V &= \frac{\tilde{b}_I}{\sqrt{C}}, & (4.4.6)\\ \lambda_V &= \lambda_I. & (4.4.7) \end{aligned}$$

Proof. For equation (4.4.3) we have that the process V satisfies

$$\begin{aligned} V(t) &= e^{-\lambda_I t} \left(CI(0) + CZ_I^*(t) \right) \\ &= e^{-\lambda_I t} \left(V(0) + CZ_I^*(t) \right) \\ &= e^{-\lambda_I t} \left(V(0) + Z_V^*(t) \right), \end{aligned} \tag{4.4.8}$$

where $Z_V^*(t) = CZ_I^*(t)$ for $t \in [0, T]$. Equation (4.4.8) shows that the distribution of $V(t)$ is determined by the distribution of $Z_V^*(t)$. Let $u \in \mathbb{R}$, from the result in (4.4.4) we have that the characteristic function of $Z_V^*(t)$ is

$$\begin{aligned} \Phi^{Z_V^*(t)}(u) &= E\left[e^{iuCZ_I^*(t)} \right] \\ &= e^{a_I\left(\sqrt{\tilde{b}_I^2 - 2iCu} - \sqrt{\tilde{b}_I^2 - 2iuCe^{\lambda_I t}}\right)} \\ &= e^{a_I\sqrt{C}\left(\sqrt{\left(\frac{\tilde{b}_I}{\sqrt{C}}\right)^2 - 2iu} - \sqrt{\left(\frac{\tilde{b}_I}{\sqrt{C}}\right)^2 - 2iue^{\lambda_I t}}\right)}. \end{aligned} \tag{4.4.9}$$

Notice that from equations (4.4.8) and (4.4.9) we have that V is an $IG(a_V, b_V)$-Ornstein-Uhlenbeck process with parameter $\lambda_V > 0$, where a_V, b_V and λ_V are defined as in equations (4.4.5), (4.4.6) and (4.4.7) respectively. □

As in Chapter 3 for pricing vanilla options with expiration date T and strike K, we need to differentiate between the cases when $T \leq \tau$ and the cases when $T > \tau$. In the case when $T \leq \tau$ we have a generalization of the Black-Scholes-Merton formula for the price of a call option (see Chapter 3 Section 3.7.3). When the expiration date $T > \tau$ we need to compute the characteristic function of $X(T)$.

Proposition 4.4.2. *Let $u \in \mathbb{R}$, the characteristic function of $X(T)$ can be expressed as:*

$$\Phi^{X(T)}(u) = e^{iurT - \frac{\sigma_P^2}{2}\epsilon^0(iu+u^2)}\Phi^{Y(T-\tau)}(u),$$

where Y is a process that satisfies the following stochastic differential equation

$$dY(t) = -\frac{1}{2}\sigma_P^2 I^-(t)dt + \sigma_P\sqrt{I^-(t)}dB_P^*(t) \text{ with } Y(0) = x,$$

where B_P^* is a Brownian motion, ϵ^0 is defined as

$$\epsilon^0 = \int_{-\tau}^{0} \phi^I(s)ds$$

and the characteristic function of $Y(t)$ is

$$\begin{aligned}\Phi^{Y(t)}(u) \\ = \exp\left\{iux + \frac{1}{2}\lambda_V^{-1}(-u^2 - iu)\left(1 - e^{-\lambda_V t}\right)V(0) + a_V\left(\sqrt{b_V^2 - 2f_1(u)} - b_V\right)\right\} \\ \exp\left\{\frac{2a_V f_2(u)}{\sqrt{2f_2(u) - b_V^2}}\left[\arctan\left(\sqrt{\frac{b_V^2}{2f_2(u) - b_V^2}}\right) - \arctan\left(\sqrt{\frac{b_V^2 - 2f_1(u)}{2f_2(u) - b_V^2}}\right)\right]\right\},\end{aligned} \tag{4.4.10}$$

with

$$\begin{aligned} f_1(u) &= -\frac{1}{2}(u^2 + iu)\left(1 - e^{-\lambda_V t}\right)\lambda_V^{-1}, \\ f_2(u) &= -\frac{1}{2}(u^2 + iu)\lambda_V^{-1}, \\ a_V &= \sigma_P a_I, && (4.4.11) \\ b_V &= \frac{\tilde{b}_I}{\sigma_P}, && (4.4.12) \\ \lambda_V &= \lambda_I, && (4.4.13) \\ V(0) &= \sigma_P^2 I(0). \end{aligned}$$

Proof. As we showed in Section 3.7.4, the random variable $X(T)$ can be written as

$$\begin{aligned} X(T) &= rT - \frac{1}{2}\int_0^{\tau} \sigma_P^2\phi^I(s-\tau)ds + \int_0^{\tau} \sigma_P\sqrt{\phi^I(s-\tau)}dW_P^*(s) \\ &\quad + x - \frac{1}{2}\int_0^{T-\tau} \sigma_P^2 I^-(s)ds + \int_0^{T-\tau} \sigma_P\sqrt{I^-(s)}dB_P^*(s), \end{aligned}$$

where this result comes from the application of Theorem 1.2.1 and the process B_P^* defined as $B_P^*(t) = W_P^*(t+\tau) - W_P^*(\tau)$ for $t \geq 0$, is a Brownian motion. Let us now define the process Y as

$$Y(t) = x - \frac{1}{2}\int_0^t \sigma_P^2 I^-(s)ds + \int_0^t \sigma_P\sqrt{I^-(s)}dB_P^*(s) \text{ for } t \geq 0. \tag{4.4.14}$$

From equation (4.4.14) we have that Y satisfies the stochastic differential equation

$$dY(t) = -\frac{1}{2}\sigma_P^2 I^-(t)dt + \sigma_P\sqrt{I^-(t)}dB_P^*(t) \text{ with } Y(0) = x.$$

Notice that the random variable $X(T)$ can be expressed in terms of the random variable $Y(T-\tau)$ as

$$X(T) = rT - \frac{1}{2}\int_0^\tau \sigma_P^2\phi^I(s-\tau)ds + \int_0^\tau \sigma_P\sqrt{\phi^I(s-\tau)}dW_P^*(s) + Y(T-\tau).$$

Because $B_P^*(t)$ is independent of $\mathcal{F}_\tau^{W_P^*}$ for all $t \geq 0$ then we have that $Y(t)$ is independent of $\mathcal{F}_\tau^{W_P^*}$ for all $t \geq 0$.

Let $u \in \mathbb{R}$, if we condition with respect to the σ-algebra $\mathcal{F}_\tau^{W_P^*}$ then we have that the characteristic function of $X(T)$ can be expressed as

$$\begin{aligned}
\Phi^{X(T)}(u) &= E\left[e^{iurT-\frac{iu}{2}\int_0^\tau \sigma_P^2\phi^I(s-\tau)ds+iu\int_0^\tau \sigma_P\sqrt{\phi^I(s-\tau)}dW_P^*(s)+iuY(T-\tau)}\right] \\
&= E_{\mathbb{Q}}\left[e^{iurT-\frac{iu}{2}\int_0^\tau \sigma_P^2\phi^I(s-\tau)ds+iu\int_0^\tau \sigma_P\sqrt{\phi^I(s-\tau)}dW_P^*(s)}\right] E_{\mathbb{Q}}\left[e^{iuY(T-\tau)}\right] \\
&= e^{iurT-\frac{\sigma_P^2}{2}\epsilon^0(iu+u^2)}E_{\mathbb{Q}}\left[e^{iuY(T-\tau)}\right].
\end{aligned}$$

For computing the characteristic function of $X(T)$ first we need to compute the characteristic function of $Y(T-\tau)$. If we define the process V as $V(t) = \sigma_P^2 I(t)$ for $t \geq 0$ then we have from Proposition 4.4.1 that V is an $IG(a_V, b_V)$-Ornstein-Uhlenbeck process with parameter $\lambda_V > 0$, where

$$\begin{aligned}
a_V &= \sigma_P a_I, \\
b_V &= \frac{\tilde{b}_I}{\sigma_P}, \\
\lambda_V &= \lambda_I.
\end{aligned}$$

The process $Y(T-\tau)$ can be written in terms of the process V as

$$Y(T-\tau) = x - \frac{1}{2}\int_0^{T-\tau} V^-(u)du + \int_0^{T-\tau} \sqrt{V^-(u)}dB_P^*(u).$$

The process (Y, V) is a Barndorff-Nielsen and Shephard model (Barndorff-Nielsen & Shephard, 2001) and the characteristic function of $Y(t)$ with $t \geq 0$, that is defined as

$$\Phi^{Y(t)}(u) = E\left[e^{iuY(t)}\right] \text{ for } u \in \mathbb{R}$$

satisfies the equation (4.4.10) (Schoutens, 2003, p. 88). We have just shown that

$$\Phi^{X(T)}(u) = e^{iurT-\frac{\sigma_P^2}{2}\epsilon^0(iu+u^2)}\Phi^{Y(T-\tau)}(u).$$

□

Now that we are able to compute the characteristic function of the log-price at time T, we would like to know when the exponential moments of the log-price are finite.

Proposition 4.4.3. *Let $\tilde{\delta} \in \mathbb{R}$, then we have that*

$$E_{\mathbb{Q}}[e^{\tilde{\delta}X(T)}] < \infty, \text{ for all } \tilde{\delta} \in (\theta_-, \theta^+),$$

where

$$\theta^+ = \frac{1}{2} + \sqrt{\frac{1}{4} + \frac{\tilde{b}_I^2}{\sigma_P^2} \frac{\lambda_I}{1 - e^{-\lambda_I(T-\tau)}}}, \tag{4.4.15}$$

$$\theta_- = \frac{1}{2} - \sqrt{\frac{1}{4} + \frac{\tilde{b}_I^2}{\sigma_P^2} \frac{\lambda_I}{1 - e^{-\lambda_I(T-\tau)}}}. \tag{4.4.16}$$

Proof. As we have shown in the proof of Proposition 4.4.2 we can express the random variable $X(T)$ as

$$X(T) = rT - \frac{1}{2}\int_0^\tau \sigma_P^2 \phi^I(u-\tau)du + \int_0^\tau \sigma_P \sqrt{\phi^I(u-\tau)}dW_P^*(u) + Y(T-\tau),$$

where

$$Y(t) = x - \frac{1}{2}\int_0^t V^-(u)du + \int_0^t \sqrt{V^-(u)}dB_P^*(u) \text{ for } t \geq 0,$$

the process B_P^* defined as $B_P^*(t) = W_P^*(t+\tau) - W_P^*(\tau)$ for $t \geq 0$, is a Brownian motion and $V(t) = \sigma_P^2 I(t)$ for $t \geq 0$. Remember that by Proposition 4.4.1 the process V is an $IG(a_V, b_V)$-Ornstein-Uhlenbeck process with parameter $\lambda_V > 0$ where a_V, b_V and λ_V satisfy the equations (4.4.11), (4.4.12) and (4.4.13) respectively.

Notice that $B_P^*(t)$ is independent of $\mathcal{F}_\tau^{W_P^*}$ for all $t \geq 0$, hence $Y(T-\tau)$ is independent of $\mathcal{F}_\tau^{W_P^*}$. Conditioning with respect to $\mathcal{F}_\tau^{W_P^*}$ we can express the exponential moment $E_{\mathbb{Q}}[e^{\tilde{\delta}X(T)}]$ as

$$\begin{aligned} E_{\mathbb{Q}}\left[e^{\tilde{\delta}X(T)}\right] &= E_{\mathbb{Q}}\left[e^{\tilde{\delta}rT - \frac{\tilde{\delta}}{2}\int_0^\tau \sigma_P^2\phi^I(u-\tau)du + \tilde{\delta}\int_0^\tau \sigma_P\sqrt{\phi^I(u-\tau)}dW_P^*(u)}\right. \\ &\quad \left. E_{\mathbb{Q}}\left[e^{\tilde{\delta}Y(T-\tau)}\middle| \mathcal{F}_\tau^{W_P^*}\right]\right] \\ &= E_{\mathbb{Q}}\left[e^{\tilde{\delta}rT - \frac{\tilde{\delta}}{2}\int_0^\tau \sigma_P^2\phi^I(u-\tau)du + \tilde{\delta}\int_0^\tau \sigma_P\sqrt{\phi^I(u-\tau)}dW_P^*(u)}\right] \\ &\quad E_{\mathbb{Q}}\left[e^{\tilde{\delta}Y(T-\tau)}\right]. \end{aligned} \tag{4.4.17}$$

The first expected value that appears in equation (4.4.17) is the moment generating function of a normal random variable, hence this expectation is finite. So $E_{\mathbb{Q}}\left[e^{\tilde{\delta}X(T)}\right]$ will be finite if and only if $E_{\mathbb{Q}}\left[e^{\tilde{\delta}Y(T-\tau)}\right]$ is finite.

From Nicolato and Venardos (2003) we have that $E_{\mathbb{Q}}\left[e^{\tilde{\delta}Y(T-\tau)}\right]$ is finite if

$$\tilde{\delta} \in (\theta_-, \theta^+),$$

where

$$\theta^+ = \inf_{0 \leq s < T-\tau} \left\{\frac{1}{2} + \sqrt{\frac{1}{4} + 2\hat{\theta}\epsilon(s, T-\tau)^{-1}}\right\},$$

$$\theta_- = \sup_{0 \leq s < T-\tau} \left\{\frac{1}{2} - \sqrt{\frac{1}{4} + 2\hat{\theta}\epsilon(s, T-\tau)^{-1}}\right\},$$

the function $\epsilon(s,t)$ is defined as

$$\epsilon(s,t) = \frac{1-e^{-\lambda_V(T-s)}}{\lambda_V} \text{ for } s \in [0,t].$$

and

$$\hat{\theta} = \sup\left\{\theta \in \mathbb{R} : \frac{a_V\theta}{\sqrt{b_V^2 - 2\theta}} < \infty\right\}.$$

It is easy to see that $\hat{\theta} = \frac{b_V^2}{2}$, so the open interval (θ_-, θ^+) can be written as

$$\begin{aligned}
\theta^+ &= \inf_{0\le s<T-\tau}\left\{\frac{1}{2} + \sqrt{\frac{1}{4} + b_V^2\frac{\lambda_V}{1-e^{-\lambda_V(T-\tau-s)}}}\right\},\\
\theta_- &= \sup_{0\le s<T-\tau}\left\{\frac{1}{2} - \sqrt{\frac{1}{4} + b_V^2\frac{\lambda_V}{1-e^{-\lambda_V(T-\tau-s)}}}\right\}.
\end{aligned}$$

Because $\epsilon(s, T-\tau)^{-1}$ is an increasing function with respect to s, when $s \in [0, T-\tau]$ we have that θ^+ and θ_- satisfy the equations (4.4.15) and (4.4.16) respectively. □

4.5 Market option prices

Using the model that is introduced in this chapter, we would like to price European options and compare the results obtained with the true market prices. We use the same data that is explained in Chapter 1. As in that chapter, we compute the root mean square error and the relative root mean square error in order to evaluate the performance of our model.

As we saw in Chapter 2, we know that the logarithm of the number of unique active addresses follows an inverse Gaussian Ornstein-Uhlenbeck process. Using this proxy for the interest, we estimate the parameters of the model using the techniques explained in Section 4.2. The estimated parameters for each temporal window are shown in Figure 4.1 and Figure 4.2. In Figure 4.1, we show the estimated parameters of the interest process, as in Chapter 2 the values change through time.

In Figure 4.2 we show the estimated parameters that appear on equation (4.1.4). The estimated values of the parameter τ are in general different from zero.

With the values of the estimated parameters shown in Figures 4.1–4.2 we compute the option prices with $\tilde{b}_I = b_I$. That is, firstly we compute the option prices with an interest process I, which has the same parameters under the probability $\mathbb{P}$ and under the probability $\mathbb{Q}$.

As we can observe in Figure 4.1, the estimated values of the parameter λ_I are above 100. For these values of λ_I the computer is not able to compute the characteristic function shown in Proposition 4.4.2. Since the value of λ_I is high, the following value that appears in the characteristic function satisfies the following:

$$\arctan\left(\sqrt{\frac{b_V^2 - 2f_1(u)}{2f_2(u) - b_V^2}}\right) \approx \arctan(i).$$

But the complex arctan function is not defined at i, hence the computer is not able to compute the characteristic function. Because of that, we need an approximation formula to be able to compute the prices of European options. This approximation is developed in Section 4.5.1.

4.5.1 Approximation

Due to the fact that it is not possible to use the characteristic function obtained in Proposition 4.4.2 for computing options with expiration date $T > \tau$, we will use a formula that approximates the value of an European call option with strike price K and expiration date T. Here, we will assume that $T > \tau$.

If we define the process V as

$$V(t) = \sigma_P^2 I(t) \text{ for } t \geq 0.$$

It is not difficult to see that the price of an European call option with strike K and expiration date $T > \tau > 0$ can be expressed as:

$$C(0) = E\left[C_{BS}(v)\right] \tag{4.5.1}$$

where

$$C_{BS}(v) = P(0)\phi\left(d_1(v)\right) - Ke^{-rT}\phi(d_2(v))$$

and

$$\begin{aligned} d_1(v) &= \frac{\log(\frac{P(0)}{K}) + (r + \frac{1}{2}v)T}{\sqrt{vT}}, \\ d_2(v) &= d_1(v) - \sqrt{vT}, \\ \phi(x) &= \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{\frac{-z^2}{2}} dz, \\ v &= \frac{1}{T}\int_0^T V(s-\tau)ds. \end{aligned}$$

Equation (4.5.1) is usually called the Hull-White formula (Hull & White, 1987).

Proceeding as in the work of Hull and White (1987), we can use the Taylor expansion to approximate equation (4.5.1) in terms of the moments of v.

Proposition 4.5.1. *The price of the call option can be expressed as:*

$$C(0) = C_{BS}(\mu_v) + \frac{1}{2}\sigma_v^2 \frac{\partial^2 C_{BS}}{\partial v^2}(\mu_v) + E[Err],$$

where

$$\begin{aligned} \mu_v &= E[v] \\ \sigma_v^2 &= Var[v] \\ Err &= \frac{1}{6}(v-\mu_v)^3 \frac{\partial^3 C_{BS}}{\partial v^3}(\epsilon), \text{ and } \epsilon \text{ is a point between } v \text{ and } \mu_v. \end{aligned}$$

Proof. Let us perform the second order Taylor expansion of the function $C_{BS}(.)$ around the value μ_v:

$$C_{BS}(v) = C_{BS}(\mu_v) + (v-\mu_v)\frac{\partial C_{BS}}{\partial v}(\mu_v) + \frac{1}{2}(v-\mu_v)^2 \frac{\partial^2 C_{BS}}{\partial v^2}(\mu_v) + Err, \tag{4.5.2}$$

where the error term is

$$Err = \frac{1}{6}(v-\mu_v)^3 \frac{\partial^3 C_{BS}}{\partial v^3}(\epsilon), \text{ and } \epsilon \text{ is a point between } v \text{ and } \mu_v.$$

From equations (4.5.2) and (4.5.1) we obtain that

$$\begin{aligned}
C(0) &= E\left[C_{BS}(v)\right] \\
&= E\left[C_{BS}(\mu_v) + (v-\mu_v)\frac{\partial C_{BS}}{\partial v}(\mu_v) + \frac{1}{2}(v-\mu_v)^2\frac{\partial^2 C_{BS}}{\partial v^2}(\mu_v) + Err\right] \\
&= C_{BS}(\mu_v) + E[(v-\mu_v)]\frac{\partial C_{BS}}{\partial v}(\mu_v) + \frac{1}{2}E[(v-\mu_v)^2]\frac{\partial^2 C_{BS}}{\partial v^2}(\mu_v) + E[Err] \\
&= C_{BS}(\mu_v) + \frac{1}{2}\sigma_v^2\frac{\partial^2 C_{BS}}{\partial v^2}(\mu_v) + E[Err].
\end{aligned}$$

□

If we are able to compute the first moment and the second moment of the random variable $v = \frac{1}{T}\int_0^T V(s-\tau)ds$ then we can apply Proposition 4.5.1 to approximate the value of an European call option. Of course, we also need to compute the derivative:

$$\frac{\partial^2 C_{BS}}{\partial v^2}.$$

Proposition 4.5.2. *It is possible to show that:*

$$\frac{\partial^2 C_{BS}}{\partial v^2}(v) = \frac{1}{4v^{3/2}}\frac{P(0)\sqrt{T}}{\sqrt{2\pi}}\exp\left\{\frac{-d_1^2}{2}\right\}(-1 + d_1(v)d_2(v))$$

Proof. In the Black-Scholes model we have that the price of call option is:

$$C_{BS}(\sigma^2) = P(0)\phi\left(d_1(\sigma^2)\right) - Ke^{-rT}\phi(d_2(\sigma^2))$$

where

$$\begin{aligned}
d_1(\sigma^2) &= \frac{\log\left(\frac{P(0)}{K}\right) + \left(r + \frac{1}{2}\sigma^2\right)T}{\sqrt{\sigma^2 T}}, \\
d_2(\sigma^2) &= d_1(\sigma^2) - \sqrt{\sigma^2 T},
\end{aligned}$$

and $\sigma > 0$ (Gulisashvili, 2012, p. 235). In addition, the first derivative with respect to σ (called vega) and the second derivative with respect to σ (called volga) have the following form:

$$\frac{\partial C_{BS}}{\partial \sigma}(\sigma^2) = \frac{P(0)\sqrt{T}}{\sqrt{2\pi}}\exp\left\{\frac{-d_1^2(\sigma^2)}{2}\right\}, \tag{4.5.3}$$

$$\frac{\partial^2 C_{BS}}{\partial \sigma^2}(\sigma^2) = \frac{P(0)\sqrt{T}}{\sqrt{2\pi}\sigma}d_1(\sigma^2)d_2(\sigma^2)\exp\left\{\frac{-d_1^2(\sigma^2)}{2}\right\}, \tag{4.5.4}$$

(Gulisashvili, 2012, p. 237). Let us define $v = \sigma^2$, then we have:

$$\frac{\partial C_{BS}}{\partial \sigma}(\sigma^2) = \frac{\partial C_{BS}}{\partial \sigma}(v) = \frac{\partial C_{BS}}{\partial v}(v)\frac{\partial v}{\partial \sigma} = \frac{\partial C_{BS}}{\partial v}(v)2\sigma$$

hence:

$$\begin{aligned}
\frac{\partial C_{BS}}{\partial v}(v) &= \frac{\partial C_{BS}}{\partial \sigma}(v)\frac{1}{2\sqrt{v}} \\
&= \frac{1}{2\sqrt{v}}\frac{P(0)\sqrt{T}}{\sqrt{2\pi}}\exp\left\{\frac{-d_1^2(v)}{2}\right\} \\
&= \frac{1}{2\sigma}\frac{P(0)\sqrt{T}}{\sqrt{2\pi}}\exp\left\{\frac{-d_1^2(\sigma^2)}{2}\right\}.
\end{aligned}$$

Let us define the function f as:

$$f(v) = \frac{1}{2\sqrt{v}}\frac{P(0)\sqrt{T}}{\sqrt{2\pi}}\exp\left\{\frac{-d_1^2(v)}{2}\right\} = \frac{1}{2\sigma}\frac{P(0)\sqrt{T}}{\sqrt{2\pi}}\exp\left\{\frac{-d_1^2(\sigma^2)}{2}\right\} = f(\sigma^2).$$

By the chain rule, we have that:

$$\frac{\partial f}{\partial \sigma}(v) = \frac{\partial f}{\partial v}(v)\frac{\partial v}{\partial \sigma} = \frac{\partial f}{\partial v}(v)2\sigma.$$

Lastly, we have that:

$$\begin{aligned}
\frac{\partial f}{\partial \sigma}(\sigma^2) &= -\frac{1}{2\sigma^2}\frac{P(0)\sqrt{T}}{\sqrt{2\pi}}\exp\left\{\frac{-d_1^2(\sigma^2)}{2}\right\} + \frac{1}{2\sigma}\frac{\partial}{\partial \sigma}\left(\frac{P(0)\sqrt{T}}{\sqrt{2\pi}}\exp\left\{\frac{-d_1^2(\sigma^2)}{2}\right\}\right) \\
&= -\frac{1}{2\sigma^2}\frac{P(0)\sqrt{T}}{\sqrt{2\pi}}\exp\left\{\frac{-d_1^2(\sigma^2)}{2}\right\} \\
&\quad +\frac{1}{2\sigma}\frac{P(0)\sqrt{T}}{\sqrt{2\pi}\sigma}d_1(\sigma^2)d_2(\sigma^2)\exp\left\{\frac{-d_1^2(\sigma^2)}{2}\right\} \\
&= \frac{1}{2\sigma^2}\frac{P(0)\sqrt{T}}{\sqrt{2\pi}}\exp\left\{\frac{-d_1^2(\sigma^2)}{2}\right\}\left(-1 + d_1(\sigma^2)d_2(\sigma^2)\right),
\end{aligned}$$

where the second equality comes from equation (4.5.4). So, at the end we have just shown that:

$$\begin{aligned}
\frac{\partial^2 C_{BS}}{\partial v^2}(v) &= \frac{\partial^2 f}{\partial v^2}(v) \\
&= \frac{1}{2\sigma}\frac{\partial f}{\partial \sigma}(\sigma^2) \\
&= \frac{1}{4\sigma^3}\frac{P(0)\sqrt{T}}{\sqrt{2\pi}}\exp\left\{\frac{-d_1^2(\sigma^2)}{2}\right\}\left(-1 + d_1(\sigma^2)d_2(\sigma^2)\right) \\
&= \frac{1}{4v^{3/2}}\frac{P(0)\sqrt{T}}{\sqrt{2\pi}}\exp\left\{\frac{-d_1^2(v)}{2}\right\}\left(-1 + d_1(v)d_2(v)\right),
\end{aligned}$$

as required. □

4.5.2 Moments

Now we would like to compute the first and second moments of the random variable v. In this case, from Proposition 4.4.1 we have that V is an $IG(a_V, b_V)$-Ornstein-Uhlenbeck process with parameter $\lambda_V > 0$ where a_V, b_V and λ_V are defined as in equations (4.4.11)–(4.4.13) respectively. We define the process $Z_V = (Z_V(t))_{t\geq 0}$ as the background driven Lévy process of the Ornstein-Uhlenbeck process V. Since V is an inverse Gaussian Ornstein-Uhlenbeck process, we have that the cumulant function of the random variable $Z_V(1)$ is:

$$k(\theta) = \log\left(E\left[e^{\theta Z_V(1)}\right]\right) = \theta a_V(b_V^2 - 2\theta)^{-1/2} \text{ for } \theta \in \mathbb{R}, \tag{4.5.5}$$

(Nicolato & Venardos, 2003, p. 449). Remember that the n-th cumulant of $Z_V(1)$ is defined as

$$k_n = \left.\frac{\partial^n k}{\partial \theta^n}(\theta)\right|_{\theta=0}, \tag{4.5.6}$$

(Pascucci, 2011, p. 460). Based in the cumulant function, we can use the following theorem to compute the moments of v.

Theorem 4.5.1. *(Barndorff-Nielsen & Shephard, 2003, p. 289) If k_1, k_2 are the first two cumulants of the random variable $Z_V(1)$, then for $t \geq 0$ we have:*

$$E\left[\int_0^t V(s)ds\right] = \frac{1 - e^{-\lambda_V t}}{\lambda_V}(V(0) - k_1) + k_1 t, \tag{4.5.7}$$

$$Var\left[\int_0^t V(s)ds\right] = \frac{k_2}{\lambda_V^2}\left(\lambda_V t - 2 + 2e^{-\lambda_V t} + \frac{1}{2} - \frac{1}{2}e^{-\lambda_V t}\right). \tag{4.5.8}$$

From equations (4.5.5) and (4.5.6) we have that:

$$\begin{aligned} k_1 &= \left. a_V(b_V^2 - 2\theta)^{-1/2} + \theta a_V(b_V^2 - \theta)^{-3/2}\right|_{\theta=0} \\ &= \frac{a_V}{b_V}, \\ k_2 &= \left. a_V(b_V^2 - \theta)^{-3/2} + a_V(b_V^2 - 2\theta)^{-3/2} + \theta a_V 3(b_V^2 - 2\theta)^{-5/2}\right|_{\theta=0} \\ &= \frac{2a_V}{b_V^3}. \end{aligned}$$

The integral v can be expressed as

$$v = \frac{1}{T}\int_0^T V(s-\tau)ds = \frac{\epsilon_1}{T} + \frac{1}{T}\int_0^{T-\tau} V(s)ds,$$

where $\epsilon_1 = \int_0^\tau V(s-\tau)ds = \sigma_P^2 \int_0^\tau \phi^I(s-\tau)ds$.

So at the end we can write:

$$\begin{aligned} \mu_v &= E\left[\frac{1}{T}\int_0^T V(s-\tau)ds\right] \\ &= \frac{\epsilon_1}{T} + \frac{1}{T}E\left[\int_0^{T-\tau} V(s)ds\right], \\ \sigma_v^2 &= Var\left[\frac{1}{T}\int_0^T V(s-\tau)ds\right] \\ &= \frac{1}{T^2}Var\left[\int_0^{T-\tau} V(s)ds\right]. \end{aligned}$$

4.5.3 Numerical experiments

For checking the reliability of the approximation, we generate European call prices using the approximation method explained in Section (4.5.1) and prices using the characteristic function method explained in Section (4.4). We pick the following values for the parameters:

$$\sigma_P = 0.5, \quad a_I = 4, \quad b_I = 10, \quad \lambda_I = 5, \quad \tau = 10\Delta,$$

$\Delta = \frac{1}{365}$ and initial values $P(0) = 100$ and

$$\phi^I(t) = 0.4 + 0.2\cos^2\left(30t + \frac{\pi}{2}\right) \text{ for } t \in [-\tau, 0].$$

As we can see in Table 4.2, the values given by the approximation formula are near to the values given by the characteristic function.

Expiration Date	Strike	Approx. values	Characteristic function
1	100	7.372011	7.372036
1	200	0.000682	0.000687
1	50	50.000341	50.000343
2	100	10.540667	10.540675
2	200	0.053607	0.053604
2	50	50.026803	50.026802
0.5	100	5.10482	5.10488
0.5	200	2.162681210^{-7}	2.7930110^{-7}

TABLE 4.2: Prices of European call options given by the approximation formula in Proposition 4.5.1 and by characteristic function obtained in Proposition 4.4.2.

4.5.4 Results

Using the approximation formula constructed in Section 4.5.1, we price the market options with $\tilde{b}_I = b_I$. That is, the interest process satisfies the same distribution under the physical measure $\mathbb{P}$ and under the risk-neutral measure $\mathbb{Q}$. The root mean square error and the relative root mean square error obtained for each date are shown in Figure 4.3 and in Figure 4.4.

As we saw in Section 4.3.1, there are different risk-neutral measures $\mathbb{Q}$ for the different values of $\tilde{b}_I$. That is, we can select different values for the parameter $\tilde{b}_I$ for pricing options. As we did in Chapter 3, we select the parameter $\tilde{b}_I$ that minimizes the relative root mean square error. The values for the root mean square error and the relative root mean square error are shown in Figures 4.3–4.4. In Figure 4.5 we show the values of the estimated b_I and the calibrated $\tilde{b}_I$. We can observe that the optimal value of $\tilde{b}_I$ is on some occasions bigger than the estimated value of b_I and on some occasions smaller. We can observe that the estimated b_I is more stable than the calibrated one.

In addition, we compute the prices given by the Black-Scholes-Merton model. The parameters of the model are estimated by the maximum likelihood estimator method using historical data. The temporal windows used for this model are the same as the ones defined above. The root mean square error and the relative root mean square error obtained by the Black-Scholes-Merton model are shown in Figure 4.3 and in Figure 4.4 respectively. The results obtained by the Black-Scholes-Merton model are similar to the proposed model. However, when we calibrate the proposed model, the model defined in Section 4.1 obtains better results. As we did in Chapter 3, we calculate the ratio between the relative root mean square error obtained by the proposed model and the relative root mean square error obtained by the Black-Scholes-Merton model. This is shown in Figure 4.6. The results shown in Figure 4.6 are in line with the results shown in Figures 4.3 and 4.4.

4.6 Conclusion and future work

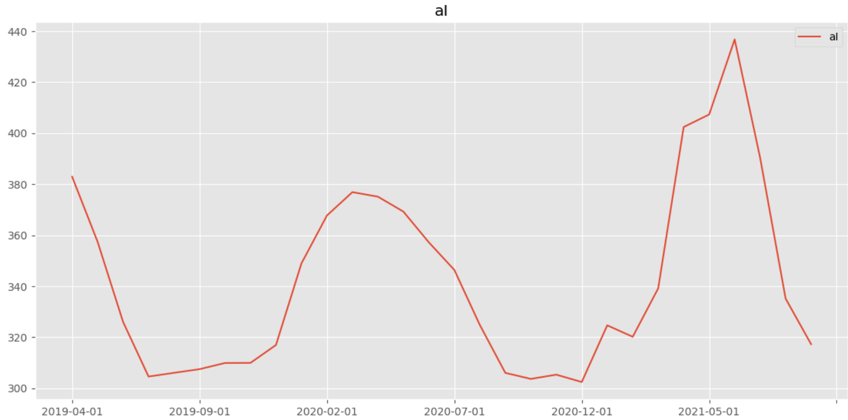
In this chapter, we proposed a model for pricing Bitcoin options which is similar to the model proposed in Chapter 3, but in this case we incorporate jumps or discontinuities into the volatility path. We showed how the estimation of the parameters of the model can be performed and how we can price options with it. In addition, we also constructed an approximation formula for European call options for the cases when the parameters are high, and it is not possible to compute the characteristic function. We also compared the prices given by our model with the real market data and showed that our model gives better results than the Black-Scholes-Merton model.

In the model developed in this chapter, we assume that the Brownian motion W_P and the Lévy process Z_I are independent. We would like to have asymmetry in the distribution of the returns. One way of obtaining asymmetry in the returns is by use of the following model

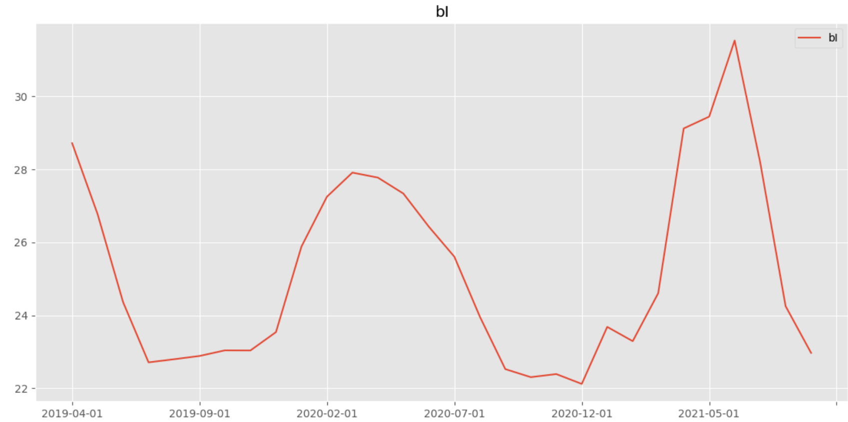
$$\begin{aligned} X(t) &= x + \mu t + \int_0^t \sigma_P \sqrt{I^-(u-\tau)}dW_P(u) \\ &\quad +\mathbb{1}_{(\tau,\infty)}(t)\rho Z_I(\lambda_I((t \vee \tau) - \tau)) \text{ with } X(0) = x \in \mathbb{R}, \qquad (4.6.1) \\ dI(t) &= -\lambda_I I(t)dt + dZ_I(\lambda_I t) \text{ when } t > 0 \qquad (4.6.2) \\ &\quad \text{with } I(t) = \phi^I(t) \text{ when } t \in [-L, 0], \end{aligned}$$

where $\mu, \rho \in \mathbb{R}$, $\sigma_P > 0$, the processes W_P, Z_I are independent and the process I is an $IG(a_I, b_I)$-Ornstein-Uhlenbeck process with parameter $\lambda_I > 0$. In future work we would like to study the model defined in equations (4.6.2)–(4.6.2).

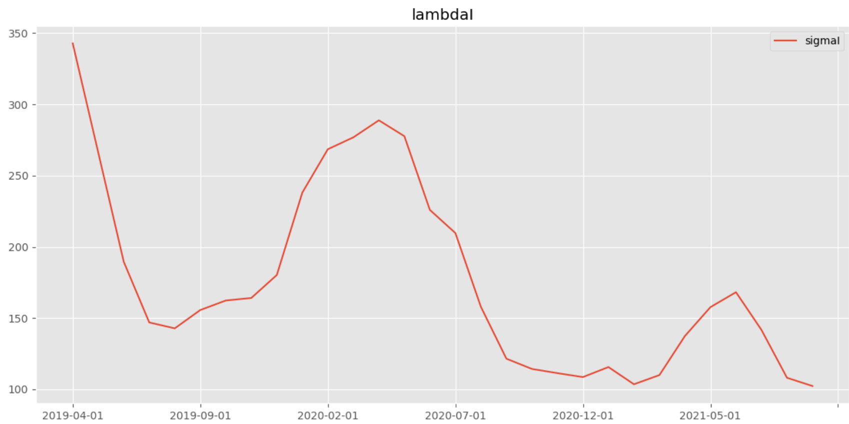
In the Barndorff-Nielsen and Shephard models (Barndorff-Nielsen & Shephard, 2001), the volatility process could be different types of Ornstein-Uhlenbeck processes. In our case, we only use an inverse Gaussian Ornstein-Uhlenbeck process. However, other types of Ornstein-Uhlenbeck processes could be used. For example, in the literature the gamma Ornstein-Uhlenbeck process or the tempered stable Ornstein-Uhlenbeck process are used(Schoutens, 2003, p.68–70).



(A) Estimation of the parameter a_I.

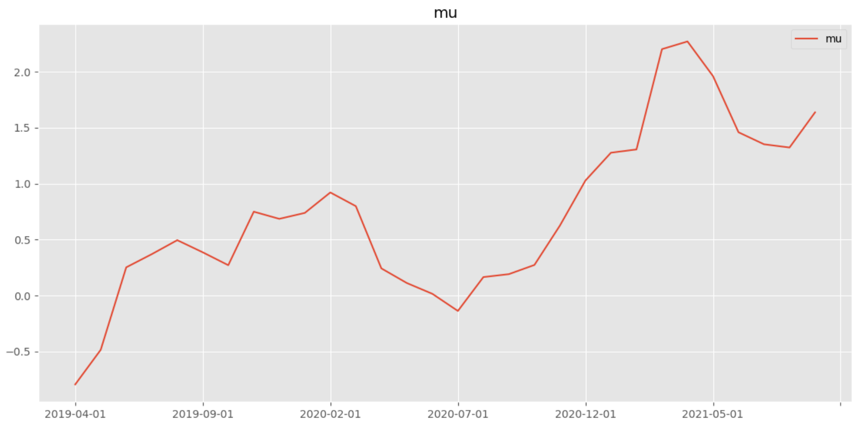


(B) Estimation of the parameter b_I.

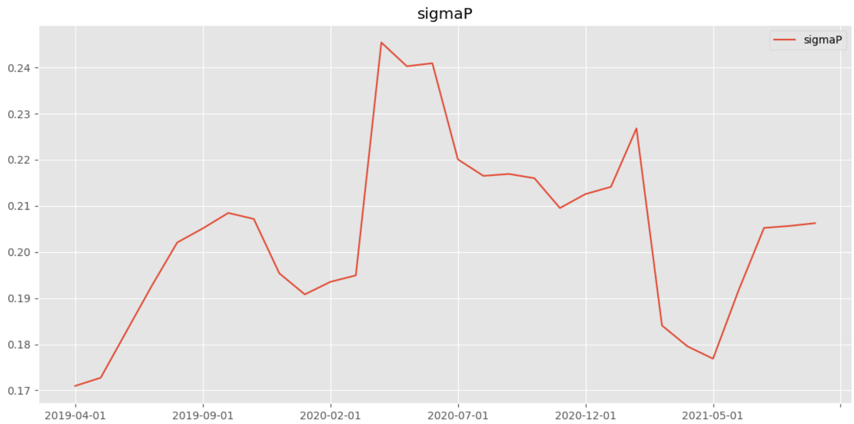


(C) Estimation of the parameter λ_I.

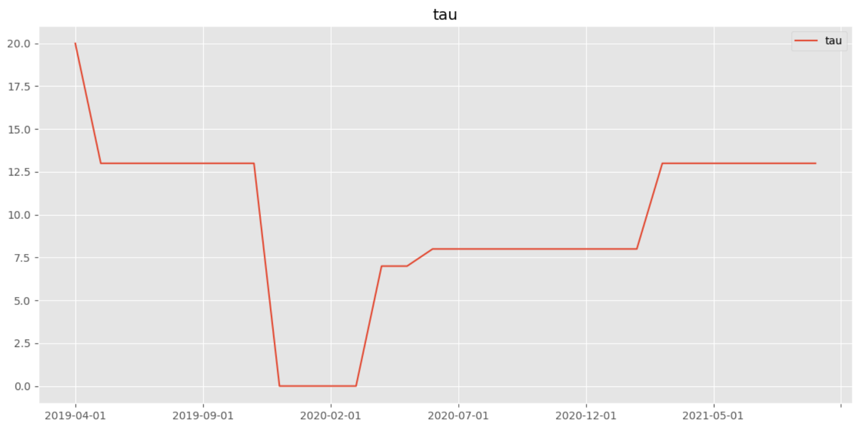
FIGURE 4.1: Estimated values for the parameters of the model defined in Section 4.1 related to the market attention process, when the proxy of the market interest is the logarithm of the number of unique active addresses.



(A) Estimation of the parameter μ.



(B) Estimation of the parameter σ_P.



(C) Estimation of the parameter τ.

FIGURE 4.2: Estimated values for the parameters of the model defined in Section 4.1 related to the price equation, when the proxy of the market interest is the logarithm of the number of unique active addresses.

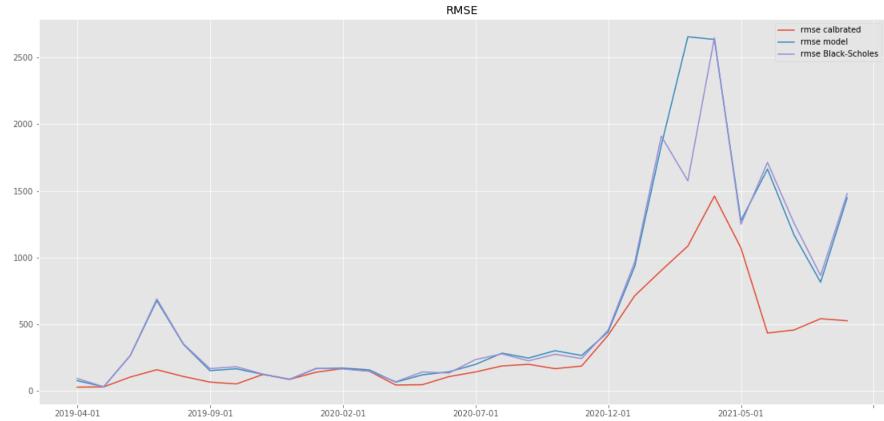


FIGURE 4.3: Root mean square error when the proxy of the interest is the logarithm of the number of unique active addresses.

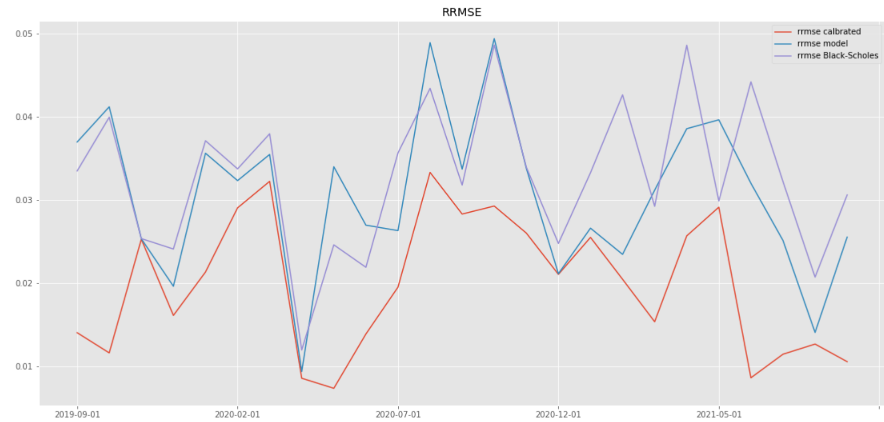


FIGURE 4.4: Relative root mean square error when the proxy of the interest is the logarithm of the number of unique active addresses.

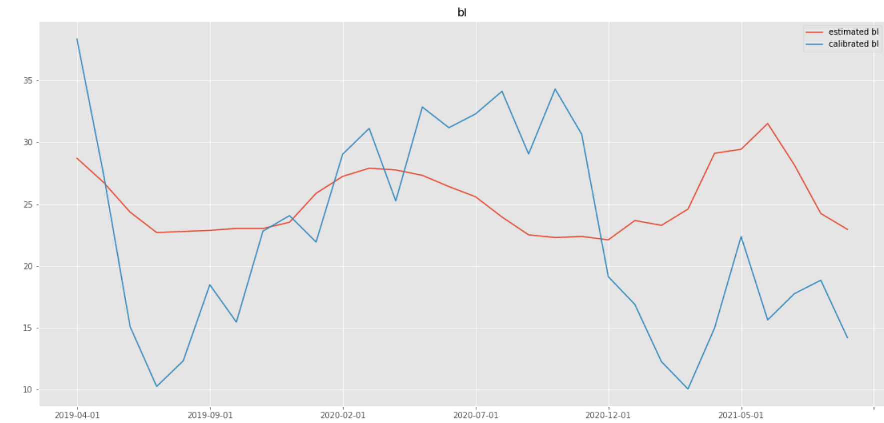


FIGURE 4.5: Calibrated values for b_I when the proxy of the interest is the logarithm of the number of unique active addresses.

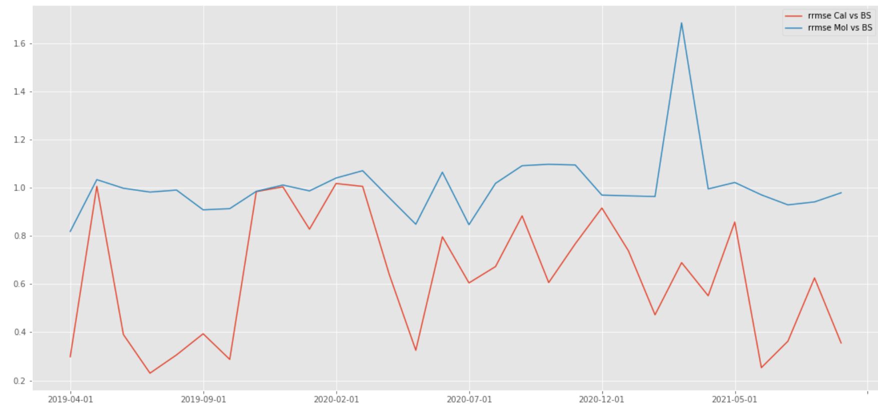


FIGURE 4.6: Relative root mean square error given by our model divided by the relative root mean square error obtained by the Black-Scholes-Merton model, when the proxy of the market interest is the logarithm of the number of unique active addresses.

Chapter 5

Time changed models

5.1 Introduction and proposed models

In the models that were proposed in Chapter 3 and in Chapter 4, the volatility of the log-prices is proportional to the market attention process. That is, in previous chapters we proposed stochastic volatility models in which the volatility is delayed by a positive parameter.

