DEGREE PROJECT IN MATHEMATICS,
SECOND CYCLE, 30 CREDITS
STOCKHOLM, SWEDEN 2017

The Rolling Window Method: Precisions of Financial Forecasting

LUDVIG HÄLLMAN

KTH ROYAL INSTITUTE OF TECHNOLOGY
SCHOOL OF ENGINEERING SCIENCES

The Rolling Window Method: Precisions of Financial Forecasting

LUDVIG HÄLLMAN

Degree Projects in Mathematical Statistics (30 ECTS credits)
Degree Programme in Applied and Computational Mathematics (120 credits)
KTH Royal Institute of Technology year 2017
Supervisors at Kidbrooke Advisory: Fredrik Davéus, Edvard Sjögren
Supervisor at KTH: Henrik Hult
Examiner at KTH: Henrik Hult

TRITA-MAT-E 2017:12
ISRN-KTH/MAT/E--17/12--SE

Royal Institute of Technology
School of Engineering Sciences
KTH SCI
SE-100 44 Stockholm, Sweden
URL: www.kth.se/sci

Abstract

In this thesis we set out to study the prediction accuracy of statistical quantities related to portfolio analysis and risk management implied by a given set of historical data. The considered forecasting procedure rely on rolling-window estimates over varying horizons where the resulting empirical return distributions can be considered the corresponding stationary distributions. By using scenarios generated from a joint interest rate-equity framework the rolling-window method allows to, empirically, study the uncertainty of return statistics as well as risk measures related to market risk. The study shows that, given the chosen models, the method is valid in predicting future statistical quantities related to portfolio return of up to one year. For risk measures, the forecasting uncertainty is found to be too significant and highlights the difficulty in foreseeing extremities of future market movements.

Sammanfattning

I detta examensarbete ämnar vi oss att studera precisionen av predikterade statistiska storheter relaterade till portföljanalys och riskhantering givet en mängd historisk data. Den använda prediktionsmetoden använder sig av rullande fönster estimeringar över varierande horisonter där de resulterande empiriska avkastningsfördelningarna kan ses som de motsvarande stationära fördelningarna. Genom att använda scenarier generade från ett ramverk för räntor och aktier, möjliggör rullande fönster metoden att, empiriskt, studera osäkerheter i skattade avkastnings statistikor och riskmått relaterade till marknads risk. Studien visar, givet de ingående modellerna, att metoden är giltig att använda för prediktering av statistiska storheter relaterade till portföljavkastningar upp till ett år. För riskmått visar sig skattningsosäkerhet vara för stor och belyser svårigheten att förutse extremiteter i framtida marknadsutfall.

Acknowledgements

There are several people whom without their help and support this thesis would not have been finalised. Firstly, I would like to thank my supervisor at the Royal Institute of Technology, Prof. Henrik Hult for introducing me to the subject and the method at hand, and for his valuable feedback and academic guidance throughout the thesis. Furthermore, I would like to extend my sincerest gratitude to my colleagues at Kidbrooke Advisory. To my two supervisors, Fredrik Davéus and Edvard Sjögren for invaluable inputs, suggestions and insight into the financial industry, and for giving me the opportunity to realise this thesis. To Sadna Sajadini for interesting and helpful discussions as well as great companionship throughout the thesis. Lastly, I would like to thank my family. To my brother, Viktor Hällman for his time and contribution in the proofreading of this thesis, and to my parents for their endless support, and constant encouragement and motivation to pursue higher education.

Ludvig Hällman
Stockholm, April 2017

Contents

Abbreviations and Notation

Probability Notations

- $(\Omega, \mathcal{F}, \mathrm{P})$: probability space.
- $\mathcal{F}$: $\sigma -$ algebra on Ω.
- P : objective probability measure.
- Q : risk-neutral probability measure.
- $\mathrm{E}(\cdot)$: expectation under P.
- $\mathrm{E}^{\mathrm{Q}}(\cdot)$: expectation under Q.
- $\mathrm{E}(\ \cdot \mid \mathcal{F}(t))$: expectation conditional on the filtration up to time t.
- $\mathrm{Var}(\cdot)$: variance under P.
- $\mathrm{Var}^{\mathrm{Q}}(\cdot)$: variance under Q.
- $\mathrm{Std}(\cdot)$: standard deviation.
- $\mathrm{Cov}(\cdot, \cdot)$: covariance.
- $\mathrm{Corr}(\cdot, \cdot)$: correlation.
- $\langle \cdot \rangle, \langle \cdot\cdot \rangle$: quadratic variation and covariation.
- $\xrightarrow{\mathrm{d.}}$: convergence in distribution.

Miscellaneous Notations

- VaR : Value-at-Risk.
- ES : Expected Shortfall.

- DE : Diversification Effect.
- $x_+ = \max(x, 0)$.
- $x \vee y = \max(x, y)$.
- $\bigvee_{i=1}^{n} x_i = x_1 \vee x_2 \vee \ldots \vee x_n$.
- $\mathbb{N}_0 = \{0, 1, 2, \ldots\}$: set of all non-negative natural numbers.

Chapter 1

Introduction

Access to reliable data is an indispensable part in the day-to-day operation and functionality of financial institutions. Historical financial data is used in numerous areas, including daily model calibration, derivative pricing, forecasting and risk calculation. However, the rapidly increasing demand for more sophisticated mathematical models within the financial industry has made the access to sufficient amount of data a recurring and increasingly more important issue. In particular, this is a problem when modelling longer time horisons[1] as the calibration of such models rely, extensively, on the access to long series of stationary historical data. Since there is a shortage of this sort of data, the absence of sufficient historical data results in poorly calibrated models and misleading future risk estimates which in turn can have major impacts on the performance and profitability of the company and its shareholders.

Following the recent financial crisis and its aftermath, the response in regulatory requirements has deemed the access to reliable and sufficient historical data a key element in the creation of stable, yet efficient financial systems. To accurately estimate future risk factors as well as to comprehensively understand the amount of uncertainty an estimate carries is essential in the mitigation of future risk and prevention of future crises.

In this thesis, we therefore set out to study the accuracy implied by a given amount of historical market data when estimating statistical quantities related to asset returns and risk management. In Section 1.1 some background related to the thesis is presented. A brief history of market risk is outlined and a review of recent regulatory framework is outlined. Section 1.2 formally outlines the main objectives of the thesis and what we aim to study, followed by scope and limitations in Section 1.3. Section 1.4 concludes the chapter by outlining how the remaining of the thesis is structured.

[1] 20+ years.

1.1 Background

To further put the thesis into context we, in this section, present some background highlighting a couple of areas where financial data is of great importance. We present market risk and outline a general overview of the regulatory frameworks in force today.