It is possible to obtain the effects of a stochastic volatility model by the use of a stochastic time changed model (Carr, Geman, Madan, & Yor, 2003). Apart from that, it has been observed that the logarithmic returns of Bitcoin prices tend to have fat tails and are not normally distributed (Chan, Chu, Nadarajah, & Osterrieder, 2017). Wang, Hou, Chen, and Härdle (2020) estimated several stochastic volatility models using Bitcoin prices. The results showed that the models that produce the best fit are the ones with jumps in the price and volatility structure. This seems to indicate that the conditional distribution of the returns given the volatility does not follow a normal distribution. This motivates the model proposed in this chapter.

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. In this space, let us define the price process P as in equation (5.1.2), and the interest process I, as the strong solution of equation (5.1.3). In addition, we define the stochastic time changed process T_I as

$$T_I(t) = \int_0^t I(s-\tau)ds, \qquad \text{for } t \geq 0. \tag{5.1.1}$$

The process T_I is positive and increasing, since we will assume that the process I is positive with probability one. So at the end we propose the following model:

$$\begin{aligned} P(t) &= p \exp\left\{ \left(\mu - \frac{\sigma^2}{2}\right) t + \sigma W_P(t) + X_P(T_I(t)) \right\}, \\ &\quad \text{for } t \geq 0 \quad \text{with } P(0) = p > 0, \end{aligned} \tag{5.1.2}$$

where $\mu \in \mathbb{R}$, $\sigma > 0$, $\tau \in [0, L]$ and $L > 0$, W_P is a Brownian motion, X_P is a Lévy process and I is the interest process. The processes W_P, X_P and I are independent. For the interest process I we have to distinguish between the behaviour when $t > 0$ and when $t \in [-L, 0]$. When $t > 0$ the process I is the strong solution of the following stochastic differential equation

$$dI(t) = a(I(t))dt + b(I(t))dX_I(t), \tag{5.1.3}$$

where X_I is a Lévy process independent of W_P and X_P, the functions $a : \mathbb{R} \to \mathbb{R}$, $b : \mathbb{R} \to (0, \infty)$ are such that the equation (5.1.3) has a strong solution. When

$t \in [-L, 0]$ the process I satisfies

$$I(t) = \phi^I(t),$$

where $\phi^I : [-L, 0] \to (0, \infty)$ is a continuous and deterministic function.

There are different possibilities for equation (5.1.3), but we impose the following requirements:

1. The strong solution of (5.1.3) has to be positive with probability 1. Hence the process T_I defined in equation (5.1.1) is increasing.
2. The integrated interest process has to be analytically tractable. That is, it should be possible to compute the characteristic function of the process $\left(\int_0^t I(s)ds\right)_{t\geq 0}$.

Having these requirements in mind, we propose two models for the interest process.

1. In one of the proposed models, the market attention process when $t > 0$ satisfies a Cox-Ingersoll-Ross process. This means that the process I is the strong solution of the following stochastic differential equation:

$$dI(t) = a_I(b_I - I(t))dt + \sigma_I \sqrt{I(t)}dW_I(t) \text{ when } t > 0, \tag{5.1.4}$$

   where W_I is a Brownian motion, $b_I \in \mathbb{R}$, $a_I, \sigma_I > 0$ and we impose the condition:

$$\frac{2a_I b_I}{\sigma_I^2} \geq 1. \tag{5.1.5}$$

   Condition (5.1.5) guarantees that I is greater than zero with probability 1 (Gulisashvili, 2012, Theorem 2.27).

2. In the other proposed model, the interest process when $t > 0$ satisfies an inverse Gaussian Ornstein-Uhlenbeck process with parameters $a_I, b_I, \lambda_I > 0$. So the interest I satisfies the stochastic differential equation:

$$dI(t) = -\lambda_I I(t)dt + dZ_I(\lambda_I t) \text{ with } I(0) = y_0, \tag{5.1.6}$$

   where Z_I is a Lévy process with Lévy triplet $(\gamma_I, 0, v_I)$ defined as

$$v_I(dx) = \frac{a_I}{2\sqrt{2\pi}} x^{-3/2}(1 + b_I^2 x)e^{-\frac{1}{2}b_I^2 x}\mathbb{1}_{(0,\infty)}(x), \tag{5.1.7}$$

$$\gamma_I = \int_{|x|<1} x v_I(dx). \tag{5.1.8}$$

In the case of the Lévy process X_P that appears in the price equation (5.1.2), different choices can be made. We will restrict ourselves to the case where X_P is a compound Poisson process, that is

$$X_P(t) = \sum_{k=1}^{N_P(t)} Z_k \text{ for } t \geq 0,$$

where N_P is a Poisson process with intensity parameter $\lambda_P > 0$ and $\{Z_k\}_{k=1}^{\infty}$ is a sequence of independent and identically distributed random variables that are independent of the Poisson process N_P.

So according to equation (5.1.2), the price process has a continuous part that behaves like a geometric Brownian motion and a discontinuous part given by a time changed compound Poisson process. Notice that when the interest process increases, the probability of a jump increases, provoking an increase in the volatility of the price. And when the interest process decreases, the probability of a jump decreases, making for a decrease in the volatility.

Different distributions can be chosen for the sequence of the random variables $\{Z_k\}_{k=1}^{\infty}$. Common choices that can be found in the literature are:

1. double exponential distribution (Kou, 2002),
2. normal distribution (Merton, 1976).

To keep things simple, let us assume that $Z_k \sim N(\eta, \delta^2)$ for all $k = 1, 2, \ldots$, with $\eta \in \mathbb{R}$ and $\delta > 0$. However, using a double exponential distribution will help us to capture the asymmetry of the jumps. Since X_P is a compound Poisson process with normally distributed jumps, we have that it has Lévy triplet $(\gamma_P, 0, \upsilon_P)$ with

$$\upsilon_P(dx) = \lambda_P f_N(x|\eta, \delta^2)dx \tag{5.1.9}$$

$$\gamma_P = \int_{|x|<1} x\upsilon_P(dx), \tag{5.1.10}$$

where $f_N(.|\eta, \delta^2)$ is the density function of a normal random variable with mean η and variance δ^2 (Cont & Tankov, 2004, Proposition 3.8 and p. 112).

Lastly, we are interested in defining the filtration with respect to the processes P and I are adapted. Let us define

$$\mathcal{F}^{W_P} = \left(\mathcal{F}_t^{W_P}\right)_{t\geq 0} \quad \text{and} \quad \mathcal{F}^{X_I} = \left(\mathcal{F}_t^{X_I}\right)_{t\geq 0}$$

to be the filtrations generated by W_P and X_I respectively. Remember that $X_I = W_I$ when the interest process satisfies a Cox-Ingersoll-Ross process (see equation (5.1.4)) and $X_I = Z_I$ when the market attention satisfies an inverse Gaussian Ornstein-Uhlenbeck process (see equation (5.1.6)). In addition define the filtration $\mathcal{F}^{X_P(T_I)} = \left(\mathcal{F}_t^{X_P(T_I)}\right)_{t\geq 0}$, where

$$\mathcal{F}_t^{X_P(T_I)} = \sigma\left(\{X_P(T_I(s)) : 0 \leq s \leq t\}\right)$$

for all $t \geq 0$. The general filtration $(\mathcal{F}_t)_{t\geq 0}$ is defined as

$$\mathcal{F}_t = \mathcal{F}_t^{W_P} \vee \mathcal{F}_t^{X_I} \vee \mathcal{F}_t^{X_P(T_I)}, \qquad \text{for } t \geq 0,$$

where $\mathcal{F}_t^{W_P} \vee \mathcal{F}_t^{X_I} \vee \mathcal{F}_t^{X_P(T_I)}$ is the smallest σ-algebra containing $\mathcal{F}_t^{W_P}$, $\mathcal{F}_t^{X_I}$ and $\mathcal{F}_t^{X_P(T_I)}$. The process I is adapted with respect to the filtration $\left(\mathcal{F}_t^{X_I}\right)_{t\geq 0}$, so I is adapted to the filtration $(\mathcal{F}_t)_{t\geq 0}$. Similarly P is adapted with respect to $\left(\mathcal{F}_t^{W_P} \vee \mathcal{F}_t^{X_P(T_I)}\right)_{t\geq 0}$, so it is adapted with respect to $(\mathcal{F}_t)_{t\geq 0}$.

Let us also define the delayed filtration $(\tilde{\mathcal{F}}_t)_{t\geq 0}$ as

$$\tilde{\mathcal{F}}_t = \begin{cases} \mathcal{F}_t^{W_P} \vee \mathcal{F}_t^{X_P(T_I)} & \text{if } t \leq \tau, \\ \mathcal{F}_t^{W_P} \vee \mathcal{F}_t^{X_P(T_I)} \vee \mathcal{F}_{t-\tau}^{X_I} & \text{if } t > \tau. \end{cases}$$

Notice that $\tilde{\mathcal{F}}_t \subseteq \mathcal{F}_t$ for $t \geq 0$, and that the process $(P(t), I(t-\tau))_{t\geq 0}$ is adapted with respect to the filtration $(\tilde{\mathcal{F}}_t)_{t\geq 0}$.

Apart from the stock price P and the interest process I, there is a bond or a market account $(B(t))_{t\geq 0}$ with known interest rate r, that satisfies:

$$B(t) = B(0)e^{rt} \text{ for } t \geq 0,$$

where $r \geq 0$.

5.2 Conditional distribution of the logarithmic returns

Before estimating the parameters of our model, let us study the conditional distribution of the logarithmic returns

$$R(s,t) = \log\left(\frac{P(t)}{P(s)}\right)$$

for $t \geq s$, given the σ-algebra $\mathcal{F}_H^{X_I}$, where $H \geq t \geq s \geq 0$. We will show that this conditional density function can be expressed as an infinite sum of normal density functions. This will allow us to estimate the parameters that appear in the price equation (5.1.2).

From equation (5.1.2), we have that the logarithmic return $R(s,t)$ can be expressed as

$$R(s,t) = \left(\mu - \frac{\sigma^2}{2}\right)(t-s) + \sigma\left(W_P(t) - W_P(s)\right) + X_P\left(T_I(t)\right) - X_P\left(T_I(s)\right)$$

for all $t \geq s \geq 0$.

Remark 5.2.1. *From Corollary C.1.1 we have that the increment $X_P\left(T_I(t)\right) - X_P\left(T_I(s)\right)$ has the same distribution as the random variable $X_P\left(T_I(t) - T_I(s)\right)$. Let us define $J_I^{s,t}$ as*

$$J_I^{s,t} = T_I(t) - T_I(s) = \int_s^t I(v-\tau)dv.$$

So at the end we have that

$$X_P\left(T_I(t)\right) - X_P\left(T_I(s)\right) \stackrel{d}{=} X_P\left(J_I^{s,t}\right), \tag{5.2.1}$$

where $\stackrel{d}{=}$ means that both sides of the equation (5.1.1) have the same distribution.

From Remark 5.2.1 we can write $R(s,t)$ as:

$$\begin{aligned} R(s,t) &= \left(\mu - \frac{\sigma^2}{2}\right)(t-s) + \sigma\left(W_P(t) - W_P(s)\right) + X_P\left(T_I(t)\right) - X_P\left(T_I(s)\right) \\ &\stackrel{d}{=} \left(\mu - \frac{\sigma^2}{2}\right)(t-s) + \sigma\left(W_P(t) - W_P(s)\right) + X_P\left(J_I^{s,t}\right). \end{aligned} \tag{5.2.2}$$

We would like to compute the conditional characteristic function of $R(s,t)$ given the σ-algebra $\mathcal{F}_H^{X_I}$.

Proposition 5.2.1. *Let $u \in \mathbb{R}$, then the conditional characteristic function of $R(s,t)$ given $\mathcal{F}_H^{X_I}$ has the following form:*

$$\begin{aligned}\Phi^{R(s,t)}\left(u\middle|\mathcal{F}_H^{X_I}\right) &= E\left[e^{iuR(s,t)}\middle|\mathcal{F}_H^{X_I}\right]\\ &= e^{iu\left(\mu-\frac{\sigma^2}{2}\right)(t-s)}e^{-\frac{u^2\sigma^2}{2}(t-s)}e^{-J_I^{s,t}\lambda_P}e^{J_I^{s,t}\lambda_P e^{iu\eta-\frac{1}{2}u^2\delta^2}}.\end{aligned} \quad (5.2.3)$$

Proof. Let $u \in \mathbb{R}$ then the conditional characteristic function of $R(s,t)$ given the σ-algebra $\mathcal{F}_H^{X_I}$ can be written as

$$\begin{aligned}\Phi_{R(s,t)}\left(u\middle|\mathcal{F}_H^{X_I}\right) &= E\left[e^{iuR(s,t)}\middle|\mathcal{F}_H^{X_I}\right]\\ &= E\left[e^{iu\left(\mu-\frac{\sigma^2}{2}\right)(t-s)+iu\sigma(W_P(t)-W_P(s))+iuX_P\left(J_I^{s,t}\right)}\middle|\mathcal{F}_H^{X_I}\right],\end{aligned}$$

where the last equality comes from (5.2.2).

First, let us study the distribution when the σ-algebra $\mathcal{F}_H^{X_I} \vee \sigma(X_P(J_I^{s,t}))$ is given. Because the random variable $W_P(t) - W_P(s)$ is independent of $\mathcal{F}_H^{X_I} \vee \sigma(X_P(J_I^{s,t}))$, we have that

$$\begin{aligned}&E\left[e^{iu\left(\mu-\frac{\sigma^2}{2}\right)(t-s)+iu\sigma(W_P(t)-W_P(s))+iuX_P(J_I^{s,t})}\middle|\mathcal{F}_H^{X_I}\vee\sigma(X_P(J_I^{s,t}))\right]\\ &\quad=e^{iu\left(\mu-\frac{\sigma^2}{2}\right)(t-s)}e^{iuX_P(J_I^{s,t})}E\left[e^{iu\sigma(W_P(t)-W_P(s))}\middle|\mathcal{F}_H^{X_I}\vee\sigma\left(X_P(J_I^{s,t})\right)\right]\\ &\quad=e^{iu\left(\mu-\frac{\sigma^2}{2}\right)(t-s)}e^{iuX_P(J_I^{s,t})}E\left[e^{iu\sigma(W_P(t)-W_P(s))}\right]\\ &\quad= e^{iu\left(\mu-\frac{\sigma^2}{2}\right)(t-s)}e^{iuX_P(J_I^{s,t})}e^{-\frac{u^2\sigma^2}{2}(t-s)},\end{aligned}$$

where the last equality comes from the fact that $W_P(t) - W_P(s) \sim N(0, t-s)$. By application of the tower property, we obtain

$$\begin{aligned}E\left[e^{iuR(s,t)}\middle|\mathcal{F}_H^{X_I}\right] &= E\left[E\left[e^{iuR(s,t)}\middle|\mathcal{F}_H^{X_I}\vee\sigma(X_P(J_I^{s,t}))\right]\middle|\mathcal{F}_H^{X_I}\right]\\ &= E\left[e^{iu\left(\mu-\frac{\sigma^2}{2}\right)(t-s)}e^{iuX_P(J_I^{s,t})}e^{-\frac{u^2\sigma^2}{2}(t-s)}\middle|\mathcal{F}_H^{X_I}\right]\\ &= e^{iu\left(\mu-\frac{\sigma^2}{2}\right)(t-s)}e^{-\frac{u^2\sigma^2}{2}(t-s)}E\left[e^{iuX_P(J_I^{s,t})}\middle|\mathcal{F}_H^{X_I}\right].\end{aligned}$$

Since X_P is a compound Poisson process with jumps given by a normal distribution, we have that the the characteristic exponent of X_P is

$$\Psi^{X_P}(u) = \lambda_P\left(e^{iu\eta-\frac{1}{2}u^2\delta^2}-1\right),$$

(Cont & Tankov, 2004, p.112). Applying Proposition C.0.1, we have that

$$\begin{aligned}E\left[e^{iuX_P(J_I^{s,t})}\middle|\mathcal{F}_H^{X_I}\right] &= e^{J_I^{s,t}\Psi^{X_P}(u)}\\ &= e^{J_I^{s,t}\lambda_P\left(e^{iu\eta-\frac{1}{2}u^2\delta^2}-1\right)}\\ &= e^{-J_I^{s,t}\lambda_P}e^{J_I^{s,t}\lambda_P\left(e^{iu\eta-\frac{1}{2}u^2\delta^2}\right)}.\end{aligned}$$

We have just shown the desired result. $\square$

Now we are ready to compute the conditional density function of $R(s,t)$ given $\mathcal{F}_H^{X_I}$.

Proposition 5.2.2. *Let $x \in \mathbb{R}$, the conditional density function of $R(s,t)$ given $\mathcal{F}_H^{X_I}$ has the following form:*

$$f_{s,t}^R\left(x\,\middle|\mathcal{F}_H^{X_I}\right) = \sum_{n=0}^{\infty} f_P\left(n\,\middle|\lambda_P J_I^{s,t}\right) f_N\left(x\,\middle|m_{R,n}, v_{R,n}^2\right), \tag{5.2.4}$$

where

$$\begin{aligned} m_{R,n} &= \left(\mu - \frac{\sigma^2}{2}\right)(t-s) + n\eta, \\ v_{R,n}^2 &= \sigma^2(t-s) + n\delta^2, \end{aligned}$$

where $f_P\left(.\,\middle|\lambda_P J_I^{s,t}\right)$ is the probability mass function of a Poisson random variable with intensity $\lambda_P J_I^{s,t}$ and $f_N\left(.\,\middle|m_{R,n}, v_{R,n}^2\right)$ is the density function of a normal random variable with mean $m_{R,n}$ and variance $v_{R,n}^2$.

Proof. From equation (5.2.3) and power expansion of the exponential function, we have that:

$$\begin{aligned} E\left[e^{iuR(s,t)}\middle|\mathcal{F}_H^{X_I}\right] &= e^{iu\left(\mu-\frac{\sigma^2}{2}\right)(t-s)} e^{-\frac{u^2\sigma^2}{2}(t-s)} e^{-J_I^{s,t}\lambda_P} e^{J_I^{s,t}\lambda_P\left(e^{iu\eta-\frac{1}{2}u^2\delta^2}\right)} \\ &= e^{iu\left(\mu-\frac{\sigma^2}{2}\right)(t-s)} e^{-\frac{u^2\sigma^2}{2}(t-s)} e^{-J_I^{s,t}\lambda_P} \sum_{n=0}^{\infty} \frac{\left(\lambda_P J_I^{s,t}\right)^n}{n!} e^{n\left(iu\eta-\frac{1}{2}u^2\delta^2\right)} \\ &= e^{iu\left(\mu-\frac{\sigma^2}{2}\right)(t-s)} e^{-\frac{u^2\sigma^2}{2}(t-s)} \sum_{n=0}^{\infty} e^{-J_I^{s,t}\lambda_P} \frac{\left(\lambda_P J_I^{s,t}\right)^n}{n!} e^{n\left(iu\eta-\frac{1}{2}u^2\delta^2\right)} \\ &= \sum_{n=0}^{\infty} \left(\frac{(\lambda_P J_I^{s,t})^n e^{-\lambda_P J_I^{s,t}}}{n!}\right. \\ &\qquad \left. e^{iu\left(\left(\mu-\frac{\sigma^2}{2}\right)(t-s)+n\eta\right)-\frac{u^2}{2}\left(\sigma^2(t-s)+\delta^2 n\right)}\right) \end{aligned} \tag{5.2.5}$$

Notice that the fraction that appears in equation (5.2.5) satisfies the following:

$$f_P\left(n\,\middle|\lambda_P J_I^{s,t}\right) = \frac{(\lambda_P J_I^{s,t})^n e^{-\lambda_P J_I^{s,t}}}{n!},$$

where $f_P\left(.\,\middle|\lambda J_I^{s,t}\right)$ is the probability mass function of a Poisson random variable with intensity parameter $\lambda_P J_I^{s,t}$. In addition we also have that

$$\Phi_N\left(u\,\middle|m_{R,n}, v_{R,n}^2\right) = e^{iu\left(\left(\mu-\frac{\sigma^2}{2}\right)(t-s)+n\eta\right)-\frac{u^2}{2}\left(\sigma^2(t-s)+\delta^2 n\right)},$$

where $\Phi_N\left(.\left|m_{R,n}, v_{R,n}^2\right.\right)$ is the characteristic function of a normal random variable with mean $m_{R,n}$ and variance $v_{R,n}^2$ where

$$\begin{aligned}
m_{R,n} &= \left(\mu - \frac{\sigma^2}{2}\right)(t-s) + n\eta, \\
v_{R,n}^2 &= \sigma^2(t-s) + n\delta^2.
\end{aligned}$$

So we can write $\Phi_{R(s,t)}\left(u\left|\mathcal{F}_H^{X_I}\right.\right)$ as:

$$\Phi_{R(s,t)}\left(u\left|\mathcal{F}_H^{X_I}\right.\right) = \sum_{n=0}^{\infty} f_P\left(n\left|\lambda_P J_I^{s,t}\right.\right)\Phi_N\left(u\left|m_{R,n}, v_{R,n}^2\right.\right). \tag{5.2.6}$$

Let $a, b \in \mathbb{R}$ such that $b > a$, then from the result of Yuan and Lei (2016, Theorem 3.2) we have that

$$F_{s,t}^R\left(b\left|\mathcal{F}_H^{X_I}\right.\right) - F_{s,t}^R\left(a\left|\mathcal{F}_H^{X_I}\right.\right) = \lim_{M\to\infty}\frac{1}{2\pi}\int_{-M}^{M}\frac{e^{-iua}-e^{-iub}}{iu}\Phi_{R(s,t)}\left(u\left|\mathcal{F}_H^{X_I}\right.\right)du, \tag{5.2.7}$$

where $F_{s,t}^R\left(.\left|\mathcal{F}_H^{X_I}\right.\right)$ is the conditional distribution function of $R(s,t)$ given $\mathcal{F}_H^{X_I}$. We can take $b = x+h$ and $a = x-h$, where $x \in \mathbb{R}$ and $h > 0$, then from equation (5.2.7) we can write

$$\begin{aligned}
&\frac{F_{s,t}^R\left(x+h\left|\mathcal{F}_H^{X_I}\right.\right) - F_{s,t}^R\left(x-h\left|\mathcal{F}_H^{X_I}\right.\right)}{2h} \\
&\qquad = \lim_{M\to\infty}\frac{1}{2\pi}\int_{-M}^{M}\frac{\sin(hu)}{hu}e^{-iux}\Phi_{R(s,t)}\left(u\left|\mathcal{F}_H^{X_I}\right.\right)du.
\end{aligned} \tag{5.2.8}$$

Using the expression in (5.2.6), the integral that appears in equation (5.2.8) can be written as:

$$\begin{aligned}
&\int_{-M}^{M}\frac{\sin(hu)}{hu}e^{-iux}\Phi_{R(s,t)}\left(u\left|\mathcal{F}_H^{X_I}\right.\right)du \\
&\quad = \int_{-M}^{M}\left(\sum_{n=0}^{\infty}\frac{\sin(hu)}{hu}e^{-iux}f_P\left(n\left|\lambda_P J_I^{s,t}\right.\right)\Phi_N\left(u\left|m_{R,n}, v_{R,n}^2\right.\right)\right)du.
\end{aligned} \tag{5.2.9}$$

We would like to change the order of integration of the integral and the summation that appears in equation (5.2.9). To do so, we would like to apply Fubini's theorem. To that end, we need to show that:

$$\int_{-M}^{M}\sum_{n=0}^{\infty}\left|\frac{\sin(hu)}{hu}e^{-iux}f_P\left(n\left|\lambda_P J_I^{s,t}\right.\right)\Phi_N\left(u\left|m_{R,n}, v_{R,n}^2\right.\right)\right|du < \infty.$$

Notice that because $\Phi_N\left(.\left|m_{R,n}, v_{R,n}^2\right.\right)$ is a characteristic function we have that:

$$\left|\Phi_N\left(u\left|m_{R,n}, v_{R,n}^2\right.\right)\right| \leq 1 \text{ for all } n = 1, 2, \ldots, \text{ and all } u \in \mathbb{R}.$$

In addition because

$$\left|e^{-iux}\right| = 1 \quad , \quad \left|\frac{\sin(hu)}{hu}\right| \leq 1$$

for all $u \in \mathbb{R}$, we can write

$$\begin{aligned}\int_{-M}^{M} \sum_{n=0}^{\infty} \left| \frac{\sin(hu)}{hu} e^{-iux} f_P\left(n \left| \lambda_P J_I^{s,t}\right.\right) \Phi_N\left(u \left| m_{R,n}, v_{R,n}^2\right.\right) \right| du \\ &\leq \int_{-M}^{M} \sum_{n=0}^{\infty} \left| f_P\left(n \left| \lambda_P J_I^{s,t}\right.\right) \right| du \\ &= \int_{-M}^{M} \sum_{n=0}^{\infty} f_P\left(n \left| \lambda_P J_I^{s,t}\right.\right) du \\ &= \int_{-M}^{M} du = 2M < \infty.\end{aligned}$$

So by Fubini's theorem, we can write equation (5.2.8) as

$$\begin{aligned}&\frac{F_{s,t}^R\left(x+h \left| \mathcal{F}_H^{X_I}\right.\right) - F_{s,t}^R\left(x-h \left| \mathcal{F}_H^{X_I}\right.\right)}{2h} \\ &= \lim_{M\to\infty} \sum_{n=0}^{\infty} f_P\left(n \left| \lambda_P J_I^{s,t}\right.\right) \frac{1}{2\pi} \int_{-M}^{M} \frac{\sin(hu)}{hu} e^{-iux} \Phi_N\left(u \left| m_{R,n}, v_{R,n}^2\right.\right) du. \qquad (5.2.10)\end{aligned}$$

Let us consider Θ to be a Poisson random variable with parameter $\lambda J_I^{s,t}$ given the σ-algebra $\mathcal{F}_H^{X_I}$, then equation (5.2.10) can be written as

$$\begin{aligned}&\frac{F_{s,t}^R\left(x+h \left| \mathcal{F}_H^{X_I}\right.\right) - F_{s,t}^R\left(x-h \left| \mathcal{F}_H^{X_I}\right.\right)}{2h} \\ &= \lim_{M\to\infty} E\left[\frac{1}{2\pi} \int_{-M}^{M} \frac{\sin(hu)}{hu} e^{-iux} \Phi_N\left(u \left| m_{R,\Theta}, v_{R,\Theta}^2\right.\right) du \right| \left. \mathcal{F}_H^{X_I} \right].\end{aligned}$$

Let us define ξ as

$$\xi(u,\Theta) = \frac{\sin(hu)}{hu} e^{-iux} \Phi_N\left(u \left| m_{R,\Theta}, v_{R,\Theta}^2\right.\right) \text{ for } u \in \mathbb{R}.$$

Since ξ takes complex values, let us differentiate between the imaginary and real part. To that end, define

$$\xi_{Re}(u,\Theta) = \Re(\xi(u,\Theta)), \quad \xi_{Im}(u,\Theta) = \Im(\xi(u,\Theta)) \text{ for } u \in \mathbb{R}.$$

So we can write

$$\begin{aligned}&\frac{F_{s,t}^R\left(x+h \left| \mathcal{F}_H^{X_I}\right.\right) - F_{s,t}^R\left(x-h \left| \mathcal{F}_H^{X_I}\right.\right)}{2h} \\ &= \lim_{M\to\infty} E\left[\frac{1}{2\pi} \int_{-M}^{M} \xi_{Re}(u,\Theta) du \right| \left. \mathcal{F}_H^{X_I} \right] + i \lim_{M\to\infty} E\left[\frac{1}{2\pi} \int_{-M}^{M} \xi_{Im}(u,\Theta) du \right| \left. \mathcal{F}_H^{X_I} \right].\end{aligned}$$

Now, let us define ξ_{Re}^{+} and ξ_{Re}^{-} to be the positive and negative parts of ξ_{Re} respectively. Similarly, we define ξ_{Im}^{+} and ξ_{Im}^{-} to be the positive and negative parts of ξ_{Im}

respectively. Then we can write

$$\begin{aligned}
&\frac{F^R_{s,t}\left(x+h\,\middle|\mathcal{F}^{X_I}_H\right)-F^R_{s,t}\left(x-h\,\middle|\mathcal{F}^{X_I}_H\right)}{2h}\\
&=\lim_{M\to\infty}E\left[\frac{1}{2\pi}\int_{-M}^{M}\xi^+_{Re}(u,\Theta)du\middle|\,\mathcal{F}^{X_I}_H\right]-\lim_{M\to\infty}E\left[\frac{1}{2\pi}\int_{-M}^{M}\xi^-_{Re}(u,\Theta)du\middle|\,\mathcal{F}^{X_I}_H\right]\\
&\quad+i\left(\lim_{M\to\infty}E\left[\frac{1}{2\pi}\int_{-M}^{M}\xi^+_{Im}(u,\Theta)du\middle|\,\mathcal{F}^{X_I}_H\right]-\lim_{M\to\infty}E\left[\frac{1}{2\pi}\int_{-M}^{M}\xi^-_{Im}(u,\Theta)du\middle|\,\mathcal{F}^{X_I}_H\right]\right).
\end{aligned} \tag{5.2.11}$$

Now let us define the sequence $(g_M)_{M=1}^{\infty}$ as

$$g_M=\frac{1}{2\pi}\int_{-M}^{M}\xi^+_{Re}(u,\Theta)du \text{ for} M=1,2,\ldots.$$

Notice that $g_M\leq g_{M+1}$ for all $M\in\mathbb{N}$. By the conditional monotone convergence theorem (Williams, 1991, Theorem 9.7), we have that:

$$\lim_{M\to\infty}E\left[\frac{1}{2\pi}\int_{-M}^{M}\xi^+_{Re}(u,\Theta)du\middle|\,\mathcal{F}^{X_I}_H\right]=E\left[\lim_{M\to\infty}\frac{1}{2\pi}\int_{-M}^{M}\xi^+_{Re}(u,\Theta)du\middle|\,\mathcal{F}^{X_I}_H\right].$$

Proceeding in a similar manner for the other limits that appear in equation (5.2.11), we have that:

$$\begin{aligned}
&\frac{F^R_{s,t}\left(x+h\,\middle|\mathcal{F}^{X_I}_H\right)-F^R_{s,t}\left(x-h\,\middle|\mathcal{F}^{X_I}_H\right)}{2h}\\
=&E\left[\lim_{M\to\infty}\frac{1}{2\pi}\int_{-M}^{M}\xi^+_{Re}(u,\Theta)du\middle|\,\mathcal{F}^{X_I}_H\right]-E\left[\lim_{M\to\infty}\frac{1}{2\pi}\int_{-M}^{M}\xi^-_{Re}(u,\Theta)du\middle|\,\mathcal{F}^{X_I}_H\right]\\
&+i\left(E\left[\lim_{M\to\infty}\frac{1}{2\pi}\int_{-M}^{M}\xi^+_{Im}(u,\Theta)du\middle|\,\mathcal{F}^{X_I}_H\right]-E\left[\lim_{M\to\infty}\frac{1}{2\pi}\int_{-M}^{M}\xi^-_{Im}(u,\Theta)du\middle|\,\mathcal{F}^{X_I}_H\right]\right)\\
=&E\left[\lim_{M\to\infty}\frac{1}{2\pi}\int_{-M}^{M}\frac{\sin(hu)}{hu}e^{-iux}\Phi_N\left(u\,\middle|m_{R,\Theta},v^2_{R,\Theta}\right)du\middle|\,\mathcal{F}^{X_I}_H\right]\\
=&\sum_{n=0}^{\infty}\frac{(\lambda_P J^{s,t}_I)^n e^{-\lambda_P J^{s,t}_I}}{n!}\left(\lim_{M\to\infty}\frac{1}{2\pi}\int_{-M}^{M}\frac{\sin(hu)}{hu}e^{-iux}\Phi_N\left(u\,\middle|m_{R,n},v^2_{R,n}\right)du\right).
\end{aligned}$$

Lastly, applying the conditional inversion theorem (Yuan & Lei, 2016, Theorem 3.2), we have that:

$$\begin{aligned}
&\lim_{M\to\infty}\frac{1}{2\pi}\int_{-M}^{M}\frac{\sin(hu)}{hu}e^{-iux}\Phi_N\left(u\,\middle|m_{R,n},v^2_{R,n}\right)du\\
=&\frac{F_N\left(x+h\,\middle|m_{R,n},v^2_{R,n}\right)-F_N\left(x-h\,\middle|m_{R,n},v^2_{R,n}\right)}{2h}\text{ given the }\sigma\text{-algebra }\mathcal{F}^{X_I}_H,
\end{aligned}$$

where $F_N\left(.\,\middle|m_{R,n},v^2_{R,n}\right)$ is the cumulative distribution function of a normal random variable with mean $m_{R,n}$ and variance $v^2_{R,n}$.

We have just shown that

$$\frac{F_{s,t}^{R}\left(x+h\left|\mathcal{F}_{H}^{X_I}\right.\right)-F_{s,t}^{R}\left(x-h\left|\mathcal{F}_{H}^{X_I}\right.\right)}{2h}$$
$$=\sum_{n=0}^{\infty} f_P\left(n\left|\lambda_P J_I^{s,t}\right.\right)\frac{F_N\left(x+h\left|m_{R,n},v_{R,n}^2\right.\right)-F_N\left(x-h\left|m_{R,n},v_{R,n}^2\right.\right)}{2h}.$$

We are interested in computing the conditional density function of $R(s,t)$ given $\mathcal{F}_{H}^{X_I}$ which can be obtained from

$$f_{s,t}^{R}\left(x\left|\mathcal{F}_{H}^{X_I}\right.\right)=\frac{\partial}{\partial x}F_{s,t}^{R}\left(x\left|\mathcal{F}_{H}^{X_I}\right.\right).$$

So we can write

$$f_{s,t}^{R}\left(x\left|\mathcal{F}_{H}^{X_I}\right.\right)=\lim_{h\to 0}\frac{F_{s,t}^{R}\left(x+h\left|\mathcal{F}_{H}^{X_I}\right.\right)-F_{s,t}^{R}\left(x-h\left|\mathcal{F}_{H}^{X_I}\right.\right)}{2h}.$$

Let us consider again the Poisson random variable Θ with parameter $\lambda_P J_I^{s,t}$ given the σ-algebra $\mathcal{F}_{H}^{X_I}$, so we can write

$$\lim_{h\to 0}\frac{F_{s,t}^{R}\left(x+h\left|\mathcal{F}_{H}^{X_I}\right.\right)-F_{s,t}^{R}\left(x-h\left|\mathcal{F}_{H}^{X_I}\right.\right)}{2h}$$
$$=\lim_{h\to 0}E\left[\left.\frac{F_N\left(x+h\left|m_{R,\Theta},v_{R,\Theta}^2\right.\right)-F_N\left(x-h\left|m_{R,\Theta},v_{R,\Theta}^2\right.\right)}{2h}\right|\mathcal{F}_{H}^{X_I}\right].$$

By the mean value theorem, we have that:

$$\frac{F_N\left(x+h\left|m_{R,\Theta},v_{R,\Theta}^2\right.\right)-F_N\left(x-h\left|m_{R,\Theta},v_{R,\Theta}^2\right.\right)}{2h}=f_N\left(\chi(h)\left|m_{R,\theta},v_{R,\Theta}^2\right.\right),$$

where $\chi(h)\in(x-h,x+h)$ and $f_N\left(.\left|m_{R,\Theta},v_{R,\Theta}^2\right.\right)$ is the density function of a normal random variable with mean $m_{R,\Theta}$ and variance $v_{R,\Theta}^2$. The density function of a normal random variable is bounded by the value obtained when it is evaluated at the mean, that is

$$\begin{aligned} f_N\left(\chi(h)\left|m_{R,\theta},v_{R,\Theta}^2\right.\right) &\leq f_N\left(m_{R,\theta}\left|m_{R,\theta},v_{R,\Theta}^2\right.\right)\\ &= \frac{1}{v_{R,\Theta}\sqrt{2\pi}}\\ &= \frac{1}{\sqrt{2\pi\left(\sigma^2(t-s)+\Theta\delta^2\right)}}. \end{aligned}$$

So we have that for all $h>0$

$$\left|\frac{F_N\left(x+h\left|m_{R,\Theta},v_{R,\Theta}^2\right.\right)-F_N\left(x-h\left|m_{R,\Theta},v_{R,\Theta}^2\right.\right)}{2h}\right|\leq\frac{1}{\sqrt{2\pi\left(\sigma^2(t-s)+\Theta\delta^2\right)}}.$$

Hence, by the conditional dominated convergence theorem (Williams, 1991, Theorem 9.7) we have that

$$\begin{aligned} f^R_{s,t}\left(x\left|\mathcal{F}^{X_I}_H\right.\right) &= E\left[\lim_{h\to 0}\frac{F_N\left(x+h\left|m_{R,\Theta},v^2_{R,\Theta}\right.\right)-F_N\left(x-h\left|m_{R,\Theta},v^2_{R,\Theta}\right.\right)}{2h}\right|\mathcal{F}^{X_I}_H\Bigg] \\ &= E\left[f_N\left(x\left|m_{R,\theta},v^2_{R,\Theta}\right.\right)\middle|\,\mathcal{F}^{X_I}_H\right] \\ &= \sum_{n=0}^{\infty}\frac{(\lambda J_I^{s,t})^n e^{-\lambda J_I^{s,t}}}{n!} f_N\left(x\left|m_{R,n},v^2_{R,n}\right.\right) \end{aligned}$$

as required. □

5.3 Estimation

For the estimation procedure, we assume that we have discrete observations of the processes X and I. We would like to estimate the parameters that appear in the price equation (5.1.2) and in the interest equation (5.1.3). We assume that the data have the same structure as in Chapter 3 and Chapter 4 (see Section 3.6 and Section 4.2).

As in previous chapters, we will estimate the parameters of the interest using the techniques explained in Chapter 2. For the parameters that appear in the price process (5.1.2), we will use the conditional likelihood estimation method.

5.3.1 Conditional likelihood estimator

We are interested in the estimation of the parameters related to the price; that is, we are interested in the parameters $\mu, \sigma, \lambda_P, \eta, \delta$ and τ. For now, let us assume that the parameter τ is given.

Let us define the vector R of logarithm returns as

$$R = \left(R(t_0,t_1), R(t_1,t_2), \ldots R(t_{N-1},t_N)\right),$$

where

$$R(t_j,t_{j+1}) = \log\left(\frac{P(t_{j+1})}{P(t_j)}\right) \text{ for } j = 0,\ldots,N-1.$$

In addition, we define the vector of random variables J_I as

$$J_I = \left(J_I^{t_0,t_1}, J_I^{t_1,t_2}, \ldots, J_I^{t_{N-1},t_N}\right),$$

where

$$J_I^{t_j,t_{j+1}} = \int_{t_j}^{t_{j+1}} I(u-\tau)du \text{ for } j = 0,1\ldots N-1.$$

From Proposition C.1.1 we have that the increments $\left(X_P\left(T_I(t_{j+1})\right) - X_P\left(T_I(t_j)\right)\right)_{j=0}^{N-1}$ are independent when the σ-algebra $\mathcal{F}^{X_I}_H$ is given. So clearly the sequence of logarithmic returns $\left(R(t_j,t_{j+1})\right)_{j=0}^{N-1}$ are independent when the σ-algebra $\mathcal{F}^{X_I}_H$ is given. So we have that the conditional density function of R

given $\mathcal{F}_H^{X_I}$ satisfies:

$$f_{1:N}^R\left(r_{1:N}\middle|\mathcal{F}_H^{X_I}\right)=\prod_{j=0}^{N-1}f_{j+1}^R\left(r_{j+1}\middle|\mathcal{F}_H^{X_I}\right),\tag{5.3.1}$$

where

$$r_{j+1}=x_{j+1}-x_j \text{ for } j=0,1,\ldots,N-1$$

and x_j is the observation of the random variable $X(t_j)$.

Let us define the set $\sigma(J_I)$ to be the σ-algebra generated by J_I. We are now interested in the conditional distribution given $\sigma(J_I)$.

Proposition 5.3.1. *Let $u_j\in\mathbb{R}$ for $j=0,\ldots,N-1$ then:*

$$\begin{aligned}
E\left[e^{\sum_{j=0}^{N-1}iu_jR(t_j,t_{j+1})}\middle|\mathcal{F}_H^{X_I}\right] &= E\left[e^{\sum_{j=0}^{N-1}iu_jR(t_j,t_{j+1})}\middle|\sigma(J_I)\right]\\
&= \prod_{j=0}^{N-1}E\left[e^{iu_jR(t_j,t_{j+1})}\middle|\sigma(J_I)\right]\\
&= \prod_{j=0}^{N-1}\left[e^{iu_j\left(\mu-\frac{\sigma^2}{2}\right)(t-s)}e^{-\frac{u_j^2\sigma^2}{2}(t-s)}\right.\\
&\qquad \left. e^{-J_I^{t_j,t_{j+1}}\lambda_P}e^{J_I^{t_j,t_{j+1}}\lambda_P\left(e^{iu_j\eta-\frac{1}{2}u^2\delta^2}\right)}\right].
\end{aligned}\tag{5.3.2}$$

A particular case of equation (5.3.2) is

$$\begin{aligned}
E\left[e^{iu_jR(t_j,t_{j+1})}\middle|\mathcal{F}_H^{X_I}\right] &= E\left[e^{iu_jR(t_j,t_{j+1})}\middle|\sigma(J_I)\right]\\
&= e^{iu_j\left(\mu-\frac{\sigma^2}{2}\right)(t-s)}e^{-\frac{u_j^2\sigma^2}{2}(t-s)}\\
&\qquad e^{-J_I^{t_j,t_{j+1}}\lambda_P}e^{J_I^{t_j,t_{j+1}}\lambda_P\left(e^{iu_j\eta-\frac{1}{2}u^2\delta^2}\right)}
\end{aligned}$$

for $j=0,1,\ldots,N-1$.

Proof. Notice that from Proposition 5.2.1 we have that

$$E\left[e^{iu_jR(t_j,t_{j+1})}\middle|\mathcal{F}_H^{X_I}\right]=e^{iu_j\left(\mu-\frac{\sigma^2}{2}\right)(t-s)}e^{-\frac{u_j^2\sigma^2}{2}(t-s)}e^{-J_I^{t_j,t_{j+1}}\lambda_P}e^{J_I^{t_j,t_{j+1}}\lambda_P\left(e^{iu_j\eta-\frac{1}{2}u^2\delta^2}\right)}\tag{5.3.3}$$

for $j=0,1,\ldots,N-1$..

Notice that the random variable J_I is measurable with respect to the σ-algebra $\mathcal{F}_H^{X_I}$, hence $\sigma(J_I)\subseteq\mathcal{F}_H^{X_I}$. By the tower property, we have that:

$$E\left[e^{\sum_{j=0}^{N-1}iu_jR(t_j,t_{j+1})}\middle|\sigma(J_I)\right]=E\left[E\left[e^{\sum_{j=0}^{N-1}iu_jR(t_j,t_{j+1})}\middle|\mathcal{F}_H^{X_I}\right]\middle|\sigma(J_I)\right].\tag{5.3.4}$$

By the conditional independence of the logarithmic returns, we obtain that

$$E\left[e^{\sum_{j=0}^{N-1}iu_jR(t_j,t_{j+1})}\middle|\mathcal{F}_H^{X_I}\right]=\prod_{j=0}^{N-1}E\left[e^{iu_jR(t_j,t_{j+1})}\middle|\mathcal{F}_H^{X_I}\right].\tag{5.3.5}$$

From the result in (5.3.3) we obtain:

$$E\left[e^{\sum_{j=0}^{N-1} iu_j R(t_j,t_{j+1})}\middle|\mathcal{F}_H^{X_I}\right] = \prod_{j=0}^{N-1}\left[e^{iu_j\left(\mu-\frac{\sigma^2}{2}\right)(t-s)}e^{-\frac{u_j^2\sigma^2}{2}(t-s)} e^{-J_I^{t_j,t_{j+1}}\lambda_P}e^{J_I^{t_j,t_{j+1}}\lambda_P\left(e^{iu_j\eta-\frac{1}{2}u^2\delta^2}\right)}\right]. \quad (5.3.6)$$

Substituting equation (5.3.6) into equation (5.3.4), we have that

$$\begin{aligned} &E\left[e^{\sum_{j=0}^{N-1} iu_j R(t_j,t_{j+1})}\middle|\sigma(J_I)\right]\\ &=E\left[\prod_{j=0}^{N-1} e^{iu_j\left(\mu-\frac{\sigma^2}{2}\right)(t-s)}e^{-\frac{u_j^2\sigma^2}{2}(t-s)}e^{-J_I^{t_j,t_{j+1}}\lambda_P}e^{J_I^{t_j,t_{j+1}}\lambda_P\left(e^{iu_j\eta-\frac{1}{2}u^2\delta^2}\right)}\middle|\sigma(J_I)\right]\\ &=\prod_{j=0}^{N-1} e^{iu_j\left(\mu-\frac{\sigma^2}{2}\right)(t-s)}e^{-\frac{u_j^2\sigma^2}{2}(t-s)}e^{-J_I^{t_j,t_{j+1}}\lambda_P}e^{J_I^{t_j,t_{j+1}}\lambda_P\left(e^{iu_j\eta-\frac{1}{2}u^2\delta^2}\right)}, \end{aligned}$$

where the last equality comes from the fact that $J_I^{t_j,t_{j+1}}$ is measurable with respect to the σ-algebra $\sigma(J_I)$. □

From Proposition 5.2.2 we have that

$$f_{j+1}^R\left(r_{j+1}\middle|\mathcal{F}_H^{X_I}\right) = \sum_{n=0}^{\infty} f_P\left(n\middle|\lambda_P J_I^{t_j,t_{j+1}}\right) f_N\left(x\middle|m_{R,n}, v_{R,n}^2\right), \text{ for } j=0,1,\dots N-1, \quad (5.3.7)$$

where

$$\begin{aligned} m_{R,n} &= \left(\mu-\frac{\sigma^2}{2}\right)\Delta + n\eta,\\ v_{R,n}^2 &= \sigma^2\Delta + n\delta^2. \end{aligned}$$

From Proposition 5.3.1 we have that

$$f_{j+1}^R\left(r_{j+1}\middle|\mathcal{F}_H^{X_I}\right) = f_{j+1}^R\left(r_{j+1}\middle|\sigma(J_I)\right) \text{ for } j=0,1,\dots,N-1,$$

and we also have that

$$f_{1:N}^R\left(r_{1:N}\middle|\mathcal{F}_H^{X_I}\right) = f_{1:N}^R\left(r_{1:N}\middle|\sigma(J)\right).$$

The random vector R depends on the vector of parameters $\theta_R = (\lambda_R, \xi_R)$, where

$$\lambda_R = (\mu, \sigma, \lambda_P, \eta, \delta), \qquad \xi_R = (a_I, b_I, \sigma_I).$$

The parameter of interest is the vector λ_R and ξ_R is the vector of nuisance parameters. So, the conditional log-likelihood can be written as

$$\begin{aligned} l^R\left(\lambda_R \left|\sigma\left(J_I\right), \tau\right.\right) &= \log f^R_{1:N}\left(r_{1:N}\left|\sigma\left(J_I\right)\right.\right) \\ &= \sum_{j=0}^{N-1} \log f^R_{j+1}\left(r_{j+1}|\sigma\left(J_I\right)\right) \\ &= \sum_{j=0}^{N-1} \log\left(\sum_{n=0}^{\infty} f_P\left(n\left|\lambda_P J_I^{t_j,t_{j+1}}\right.\right) f_N\left(r_{j+1}\left|m_{R,n}, v^2_{R,n}\right.\right)\right) \end{aligned} \tag{5.3.8}$$

We will maximize the function $l^R\left(.\left|\sigma\left(J_I\right), \tau\right.\right)$ to estimate the parameter vector λ_R. In Chapters 3 and 4 we were able to find an analytical expression for the parameter estimates that maximizes the conditional log-likelihood. In this case, we are not able to find an analytical formula for the parameter estimates that maximizes the expression (5.3.8). Because of that we have to rely on numerical methods to maximize the conditional log-likelihood defined in equation (5.3.8).