1.1.1 Market Risk

Within banking, the most prevalent and well known type of risk is market risk. Market risk refers to the risk of a change in the value of financial positions due to price fluctuations in the underlying components of which the position depends. Because of the multitude of asset classes such a financial position can depend, market risk is usually considered an umbrella term encompassing multiple types of risk. The major risk factors is reflected in the movements of equities, commodities, interest rates and currencies. As a simple example, consider an investor taking a long position in a zero-coupon bond maturing at a future date. Now, assume the market observes an unexpected increase in interest rates. This could trigger companies to issue new bonds with a higher yield that incorporates the realised increase in interest rate. This would make the initial bond less attractive to new investors resulting in a decrease in its value. If the investor were to sell the bond before its maturity, this could mean that the investor had to sell the bond for less than it was originally traded for. This interest rate movement is clearly unfavorable for the bond holder, reflecting interest rate risk that resulted in a capital loss for the investor.

Another example would be to consider an investor taken a long position in an *in-the-money* call option at time of purchases. Assume further that, over the contract period, the underlying stock develop negatively with a significantly decrease in the value of the underlying so that, at maturity the option is *out-of-the-money*. This is, in the same way as in the bond example, an unfavorable outcome for the investor. Here the decrease of the underlying stock reflect equity risk that resulted in an unrealised loss for the investor.

The large number of risk factors associated with various positions have been a driving force for extensive studies in the measurement of market risk. Unfortunately this is not trivial, extreme outcomes commonly associated with risk measures[2] occurs with low probability making the number of observations limited. This makes it all the more important to understand the limitations that a series of historical data has in the estimation of future risk.

1.1.2 Regulatory Frameworks

Following the aftermath of the late 2000s financial crisis, there is a growing consensus among policymakers and economic researchers about the need of tougher, uniformed regulatory frameworks and a re-orientation in their focus towards a macroprudential perspective. In

[2] Usually 95-99.5th quantiles of the loss distribution are of interest.

response, two of the recently introduced regulatory frameworks enforced in the financial sector today are Solvency II and Basel III.

Basel III, is the third and latest regulatory directive issued by the Basel Committee on Banking Supervision (BCBS). It came into force in 2014 and is expected to be finalised in the beginning of 2019. The directive targets banks and stipulates requirements in areas regarding minimum capital requirement standards, liquidity requirement standards and the choice of risk measures in the estimation of financial risk.

The Solvency II directive refers to the latest regulatory directive affecting the EU insurance sector. Solvency II came into force in 2014 and contains, among other, regulations concerning solvency capital requirements. Insurance companies are obliged to hold sufficient buffer of funds to ensure that they can meet financial obligations over 12 months with a probability of at least 99.5%. Plainly speaking, this is suppose to reflect the fact that the risk of a company defaulting should be strictly less than one in two hundred years.

The calculation of capital requirement and estimation of future risk exposure are two of many procedures within regulatory compliance that extensively rely on the access to sufficient and reliable data. Here, accurately calibrated models, risk factor simulation and forecasting of statistical quantities are key procedures in the calculation of capital requirement.

1.2 Thesis Objectives

The objective of this thesis is to investigate the predictive power implied by historical data. In doing so, the basis of this study will rely on a joint modelling framework for interest rates equities. The framework will serve as a scenario generator from which statistical quantities related to portfolio returns and risk management can be estimated. The development of a flexibility framework that allows to vary the complexity of the included models will enable us to study the accuracy of the estimated statistics when introducing commonly known behaviors and stylized facts seen in equities and interest rates.

1.3 Scope and Limitations

As stated in Section 1.2, the main focus of this work is to study the accuracy of estimated statistics related to risk factors implied by historical data. However, as this does not include the calibration of the framework generating the scenarios itself, no insight into possible calibration procedures regarding the models will be reviewed. Instead, justifiable model parameters will be used such that the generated scenarios can be considered comparable to the historical behavior of equities, interest rates and their dependence.

The sources of risk that will be analysed will be limited to market risk, and in particular equity and interest rate risk. The equity risk will be represented by modelling the equity process, which will serve as a proxy for a stock index. For the interest rate risk factor,

a zero-coupon bond will be considered, with the underlying interest rate being driven by a stochastic process. Since credit risk is not within the scope of the thesis, the bond considered will not be susceptible to defaults and no form of hazard modelling will therefore be discussed. Instead we will model so-called non-defaultable zero-coupon bonds where the investor is guaranteed payoff at maturity. Lastly, all processes will be modelled as excess returns above the risk-free rate.

1.4 Disposition

The remainder of the thesis is structured as follows. In Chapter 2, the underlying mathematics of the thesis is presented. A review of some of the most significant findings within resent financial literature will be presented. Market relevant models connected to the modelling of equity and interest rate in continuous time are presented. Methods concerning the modelling of dependence between asset classes are reviewed in Section 2.3 and includes copula methods and tail dependence. The chapter concludes with Section 2.4 where a review in risk management is presented. The concept of risk measure is defined along with examples of some widely used measures such as Value-at-Risk and Expected Shortfall.

Chapter 3 deals with the methodology behind the estimation analysis. Estimation procedures of sample statistics and risk measures is reviewed along with the simulation procedures of the presented models in Chapter 2 and their dependence.

The results of the uncertainty study are presented in Chapter 4 and includes the result connected to the predictability of both asset classes separately as well as a review of the predictability of the joint multivariate framework.

Conclusions are presented in Chapter 5 followed by a discussion on the chosen models and possible extensions and future work. Appendices, containing complementary information, are referred to as needed throughout the thesis.

Chapter 2

Mathematical Background

In this chapter the underlying mathematical background of the thesis is presented. The chapter commences by introducing the concept of interest rate and equity modelling. The considered models composes some of the most relevant models seen in recent literature and industry. The chapter continues by presenting dependence modelling and how we aim to model a joint equity-interest rate framework and finalises with a review in risk management and its mathematical applications.

The reader is assumed to be familiar with concepts within probability theory, stochastic calculus and option pricing. Should that not be the case, rigorous treatment of the material presented in this chapter can be found in e.g., Karatzas and Shreve (1991), Øksendal (2003), and Protter (2004).

We start by fixing a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}(t))_{t\geq 0}, \mathrm{P})$, where Ω denotes the sample space containing outcome elements ω, $\mathcal{F}$ denotes the σ-algebra on Ω, and P denotes the probability space on $(\Omega, \mathcal{F})$. P will be referred to as the *objective*[1] probability measure. $\{\mathcal{F}(t) : t \geq 0\}$ denotes the filtration of information available up to time t such that whenever $s < t < \infty$, $\mathcal{F}(s) \subseteq \mathcal{F}(t) \subseteq \mathcal{F}$, is a sub-$\sigma$-algebra of $\mathcal{F}$. A priori, no finite time horizon is fixed such that $\mathcal{F} = \mathcal{F}(\infty) = \bigvee_{t\geq 0} \mathcal{F}(t)$. Furthermore, we assume all stochastic processes, say $X(t)$, to be adapted to $\{\mathcal{F}(t)\}$, i.e., $X(t)$ is fully observable at time t, and that the filtration satisfies the usual conditions[2].