Remark 5.3.1. *As we did in Section 3.6.2 we reduce the problem of estimating the parameter τ to a problem of model selection. We select the value of τ that maximizes the following expression*

$$l^R\left(\hat{\lambda}_R\left|\sigma\left(J_I\right), \tau = r\Delta\right.\right) \text{ for } r = 0, 1, \ldots M,$$

where $\hat{\lambda}_R = (\hat{\mu}, \hat{\sigma}, \hat{\lambda}_P, \hat{\eta}, \hat{\delta})$ is the vector of parameters that maximizes the conditional log-likelihood defined in equation (5.3.8).

5.3.2 Numerical experiments

We would like to check the reliability of our estimation method proposed in Section 5.3. To that end, we perform numerical simulations of the two proposed models. We first focus on the case when the interest is a Cox-Ingersoll-Ross process and then on the case when the interest is an inverse Gaussian Ornstein-Uhlenbeck process.

When the interest is a Cox-Ingersoll-Ross process, we generate 100 realizations of $N = 1000$ steps with the following values for the parameters

$$a_I = 0.5, \quad b_I = 0.06, \quad \sigma_I = 0.15, \quad \mu = 0.05, \quad \sigma_P = 0.2,$$

$$\lambda_P = 20, \quad \eta = 0, \quad \delta = 0.2 \quad \tau = 5\Delta,$$

with time step $\Delta = 0.1$, $L = 10\Delta$ and initial values $P(0) = 100$ and

$$\phi^I(t) = 0.05 + 0.1\cos^2\left(4t + \frac{\pi}{2}\right) \text{ for } t \in [-L, 0].$$

For each of the realizations we estimate the parameters of the model using the techniques introduced in Section 5.3. Once we estimate the parameters of all of the simulations, we compute the mean of the estimated parameters and its standard deviation. The results are shown in Table 5.1.

In the case when the interest is an inverse Gaussian Ornstein-Uhlenbeck process, we generate 20 realizations of $N = 1000$ steps with the following values for the parameters

$$a_I = 4, \quad b_I = 10, \quad \lambda_I = 5, \quad \mu = 0.05, \quad \sigma_P = 0.2,$$

$$\lambda_P = 20, \quad \eta = 0, \quad \delta = 0.2 \quad \tau = 5\Delta,$$

Variable	True value	Mean	Standard deviation
a_I	0.5	0.546785	0.113987
b_I	0.06	0.059973	0.007607
σ_I	0.15	0.150301	0.003112
μ	0.05	0.0490428	0.020118
σ	0.2	0.199536	0.007402
λ_P	20	20.238023	4.507455
η	0	-0.003124	0.023025
δ	0.2	0.197553	0.024317
τ	5 Δ	4.8 Δ	2.198727 Δ

TABLE 5.1: Estimated values for the parameters for the simulated model defined in Section 5.1 when the interest is a Cox-Ingersoll-Ross process.

with time step $\Delta = 0.01$, $L = 20\Delta$ and initial values $P(0) = 100$ and

$$\phi^I(t) = 0.4 + 0.2\cos^2\left(30t + \frac{\pi}{2}\right) \quad \text{for } t \in [-L, 0].$$

As we did above, we estimate the parameters of the model using the techniques explained in Section 5.3 and compute the mean of the estimated parameters and its standard deviation. The results are shown in Table 5.2.

Variable	True value	Mean	Standard deviation
a_I	4	4.134514	0.547405
b_I	10	10.226795	1.231333
λ_I	5	5.467092	1.263248
μ	0.05	0.028039	0.065766
σ	0.2	0.200946	0.005251
λ_P	20	19.655269	2.484821
η	0	0.000198	0.021476
δ	0.2	0.191985	0.018088
τ	5 Δ	4.5 Δ	3.278719 Δ

TABLE 5.2: Estimated values for the parameters for the simulated model defined in Section 5.1 when the market attention is an inverse Gaussian Ornstein-Uhlenbeck process.

As we can observe in Tables 5.1 and 5.2 these methods give good estimates for the parameters.

5.4 Option Pricing

Now we would like to price plain vanilla options with strike price K and expiration date T. For this case, we will consider $t \in [0, T]$ and $\mathcal{F} = \mathcal{F}_T$. Options are priced under a risk–neutral measure $\mathbb{Q}$. Under this measure, the discounted stock prices $\tilde{P}$, defined as $\tilde{P}(t) = e^{-rt}P(t)$ for $t \in [0, T]$, is a martingale. We will show that there is a risk–neutral measure $\mathbb{Q}$ in which the price satisfies the equation:

$$P(t) = p\exp\left\{\left(r - \frac{\sigma^2}{2}\right)t + \sigma W_P^*(t) + X_P(T_I(t)) - T_I(t)\xi^{X_P}\right\} \text{ for } t \in [0,T], \tag{5.4.1}$$

where W_P^* is a Brownian motion under $\mathbb{Q}$ and

$$\xi^{X_P} = \lambda_P\left(e^{\eta+\frac{\delta^2}{2}} - 1\right). \tag{5.4.2}$$

When the price satisfies equation (5.4.1), the discounted Bitcoin price is a martingale (see Proposition 5.4.2 below).

To obtain equation (5.4.1), we need to perform a change of measure. Depending on the process we use for the market attention, the change of measure is done in a different manner. Because of that, we need to distinguish between the case when the interest process satisfies a Cox-Ingersoll-Ross process and when the interest process is an inverse Gaussian Ornstein-Uhlenbeck process.

5.4.1 Interest process is a Cox-Ingersoll-Ross process

When the market attention process is a Cox-Ingersoll-Ross process, we would like to obtain a risk–neutral measure $\mathbb{Q}$ under which the price equation satisfies equation (5.4.1) and the interest is still a Cox-Ingersoll-Ross process with possibly different parameters. Remember that in this case the Lévy process X_I of equation (5.1.3) is the Brownian motion W_I.

We are interested in obtaining an equivalent probability measure $\mathbb{Q}$ with respect to $\mathbb{P}$, under which the price process satisfies the equation (5.4.1). To obtain this equivalent measure $\mathbb{Q}$, let us define the following processes

$$\begin{aligned} W_P^*(t) &= W_P(t) + \int_0^t \theta_P(s)ds \text{ for } t \in [0,T], \\ W_I^*(t) &= W_I(t) + \int_0^t \theta_I(s)ds \text{ for } t \in [0,T], \end{aligned}$$

and

$$Z(t) = \exp\left\{-\int_0^t \theta_P(s)dW_P(s) - \int_0^t \theta_I(s)dW_I(s) - \frac{1}{2}\int_0^t \theta_P^2(s)ds - \frac{1}{2}\int_0^t \theta_I^2(s)ds\right\} \tag{5.4.3}$$

for $t \in [0,T]$, where $(\theta_P(t))_{t\in[0,T]}$ and $(\theta_I(t))_{t\in[0,T]}$ are two adapted processes with respect to $(\mathcal{F}_t)_{t\in[0,T]}$ (these processes will be determined below). If the process $(Z(t))_{t\in[0,T]}$ is a martingale with respect to $(\mathcal{F}_t)_{t\in[0,T]}$, then we can apply the Girsanov Theorem. By the Girsanov Theorem, the process $(W_P^*(t), W_I^*(t))_{t\in[0,T]}$ is a two–dimensional Brownian motion under the measure $\mathbb{Q}$, where:

$$\mathbb{Q}(A) = \int_A Z(T)d\mathbb{P},\ \forall A \in \mathcal{F}.$$

To obtain equation (5.4.1), different values for the adapted processes θ_P and θ_I can be taken. In this case, we choose the following values for the processes θ_P and θ_I:

$$\theta_P(t) = \frac{\mu - r + \xi^{X_P} I(t-\tau)}{\sigma} \text{ for } t \in [0,T], \tag{5.4.4}$$

$$\theta_I(t) = \frac{\lambda_1}{\sigma_I\sqrt{I(t)}} + \frac{\lambda_2}{\sigma_I}\sqrt{I(t)} \text{ for } t \in [0,T], \tag{5.4.5}$$

where ξ^{X_P} is defined as in (5.4.2) and $\lambda_1, \lambda_2 \in \mathbb{R}$.

If we assume that $(Z(t))_{t\geq 0}$ is a martingale (this is shown in Theorem 5.4.1 below), then we can apply the Girsanov Theorem. Under the measure $\mathbb{Q}$, we have that the price P satisfies equation (5.4.1) and the interest I satisfies the following stochastic differential equation:

$$dI(t) = (a_I b_I - \lambda_1 - (a_I + \lambda_2)I(t))dt + \sigma_I\sqrt{I(t)}dW_I^*(t) \text{ when } t \in (0,T], \tag{5.4.6}$$

and $I(t) = \phi^I(t)$ when $t \in [-L, 0]$.

Equation (5.4.6) can be rewritten as

$$dI(t) = \tilde{a}_I(\tilde{b}_I - I(t))dt + \sigma_I\sqrt{I(t)}dW_I^*(t) \text{ when } t \in (0,T] \tag{5.4.7}$$

and $I(t) = \phi^I(t)$ when $t \in [-L, 0]$, where $\tilde{a}_I = a_I + \lambda_2$ and $\tilde{b}_I = \frac{a_I b_I - \lambda_1}{a_I + \lambda_2}$. If we impose the conditions

$$\frac{2\tilde{a}_I\tilde{b}_I}{\sigma_I^2} \geq 1 \tag{5.4.8}$$

$$\tilde{a}_I > 0, \tag{5.4.9}$$

then equation (5.4.7) has a strong solution (Gulisashvili, 2012, p. 44) and it is greater than zero almost surely (Gulisashvili, 2012, Theorem 2.27). It is possible to show that the process Z defined as in (5.4.3) is a martingale.

Theorem 5.4.1. *If conditions (5.4.8) and (5.4.9) are satisfied, then $(Z(t))_{t\in[0,T]}$ is a martingale with respect to $(\mathcal{F}_t)_{t\in[0,T]}$.*

Proof. The proof of this theorem is very similar to the proof of Theorem 3.7.1. □

We have just shown that under $\mathbb{Q}$ the process (W_P^*, W_I^*) is a two–dimensional Brownian motion. It is also possible to show that under $\mathbb{Q}$ the process X_P is still a compound Poisson process with Lévy triplet $(\gamma_P, 0, \upsilon_P)$ and it is independent of the process (W_P^*, W_I^*). The proofs of these results are similar to ones shown in Appendix B.

In summary, we have just shown the existence of a probability measure $\mathbb{Q}$ with respect to the measure $\mathbb{P}$ such that:

1. the interest process satisfies equation (5.4.7),
2. the price process satisfies the process defined in (5.4.1),
3. and the processes W_P^*, W_I^* and X_P are independent under the measure $\mathbb{Q}$.

5.4.2 Interest is an inverse Gaussian Ornstein-Uhlenbeck process

When the market attention follows an inverse Gaussian Ornstein-Uhlenbeck process, we construct an equivalent measure $\mathbb{Q}$ under which the price process P satisfies equation (5.4.1) in two different steps:

1. First, we will construct a measure under which the background driven Lévy process Z_I that appears in equation (5.1.6) could have other parameters;
2. second, we do a change of measure such that the price process satisfies equation (5.4.1).

Let us find an equivalent measure $\mathbb{Q}^*$ with respect to $\mathbb{P}$, such that the Lévy process Z_I has the Lévy triplet $(\tilde{\gamma}_I, 0, \tilde{v}_I)$ where

$$\tilde{v}_I(dx) = \tilde{w}_I(x)dx = \frac{\tilde{a}_I}{2\sqrt{2\pi}} x^{-3/2}(1+\tilde{b}_I^2 x)e^{-\frac{1}{2}\tilde{b}_I^2 x}\mathbb{1}_{(0,\infty)}(x)dx, \tag{5.4.10}$$

$$\tilde{\gamma}_I = \int_{|x|<1} x\tilde{v}_I(dx), \tag{5.4.11}$$

with $\tilde{a}_I, \tilde{b}_I > 0$. That is, under this new proposed probability measure $\mathbb{Q}^*$ the market attention process could have different parameters.

Proposition 5.4.1. *If $\tilde{a}_I = a_I$ then there exists an equivalent probability measure $\mathbb{Q}^*$ with respect to the measure $\mathbb{P}$ such that the Lévy process Z_I has Lévy triplet $(\tilde{\gamma}_I, 0, \tilde{v}_I)$ where $\tilde{v}_I$ and $\tilde{\gamma}_I$ satisfy equations (5.4.10) and (5.4.11) respectively.*

And the probability measure $\mathbb{Q}^*$ can be expressed as

$$\mathbb{Q}^*(A) = \int_A Z^*(T)d\mathbb{P} \text{ for } A \in \mathcal{F},$$

where the process Z^* is defined as

$$Z^*(t) = \exp\left\{\lim_{\epsilon\downarrow 0}\left(\sum_{\substack{s\le t\\ |\Delta Z_I(s)|>\epsilon}} \log(H(\Delta Z_I(s))) - t\int_{|x|>\epsilon}(H(x)-1)v_I(dx)\right)\right\}$$

for $t \in [0,T]$, with

$$H(x) = \frac{1+\tilde{b}_I^2 x}{1+b_I^2 x} e^{-\frac{x}{2}(\tilde{b}_I^2 - b_I^2)}\mathbb{1}_{(0,\infty)}(x).$$

Proof. See proof of Proposition 4.3.1. □

So, if we assume that $\tilde{a}_I = a_I$ then under the equivalent measure $\mathbb{Q}^*$ the Lévy process Z_I has Lévy triplet $(\tilde{\gamma}_I, 0, \tilde{v}_I)$. Also, under the probability measure $\mathbb{Q}^*$ the processes W_P and X_P are still a Brownian motion and a compound Poisson process with Lévy triplet $(\gamma_P, 0, v_P)$ respectively. It is possible to show that under $\mathbb{Q}^*$ the property of independence between W_P, Z_I and X_P is still maintained. The proofs of these results are similar to ones shown in Appendix B.

Lastly, we would like to find a measure $\mathbb{Q}$ equivalent to $\mathbb{Q}^*$ such that the price process P satisfies equation (5.4.1). To obtain this equivalent measure $\mathbb{Q}$, let us define the following process

$$W_P^*(t) = W_P(t) + \int_0^t \theta_P(s)ds \text{ for } t \in [0,T], \tag{5.4.12}$$

and

$$Z(t) = \exp\left\{ -\int_0^t \theta_P(s)dW_P(s) - \frac{1}{2}\int_0^t \theta_P^2(s)ds \right\} \text{ for } t \in [0,T], \tag{5.4.13}$$

where $(\theta_P(t))_{t\in[0,T]}$ is an adapted process with respect to $(\mathcal{F}_t)_{t\in[0,T]}$ (to be defined below). If the process Z is a martingale with respect to $(\mathcal{F}_t)_{t\in[0,T]}$ (this is shown below) then we can apply the Girsanov Theorem. By the Girsanov Theorem, $(W_P^*(t))_{t\in[0,T]}$ is a Brownian motion under the measure $\mathbb{Q}$, where:

$$\mathbb{Q}(A) = \int_A Z(T)d\mathbb{Q}^*, \forall A \in \mathcal{F}.$$

To obtain equation (5.4.1), we chose the process θ_P as

$$\theta_P(t) = \frac{\mu - r + \xi^{X_P} I^-(t-\tau)}{\sigma} \text{ for } t \in [0,T], \tag{5.4.14}$$

where ξ^{X_P} is defined as in (5.4.2) and

$$I^-(t) = \lim_{s\to t^-} I(s) \text{ almost surely.} \tag{5.4.15}$$

We know that the limit (5.4.15) converges almost surely, because the process I has almost surely càdlàg paths (Valdivieso, Schoutens, & Tuerlinckx, 2009, p. 4).

Remark 5.4.1. *Notice that because the process I has almost surely càdlàg paths we have that*

$$T_I(t) = \int_0^t I(s-\tau)ds = \int_0^t I^-(s-\tau)ds$$

almost surely.

From Remark 5.4.1 we know that if we use the change of measure proposed in equations (5.4.12) and (5.4.14), we obtain the price equation (5.4.1). Finally, we will show that the process Z is a martingale.

Theorem 5.4.2. *The process $(Z(t))_{t\in[0,T]}$ defined in (5.4.13) is a martingale with respect to $(\mathcal{F}_t)_{t\in[0,T]}$.*

Proof. The proof of this Theorem is similar to the proof of Theorem 4.3.1. □

We have just shown that W_P^* is a Brownian measure under the probability measure $\mathbb{Q}$. Also, under the measure $\mathbb{Q}$ the processes Z_I and X_P are still a Lévy process with triplet $(\tilde{\gamma}_I, 0, \tilde{\upsilon}_I)$ and a compound Poisson process with triplet $(\gamma_P, 0, \upsilon_P)$, respectively. Further, the processes W_P^*, X_P and Z_I are independent. The proof of these results is similar to the ones shown in Appendix B.

Because $\mathbb{Q}$ is equivalent to $\mathbb{Q}^*$ and $\mathbb{Q}^*$ is equivalent to $\mathbb{P}$, we arrive at the fact that $\mathbb{Q}$ is equivalent to $\mathbb{P}$.

In summary, we have just shown the existence of a probability measure $\mathbb{Q}$ equivalent to $\mathbb{P}$ such that:

1. the market attention process satisfies an $IG(a_I, \tilde{b}_I)$-Ornstein-Uhlenbeck process with parameter $\lambda_I > 0$,
2. the price process satisfies the process defined in (5.4.1),
3. and the processes W_P^*, Z_I and X_P are independent under the measure $\mathbb{Q}$.

5.4.3 Martingale property

We have proved the existence of an equivalent probability $\mathbb{Q}$ under which the price process P satisfies the equation (5.4.1). Notice that the price process under $\mathbb{Q}$ satisfies equation (5.4.1) independently of whether the market attention process satisfies a Cox-Ingersoll-Ross process or an inverse Gaussian Ornstein-Uhlenbeck process.

We still have to show that the discounted price process $\tilde{P}$, is a martingale with respect to the filtration $(\mathcal{F}_t)_{t\in[0,T]}$ under the measure $\mathbb{Q}$. That is, we would like to show that

$$E_{\mathbb{Q}}\left[\tilde{P}(t)\middle|\mathcal{F}_s\right] = \tilde{P}(s) \text{ for } t,s \in [0,T] \text{ such that } t \geq s,$$

where $E_{\mathbb{Q}}$ symbolizes the expected value under the measure $\mathbb{Q}$. We will denote X_I^* as $X_I^* = W_I^*$ when the interest process satisfies a Cox-Ingersoll-Ross process and $X_I^* = Z_I$ when the market attention is an inverse Gaussian Ornstein-Uhlenbeck process. Before proving the martingale property, we need to prove the following Lemma.

Lemma 5.4.1. *If for some $\theta > 2$ we have that*

$$E_{\mathbb{Q}}\left[e^{T_I(T)\left(\xi_\theta^{X_P} - \theta\xi^{X_P}\right)}\right] < \infty, \tag{5.4.16}$$

where

$$\begin{aligned} \xi_\theta^{X_P} &= \Psi^{X_P}(-i\theta) = \lambda_P\left(e^{\theta\eta + \frac{1}{2}\theta^2\delta^2} - 1\right), \\ \xi^{X_P} &= \Psi^{X_P}(-i) = \lambda_P\left(e^{\eta + \frac{\delta^2}{2}} - 1\right), \end{aligned}$$

then

$$E\left[e^{X_P(T_I(t)) - T_I(t)\xi^{X_P}}\middle|\mathcal{F}_s\right] = e^{X_P(T_I(s)) - T_I(s)\xi^{X_P}} \text{ for } T \geq t \geq s \geq 0.$$

Proof. This proof is the same as the proof done by Shojaee (2018, Theorem 2.41), but with minor modifications.

Let us define the process M as

$$M(v) = e^{X_P(v) - v\xi^{X_P}} \text{ for } v \geq 0.$$

Since the process X_P is a compound Poisson process with normally distributed jumps we have that all the exponential moments are finite, that is

$$E_{\mathbb{Q}}\left[e^{\theta X_P(v)}\right] < \infty \text{ for } \theta \in \mathbb{R} \text{ and } v \geq 0 \tag{5.4.17}$$

(Cont & Tankov, 2004, p. 112). Hence from Proposition 1.1.2 we have that

$$E_{\mathbb{Q}}\left[e^{\theta X_P(v)}\right] = e^{v\Psi^{X_P}(-i\theta)}, \tag{5.4.18}$$

where

$$\Psi^{X_P}(u) = \lambda_P\left(e^{iu\eta - \frac{1}{2}u^2\delta^2} - 1\right).$$

From equation (5.4.18) we have that the process M can be written as:

$$\begin{aligned} M(v) &= e^{X_P(v)-v\xi^{X_P}} \\ &= e^{X_P(v)-v\Psi^{X_P}(-i)} \\ &= \frac{e^{X_P(v)}}{E_{\mathbb{Q}}\left[e^{X_P(v)}\right]}. \end{aligned}$$

So, from Theorem 1.1.6 we have that M is a martingale with respect to its natural filtration $\left(\mathcal{F}_v^M\right)_{v\geq 0}$.

From Proposition C.0.1 and Theorem C.1.1 we have that the process $(X_P(T_I(t)))_{t\in[0,T]}$ is a compound Poisson process with normally distributed jumps given the σ-algebra $\mathcal{F}_T^{X_I^*}$. Hence, from Proposition 1.1.2 we have that

$$\begin{aligned} E_{\mathbb{Q}}\left[e^{\theta X_P(T_I(T))}\middle|\mathcal{F}_T^{X_I^*}\right] &= e^{T_I(T)\Psi^{X_P}(-i\theta)} \\ &= e^{T_I(T)\xi_\theta^{X_P}}, \end{aligned}$$

where

$$\xi_\theta^{X_P} = \lambda_P\left(e^{\theta\eta+\frac{1}{2}\theta^2\delta^2}-1\right).$$

Conditioning with respect to the σ-algebra $\mathcal{F}_T^{X_I^*}$ we have that the moment $E_{\mathbb{Q}}[M^\theta(T_I(T))]$ can be expressed as

$$\begin{aligned} E_{\mathbb{Q}}[M^\theta(T_I(T))] &= E_{\mathbb{Q}}\left[e^{\theta X_P(T_I(T))-\theta T_I(T)\xi^{X_P}}\right] \\ &= E_{\mathbb{Q}}\left[e^{-\theta T_I(T)\xi^{X_P}}E_{\mathbb{Q}}\left[e^{\theta X_P(T_I(T))}\middle|\mathcal{F}_T^{X_I^*}\right]\right] \\ &= E_{\mathbb{Q}}\left[e^{T_I(T)\left(\xi_\theta^{X_P}-\theta\xi^{X_P}\right)}\right] < \infty, \end{aligned} \tag{5.4.19}$$

for some $\theta > 2$ by condition (5.4.16).

Let us now prove that the process $(M(T_I(u)))_{u\in[0,T]}$ is a martingale with respect to the filtration

$$(\mathcal{F}_u^*)_{u\in[0,T]} = \left(\mathcal{F}_u^{M(T_I)} \vee \mathcal{F}_u^{X_I^*} \vee \mathcal{F}_u^{W_P^*}\right)_{u\in[0,T]},$$

where

$$\mathcal{F}_u^{M(T_I)} = \sigma\left(\{M(T_I(s)) : 0\leq s\leq u\}\right).$$

First we define the filtration $\left(\bar{\mathcal{F}}_u\right)_{u\geq 0}$ as

$$\bar{\mathcal{F}}_u = \mathcal{F}_u^M \vee \mathcal{F}_T^{X_I^*} \vee \mathcal{F}_T^{W_P^*} \text{ for all } u\geq 0,$$

Notice that if $T\geq t\geq s\geq 0$ then the sets $\{T_I(T)\leq u\}, \{T_I(t)\leq u\}, \{T_I(s)\leq u\} \in \bar{\mathcal{F}}_u$ for all $u\geq 0$. So $T_I(T)$, $T_I(t)$ and $T_I(s)$ are stopping times under the filtration $\left(\bar{\mathcal{F}}_u\right)_{u\geq 0}$. Due to the fact that the filtration $\left(\mathcal{F}_u^M\right)_{u\geq 0}$ and the σ-algebra $\mathcal{F}_T^{X_I^*}\vee\mathcal{F}_T^{W_P^*}$ are independent we have that the $\left(\mathcal{F}_u^M\right)_{u\geq 0}$-martingale M is also a martingale with respect to the filtration $\left(\bar{\mathcal{F}}_u\right)_{u\geq 0}$ (Shojaee, 2018, Theorem 2.39). Since M is a martingale and $T_I(T)$ is a stopping time we have that the process $\tilde{M}$ defined as

$$\tilde{M}(u) = M(u\wedge T_I(T)) \text{ for } u\geq 0$$

is a martingale with respect to $\left(\bar{\mathcal{F}}_u\right)_{u\geq 0}$ (Klebaner, 2012, Theorem 7.14). Next we will show that $\tilde{M}$ is a uniform integrable martingale. To that end let us show that $\sup_{u\geq 0} E[\tilde{M}^{\theta/2}(u)] < \infty$ for some $\theta > 2$. The moment $E[\tilde{M}^{\theta/2}(u)]$ can be written as

$$\begin{aligned}
E[\tilde{M}^{\theta/2}(u)] &= E_{\mathbb{Q}}\left[e^{\frac{\theta}{2}X_P(u\wedge T_I(T))-\frac{\theta}{2}(T_I(T)\wedge u)\xi^{X_P}}\right] \\
&= E_{\mathbb{Q}}\left[e^{\frac{\theta}{2}X_P(u\wedge T_I(T))-\frac{T_I(T)\wedge u}{2}\xi_\theta^{X_P}}e^{\frac{T_I(T)\wedge u}{2}\xi_\theta^{X_P}-\frac{\theta}{2}(T_I(T)\wedge u)\xi^{X_P}}\right] \\
&\leq \sqrt{E_{\mathbb{Q}}\left[e^{\theta X_P(u\wedge T_I(T))-(T_I(T)\wedge u)\xi_\theta^{X_P}}\right]} \\
&\quad \sqrt{E_{\mathbb{Q}}\left[e^{(T_I(T)\wedge u)\xi_\theta^{X_P}-\theta(T_I(T)\wedge u)\xi^{X_P}}\right]} \\
&= \sqrt{E_{\mathbb{Q}}\left[e^{(T_I(T)\wedge u)(\xi_\theta^{X_P}-\theta\xi^{X_P})}\right]},
\end{aligned}$$

where the last equality comes from the fact that:

$$\begin{aligned}
E_{\mathbb{Q}}\left[e^{\theta X_P(u\wedge T_I(T))-(T_I(T)\wedge u)\xi_\theta^{X_P}}\right] &= E_{\mathbb{Q}}\left[e^{-(T_I(T)\wedge u)\xi_\theta^{X_P}}E_{\mathbb{Q}}\left[e^{\theta X_P(u\wedge T_I(T))}\middle|\mathcal{F}_T^{X_I^*}\right]\right] \\
&= E_{\mathbb{Q}}\left[e^{-(T_I(T)\wedge u)\xi_\theta^{X_P}}e^{(T_I(T)\wedge u)\Psi^{X_P}(-i\theta)}\right] = 1.
\end{aligned}$$

Notice that the constant $\xi_\theta^{X_P} - \theta\xi^{X_P} \in \mathbb{R}$ and hence we can write

$$\begin{aligned}
E[\tilde{M}^{\theta/2}(u)] &\leq \sqrt{E_{\mathbb{Q}}\left[e^{(T_I(T)\wedge u)(\xi_\theta^{X_P}-\theta\xi^{X_P})}\right]} \\
&\leq \begin{cases} \sqrt{E_{\mathbb{Q}}\left[e^{T_I(T)(\xi_\theta^{X_P}-\theta\xi^{X_P})}\right]} < \infty & \text{if } \xi_\theta^{X_P}-\theta\xi^{X_P} > 0 \\ 1 & \text{if } \xi_\theta^{X_P}-\theta\xi^{X_P} \leq 0 \end{cases}
\end{aligned} \tag{5.4.20}$$

by condition (5.4.16). Noticing that the right hand side of equation (5.4.20) does not depend on u, we have just shown that

$$\sup_{u\geq 0} E[\tilde{M}^{\theta/2}(u)] < \infty \text{ for some } \theta > 2$$

and hence $\tilde{M}$ is a uniform integrable martingale with respect to $\left(\bar{\mathcal{F}}_u\right)_{u\geq 0}$ (Klebaner, 2012, Theorem 7.7).

Because the process T_I is non-decreasing we have that $T_I(T) \geq T_I(t) \geq T_I(s)$ and by application of the optional sampling theorem (Klebaner, 2012, Theorem 7.18) we have that

$$E\left[M(T_I(t))\left|\bar{\mathcal{F}}_{T(s)}\right.\right] = E\left[\tilde{M}(T_I(t))\left|\bar{\mathcal{F}}_{T_I(s)}\right.\right] = \tilde{M}(T_I(s)) = M(T_I(s)) \text{ almost surely.}$$

The random variable $\tilde{M}(T_I(u)) = M(T_I(u))$ is measurable with respect to $\bar{\mathcal{F}}_{T_I(u)}$ for all $u \in [0,T]$ (Le Gall, 2016, Theorem 3.7). Since the stopping times $T_I(v) \geq T_I(u)$ for all $v, u \in [0,T]$ with $v \geq u$, we have that $\bar{\mathcal{F}}_{T_I(u)} \subset \bar{\mathcal{F}}_{T_I(v)}$ (Le Gall, 2016, p. 45). So we have that $\mathcal{F}_s^{M(T_I)} \subseteq \bar{\mathcal{F}}_{T_I(s)}$ and hence we can write

$$\mathcal{F}_s^* = \mathcal{F}_s^{M(T_I)} \vee \mathcal{F}_s^{X_I^*} \vee \mathcal{F}_s^{W_P^*} \subseteq \bar{\mathcal{F}}_{T_I(s)}.$$

By application of the tower property we obtain

$$
\begin{aligned}
E\left[M(T_I(t))\,|\mathcal{F}_s^*\right] &= E\left[E\left[M(T_I(t))\,\middle|\bar{\mathcal{F}}_{T_I(s)}\right]|\mathcal{F}_s^*\right] \\
&= E\left[M(T_I(s))\,|\mathcal{F}_s^*\right] \\
&= M(T_I(s)).
\end{aligned}
$$

We have just shown that the process $(M(T_I(u)))_{u\in[0,T]}$ is a martingale with respect to the filtration $(\mathcal{F}_u^*)_{u\geq 0}$. Notice that $X_P(T_I(s)) = \log(M(T_I(s))) + \xi^{X_P}T_I(s)$ is $\mathcal{F}_s^*$-measurable, so we can write:

$$
\mathcal{F}_s = \mathcal{F}_s^{X_P(T_I)} \vee \mathcal{F}_s^{X_I^*} \vee \mathcal{F}_s^{W_P^*} \subseteq \mathcal{F}_s^*,
$$

and finally we can write

$$
\begin{aligned}
E\left[M(T_I(t))\,|\mathcal{F}_s\right] &= E\left[E\left[M(T_I(t))\,|\mathcal{F}_s^*\right]|\mathcal{F}_s\right] \\
&= E\left[M(T_I(s))\,|\mathcal{F}_s\right] \\
&= M(T_I(s)).
\end{aligned}
$$

□

With the use of Lemma 5.4.1 we can now prove the martingale property of the discounted Bitcoin price.

Proposition 5.4.2. *If for some $\theta > 2$ we have that*

$$
E_{\mathbb{Q}}\left[e^{T_I(T)\left(\xi_\theta^{X_P}-\theta\xi^{X_P}\right)}\right] < \infty, \tag{5.4.21}
$$

where

$$
\begin{aligned}
\xi_\theta^{X_P} &= \lambda_P\left(e^{\theta\eta+\frac{1}{2}\theta^2\delta^2}-1\right), \\
\xi^{X_P} &= \lambda_P\left(e^{\eta+\frac{\delta^2}{2}}-1\right),
\end{aligned}
$$

then the discounted price process $\tilde{P}$ is a martingale with respect to the filtration $(\mathcal{F}_t)_{t\in[0,T]}$ in the probability space $(\Omega, \mathcal{F}, \mathbb{Q})$.

Proof. From equation (5.4.1) we have that the discounted price process $\tilde{P}$ can be written as

$$
\tilde{P}(t) = p\exp\left\{-\frac{\sigma^2}{2}t + \sigma W_P^*(t) + X_P(T_I(t)) - T_I(t)\xi^{X_P}\right\} \text{ for } t\in[0,T].
$$

Let $t, s \in [0,T]$ such that $t \geq s$ and let us compute the conditional expectation of $\tilde{P}(t)$ with respect to the σ-algebra $\mathcal{F}_s \vee \sigma(T_I(t)) \vee \sigma(X_P(T_I(t)))$. By independence of the increment $(W_P^*(t) - W_P^*(s))$ and the σ-algebra $\mathcal{F}_s \vee \sigma(T_I(t)) \vee \sigma(X_P(T_I(t)))$ we have

$$\begin{aligned}
&E_{\mathbb{Q}}\left[\tilde{P}(t)\middle|\mathcal{F}_s\vee\sigma(T_I(t))\vee\sigma(X_P(T_I(t)))\right]\\
&\qquad=pe^{-\frac{\sigma}{2}t+X_P(T_I(t))-T_I(t)\xi^{X_P}}\\
&\qquad\quad E_{\mathbb{Q}}\left[pe^{\sigma W_P^*(t)-\sigma W_P^*(s)+\sigma W_P^*(s)}\middle|\mathcal{F}_s\vee\sigma(T_I(t))\vee\sigma(X_P(T_I(t)))\right]\\
&\qquad=pe^{-\frac{\sigma}{2}t+X_P(T_I(t))-T_I(t)\xi^{X_P}+\sigma W_P^*(s)}\\
&\qquad\quad E_{\mathbb{Q}}\left[e^{\sigma\left(W_P^*(t)-W_P^*(s)\right)}\middle|\mathcal{F}_s\vee\sigma(T_I(t))\vee\sigma(X_P(T_I(t)))\right]\\
&\qquad=pe^{-\frac{\sigma}{2}t+X_P(T_I(t))-T_I(t)\xi^{X_P}+\sigma W_P^*(s)}E_{\mathbb{Q}}\left[e^{\sigma\left(W_P^*(t)-W_P^*(s)\right)}\right]\\
&\qquad=pe^{-\frac{\sigma}{2}t+X_P(T_I(t))-T_I(t)\xi^{X_P}+\sigma W_P^*(s)}e^{\frac{\sigma^2}{2}(t-s)}\\
&\qquad=pe^{-\frac{\sigma}{2}s+\sigma W_P^*(s)+X_P(T_I(t))-T_I(t)\xi^{X_P}}.
\end{aligned}$$

By application of the tower property, we have

$$\begin{aligned}
E_{\mathbb{Q}}\left[\tilde{P}(t)\middle|\mathcal{F}_s\right] &= E_{\mathbb{Q}}\left[E_{\mathbb{Q}}\left[\tilde{P}(t)\middle|\mathcal{F}_s\vee\sigma(T_I(t))\vee\sigma(X_P(T_I(t)))\right]\middle|\mathcal{F}_s\right]\\
&= E_{\mathbb{Q}}\left[pe^{-\frac{\sigma}{2}s+\sigma W_P^*(s)+X_P(T_I(t))-T_I(t)\xi^{X_P}}\middle|\mathcal{F}_s\right]\\
&= pe^{-\frac{\sigma}{2}s+\sigma W_P^*(s)}E_{\mathbb{Q}}\left[e^{X_P(T_I(t))-T_I(t)\xi^{X_P}}\middle|\mathcal{F}_s\right]. \qquad (5.4.22)
\end{aligned}$$

By condition (5.4.21) we can apply Lemma 5.4.1 and hence we have that:

$$\begin{aligned}
E_{\mathbb{Q}}\left[\tilde{P}(t)\middle|\mathcal{F}_s\right]= &= pe^{-\frac{\sigma}{2}s+\sigma W_P^*(s)}E_{\mathbb{Q}}\left[e^{X_P(T_I(t))-T_I(t)\xi^{X_P}}\middle|\mathcal{F}_s\right]\\
&= e^{-\frac{\sigma}{2}s+\sigma W_P^*(s)}e^{X_P(T_I(s))-T_I(s)\xi^{X_P}}=\tilde{P}(s),
\end{aligned}$$

as required. □

Since $\tilde{\mathcal{F}}_t\subseteq\mathcal{F}_t$ for all $t\in[0,T]$, by the tower property we have that:

$$\begin{aligned}
E\left[\tilde{P}(t)\middle|\tilde{\mathcal{F}}_s\right] &= E\left[E\left[\tilde{P}(t)\middle|\mathcal{F}_s\right]\middle|\tilde{\mathcal{F}}_s\right]\\
&= E\left[\tilde{P}(t)\middle|\tilde{\mathcal{F}}_s\right]\\
&= \tilde{P}(s) \text{ for } t\geq s \text{ with } t,s\in[0,T],
\end{aligned}$$

where the last equality comes from the fact that $\tilde{P}$ is adapted with respect to the delayed filtration $(\tilde{\mathcal{F}}_t)_{t\in[0,T]}$. So, we have just shown that $\tilde{P}$ is also a martingale with respect to the delayed filtration $(\tilde{\mathcal{F}}_t)_{t\in[0,T]}$.

5.4.4 Characteristic function

We are interested in computing the characteristic function of the natural logarithm of the price process. Using the characteristic function of the natural logarithm of the price, we can compute the price of vanilla options using the result in Carr and Madan (1999, p. 6). Of course, we require the discounted Bitcoin price to be a martingale. Because of that, we assume that condition (5.4.16) is satisfied.

Let us define the process X as:

$$X(t)=\log(P(t)) \text{ for } t\in[0,T].$$

From equation (5.4.1) we have that:

$$X(t) = x + \left(r - \frac{\sigma^2}{2}\right) t + \sigma W_P^*(t) + X_P(T_I(t)) - T_I(t)\xi^{X_P} \text{ for } t \in [0,T], \quad (5.4.23)$$

where $x = \log(p)$. The option has an expiration date T and because of that we are interested in computing the characteristic function of the random variable $X(T)$. The characteristic function of $X(T)$ is defined as

$$\Phi^{X(T)}(u) = E_{\mathbb{Q}}\left[e^{iuX(T)}\right] \text{ for all } u \in \mathbb{R}.$$

Proposition 5.4.3. *The characteristic function of $X(T)$ can be expressed as:*

$$\Phi^{X(T)}(u) = e^{iux+iu\left(r-\frac{\sigma^2}{2}\right)T-\frac{\sigma^2 u^2 T}{2}} E_{\mathbb{Q}}\left[e^{T_I(T)\left(\Psi^{X_P}(u)-iu\xi^{X_P}\right)}\right],$$

where

$$\begin{aligned}
\Psi^{X_P}(u) &= \lambda_P\left(e^{iu\eta-\frac{1}{2}u^2\delta^2}-1\right),\\
\xi^{X_P} &= \lambda_P\left(e^{\eta+\frac{\delta^2}{2}}-1\right).
\end{aligned}$$

Proof. Conditioning with respect to $\mathcal{F}_T^{W_P^*} \vee \mathcal{F}_T^{X_I^*}$ and applying Proposition C.0.1, we have that:

$$\begin{aligned}
&E_{\mathbb{Q}}\left[e^{iuX(T)}\Big|\mathcal{F}_T^{W_P^*} \vee \mathcal{F}_T^{X_I^*}\right]\\
&\quad= E_{\mathbb{Q}}\left[e^{iux+iu\left(r-\frac{\sigma^2}{2}\right)T+iu\sigma W_P^*(T)+iuX_P(T_I(T))-iuT_I(T)\xi^{X_P}}\Big|\mathcal{F}_T^{W_P^*} \vee \mathcal{F}_T^{X_I^*}\right]\\
&\quad= e^{iux+iu\left(r-\frac{\sigma^2}{2}\right)T+iu\sigma W_P^*(T)-iuT_I(T)\xi^{X_P}} E_{\mathbb{Q}}\left[e^{iuX_P(T_I(T))}\Big|\mathcal{F}_T^{W_P^*} \vee \mathcal{F}_T^{X_I^*}\right]\\
&\quad= e^{iux+iu\left(r-\frac{\sigma^2}{2}\right)T+iu\sigma W_P^*(T)-iuT_I(T)\xi^{X_P}} e^{\Psi^{X_P}(u)T_I(T)},
\end{aligned}$$

where $\Psi^{X_P}(u)$ is the characteristic exponent of the compound Poisson process X_P. Now, by application of the tower property we have:

$$\begin{aligned}
E_{\mathbb{Q}}\left[e^{iuX(T)}\right] &= E_{\mathbb{Q}}\left[E_{\mathbb{Q}}\left[e^{iuX(T)}\Big|\mathcal{F}_T^{W_P^*} \vee \mathcal{F}_T^{X_I^*}\right]\right]\\
&= E_{\mathbb{Q}}\left[e^{iux+iu\left(r-\frac{\sigma^2}{2}\right)T+iu\sigma W_P^*(T)+\Psi^{X_P}(u)T_I(T)-iuT_I(T)\xi^{X_P}}\right].
\end{aligned}$$

Conditioning now with respect to $\mathcal{F}_T^{X_I^*}$ we obtain:

$$\begin{aligned}
&E_{\mathbb{Q}}\left[e^{iux+iu\left(r-\frac{\sigma^2}{2}\right)T+iu\sigma W_P^*(T)+\Psi^{X_P}(u)T_I(T)-iuT_I(T)\xi^{X_P}}\Big|\mathcal{F}_T^{X_I^*}\right]\\
&\qquad= e^{iux+iu\left(r-\frac{\sigma^2}{2}\right)T+\Psi^{X_P}(u)T_I(T)-iuT_I(T)\xi^{X_P}} E_{\mathbb{Q}}\left[e^{iu\sigma W_P^*(T)}\Big|\mathcal{F}_T^{X_I^*}\right]\\
&\qquad= e^{iux+iu\left(r-\frac{\sigma^2}{2}\right)T+\Psi^{X_P}(u)T_I(T)-iuT_I(T)\xi^{X_P}} E_{\mathbb{Q}}\left[e^{iu\sigma W_P^*(T)}\right]\\
&\qquad= e^{iux+iu\left(r-\frac{\sigma^2}{2}\right)T+\Psi^{X_P}(u)T_I(T)-iuT_I(T)\xi^{X_P}} e^{-\frac{\sigma^2 u^2 T}{2}}.
\end{aligned}$$

Lastly, by application of the tower property we obtain:

$$
\begin{aligned}
&E_{\mathbb{Q}}\left[e^{iuX(T)}\right] \\
&\qquad = E_{\mathbb{Q}}\left[E_{\mathbb{Q}}\left[e^{iux+iu\left(r-\frac{\sigma^2}{2}\right)T+iu\sigma W_P^*(T)+\Psi^{X_P}(u)T_I(T)-iuT_I(T)\xi^{X_P}} \,\middle|\, \mathcal{F}_T^{X_I^*}\right]\right] \\
&\qquad = E_{\mathbb{Q}}\left[e^{iux+iu\left(r-\frac{\sigma^2}{2}\right)T-\frac{\sigma^2u^2T}{2}+\Psi^{X_P}(u)T_I(T)-iuT_I(T)\xi^{X_P}}\right] \\
&\qquad = e^{iux+iu\left(r-\frac{\sigma^2}{2}\right)T-\frac{\sigma^2u^2T}{2}} E_{\mathbb{Q}}\left[e^{T_I(T)\left(\Psi^{X_P}(u)-iu\xi^{X_P}\right)}\right].
\end{aligned}
$$

□

Notice that the random variable $T_I(T)$ can be written as:

$$
T_I(T) = \int_0^T I(s-\tau)ds = \int_{-\tau}^{T-\tau} I(s)ds = \begin{cases} \int_{-\tau}^0 \phi^I(s)ds + \int_0^{T-\tau} I(s)ds & \text{if } T > \tau \\ \int_{-\tau}^{T-\tau} \phi^I(s)ds & \text{if } T \leq \tau. \end{cases} \tag{5.4.24}
$$

From equation (5.4.24) it is clear that we have to differentiate between the cases when $T \leq \tau$ and $T > \tau$.

5.4.5 Characteristic function when $T \leq \tau$

From equation (5.4.24) and Proposition 5.4.3 we have that the characteristic function of $X(T)$ when $T \leq \tau$ is

$$
\begin{aligned}
\Phi^{X(T)}(u) &= e^{iux+iu\left(r-\frac{\sigma^2}{2}\right)T-\frac{\sigma^2u^2T}{2}} E_{\mathbb{Q}}\left[e^{\left(\int_{-\tau}^{T-\tau}\phi^I(s)ds\right)\left(\Psi^{X_P}(u)-iu\xi^{X_P}\right)}\right] \\
&= e^{iux+iu\left(r-\frac{\sigma^2}{2}\right)T-\frac{\sigma^2u^2T}{2}+\left(\int_{-\tau}^{T-\tau}\phi^I(s)ds\right)\left(\Psi^{X_P}(u)-iu\xi^{X_P}\right)}.
\end{aligned} \tag{5.4.25}
$$

Notice that from equation (5.4.25) we have that the price of the option does not depend on the interest process chosen when $T \leq \tau$. That is, we have the same formula if we choose a Cox-Ingersoll-Ross process or an inverse Gaussian Ornstein-Uhlenbeck process. However, this is not the case for the options with expiration date $T > \tau$; in those cases the price of the option is affected by the market attention process.