In addition to P, we assume that there exist an equivalent martingale measure[3] (EMM) $\mathrm{Q} \sim \mathrm{P}$, on the form (B.1), defined on $(\Omega, \mathcal{F})$. We will refer to Q as the *risk-neutral* probability measure. Finally, $\mathcal{F}(t)$ is assumed to be a Brownian filtration, i.e., $\mathcal{F}(t) = \sigma(W(s) : 0 < s < t)$, where W is a Brownian motion under P such that there exists a Brownian motion W^* defined under Q by (B.2).

[1] Commonly also referred to as the *real world* or *physical* measure.

[2] The usual conditions are; completeness: $\mathcal{F}(0)$ contains all null-sets of $\mathcal{F}$, and right-continuity: $\mathcal{F}(t) = \bigcap_{s>t} \mathcal{F}(s),\ \forall t \geq 0$.

[3] See Appendix B for more details.

2.1 Interest Rate Models

In this section the concept of interest rate modelling will be introduced as well as how we aim to model it throughout the thesis. We also present its connection to bond pricing.

The basis of interest rate modelling relies on the existence of the so-called *money-market account*[4] and represents a risk-less investment, where profit is accrued continuously at the risk-free rate in the market. The money-market account will be denoted B and is defined by the differential equation

$$dB(t) = r(t)B(t)dt, \quad B(0) = 1, \tag{2.1}$$

where r is an assumed positive function of time referred to as the instantaneous rate.

The solution to the differential equation of the bank account in (2.1) implies that

$$B(t) = \exp\left\{\int_0^t r(s)ds\right\}, \tag{2.2}$$

meaning that, by investing an amount of 1 at time 0 yields a risk-less profit at time t equal to the value in (2.2).

In option pricing theory, the money-market account plays a central role where it serves as the *numéraire* under Q. From the existence of $\mathrm{Q} \sim \mathrm{P}$, the first fundamental theorem of asset pricing ensures that the market is free of arbitrage such that the time t price of any contingent claim with payoff $V(T)$ at time T is given by

$$V(t) = \mathrm{E}^{\mathrm{Q}}\left(\frac{B(t)}{B(T)}V(T) \;\middle|\; \mathcal{F}(t)\right) \tag{2.3}$$

(e.g., Björk (2009, ch. 10, 23), Filipović (2009, ch. 4, 5)). The contingent claim that will be considered is that of $V(T) = 1$, known as the zero-coupon bond.

The price of the zero-coupon bond will be denoted by $P(t,T)$ and paying unity at maturity, i.e., $P(T,T) = 1$. By means of (2.3), the risk-free zero-coupon bond price at time t will be given by

$$P(t,T) = \mathrm{E}^{\mathrm{Q}}\left(\frac{B(t)}{B(T)} \cdot 1 \;\middle|\; \mathcal{F}(t)\right) = \mathrm{E}^{\mathrm{Q}}\left(\exp\left\{-\int_t^T r(s)ds\right\} \;\middle|\; \mathcal{F}(t)\right) \tag{2.4}$$

and ensures that every discounted T-bond price process $\frac{P(t,T)}{B(t)}$ is a martingale under Q.

The zero-coupon bond will be the main financial instrument subject to interest rate risk that will be addressed in this thesis. This means that r needs to be modelled. In doing so, a family of interest rate models called short-rate models will be considered.

The short-rate is a mathematical quantity representing the interest rate valid for an infinitesimally short period of time. These kind of models were developed to directly model

[4] Also refereed to as the *bank account*.

the instantaneous spot rate as an Itô process under Q. The advantage with this method is that r, under P, is not fully determined without the exogenous specification of the market price of risk[5]. The existence of $\mathrm{Q} \sim \mathrm{P}$ only infer absence of arbitrage and not uniqueness, since the money-market account alone cannot replicate contingent claims. The lack of uniqueness means that Q can be any EMM and the market price of risk is not accessible. To solve this, the specification of the dynamics of r under the objective measure rely on statistical methods applied to historical observations of price movements. For that reason, it is instead much more convenient to model the short-rate directly under Q where bond prices are given by (2.4) (e.g., Filipović (2009, ch. 5), Brigo and Mercurio (2012, ch. 3)).

2.1.1 Short-Rate Models

The short-rate models that will be considered are the one factor models given by the one-dimension diffusion process

$$\begin{aligned} dr(t) &= b(t, r(t))dt + \sigma(t, r(t))dW^*(t), \quad t > 0, \\ r(0) &= r_0, \end{aligned} \tag{2.5}$$

for some functions $b, \sigma : \mathbb{R}_+ \times \mathbb{R} \to \mathbb{R}$. Here, dW^* denotes the Brownian motion defined under Q.

The special class of short-rate models that will be considered in this thesis are the class referred to as affine models. In the case of one-factor models, this class of short-rate models are widely popular due to their ability of providing closed-form expressions for bond prices. Given an affine short-rate r, the time t price of the corresponding zero-coupon bond price is give by

$$P(t, T) = e^{-A(t,T)-B(t,T)r(t)}. \tag{2.6}$$

As there are many short-rate models not belonging to the family of affine process, a proper definition of their structure is outlined below.

Affine Term-Structure

An interest rate model is said to be an affine term-structure (ATS) model if the continuously-compounded spot rate R is an affine function in r, (Brigo and Mercurio, 2012, ch. 3), i.e.,

$$R(t, T) = \xi(t, T) + \zeta(t, T)r(t), \tag{2.7}$$

for some deterministic functions ξ and ζ. From the definition of R,

$$R(t, T) := -\frac{\log P(t, T)}{T - t}, \tag{2.8}$$

[5] See Appendix B for more details.

it is clear that the affine relationship between R and r always hold for bond prices of the form in (2.6) as it suffices to set

$$\xi(t,T) = \frac{A(t,T)}{T-t}, \quad \text{and} \quad \zeta(t,T) = \frac{B(t,T)}{T-t},$$

given that $A(T,T) = B(T,T) \equiv 0$.

Here we would like to stress the importance of the bond price being of the form in (2.6) as there are many short-rate models that do not provide an ATS. The key is for the drift and diffusion of the short-rate to be affine themselves (e.g., Björk (1997), Duffie (2001, ch. 7)).