5.4.6 Characteristic function when $T > \tau$

In the case when $T > \tau$, from equation (5.4.24) and Proposition 5.4.3 we have that

$$
\begin{aligned}
\Phi^{X(T)}(u) &= e^{iux+iu\left(r-\frac{\sigma^2}{2}\right)T-\frac{\sigma^2u^2T}{2}} E_{\mathbb{Q}}\left[e^{\left(\int_{-\tau}^{0}\phi^I(s)ds+\int_0^{T-\tau}I(s)ds\right)\left(\Psi^{X_P}(u)-iu\xi^{X_P}\right)}\right] \\
&= e^{iux+iu\left(r-\frac{\sigma^2}{2}\right)T-\frac{\sigma^2u^2T}{2}+\epsilon_0\left(\Psi^{X_P}(u)-iu\xi^{X_P}\right)} \\
&\quad E_{\mathbb{Q}}\left[e^{\left(\int_0^{T-\tau}I(s)ds\right)\left(\Psi^{X_P}(u)-iu\xi^{X_P}\right)}\right],
\end{aligned} \tag{5.4.26}
$$

where $\epsilon_0 = \int_{-\tau}^0 \phi^I(u)du$. Define the process Y as

$$
Y(t) = \int_0^t I(s)ds \text{ for } t \geq 0. \tag{5.4.27}
$$

The expectation that appears in equation (5.4.26) can be written as

$$E_{\mathbb{Q}}\left[e^{\left(\int_0^{T-\tau} I(s)ds\right)\left(\Psi^{X_P}(u)-iu\xi^{X_P}\right)}\right] = E_{\mathbb{Q}}\left[e^{Y(T-\tau)\left(\Psi^{X_P}(u)-iu\xi^{X_P}\right)}\right]. \tag{5.4.28}$$

Notice that from the proof of Proposition 5.4.3 we have that the expectation in (5.4.28) is well defined since

$$\begin{aligned}\left|E_{\mathbb{Q}}\left[e^{Y(T-\tau)\left(\Psi^{X_P}(u)-iu\xi^{X_P}\right)}\right]\right| &= \left|E_{\mathbb{Q}}\left[e^{iuX_P(Y(T-\tau))-iu\xi^{X_P}Y(T-\tau)}\right]\right| \\ &\leq E_{\mathbb{Q}}\left[\left|e^{iuX_P(Y(T-\tau))-iu\xi^{X_P}Y(T-\tau)}\right|\right] = 1.\end{aligned}$$

So proceeding as in Carr, Geman, Madan, and Yor (2003, p. 359), we have that

$$E_{\mathbb{Q}}\left[e^{Y(T-\tau)\left(\Psi^{X_P}(u)-iu\xi^{X_P}\right)}\right] = \Phi^{Y(T-\tau)}\left(-i\Psi^{X_P}(u) - u\xi^{X_P}\right),$$

where $\Phi^{Y(T-\tau)}$ is the characteristic function of the random variable $Y(T-\tau)$.

Hence the characteristic function of $X(T)$ can be written as

$$\Phi^{X(T)}(u) = e^{iux+iu\left(r-\frac{\sigma^2}{2}\right)T-\frac{\sigma^2u^2T}{2}+\epsilon_0\left(\Psi^{X_P}(u)-iu\xi^{X_P}\right)}\Phi^{Y(T-\tau)}\left(-i\Psi^{X_P}(u) - u\xi^{X_P}\right). \tag{5.4.29}$$

From equation (5.4.29) we have that the characteristic function of $X(T)$ depends on the characteristic function of $Y(T-\tau)$. So in this case, the price of the option will depend on the interest process. Hence we have different formulas depending on the interest process that we choose.

1. If the interest process I is an Cox-Ingersoll-Ross process and satisfies equation (5.4.7), then the characteristic function of the random variable $Y(t)$ has the form:

$$\Phi^{Y(t)}(u) = \frac{e^{\frac{\tilde{a}_I^2\tilde{b}_I t}{\sigma_I^2}}\, e^{\frac{2I(0)iu}{\tilde{a}_I+\gamma\coth(\gamma t/2)}}}{\left(\cosh\left(\frac{\gamma t}{2}\right)+\frac{a_I}{\gamma}\sinh\left(\frac{\gamma t}{2}\right)\right)^{\frac{2\tilde{a}_I\tilde{b}_I}{\sigma_I^2}}},$$

where $\gamma = \sqrt{\tilde{a}_I^2 - 2\sigma_I^2 iu}$, $u \in \mathbb{R}$ and $t \geq 0$ (Schoutens, 2003, p. 89).

2. In the case that the market attention process I satisfies an $IG(a_I, \tilde{b}_I)$-Ornstein-Uhlenbeck process with parameter $\lambda_I > 0$, the characteristic function of $Y(t)$ can be written as

$$\Phi^{Y(t)}(u) = \exp\left\{\frac{iuI(0)}{\lambda_I}\left(1-e^{-\lambda_I t}\right) + \frac{2a_I iu}{\tilde{b}_I\lambda_I}A(u,t)\right\}, \tag{5.4.30}$$

where

$$\begin{aligned}A(u,t) &= \frac{1-\sqrt{1+\kappa\left(1-e^{-\lambda_I t}\right)}}{\kappa} \\ &\quad +\frac{1}{\sqrt{1+\kappa}}\left(arctanh\left(\frac{\sqrt{1+\kappa\left(1-e^{-\lambda_I t}\right)}}{\sqrt{1+\kappa}}\right) - arctanh\left(\frac{1}{\sqrt{1+\kappa}}\right)\right) \\ \kappa &= -\frac{2iu}{\tilde{b}_I^2\lambda_I},\end{aligned}$$

for $u \in \mathbb{R}$ and $t \geq 0$ (Schoutens, 2003, p. 91).

5.5 Exponential moments

As we have seen, we need to study the exponential moments of log-price to see if the price is a martingale (see Proposition 5.4.2). We also need to study the exponential moments because we use the characteristic function for pricing options. Let $\theta \in \mathbb{R}$; the exponential moments of the random variable $X(T)$ defined in (5.4.23) can be expressed as

$$E_{\mathbb{Q}}\left[e^{\theta X(T)}\right] = E_{\mathbb{Q}}\left[e^{\theta x+\theta\left(r-\frac{\sigma^2}{2}\right)T+\theta\sigma W_P^*(T)+\theta X_P(T_I(T))-\theta T_I(T)\xi^{X_P}}\right].$$

It is easy to see that the exponential moment of $X(T)$ can be written as

$$E_{\mathbb{Q}}\left[e^{\theta X(T)}\right] = E_{\mathbb{Q}}\left[e^{\theta x+\theta\left(r-\frac{\sigma^2}{2}\right)T+\theta\sigma W_P^*(T)}\right] E_{\mathbb{Q}}\left[e^{\theta X_P(T_I(T))-\theta T_I(T)\xi^{X_P}}\right]. \tag{5.5.1}$$

The first expected value that appears in equation (5.5.1) is clearly finite, since it is the exponential moment of a normal random variable. So to study the exponential moment of $X(T)$, we need to study the expected value $E_{\mathbb{Q}}\left[e^{\theta X_P(T_I(T))-\theta T_I(T)\xi^{X_P}}\right]$. In the proof of Proposition 5.4.2 (see equation (5.4.19)) we showed that

$$E_{\mathbb{Q}}\left[e^{\theta X_P(T_I(T))-\theta T_I(T)\xi^{X_P}}\right] = E_{\mathbb{Q}}\left[e^{T_I(T)\left(\xi_\theta^{X_P}-\theta\xi^{X_P}\right)}\right], \tag{5.5.2}$$

where

$$\begin{aligned}
\xi_\theta^{X_P} &= \lambda_P\left(e^{\theta\eta+\frac{1}{2}\theta^2\delta^2}-1\right),\\
\xi^{X_P} &= \lambda_P\left(e^{\eta+\frac{\delta^2}{2}}-1\right).
\end{aligned}$$

From equation (5.4.24) we have to distinguish between the case when $T \le \tau$ and when $T > \tau$.

When $T \le \tau$ we have that equation (5.5.2) can be expressed as

$$\begin{aligned}
E_{\mathbb{Q}}\left[e^{\theta X_P(T_I(T))-\theta T_I(T)\xi^{X_P}}\right] &= E_{\mathbb{Q}}\left[e^{T_I(T)\left(\xi_\theta^{X_P}-\theta\xi^{X_P}\right)}\right]\\
&= E_{\mathbb{Q}}\left[e^{\left(\int_{-\tau}^{T-\tau}\phi^I(s)ds\right)\left(\xi_\theta^{X_P}-\theta\xi^{X_P}\right)}\right]\\
&= e^{\left(\int_{-\tau}^{T-\tau}\phi^I(s)ds\right)\left(\xi_\theta^{X_P}-\theta\xi^{X_P}\right)} < \infty.
\end{aligned}$$

So when $T \le \tau$ all the exponential moments of $X(T)$ are finite.

A more interesting case is when $T > \tau$. When $T > \tau$ we have that equation (5.5.2) can be expressed as

$$\begin{aligned}
E_{\mathbb{Q}}\left[e^{\theta X_P(T_I(T))-\theta T_I(T)\xi^{X_P}}\right] &= E_{\mathbb{Q}}\left[e^{T_I(T)\left(\xi_\theta^{X_P}-\theta\xi^{X_P}\right)}\right]\\
&= E_{\mathbb{Q}}\left[e^{\left(\int_{-\tau}^{0}\phi^I(s)ds+\int_0^{T-\tau}I(s)ds\right)\left(\xi_\theta^{X_P}-\theta\xi^{X_P}\right)}\right]\\
&= e^{\epsilon_0\left(\xi_\theta^{X_P}-\theta\xi^{X_P}\right)} E_{\mathbb{Q}}\left[e^{Y(T-\tau)\tilde{\theta}}\right], \qquad (5.5.3)
\end{aligned}$$

where $\epsilon_0 = \int_{-\tau}^{0} \phi^I(s)ds$, Y is the process defined by (5.4.27) and

$$\tilde{\theta} \quad = \quad \left(\xi_\theta^{X_P} - \theta\xi^{X_P}\right) \in \mathbb{R}.$$

From equation (5.5.3) we have that the exponential moment of $X(T)$ is finite if and only if $E_{\mathbb{Q}}\left[e^{Y(T-\tau)\tilde{\theta}}\right]$ is finite. Let us define $T^* = T - \tau$ and notice that when $\tilde{\theta} \leq 0$ we have that

$$E_{\mathbb{Q}}\left[e^{Y(T^*)\tilde{\theta}}\right] \leq 1$$

since $Y(T^*) > 0$. So any explosions of the moment would only appear when $\tilde{\theta} > 0$.

To study the exponential moment of $Y(T^*)$, we need to differentiate between the case when I is a Cox-Ingersoll-Ross process and when I is an inverse Gaussian Ornstein-Uhlenbeck process. In the following subsections, we will focus on the cases when $\tilde{\theta} > 0$.

5.5.1 Interest is a Cox-Ingersoll-Ross process

In the case that the market attention follows a Cox-Ingersoll-Ross process, we have that the exponential moment of $Y(T^*)$ can be expressed as

$$E_{\mathbb{Q}}\left[e^{\tilde{\theta}Y(T^*)}\right] = \exp\left\{A(\tilde{\theta}, T^*) + I(0)B(\tilde{\theta}, T^*)\right\}, \tag{5.5.4}$$

where

$$\begin{aligned} A(\tilde{\theta}, T^*) &= \frac{\tilde{a}_I^2\tilde{b}_I t}{\sigma_I^2} - \frac{2\tilde{a}_I\tilde{b}_I}{\sigma_I^2}\log\left[\sinh\left(\frac{\gamma(\tilde{\theta})T^*}{2}\right)\left(\coth\left(\frac{\gamma(\tilde{\theta})T^*}{2}\right) + \frac{\tilde{a}_I}{\gamma(\tilde{\theta})}\right)\right], \\ B(\tilde{\theta}, T^*) &= \frac{2\tilde{\theta}}{\tilde{a}_I + \gamma(\tilde{\theta})\coth\left(\frac{\gamma(\tilde{\theta})T^*}{2}\right)}, \\ \gamma(\tilde{\theta}) &= \sqrt{\tilde{a}_I^2 - 2\sigma_I^2\tilde{\theta}}, \end{aligned} \tag{5.5.5}$$

(Friz & Keller-Ressel, 2010, p. 5).

From Friz and Keller-Ressel (2010, p. 5) we have that equation (5.5.5) and hence (5.5.4), is well defined for $\tilde{\theta} < \tilde{\theta}^*$ and explodes when $\tilde{\theta} \uparrow \tilde{\theta}^*$, where $\tilde{\theta}^* > 0$ is the solution of the equation $H(\tilde{\theta}^*) = 0$ where

$$H(l) = \tilde{a}_I + \gamma(l)\coth\left(\frac{\gamma(l)T^*}{2}\right).$$

Notice that since $\coth(x) > 0$ for $x > 0$ and $\tilde{a}_I > 0$ we have that when $\tilde{a}_I^2 - 2\sigma_I^2 l > 0$ then $H(l) > 0$. Hence if

$$\tilde{\theta} < \frac{\tilde{a}_I^2}{2\sigma_I^2} \tag{5.5.6}$$

then the exponential moment of $Y(T^*)$ is finite for all $T^* \geq 0$.

5.5.2 Interest is an inverse Gaussian Ornstein-Uhlenbeck process

If the interest process satisfies an $IG(a_I, \tilde{b}_I)$-Ornstein-Uhlenbeck process with parameter $\lambda_I > 0$, we will study the exponential moments of the following Itô integral:

$$\int_0^{T^*} \sqrt{I^-(u)} dB(u),$$

where B is a Brownian motion independent of the background driving Lévy process Z_I. From Proposition 1.4.1 we have that

$$\int_0^{T^*} \sqrt{I^-(u)} dB(u) \bigg| \mathcal{F}_{T^*}^{Z_I} \sim N\left(0, \int_0^{T^*} I^-(u) du\right).$$

Hence we can write

$$\begin{aligned} E_{\mathbb{Q}}\left[e^{\sqrt{2\tilde{\theta}} \int_0^{T^*} \sqrt{I^-(u)} dB(u)}\right] &= E_{\mathbb{Q}}\left[E_{\mathbb{Q}}\left[e^{\sqrt{2\tilde{\theta}} \int_0^{T^*} \sqrt{I^-(u)} dB(u)} \bigg| \mathcal{F}_{T^*}^{Z_I}\right]\right] \\ &= E_{\mathbb{Q}}\left[e^{\tilde{\theta} \int_0^{T^*} I^-(u) du}\right] \\ &= E_{\mathbb{Q}}\left[e^{\tilde{\theta} Y(T^*)}\right]. \end{aligned}$$

From the result by Nicolato and Venardos (2003, Theorem 2.2) we have that the moment

$$E_{\mathbb{Q}}\left[e^{\sqrt{2\tilde{\theta}} \int_0^{T^*} \sqrt{I^-(u)} dB(u)}\right]$$

is finite if

$$\sqrt{2\tilde{\theta}} < \inf_{0 \le s < T^*} \sqrt{\tilde{b}_I^2 \epsilon(s, T^*)^{-1}}, \tag{5.5.7}$$

where

$$\epsilon(s, T^*) = \frac{1 - e^{-\lambda_I (T^* - s)}}{\lambda_I} \quad \text{for } s \in [0, T^*].$$

So condition (5.5.7) can be written as

$$\sqrt{2\tilde{\theta}} < \inf_{0 \le s < T^*} \sqrt{\tilde{b}_I^2 \frac{\lambda_I}{1 - e^{-\lambda_I (T^* - s)}}}, \tag{5.5.8}$$

Because $\epsilon(s, T^*)^{-1}$ is an increasing function with respect to s when $s \in [0, T^*]$, we have that the condition (5.5.8) can be expressed as

$$\sqrt{2\tilde{\theta}} < \sqrt{\tilde{b}_I^2 \frac{\lambda_I}{1 - e^{-\lambda_T T^*}}}. \tag{5.5.9}$$

Hence, if condition (5.5.9) is satisfied, then the moment $E_{\mathbb{Q}}\left[e^{\tilde{\theta} Y(T^*)}\right]$ is finite.

5.6 Market option prices

With the models that are introduced in this chapter, we would like to price European options and compare the results obtained with the true market prices. The data used in this chapter are the same as the one used in Chapter 3. For evaluating the performance of our model, we again compute the root mean square error and the relative root mean square error.

In this chapter, we consider that the interest process can satisfy a Cox-Ingersoll-Ross process or an inverse Gaussian Ornstein-Uhlenbeck process. Because of that, we split the analysis when the interest follows a Cox-Ingersoll-Ross process and when the interest follows an inverse Gaussian Ornstein-Uhlenbeck process.

5.6.1 Cox-Ingersoll-Ross process

In Chapter 2, we saw that we can use two proxies for the interest when it follows a Cox-Ingersoll-Ross process. The two proxies are the unique number of active addresses and the number of views in Wikipedia of the keyword "Altcoin". We first do the analysis when the selected proxy is the unique number of active addresses and then when the selected proxy is the number of views in Wikipedia.

When the proxy of the interest is the number of unique active addresses, the estimated parameters are shown in Figure 5.1 and Figure 5.2. The results show that the parameters change through time and that the delay parameter is always greater than zero.

With the estimated values of the parameters, we would like to compute European option prices and compare them with the true market prices. We compute the root mean square error and the relative root mean square error; these are shown in Figure 5.3 and in Figure 5.4 respectively. As we saw in Section 5.4.1, we can select different values for the parameters λ_1 and λ_2 defined in equation (5.4.5); that is, we can select different risk–neutral probabilities. In our case, we are interested in the risk–neutral probability that minimizes the relative root mean square error. That is, we select values of λ_1 and λ_2 that minimize the relative root mean square error. For these new values of λ_1 and λ_2 we compute new values for the root mean square error and the relative root mean square error. These are shown in Figure 5.3 and in Figure 5.4 respectively. As we can observe in Figure 5.5 the values of the calibrated $\tilde{a}_I$ and $\tilde{b}_I$ are in general bigger than the estimated values of the parameters. In addition, we compute the root mean square error and the relative root mean square error obtained by the Black-Scholes-Merton model. Again, these are shown in Figures 5.3–5.4. As we can observe when we calibrate the proposed model, we obtain better results than the Black-Scholes-Merton model. These results can also be observed in Figure 5.6, where the ratio between the relative root mean square error given by our model and the relative root mean square error given by the Black-Scholes-Merton model is given.

In the case that the interest proxy is the number of views of the word "Altcoin", we obtain the estimated parameters that appear in Figures 5.7–5.8. With these values for the parameters, we calculate the values of European options and compare them with their respective market values. The root mean square error and the relative root mean square error are shown in Figure 5.9 and in Figure 5.10 respectively. Again, we can choose different values for the parameters λ_1 and λ_2. As we did before, we select the values of λ_1 and λ_2 that minimize the relative root mean square error. For these new values of λ_1 and λ_2 we have new values for the parameters $\tilde{a}_I$ and $\tilde{b}_I$, these are shown in Figure 5.11. As we can observe, the estimated values of the parameters $\tilde{a}_I$ and $\tilde{b}_I$ are in general greater than the calibrated ones. With these new values of the parameters $\tilde{a}_I$ and $\tilde{b}_I$ we price European options and compute the root mean square error and the relative root mean square error, these are shown in Figures 5.9–5.10. Again, we compute the root mean square error and the relative root mean square error given by the Black-Scholes-Merton model; these appear in Figures 5.9–5.10. For this case, the calibrated proposed model obtains better results than the Black-Scholes-Merton model. As we did before, we show the relative root mean square error given by our model divided by the relative root mean square error given by the Black-Scholes-Merton model in Figure 5.12.

5.6.2 Inverse Gaussian Ornstein-Uhlenbeck process

As we saw in Chapter 2, the logarithm of the number of unique active addresses satisfies an inverse Gaussian Ornstein-Uhlenbeck process. The estimated parameters obtained for the different temporal windows are shown in Figures 5.13–5.14.

As we can observe, the values we obtain for the parameter λ_I are all above 100. As discussed in Chapter 4 for these values of λ_I, the computer is not able to compute the characteristic function that appears in equation (5.4.30). When the value of λ_I is high, the following value that appears on the characteristic function satisfies the following

$$\operatorname{arctanh}\left(\frac{\sqrt{1+\kappa\left(1-e^{-\lambda_I t}\right)}}{\sqrt{1+\kappa}}\right) \approx \operatorname{arctanh}(1).$$

But the arctanh is not defined for the value 1 and hence we are not able to compute the characteristic function. Because of that, we would need an approximation formula for computing European call options, but the computation of that approximation formula is outside of the scope of this thesis.

5.7 Summary of models

We are interested in comparing the results obtained by the models presented in Chapter 3, Chapter 4 and Chapter 5. We first compare the relative root mean square error given by our models with the relative root mean square error given by the Black-Scholes-Merton model. As we did before, we divide the relative root mean square error given by the presented models by the relative root mean square error given by the Black-Scholes-Merton model. The results are shown in Figure 5.15. As we can observe, at least one of our models always gets better results than the Black-Scholes-Merton model for each date. In addition, it is interesting to see that the models that use the number of Wikipedia views as proxy for the interest are the ones that obtain the best results.

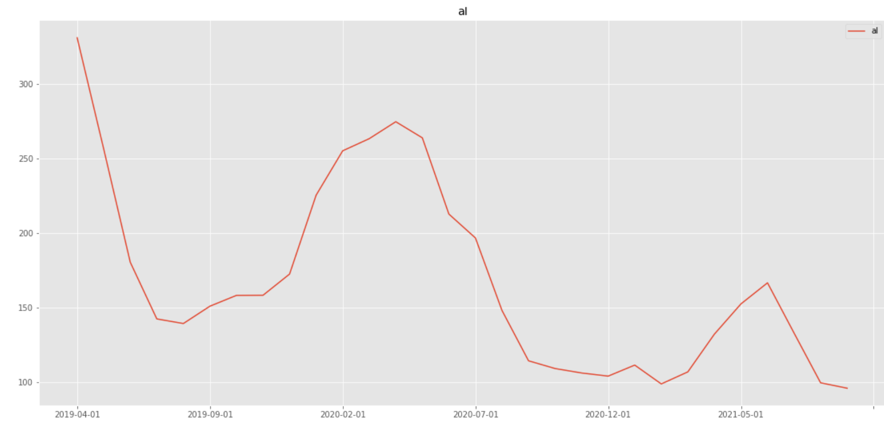
All the models shown in Chapter 3, Chapter 4 and Chapter 5 have several risk neutral probability measures. Because of that, we can select the risk neutral measure that minimize the relative root mean square error. Now we would like to compare the relative root mean square error given by our models when they are calibrated against options market data. To that end, we divide the relative root mean square error given by our calibrated models by the root mean square error given by the Black-Scholes-Merton model. We can see in Figure 5.16 that our calibrated models give better results than the Black-Scholes-Merton model. When comparing Figure 5.15 and Figure 5.16 it is clear that the calibrated models give better results than the models without calibration. As happened before, the models that have the number of Wikipedia views as a proxy for the interest give better results.

5.8 Conclusion and future work

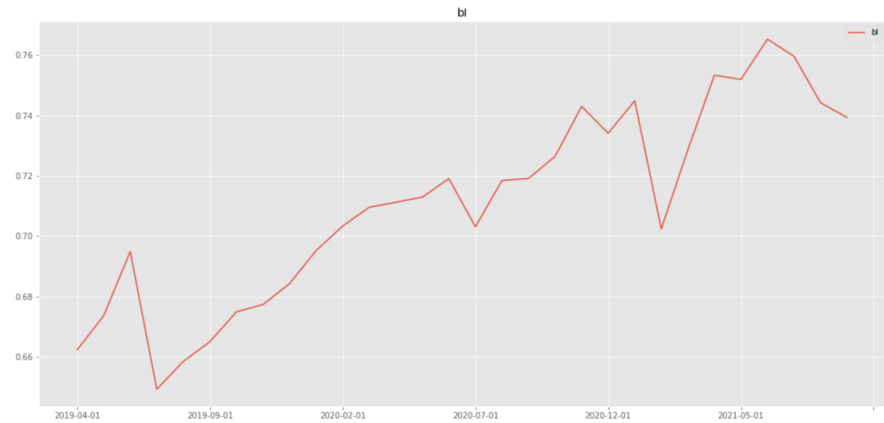
In this chapter, we developed time changed models in which the probability of a price jump increases when the market attention is high. On the contrary, if the market attention is low, the probability of having a jump decreases. We showed how we can estimate the parameters of these models and how we can price European options.

One of the issues of the presented models is that these models do not capture the asymmetry of the jumps. One solution would be having a time changed compound Poisson process with double exponential distributed jumps (Kou, 2002). Another solution proposed in the literature is the use of a time changed tempered stable process (Klingler, Kim, Rachev, & Fabozzi, 2013).

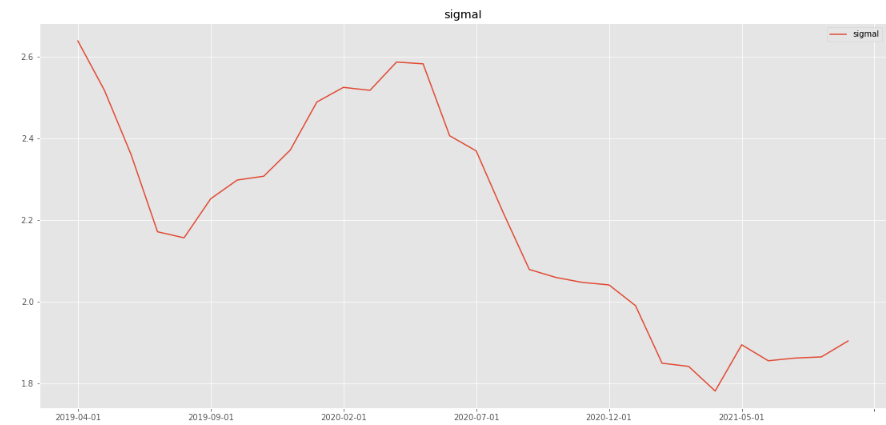
Furthermore, it could be interesting to construct models that have several market attention proxies instead of having just one proxy. In this case, our price process would be formed by several time changed compound Poisson processes. The problem with this approach is that the estimation of parameters becomes more difficult.



(A) Estimation of the parameter a_I.

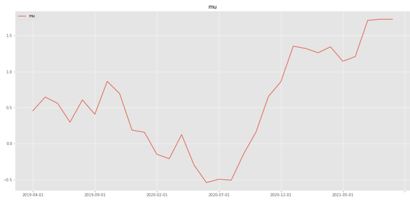


(B) Estimation of the parameter b_I.

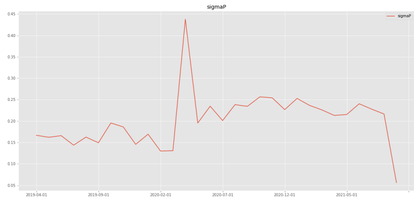


(C) Estimation of the parameter σ_I.

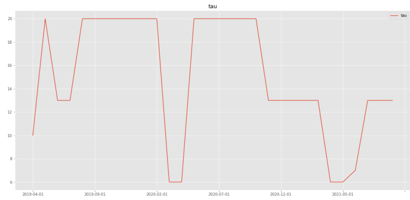
FIGURE 5.1: Estimated values for the parameters of the model defined in Section 5.1 when the proxy of the interest is the number of unique active addresses and the interest is a Cox-Ingersoll-Ross process.



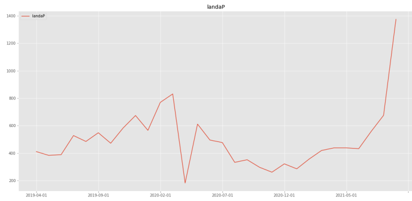
(A) Estimation of the parameter μ.



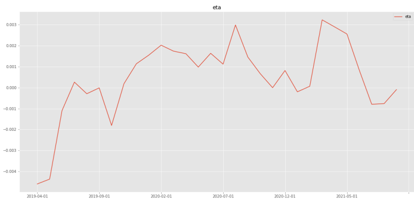
(B) Estimation of the parameter σ_P.



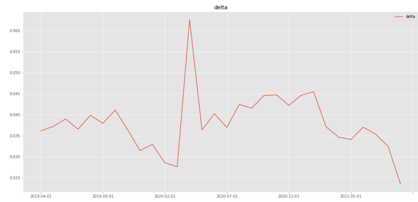
(C) Estimation of the parameter τ.



(D) Estimation of the parameter λ_P.



(E) Estimation of the parameter η.



(F) Estimation of the parameter δ.

FIGURE 5.2: Estimated values for the parameters of the model defined in Section 5.1 when the proxy of the interest is the number of unique active addresses and the interest is a Cox-Ingersoll-Ross process.

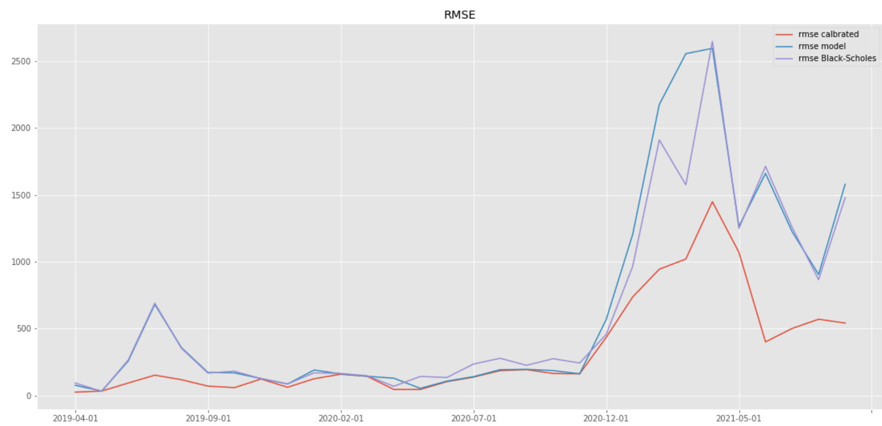


FIGURE 5.3: Root mean square error when the proxy of the interest is the number of unique active addresses.

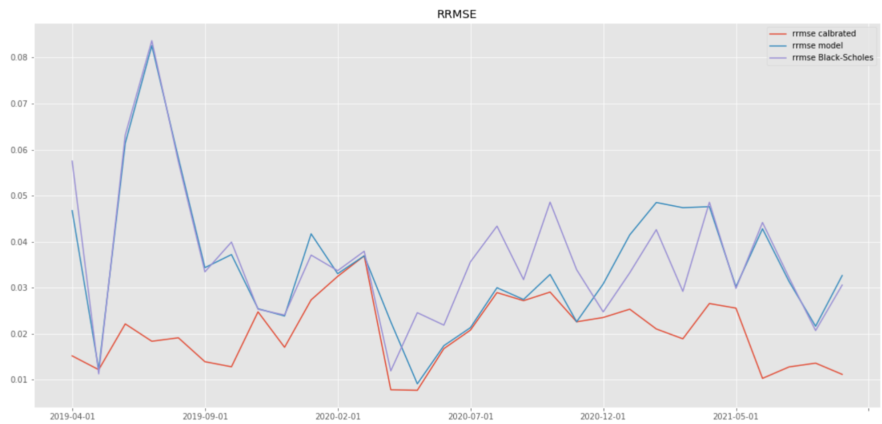
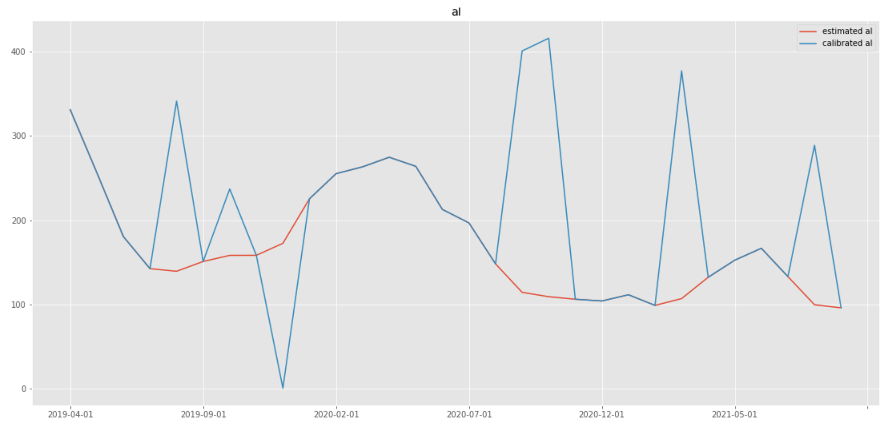
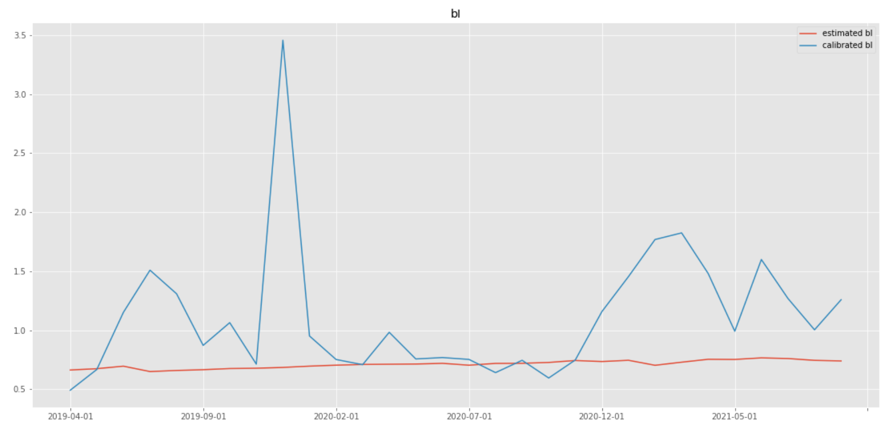


FIGURE 5.4: Relative root mean square error when the proxy of the interest is the number of unique active addresses and the interest is a Cox-Ingersoll-Ross process.



(A) Estimated and calibrated values for the parameter a_I.



(B) Estimated and calibrated values for the parameter b_I.

FIGURE 5.5: Calibrated values for the parameters a_I and b_I when the proxy of the interest is the number of unique active addresses and the interest is a Cox-Ingersoll-Ross process.

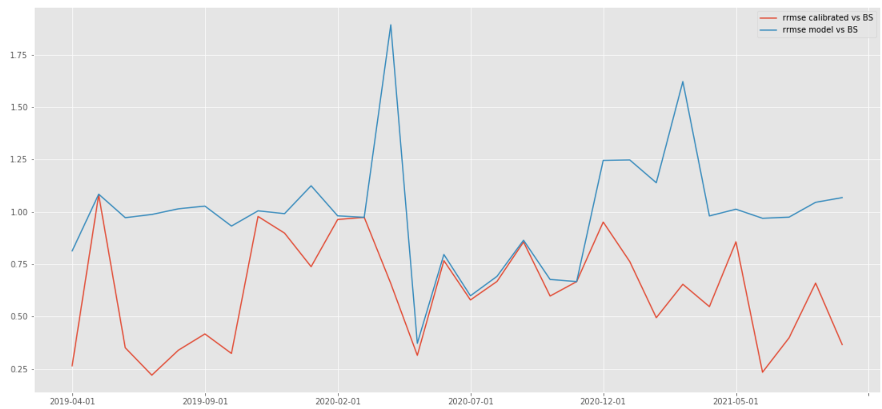
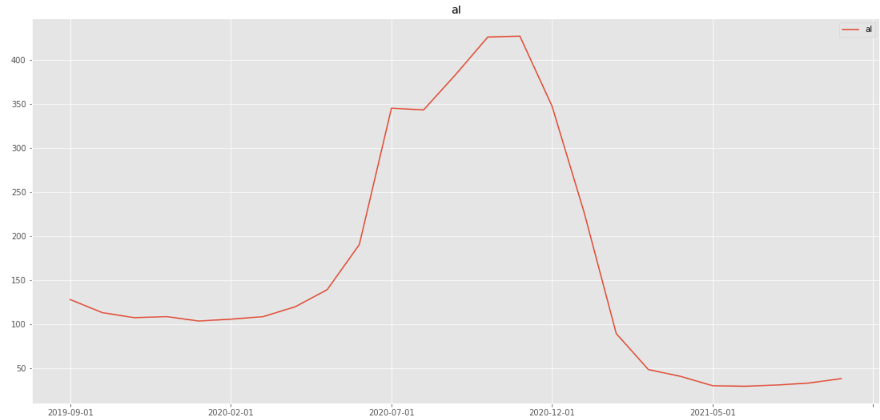
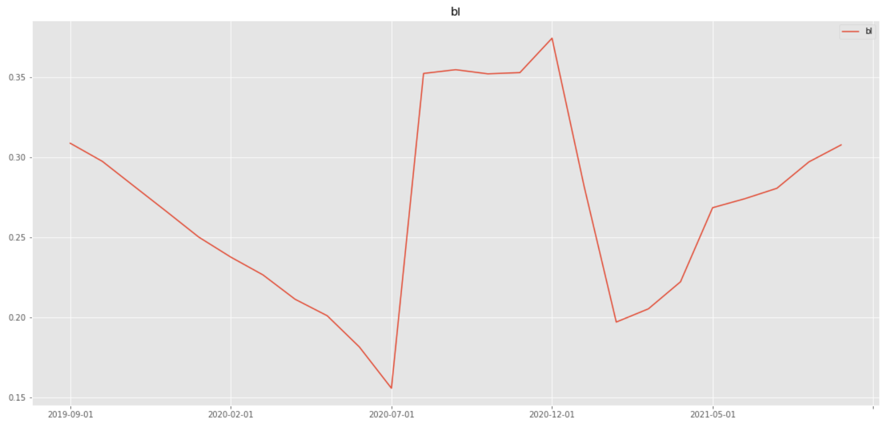


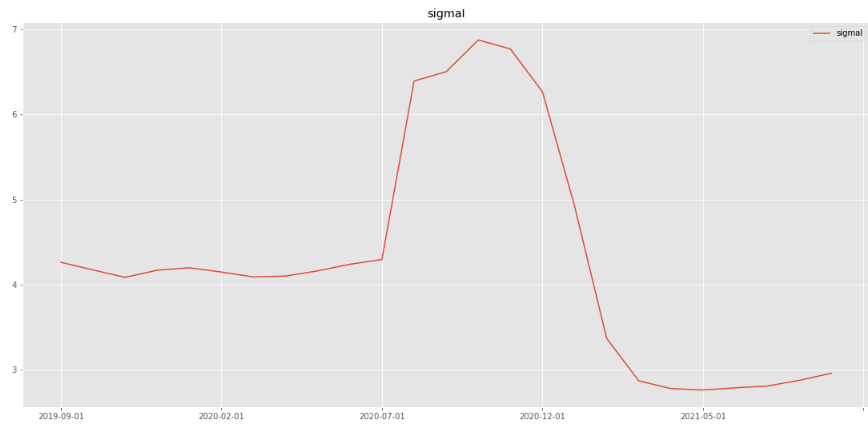
FIGURE 5.6: Relative root mean square error given by our model divided by the relative root mean square error obtained by the Black-Scholes-Merton model, when the proxy of the market interest is the number of unique active addresses.



(A) Estimation of the parameter a_I.

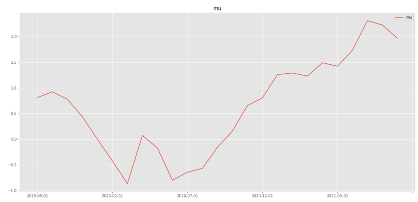


(B) Estimation of the parameter b_I.

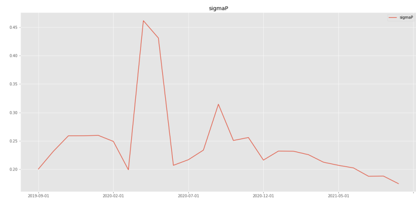


(C) Estimation of the parameter σ_I.

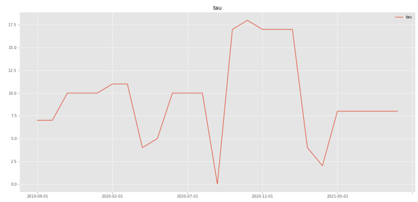
FIGURE 5.7: Estimated values for the parameters of the model defined in Section 5.1 when the proxy of the interest is the number of views in Wikipedia of the keyword "Altcoin" and the interest is a Cox-Ingersoll-Ross process.



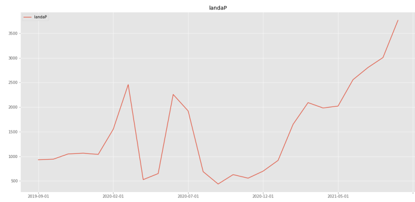
(A) Estimation of the parameter μ.



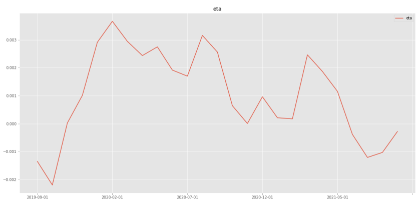
(B) Estimation of the parameter σ_P.



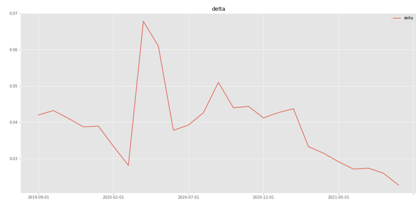
(C) Estimation of the parameter τ.



(D) Estimation of the parameter λ_P.



(E) Estimation of the parameter η.



(F) Estimation of the parameter δ.

FIGURE 5.8: Estimated values for the parameters of the model defined in Section 5.1 when the proxy of the interest is the number of views in Wikipedia of the keyword "Altcoin" and the interest is a Cox-Ingersoll-Ross process.

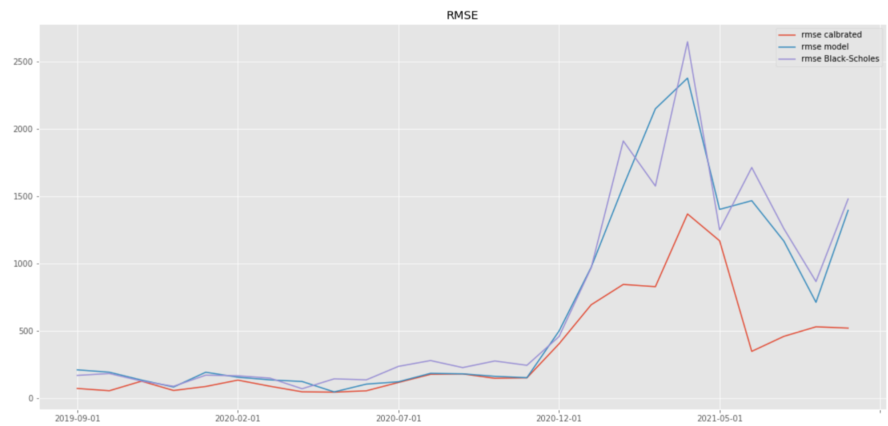


FIGURE 5.9: Root mean square error when the proxy of the interest is the number of views in Wikipedia of the keyword "Altcoin" and the interest is a Cox-Ingersoll-Ross process.

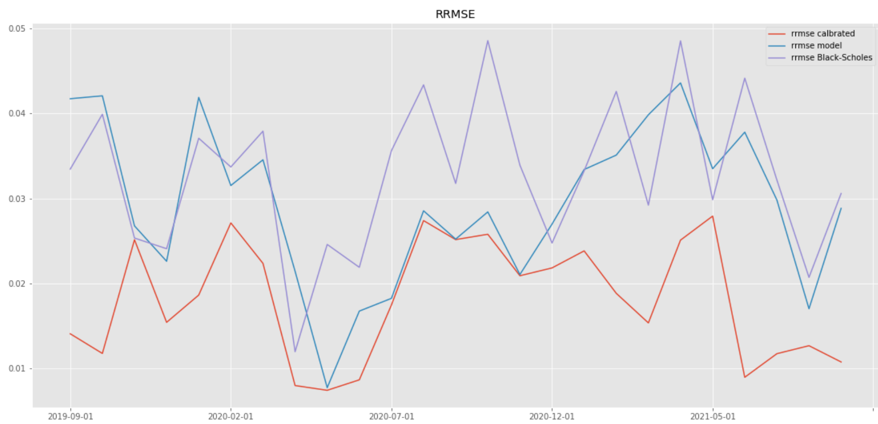
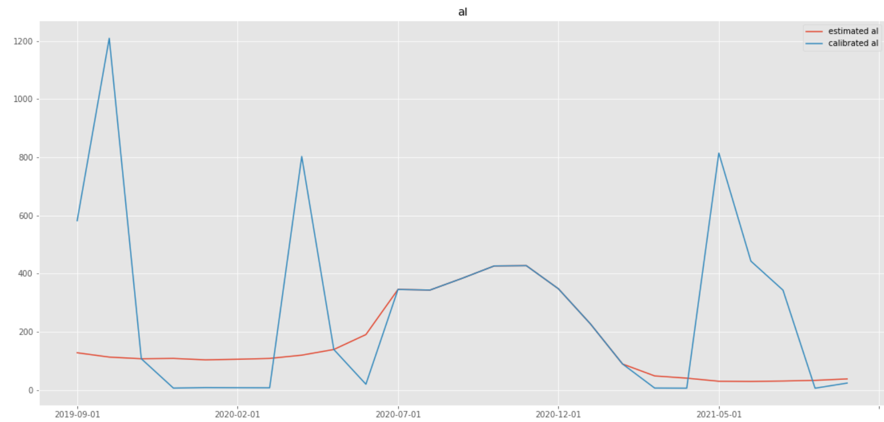
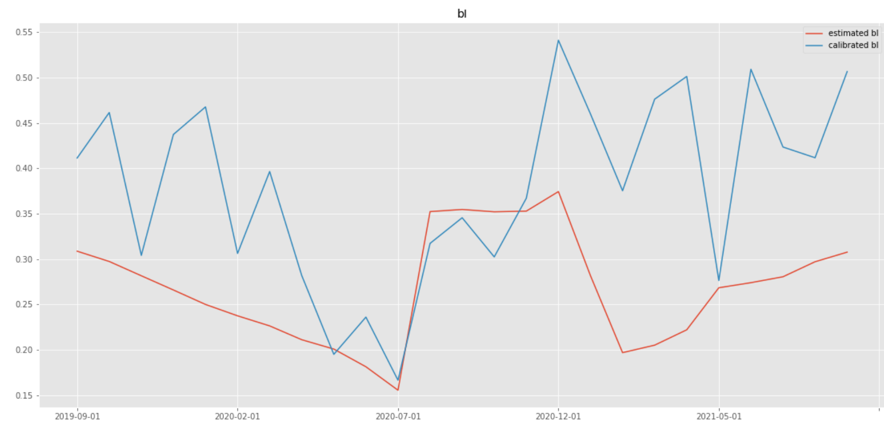


FIGURE 5.10: Relative root mean square error when the proxy of the interest is the number of views in Wikipedia of the keyword "Altcoin" and the interest is a Cox-Ingersoll-Ross process.



(A) Estimated and calibrated values for the parameter a_I.



(B) Estimated and calibrated values for the parameter b_I.

FIGURE 5.11: Calibrated values for the parameters a_I and b_I when the proxy of the interest is the number of views in Wikipedia of the keyword “Altcoin” and the interest is a Cox-Ingersoll-Ross process.

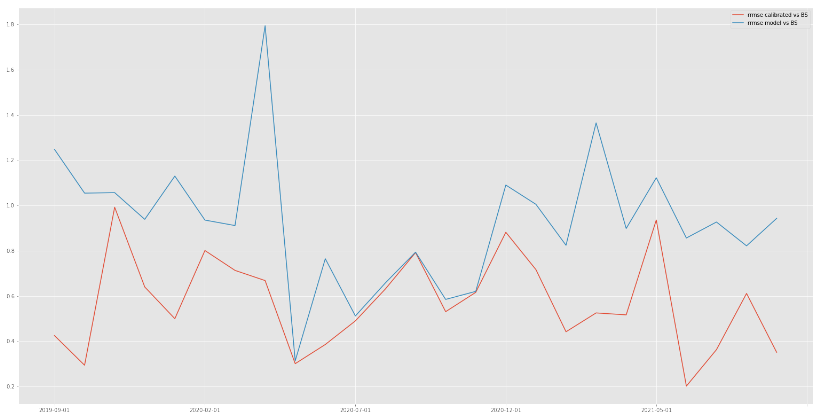
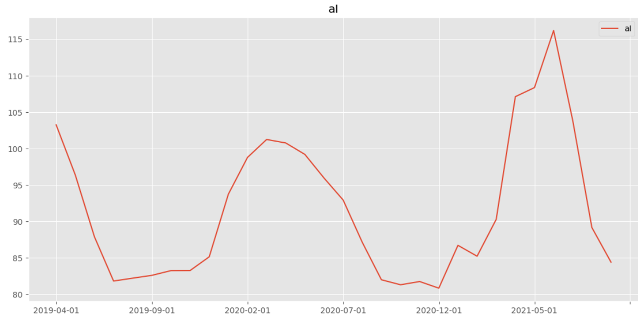
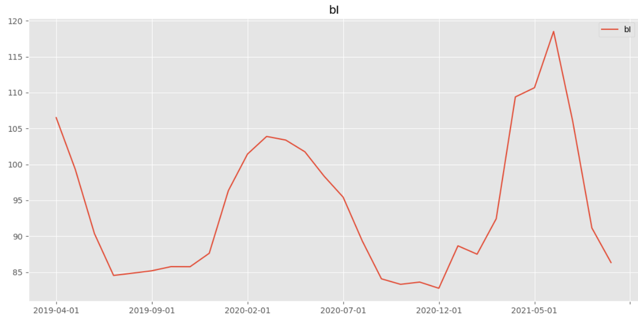


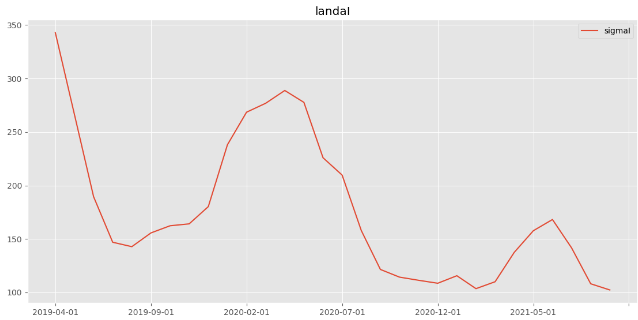
FIGURE 5.12: Relative root mean square error given by our model divided by the relative root mean square error obtained by the Black-Scholes-Merton model, when the proxy for market attention is the number of Wikipedia views of the word "Altcoin".



(A) Estimation of the parameter a_I.

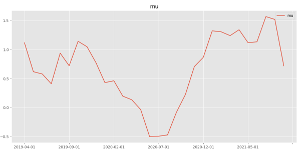


(B) Estimation of the parameter b_I.

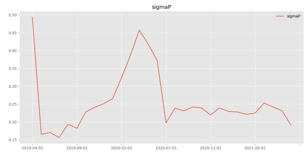


(C) Estimation of the parameter λ_I.

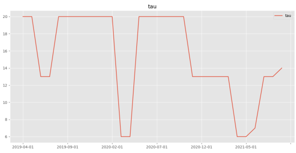
FIGURE 5.13: Estimated values for the parameters of the model defined in Section 5.1 when the proxy of the interest is the logarithm of the number of unique active addresses and the interest is an inverse Gaussian Ornstein-Uhlenbeck process.



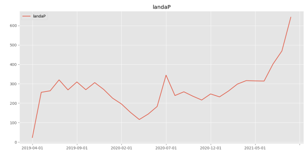
(A) Estimation of the parameter μ.



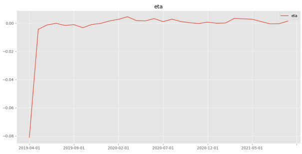
(B) Estimation of the parameter σ_P.



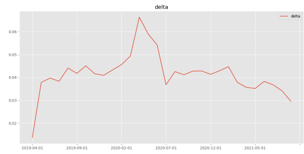
(C) Estimation of the parameter τ.



(D) Estimation of the parameter λ_P.



(E) Estimation of the parameter η.



(F) Estimation of the parameter δ.

FIGURE 5.14: Estimated values for the parameters of the model defined in Section 5.1 when the proxy of the interest is the logarithm of the number of unique active addresses and the interest is an inverse Gaussian Ornstein-Uhlenbeck process.

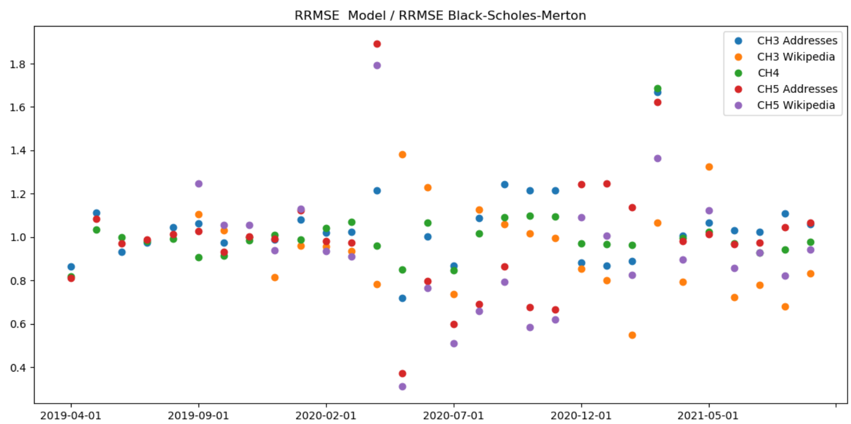


FIGURE 5.15: Relative root mean square errors given by our models divided by the relative root mean square error given by Black-Scholes-Merton model.

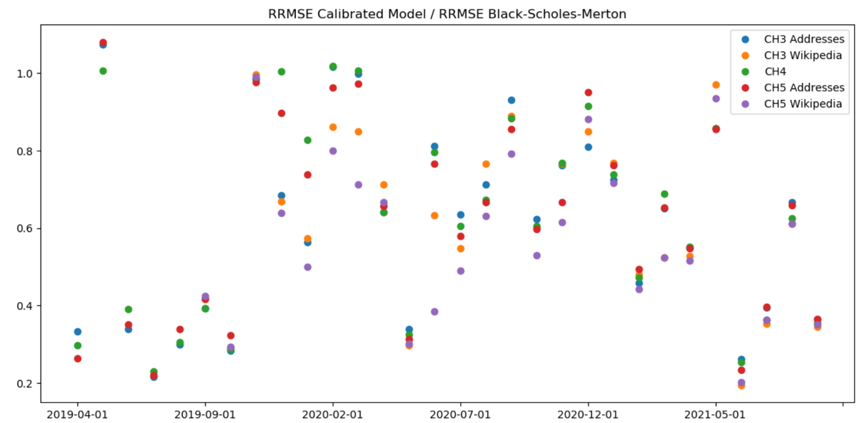


FIGURE 5.16: Relative root mean square errors given by our calibrated models divided by the relative root mean square error given by Black-Scholes-Merton model.

Chapter 6

Stochastic volatility models with an exogenous variable

In previous chapters, we have assumed that price volatility is fully explained by an observed exogenous variable. We have taken this observed variable to be market attention. In addition, we have assumed that the volatility of Bitcoin is proportional to the square root of the market attention. Due to the complex nature of financial markets, it is unlikely that the volatility is truly described by just one variable. Whilst identifying the main factors that affect the volatility of cryptocurrencies is beyond of the scope of this thesis, in this chapter we proposed models in which the volatility is described by an observed exogenous process and an unobserved component that follows a certain latent model. Again, the observed variable will be market attention and we expect that the hidden process will capture the effect of the other variables.

To work with models of this kind, we will need new tools. These new tools are sequential Monte Carlo methods, and they will help us to deal with the unobserved component of the volatility. Sequential Monte Carlo methods have been used previously when working with stochastic volatility models; the works of Al-Saadony (2013) and Yang (2015) contain more information.

6.1 State space models and particle filters

Before studying the models of interest, we will study models in which the volatility is completely explained by a hidden process. These models are usually called stochastic volatility models. But first let us introduce the concept of state space model. The introduction of state space models presented here follows the exposition done in the work of Chopin and Papaspiliopoulos (2020, Chapter 4), including Proposition 6.1.1.

Notation 6.1.1. *Here we will use capital letters to denote random variables and lower-case letters for the realisations of the random variables. In addition we will use the semicolon notation to indicate vectors, that is for $t = 1, 2, \ldots$ we define the vectors*

$$Y_{0:t} = (Y_0, Y_1, \ldots, Y_t), \qquad Y_{1:t} = (Y_1, Y_2, \ldots, Y_t), \tag{6.1.1}$$

$$y_{0:t} = (y_0, y_1, \ldots, y_t), \qquad y_{1:t} = (y_1, y_2, \ldots, y_t), \tag{6.1.2}$$

where the vectors in (6.1.1) are vectors of random variables and the vectors in (6.1.2) are vectors of realizations. In addition, we will write $\mathbb{P}(dx)$ to indicate the law of the random variable X and $\mathbb{P}(dx_{0:t})$ to represent the law of $(X_0, X_1, \ldots, X_t)$.

Remark 6.1.1. *We will consider $\mathcal{X}$ to be a set and $\mathcal{B}(\mathcal{X})$ to be the σ-algebra (usually a Borel σ-algebra) of the set $\mathcal{X}$; in this thesis the set $\mathcal{X}$ is $\mathbb{R}^d$ with $d \geq 1$.*

State space models are related to the concept of probability kernels.

Definition 6.1.1. *Let $(\mathcal{X}, \mathcal{B}(\mathcal{X}))$ and $(\mathcal{Y}, \mathcal{B}(\mathcal{Y}))$ be two measurable spaces. A probability kernel from $(\mathcal{X}, \mathcal{B}(\mathcal{X}))$ to $(\mathcal{Y}, \mathcal{B}(\mathcal{Y}))$, $P(x, dy)$ is a function from $(\mathcal{X}, \mathcal{B}(\mathcal{Y}))$ to $[0, 1]$ such that:*

1. *for every $x \in \mathcal{X}$, $P(x, .)$ is a probability measure on $(\mathcal{Y}, \mathcal{B}(\mathcal{Y}))$,*
2. *for every $A \in \mathcal{B}(\mathcal{Y})$, $x \to P(x, A)$ is a measurable function on $\mathcal{X}$.*

The definition of probability kernels allows us to work with conditional probabilities, and we can use them to define discrete Markov processes. To that end, we consider a sequence of probability kernels $\{P_t\}_{t=1}^T$ from $(\mathcal{X}, \mathcal{B}(\mathcal{X}))$ to $(\mathcal{X}, \mathcal{B}(\mathcal{X}))$, for some $T \in \mathbb{Z}_+$ and a probability measure $\mathbb{P}(dx_0)$ on $(\mathcal{X}, \mathcal{B}(\mathcal{X}))$.

Definition 6.1.2. *A sequence of random variables $X_{0:T}$ with joint distribution given by*

$$\mathbb{P}_T(X_{0:T} \in dx_{0:T}) = \mathbb{P}_0(dx_0) \prod_{s=1}^{T} P_s(x_{s-1}, dx_s) \tag{6.1.3}$$

is called a (discrete-time) Markov process with state space $\mathcal{X}$, initial distribution $\mathbb{P}_0$ for the random variable X_0 and transition kernel at time t, P_t.

It can be shown that equation (6.1.3) implies that

$$\begin{aligned} \mathbb{P}_T(X_t \in dx_t | X_{0:t-1} = x_{0:t-1}) &= \mathbb{P}_T(X_t \in dx_t | X_{t-1} = x_{t-1}) \\ &= P_t(x_{t-1}, dx_t) \text{ for all } t \leq T. \end{aligned}$$

Notice that this definition of Markov process is given in terms of the probability measure $\mathbb{P}_T$, however on some occasions we would like to have a sequence of probability measures $\mathbb{P}_t$ for $t \leq T$.

Proposition 6.1.1. *Consider a sequence of random variables $X_{0:T}$ and a sequence of probability measures $\{\mathbb{P}_t\}_{t=0}^T$, defined as*

$$\mathbb{P}_t(X_{0:t} \in dx_{0:t}) = \mathbb{P}_0(dx_0) \prod_{s=1}^{t} P_s(x_{s-1}, dx_s),$$

where P_s is a probability kernel for $s = 1, 2, \ldots, T$. Then for any $t \leq T$,

$$\mathbb{P}_T(dx_{0:t}) = \mathbb{P}_t(dx_{0:t}).$$

Proposition 6.1.1 tells us that for every bounded measurable function $\varphi : \mathcal{X}^{t+1} \to \mathbb{R}$ we have that

$$E_{\mathbb{P}_T}[\varphi(X_{0:t})] = E_{\mathbb{P}_t}[\varphi(X_{0:t})],$$

where $E_{\mathbb{P}_T}[.]$ and $E_{\mathbb{P}_t}[.]$ are the expected values with respect to the measures $\mathbb{P}_T$ and $\mathbb{P}_t$, respectively. We have just seen that we can use probability kernels to define Markov processes, but we can go further and use them to define state space models.

Definition 6.1.3. *Let $X = \{X\}_{t=0}^T$ and $Y = \{Y\}_{t=0}^T$ be two stochastic processes such that $X_t \in \mathcal{X}$ and $Y_t \in \mathcal{Y}$, a state space model is the stochastic process $\{(X_t, Y_t)\}_{t=0}^T$ on the measurable space $(\mathcal{X} \times \mathcal{Y}, \mathcal{B}(\mathcal{X}) \times \mathcal{B}(\mathcal{Y}))$, where $\mathcal{B}(\mathcal{X}) \times \mathcal{B}(\mathcal{Y})$ is the product σ-algebra. The distribution of the state space model is defined in terms of an initial distribution $\mathbb{P}_0(dx_0)$ of the random variable X_0 and two sequences of probability kernels $\{P_t(x_{t-1}, dx_t)\}_{t=1}^T$, $\{F_t(x_t, dy_t)\}_{t=0}^T$ where $P_t(x_{t-1}, dx_t)$ is a probability kernel* *from* $(\mathcal{X}, \mathcal{B}(\mathcal{X}))$ *to* $(\mathcal{X}, \mathcal{B}(\mathcal{X}))$ *and* $F_t(x_t, dy_t)$ *is a probability kernel from* $(\mathcal{X}, \mathcal{B}(\mathcal{X}))$ *to* $(\mathcal{Y}, \mathcal{B}(\mathcal{Y}))$ *and the joint distribution of* $(X_{0:T}, Y_{0:T})$ *is defined by*

$$\begin{aligned}
\mathbb{P}_T(X_{0:T} \in dx_{0:T}, Y_{0:T} \in dy_{0:T}) &= \mathbb{P}_0(dx_0) \prod_{t=1}^{T} P_t(x_{t-1}, dx_t) \prod_{t=0}^{T} F_t(x_t, dy_t) \\
&= \mathbb{P}_0(dx_0) F_0(x_0, dy_0) \\
&\quad \prod_{t=1}^{T} P_t(x_{t-1}, dx_t) F_t(x_t, dy_t) \qquad (6.1.4) \\
&= \mathbb{P}_T(dx_{0:T}) \prod_{t=0}^{T} F_t(x_t, dy_t). \qquad (6.1.5)
\end{aligned}$$

From equation (6.1.4) we have that the process $\{(X_t, Y_t)\}_{t=1}^{T}$ is a Markov process with initial distribution $\mathbb{P}_0(dx_0)F_0(x_0, dy_0)$ for the random variable (X_0, Y_0) and transition kernels $\{P_t(x_{t-1}, dx_t)F_t(x_t, dy_t)\}_{t=1}^{T}$. From equation (6.1.5) we have that the process X is a Markov process with initial distribution $\mathbb{P}_0(dx_0)$ for the random variable X_0 and transition kernels $\{P_t(x_{t-1}, dx_t)\}_{t=1}^{T}$.

These types of models are used to model time series that have been generated by the process Y. That is, only the realizations of the process Y are observed, while the realizations of the Markov process X remain unobserved. When the set $\mathcal{X}$ is finite, state space models are also called hidden Markov processes.

Remark 6.1.2. *In this thesis, we are interested in an special case of state space models, in which there are two measures* μ *and* ν *on the measurable spaces* $(\mathcal{X}, \mathcal{B}(\mathcal{X}))$ *and* $(\mathcal{Y}, \mathcal{B}(\mathcal{Y}))$*, respectively such that:*

$$\begin{aligned}
\mathbb{P}_0(dx_0) &= f_0^X(x_0)\mu(dx_0), \\
P_t(x_{t-1}|dx_t) &= f_{t|t-1}^{X|X}(x_t|x_{t-1})\mu(dx_t), \\
F_t(x_t|dy_t) &= f_{t|t}^{Y|X}(y_t|x_t)\nu(dy_t),
\end{aligned}$$

where f_0^X *represents the marginal density of* X_0*,* $f_{t|t-1}^{X|X}$ *is the conditional density of* X_t *given* X_{t-1} *and* $f_{t|t}^{Y|X}$ *is the density of* Y_t *given* X_t*. Usually the measures* μ *and* ν *are the Lebesgue measure. That is, we focus on state space models whose probability kernels satisfy continuous distributions.*

If we assume that

$$F_t(x_t|dy_t) = f_{t|t}^{Y|X}(y_t|x_t)\nu(dy_t) \quad \forall t,$$

then the equation (6.1.5) can be rewritten as

$$\mathbb{P}_T(X_{0:T} \in dx_{0:T}, Y_{0:T} \in dy_{0:T}) = \mathbb{P}_T(dx_{0:T}) \prod_{t=0}^{T} f_{t|t}^{Y|X}(y_t|x_t)\nu(dy_t). \qquad (6.1.6)$$

On some occasions we will be interested in computing the marginal distribution of $Y_{0:t}$, which can be computed as

$$\mathbb{P}_T(Y_{0:t} \in dy_{0:t}) = E_{\mathbb{P}_t}\left[\prod_{s=0}^{t} f_{s|s}^{Y|X}(y_s|X_s)\right] \prod_{s=0}^{t} \nu(dy_s), \qquad (6.1.7)$$

where $E_{\mathbb{P}_t}[.]$ is the expected value with respect to the probability measure $\mathbb{P}_t$. From equation (6.1.7) we have that the marginal density of $Y_{0:t}$ is equal to

$$f^Y_{0:t}(y_{0:t}) = E_{\mathbb{P}_t}\left[\prod_{s=0}^{t} f^{Y|X}_{s|s}(y_s|X_s)\right]. \tag{6.1.8}$$

Using the result in (6.1.8) we can construct the marginal densities $f^Y_{0:t}$ for different values of t. In addition, the conditional density of Y_t given $Y_{0:t-1}$ can be defined in terms of two consecutive marginal densities, that is

$$f^{Y|Y}_{t|0:t-1}(y_t|y_{0:t-1}) = \frac{f^Y_{0:t}(y_{0:t})}{f^Y_{0:t-1}(y_{0:t-1})}. \tag{6.1.9}$$

The ratios defined in equation (6.1.9) are called likelihood ratios since the likelihood can be defined in terms of them as

$$f^Y_{0:T}(y_{0:T}) = f^Y_0(y_0)\prod_{s=1}^{T} f^{Y|Y}_{s|0:s-1}(y_s|y_{0:s-1}). \tag{6.1.10}$$

Remark 6.1.3. *Usually, state space models will depend on a vector of parameters θ that can take values in a set Θ. When we want to indicate that our model depends on a vector of parameters θ we will use the notation $f^Y_{0:T}(y_{0:T}|\theta)$ rather than $f^Y_{0:T}(y_{0:T})$. Using this notation, the likelihood in (6.1.10) can be rewritten as:*

$$L^Y(\theta) = f^Y_{0:T}(y_{0:T}|\theta) = f^Y_0(y_0|\theta)\prod_{s=1}^{T} f^Y_{s|s-1}(y_s|y_{s-1},\theta). \tag{6.1.11}$$

The log-likelihood is then

$$l^Y(\theta) = \log\left(L^Y(\theta)\right) = \log\left(f^Y_0(y_0|\theta)\right) + \sum_{s=1}^{T}\log\left(f^Y_{s|s-1}(y_s|y_{s-1},\theta)\right). \tag{6.1.12}$$

For estimating the parameter θ, one can try to maximize the function l^Y, but the problem is that in general, state space models do not have a close formula for the likelihood.

In addition, to the likelihood, we are also concerned with the distribution of $X_{0:t}$ given $Y_{0:t}$, which can be expressed as

$$\mathbb{P}_t\left(X_{0:t} \in dx_{0:t}|Y_{0:t} = y_{0:t}\right) = \frac{1}{f^Y_{0:t}(y_{0:t})}\left[\prod_{s=0}^{t} f^{Y|X}_{s|s}(y_s|x_s)\right]\mathbb{P}_t(dx_{0:t}). \tag{6.1.13}$$

In this thesis, we are interested in determining the distribution of X_t given $Y_{0:t}$; this is known as filtering.

6.1.1 Examples

In this subsection, we will see two examples of stochastic volatility models and their representation as state space models.

Remark 6.1.4. *In all the models to be studied here, the density functions $f^{Y|X}_{t|t}$ and $f^{X|X}_{t|t-1}$ are going to be normal density functions. Because of this, we introduce the* *following notation to indicate the density function of a normal random variable with mean μ and variance σ^2:*

$$f_N(z|\mu,\sigma^2) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\frac{(z-\mu)^2}{\sigma^2}} \quad \text{for all } z \in \mathbb{R}. \tag{6.1.14}$$

Example 6.1.1. *Simple stochastic volatility model.*

This is one of the first stochastic volatility models and it was proposed by Taylor (1982). Let $\{Y_t\}_{t=0}^{\infty}$ and $\{X_t\}_{t=0}^{\infty}$ represent the logarithmic returns of the price and the logarithmic volatility squared, respectively. This model is written as

$$Y_t = \exp\left\{\frac{X_t}{2}\right\}\epsilon_t, \tag{6.1.15}$$

$$X_t = \rho_0 + \rho_1 X_{t-1} + \tau\eta_t, \tag{6.1.16}$$

where $\rho_0 \in \mathbb{R}$, $\rho_1 \in (-1,1)$, $\tau > 0$ and where $\{\eta_t\}_{t=1}^{\infty}$ and $\{\epsilon_t\}_{t=0}^{\infty}$ are two independent sequences of independent standard normal random variables. We assume that the initial value of the volatility X_0 follows a normal distribution with mean μ_0 and variance σ_0^2. Notice that the probability kernels of this state space model are

$$\begin{aligned} F_t(x_t, dy_t) &= f_N(y_t|0, e^{x_t})m(dy_t), \\ P_t(x_{t-1}, dx_t) &= f_N(x_t|\rho_0 + \rho_1 x_{t-1}, \tau^2)m(dx_t), \end{aligned}$$

where m is the Lebesgue measure.

Example 6.1.2. *Heston model.*

The Heston model (Heston, 1993) is one of the most popular stochastic volatility models and it is used in the context of option pricing. This model is a continuous time model and because of that we have to discretize it. We use the discretization used by Mrázek and Pospíšil (2017, p. 694). In this case, $\{Y_t\}_{t=1}^{\infty}$ and $\{V_t\}_{t=0}^{\infty}$ represent the logarithmic returns of the price and the volatility, respectively. This model can be written as

$$Y_t = \left(\mu - \frac{1}{2}V_{t-1}^{+}\right)\Delta t + \sqrt{V_{t-1}^{+}\Delta t}Z_t^Y, \tag{6.1.17}$$

$$V_t = V_{t-1} + a_I(b_I - V_{t-1}^{+})\Delta t + \sigma_I\sqrt{V_{t-1}^{+}\Delta t}Z_t^V, \tag{6.1.18}$$

where $x^+ = \max\{x, 0\}$, $\mu \in \mathbb{R}$, $a_I, b_I, \sigma_I, \Delta t > 0$, and where $\frac{2a_I b_I}{\sigma_I^2} \geq 1$. In addition, $\{Z_t^Y\}_{t=1}^{\infty}$ and $\{Z_t^V\}_{t=1}^{\infty}$ are two correlated sequences of independent standard normal random variables with correlation corr$[Z_t^Y, Z_t^V] = \rho \in (-1,1)$ for all $t \geq 0$. Here Δt is not a parameter and it represents the time step. Again, the distribution of V_0 is a normal distribution with mean μ_0 and variance σ_0^2.

The correlated sequences $\{Z_t^V\}_{t=1}^{\infty}$ and $\{Z_t^Y\}_{t=1}^{\infty}$ can be expressed in terms of two independent sequences of independent standard normal random variables $\{B_t^V\}_{t=1}^{\infty}$ and $\{B_t^Y\}_{t=1}^{\infty}$ as:

$$\begin{aligned} Z_t^Y &= \rho B_t^V + \sqrt{1-\rho^2}B_t^Y \\ Z_t^V &= B_t^V, \end{aligned}$$

(Hirsa, 2012, p. 233). So, the model introduced in equations (6.1.17)–(6.1.18) can be rewritten as

$$Y_t = \left(\mu - \frac{1}{2}V_{t-1}^+\right)\Delta t + \sqrt{V_{t-1}^+\Delta t}\left(\rho B_t^V + \sqrt{1-\rho^2}B_t^Y\right) \quad (6.1.19)$$

$$V_t = V_{t-1} + a_I(b_I - V_{t-1}^+)\Delta t + \sigma_I\sqrt{V_{t-1}^+\Delta t}B_t^V. \quad (6.1.20)$$

From equation (6.1.20) we have that

$$\sqrt{V_{t-1}^+\Delta t}B_t^V = \frac{1}{\sigma_I}\left(V_t - V_{t-1} - a_I(b_I - V_{t-1}^+)\Delta t\right). \quad (6.1.21)$$

Substituting equation (6.1.21) into equation (6.1.19) we arrive at

$$Y_t = \left(\mu - \frac{1}{2}V_{t-1}^+\right)\Delta t + \frac{\rho}{\sigma_I}\left(V_t - V_{t-1} - a_I(b_I - V_{t-1}^+)\Delta t\right) + \sqrt{1-\rho^2}\sqrt{V_{t-1}^+\Delta t}B_t^Y. \quad (6.1.22)$$

In this case, the hidden process has a dimension of two. We define the process $\{X_t\}_{t=1}^\infty$ as

$$X_t = \begin{pmatrix} X_t(1) \\ X_t(2) \end{pmatrix} = \begin{pmatrix} V_t \\ V_{t-1} \end{pmatrix}.$$

So equations (6.1.20) and (6.1.22) can be rewritten in terms of X:

$$\begin{aligned} Y_t &= \left(\mu - \frac{1}{2}X_t^+(2)\right)\Delta t + \frac{\rho}{\sigma_I}\left(X_t(1) - X_t(2) - a_I(b_I - X_t^+(2))\Delta t\right) \\ &\quad + \sqrt{1-\rho^2}\sqrt{X_t^+(2)\Delta t}B_t^Y, \\ \begin{pmatrix} X_t(1) \\ X_t(2) \end{pmatrix} &= \begin{pmatrix} X_{t-1}(1) + a_I(b_I - X_{t-1}^+(1))\Delta t + \sigma_I\sqrt{X_{t-1}^+(1)\Delta t}B_t^V \\ X_{t-1}(1) \end{pmatrix}. \end{aligned}$$

So, we have just rewritten our model in terms of a Markov process X and an observed process Y, whose value at time t depends on X_t. That is, we have expressed our model as a state space model. In this case, the probability kernels are

$$\begin{aligned} F_t(x_t, dy_t) &= f_N\left(y_t \,|\mu_F, \sigma_F^2\right)m(dy_t), \\ P_t(x_{t-1}, dx_t) &= f_N(x_t(1)|\mu_P, \sigma_P^2)\delta_{x_{t-1}(1)}(dx_t(2))m(dx_t(1)), \end{aligned}$$

where m is the Lebesgue measure, $\delta_{x_{t-1}(1)}$ is the Dirac measure concentrated on $x_{t-1}(1)$ and

$$\begin{aligned} \mu_F &= (\mu - 0.5x_t^+(2))\Delta t + \frac{\rho}{\sigma_I}\left(x_t(1) - x_t(2) - a_I(b_I - x_t^+(2))\Delta t\right), \\ \sigma_F^2 &= (1-\rho^2)x_t^+(2)\Delta t, \\ \mu_P &= x_{t-1}(1) + a_I(b_I - x_{t-1}^+(1))\Delta t, \\ \sigma_P^2 &= \sigma_I^2 x_{t-1}^+(1)\Delta t. \end{aligned}$$

6.1.2 Particle filter

Particle filters are filtering algorithms for state space models. There are different types of particle filters: auxiliary filters, guided filters or bootstrap filters. Remember that filtering consists of deriving the distribution of X_t conditional on $Y_{1:t} = y_{1:t}$, for $t = 1, \dots, T$. Particle filter algorithms approximate the distribution $\mathbb{P}_t(X_t \in dx_t | Y_{0:t} = y_{0:t})$ sequentially.

Particle filters are related to Monte Carlo methods and they are included in a set of techniques called sequential Monte Carlo methods.

6.1.3 Importance sampling

Particle filters are based on the importance sampling algorithm. This algorithm works as follows.

Let us imagine that we have a random variable X that takes values in the set $\mathcal{X}$, and the distribution of this random variable is denoted by $\mathbb{Q}$. We are interested in computing the expected value

$$E_{\mathbb{Q}}\left[\varphi(X)\right] := \int_{\mathcal{X}} \varphi(x)\mathbb{Q}(dx), \tag{6.1.23}$$

where $\varphi : \mathcal{X} \to \mathbb{R}$ is a bounded measurable function and $E_{\mathbb{Q}}[.]$ is the expected value with respect to the measure $\mathbb{Q}$.

This integral can be approximated by sampling from the distribution $\mathbb{Q}$. On some occasions we will not be able to draw samples from $\mathbb{Q}$. Let us consider a new distribution $\mathbb{M}$, which we know how to sample from, and assume that $\mathbb{Q}$ is absolutely continuous with respect to $\mathbb{M}$. By the Radon-Nikodym theorem, we know that there exists a non-negative measurable function w such that

$$\mathbb{Q}(dx) = w(x)\mathbb{M}(dx). \tag{6.1.24}$$

Using equation (6.1.24), we can rewrite the expected value in (6.1.23) as

$$\begin{aligned} E_{\mathbb{Q}}\left[\varphi(X)\right] &= \int_{\mathcal{X}} \varphi(x)w(x)\mathbb{M}(dx) \\ &= \int_{\mathcal{X}} \varphi(x)\frac{w(x)}{E_{\mathbb{M}}[w(X)]}\mathbb{M}(dx), \end{aligned} \tag{6.1.25}$$

where the last equality comes from the fact that:

$$E_{\mathbb{M}}[w(X)] := \int_{\mathcal{X}} w(x)\mathbb{M}(dx) = \int_{\mathcal{X}} \mathbb{Q}(dx) = 1.$$

Now we can approximate the expectation in equation (6.1.25) by generating samples from the distribution $\mathbb{M}$. If we generate a sequence $\{X^i\}_{i=1}^{N_{par}}$ of $N_{par} \in \mathbb{N}$ independent and identically distributed random variables with distribution $\mathbb{M}$, then the distribution $\mathbb{Q}$ can be approximated by the distribution

$$\mathbb{Q}^{N_{par}}(dx) := \sum_{i=1}^{N_{par}} W^i \delta_{X^i}(dx),$$

where

$$W^i = \frac{w(X^i)}{\sum_{j=1}^{N_{par}} w(X^j)} \quad \text{for } i = 1, \dots, N_{par},$$

and the expected value (6.1.23) can be approximated by

$$E_{\mathbb{Q}^{N_{par}}}[\varphi(X^{1:N_{par}})] = \sum_{i=1}^{N_{par}} \varphi(X^i)W^i,$$

(Chopin & Papaspiliopoulos, 2020, p. 89).

6.1.4 Feynman-Kac models and sequential importance sampling

Particle filter algorithms are based on the importance sampling algorithm, but the samples are generated sequentially. To construct this type of algorithm, we need to introduce the concept of a Feynman-Kac model.

Definition 6.1.4. *(Chopin & Papaspiliopoulos, 2020, p. 51)*
Let us consider a Markov process $\{X_t\}_{t=0}^T$ such that $X_t \in \mathcal{X}$ for all $0 \leq t \leq T$ with initial distribution $\mathbb{M}_0$ for the random variable X_0 and transition kernels $M_{1:T}$, that satisfies:

$$\mathbb{M}_T(dx_{0:T}) = \mathbb{M}_0(dx_0) \prod_{t=1}^{T} M_t(x_{t-1}, dx_t).$$

Consider also a sequence of functions $G_0 : \mathcal{X} \to \mathbb{R}^+$, and $G_t : \mathcal{X}^2 \to \mathbb{R}^+$ for $1 \leq t \leq T$ called potential functions. A sequence of Feynman-Kac models is given by a sequence of probability measures on $\left(\mathcal{X}^{t+1}, \mathcal{B}(\mathcal{X})^{t+1}\right)$ for $0 \leq t \leq T$, defined as the following changes of measure from $\mathbb{M}_t$:

$$\mathbb{Q}_t(dx_{0:t}) = \frac{1}{L_t} G_0(x_0) \left[\prod_{s=1}^{t} G_s(x_{s-1}, x_s)\right] \mathbb{M}_t(dx_{0:t}), \tag{6.1.26}$$

where L_t is the normalising constant needed for $\mathbb{Q}_t$ to be a probability measure,

$$L_t = \int_{\mathcal{X}^{t+1}} G_0(x_0) \prod_{s=1}^{t} G_s(x_{s-1}, x_s) \mathbb{M}_t(dx_{0:t}) = E_{\mathbb{M}_t}\left[G_0(X_0) \prod_{s=1}^{t} G_s(X_{s-1}, X_s)\right].$$

For these models to be well defined, we require $0 < L_t < \infty$ for all $t \leq T$.

The normalizing constant L_t is called the partition function or likelihood function and we denote the ratios for the successive normalizing constants as

$$l_t = \frac{L_t}{L_{t-1}}.$$

The Feynman-Kac model defined in equation (6.1.26) is just a change of measure from the measure $\mathbb{M}_t(dx_{0:t})$ to the measure $\mathbb{Q}_t(dx_{0:t})$. From equation (6.1.26) we have that if we know how to sample from the distribution $\mathbb{M}_t(dx_{0:t})$, then we can approximate the distribution $\mathbb{Q}_t(dx_{0:t})$ using the importance sampling algorithm. Based on this, it is possible to construct the sequential importance sampling algorithm, see Algorithm 1 (Chopin & Papaspiliopoulos, 2020, p. 132).

Algorithm 1: Sequential importance sampling

$X_0^i \sim \mathbb{M}_0(dx_0)$ for $i = 1, \ldots, N_{par}$;
$w_0^i \leftarrow G_0\left(X_0^i\right)$ for $i = 1, \ldots, N_{par}$;
$W_0^i \leftarrow \frac{w_0^i}{\sum_{m=1}^{N_{par}} w_0^m}$ for $i = 1, \ldots, N_{par}$;
for $t = 1, 2, \ldots T$ **do**
 $X_t^i \sim M_t(X_{t-1}^i, dx_t)$ for $i = 1, \ldots, N_{par}$;
 $w_t^i \leftarrow w_{t-1}^i G_t\left(X_{t-1}^i, X_t^i\right)$ for $i = 1, \ldots, N_{par}$;
 $W_t^i = \frac{w_t^i}{\sum_{m=1}^{N_{par}} w_t^m}$ for $i = 1, \ldots, N_{par}$;
end

Algorithm 1 allow us to approximate $\mathbb{Q}_t(dx_t)$ by

$$\mathbb{Q}_t^{N_{par}}(dx_t) = \sum_{i=1}^{N_{par}} W_t^i \delta_{X_t^i}(dx_t),$$

and the ratio l_t is approximated by

$$l_t^{N_{par}} = \frac{\sum_{i=1}^{N_{par}} w_t^i}{\sum_{i=1}^{N_{par}} w_{t-1}^i}.$$

The problem with this algorithm is that some of the weights W_t^i that appear in Algorithm 1 can become zero or near to zero. This is called particle degeneracy. Because of the recursive nature of the algorithm, we can lose several particles when T is high. To avoid this, we can perform a re-sampling with replacement at each iteration of the algorithm. The problem of re-sampling at each step is that it increases the computing time of the algorithm. Because of that, it is defined the effective sample size for the weights $W_t^{1:N_{par}}$ as

$$ESS\left(W_t^{1:N_{par}}\right) = \frac{1}{\sum_{m=1}^{N_{par}} (W_t^m)^2}. \tag{6.1.27}$$

Notice that when $W_t^i = \frac{1}{N_{par}}$ for all $i = 1, \ldots, N_{par}$ we have that $ESS\left(W_t^{1:N_{par}}\right) = N_{par}$ meaning that all particles in the sample are contributing equally. In the case when all the weights $W_t^{1:N_{par}}$ are equal to zero except one which is equal to one, we have that $ESS\left(W_t^{1:N_{par}}\right) = 1$. A common policy is to resample when the effective sample size is smaller than a certain value that we call N_{tol}. That is, a resampling will occur at time t if

$$ESS\left(W_t^{1:N_{par}}\right) < N_{tol}.$$

A usual value for the tolerance is $N_{tol} = N_{par}/2$.

Using the concept of effective sample size, it is possible to construct a new algorithm based on Algorithm 1, called sequential importance sampling with adaptive resampling. This appears in Algorithm 2 below (Chopin & Papaspiliopoulos, 2020, p. 134).

Algorithm 2: Sequential importance sampling with adaptive resampling

$X_0^i \sim \mathbb{M}_0(dx_0)$ for $i = 1, \ldots, N_{par}$;
$w_0^i \leftarrow G_0\left(X_0^i\right)$ for $i = 1, \ldots, N_{par}$;
$W_0^i \leftarrow \frac{w_0^i}{\sum_{m=1}^{N_{par}} w_0^m}$ for $i = 1, \ldots, N_{par}$;
for $t = 1, 2, \ldots T$ **do**
    if $ESS\left(W_{t-1}^{1:N_{par}}\right) < N_{tol}$ **then**
        Draw (with replacement) N_{par} indices I_t^i for $i = 1, \ldots, N_{par}$ using the normalized weights $W_{t-1}^{1:N_{par}}$;
        $\hat{w}_{t-1}^i \leftarrow 1$ for $i = 1, \ldots, N_{par}$;
    else
        $I_t^i \leftarrow i$ for $i = 1, \ldots, N_{par}$;
        $\hat{w}_{t-1}^i \leftarrow w_{t-1}^i$ for $i = 1, \ldots, N_{par}$;
    end
    $X_t^i \sim M_t(X_{t-1}^{I_t^i}, dx_t)$ for $i = 1, \ldots, N_{par}$;
    $w_t^i \leftarrow \hat{w}_{t-1}^i G_t\left(X_{t-1}^{I_t^i}, X_t^i\right)$ for $i = 1, \ldots, N_{par}$;
    $W_t^i = \frac{w_t^i}{\sum_{m=1}^{N_{par}} w_t^m}$ for $i = 1, \ldots, N_{par}$;
end

From Algorithm 2 we have that the distribution $\mathbb{Q}_t(dx_t)$ is approximated by

$$\mathbb{Q}_t^{N_{par}}(dx_t) = \sum_{i=1}^{N_{par}} W_t^i \delta_{X_t^i}(dx_t), \tag{6.1.28}$$

and the ratio l_t is approximated by

$$l_t^{N_{par}} = \begin{cases} \frac{1}{N_{par}} \sum_{i=1}^{N_{par}} w_t^i, & \text{if resampling occurred at time } t, \\ \frac{\sum_{i=1}^{N_{par}} w_t^i}{\sum_{i=1}^{N_{par}} w_{t-1}^i}, & \text{otherwise.} \end{cases} \tag{6.1.29}$$

6.1.5 Bootstrap filter

Returning to the definition of state space models, consider a state space model with initial distribution $\mathbb{P}_0(dx_0)$, probability kernels $\{P_t(x_{t-1}, dx_t)\}_{t=1}^{\infty}$ for the hidden process X and a sequence of probability kernels $\{F_t(x_t, dy_t)\}_{t=o}^{\infty}$ for the observed process Y. In this case, we assume that the probability kernels $\{F_t(x_t, dy_t)\}_{t=0}^{\infty}$ satisfy

$$F_t(x_t, dy_t) = f_{t|t}^{Y|X}(y_t|x_t)m(dy_t) \quad \forall t \geq 0,$$

where m is Lebesgue measure.

Using Feynman-Kac models defined in Section 6.1.4 we can construct particle filter algorithms for state space models. Depending on how we define the components of the Feynman-Kac models, we can build different types of filtering algorithms. In this thesis we focus only on the bootstrap algorithm, but there are other types of filtering algorithms such as the guided particle filter or the auxiliary particle filter.

Definition 6.1.5. *(Chopin & Papaspiliopoulos, 2020, p. 53) The bootstrap Feynman-Kac formalism is the Feynman-Kac model with the following components*

$$\mathbb{M}_0(dx_0) = \mathbb{P}(dx_0), \quad G_0(x_0) = f_{0|0}^{Y|X}(y_0|x_0),$$
$$M_t(x_{t-1}, dx_t) = P_t(x_{t-1}, dx_t), \quad G_t(x_{t-1}, x_t) = f_{t|t}^{Y|X}(y_t|x_t).$$

Under the bootstrap Feynman-Kac formalism, we have that

$$\begin{aligned} \mathbb{Q}_t(dx_{0:t}) &= \mathbb{P}_t(X_{0:t} \in dx_{0:t}|Y_{0:t} = y_{0:t}), \\ L_t &= f_{0:T}^{Y}(y_{0:T}), \\ l_t &= \frac{L_t}{L_{t-1}} = f_{t|0:t-1}^{Y|Y}(y_t|y_{0:t-1}). \end{aligned}$$

From Definition 6.1.5 and Algorithm 2 the bootstrap filter algorithm is constructed as shown in Algorithm 3 (Chopin & Papaspiliopoulos, 2020, p. 136).

Algorithm 3: Bootstrap filter

$X_0^i \sim \mathbb{P}_0(dx_0)$ for $i = 1, \ldots, N_{par}$;
$w_0^i \leftarrow f_{0|0}^{Y|X}(y_0|X_0^i)$ for $i = 1, \ldots, N_{par}$;
$W_0^i \leftarrow \frac{w_0^i}{\sum_{m=1}^{N_{par}} w_0^m}$ for $i = 1, \ldots, N_{par}$;
for $t = 1, 2, \ldots T$ **do**
  if $ESS\left(W_{t-1}^{1:N_{par}}\right) < N_{tol}$ **then**
    Draw (with replacement) N_{par} indices I_t^i for $i = 1, \ldots, N_{par}$ using the normalized weights $W_{t-1}^{1:N_{par}}$;
    $\hat{w}_{t-1}^i \leftarrow 1$ for $i = 1, \ldots, N_{par}$;
  else
    $I_t^i \leftarrow i$ for $i = 1, \ldots, N_{par}$;
    $\hat{w}_{t-1}^i \leftarrow w_{t-1}^i$ for $i = 1, \ldots, N_{par}$;
  end
  $X_t^i \sim P_t\left(X_{t-1}^{I_t^i}, dx_t\right)$ for $i = 1, \ldots, N_{par}$;
  $w_t^i \leftarrow \hat{w}_{t-1}^i f_{t|t}^{Y|X}\left(y_t|X_t^i\right)$ for $i = 1, \ldots, N_{par}$;
  $W_t^i = \frac{w_t^i}{\sum_{m=1}^{N_{par}} w_t^m}$ for $i = 1, \ldots, N_{par}$;
end

Notice that the bootstrap filter algorithm (Algorithm 3) is the same as Algorithm 2 with the components of the Feynman-Kac model as in Definition 6.1.5. Using the bootstrap filter algorithm, we can approximate the distribution $\mathbb{P}\left(X_t \in dx_t|Y_{0:t} = y_{0:t}\right)$ by $\mathbb{Q}_t^{N_{par}}(dx_t)$ defined as in (6.1.28). And the likelihood $L^Y(\theta) = f_{0:T}^Y(y_{0:T}|\theta)$ is approximated by

$$L_T^{N_{par}}(\theta) = \prod_{t=1}^{T} l_t^{N_{par}}(\theta), \tag{6.1.30}$$

where $l_t^{N_{par}}(\theta)$ is defined as in equation (6.1.29).

6.1.6 Particle marginal Metropolis-Hastings algorithm

In the previous section, we saw how to use the bootstrap filter algorithm to generate samples from the distribution $\mathbb{P}_t\left(X_t \in dx_t|Y_{0:t} = y_{0:t}\right)$. In Algorithm 3, the vector of parameters θ is assumed to be known. In general, this vector θ is unknown and we would like to estimate it. To that end, we will use Bayesian estimation techniques. Before introducing the particle marginal Metropolis-Hastings algorithm, let us first introduce the Metropolis-Hastings algorithm.

Let us assume that the vector of parameters θ can take values on the set Θ. In the Bayesian setting, the parameters are assumed to satisfy a prior distribution; that is, the vector of parameters θ is a random variable with distribution $\mathbb{P}(d\theta)$. We will assume that the prior distribution can be expressed as

$$\mathbb{P}(d\theta) = f^{\theta}(\theta)m(d\theta),$$

where m is Lebesgue measure.

The density function f^{θ} is called the prior density function. The objective of the Bayesian estimation technique is to sample from the conditional distribution of θ given $Y_{0:N}$; that is, we are interested in the distribution

$$\mathbb{P}\left(\theta \in d\theta | Y_{0:T} = y_{0:T}\right).$$

If we have that

$$\mathbb{P}\left(Y_{0:T} \in dy_{0:t} | \theta\right) = f^{Y}_{0:t}(y_{0:T}|\theta)m(dy_{0:T}),$$

then by the Bayes theorem we can express the posterior distribution as

$$\mathbb{P}\left(\theta \in d\theta | Y_{0:T} = y_{0:T}\right) = \frac{1}{L} f^{\theta}(\theta) f^{Y}_{0:T}(y_{0:T}|\theta)m(d\theta),$$

where

$$L = \int_{\Theta} f^{\theta}(\theta) f^{Y}_{0:T}(y_{0:T}|\theta)m(d\theta)$$

is the normalizing constant (Hautsch & Ou, 2008, p. 257). We have expressed the posterior density in terms of the prior density and the likelihood. In the case that the likelihood has an analytical form and we can compute it, we can use the Metropolis-Hastings algorithm to sample from the posterior distribution. This algorithm is based on the rejection sampling algorithm (Rachev, Hsu, Bagasheva, & Fabozzi, 2008, p. 64). In the rejection sampling algorithm, we want to sample from an objective distribution, but we can not sample directly from it, we must use a proposal distribution that we know how to sample from. We generate samples from the proposal distribution and then those samples are accepted or rejected according to a certain rule.

The objective of the Metropolis-Hastings algorithm is to generate $N_{sim} \in \mathbb{N}$ samples from the posterior distribution. That is, we would like to generate a sequence $\theta^{1:N_{sim}} = \left(\theta^{(1)}, \theta^{(2)}, \ldots, \theta^{(N_{sim})}\right)$ where $\theta^{(i)}$ has been sampled from the posterior distribution $\mathbb{P}\left(\theta \in d\theta | Y_{0:T} = y_{0:T}\right)$ for $i = 1, 2, \ldots, N_{sim}$. The problem is that in general it is not possible to sample from the posterior distribution directly. Because of that, we generate samples in a sequential manner from a proposal probability kernel. That is given $\theta^{(i-1)}$ we generate the proposed sample $\tilde{\theta}$ from the probability kernel $\tilde{M}(\theta^{(i-1)}, d\tilde{\theta})$, then $\tilde{\theta}$ is accepted with probability

$$\alpha\left(\theta^{(i-1)}, \tilde{\theta}\right) = \min\left\{1, r\left(\theta^{(i-1)}, \tilde{\theta}\right)\right\},$$

where

$$
\begin{aligned}
r\left(\theta^{(i-1)},\tilde{\theta}\right) &= \frac{\mathbb{P}\left(\tilde{\theta}\in d\tilde{\theta}|Y_{0:T}=y_{0:T}\right)\tilde{M}\left(\tilde{\theta},d\theta^{(i-1)}\right)}{\mathbb{P}\left(\theta^{(i-1)}\in d\theta^{(i-1)}|Y_{0:T}=y_{0:T}\right)\tilde{M}\left(\theta^{(i-1)},d\tilde{\theta}\right)}\\
&\propto \frac{f^{\theta}\left(\tilde{\theta}\right)f^{Y}_{0:T}\left(y_{0:T}|\tilde{\theta}\right)m\left(d\tilde{\theta}\right)\tilde{M}\left(\tilde{\theta},d\theta^{(i-1)}\right)}{f^{\theta}\left(\theta^{(i-1)}\right)f^{Y}_{0:T}\left(y_{0:T}|\theta^{(i-1)}\right)m\left(d\theta^{(i-1)}\right)\tilde{M}\left(\theta^{(i-1)},d\tilde{\theta}\right)}.
\end{aligned}
$$

If the proposal probability kernel $\tilde{M}\left(\theta^{(i-1)},d\tilde{\theta}\right)$ can be written in terms of a density function $\tilde{m}$ as

$$\tilde{M}\left(\theta^{(i-1)},d\tilde{\theta}\right)=\tilde{m}\left(\tilde{\theta}|\theta^{(i-1)}\right)m\left(d\tilde{\theta}\right),$$

then the Metropolis-Hastings algorithm works as it is shown in Algorithm 4 (Rachev, Hsu, Bagasheva, & Fabozzi, 2008, p. 68).