For instance, given a short-rate model as the one defined in (2.5), then r is said to provide an ATS with bond price given by (2.6) if and only if the drift and diffusion are of the form

$$b(t,x) = \alpha(t) + \beta(t)x \quad \text{and} \quad \sigma^2(t,x) = \gamma(t) + \delta(t)x \tag{2.9}$$

for some continuous functions $\alpha, \beta, \gamma, \delta$ and the functions A and B satisfies the system of ordinary differential equations (ODEs), for all $t < T$, (e.g., Filipović (2009, ch. 5), Brigo and Mercurio (2012, ch. 3))

$$\begin{aligned} \frac{\partial}{\partial t}A(t,T) &= \frac{1}{2}\gamma(t)B^2(t,T) - \alpha(t)B(t,T), \quad A(T,T) = 0 \\ \frac{\partial}{\partial t}B(t,T) &= \frac{1}{2}\delta(t)B^2(t,T) - \beta(t)B(t,T) - 1, \quad B(T,T) = 0. \end{aligned} \tag{2.10}$$

An example of an affine short-rate model is the one-factor Vasicek model given by

$$dr(t) = \lambda(\theta - r(t))dt + \eta dW^*(t). \tag{2.11}$$

From coefficient matching between (2.11) and (2.9), it is easy to see that $\alpha(t) = \lambda\theta, \beta(t) = \lambda, \gamma(t) = \eta, \delta(t) = 0$, which by solving the corresponding system of ODEs in (2.10) yields A and B equal to

$$A(t,T) = \left(\theta - \frac{\eta^2}{2\lambda^2}\right)(T - t - B(t,T)) + \frac{\eta^2}{4\lambda}B(t,T)^2, \tag{2.12}$$

$$B(t,T) = \frac{1}{\lambda}\left(1 - e^{-\lambda(T-t)}\right). \tag{2.13}$$

With A and B known, the bond price is simply calculated using (2.6).

The Vasicek model has been widely used in finance due to its affine structure and Gaussian analytical tractability giving closed form expression for bond prices. However, because of its poor fit to the initial term structure and positive theoretical probability of reaching negative levels, multiple alternative short-rate models have been introduced. Two alternative affine short-rate models that will be considered in more detail are that of Cox, Ingersoll and Ross, and Hull and White's extended Vasicek model.

The Hull-White (1994) Model

In 1994, Hull and White proposed an extension to the Vasicek by allowing θ to be a function of time (Hull and White, 1994). By doing so, this enables the dynamics of the short-rate to exactly fit the term structure of interest rates being currently observed in the market. The extended Vasicek model is given by

$$dr(t) = \lambda(\theta(t) - r(t))dt + \eta dW^*(t), \tag{2.14}$$

where η determines the overall level of volatility and the reversion rate parameter λ determines the relative volatilities. A high value of λ causes the short-term rate movements to damp out quickly, reducing the long-term volatility.

To fit the initial term structure, θ is taken equal to

$$\theta(t) = f(0,t) + \frac{1}{\lambda}\frac{\partial}{\partial t}f(0,t) + \frac{\eta^2}{2\lambda^2}\left(1 - e^{-2\lambda t}\right). \tag{2.15}$$

where $f(0,t) = -\frac{\partial}{\partial t}\log P(0,t)$ denotes the market instantaneous forward rate at time 0 for the maturity t. This expression ensure a perfect fit of the initial yield curve at each maturity.

Just as for the regular one-factor Vasicek model, the Hull-White model also has an explicit solution for $t \geq 0$ given by

$$\begin{aligned} r(t) &= e^{-\lambda t}r(0) + \lambda\int_0^t e^{-\lambda s}\theta(s)ds + \eta\int_0^t e^{-\lambda s}dW^*(s) \\ &= \psi(t) + \eta\int_0^t e^{-\lambda s}dW^*(s), \end{aligned} \tag{2.16}$$

where

$$\psi(t) = f(0,t) + \frac{\eta^2}{2\lambda^2}\left(1 - e^{-\lambda t}\right)^2. \tag{2.17}$$

Therefore, $r(t)$ conditional of $\mathcal{F}(0)$ is normally distributed with its mean and variance given, respectively, by

$$\mathrm{E}^{\mathrm{Q}}(r(t)|\mathcal{F}(0)) = \psi(t) \tag{2.18}$$

$$\mathrm{Var}^{\mathrm{Q}}(r(t)|\mathcal{F}(0)) = \frac{\eta^2}{2\lambda}\left(1 - e^{-2\lambda t}\right). \tag{2.19}$$

By inspection it is possible to conclude that the final term on the right hand side in (2.16) is the explicit solution of a regular zero mean Vasicek model. This makes it possible to rewrite (2.14) (e.g., Brigo and Mercurio (2012)) as

$$r(t) = \tilde{r}(t) + \psi(t), \tag{2.20}$$

where ψ is given in (2.17) and

$$d\tilde{r}(t) = -\lambda\tilde{r}(t)dt + \eta dW^*(t), \quad \tilde{r}(0) = 0. \tag{2.21}$$

The advantage of this transformation is the reduction of the Hull-White model to a regular mean-reverting Vasicek model accompanied by the independent deterministic process ψ. This also allows the Hull-White bond price to be expressed in terms of market observable bond prices and the corresponding zero mean Vasicek bond price.

By solving (2.10), the ATS functions A and B is given by

$$A(t,T) = -\frac{\eta^2}{2}\int_t^T B^2(s,T)ds + \lambda\int_t^T \theta(s)B(s,T)ds,$$
$$B(t,T) = \frac{1}{\lambda}\left(1 - e^{-\lambda(T-t)}\right),$$

with θ given by (2.15). The substitution of A and B into (2.6) leads to the Hull-White implied bond price of the form[6] (2.6) and can be written as

$$P(t,T) = \frac{P(0,T)}{P(0,t)}\exp\left\{-\frac{\eta^2}{2\lambda^2}\int_t^T\left(1 - e^{-\lambda s}\right)^2\right\}\tilde{P}(t,T), \tag{2.22}$$

where $\tilde{P}$ denotes the respective bond price of the Vasicek model $\tilde{r}$.

Despite the improvement of being able to perfectly fit the initial yield curve, the Hull-White still allows for r to become negative. A more detailed discussion as to whether this is a problem or not will be discussed further in Section 2.1.2.

The Cox, Ingersoll, and Ross (1985) Model

The Cox, Ingersoll, and Ross process (CIR) introduce by Cox et al. (1985) proposed a short-rate model including a square-root term into the diffusion coefficient of the Vasicek model in order to ensure positivity of the process. This phenomena was first studied by Feller (1951), who discovered that by choosing the parameters of the model in a certain way would allow the model to stay positive at all times (considering the initial condition is strictly positive). The positivity and the models analytically tractability has made the model widely popular when modelling interest rates and volatility.

In this thesis the CIR model will be used to model volatility. However, since it was first proposed as a short-rate model, the model and its key features will be presented here.

The risk-neutral dynamics of the CIR process is given by

$$dr(t) = \lambda(\theta - r(t))dt + \eta\sqrt{r(t)}dW^*(t), \quad r(0) = r_0, \tag{2.23}$$

for some positive constants $r_0, \lambda, \theta, \eta$. By choosing the parameters such that $2\lambda\theta \geq \eta^2$, Feller (1951) proved the drift to be sufficiently large to make the origin inaccessible. In

[6] A full derivation of the Hull-White model and its parameters can be found in Appendix A.

some literature this condition is therefore referred to as the *Feller property* where the model is said to fulfill the Feller property if $2\lambda\theta \geq \eta^2$. If the condition is not fulfilled, i.e., $2\kappa\theta < \eta^2$, it can be shown that the origin is accessible making it possible for the process to reach zero.