Algorithm 4: Metropolis-Hastings algorithm

Initialize the sequence with a value $\theta^{(0)}$;
for $i=1,2,\dots N_{sim}$ **do**
  Simulate $\tilde{\theta}\sim\tilde{M}\left(\theta^{(i-1)},d\tilde{\theta}\right)$;
  Set $\alpha=\min\left\{1,\frac{f^{\theta}(\tilde{\theta})f^{Y}_{0:T}(y_{0:T}|\tilde{\theta})\tilde{m}(\theta^{(i-1)}|\tilde{\theta})}{f^{\theta}(\theta^{(i-1)})f^{Y}_{0:T}(y_{0:T}|\theta^{(i-1)})\tilde{m}(\tilde{\theta}|\theta^{(i-1)})}\right\}$;
  Simulate $u\sim Unif([0,1])$;
  if $u\leq\alpha$ **then**
    Set $\theta^{(i)}=\tilde{\theta}$;
  else
    Set $\theta^{(i)}=\theta^{(i-1)}$;
  end
end

A special case of the Metropolis-Hastings algorithm is the random walk Metropolis-Hastings algorithm, where the proposal samples are generated according to the following probability kernel

$$\tilde{M}\left(\theta^{(i-1)},d\tilde{\theta}\right)=f_{MN}\left(\tilde{\theta}|\theta^{(i-1)},\delta\Sigma\right)m\left(d\tilde{\theta}\right) \tag{6.1.31}$$

where $\delta>0$ and $f_{MN}(.|\theta^{(i-1)},\delta\Sigma)$ is the density function of a multivariate normal distribution with vector mean $\theta^{(i-1)}$ and covariance matrix $\delta\Sigma$ (Chopin & Papaspiliopoulos, 2020, p. 282). Because the normal distribution is symmetric around the mean we have that

$$\tilde{m}\left(\tilde{\theta}|\theta^{(i-1)}\right)=\tilde{m}\left(\theta^{(i-1)}|\tilde{\theta}\right).$$

In this particular case we have that

$$\alpha=\min\left\{1,\frac{f^{\theta}\left(\tilde{\theta}\right)f^{Y}_{0:T}\left(y_{0:T}|\tilde{\theta}\right)}{f^{\theta}\left(\theta^{(i-1)}\right)f^{Y}_{0:T}\left(y_{0:T}|\theta^{(i-1)}\right)}\right\}.$$

Remark 6.1.5. *The density function of a multivariate normal random variable with vector mean $\mu \in \mathbb{R}^n$ and $n \times n$ covariance matrix Σ is*

$$f_{MN}(x|\mu,\Sigma) = \frac{\exp\left\{-\frac{1}{2}(x-\mu)'\Sigma^{-1}(x-\mu)\right\}}{\sqrt{(2\pi)^n|\Sigma|}} \text{ for } x \in \mathbb{R}^n,$$

where $|\Sigma|$ is the determinant of Σ and $(x-\mu)'$ is the transpose of $(x-\mu)$.

The sequence $\theta_{1:N_{sim}}$ generated by the Metropolis-Hastings algorithm converges to the posterior distribution $\mathbb{P}(\theta \in d\theta | Y_{0:T} = y_{0:T})$ in theory when $N_{sim} \to \infty$. Of course, we can not generate a sequence of infinite size. But this means that the sequence takes some time to converge to the posterior. So the initial values of the sequence $\theta^{1:N_{sim}}$ may not have converged to the posterior. Because of that we will get rid of some of the initial values of the sequence $\theta^{1:N_{sim}}$. We pick a positive integer N_{burn} with $N_{burn} < N_{sim}$, called the burn-in period, and we will eliminate all of the $\theta^{(i)}$ with $i < N_{burn}$. So at the end, we are only interested in the sequence $\theta^{N_{burn}:N_{sim}}$ (Rachev, Hsu, Bagasheva, & Fabozzi, 2008, p. 75).

In general, for state space models, we do not have an analytical formula for the likelihood $f^Y_{0:T}$, but we can approximate it using the bootstrap filter algorithm (Algorithm 3). From the result in (6.1.30), we have that the likelihood can be approximated by $L_T^{N_{par}}(\theta)$. The particle marginal Metropolis-Hastings algorithm is based on the Metropolis-Hastings algorithm and on the bootstrap filter. This algorithm allows us to sample from the posterior distribution and perform Bayesian estimation of the vector of parameters of the state space model. The particle marginal Metropolis-Hastings algorithm works as it is explained in Algorithm 5 (Chopin & Papaspiliopoulos, 2020, p. 308).

Algorithm 5: Particle marginal Metropolis-Hastings algorithm

Initialize the sequence with a value $\theta^{(0)}$;
Run the bootstrap filter (Algorithm 3) using $\theta^{(0)}$ and obtain $L_T^{N_{par}}\left(\theta^{(0)}\right)$;
for $i = 1, 2, \ldots N_{sim}$ **do**
  Simulate $\tilde{\theta} \sim \tilde{M}\left(\theta^{(i-1)}, d\tilde{\theta}\right)$;
  Run the bootstrap filter (Algorithm 3) using $\tilde{\theta}$ and obtain $L_T^{N_{par}}\left(\tilde{\theta}\right)$;
  Set $\alpha = \min\left\{1, \frac{f^\theta(\tilde{\theta}) L_T^{N_{par}}(\tilde{\theta}) \tilde{m}(\theta^{(i-1)}|\tilde{\theta})}{f^\theta\left(\theta^{(i-1)}\right) L_T^{N_{par}}\left(\theta^{(i-1)}\right) \tilde{m}(\tilde{\theta}|\theta^{(i-1)})}\right\}$;
  Simulate $U \sim Unif([0,1])$;
  if $U \leq \alpha$ **then**
    Set $\theta^{(i)} = \tilde{\theta}$;
  else
    Set $\theta^{(i)} = \theta^{(i-1)}$;
  end
end

6.1.7 Numerical examples

Now that we know how to estimate the vector of parameters θ, we would like to try the methods introduced in sections 6.1.2 and 6.1.6. We will apply these methods to synthetic data that has been generated using the models introduced in section 6.1.1.

In all the experiments done in this chapter we use a tolerance of $N_{tol} = 0.5 N_{par}$. The policy used to generate the proposed parameters is the one that is explained in equation (6.1.31). That is, we use a random walk particle marginal Metropolis-Hastings algorithm. So in this case, the value α that appears in Algorithm 5 can be written as

$$\alpha = \min\left\{1, \frac{f^{\theta}\left(\tilde{\theta}\right) L_T^{N_{par}}\left(\tilde{\theta}\right)}{f^{\theta}\left(\theta^{(i-1)}\right) L_T^{N_{par}}\left(\theta^{(i-1)}\right)}\right\}.$$

Example 6.1.3. *Simple stochastic volatility model.*

For the model introduced in Example 6.1.1, we simulate 1100 *time steps (approximately three years of daily data) with the following values for the parameters*

$$\rho_0 = 0.4, \quad \rho_1 = 0.8, \quad \tau = 0.5$$

and an initial distribution which is a normal distribution with mean $\mu_0 = 1$ *and variance* $\sigma_0^2 = 0.05$*. Using the synthetic data of the logarithmic returns, we would like to apply the particle marginal Metropolis-Hastings algorithm (Algorithm 5) and see if we can estimate the parameters. The prior distributions for the parameters are the ones that appear in Table 6.1. The covariance matrix of the multivariate normal density of equation (6.1.31) is*

$$\delta\Sigma = 0.1^2 I_3,$$

where I_3 *is the identity matrix of dimension* 3*.*

Parameter	Prior Distribution
ρ_0	N(0,10)
ρ_1	N(0,1)
τ	Gamma(1,1/2)

TABLE 6.1: Table showing prior distributions of the parameters of the model introduced in Example 6.1.1.

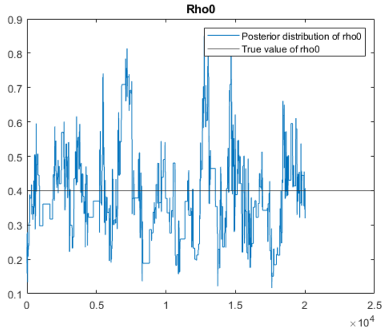
We use $N_{par} = 200$ *particles for the bootstrap filter part and* $N_{sim} = 20000$ *simulations for the Metropolis-Hastings algorithm part. We can observe in Figure 6.1 that the algorithm gives estimates that are close to the true values of the parameters.*

In addition, we have computed the mean and the 95% *confidence interval of the sequence of the parameters obtained in the particle Metropolis-Hastings algorithm with a burn-in period of* $N_{burn} = 2000$*. The values that we have obtained appear in Table 6.2 and we observe that the true value of the parameter is always inside the* 95% *confidence interval.*

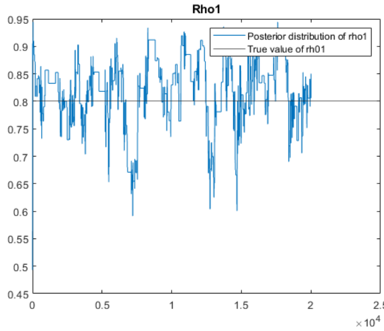
Parameter	True value	Mean	95% confidence interval
ρ_0	0.4	0.389004	(0.184163,0.709051)
ρ_1	0.8	0.818377	(0.667205,0.919404)
τ	0.4	0.388147	(0.240680,0.554056)

TABLE 6.2: Table containing the mean and the 95% confidence interval of the posterior distribution of the parameters, obtained using Algorithm 5 of the model introduced in Example 6.1.1.

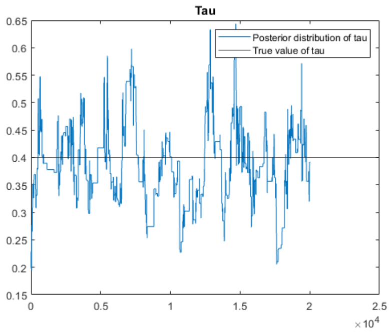
Example 6.1.4. *Heston model.*



(A) Estimation of the parameter ρ_0.



(B) Estimation of the parameter ρ_1.



(C) Estimation of the parameter τ.

FIGURE 6.1: Posterior distribution of the parameters of the model introduced in Example 6.1.1.

In the case of the discretized Heston model introduced in Example 6.1.2, we simulate 1100 *steps with the following values of the parameters*

$$a_I = 0.5, \quad b_I = 0.06, \quad \sigma_I = 0.15, \quad \mu = 0.1, \quad \rho = 0.5,$$

the initial distribution is a normal distribution with mean $\mu_0 = 0.1$ *and variance* $\sigma_0^2 = 0.01$ *and a time step of* $\Delta t = 0.1$.

Using the synthetic data of the logarithmic returns, we would like to apply the particle marginal Metropolis-Hastings algorithm (Algorithm 5) and see if we can estimate the parameters. The prior distributions for the parameters are shown in Table 6.3. The covariance matrix of the multivariate normal density of equation (6.1.31) is

$$\delta\Sigma = 0.05^2 I_5,$$

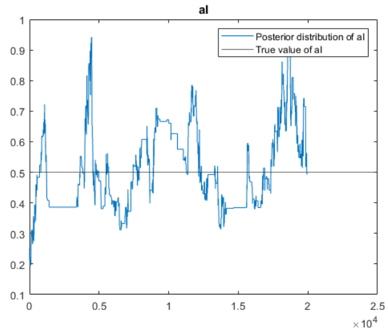
where I_5 *is the identity matrix of dimension* 5.

Parameter	Prior Distribution
a_I	Gamma(1,2)
b_I	Gamma(1,2)
σ_I	Gamma(1,1/2)
μ	N(0,1)
ρ	Unif(-1,1)

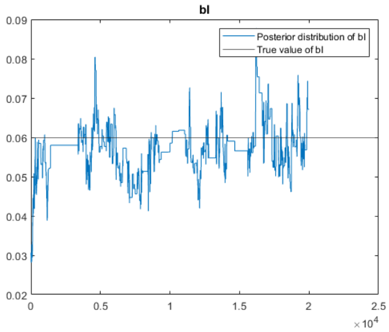
TABLE 6.3: Table showing the prior distributions of the parameters of the model introduced in Example 6.1.2.

Again, we use $N_{par} = 200$ *particles and* $N_{sim} = 20000$ *simulations for the particle marginal Metropolis-Hastings algorithm. Figure 6.2 shows the results of the posterior distribution of the parameters obtained by the particle marginal Metropolis-Hastings algorithm.*

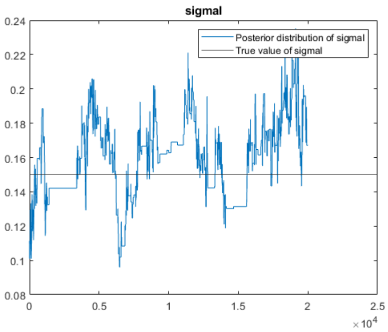
In Table 6.4, we observe the mean and the 95% *confidence interval obtained by the algorithm with a burn-in period of* $N_{burn} = 2000$. *The true values of the parameters are inside the* 95% *confidence interval.*



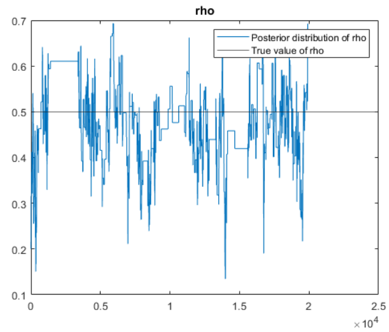
(A) Estimation of the parameter a_I.



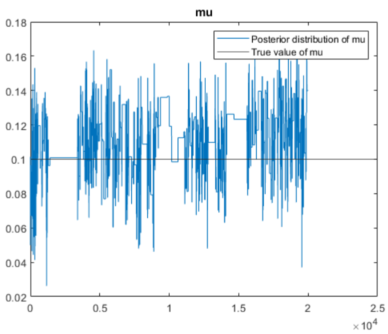
(B) Estimation of the parameter b_I.



(C) Estimation of the parameter σ_I.



(D) Estimation of the parameter ρ.



(E) Estimation of the parameter μ.

FIGURE 6.2: Posterior distribution of the parameters of the model introduced in Example 6.1.2.

6.2 State space models with an exogenous process

In Section 6.1, we introduced state space models. These models have a latent discrete process $X = \{X_t\}_{t=0}^{\infty}$ and an observed discrete process $Y = \{Y_t\}_{t=0}^{\infty}$, that is affected by X. Now we are interested in models in which the observed process Y is affected by the latent process X and an exogenous discrete process $A = \{A_t\}_{t=0}^{\infty}$. Since the process A is observed, we will assume that we know the values of the following sequence of

Parameter	True value	Mean	95% confidence interval
a_I	0.5	0.516925	(0.332549,0.816280)
b_I	0.06	0.057649	(0.045616,0.071305)
σ_I	0.15	0.161564	(0.113633,0.205397)
μ	0.1	0.111692	(0.075767,0.145486)
ρ	0.5	0.475196	(0.297232,0.623241)

TABLE 6.4: Table containing the mean and the 95% confidence interval of the posterior distribution of the parameters, obtained using Algorithm 5 of the model introduced in Example 6.1.2.

realizations $\{a_t\}_{t=0}^{\infty}$ of the process A. We assume that the processes X and Y form a state space model as in Definition 6.1.3 with a vector of parameters θ such that $\{a_t\}_{t=0}^{\infty} \subseteq \theta$. We can do this due to the fact that the realizations of the process A are observed. So, we treat these types of models as states space models whose vector of parameters contains the realizations of the exogenous process. Since we have a state space model, we can use the techniques introduced in Section 6.1.

Based on the stochastic volatility models introduced in Section 6.1.1, we would like to construct two different types of models. In the first type of model, the unobserved volatility of log-returns is partially explained by the market attention. These types of models have been studied previously by Li, Yang, and Wang (2019) and by Balash (2013). In the other type of model, the unobserved part of the volatility is independent of the interest process. That is, the volatility of the log-returns is explained by two processes, one an unobserved process and the other, the market attention.

6.2.1 Stochastic volatility model with exogenous variable

Based on the model introduced in Example 6.1.1 and on the work of Balash (2013, p. 35), we construct the following model,

$$Y_t = \exp\left\{\frac{X_t}{2}\right\}\epsilon_t \tag{6.2.1}$$

$$X_t = \rho_0 + \rho_1 X_{t-1} + \rho_2 A_t + \tau\eta_t, \tag{6.2.2}$$

where $\rho_0, \rho_2 \in \mathbb{R}$, $\tau > 0$, $\rho_1 \in (-1, 1)$ and $\{\eta_t\}_{t=1}^{\infty}$, $\{\epsilon_t\}_{t=1}^{\infty}$ are two sequences of independent standard normal random variables. The distribution of X_0 is a normal distribution with known mean μ_0 and known variance σ_0^2.

The probability kernels of this state space model are

$$\begin{aligned} F_t(x_t, dy_t) &= f_N(y_t|0, e^{x_t})m(dy_t), \\ P_t(x_{t-1}, dx_t) &= f_N(x_t|\rho_0 + \rho_1 x_{t-1} + \rho_2 a_t, \tau^2)m(dx_t). \end{aligned}$$

From equation (6.2.1), we have that the unobserved volatility at time t is not only explained by its value at time $t-1$, but also by the market attention at time t.

For this model, we would like to check that the particle marginal Metropolis-Hastings algorithm; can accurately estimate the parameters of this model.

For the model introduced in equations (6.2.1)-(6.2.2), we simulate 1100 time steps with the following values for the parameters

$$\rho_0 = 0.4, \quad \rho_1 = 0.8, \quad \rho_2 = 0.5, \quad \tau = 0.4$$

and an initial normal distribution with mean $\mu_0 = 2$ and variance $\sigma_0^2 = 0.2$. Using the synthetic data of the logarithmic returns, we would like to apply the particle marginal Metropolis-Hastings algorithm (Algorithm 5) and see if we can estimate the parameters. The prior distributions for the parameters are the ones that appear in Table 6.5. Again, the covariance matrix of the multivariate normal density of equation (6.1.31) is

$$\delta\Sigma = 0.05^2 I_4,$$

where I_4 is the identity matrix of dimension 4.

Remark 6.2.1. *We still have to specify how we generate the synthetic realizations of the interest process A. In this case, we assume that the attention follows the following autoregressive process:*

$$A_t = \phi_0 + \phi_1 A_{t-1} + \mu\epsilon_t^A, \tag{6.2.3}$$

where $\phi_0 \in \mathbb{R}$, $\mu > 0$, $\phi_1 \in (-1, 1)$, $\{\epsilon_t^A\}_{t=1}^{\infty}$ is a sequence of independent standard normal random variables and A_0 is a constant random variable. For the generation of the synthetic realizations, we use the following values for the parameters

$$\phi_0 = 0.3, \quad \phi_1 = 0.4, \quad \mu = 0.3$$

and an initial value of $A_0 = 0.2$.

Parameter	Prior Distribution
ρ_0	N(0,1)
ρ_1	N(0,1)
ρ_2	N(0,1)
τ	Gamma(1,1/2)

TABLE 6.5: Table showing the prior distributions of the parameters of model defined in equations (6.2.1)-(6.2.2).

As in Example 6.1.3, we use $N_{par} = 200$ particles and $N_{sim} = 20000$ simulations for the particle marginal Metropolis-Hastings algorithm. We can observe in Figure 6.3 that the algorithm manages to estimate the true value of the parameters.

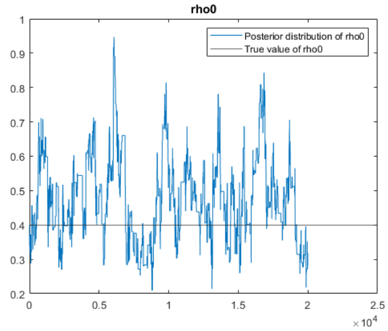
As shown in Table 6.6, the 95% confidence interval obtained by the algorithm with a burn-in period of $N_{burn} = 2000$ contains the true value of the parameters.

Parameter	True value	Mean	95% confidence interval
ρ_0	0.4	0.470852	(0.273000,0.755202)
ρ_1	0.8	0.792326	(0.703900,0.856203)
τ	0.4	0.406326	(0.329099,0.520524)
ρ_2	0.5	0.474613	(0.278493,0.641246)

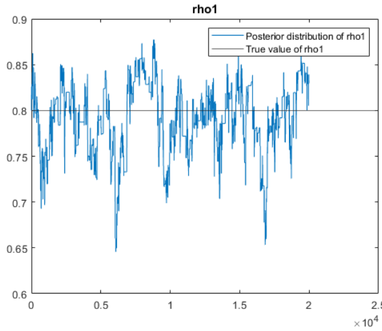
TABLE 6.6: Table containing the mean and the 95% confidence interval of the posterior distribution of the parameters, obtained using Algorithm 5, of the model defined in equations (6.2.1)-(6.2.2).

6.2.2 Partially observed volatility model

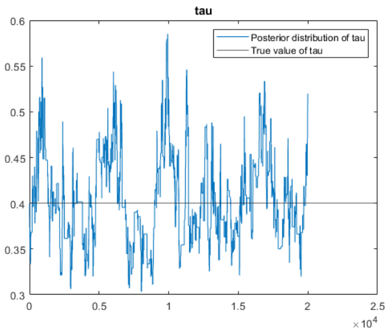
We are now interested in a model in which the volatility is composed of two parts. One part is unobserved and the other one is observed. Based on the model introduced



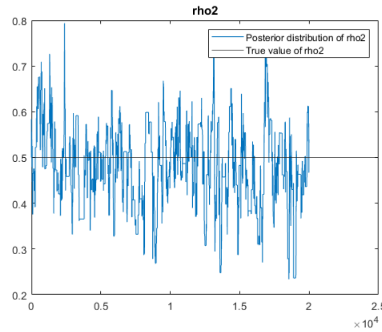
(A) Estimation of the parameter ρ_0.



(B) Estimation of the parameter ρ_1.



(C) Estimation of the parameter τ.



(D) Estimation of the parameter ρ_2.

FIGURE 6.3: Posterior distribution of the parameters of the model introduced in equations (6.2.1)-(6.2.2).

in Example 6.1.1, we propose the following model,

$$Y_t = \exp\left\{\frac{A_t}{2} + \frac{X_t}{2}\right\}\epsilon_t \qquad (6.2.4)$$

$$X_t = \rho_0 + \rho_1 X_{t-1} + \tau\eta_t, \qquad (6.2.5)$$

where $\rho_0 \in \mathbb{R}$, $\tau > 0$, $\rho_1 \in (-1, 1)$ and where $\{\eta_t\}_{t=1}^{\infty}$, $\{\epsilon_t\}_{t=1}^{\infty}$ are two sequences of independent standard normal random variables. The distribution of X_0 is a normal distribution with known mean μ_0 and known variance σ_0^2. From equation (6.2.4), we observe that in this case the attention and the unobserved part of the volatility are two separated processes. The probability kernels of this state space model are

$$F_t(x_t, dy_t) = f_N(y_t|0, e^{a_t+x_t})m(dy_t),$$
$$P_t(x_{t-1}, dx_t) = f_N(x_t|\rho_0 + \rho_1 x_{t-1}, \tau^2)m(dx_t).$$

As in previous examples, we would like to apply the particle marginal Metropolis-Hastings algorithm to the model introduced in equations (6.2.4)–(6.2.5). To that end, we simulate 1100 time steps with the following values of the parameters:

$$\rho_0 = 0.4, \quad \rho_1 = 0.8, \quad \tau = 0.4$$

and an initial normal distribution with mean $\mu_0 = 2$ and variance $\sigma_0^2 = 0.07$.

We choose $N_{par} = 200$ and $N_{sim} = 20000$ as the number of particles and simulations respectively, that are going to be used in the particle marginal Metropolis-Hastings algorithm. The priors that have been chosen for the parameters appear in Table 6.7. The covariance matrix of the multivariate normal density of equation

(6.1.31) is

$$\delta\Sigma = 0.05^2 I_3,$$

where I_3 is the identity matrix of dimension 3.

Remark 6.2.2. *Again, for the generation of the synthetic data of the interest process, we use the autoregressive model introduced in equation (6.2.3) with the following values for the parameters:*

$$\phi_0 = 0.4, \quad \phi_1 = 0.8, \quad \mu = 0.4$$

and an initial value of $A_0 = 0.2$.

Parameter	Prior Distribution
ρ_0	N(0,1)
ρ_1	N(0,1)
τ	Gamma(1,1/2)

Table 6.7: Table showing the prior distributions of the parameters of model defined in equations (6.2.4)-(6.2.5).

Parameter	True value	Mean	95% confidence interval
ρ_0	0.4	0.354721	(0.222827,0.559826)
ρ_1	0.8	0.826112	(0.721924,0.891564)
τ	0.4	0.438589	(0.335611,0.567222)

Table 6.8: Table containing the mean and the 95% confidence interval of the posterior distribution of the parameters, obtained using Algorithm 5, of the model defined in equations (6.2.4)-(6.2.5).

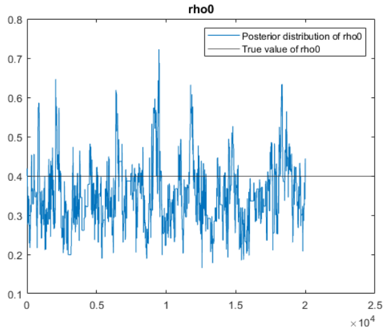
Figure 6.4 and Table 6.8 show that the algorithm manages to give a good estimation for the true value of the parameters. The mean and the confidence interval that appear in Table 6.8 are computed using a burn-in period of $N_{burn} = 2000$.

6.3 Real data

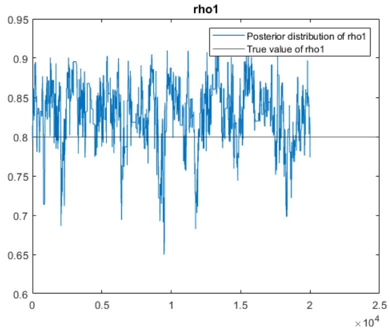
In this section, we would like to fit the models introduced in Section 6.2 to Bitcoin price data. The price data is taken from https://charts.coinmetrics.io/network-data/ for the dates from 19/09/2018 to 20/09/2021. We will first fit the data to the model introduced in Section 6.2.1 and then to the model introduced in Section 6.2.2. As in previous chapters, the proxies for the market attention are going to be the number of Wikipedia views of the keyword "altcoin" and the unique number of active addresses. The number of Wikipedia views are taken from https://pageviews.toolforge.org and the unique number of active addresses is taken from https://charts.coinmetrics.io/network-data/ for the dates 20/09/2018 to 20/09/2021. The selected proxies for the market attention are presented in Figure 6.5.

Remark 6.3.1. *Due to the fact that the proxies for the attention have large values, we scale them to fit on the interval* $[0, 1]$*. That is, if* T *is the horizon and* $\{a_t\}_{t=0}^T$ *is the sequence of realizations of the data of the attention, we construct the scaled version of the attention as:*

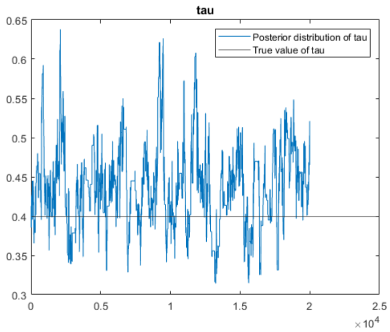
$$\tilde{a}_t = \frac{a_t}{\max\{a_s : s \in \{0, 1, \dots, T\}\}} \text{ for } t = 0, 2, \dots, T.$$



(A) Estimation of the parameter ρ_0.



(B) Estimation of the parameter ρ_1.



(C) Estimation of the parameter τ.

FIGURE 6.4: Posterior distribution of the parameters of the model defined in equations (6.2.4)-(6.2.5).

We use this scaled version of the attention to fit the models presented in Section 6.2.

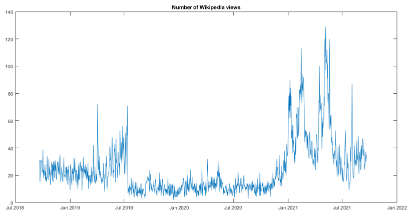
6.3.1 Benchmark model

We would like to compare the models introduced in Section 6.2 to a benchmark model. In this case, the benchmark model is the simple stochastic volatility model presented in Example 6.1.1. For fitting this model, we only use the log-returns of Bitcoin; we do not use any proxy for the attention. The idea is to compare this model with the models that contain an exogenous variable. We use the particle marginal Metropolis-Hastings algorithm with $N_{par} = 200$ particles and $N_{sim} = 50000$ simulations. The mean and the confidence interval of the posterior distribution of the parameters are shown in Table 6.9, which are calculated using a burn-in period of $N_{burn} = 5000$. The posterior distribution of the parameters are shown in Figure 6.6.

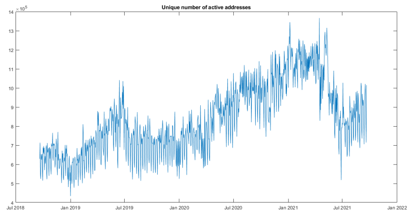
Parameter	Mean	95% confidence interval
ρ_0	-1.124007	(-1.294804,-0.972485)
ρ_1	0.842064	(0.819718,0.861390)
τ	0.681720	(0.597685,0.837154)

TABLE 6.9: Table containing the mean and the 95% confidence interval of the posterior distribution of the parameters, obtained using Algorithm 5. For the model defined in equations (6.1.15)-(6.1.16).

Using the mean of the posterior distribution of the parameters shown in Table 6.9, we use the bootstrap filter algorithm to filter the hidden process X. The filtered distribution of the hidden process X is shown in Figure 6.7.



(A) Number of Wikipedia views of the word "Altcoin" from 20/09/2018 to 20/09/2021.



(B) Number of unique active addresses taken from 20/09/2018 to 20/09/2021.

FIGURE 6.5: Number of Wikipedia views and number of unique active addresses.

The objective in this section is to compare the volatility given by this benchmark model with the volatility given by the models with an exogenous variable. Notice that from equation (6.1.15) we have that the conditional standard deviation of the return Y_t given X_t is:

$$\sqrt{Var[Y_t|X_t]} = e^{0.5X_t}. \tag{6.3.1}$$

We would like to compare the conditional standard deviation of this model with the weekly historical volatility, that is defined as

$$h_t = \sqrt{\frac{1}{7}\sum_{j=t-7}^{t} y_j^2 - \left(\frac{1}{7}\sum_{j=t-7}^{t} y_j\right)^2} \text{ for } t = 7, 8, \ldots T.$$

In Figure 6.8, we show the weekly historical volatility and the estimated conditional standard deviation defined in equation (6.3.1). For each $t = 0, 1, 2, \ldots, T$ the random variable X_t is estimated as the mean of the posterior distribution obtained from the bootstrap filter. We observe that the weekly historical volatility has spikes that are not captured by the estimated volatility. We hope that the models that incorporate the market attention as an exogenous variable will be able to capture these spikes.

Remark 6.3.2. *The prior distributions of the parameters used in the particle marginal Metropolis-Hastings algorithm are the same as the ones used in the numerical example presented in Example 6.1.3.*

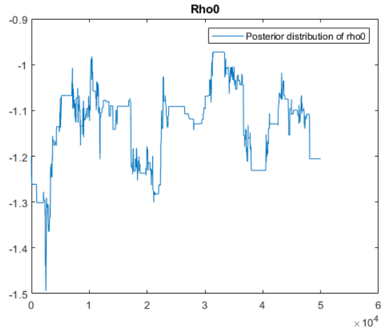
As it is done by A. Hou, Wang, Chen, and Härdle (2020) we can compute the residuals, defined as

$$\epsilon_t = \exp\left\{-\frac{X_t}{2}\right\} Y_t \text{ for } t = 1, 2, \ldots T, \tag{6.3.2}$$

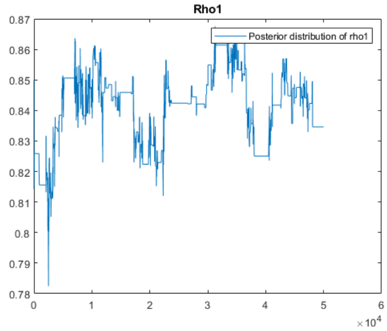
to validate the model. Under the assumption that the data have been generated by the model defined in Example 6.1.1 the sequence $\{\epsilon_t\}_{t=1}^{T}$ is a sequence of independent standard normal random variables. To check this assumption, we can use the Kolmogorov–Smirnov test (Massey, 1951). This test gives us a p–value of 0.0061, since the p–vale is less than 0.05 we can reject the null hypotheses with a 5% confidence level.

6.3.2 Stochastic volatility model with exogenous variable

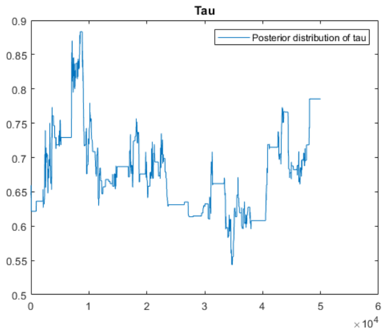
For the model introduced in Section 6.2.1, we first fit the model using the number of Wikipedia views as proxy for the interest process. We use $N_{par} = 200$ particles and



(A) Estimation of the parameter ρ_0.



(B) Estimation of the parameter ρ_1.



(C) Estimation of the parameter τ.

FIGURE 6.6: Posterior distribution of the parameters of the model defined in Example 6.1.1 using the log-returns of Bitcoin for fitting the model.

$N_{sim} = 50000$ simulations for the particle marginal Metropolis-Hastings algorithm.

The results related to the posterior distribution of the parameters are shown in Table 6.10 and in Figure 6.9. The mean and the confidence interval shown in Table 6.10 are computed using a burn-in period of $N_{burn} = 5000$.

Parameter	Mean	95% confidence interval
ρ_0	-2.024311	(-2.341229,-1.754662)
ρ_1	0.738657	(0.689649,0.777086)
ρ_2	0.798545	(0.584970,0.957000)
τ	0.845450	(0.710839,0.945651)

TABLE 6.10: Table containing the mean and the 95% confidence interval of the posterior distribution of the parameters, obtained using Algorithm 5. For the model defined in equations (6.2.1)-(6.2.2) and where the proxy of attention is the number of Wikipedia views.

We use the mean of the posterior distribution of the parameters shown in Table 6.10 as estimators for the parameters. With those estimators, we use the bootstrap filter algorithm to filter the hidden process X. The filtered distribution of the hidden process X appears in Figure 6.10.

From equation (6.2.1), we have that the conditional standard deviation of the return Y_t given X_t is:

$$\sqrt{Var[Y_t|X_t]} = e^{0.5X_t}. \tag{6.3.3}$$

As we did in Section 6.3.1, we compare the conditional standard deviation and the

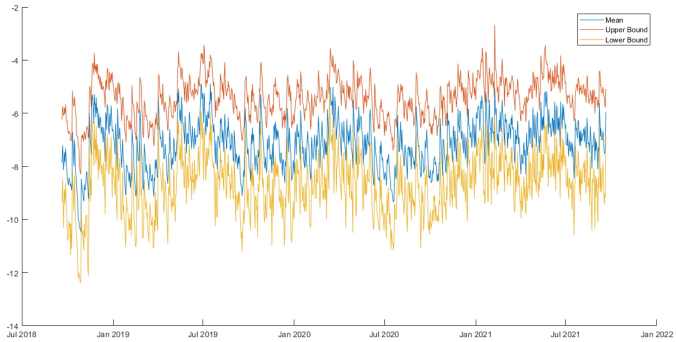


FIGURE 6.7: Filtered distribution of the hidden process X of the model defined in Example 6.1.1 using the log-returns of Bitcoin for fitting the model. The mean and the 95% confidence interval of the posterior distribution are shown.

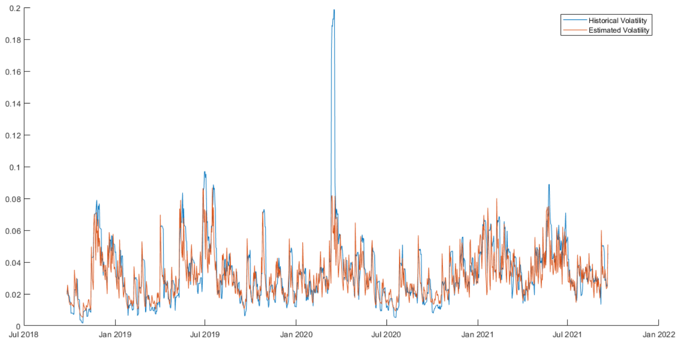
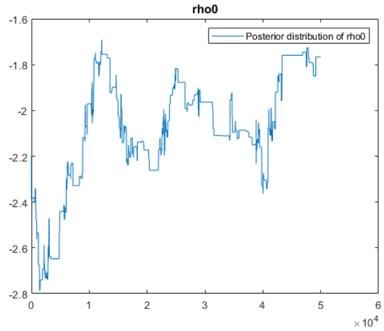


FIGURE 6.8: Weekly historical volatility and estimated volatility of the model defined in Example 6.1.1 using the log-returns of Bitcoin for fitting the model.

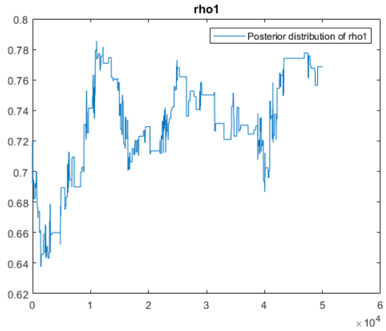
weekly historical volatility. This is shown in Figure 6.11. We observe that this model captures the spikes better than the benchmark model. As we did before, we compute the residuals for this model. In this case the p–value given by the Kolmogorov-Smirnov test to check the normality assumption is 0.0227. So, we can reject the normality assumption with a confidence level of 5% but we cannot reject the assumption with a confidence level of 1%.

Proceeding in a similar manner and using the number of unique active addresses as the proxy of attention, we compare the conditional standard deviation defined in equation (6.3.3) with the weekly historical volatility. This is shown in Figure 6.12. Again, the proposed model captures the spikes better than the benchmark model. In this case the p–value of the Kolmogorov-Smirnov test is 0.0064.

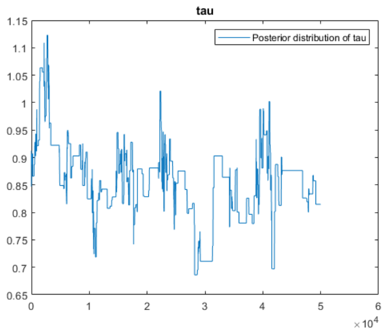
Remark 6.3.3. *The prior distributions of the parameters used in the particle marginal Metropolis-Hastings algorithm are the same as the ones used in the numerical example presented in Section 6.2.1.*



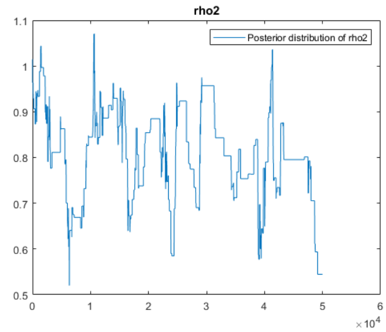
(A) Estimation of the parameter ρ_0.



(B) Estimation of the parameter ρ_1.



(C) Estimation of the parameter τ.



(D) Estimation of the parameter ρ_2.

FIGURE 6.9: Posterior distribution of the parameters of the model introduced in equations (6.2.1)-(6.2.2) and where the proxy of attention is the number of Wikipedia views.

6.3.3 Partially observed volatility model

Let us now focus on the model introduced in Section 6.2.2. For this case, we use $N_{par} = 200$ particles and $N_{sim} = 50000$ simulations for the particle marginal Metropolis-Hastings algorithm.

The results of the posterior distribution of the parameters are shown in Table 6.11 and in Figure 6.13. The mean and the confidence interval shown in Table 6.11 are computed using a burn-in period of $N_{burn} = 5000$.

Parameter	Mean	95% confidence interval
ρ_0	-1.602500	(-2.056705,-1.125938)
ρ_1	0.782517	(0.716372,0.848551)
τ	0.814440	(0.703566,0.956670)

TABLE 6.11: Table containing the mean and the 95% confidence interval of the posterior distribution of the parameters, obtained using Algorithm 5. For the model defined in equations (6.2.4)-(6.2.5) and where the proxy of attention is the number of Wikipedia views.

Again, we use the mean of the posterior distribution of the parameters that appears in Table 6.11, as estimators of the parameters. With those estimators we use the bootstrap filter algorithm to filter the Markov process X. The filtered distribution of the hidden process X appears in Figure 6.14.

We are again interested in comparing the volatility given by this model with the

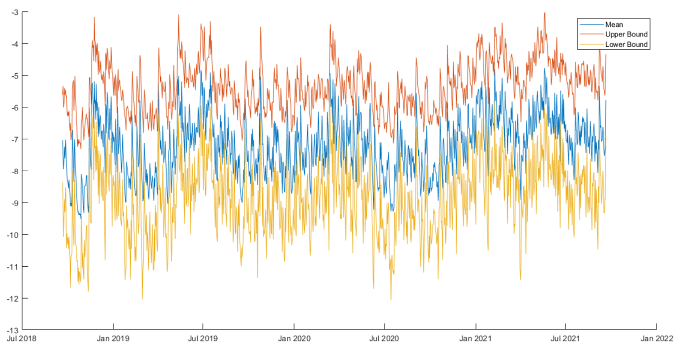


FIGURE 6.10: Filtered distribution of the hidden process X of the model introduced in equations (6.2.1)-(6.2.2) and where the proxy of attention is the number of Wikipedia views. The mean and the 95% confidence interval of the posterior distribution are shown.

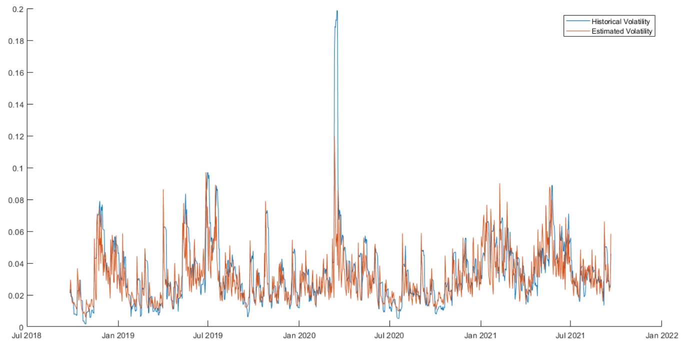


FIGURE 6.11: Weekly historical volatility and estimated volatility of the model defined in equations (6.2.1)-(6.2.2) and where the proxy of attention is the number of Wikipedia views.

weekly historical volatility. From equation (6.2.4), we have that the conditional standard deviation of the return Y_t given X_t and A_t is:

$$\sqrt{Var[Y_t|X_t, A_t]} = e^{0.5A_t+0.5X_t}. \tag{6.3.4}$$

In Figure 6.15, the historical weekly volatility and the conditional standard deviation defined in equation 6.3.4 are shown. Again, the proposed model captures the spikes better than the benchmark model. In this case, the residuals are defined as

$$\epsilon_t = \exp\left\{-\frac{A_t}{2} - \frac{X_t}{2}\right\} Y_t \text{ for } t = 1, \dots, T.$$

For checking the normality assumption of the residuals we compute the p–value of the Kolmogorov-Smirnov test. In the case the p–value has a value of 0.0280. So, we can reject the normality assumption with a confidence level of 5% but we cannot reject the assumption with a confidence level of 1%.

For the unique number of active addresses, we proceed in a similar way and we compare the conditional standard deviation of the model defined in equation (6.3.4) with the weekly historical volatility. This is shown in Figure 6.16. We observe that the results are in line with previous results shown in this section and in Section 6.3.2.

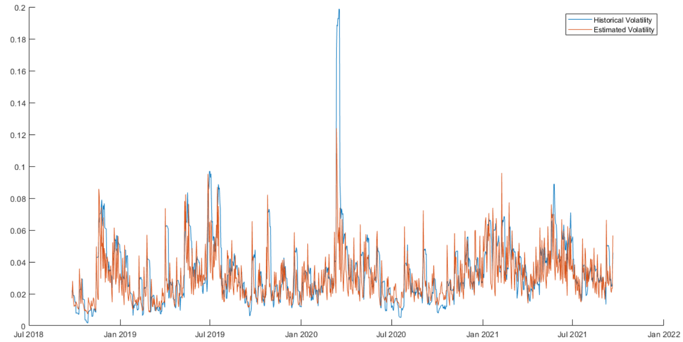


FIGURE 6.12: Weekly historical volatility and estimated volatility of the model defined in equations (6.2.1)-(6.2.2) and where the proxy of attention is the number of unique active addresses.

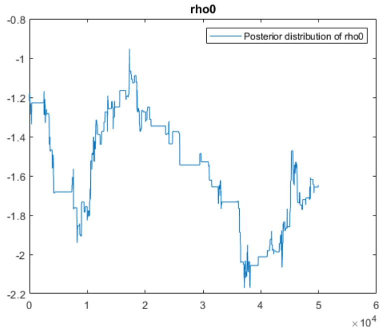
For this case, the p-value of the Kolmogorov-Smirnov test is 0.0177.

Remark 6.3.4. *The prior distributions of the parameters used in the particle marginal Metropolis-Hastings algorithm are the same as the ones used in the numerical example presented in Section 6.2.2.*

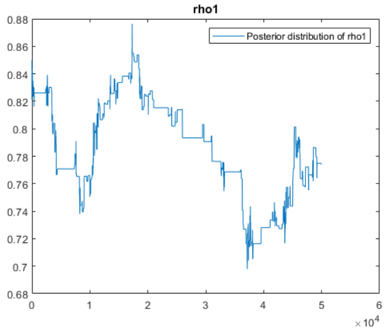
6.4 Conclusion and future work

In this chapter, we suggested models in which the volatility is explained by an unobserved process and by market attention. We observe that the volatility of the proposed models explains the spikes that appear on the historical volatility better than the benchmark model introduced in Example 6.1.1. These results need to be expanded by incorporating the study of more proxies for market attention, aside from those that we studied in this chapter. Also, would be interested to study models that introduce jumps in price and volatility structure.