From the result of Feller (1951), Cox et al. (1985) showed that the conditional distribution of $r(t)$ given $r(s)$ for some $t > s$ involves a non-central chi-squared distribution (e.g., Brigo and Mercurio (2012), Glasserman (2004), Andersen (2007)). The fact that the density of r is known has since then been utilised to develop efficient simulation methods of the process, as will be seen in Chapter 3.

2.1.2 Displacement and Negative Rates

As mentioned above, one of the arguable drawbacks with the Hull-White short-rate is its positive probability of reaching negative rates. Using the fact that under the Hull-White model, r is normally distributed it is possible to show that the probability of r being negative at time t is given by

$$\mathrm{Q}(r(t) < 0) = \Phi\left(-\frac{\psi(t)}{\sqrt{\frac{\eta^2}{2\lambda}\left(1 - e^{-2\lambda t}\right)}}\right), \tag{2.24}$$

where Φ denotes the standard normal distribution function and $\mathrm{Q}(\cdot)$ that the probability is taken under the risk-neutral measure. Historically this has been seen as a major drawback of the Hull-White model due to the fact that zero and negative rates has not been observed in the market. However, events following the 2008 financial crisis has challenged the assumption of a strictly positive interest rate. Since then, small negative rates have been observed in multiple countries, including the United States, Switzerland and Sweden. This has sparked the interest in studying the effect of incorporating negative rates in the modelling of interest rate.

One method allowing for negative rates is the so-called displaced diffusion (e.g., Rebonato (2004)). The intuition behind displacements is the fact that by shifting a process $r(t)$ with a fixed constant δ does not change the dynamics of the model; clearly since it holds that $d(r(t) + \delta) = dr(t)$. Consider the CIR process as an example, by assuming that the Feller property is fulfilled, it is known that the CIR process cannot enter into negative territories. However, by defining $x(t) = r(t) + \delta$, the square-root of the diffusion part of $x(t)$ will not reach zero until $r(t) = -\delta$ which would allow $r(t)$ to take negative values up to a level δ. Albeit, since the displacement coefficient is an exogenous parameter it needs to be estimated through market data in order to be implemented. In comparison, the Hull-White model naturally allows for the interest rate to reach negative levels, and with the time dependent mean-level parameter, the Hull-White model ensures a perfect fit to the initial yield curve.

2.2 Equity Models

The second asset class that will be considered in this thesis is equity. One of the first models used to model stock dynamics is the one used in the seminal paper of Black and Myron (1973). They assumed the dynamics of the stock to follow a regular geometric Brownian motion given by

$$\frac{dS(t)}{S(t)} = \mu dt + \sigma dW(t), \tag{2.25}$$

with μ and σ being constant quantities and W a regular Brownian motion under P, leading to prices being log-normally distributed. Its analytically tractability and explicit pricing formula for European options has made the model widely used both in finance and research.

However, over time multiple assumptions in the Black-Scholes model has been found to be to crude and restrictive. In particular, studies have shown that there are at least three systematic and persistent distinctions between Black-Scholes and what is observed in market returns (e.g., Carr and Wu (2004), Kou (2002));

(I) Asset prices jump, resulting in leptokurtic (heavy tailed) return distributions different from the one seen in the Black-Scholes model.

(II) Return volatilities tends to display significant variability over time, resulting in cluster formations.

(III) Returns and volatilities are correlated, often negative for equities, leading to asymmetric empirical distributions with skewed distribution peak relative to its mean.

Due to these shortcomings, many studies have attempted to modify the Black-Scholes model in order to address the above three empirical stylized facts.

2.2.1 Stochastic Volatility

The first class of equity models that will be considered, extending the model of Black and Scholes, is the stochastic volatility model. The problem with (2.25), as mentioned above, is the assumption that volatility stays constant over time, which is not the case for market returns, (II). The stochastic volatility models 'cope' with this problem by directly modelling the instantaneous variance of the stock as a diffusion process itself. By doing so, the stochastic volatility extension manage, not only to capture the time variability of the volatility, (II), but it also allows for dependence between the stock and its volatility, (III), by introducing a linear combination between the two driving processes.

The Heston (1993) model

One of the more popular stochastic volatility models used to model equities is the Heston model where the instantaneous variance is taken to follow a CIR process (Heston, 1993). The objective dynamics of the model reads

$$\begin{cases} \frac{dS(t)}{S(t)} & = \mu dt + \sqrt{V(t)} dW_S(t) \\ dV(t) & = \kappa(\theta - V(t))dt + \nu\sqrt{V(t)} dW_V(t), \\ \langle dW_S(t) dW_V(t) \rangle & = \rho dt \end{cases} \tag{2.26}$$

where, V denotes the instantaneous variance, κ its speed of reversion and θ the long-term variance, ν denotes the volatility of variance and ρ the correlation between the two driving Brownian motions.

2.2.2 Jump-diffusions

Financial models with jumps fall into two categories. In the first category, called jump-diffusion models or finite-activity jump processes, the evolution of the model is given by a diffusion process, punctuated by jumps at random times. The jumps are prominently modeled through a compounded Poisson process with finitely many jumps in each time interval. The second category consist of models with an infinite number of jumps in a time interval, and are therefore usually referred to as infinitely activity models. In this thesis attention will be devoted to models belonging to the first category, in particular the Merton model. The main reason for this is their close connection to regular diffusion and stochastic volatility models.

The Merton (1976) Model

The model introduced by Merton (1976) was the first model to consider an additional jump process to the regular diffusion models. Merton applied this to option pricing in order to model the idiosyncratic shocks affecting the price of the asset underlying the option.

The Merton model can be specified though the SDE

$$\frac{dS(t)}{S(t-)} = \mu dt + \sigma dW(t) + dJ(t), \tag{2.27}$$

where, just as in (2.25), μ and σ are assumed to be constant and W a regular Brownian motion, $S(t-)$ denotes the value right before the jump defined as the left limit as $S(t-) := \lim_{u \uparrow t} S(u)$. The jump term J is given by a compounded Poisson process defined as

$$J(t) = \sum_{j=1}^{N(t)} Y_j - 1. \tag{2.28}$$

Here, $\{Y_j\}_{j\in N(t)}$ is a series of log-normally distributed random variables controlling the magnitude of the jump through its log-mean and standard deviation given by μ_J and σ_J respectively. N is a regular Poisson process with fixed intensity λ controlling the interarrival time of the jumps where a larger λ corresponds to a larger number of expected jumps observed within a time interval. Moreover, W, N and Y are all assumed to be mutually independent of one another.

Another model part of the finite-activity jump category is the one studied by Kou (2002), proposing $\log Y_j$ to be double-exponentially distributed. The Kou and Merton model are in many cases equal, they both ensure a closed-form solution for standard put and call options, and capture the leptokurtic behavior seen in market data. The main advantage with the Kou model compared to the Merton model is the memoryless property that the exponential random variable brings, enabling analytical expressions for expectations involving first passage times, i.e., path-dependent options (Kou (2002), Cont and Tankov (2004, ch. 4)). The reason for considering the Merton model over that of Kou, is Mertons' practical popularity and its convenient transition to the stochastic volatility jump-diffusion models (see Figure 2.1).

2.2.3 Stochastic Volatility Jump-diffusions

So far it has been shown that the stochastic volatility model successfully solves the problem with constant volatility and in turn the correlation between the volatility and the stock. However, it fails to capture market extreme events, something that the jump-diffusion model does. The stochastic volatility jump-diffusion model introduced by Bates (1996) copes with this problem by adding a Merton based jump component to the Heston stochastic volatility model. In doing so, the Bates model allows for rich return dynamics encompassing all three shortcomings of Black and Scholes.

The Bates (1996) Model

The objective dynamics proposed by Bates has the following form

$$\begin{cases} \frac{dS(t)}{S(t-)} & = \mu dt + \sqrt{V(t)}dW_S(t) + dJ(t), \\ dV(t) & = \kappa(\theta - V(t))dt + \nu\sqrt{V(t)}dW_V(t), \\ \langle dW_S(t)dW_V(t)\rangle & = \rho dt, \end{cases} \tag{2.29}$$

where the model parameters are equal to those in (2.26) and (2.27).

Since the introduction of the Bates model, multiple of studies have been put forward, extending the model by introducing a jump specification into the volatility process as well (e.g., Eraker et al. (2003), Eraker (2004), Broadie et al. (2007)). The justification for an additional jump term originate from empirical evidence supporting misspecification of the volatility process in (2.29) studied by Eraker et al. (2003). Inevitably, the complexity of

such a model has induced complex calibration procedures and hence not gained the same practical popularity as the Bates model (Gatheral, 2006, ch. 5).

The connection between the presented equity model is outlined in Figure 2.1 with arrows and attached parameter equalities to transition between the models.

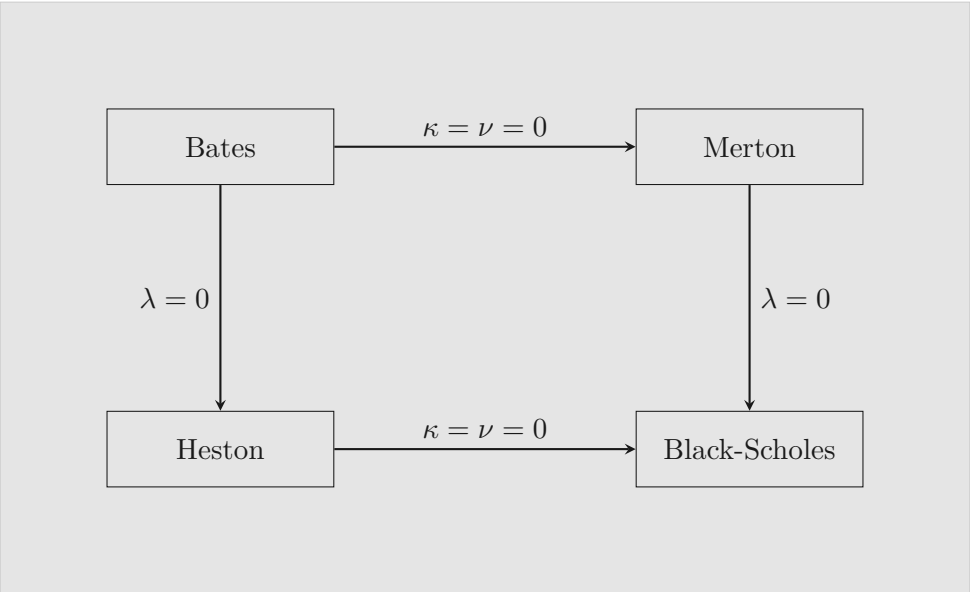


Figure 2.1: Overview of the connection between the equity models. Arrows with attached parameter equalities indicates transition to another model specification. λ - jump intensity, κ - mean reversion speed of variance and ν - volatility of variance.

2.3 Dependence Modelling

When working with multiple time series, it is often of great interest to understand how time series comove and depend on each other. In this section, statistical measures connected to model dependence will be presented. The section covers the more basic measures such as correlation and autocorrelation to more complex dependence modelling involving copulas and tail dependence and how they will be relevant for this thesis.

2.3.1 Correlation and Dependence

The most familiar measure of dependence is the Pearson correlation coefficient. By considering two random variables X and Y, the correlation coefficient is defined as the quotient of the covariance of the two variables and the product of their standard deviations. We

formally define it as

$$\rho_{X,Y} \coloneqq \mathrm{Corr}(X,Y) = \frac{\mathrm{Cov}(X,Y)}{\sqrt{\mathrm{Var}(X)}\sqrt{\mathrm{Var}(Y)}}. \tag{2.30}$$

This definition of correlation measures the strength of linear dependence between two random variables. It can be shown that $\rho_{X,Y} \in [-1,1]$, with $\rho_{X,Y} = 0$ meaning that X and Y are uncorrelated. Here we emphasize the fact that zero correlation does not imply independence whilst the converse is always true. If $|\rho_{X,Y}| = 1$ then it is equivalent to saying that X and Y are perfectly linearly correlated, i.e., $X = a + bY$ a.s. for some $a \in \mathbb{R}$, $b \neq 0$ and $b > (<)0$ for positive (negative) linear dependence.

In time series analysis it is often desirable to examine the linear dependence between a process and its past. This is refereed to as autocorrelation, and by considering the time varying stochastic process $\{X_t\}$, the autocorrelation between the two time points t and $t+\ell$ is defined as

$$\rho(\ell) = \frac{\mathrm{Cov}(X(t), X(t+\ell))}{\sqrt{\mathrm{Var}(X(t))}\sqrt{\mathrm{Var}(X(t+\ell))}}, \tag{2.31}$$

where ℓ is called the time lag.

2.3.2 Copulas

The primary idea of a copula is to model the dependence of multiple random variables using a joint multivariate distribution function. In doing so, the copula relies on two important distributional properties. The first one, known as the *probability integral transform*, states that, given a random variable X with continuous distribution function F it holds that $F(X)$ is uniformly distributed on $(0,1)$. The second property called the *quantile-transform* states that if U is a standard uniform random variable and F is any continuous and strictly monotonic distribution function, then $F^{-1}(U)$ has a F-distribution. In any dimension d, consider the random vector $(X_1, \ldots, X_d)^\top$ with marginal distribution functions implied by the *quantile-transform*, then it holds that

$$(U_1, \ldots, U_d)^\top = (F_1(X_1), \ldots, F_d(X_d))^\top, \tag{2.32}$$

with $U_i,\ i = 1, \ldots, d$ being standard uniforms.