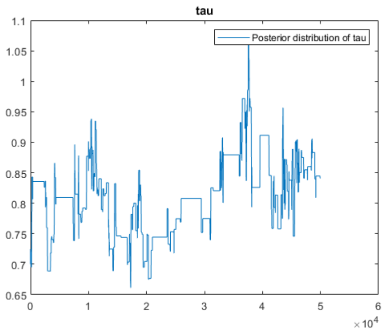
Another question that remains is how we can use the models introduced in Section 6.2 for pricing options. One way is the construction of the stochastic volatility models. For example, Barunik, Chen, and Vecer (2019) proposed a continuous stochastic volatility model in which the sentiment is incorporated in the structure of the volatility. Another way consists of assuming that the unobserved process is stochastic, but no underlying process is specified. This approach is applied in uncertain volatility models, introduced by Avellaneda, Levy, and Parás (1995). For example, in the model introduced in Section 6.2.2, instead of assuming that the hidden process X follows an autoregressive process, we will assume that X is bounded on an interval and X does not follow a specific process. One of the benefits of introducing uncertainty into the hidden process is that the option prices can be specified in terms of the bid and ask prices. Instead of having just one price, we can have an interval of prices. So, we can have a valid range of Bitcoin option prices which are also influenced by market attention.



(A) Estimation of the parameter ρ_0.



(B) Estimation of the parameter ρ_1.



(C) Estimation of the parameter τ.

FIGURE 6.13: Posterior distribution of the parameters of the model defined in equations (6.2.4)-(6.2.5) and where the proxy of attention is the number of Wikipedia views.

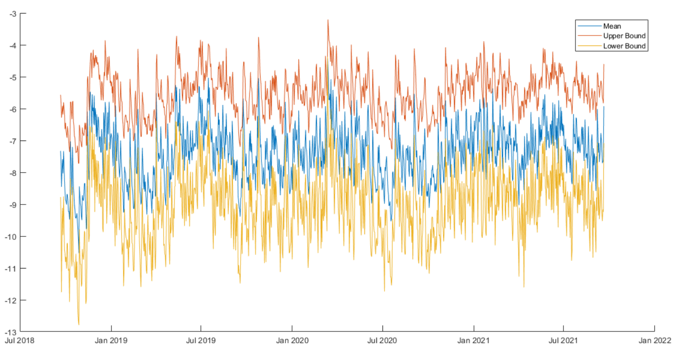


FIGURE 6.14: Filtered distribution of the hidden process X of the model introduced in equations (6.2.4)-(6.2.5) and where the proxy of attention is the number of Wikipedia views. The mean and the 95% confidence interval of the posterior distribution are shown.

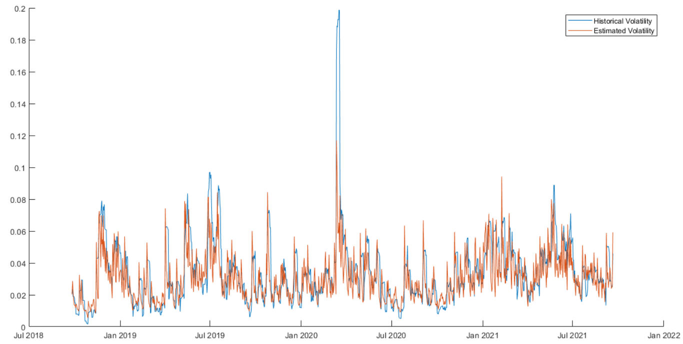


FIGURE 6.15: Weekly historical volatility and estimated volatility of the model introduced in equations (6.2.4)-(6.2.5) and where the proxy of attention is the number of Wikipedia views.

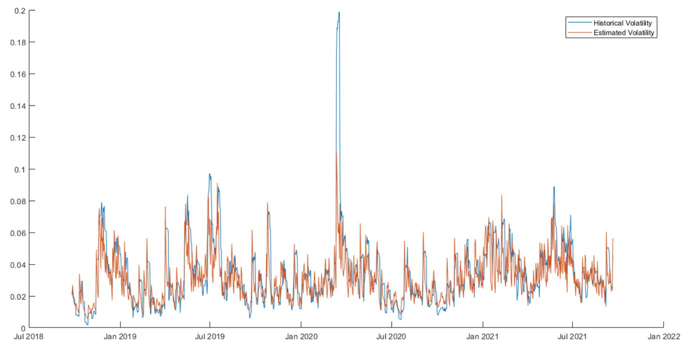


FIGURE 6.16: Weekly historical volatility and estimated volatility of the model introduced in equations (6.2.4)-(6.2.5) and where the proxy of attention is the number of unique active addresses.

Appendix A

Correlated model

In this section, we focus ourselves on a modified version of the model proposed in Section 3.7.2. This new model will take into account the correlation between the underlying Brownian motions of the price and interest processes. Let us assume that under the risk neutral probability space $(\Omega, \mathcal{F}, \mathbb{Q})$ the logarithmic price process X satisfies the following equation:

$$\begin{aligned} X(t) &= x + rt - \frac{1}{2}\int_0^t \sigma_P^2 I(u-\tau)du \\ &+ \mathbb{1}_{(\tau,\infty)}(t)\left(\int_0^{t-\tau} \sigma_P \rho \sqrt{I(u)}dW_I^*(u)\right) \\ &+ \int_0^t \sigma_P \sqrt{1-\rho^2}\sqrt{I(u-\tau)}dW_P^*(u) \end{aligned}$$

with $X(0) = x \in \mathbb{R}$, the market attention process is a Cox-Ingersoll-Ross process and satisfies the following stochastic differential equation

$$dI(t) = \tilde{a}_I(\tilde{b}_I - I(t))dt + \sigma_I \sqrt{I(t)}dW_I^*(t) \text{ when } t > 0 \text{ with } I(t) = \phi^I(t),\ t \in [-L, 0].$$

where $\tilde{b}_I \in \mathbb{R}$, $L, \sigma_P, \sigma_I, \tilde{a}_I > 0$, $\tau \in [0, L]$, $\rho \in (-1, 1)$, $r \geq 0$ is the known interest rate, W_P^* and W_I^* are two independent Brownian motions, we have the condition $\frac{2\tilde{a}_I\tilde{b}_I}{\sigma_I^2} \geq 1$ and $\phi^I : \mathbb{R} \to (0, \infty)$ is a deterministic continuous function.

If we assume that the discounted price process is a martingale, then we have that the price of a call option with expiry date T and strike K is

$$C(0) = E_{\mathbb{Q}}\left[e^{-rT}\left(e^{X(T)} - K\right)^+\right].$$

When $T \leq \tau$, the formula for the price of the call option is similar to the price formula obtained in Section 3.7.3. In the case when $T > \tau$ we have to price the option using the characteristic function of the random variable $X(T)$.

Theorem A.0.1. *Let $\lambda \in \mathbb{R}$, the characteristic function of $X(T)$ can be expressed as*

$$\Phi^{X(T)}(\lambda) = E_{\mathbb{Q}}\left[e^{i\lambda X(T)}\right] = e^{i\lambda rT - \frac{\sigma_P^2}{2}\epsilon^0(i\lambda + (1-\rho^2)\lambda^2)}\Phi^{Z(T-\tau)}(\lambda),$$

where

$$
\begin{aligned}
\epsilon^0 &= \int_{-\tau}^{0} \phi^I(u)du, \\
\Phi^{Z(t)}(\lambda) &= \exp\{i\lambda x\} \\
& \exp\left\{b_V a_V \sigma_V^{-2}\left[(a_V - \rho\sigma_V\lambda i - d)t - 2log\left(\frac{1-ge^{-dt}}{1-g}\right)\right]\right\} \\
& \exp\left\{V(0)\sigma_V^{-2}\left[\frac{(a_V - \rho\sigma_V\lambda i - d)(1-e^{-dt})}{1-ge^{-dt}}\right]\right\},
\end{aligned}
$$

with

$$
\begin{aligned}
d &= \left((\rho\sigma_V\lambda i - a_V)^2 - \sigma_V^2(-i\lambda - \lambda^2)\right)^{1/2}, \\
g &= \frac{a_V - \rho\sigma_V\lambda i - d}{a_V - \rho\sigma_V\lambda i + d}, \\
V(0) &= \sigma_P^2\phi^I(0), \\
a_V &= \tilde{a}_I, \\
b_V &= \sigma_P^2\tilde{b}_I, \\
\sigma_V &= \sigma_P\sigma_I.
\end{aligned}
$$

Proof. The logarithmic price process at time T can be expressed as

$$
\begin{aligned}
X(T) &= x + rT - \frac{1}{2}\int_0^T \sigma_P^2 I(u-\tau)du + \int_0^{T-\tau} \sigma_P\rho\sqrt{I(u)}dW_I^*(u) \\
&+ \int_0^T \sigma_P\sqrt{1-\rho^2}\sqrt{I(u-\tau)}dW_P^*(u) \\
&= x + rT - \frac{1}{2}\int_0^\tau \sigma_P^2 I(u-\tau)du - \frac{1}{2}\int_\tau^T \sigma_P^2 I(u-\tau)du \\
&+ \int_0^{T-\tau} \sigma_P\rho\sqrt{I(u)}dW_I^*(u) \\
&+ \int_0^\tau \sigma_P\sqrt{1-\rho^2}\sqrt{I(u-\tau)}dW_P^*(u) + \int_\tau^T \sigma_P\sqrt{1-\rho^2}\sqrt{I(u-\tau)}dW_P^*(u) \\
&= x + rT - \frac{1}{2}\int_0^\tau \sigma_P^2 \phi^I(u-\tau)du - \frac{1}{2}\int_0^{T-\tau} \sigma_P^2 I(u)du \\
&+ \int_0^{T-\tau} \sigma_P\rho\sqrt{I(u)}dW_I^*(u) \\
&+ \int_0^\tau \sigma_P\sqrt{1-\rho^2}\sqrt{I(u-\tau)}dW_P^*(u) + \int_\tau^T \sigma_P\sqrt{1-\rho^2}\sqrt{I(u-\tau)}dW_P^*(u) \\
&= rT - \frac{1}{2}\int_0^\tau \sigma_P^2 \phi^I(u-\tau)du + \int_0^\tau \sigma_P\sqrt{1-\rho^2}\sqrt{\phi^I(u-\tau)}dW_P^*(u) \\
&+ x - \frac{1}{2}\int_0^{T-\tau} \sigma_P^2 I(u)du + \int_0^{T-\tau} \sigma_P\rho\sqrt{I(u)}dW_I^*(u) \\
&+ \int_0^{T-\tau} \sigma_P\sqrt{1-\rho^2}\sqrt{I(u)}dB_P^*(u),
\end{aligned}
$$

where the last equality comes from the application of Theorem 1.2.1 and $B_P^*(t) = W_P^*(t+\tau) - W_P^*(\tau)$ for $t \geq 0$. Let us define the process Z as

$$Z(t) = x - \frac{1}{2}\int_0^t \sigma_P^2 I(u)du + \int_0^t \sigma_P\rho\sqrt{I(u)}dW_I^*(u) + \int_0^t \sigma_P\sqrt{1-\rho^2}\sqrt{I(u)}dB_P^*(u)$$

for $t \geq 0$. The process Z satisfies the stochastic differential equation

$$dZ(t) = -\frac{1}{2}\sigma_P^2 I(t)dt + \sigma_P\rho\sqrt{I(t)}dW_I^* + \sigma_P\sqrt{1-\rho^2}\sqrt{I(t)}dB_P^*(t) \text{ with } Z(0) = x.$$

The random variable $X(T)$ can be expressed as

$$X(T) = rT - \frac{1}{2}\int_0^\tau \sigma_P^2\phi^I(u-\tau)du + \int_0^\tau \sigma_P\sqrt{1-\rho^2}\sqrt{\phi^I(u-\tau)}dW_P^*(u) + Z(T-\tau).$$

Because $B_P^*(t)$ is independent of $\mathcal{F}_\tau^{W_P^*}$ for all $t \geq 0$ then we have that $Z(t)$ is independent of $\mathcal{F}_\tau^{W_P^*}$ for all $t \geq 0$. Applying a similar argument as in Section 3.7.4, we have that

$$\Phi_{X(T)}(\lambda) = e^{i\lambda rT - \frac{\sigma_P^2}{2}\epsilon^0(i\lambda+(1-\rho^2)\lambda^2)}E_{\mathbb{Q}}\left[e^{i\lambda Z(T-\tau)}\right],$$

where $\epsilon^0 = \int_{-\tau}^0 \phi^I(u)du$. It is clear that for computing the characteristic function of $X(T)$, we have to compute the characteristic function $Z(T-\tau)$. If we define the process V as $V(t) = \sigma_P^2 I(t)$ for $t \geq 0$, then $Z(T-\tau)$ can be written as

$$\begin{aligned} Z(T-\tau) &= x - \frac{1}{2}\int_0^{T-\tau} V(u)du + \int_0^{T-\tau} \rho\sqrt{V(u)}dW_I^*(u) &\text{(A.0.1)}\\ &+ \int_0^{T-\tau} \sqrt{1-\rho^2}\sqrt{V(u)}dB_P^*(u). &\text{(A.0.2)} \end{aligned}$$

By the Itô formula, we have that V satisfies the stochastic differential equation

$$dV(t) = a_V(b_V - V(t))dt + \sigma_V\sqrt{V(t)}dW_I^*(t) \text{ with } V(0) = \sigma_P^2\phi^I(0),$$

where

$$\begin{aligned} a_V = \tilde{a}_I, \qquad & b_V = \sigma_P^2\tilde{b}_I, \\ \text{and} \qquad & \sigma_V = \sigma_P\sigma_I. \end{aligned}$$

We also have that

$$\frac{2a_V b_V}{\sigma_V^2} = \frac{2\tilde{a}_I\tilde{b}_I}{\sigma_I^2} \geq 1.$$

So the Feller condition is satisfied and the process V is greater than zero with probability one. The process (Z, V) is the correlated Heston model (Heston, 1993) and the characteristic function of $Z(t)$ is known and it has the form

$$\begin{aligned} \Phi^{Z(t)}(\lambda) &= E_{\mathbb{Q}}\left[e^{i\lambda Z(t)}\right] \\ &= \exp\{i\lambda x\} \\ &\quad \exp\left\{b_V a_V \sigma_V^{-2}\left[(a_V - \rho\sigma_V\lambda i - d)t - 2log\left(\frac{1-ge^{-dt}}{1-g}\right)\right]\right\} \\ &\quad \exp\left\{V(0)\sigma_V^{-2}\left[\frac{(a_V - \rho\sigma_V\lambda i - d)(1-e^{-dt})}{1-ge^{-dt}}\right]\right\}, \end{aligned}$$

with

$$\begin{aligned} d &= \left((\rho\sigma_V\lambda i - a_V)^2 - \sigma_V^2(-i\lambda - \lambda^2)\right)^{1/2}, \\ g &= \frac{a_V - \rho\sigma_V\lambda i - d}{a_V - \rho\sigma_V\lambda i + d} \end{aligned}$$

(Madan, Reyners, & Schoutens, 2019). We have just shown that

$$\Phi^{X(T)}(\lambda) = e^{i\lambda rT - \frac{\sigma_P^2}{2}\epsilon^0(i\lambda + (1-\rho^2)\lambda^2)}\Phi^{Z(T-\tau)}(\lambda).$$

□

This section is just a discussion about how a model with correlation could be constructed and how we could price options with it. More work needs to be done with the model proposed in this section. For example, we still have to prove the existence of the risk neutral measure and the martingale property of the discounted stock price.

Appendix B

Change of measure

In this section, we will show some results that are stated in Section 4.3.1. These results are quite important because they state that some distributional properties that we have in the physical measure $\mathbb{P}$ are maintained in the equivalent measures $\mathbb{Q}^*$ and $\mathbb{Q}$.

B.1 Change of measure from $\mathbb{P}$ to $\mathbb{Q}^*$

Let us consider the change of measure that is defined in Proposition 4.3.1. We would like to show that under the probability $\mathbb{Q}^*$ the process W_P is still a Brownian motion and that the processes Z_I and W_P are still independent. Let us first show that W_P is still a Brownian motion.

Proposition B.1.1. *The process W_P is a Brownian motion under the probability measure $\mathbb{Q}^*$.*

Proof. To show that W_P is a Brownian motion, we will have to show that the following holds true:

1. $W_P(0) = 0$ almost surely.
2. W_P has independent increments.
3. $W_P(t) - W_P(s) \sim N(0, t-s)$ for $0 \leq s \leq t$.
4. W_P has almost surely continuous sample paths.

First, let us check that $W_P(0) = 0$ almost surely. To that end, define the set $\mathcal{N}_0$ as

$$\mathcal{N}_0 = \{\omega \in \Omega : W_P(0, \omega) = 0\}.$$

Since $\mathbb{P}$ and $\mathbb{Q}^*$ are equivalent probability measures, we have that

$$\mathbb{Q}^*(\mathcal{N}_0) = \mathbb{P}(\mathcal{N}_0) = 1.$$

To show the normality property, let us compute the characteristic function of $W_P(t) - W_P(s)$ for $0 \leq s \leq t$. Let $u \in \mathbb{R}$ then the characteristic function of $W_P(t) - W_P(s)$ is:

$$\begin{aligned} E_{\mathbb{Q}^*}\left[e^{iu(W_P(t)-W_P(s))}\right] &= E_{\mathbb{P}}\left[Z^*(T)e^{iu(W_P(t)-W_P(s))}\right] \\ &= E_{\mathbb{P}}\left[Z^*(T)\right] E_{\mathbb{P}}\left[e^{iu(W_P(t)-W_P(s))}\right] \\ &= e^{-\frac{1}{2}u^2(t-s)}. \end{aligned} \tag{B.1.1}$$

So, the random variable $W_P(t) - W_P(s)$ has the characteristic function of a normal random variable with mean 0 and variance $t - s$.

To prove the independence of increments, let us take the times $0 \leq t_0 < t_1 < \ldots < t_n \leq T$ and define the increments

$$\Delta^i W_P = W(t_i) - W(t_{i-1}) \text{ for } i = 1, \ldots, n.$$

Let $u_1, u_2, \ldots, u_n \in \mathbb{R}$ then the joint characteristic function of $\left(\Delta^1 W_P, \ldots, \Delta^n W_P\right)$ can be written as

$$\begin{aligned} E_{\mathbb{Q}^*}\left[e^{\sum_{j=1}^n iu_j \Delta^j W_P}\right] &= E_{\mathbb{P}}\left[Z^*(T) e^{\sum_{j=1}^n iu_j \Delta^j W_P}\right] \\ &= E_{\mathbb{P}}\left[Z^*(T)\right] E_{\mathbb{P}}\left[e^{\sum_{j=1}^n iu_j \Delta^j W_P}\right] \\ &= \prod_{j=1}^n E_{\mathbb{P}}\left[e^{iu_j \Delta^j W_P}\right]. \end{aligned} \tag{B.1.2}$$

From equation (B.1.1) we have that:

$$E_{\mathbb{P}}\left[e^{iu_j \Delta^j W_P}\right] = E_{\mathbb{Q}^*}\left[e^{iu_j \Delta^j W_P}\right] \text{ for } j = 1, \ldots, n. \tag{B.1.3}$$

Combining equations (B.1.2) and (B.1.3), we arrive at

$$E_{\mathbb{Q}^*}\left[e^{\sum_{j=1}^n iu_j \Delta^j W_P}\right] = \prod_{i=1}^n E_{\mathbb{Q}^*}\left[e^{iu_j \Delta^j W_P}\right].$$

Hence the increments are independent under the measure $\mathbb{Q}^*$.

Lastly, we will show that W_P has almost surely continuous sample paths, to that end let us define the set $\mathcal{N}_C$ as

$$\mathcal{N}_C = \{\omega \in \Omega : \text{ the map } t \to W_P(t, \omega) \text{ is continuous}\}.$$

The probability of $\mathcal{N}_C$ under $\mathbb{Q}$ can be written as

$$\begin{aligned} \mathbb{Q}^*(\mathcal{N}_C) &= E_{\mathbb{Q}^*}\left[\mathbb{1}_{\mathcal{N}_C}\right] \\ &= E_{\mathbb{P}}\left[\mathbb{1}_{\mathcal{N}_C} Z^*(T)\right] \\ &= E_{\mathbb{P}}\left[\mathbb{1}_{\mathcal{N}_C}\right] E_{\mathbb{P}}\left[Z^*(T)\right] \\ &= \mathbb{P}(\mathcal{N}_C) = 1. \end{aligned}$$

□

We have just shown that W_P is a Brownian motion under $\mathbb{Q}^*$, but we still have to prove the independence of W_P and Z_I under $\mathbb{Q}^*$. But before this, let us define the concept of independence between two stochastic processes.

Definition B.1.1. *(Lapidoth, 2017, Definition 25.2.3) Two stochastic processes $(X(t))_{t\in\mathbb{R}}$ and $(Y(t))_{t\in\mathbb{R}}$ defined on the same probability space are said to be independent if for every $n \in \mathbb{N}$ and any choice of $t_1, \ldots t_n \in \mathbb{R}$ the random vectors $(X(t_1), \ldots, X(t_n))$ and $(Y(t_1), \ldots, Y(t_n))$ are independent.*

Proposition B.1.2. *The processes W_P and Z_I are independent under the probability measure $\mathbb{Q}^*$.*

Proof. Let us consider the times $0 \leq t_1 \leq \ldots \leq t_n \leq T$ and let $u_1, \ldots u_n, v_1, \ldots v_n \in \mathbb{R}$, then the joint characteristic function of $(W_P(t_1), \ldots, W_P(t_n), Z_I(t_1), \ldots, Z_I(t_n))$

can be written as

$$\begin{aligned}E_{\mathbb{Q}^*}\left[e^{\sum_{j=1}^n iu_jW_P(t_j)}e^{\sum_{j=1}^n iv_jZ_I(t_j)}\right] &= E_{\mathbb{P}}\left[Z^*(T)e^{\sum_{j=1}^n iu_jW_P(t_j)}e^{\sum_{j=1}^n iv_jZ_I(t_j)}\right]\\ &= E_{\mathbb{P}}\left[e^{\sum_{j=1}^n iu_jW_P(t_j)}\right]\\ &\quad E_{\mathbb{P}}\left[Z^*(T)e^{\sum_{j=1}^n iv_jZ_I(t_j)}\right]\\ &= E_{\mathbb{P}}\left[e^{\sum_{j=1}^n iu_jW_P(t_j)}\right]E_{\mathbb{Q}^*}\left[e^{\sum_{j=1}^n iv_jZ_I(t_j)}\right].\end{aligned}$$

From Proposition B.1.1 we have that

$$E_{\mathbb{P}}\left[e^{\sum_{j=1}^n iu_jW_P(t_j)}\right] = E_{\mathbb{Q}^*}\left[e^{\sum_{j=1}^n iu_jW_P(t_j)}\right].$$

At the end we can write

$$E_{\mathbb{Q}^*}\left[e^{\sum_{j=1}^n iu_jW_P(t_j)}e^{\sum_{j=1}^n iv_jZ_I(t_j)}\right] = E_{\mathbb{Q}^*}\left[e^{\sum_{j=1}^n iu_jW_P(t_j)}\right]E_{\mathbb{Q}^*}\left[e^{\sum_{j=1}^n iv_jZ_I(t_j)}\right]$$

as required. □

B.2 Change of measure from $\mathbb{Q}^*$ to $\mathbb{Q}$

In this case, we are interested in the change of measure defined in Section 4.3.1 from the measure $\mathbb{Q}^*$ to the measure $\mathbb{Q}$. Recall from equations (4.3.5)–(4.3.6) that the measure $\mathbb{Q}$ is defined as

$$\mathbb{Q}(A) = \int_A Z(T)d\mathbb{Q}^* \text{ for } A \in \mathcal{F},$$

where the process Z is defined as

$$Z(t) \quad = \quad \exp\left\{-\int_0^t \theta_P(s)dW_P(s) - \frac{1}{2}\int_0^t \theta_P^2(s)ds\right\} \text{ for } t \in [0,T],$$

where θ_P is an adapted process that satisfies

$$\theta_P(t) = \frac{\mu + \frac{\sigma_P^2}{2}I^-(t-\tau) - r}{\sigma_P\sqrt{I^-(t-\tau)}} \text{ for } t \in [0,T].$$

We have shown in Section 4.3.1 that the process W_P^* that is defined as

$$W_P^*(t) = W_P(t) + \int_0^t \theta_P(s)ds \text{ for } t \in [0,T],$$

is a Brownian motion under the measure $\mathbb{Q}$. We still have to show that Z_I is a Lévy process with the same distribution as under $\mathbb{Q}^*$ and that the processes W_P and Z_I are independent under $\mathbb{Q}$.

Proposition B.2.1. *Under the probability measure $\mathbb{Q}$, the process Z_I is a Lévy process with Levy triplet $(\tilde{\gamma}, 0, \tilde{v})$ given by equations (4.3.1)–(4.3.2), that is the distribution of Z_I is the same under $\mathbb{Q}$ as under $\mathbb{Q}^*$.*

Proof. First we will show that the distribution of $Z_I(t)$ under the measure $\mathbb{Q}$ is the same as the distribution of $Z_I(t)$ under the equivalent measure $\mathbb{Q}^*$ for $t \in [0,T]$. Take any $u \in \mathbb{R}$. If we condition with respect to the σ-algebra $\mathcal{F}_T^{Z_I}$ and apply Proposition 1.4.2, then the characteristic function of $Z_I(t)$ under $\mathbb{Q}$ can be expressed as

$$\begin{aligned} E_{\mathbb{Q}}\left[e^{iuZ_I(t)}\right] &= E_{\mathbb{Q}^*}\left[Z(T)e^{iuZ_I(t)}\right] \\ &= E_{\mathbb{Q}^*}\left[e^{-\frac{1}{2}\int_0^T \theta_P^2(s)ds}e^{iuZ_I(t)}E_{\mathbb{Q}^*}\left[e^{-\int_0^T \theta_P(s)dW_P(s)}\middle|\mathcal{F}_T^{Z_I}\right]\right] \\ &= E_{\mathbb{Q}^*}\left[e^{iuZ_I(t)}\right]. \end{aligned} \tag{B.2.1}$$

From equation (B.2.1) we have that under the measure $\mathbb{Q}$ the random variable $Z_I(t)$ has the same characteristic function as $Z_I(t)$ under the measure $\mathbb{Q}^*$. Hence, if we show that Z_I is a Lévy process under the measure $\mathbb{Q}$, then Z_I will have Lévy triplet $(\tilde{\gamma}, 0, \tilde{v})$. To show that Z_I is a Lévy process, we need to show that

1. $Z_I(0) = 0$ almost surely.
2. Z_I has independent and stationary increments.
3. Z_I is stochastically continuous.

Let us first show that $Z_I(0) = 0$ to that end, let us define the the set

$$\mathcal{N}_0 = \{\omega \in \Omega : Z_I(0,\omega) = 0\}.$$

Because $\mathbb{Q}$ and $\mathbb{Q}^*$ are equivalent probability measures, we have that:

$$\mathbb{Q}(\mathcal{N}_0) = \mathbb{Q}^*(\mathcal{N}_0) = 1.$$

For the stationarity property, by conditioning with respect to $\mathcal{F}_T^{Z_I}$ and applying Proposition 1.4.2, we have that the characteristic function of $Z_I(t) - Z_I(s)$ for $t \geq s$ can be written for all $u \in \mathbb{R}$ as

$$\begin{aligned} E_{\mathbb{Q}}\left[e^{iu(Z_I(t)-Z_I(s))}\right] &= E_{\mathbb{Q}^*}\left[Z(T)e^{iu(Z_I(t)-Z_I(s))}\right] \\ &= E_{\mathbb{Q}^*}\left[e^{iu(Z_I(t)-Z_I(s))}\right] \\ &= E_{\mathbb{Q}^*}\left[e^{iuZ_I(t-s)}\right] \\ &= E_{\mathbb{Q}}\left[e^{iuZ_I(t-s)}\right], \end{aligned}$$

where the last equality comes from equation (B.2.1).

For the independence property, let us consider a $n \in \mathbb{N}$ and the time steps $0 \leq t_1 < t_2 < \ldots < t_n \leq T$ and define the following increments of Z_I:

$$\Delta^j Z_I = Z_I(t_j) - Z_I(t_{j-1}) \text{ for } j = 1, \ldots, n.$$

For the values $u_1, \ldots, u_n \in \mathbb{R}$ we would like to compute the characteristic function of $(\Delta^1 Z_I, \ldots, \Delta^n Z_I)$. Again, conditioning with respect to the σ-algebra $\mathcal{F}_T^{Z_I}$ and using

Proposition 1.4.2, we can write the characteristic function of the increments as

$$\begin{aligned} E_{\mathbb{Q}}\left[e^{\sum_{j=1}^n iu_j \Delta^j Z_I}\right] &= E_{\mathbb{Q}^*}\left[Z(T)e^{\sum_{j=1}^n iu_j \Delta^j Z_I}\right] \\ &= E_{\mathbb{Q}^*}\left[e^{\sum_{j=1}^n iu_j \Delta^j Z_I}\right] \\ &= \prod_{j=1}^n E_{\mathbb{Q}^*}\left[e^{iu_j \Delta^j Z_I}\right] \\ &= \prod_{j=1}^n E_{\mathbb{Q}}\left[e^{iu_j \Delta^j Z_I}\right], \end{aligned}$$

where the last equality comes from equation (B.2.1).

Finally, we would like to show that Z_I is stochastically continuous. Let us take an $\epsilon > 0$ and an $s \geq 0$ then we have that

$$\begin{aligned} \lim_{t\to s} \mathbb{Q}\left(|Z_I(t) - Z_I(s)| > \epsilon\right) &= \lim_{t\to s} E_{\mathbb{Q}}\left[\mathbb{1}_{|Z_I(t)-Z_I(s)|>\epsilon}\right] \\ &= \lim_{t\to s} E_{\mathbb{Q}^*}\left[Z(T)\mathbb{1}_{|Z_I(t)-Z_I(s)|>\epsilon}\right]. \end{aligned}$$

Again conditioning with respect to $\mathcal{F}_T^{Z_I}$ and using Proposition 1.4.2, we have that

$$\begin{aligned} \lim_{t\to s} \mathbb{Q}\left(|Z_I(t) - Z_I(s)| > \epsilon\right) &= \lim_{t\to s} E_{\mathbb{Q}^*}\left[\mathbb{1}_{|Z_I(t)-Z_I(s)|>\epsilon}\right] \\ &= \lim_{t\to s} \mathbb{Q}^*\left(|Z_I(t) - Z_I(s)| > \epsilon\right) = 0, \end{aligned}$$

where the last equality comes from the fact that Z_I is a Lévy process under the measure $\mathbb{Q}^*$.

□

Before showing the independence between Z_I and W_P^* under $\mathbb{Q}$, let us show the following result.

Lemma B.2.1. *Let us consider the times $0 \leq t_1 < t_2 < \ldots, t_n \leq T$ and the values $u_1, \ldots, u_n \in \mathbb{R}$, then:*

$$\begin{aligned} &E_{\mathbb{Q}^*}\left[e^{\sum_{j=1}^n iu_j W_P(t_j) - \int_0^T \theta_P(s)dW_P(s)} \middle| \mathcal{F}_T^{Z_I}\right] \\ &\qquad = E_{\mathbb{Q}}\left[e^{i\sum_{j=1}^n u_j W_P^*(t_j)}\right] e^{\frac{1}{2}\int_0^T \theta_P^2(s)ds} e^{-i\sum_{j=1}^n u_j \int_0^{t_j} \theta_P(s)ds}. \end{aligned}$$

Proof. Let us define the deterministic function $f : \mathbb{R} \to \mathbb{R}$ as:

$$f(s) = \sum_{j=1}^n \mathbb{1}_{(0,t_j]}(s)u_j, \text{ for } s \in \mathbb{R}.$$

Using the function f, we can write:

$$
\begin{aligned}
\sum_{j=1}^{n} iu_j W_P(t_j) - \int_0^T \theta_P(s)dW_P(s) &= \sum_{j=1}^{n} i\int_0^T \mathbb{1}_{(0,t_j]}(s)u_j dW_P(s) \\
&\quad - \int_0^T \theta_P(s)dW_P(s) \\
&= i\int_0^T \left(\sum_{j=1}^{n} \mathbb{1}_{(0,t_j]}(s)u_j\right) dW_P(s) \\
&\quad - \int_0^T \theta_P(s)dW_P(s) \\
&= i\int_0^T f(s)dW_P(s) \\
&\quad - \int_0^T \theta_P(s)dW_P(s) \qquad \text{(B.2.2)}
\end{aligned}
$$

Using Equation B.2.2 we can write

$$
\begin{aligned}
&E_{\mathbb{Q}^*}\left[e^{\sum_{j=1}^{n} iu_j W_P(t_j) - \int_0^T \theta_P(s)dW_P(s)} \middle| \mathcal{F}_T^{Z_I}\right] \\
&\qquad = E_{\mathbb{Q}^*}\left[e^{i\int_0^T f(s)dW_P(s) - \int_0^T \theta_P(s)dW_P(s)} \middle| \mathcal{F}_T^{Z_I}\right]. \qquad \text{(B.2.3)}
\end{aligned}
$$

Firstly, let us take $\lambda_1, \lambda_2 \in \mathbb{R}$ and we will compute the joint characteristic function of $\int_0^T f(s)dW_P(s)$ and $\int_0^T \theta_P(s)dW_P(s)$ given $\mathcal{F}_T^{Z_I}$ as:

$$
\begin{aligned}
&E\left[e^{i\lambda_1\int_0^T f(s)dW_P(s) + i\lambda_2\int_0^T \theta_P(s)dW_P(s)} \middle| \mathcal{F}_T^{Z_I}\right] \\
=&E\left[e^{i\int_0^T (\lambda_1 f(s) + \lambda_2\theta_P(s))dW_P(s)} \middle| \mathcal{F}_T^{Z_I}\right] \\
=&e^{\frac{-1}{2}\int_0^T (\lambda_1 f(s) + \lambda_2\theta_P(s))^2 ds} \\
=&e^{-\frac{\lambda_1^2}{2}\int_0^T f^2(s)ds - \frac{\lambda_2^2}{2}\int_0^T \theta_P^2(s)ds - \lambda_1\lambda_2\int_0^T f(s)\theta_P(s)ds}, \qquad \text{(B.2.4)}
\end{aligned}
$$

where the second equality comes from the use of Proposition 1.4.2. Equation (B.2.4) is the characteristic function of a bi-variate normal distribution. Let us define the random variables X and Y as:

$$
\begin{aligned}
X &= \int_0^T f(s)dW_P(s), \\
Y &= \int_0^T \theta_P(s)dW_P(s),
\end{aligned}
$$

hence we can write

$$(X, Y)|\, \mathcal{F}_T^{Z_I} \sim N_M(0, \Sigma),$$

with

$$\Sigma = \begin{pmatrix} \sigma_X^2 & \sigma_{XY} \\ \sigma_{YX} & \sigma_Y^2 \end{pmatrix}$$

and

$$\begin{aligned}
\sigma_X^2 &= \int_0^T f^2(s)ds, \\
\sigma_Y^2 &= \int_0^T \theta^2(s)ds, \\
\sigma_{XY} &= \sigma_{YX} = \int_0^T \theta(s)f(s)ds.
\end{aligned}$$

Equation (B.2.3) can be written as:

$$\begin{aligned}
&E_{\mathbb{Q}^*}\left[e^{i\int_0^T f(s)dW_P(s)-\int_0^T \theta_P(s)dW_P(s)}\middle|\,\mathcal{F}_T^{Z_I}\right] \\
=&E_{\mathbb{Q}^*}\left[e^{iX-Y}\middle|\,\mathcal{F}_T^{Z_I}\right] \\
=&E_{\mathbb{Q}^*}\left[e^{-Y}\ E\left[e^{iX}\middle|\,\mathcal{F}_T^{Z_I}\vee\sigma(Y)\right]\middle|\,\mathcal{F}_T^{Z_I}\right].
\end{aligned} \tag{B.2.5}$$

By properties of the bi-variate normal distribution, we have that:

$$X\left|\mathcal{F}_T^{Z_I}\vee\sigma(Y)\right. \sim N(m_{X|Y}, V_{X|Y}^2),$$

where

$$\begin{aligned}
m_{X|Y} &= \frac{\sigma_{XY}}{\sigma_Y^2}Y, \\
V_{X|Y}^2 &= \sigma_X^2 - \frac{\sigma_{XY}}{\sigma_{Y^2}^2},
\end{aligned}$$

(Murphy, 2012, p. 111). Hence we can write

$$E_{\mathbb{Q}}^*\left[e^{iX}\middle|\,\mathcal{F}_T^{Z_I}\vee\sigma(Y)\right] = e^{i\frac{\sigma_{XY}}{\sigma_Y^2}Y-\frac{\sigma_X^2}{2}+\frac{\sigma_{XY}^2}{2\sigma_Y^2}}.$$

So equation (B.2.5) can be written as:

$$E_{\mathbb{Q}^*}\left[e^{iX-Y}\middle|\,\mathcal{F}_T^{Z_I}\right] = E_{\mathbb{Q}^*}\left[e^{Y\left(i\frac{\sigma_{XY}}{\sigma_Y^2}-1\right)}\middle|\,\mathcal{F}_T^{Z_I}\right]e^{-\frac{\sigma_X^2}{2}+\frac{\sigma_{XY}^2}{2\sigma_Y^2}}. \tag{B.2.6}$$

Since the random variable $Y|\mathcal{F}_T^{Z_I} \sim N(0, \sigma_Y^2)$ we can write:

$$\begin{aligned}
E_{\mathbb{Q}^*}\left[e^{Y\left(i\frac{\sigma_{XY}}{\sigma_Y^2}-1\right)}\middle|\,\mathcal{F}_T^{Z_I}\right] &= e^{\frac{\sigma_Y^2}{2}\left(i\frac{\sigma_{XY}}{\sigma_Y^2}-1\right)^2} \\
&= e^{\frac{\sigma_Y^2}{2}\left(-\frac{\sigma_{XY}^2}{\sigma_Y^4}+1-2i\frac{\sigma_{XY}}{\sigma_Y^2}\right)} \\
&= e^{-\frac{\sigma_{XY}^2}{2\sigma_Y^2}+\frac{\sigma_Y^2}{2}-i\sigma_{XY}}.
\end{aligned} \tag{B.2.7}$$

Combining equation (B.2.6) with equation (B.2.7), we arrive to:

$$\begin{aligned}
E_{\mathbb{Q}^*}\left[e^{iX-Y}\middle|\mathcal{F}_T^{Z_I}\right] &= e^{-\frac{\sigma_{XY}^2}{2\sigma_Y^2}+\frac{\sigma_Y^2}{2}-i\sigma_{XY}}e^{-\frac{\sigma_X^2}{2}+\frac{\sigma_{XY}^2}{2\sigma_Y^2}} \\
&= e^{\frac{\sigma_Y^2}{2}-\frac{\sigma_X^2}{2}-i\sigma_{XY}} \\
&= e^{\frac{1}{2}\int_0^T\theta_P^2(s)ds}e^{-\frac{1}{2}\int_0^T f^2(s)ds}e^{-i\int_0^T f(s)\theta_P(s)ds} \\
&= e^{\frac{1}{2}\int_0^T\theta_P^2(s)ds}e^{-\frac{1}{2}\int_0^T\left(\sum_{j=1}^n \mathbb{1}_{(0,t_j]}(s)u_j\right)^2 ds} \\
&\quad e^{-i\sum_{j=1}^n u_j\int_0^{t_j}\theta_P(s)ds}
\end{aligned}$$

Lastly, since W_P^* is a Brownian motion under the probability measure $\mathbb{Q}$ we have that:

$$\begin{aligned}
E_{\mathbb{Q}}\left[e^{\sum_{j=1}^n iu_jW_P^*(t_j)}\right] &= E_{\mathbb{Q}}\left[e^{i\int_0^T\left(\sum_{j=1}^n \mathbb{1}_{(0,t_j]}(s)u_j\right)dW_P^*(s)}\right] \\
&= e^{\frac{-1}{2}\int_0^T\left(\sum_{j=1}^n \mathbb{1}_{(0,t_j]}(s)u_j\right)^2 ds}.
\end{aligned}$$

We have just shown the desired result. □

Now we are ready to show the independence of the processes W_P^* and Z_I.

Proposition B.2.2. *The processes W_P^* and Z_I are independent under the equivalent measure $\mathbb{Q}$.*

Proof. For the times $0 \le t_1 < t_2 < \ldots, t_n \le T$ we will show that the random vectors $(W_P^*(t_1),\ldots,W_P^*(t_n))$ and $(Z_I(t_1),\ldots,Z_I^*(t_n))$ are independent under the measure $\mathbb{Q}$. To that end, we will compute the joint characteristic function of $(W_P^*(t_1),\ldots,W_P^*(t_n),Z_I(t_1),\ldots,Z_I^*(t_n))$.

Let us consider the values $u_1,\ldots,u_n,v_1,\ldots,v_n \in \mathbb{R}$, then the joint characteristic function of $(W_P^*(t_1),\ldots,W_P^*(t_n),Z_I(t_1),\ldots,Z_I^*(t_n))$ can be written as:

$$\begin{aligned}
&E_{\mathbb{Q}}\left[e^{\sum_{j=1}^n iu_jW_P^*(t_j)}e^{\sum_{j=1}^n iv_jZ_I(t_j)}\right] \\
&\quad = E_{\mathbb{Q}^*}\left[Z(T)e^{\sum_{j=1}^n iu_jW_P(t_j)}e^{\sum_{j=1}^n iu_j\int_0^{t_j}\theta_P(s)ds}e^{\sum_{j=1}^n iv_jZ_I(t_j)}\right] \\
&\quad = E_{\mathbb{Q}^*}\left[e^{-\frac{1}{2}\int_0^T\theta_P^2(s)ds-\int_0^T\theta_P(s)dW_P(s)}\right. \\
&\qquad \left. e^{\sum_{j=1}^n iu_jW_P(t_j)}e^{\sum_{j=1}^n iu_j\int_0^{t_j}\theta_P(s)ds}e^{\sum_{j=1}^n iv_jZ_I(t_j)}\right].
\end{aligned}$$

Conditioning with respect to the σ-algebra $\mathcal{F}_T^{Z_I}$ we have that:

$$\begin{aligned}
&E_{\mathbb{Q}}\left[e^{\sum_{j=1}^n iu_jW_P^*(t_j)}e^{\sum_{j=1}^n iv_jZ_I(t_j)}\right] \\
&\quad = E_{\mathbb{Q}^*}\left[e^{-\frac{1}{2}\int_0^T\theta_P^2(s)ds}e^{\sum_{j=1}^n iu_j\int_0^{t_j}\theta_P(s)ds}e^{\sum_{j=1}^n iv_jZ_I(t_j)}\right. \\
&\qquad \left. E_{\mathbb{Q}^*}\left[e^{-\int_0^T\theta_P(s)dW_P(s)}e^{\sum_{j=1}^n iu_jW_P(t_j)}\middle|\mathcal{F}_T^{Z_I}\right]\right].
\end{aligned}$$

Using Lemma B.2.1 we can write:

$$
\begin{aligned}
&E_{\mathbb{Q}}\left[e^{\sum_{j=1}^{n} iu_j W_P^*(t_j)} e^{\sum_{j=1}^{n} iv_j Z_I(t_j)}\right] \\
&\qquad = E_{\mathbb{Q}^*}\left[e^{-\frac{1}{2}\int_0^T \theta_P^2(s)ds} e^{\sum_{j=1}^{n} iu_j \int_0^{t_j} \theta_P(s)ds} e^{\sum_{j=1}^{n} iv_j Z_I(t_j)}\right. \\
&\qquad\qquad \left. E_{\mathbb{Q}}\left[e^{\sum_{j=1}^{n} iu_j W_P^*(t_j)}\right] e^{\frac{1}{2}\int_0^T \theta_P^2(s)ds} e^{-\sum_{j=1}^{n} iu_j \int_0^{t_j} \theta_P(s)} ds\right] \\
&\qquad = E_{\mathbb{Q}}\left[e^{\sum_{j=1}^{n} iu_j W_P^*(t_j)}\right] E_{\mathbb{Q}^*}\left[e^{\sum_{j=1}^{n} iv_j Z_I(t_j)}\right].
\end{aligned}
$$

Finally from Proposition B.2.1 we have that:

$$
E_{\mathbb{Q}^*}\left[e^{\sum_{j=1}^{n} iv_j Z_I(t_j)}\right] = E_{\mathbb{Q}}\left[e^{\sum_{j=1}^{n} iv_j Z_I(t_j)}\right].
$$

We have just proved that

$$
E_{\mathbb{Q}}\left[e^{\sum_{j=1}^{n} iu_j W_P^*(t_j)} e^{\sum_{j=1}^{n} iv_j Z_I(t_j)}\right] = E_{\mathbb{Q}}\left[e^{\sum_{j=1}^{n} iu_j W_P^*(t_j)}\right] E_{\mathbb{Q}}\left[e^{\sum_{j=1}^{n} iv_j Z_I(t_j)}\right]
$$

and hence the processes Z_I and W_P^* are independent under $\mathbb{Q}$. $\square$

Appendix C

Stochastic time changed process: conditional properties

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and let $(X(t))_{t\geq 0}$ be a Lévy process. Consider a non-negative and non-decreasing process $(T(t))_{t\geq 0}$ which has almost surely continuous sample paths. This process T is adapted with respect to a filtration $(\mathcal{G}_t)_{t\geq 0}$. We also assume that the process T and the filtration $(\mathcal{G}_t)_{t\geq 0}$ are independent of the process X. This process T will be used to change the time of the Lévy process X. We define the time changed Lévy process $(Z(t))_{t\geq 0}$ as

$$Z(t) = X(T(t)) \text{ for } t \geq 0.$$

We are interested in the distribution of the random variable $Z(t)$ given the σ-algebra $\mathcal{G}_t$. To that end, we will compute the conditional characteristic function of $Z(t)$ given $\mathcal{G}_t$, defined as

$$\Phi_{Z(t)}\left(\eta \middle| \mathcal{G}_t\right) = E\left[e^{i\eta Z(t)} \middle| \mathcal{G}_t\right], \text{ for all } \eta \in \mathbb{R}.$$

But before computing the conditional characteristic function, let us define the following function:

$$k_n(s) = \sum_{j=1}^{n^2} \frac{j}{n} \mathbb{1}_{\left[\frac{j-1}{n}, \frac{j}{n}\right)}(s), \text{ for } s \geq 0 \text{ and } n \in \mathbb{N}. \tag{C.0.1}$$

This function will be used repeatedly in the proofs of Proposition C.0.1 and Theorem C.1.1.

Proposition C.0.1. *The conditional characteristic function of $Z(t)$ given $\mathcal{G}_t$ satisfies:*

$$\Phi^{Z(t)}\left(\eta \middle| \mathcal{G}_t\right) = \exp\left\{T(t)\Psi_X(\eta)\right\}, \ \forall \eta \in \mathbb{R}, \tag{C.0.2}$$

where Ψ_X is the characteristic exponent of X.