The copula, denoted C, of $(X_1, \ldots, X_d)^\top$ is now defined as the joint distribution function of (2.32) such that $C : [0,1]^d \to [0,1]$ given by

$$C(\mathbf{u}) = C(u_1, \ldots, u_d) = \mathrm{P}(F_1(X_1) < u_1, \ldots, F_d(X_d) < u_d), \quad \mathbf{u} \in [0,1]^d. \tag{2.33}$$

Gaussian Copula

In finance, one of the most used copulas is the Gaussian copula. Given a dimension d, the d-dimensional Gaussian copula is given by

$$C_R^{Ga}(\mathbf{u}) = \Phi_R(\Phi^{-1}(u_1), \ldots, \Phi^{-1}(u_d)), \quad \mathbf{u} \in [0,1]^d, \tag{2.34}$$

where Φ is the inverse distribution function of a standard normal random variable and Φ_R is the joint distribution function with zero mean and covariance matrix equal to the correlation matrix $R = \rho(\Sigma)$, that is, $R_{ij} = \rho_{ij}$.

Since the Gaussian distribution is fully characterized by its first two moments we see that C_R^{Ga} is fully parametrized by the $\frac{1}{2}d(d-1)$ value of R.

Student's t Copula

In the same fashion as for the Gaussian copula, it is possible to define a copula related to the Student's t distribution. Consequently, the d-dimensional t copula takes the form

$$C_{\nu,R}^t(\mathbf{u}) = t_{\nu,R}(t_\nu^{-1}(u_1), \ldots, t_\nu^{-1}(u_d)), \quad \mathbf{u} \in [0,1]^d, \tag{2.35}$$

where, just as in the Gaussian case, t_ν denotes the distribution function of a standard univariate t distribution with ν degrees of freedom and $t_{\nu,R}$ is the joint distribution function with zero mean and correlation matrix $R = \rho(\Sigma)$. Just as for the univariate case, it holds that the t copula coincides with the Gaussian in the limit of $\nu \to \infty$.

Archimedean Copulas

Both the Gaussian and the t copula are so-called *implicit copulas*. This is because, even though they have well-known multivariate distribution functions, they themselves do not offer closed form expressions. An associative class of copulas with explicit algebraic expression are the Archimedean copulas defined[7] as

$$C(\mathbf{u}) = C(u_1, \ldots, u_d; \psi) = \psi(\psi^{-1}(u_1) + \ldots + \psi^{-1}(u_d)), \quad \mathbf{u} \in [0,1]^d. \tag{2.36}$$

Here ψ is known as the Archimedean copula generator and ψ^{-1} its inverse. $\psi : [0,\infty) \to [0,1]$ is a continuous, strictly decreasing and convex function on $[0, \psi^{-1}(0)]$ with $\psi(0) = 1$ and $\psi(\infty) = 0$ with $\psi^{-1} \coloneqq \inf\{u : \psi(u) \leq t\}$ (McNeil et al., 2015, ch. 7). Moreover, Kimberling (1974) gave the necessary and sufficient condition ensuring that ψ generates a copula of the form (2.36) in any dimension $d \geq 2$ if ψ is completely monotonic on $[0,\infty)$, that is, $\psi \in \mathrm{C}^\infty(0,\infty)$ with $(-1)^k \frac{d^k}{dt^k}\psi(t) \geq 0$ for any $k \in \mathbb{N}_0,\ t \in (0,\infty)$.

The complete monotonicity of ψ is here key to ensure that (2.36) forms a true copula. Following the Berstein's theorem on completely monotone functions (e.g., Feller (1971, ch. 13)), allows to construct a relationship between ψ and the Laplace-Stieltjes transform. By letting ψ be defined as above and considering a strictly positive random variable X with distribution function F on $\mathbb{R}^+$ with $F(0) = 0$, then the Laplace-Stiltjes transform is given by

$$\mathcal{L}_t\{F(x)\} = \int_0^\infty e^{-tx} dF(x) = \psi(t), \tag{2.37}$$

[7] Here we, in contrast to other literature, define the Archimedean copula through ψ rather than its inverse, see McNeil et al. (2015)

ensuring that, given a uniformly distributed random variable U on $(0,1)$ then $\psi\left(-\log(U)/X\right)$ is also uniformly distributed on $(0,1)$,

$$\begin{aligned}
\mathrm{P}\left(\psi\left(\frac{-\log U}{X}\right) < u\right) &= \mathrm{E}\left(\mathrm{P}\left(\psi\left(\frac{-\log U}{X}\right) < u \middle| X\right)\right) \\
&= \mathrm{E}\left(\mathrm{P}\left(U < \exp\left(-\psi^{-1}(u)X\right) \middle| X\right)\right) \\
&= \mathrm{E}\left(\exp(-\psi^{-1}(u)X)\right) \\
&= \psi(\psi^{-1}(u)) = u,
\end{aligned}$$

(Hult et al., 2012).

The Archimedean copula that will be considered in this thesis is the Clayton copula, where the generator is given by

$$\psi_\theta(t) = (1+\theta t)_+^{-1/\theta}, \quad t \in (0,\infty) \tag{2.38}$$

and is completely monotonic for any $\theta \geq 0$, where, $\theta = 0$ should be interpreted as the limit (from either side) of $\psi_0(t) = \lim_{\theta\to 0}\psi_\theta(t)$ (McNeil and Nešlehová, 2009). Using the generator given in (2.38), the explicit expression for the Clayton copula in dimension d thus becomes

$$C_\theta^{Cl}(\mathbf{u}) = \left(1 + \sum_{i=1}^{d}\left(u_i^{-\theta} - 1\right)\right)_+^{-1/\theta}, \quad \mathbf{u} \in [0,1]^d. \tag{2.39}$$

When $d = 2$, Genest and MacKay (1986) proved that θ is related to the rank correlation coefficient Kendall's tau as $\theta = \frac{2\tau}{1-\tau}$. For the special case of $(X_1, X_2)^\top$ belonging to an elliptical distribution[8], e.g., Gaussian or Student's t, Kendall's tau is related to Pearson's rho as

$$\tau = \frac{2}{\pi}\arcsin\rho. \tag{2.40}$$

(Hult et al., 2012, ch. 9).

Coefficient of Tail Dependence

Unlike linear correlation that depends on the outcome of two random variables, the coefficient of tail dependence only depends on the copula of the two random variables. More precisely, the coefficient of tail dependence provides a measure of the dependence between two random variables in the tails of the distribution. Roughly speaking, the coefficient of tail dependence measure the probability that a random variable exceeds a certain threshold given that another random variable has already exceeded such a threshold.

[8] For a detailed review of elliptical distributions see e.g., Cambanis et al. (1981)

In risk management, a lower tail dependence is often of particular interest as it measures the probability of experiencing an extreme loss in one asset, given that a substantial loss has already occurred in another. For that reason, the exploration of tail dependence will be narrowed down to lower tail dependence.