Proof. This proof is inspired by the proof of Sato (1999, Theorem 30.1). Fix any $n \in \mathbb{N}$ and define the function k_n as in (C.0.1). Notice that we can write $e^{i\eta X(k_n(T(t)))}$ as

$$e^{i\eta X(k_n(T(t)))} = \sum_{j=1}^{n^2} e^{i\eta X\left(\frac{j}{n}\right)} \mathbb{1}_{\left[\frac{j-1}{n}, \frac{j}{n}\right)}(T(t)).$$

By linearity of the conditional expectation and because $\mathbb{1}_{\left[\frac{j-1}{n},\frac{j}{n}\right)}(T(t))$ is $\mathcal{G}_t$-measurable, we have that

$$\begin{aligned} E\left[e^{i\eta X(k_n(T(t)))}\middle|\,\mathcal{G}_t\right] &= \sum_{j=1}^{n^2} E\left[e^{i\eta X\left(\frac{j}{n}\right)}\mathbb{1}_{\left[\frac{j-1}{n},\frac{j}{n}\right)}(T(t))\middle|\,\mathcal{G}_t\right] \\ &= \sum_{j=1}^{n^2} \mathbb{1}_{\left[\frac{j-1}{n},\frac{j}{n}\right)}(T(t))E\left[e^{i\eta X\left(\frac{j}{n}\right)}\middle|\,\mathcal{G}_t\right]. \end{aligned}$$

By independence of the process X and the σ-algebra $\mathcal{G}_t$ we obtain for all $j = 1, 2, \dots, n^2$ that

$$\begin{aligned} E\left[e^{i\eta X\left(\frac{j}{n}\right)}\middle|\,\mathcal{G}_t\right] &= E\left[e^{i\eta X\left(\frac{j}{n}\right)}\right] \\ &= e^{\frac{j}{n}\Psi_X(\eta)}, \end{aligned}$$

where the last equality comes from Theorem 1.1.2. So we have just shown that

$$E\left[e^{i\eta X(k_n(T(t)))}\middle|\,\mathcal{G}_t\right] = \sum_{j=1}^{n^2} \mathbb{1}_{\left[\frac{j-1}{n},\frac{j}{n}\right)}(T(t))e^{\frac{j}{n}\Psi_X(\eta)}.$$

We will show below (see Proposition C.1.1 and Proposition C.1.2) that for all $t \geq 0$ we have:

$$\lim_{n\to\infty} \sum_{j=1}^{n^2} \mathbb{1}_{\left[\frac{j-1}{n},\frac{j}{n}\right)}(T(t))e^{\frac{j}{n}\Psi_X(\eta)} = e^{T(t)\Psi_X(\eta)} \text{ almost surely,}$$

and

$$\begin{aligned} \lim_{n\to\infty} e^{i\eta X(k_n(T(t)))} &= \lim_{n\to\infty} \sum_{j=1}^{n^2} e^{i\eta X\left(\frac{j}{n}\right)}\mathbb{1}_{\left[\frac{j-1}{n},\frac{j}{n}\right)}(T(t)) \\ &= e^{i\eta X(T(t))} \text{ almost surely.} \end{aligned}$$

By application of the conditional dominated convergence theorem (Williams, 1991, Theorem 9.7), we have that

$$\begin{aligned} E\left[e^{i\eta Z(t)}\middle|\,\mathcal{G}_t\right] &= \lim_{n\to\infty} E\left[e^{i\eta X(k_n(T(t)))}\middle|\,\mathcal{G}_t\right] \\ &= \lim_{n\to\infty} \sum_{j=1}^{n^2} \mathbb{1}_{\left[\frac{j-1}{n},\frac{j}{n}\right)}(T(t))e^{\frac{j}{n}\Psi_X(\eta)} \\ &= e^{T(t)\Psi_X(\eta)} \end{aligned}$$

as required. $\square$

C.1 Limits

In this subsection, we will show that:

$$\lim_{n\to\infty}\sum_{j=1}^{n^2}\mathbb{1}_{\left[\frac{j-1}{n},\frac{j}{n}\right)}(T(t))e^{\frac{j}{n}\Psi_X(\eta)} = e^{T(t)\Psi_X(\eta)} \text{ almost surely,} \tag{C.1.1}$$

$$\lim_{n\to\infty}\sum_{j=1}^{n^2}e^{i\eta X\left(\frac{j}{n}\right)}\mathbb{1}_{\left[\frac{j-1}{n},\frac{j}{n}\right)}(T(t)) = e^{i\eta X(T(t))} \text{ almost surely.} \tag{C.1.2}$$

Define two sets $\mathcal{N}_X$ and $\mathcal{N}_T$ as

$$\begin{aligned}\mathcal{N}_X &= \{w\in\Omega:\ \text{the map } t:\to X(t,\omega) \text{ is not a càdlàg path}\},\\ \mathcal{N}_T &= \{w\in\Omega:\ \text{the map } t:\to T(t,\omega) \text{ is not a continuous path}\},\end{aligned}$$

We know that $\mathbb{P}(\mathcal{N}_X)=0$ and $\mathbb{P}(\mathcal{N}_T)=0$. Define the set $\mathcal{N}$ as

$$\mathcal{N}=\mathcal{N}_X\cup\mathcal{N}_T,$$

and notice that $\mathbb{P}(\mathcal{N})\le\mathbb{P}(\mathcal{N}_T)+\mathbb{P}(\mathcal{N}_X)=0$. So if we pick an event that does not belong to the set $\mathcal{N}$, then the paths of X and T for this event are càdlàg and continuous respectively.

Now let us pick a scenario $\omega\in\mathcal{N}^c$ and let us show that

$$\lim_{n\to\infty}k_n(T(t,\omega))=T(t,\omega) \text{ almost surely,}$$

where k_n is defined as in equation (C.0.1).

Because $\omega\in\mathcal{N}^c$ we know that the path of T is continuous for this ω, we have that $T(t,\omega)$ is bounded from above in the interval $[0,t]$. For this ω we can define the integer $M(\omega)$ as

$$M(\omega)=\min\{m\in\mathbb{N}:m\ge T(t,\omega)\}. \tag{C.1.3}$$

Pick an $n(\omega)\in\mathbb{N}$ such that $n(\omega)>M(\omega)$, so we can find an integer $q^n(\omega)$ such that:

$$q^n(\omega)=\min\left\{j\in\{1,\dots,n^2(\omega)\}:T(t,\omega)\in\left[\frac{j-1}{n(\omega)},\frac{j}{n(\omega)}\right)\right\}. \tag{C.1.4}$$

We can write

$$k_n(T(t,\omega))=\frac{q^n(\omega)}{n(\omega)}.$$

Due to the fact that T is non-decreasing and $T(t,\omega)\in\left[\frac{q^n(\omega)-1}{n(\omega)},\frac{q^n(\omega)}{n(\omega)}\right)$ we have that

$$\begin{aligned}|k_n(T(t,\omega))-T(t,\omega)| &= \frac{q^n(\omega)}{n(\omega)}-T(t,\omega)\\ &\le \frac{q^n(\omega)}{n(\omega)}-\frac{q^n(\omega)-1}{n(\omega)}\\ &= \frac{1}{n(\omega)}.\end{aligned}$$

Lastly, letting $n(\omega)$ goes to infinite we have that $|k_n(T(t,\omega)) - T(t,\omega)|$ goes to zero. We have just show that

$$\lim_{n\to\infty} k_n(T(t,\omega)) = T(t,\omega) \text{ for every } \omega \in \mathcal{N}^c.$$

Since the set $\mathcal{N}^c$ has probability 1. We showed that

$$\lim_{n\to\infty} k_n(T(t)) = T(t) \text{ almost surely.}$$

We are now ready to prove the results in (C.1.1) and (C.1.2).

Proposition C.1.1. *For all $\omega \in \mathcal{N}^c$ we have that:*

$$\lim_{n\to\infty} \sum_{j=1}^{n^2} \mathbb{1}_{\left[\frac{j-1}{n},\frac{j}{n}\right)}(T(t,\omega)) e^{\frac{j}{n}\Psi_X(\eta)} = e^{T(t,\omega)\Psi_X(\eta)}.$$

Proof. Let $\omega \in \mathcal{N}^c$ and notice that

$$\sum_{j=1}^{n^2} \mathbb{1}_{\left[\frac{j-1}{n},\frac{j}{n}\right)}(T(t,\omega)) e^{\frac{j}{n}\Psi_X(\eta)} = e^{k_n(T(t,\omega))\Psi_X(\eta)}.$$

The function $f : [0,\infty) \to \mathbb{C}$ defined as

$$f(x) \quad = \quad e^{x\Psi_X(\eta)},$$

is continuous (Pascucci, 2011, Lemma 13.18).

Because f is a continuous function and $\lim_{n\to\infty} k_n(T(t,\omega)) = T(t,\omega)$, applying the result from Abbott (2015, Theorem 4.3.2) we have that

$$\lim_{n\to\infty} e^{k_n(T(t,\omega))\Psi_X(\eta)} = \lim_{n\to\infty} f(k_n(T(t,\omega))) = f(T(t,\omega)).$$

$\square$

Proposition C.1.2. *For all $\omega \in \mathcal{N}^c$ we have that:*

$$\lim_{n\to\infty} \sum_{j=1}^{n^2} e^{i\eta X\left(\frac{j}{n},\omega\right)} \mathbb{1}_{\left[\frac{j-1}{n},\frac{j}{n}\right)}(T(t,\omega)) = e^{i\eta X(T(t,\omega),\omega)}.$$

Proof. Let $\omega \in \mathcal{N}^c$. First, we will show that $X(k_n(T(t,\omega)),\omega) \to X(T(t,\omega),\omega)$ when $n \to \infty$. We have already proved that $k_n(T(t,\omega)) \to T(t,\omega)$ from above when $n \to \infty$.

Due to the fact that $\omega \in \mathcal{N}^c$, we have that the path of X is càdlàg for this ω. So, we have that

$$\lim_{n\to\infty} X(k_n(T(t,\omega)),\omega) = X(T(t,\omega),\omega),$$

(Applebaum, 2009, p. 117).

The function $g : [0,\infty) \to \mathbb{C}$ defined as

$$\begin{aligned} g(x) &= e^{i\eta x} \\ &= \cos(\eta x) + i\sin(\eta x) \end{aligned}$$

is continuous. Finally, we obtain that

$$\lim_{n\to\infty} g(Y_n(\omega)) = g(X(T(t,\omega),\omega)),$$

where $Y_n(\omega) = X(k_n(T(t,\omega)),\omega)$ (Abbott, 2015, Theorem 4.3.2). □

C.1.1 Independence of increments

Let $N \in \mathbb{N}$. We would like to show that the increments $(Z(t_j) - Z(t_{j-1}))_{j=1}^{N}$ where $0 \le t_0 < t_1 < \ldots < t_N$ are independent when conditioned with respect to the sigma algebra $\mathcal{G}_H$ where $H \ge t_N$. To prove this fact, we will use the following result.

Theorem C.1.1. *Let $\eta_1, \eta_2, \ldots, \eta_N \in \mathbb{R}$, then:*

$$\begin{aligned} E\left[e^{\sum_{j=1}^{N} i\eta_j [X(T(t_j)) - X(T(t_{j-1}))]} \middle| \mathcal{G}_H \right] &= e^{\sum_{j=1}^{N} (T(t_j)-T(t_{j-1}))\Psi_X(\eta_j)} \\ &= \prod_{j=1}^{N} e^{(T(t_j)-T(t_{j-1}))\Psi_X(\eta_j)}, \end{aligned} \tag{C.1.5}$$

where Ψ_X is the characteristic exponent of X.

Proof. For every $n \in \mathbb{N}$ we define the function k_n as in (C.0.1). Consider the integer $M \in \mathbb{N}$. We approximate

$$e^{\sum_{j=1}^{N} i\eta_j [X(T(t_j)) - X(T(t_{j-1}))]}$$

by

$$e^{\sum_{j=1}^{N} i\eta_j [X(k_M(T(t_j))) - X(k_M(T(t_{j-1})))]}. \tag{C.1.6}$$

Approximation (C.1.6) can be written as

$$\begin{aligned} &e^{\sum_{j=1}^{N} i\eta_j [X(k_M(T(t_j))) - X(k_M(T(t_{j-1})))]} \\ &\qquad = \sum_{h=0}^{N} \sum_{l_h=1}^{M^2} \left[e^{\sum_{j=1}^{N} i\eta_j \left[X\left(\frac{l_j}{M}\right) - X\left(\frac{l_{j-1}}{M}\right)\right]} \prod_{j=0}^{N} \mathbb{1}_{A_{l_j}^M}(T(t_j)) \right], \end{aligned}$$

where $A_{l_j}^M = \left[\frac{l_j-1}{M}, \frac{l_j}{M}\right)$. Taking the conditional expectation on both sides of the previous equation, we have

$$\begin{aligned} &E\left[e^{\sum_{j=1}^{N} i\eta_n [X(k_M(T(t_j))) - X(k_M(T(t_{j-1})))]} \middle| \mathcal{G}_H \right] \\ &\qquad = \sum_{h=0}^{N} \sum_{l_h=1}^{M^2} E\left[e^{\sum_{j=1}^{N} i\eta_j \left[X\left(\frac{l_j}{M}\right) - X\left(\frac{l_{j-1}}{M}\right)\right]} \prod_{j=0}^{N} \mathbb{1}_{A_{l_j}^M}(T(t_j)) \middle| \mathcal{G}_H \right] \\ &\qquad = \sum_{h=0}^{N} \sum_{l_h=1}^{M^2} \left[E\left[e^{\sum_{j=1}^{N} i\eta_j \left[X\left(\frac{l_j}{M}\right) - X\left(\frac{l_{j-1}}{M}\right)\right]} \middle| \mathcal{G}_H \right] \prod_{j=0}^{N} \mathbb{1}_{A_{l_j}^M}(T(t_j)) \right] \\ &\qquad = \sum_{h=0}^{N} \sum_{l_h=1}^{M^2} \left[E\left[e^{\sum_{j=1}^{N} i\eta_j \left[X\left(\frac{l_j}{M}\right) - X\left(\frac{l_{j-1}}{M}\right)\right]} \right] \prod_{j=0}^{N} \mathbb{1}_{A_{l_j}^M}(T(t_j)) \right] \end{aligned}$$

For each of the terms in this summation, we differentiate between the following cases:

1. When $\frac{l_N}{M} \geq \frac{l_{N-1}}{M} \geq \ldots \geq \frac{l_1}{M} \geq \frac{l_0}{M}$ is satisfied. In this case, we have

$$\begin{aligned} E\left[e^{\sum_{j=1}^{N} i\eta_j \left[X\left(\frac{l_j}{M}\right)-X\left(\frac{l_{j-1}}{M}\right)\right]}\right] &= \prod_{j=1}^{N} E\left[e^{i\eta_j \left[X\left(\frac{l_j}{M}\right)-X\left(\frac{l_{j-1}}{M}\right)\right]}\right] \quad \text{(C.1.7)} \\ &= \prod_{j=1}^{N} E\left[e^{i\eta_j \left[X\left(\frac{l_j}{M}-\frac{l_{j-1}}{M}\right)\right]}\right] \quad \text{(C.1.8)} \\ &= \prod_{j=1}^{N} e^{\left(\frac{l_j}{M}-\frac{l_{j-1}}{M}\right)\Psi_X(\eta_j)} \quad \text{(C.1.9)} \\ &= e^{\sum_{j=1}^{N}\left(\frac{l_j}{M}-\frac{l_{j-1}}{M}\right)\Psi_X(\eta_j)}, \end{aligned}$$

equalities (C.1.7) and (C.1.8) come from the fact that the increments of a Lévy process are independent and stationary, respectively. Equality (C.1.9) is obtained by application of the result in Pascucci (2011, Theorem 13.15).

2. When $\frac{l_N}{M} \geq \frac{l_{N-1}}{M} \geq \ldots \geq \frac{l_1}{M} \geq \frac{l_0}{M}$ is not satisfied. For this case, it is possible to find $q, p \in \{0, 1, \ldots, N\}$ with $p > q$ such that

$$\frac{l_p}{M} < \frac{l_q}{M}.$$

Hence we have that $\frac{l_p}{M} \leq \frac{l_q - 1}{M}$. So we have

$$A_{l_p}^M \cap A_{l_q}^M = \left[\frac{l_p - 1}{M}, \frac{l_p}{M}\right) \bigcap \left[\frac{l_q - 1}{M}, \frac{l_q}{M}\right) = \emptyset.$$

Because the process T is non-decreasing we have that $T(t_p) \geq T(t_q)$, so then

$$\mathbb{1}_{A_{l_p}^M}(T(t_p))\mathbb{1}_{A_{l_q}^M}(T(t_q)) = 0.$$

We have just proved that when $\frac{l_N}{M} \geq \frac{l_{N-1}}{M} \geq \ldots \geq \frac{l_1}{M} \geq \frac{l_0}{M}$ is not satisfied we have

$$\begin{aligned} E\left[e^{\sum_{j=1}^{N} i\eta_j \left[X\left(\frac{l_j}{M}\right)-X\left(\frac{l_{j-1}}{M}\right)\right]}\right] \prod_{j=0}^{N} \mathbb{1}_{A_{l_j}^M}(T(t_j)) &= 0 \\ &= e^{\sum_{j=1}^{N}\left(\frac{l_j}{M}-\frac{l_{j-1}}{M}\right)\Psi_X(\eta_j)} \prod_{j=0}^{N} \mathbb{1}_{A_{l_j}^M}(T(t_j)). \end{aligned}$$

We have just shown that

$$\begin{aligned} \sum_{h=0}^{N} \sum_{l_h=1}^{M^2} &\left[E\left[e^{\sum_{j=1}^{N} i\eta_j \left[X\left(\frac{l_j}{M}\right)-X\left(\frac{l_{j-1}}{M}\right)\right]}\right] \prod_{j=0}^{N} \mathbb{1}_{A_{l_j}^M}(T(t_j))\right] \\ &= \sum_{h=0}^{N} \sum_{l_h=1}^{M^2} \left[e^{\sum_{j=1}^{N}\left(\frac{l_j}{M}-\frac{l_{j-1}}{M}\right)\Psi_X(\eta_j)} \prod_{j=0}^{N} \mathbb{1}_{A_{l_j}^M}(T(t_j))\right] \\ &= e^{\sum_{j=1}^{N}(k_M(T(t_j))-k_M(T(t_{j-1})))\Psi_X(\eta_j)}. \end{aligned}$$

Applying similar results to the ones showed in Proposition C.1.1 and in Proposition C.1.2, we obtain that

$$\begin{aligned}\lim_{M\to\infty} & e^{\sum_{j=1}^{N}(k_M(T(t_j))-k_M(T(t_{j-1})))\Psi_X(\eta_j)} \\ &= e^{\sum_{j=1}^{N}(T(t_j)-T(t_{j-1}))\Psi_X(\eta_j)} \\ &= \prod_{j=1}^{N} e^{(T(t_j)-T(t_{j-1}))\Psi_X(\eta_j)}, \text{ almost surely}\end{aligned}$$

and

$$\begin{aligned}\lim_{M\to\infty} & e^{\sum_{j=1}^{N} i\eta_j[X(k_M(T(t_j)))-X(k_M(X(t_{j-1})))]} \\ &= e^{\sum_{j=1}^{N} i\eta_j[X(T(t_j))-X(T(t_{j-1}))]}, \text{ almost surely.}\end{aligned}$$

By application of the conditional dominated convergence theorem (Williams, 1991, Theorem 9.7), we have that

$$\begin{aligned}E & \left[e^{\sum_{j=1}^{N} i\eta_j[X(T(t_j))-X(T(t_{j-1}))]} \middle| \mathcal{G}_H \right] \\ &= \lim_{M\to\infty} E\left[e^{\sum_{j=1}^{N} i\eta_j[X(k_M(T(t_j)))-X(k_M(T(t_{j-1})))]} \middle| \mathcal{G}_H \right] \\ &= \prod_{j=1}^{N} e^{(T(t_j)-T(t_{j-1}))\Psi_X(\eta_j)}\end{aligned}$$

□

C.1.2 Pseudo-stationary increments

We would like to show that for given $t, s \in [0, \infty)$ such that $t \geq s$, the random variable $(X(T(t))-X(T(s)))$ has the same distribution as the random variable $X(T(t)-T(s))$. First we will show that the distributions of $(X(T(t)) - X(T(s)))$ and $X(T(t) - T(s))$ are the same when conditioned with respect to $\mathcal{G}_H$ where $H \geq t$.

Proposition C.1.3. *Let $\eta \in \mathbb{R}$, then:*

$$\begin{aligned} E\left[e^{i\eta(X(T(t))-X(T(s)))} \middle| \mathcal{G}_H \right] &= E\left[e^{i\eta(X(T(t)-T(s)))} \middle| \mathcal{G}_H \right] \\ &= e^{(T(t)-T(s))\Psi_X(\eta)}. \end{aligned} \tag{C.1.10}$$

Proof. Applying Theorem C.1.1 for just one increment we have that

$$E\left[e^{i\eta(X(T(t))-X(T(s)))} \middle| \mathcal{G}_H \right] = e^{(T(t)-T(s))\Psi_X(\eta)}.$$

Define the process $(R(s,u))_{u\geq 0}$ as

$$R(s,u) = \mathbb{1}_{\{u\geq s\}}\left(T(u) - T(s)\right) \text{ for } u \geq 0.$$

The process $(R(s,u))_{u\geq 0}$ is non-decreasing, non-negative and has continuous sample paths. Also $R(s,u)$ is $\mathcal{G}_u$-measurable for all $u \geq 0$. So, by application of Proposition

C.0.1 we have:

$$\begin{aligned} E\left[e^{i\eta(X(T(t)-T(s)))}\middle|\mathcal{G}_H\right] &= E\left[e^{i\eta X(R(s,t))}\middle|\mathcal{G}_H\right] \\ &= e^{R(s,t)\Psi_X(\eta)} \\ &= e^{(T(t)-T(s))\Psi_X(\eta)}. \end{aligned}$$

□

Corollary C.1.1. *Let $\eta \in \mathbb{R}$, then:*

$$E\left[e^{i\eta(X(T(t))-X(T(s)))}\right] = E\left[e^{i\eta(X(T(t)-T(s)))}\right]$$

Proof. From Proposition C.1.3 we have that

$$E\left[e^{i\eta(X(T(t))-X(T(s)))}\middle|\mathcal{G}_H\right] = E\left[e^{i\eta(X(T(t)-T(s)))}\middle|\mathcal{G}_H\right]. \tag{C.1.11}$$

Taking the expectation on both sides of equation (C.1.11) and applying the tower property, we arrive at the desired result. □

From Corollary C.1.1 we have that the random variables $(X(T(t))-X(T(s)))$ and $X(T(t)-T(s))$ have the same distribution.

Appendix D

Simulation

In this section of the appendix, we explain the techniques for simulating the processes introduced in Section 2.2 and Section 2.3.

D.1 Cox-Ingersoll-Ross process

One of the advantages of the Cox-Ingersoll-Ross process is that its transition density function is known and can be sampled from. For simulating N steps of a Cox-Ingersoll-Ross process with a vector of parameters $\theta_I = (a_I, b_I, \sigma_I)$, time step $\Delta > 0$ and an initial value of $I(0) = y_0$ we use algorithm taken from Iacus (2009, p. 83) (see Algorithm 6). This algorithm generates a realization $\{y_j\}_{j=1}^N$ of the Cox-Ingersoll-Ross process for a given vector of parameters θ_I, time step Δ and initial value y_0.

Algorithm 6: Simulation of the Cox-Ingersoll-Ross process.

$c = \frac{2a_I}{\sigma_I^2(1-\exp(-a_I\Delta))}$;
$q = \frac{2a_I b_I}{\sigma_I^2} - 1$;
for $j = 1, 2, \ldots, N$ **do**
  $u_{j-1} = c y_{j-1} e^{-a_I \Delta}$;
  Draw a sample x_j from a non central chi-square with $2q+2$ degrees freedom and non-centrality parameter $2u_{j-1}$;
  $y_j = \frac{x_j}{2c}$;
end

D.2 Inverse Gaussian Ornstein-Uhlenbeck process

We use the Euler method for the simulation of the inverse Gaussian Ornstein-Uhlenbeck process introduced in Section 2.3.2. For using the Euler method, first we need to know how to simulate the background driving Lévy process Z_I that appears in equation (2.3.1).

The background driving Lévy process Z_I of an $IG(a_I, b_I)$-Ornstein-Uhlenbeck process can be expressed as the sum of two independent possesses $Z_I^{(1)}$ and $Z_I^{(2)}$, where $Z_I^{(1)}$ is an inverse Gaussian Lévy process with parameters $a_I/2$ and b_I, and

$$Z_I^{(2)}(t) = \frac{1}{b_I^2} \sum_{n=1}^{N_I(t)} X_i^2 \quad \text{for } t \geq 0,$$

where N_I is a Poisson process with intensity parameter $\frac{a_I b_I}{2}$, $\{X_i\}_{i=1}^\infty$ is a sequence of independent and identically distributed standard normal random variables and the sequence $\{X_i\}_{i=1}^{\infty}$ and the Poisson process N_I are independent (Schoutens, 2003, p. 69).

Since the background driving Lévy process Z_I can be decomposed as the sum of two other processes that we know how to simulate, we can use the Euler method to simulate the $IG(a_I, b_I)$-Ornstein-Uhlenbeck process:

$$I(t+\Delta) \approx (1-\lambda_I \Delta) I(t) + Z_I(\lambda_I(t+\Delta)) - Z_I(\lambda_I t),$$

for a small time step $\Delta > 0$ (Protter & Talay, 1997). We use Algorithm 7 to generate a realization $\{y_j\}_{j=1}^{N}$ of an inverse Gaussian Ornstein-Uhlenbeck process with given vector of parameters $\theta_I = (a_I, b_I, \lambda_I)$, time step Δ and initial value y_0.

Algorithm 7: Simulation of the inverse Gaussian Ornstein-Uhlenbeck process.

for $j = 1, 2, \ldots, N$ **do**
  Draw a sample Δz_j from the random variable $Z_I(\lambda_I t_j) - Z_I(\lambda_I t_{j-1})$;
  $y_j = (1-\lambda_I \Delta) y_{j-1} + \Delta z_j$;
end

Appendix E

Initial values

As we have seen in Sections 2.2.1 and 2.3.3, we need to maximize the likelihood function to estimate the parameters of the model. The maximization of the likelihood function relies on numerical techniques. These numerical techniques need a starting point or an initial estimate so they can work.

E.1 Cox-Ingersoll-Ross process

For the Cox-Ingersoll-Ross process, we discretize the model and use the ordinary least square estimator to get an initial estimate of a_I, b_I and σ_I. This technique is also used by Kladívko (2007).

For a time step $\Delta > 0$, the Euler method produces the following approximation on the interval $[t, t+\Delta)$ for the Cox-Ingersoll-Ross process:

$$I(t+\Delta) \approx I(t) + a_I(b_I - I(t))\Delta + \sigma_I\sqrt{I(t)}\sqrt{\Delta}\epsilon_t, \tag{E.1.1}$$

where $\epsilon_t \sim N(0,1)$ (Iacus, 2009, p. 122). Now if we have an equally time–spaced sample $\{y_j\}_{j=0}^N$ of the sequence $\{I(t_j)\}_{j=0}^N$, where each y_j is a realization of $I(t_j)$, $t_j = j\Delta$ and each $I(t_j)$ satisfies the discretized equation (E.1.1), then by the Euler method we have that

$$I(t_{j+1}) - I(t_j) = a_I(b_I - I(t_j))\Delta + \sigma_I\sqrt{I(t_j)}\sqrt{\Delta}\epsilon_j \quad \text{for } j = 0, 1, \ldots, N-1, \tag{E.1.2}$$

where $\{\epsilon_j\}_{j=0}^{N-1}$ is a sequence of independent and standard normal random variables. Dividing both sides of equation (E.1.2) by $\sqrt{I(t_j)\Delta}$, we arrive at:

$$\frac{I(t_{j+1}) - I(t_j)}{\sqrt{I(t_j)\Delta}} = a_I b_I\sqrt{\frac{\Delta}{I(t_j)}} - a_I\sqrt{I(t_j)\Delta} + \sigma_I\epsilon_j \quad \text{for } j = 0, 1, \ldots, N-1. \tag{E.1.3}$$

Now let us define

$$\tilde{Y}_j = \frac{I(t_{j+1}) - I(t_j)}{\sqrt{I(t_j)\Delta}}, \quad X_{1,j} = \sqrt{\frac{\Delta}{I(t_j)}}, \quad X_{2,j} = \sqrt{I(t_j)\Delta}, \quad \tau_j = \sigma_I\epsilon_j$$

for $j = 0, 1, \ldots, N-1$, and

$$\beta_1 = a_I b_I, \quad \beta_2 = -a_I. \tag{E.1.4}$$

Then equation (E.1.3) can be rewritten as

$$\tilde{Y}_j = \beta_1 X_{1,j} + \beta_2 X_{2,j} + \tau_j \quad \text{for } j = 0, 1, \ldots, N-1, \tag{E.1.5}$$

where $\{\tau_j\}_{j=0}^{N-1}$ is a sequence of independent normal random variables with mean 0 and variance σ_I^2. Notice that in equation (E.1.5) we have a linear regression model. Since we have the sample $\{y_j\}_{j=0}^{N}$, we can compute the realization of the sequences $\{\tilde{Y}_j\}_{j=0}^{N-1}$, $\{X_{1,j}\}_{j=0}^{N-1}$ and $\{X_{2,j}\}_{j=0}^{N-1}$. The parameters β_1, β_2 and σ_I are then estimated using the ordinary least square estimation method. For linear models, this estimator has an analytical form. From (E.1.4) we can compute the initial estimators of a_I and b_I, using the results obtained for β_1 and β_2.

E.2 Inverse Gaussian Ornstein Uhlenbeck process

To optimize the log-likelihood function defined in (2.3.14), we use a numerical optimization method, but this method needs a starting estimate for the parameters a_I, b_I and λ_I. Since the $IG(a_I, b_I)$-Ornstein-Uhlenbeck process is a stationary process we have the following result.

Proposition E.2.1. *(Valdivieso, 2005, Proposition 3.5) For any $t \geq 0$ and $\Delta \in \mathbb{R}$, the auto-correlation function has the form:*

$$\rho(\Delta) = \frac{Cov\left[I(t), I(t+\Delta)\right]}{\sqrt{Var[I(t)]Var[I(t+\Delta)]}} = e^{-\lambda_I|\Delta|}. \tag{E.2.1}$$

For the parameter λ_I the auto-correlation equation (E.2.1) suggests the initial condition:

$$\hat{\lambda}_I^0 = \frac{-\log(a\hat{c}f(1))}{\Delta},$$

where $a\hat{c}f(1)$ is the empirical auto-correlation function of lag 1, computed from the sample data $\{y_j\}_{j=0}^{N}$ (Valdivieso, Schoutens, & Tuerlinckx, 2009, p. 9).

In Section 2.3.4 we have defined the sequence of residuals $\{M_j\}_{j=1}^{N}$. We know that $\{M_j\}_{j=1}^{N}$ is an independent and identically distributed sequence of random variables, and the random variable M_j has the same distribution as $Z_I^*(\Delta)$ for all $j = 1, \ldots, N$. From the observed sample $\{y_j\}_{j=0}^{N}$ we can compute a realization of the sequence $\{M_j\}_{j=1}^{N}$, that we call $\{m_j\}_{j=1}^{N}$. Since we can compute the first and second moment of the random variable $Z_I^*(\Delta)$ we can use the method of moments for obtaining initial estimates for the parameters a_I and b_I. The method of moments gives the following initial estimators

$$\hat{a}_I^0 = \frac{\hat{b}_I^0 \bar{m}}{(e^{\lambda_I^0 \Delta} - 1)}, \quad \hat{b}_I^0 = \frac{1}{s_m}\sqrt{\frac{\bar{m}(e^{2\lambda_I^0 \Delta} - 1)}{e^{\lambda_I^0 \Delta} - 1}},$$

where

$$\bar{m} = \frac{1}{N}\sum_{j=1}^{N} m_j, \quad s_m^2 = \frac{1}{N}\sum_{j=1}^{M}(m_j - \bar{m})^2,$$

(Valdivieso, 2005, p. 100).

Bibliography

Aalborg, H. A., Molnár, P., & de Vries, J. E. (2019). What can explain the price, volatility and trading volume of Bitcoin? *Finance Research Letters*, *29*, 255–265.

Abbott, S. (2015). *Understanding analysis*. Springer.

Al Guindy, M. (2021). Cryptocurrency price volatility and investor attention. *International Review of Economics & Finance*, *76*, 556–570.

Alexander, C., & Imeraj, A. (2021). *Inverse options in a Black-Scholes world* (tech. rep.) [Online accessed 05-Jul-2022]. https://doi.org/10.48550/arXiv.2107.12041

Al-Saadony, M. (2013). *Bayesian stochastic differential equation modelling with application to finance* (Doctoral dissertation). University of Plymouth.

Andersen, E. B. (1970). Asymptotic properties of conditional maximum-likelihood estimators. *Journal of the Royal Statistical Society: Series B (Methodological)*, *32*(2), 283–301.

Applebaum, D. (2009). *Lévy processes and stochastic calculus*. Cambridge University Press.

Aslanidis, N., Bariviera, A. F., & López, Ó. G. (2021). *The link between Bitcoin and Google Trends attention* (tech. rep.) [Online accessed 05-Jul-2022]. https://doi.org/10.48550/arXiv.2106.07104

Avellaneda, M., Levy, A., & Parás, A. (1995). Pricing and hedging derivative securities in markets with uncertain volatilities. *Applied Mathematical Finance*, *2*(2), 73–88.

Balash, V. (2013). *Stochastic volatility model with an exogenous control process of news flow* (Master's thesis). Brunel University, School of Information Systems, Computing and Mathematics.

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

Barndorff-Nielsen, O. E., & Shephard, N. (2003). Integrated Ornstein–Uhlenbeck processes and non-Gaussian Ornstein–Uhlenbeck based stochastic volatility models. *Scandinavian Journal of statistics*, *30*(2), 277–295.

Barunik, J., Chen, C. Y., & Vecer, J. (2019). *Sentiment-driven stochastic volatility model: A high-frequency textual tool for economists* (tech. rep.) [Online accessed 05-Jul-2022]. https://doi.org/10.48550/arXiv.1906.00059

Benth, F. E., Kallsen, J., & Meyer-Brandis, T. (2007). A non-Gaussian Ornstein–Uhlenbeck process for electricity spot price modeling and derivatives pricing. *Applied Mathematical Finance*, *14*(2), 153–169.

Cao, M., & Celik, B. (2021). Valuation of Bitcoin options. *Journal of Futures Markets*, *41*, 1007–1026.

Carr, P., Geman, H., Madan, D. B., & Yor, M. (2003). Stochastic volatility for Lévy processes. *Mathematical Finance*, *13*(3), 345–382.

Carr, P., & Madan, D. B. (1999). Option valuation using the fast Fourier transform. *Journal of Computational Finance*, *2*(4), 61–73.

Chan, S., Chu, J., Nadarajah, S., & Osterrieder, J. (2017). A statistical analysis of cryptocurrencies. *Journal of Risk and Financial Management*, *10*(2), 1–23.

Chen, J., Tang, G., Yao, J., & Zhou, G. (2022). Investor attention and stock returns. *Journal of Financial and Quantitative Analysis*, *57*(2), 455–484.

Chen, K., & Huang, Y. (2021). Detecting jump risk and jump-diffusion model for Bitcoin options pricing and hedging. *Mathematics*, *9*(20), 1–24.

Cherny, A. S., & Engelbert, H. (2005). *Singular stochastic differential equations*. Springer Science & Business Media.

Chopin, N., & Papaspiliopoulos, O. (2020). *An introduction to sequential Monte Carlo*. Springer.

Christoffersen, P., Heston, S., & Jacobs, K. (2009). The shape and term structure of the index option smirk: Why multifactor stochastic volatility models work so well. *Management Science*, *55*(12), 1914–1932.

Ciaian, P., Rajcaniova, M., & Kancs, D. A. (2016). The economics of Bitcoin price formation. *Applied Economics*, *48*(19), 1799–1815.

Cont, R., & Tankov, P. (2004). *Financial modelling with jump processes*. Chapman & Hall/CRC.

Cox, J. C., Ingersoll, J. E., & Ross, S. A. (1985). A theory of the term structure of interest rates. *Econometrica*, *53*(2), 385–407.

Cretarola, A., & Figà-Talamanca, G. (2021). Detecting bubbles in Bitcoin price dynamics via market exuberance. *Annals of Operations Research*, *299*, 1–21.

Cretarola, A., Figà-Talamanca, G., & Patacca, M. (2020). Market attention and Bitcoin price modeling: Theory, estimation and option pricing. *Decisions in Economics and Finance*, *43*, 187–228.

Da, Z., Engelberg, J., & Gao, P. (2011). In search of attention. *The Journal of Finance*, *66*(5), 1461–1499.

Davison, A. C. (2003). *Statistical models*. Cambridge University Press.

deLeeuw, J. (1992). Introduction to Akaike (1973) information theory and an extension of the maximum likelihood principle. In S. Kotz & N. L. Johnson (Eds.), *Breakthroughs in statistics: Foundations and basic theory* (pp. 599–609). Springer New York.

Eom, C., Kaizoji, T., Kang, S. H., & Pichl, L. (2019). Bitcoin and investor sentiment: Statistical characteristics and predictability. *Physica A: Statistical Mechanics and its Applications*, *514*, 511–521.

Figà-Talamanca, G., & Patacca, M. (2019). Does market attention affect Bitcoin returns and volatility? *Decisions in Economics and Finance*, *42*, 135–155.

Figà-Talamanca, G., & Patacca, M. (2020). Disentangling the relationship between Bitcoin and market attention measures. *Journal of Industrial and Business Economics*, *47*, 71–91.

Friz, P., & Keller-Ressel, M. (2010). Moment explosions in stochastic volatility models. In R. Cont (Ed.), *Encyclopedia of quantitative finance* (pp. 1247–1253). John Wiley & Sons.

Gulisashvili, A. (2012). *Analytically tractable stochastic stock price models*. Springer Berlin, Heidelberg.

Hautsch, N., & Ou, Y. (2008). Stochastic volatility estimation using markov chain simulation. In W. K. Härdle, N. Hautsch, & L. Overbeck (Eds.), *Applied quantitative finance* (pp. 249–274). Springer Berlin, Heidelberg.

Heston, S. L. (1993). 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.

Hirsa, A. (2012). *Computational methods in finance*. Chapman & Hall/CRC.

Hou, A., Wang, W., Chen, C., & Härdle, W. (2020). Pricing cryptocurrency options. *Journal of Financial Econometrics*, *18*(2), 250–279.

Hou, K., Xiong, W., & Peng, L. (2009). *A tale of two anomalies: The implications of investor attention for price and earnings momentum* (tech. rep.) [Online accessed 05-Jul-2022]. http://dx.doi.org/10.2139/ssrn.976394

Hull, J., & White, A. (1987). The pricing of options on assets with stochastic volatilities. *The journal of Finance*, *42*(2), 281–300.

Iacus, S. M. (2009). *Simulation and inference for stochastic differential equations: with R examples*. Springer New York, NY.

Jondeau, E., Poon, S., & Rockinger, M. (2007). *Financial modeling under non-Gaussian distributions*. Springer London.

Kalbfleisch, J. D., & Sprott, D. A. (1970). Application of likelihood methods to models involving large numbers of parameters. *Journal of the Royal Statistical Society: Series B (methodological)*, *32*(2), 175–194.

Karatzas, I., & Shreve, S. E. (1998). *Brownian motion and stochastic calculus*. Springer New York, NY.

Kladívko, K. (2007). Maximum likelihood estimation of the Cox-Ingersoll-Ross process: the Matlab implementation. *Technical Computing Prague*, *7*(8).

Klebaner, F. C. (2012). *Introduction to stochastic calculus with applications*. Imperial College Press.

Klingler, S., Kim, Y. S., Rachev, S. T., & Fabozzi, F. J. (2013). Option pricing with time-changed Lévy processes. *Applied Financial Economics*, *23*(15), 1231–1238.

Kopp, E., Malczak, J., & Zastawniak, T. (2014). *Probability for finance*. Cambridge University Press.

Kou, S. G. (2002). A jump-diffusion model for option pricing. *Management Science*, *48*(8), 1086–1101.

Kristoufek, L. (2015). What are the main drivers of the Bitcoin price? Evidence from wavelet coherence analysis. *PloS ONE*, *10*(4).

Lapidoth, A. (2017). *A foundation in digital communication*. Cambridge University Press.

Le Gall, J.-F. (2016). *Brownian motion, martingales, and stochastic calculus* (Vol. 274). Springer.

Li, H., Yang, K., & Wang, D. (2019). A threshold stochastic volatility model with explanatory variables. *Statistica Neerlandica*, *73*(1), 118–138.

Liang, C., Zhang, Y., Li, X., & Ma, F. (2022). Which predictor is more predictive for Bitcoin volatility? And why? *International Journal of Finance & Economics*, *27*(2), 1947–1961.

Lindström, E. (2004). *Statistical modeling of diffusion processes with financial applications* (Doctoral dissertation). Lund University.

Ljung, G. M., & Box, G. E. (1978). On a measure of lack of fit in time series models. *Biometrika*, *65*(2), 297–303.

López-Cabarcos, M. Á., Pérez-Pico, A. M., Piñeiro-Chousa, J., & Šević, A. (2021). Bitcoin volatility, stock market and investor sentiment. are they connected? *Finance Research Letters*, *38*, 1–7.

Madan, D. B., Reyners, S., & Schoutens, W. (2019). Advanced model calibration on Bitcoin options. *Digital Finance*, *1*, 117–137.

Massey, F. J. (1951). The Kolmogorov-Smirnov test for goodness of fit. *Journal of the American Statistical Association*, *46*(253), 68–78.

Merton, R. C. (1976). Option pricing when underlying stock returns are discontinuous. *Journal of Financial Economics*, *3*(1), 125–144.

Mijatović, A., & Urusov, M. (2012). On the martingale property of certain local martingales. *Probability Theory and Related Fields*, *152*(1), 1–30.

Mishura, Y., & Yurchenko-Tytarenko, A. (2018). Fractional Cox–Ingersoll–Ross process with non-zero «mean». *Modern Stochastics: Theory and Applications*, *5*(1), 99–111.

Mrázek, M., & Pospíšil, J. (2017). Calibration and simulation of Heston model. *Open Mathematics*, *15*(1), 679–704.

Murphy, K. P. (2012). *Machine learning: A probabilistic perspective*. MIT press.

Nakamoto, S. (2008). *Bitcoin: A peer-to-peer electronic cash system* (tech. rep.) [Online accessed 05-Jul-2022]. https://bitcoinwhitepaper.co/bitcoin.pdf

Neath, A. A., & Cavanaugh, J. E. (2012). The Bayesian information criterion: Background, derivation, and applications. *Wiley Interdisciplinary Reviews: Computational Statistics*, *4*(2), 199–203.

Nicolato, E., & Venardos, E. (2003). Option pricing in stochastic volatility models of the Ornstein–Uhlenbeck type. *Mathematical Finance: An International Journal of Mathematics, Statistics and Financial Economics*, *13*(4), 445–466.

Olivares, P. (2020). *Pricing Bitcoin derivatives under jump-diffusion models* (tech. rep.) [Online accessed 05-Jul-2022]. https://doi.org/10.48550/arXiv.2002.07117

Pascucci, A. (2011). *PDE and martingale methods in option pricing*. Springer Milano.

Philippas, D., Rjiba, H., Guesmi, K., & Goutte, S. (2019). Media attention and Bitcoin prices. *Finance Research Letters*, *30*, 37–43.

Pipiras, V., & Taqqu, M. S. (2017). *Long-range dependence and self-similarity*. Cambridge University Press.

Protter, P., & Talay, D. (1997). The Euler scheme for Lévy driven stochastic differential equations. *The Annals of Probability*, *25*(1), 393–423.

Rachev, S. T., Hsu, J. S., Bagasheva, B. S., & Fabozzi, F. J. (2008). *Bayesian methods in finance*. John Wiley & Sons.

Ranco, G., Aleksovski, D., Caldarelli, G., Grčar, M., & Mozetič, I. (2015). The effects of Twitter sentiment on stock price returns. *PloS one*, *10*(9), 1–21.

Sato, K. (1999). *Lévy processes and infinitely divisible distributions*. Cambridge University Press.

Schoutens, W. (2003). *Lévy processes in finance: Pricing financial derivatives*. John Wiley & Sons.

Shirvani, A., Mittnik, S., Lindquist, W. B., & Rachev, S. T. (2021). *Bitcoin volatility and intrinsic time using double subordinated Lévy processes* (tech. rep.) [Online accessed 05-Jul-2022]. https://doi.org/10.48550/arXiv.2109.15051

Shojaee, S. (2018). *Stochastic volatility enhanced Lévy processes in financial asset pricing* (Master's thesis). Chalmers: University of Technology.

Siu, T. K., & Elliott, R. J. (2021). Bitcoin option pricing with a SETAR–GARCH model. *The European Journal of Finance*, *27*(6), 564–595.

Skiadas, C. H. (2010). Exact solutions of stochastic differential equations: Gompertz, generalized logistic and revised exponential. *Methodology and Computing in Applied Probability*, *12*(2), 261–270.

Smales, L. A. (2022). Investor attention in cryptocurrency markets. *International Review of Financial Analysis*, *79*, 1–17.

Sokol, A., & Rønn-Nielsen, A. (2013). *Advanced probability* (Lecture Notes) [Online accessed 21-Jan-2020]. University of Copenhagen: Department of Mathematical Sciences. http://web.math.ku.dk/noter/filer/vidsand12.pdf

Suardi, S., Rasel, A. R., & Liu, B. (2022). On the predictive power of tweet sentiments and attention on Bitcoin. *International Review of Economics & Finance*, *79*, 289–301.

Taylor, S. J. (1982). Financial returns modelled by the product of two stochastic processes-a study of the daily sugar prices 1961-75. In O. D. Anderson (Ed.), *Time series analysis: Theory and practice 1* (pp. 203–226). North–Holland.

Valdivieso, L. (2005). Parameter estimation for an inverse Gaussian Ornstein–Uhlenbeck stochastic volatility model. In M. Vanmaele, A. Schepper, J. Dhaene, H. Reynaerts, W. Schoutens, & P. Goethem (Eds.), *Proceedings of 3rd actuarial and financial mathematics day* (pp. 95–104). Royal Flemish Academy of Belgium for Science; Arts.

Valdivieso, L., Schoutens, W., & Tuerlinckx, F. (2009). Maximum likelihood estimation in processes of Ornstein–Uhlenbeck type. *Statistical Inference for Stochastic Processes*, *12*(1), 1–19.

Venter, P. J., Mare, E., & Pindza, E. (2020). Price discovery in the cryptocurrency option market: A univariate GARCH approach. *Cogent Economics & Finance*, *8*(1), 1–9.

Wang, W., Hou, A. J., Chen, C., & Härdle, W. K. (2020). Pricing cryptocurrency options. *Journal of Financial Econometrics*, *18*(2), 250–279.

Wei, W., Mao, Y., & Wang, B. (2016). Twitter volume spikes and stock options pricing. *Computer Communications*, *73*, 271–281.

Williams, D. (1991). *Probability with martingales*. Cambridge University Press.

Wilmott, P. (2006). *Paul Wilmott on quantitative finance, 3 volume set*. John Wiley & Sons.

Yang, H. (2015). *Markov chain Monte Carlo estimation of stochastic volatility models with finite and infinite activity Lévy jumps: Evidence for efficient models and algorithms* (Doctoral dissertation). Tampere University of Technology.

Yuan, D., & Lei, L. (2016). Some results following from conditional characteristic functions. *Communications in Statistics–Theory and Methods*, *45*(12), 3706–3720.

Zhang, X., Lu, F., Tao, R., & Wang, S. (2021). The time-varying causal relationship between the Bitcoin market and internet attention. *Financial Innovation*, *7*(1), 1–19.