Given two variables X and Y with distribution functions given by F_X and F_Y respectively, then the coefficient of lower tail dependence is defined as

$$\lambda_l \coloneqq \lambda_l(X,Y) = \lim_{q\downarrow 0} \mathrm{P}\left(X < F_X^{-1}(q) | Y < F_Y^{-1}(q)\right) = \lim_{q\downarrow 0} \frac{C(q,q)}{q}, \tag{2.41}$$

provided the limit $\lambda_l \in [0,1]$ exists[9]. If $\lambda_l \in (0,1]$, X and Y are said to show lower tail dependence. If, on the other hand, $\lambda_l = 0$ we say that X and Y has no tail dependence meaning that they are *asymptotically independent*.

It can be shown (e.g., McNeil et al. (2015, ch. 7)) that the Gaussian copula is asymptotically independent for every $\rho \neq 1$. In contrast, by considering the t copula with off-diagonal correlation element ρ, then

$$\lambda_l = 2t_{\nu+1}\left(-\sqrt{\frac{(\nu+1)(1-\rho)}{(1+\rho)}}\right) \tag{2.42}$$

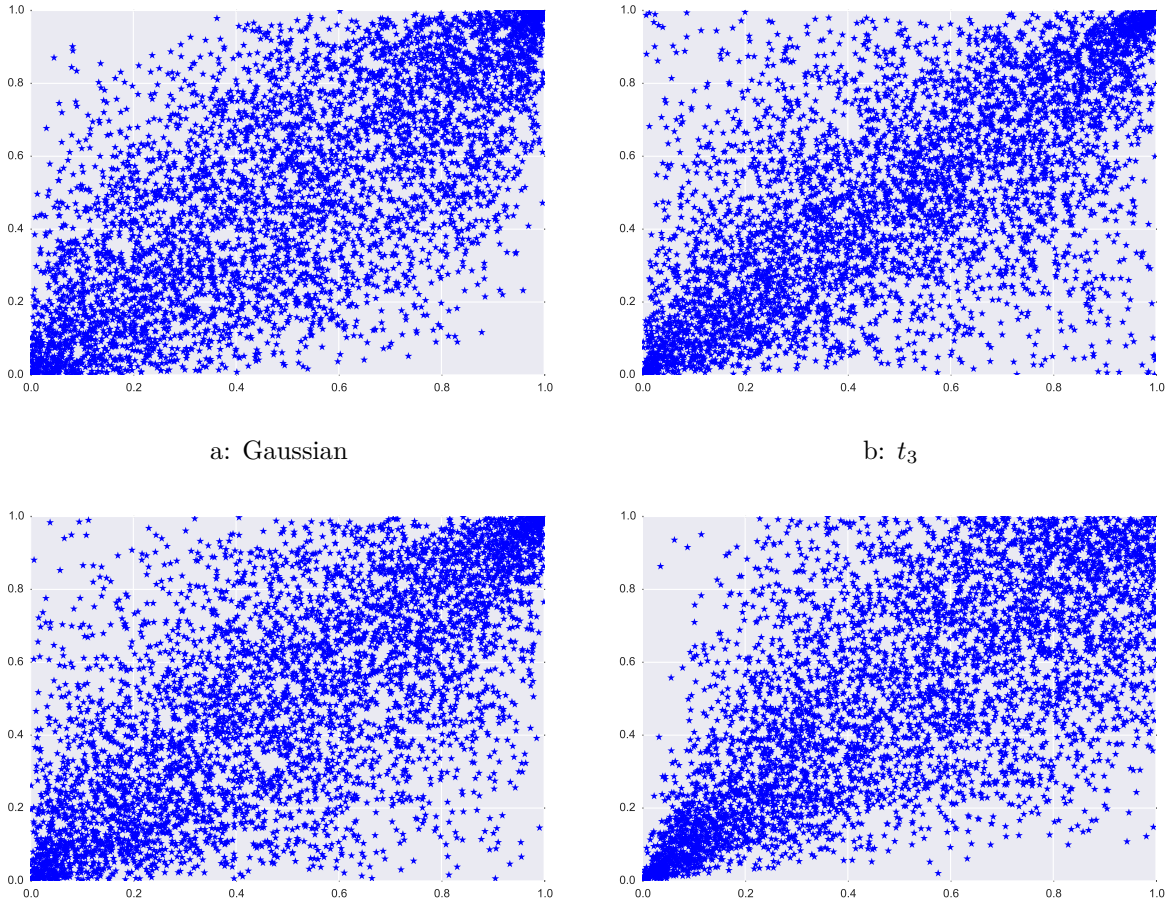
for all $\rho \neq -1$. Perhaps counterintuitively, even if $\rho < 0$, the t copula experience tail dependence. Finally, for the Clayton copula, easy calculations shows that the lower tail dependence of the Clayton copula is given by,

$$\lambda_l = \lim_{q\downarrow 0} \frac{C_\theta^{Cl}(q,q)}{q} = \lim_{q\downarrow 0} \frac{1}{q}\left(2q^{-\theta}-1\right)^{-1/\theta} = \lim_{q\downarrow 0}\left(2-q^{\theta}\right)^{-1/\theta} = 2^{-1/\theta}, \tag{2.43}$$

(McNeil et al., 2015).

A visualisation of the three presented copulas can be seen in Figure 2.2 with corresponding value of the coefficient of lower tail dependence.

[9] Given that, there in practice only exist a finite number of observed outcomes of X and Y, it is common practice to consider λ_l for some fixed quantile $q > 0$ rather than the limit.



a: Gaussian

b: t_3

c: t_7

d: Clayton

Figure 2.2: a: Simulated bivariate Gaussian copula with $\rho = 0.7$ and $\lambda_l = 0$. b: Simulated bivariate t copula with $\nu = 3$ and $\lambda_l = 0.45$. c: Simulated t bivariate copula with $\nu = 7$ and $\lambda_l = 0.27$. d: Simulated bivariate Clayton copula with $\theta = 2$ and $\lambda_l = 0.7$. All simulations were done using a sample size of $N = 5000$.

2.3.3 Stationarity and Ergodicity

Two of the fundamental concepts in time series analysis are stationarity and ergodicity. Both of these properties will lie as a base in the analysis to come and are therefore being formally introduced below.

Stationarity

A stochastic process $\{X(t)\}$ is said to be strongly or strictly stationary if the joint distribution function of $(X(t_1), \ldots, X(t_n))$ is identical to that of $(X(t_1 + \ell), \ldots, X(t_n + \ell))$ for all n, ℓ and $t_1, \ldots, t_n$. That is,

$$(X(t_1), \ldots, X(t_n)) \overset{\text{d.}}{=} (X(t_1 + \ell), \ldots, X(t_n + \ell)), \tag{2.44}$$

meaning that the distribution of $\{X(t)\}$ is invariant under time shift.

This behavior is desirable when, for instance, working with scenario simulations as it allows to interpret the time t distribution of $\{X(t)\}$ as that of $\{X(t + \ell)\}$. Consequently, if the true time t stationary distribution is known and accessible, it is possible to forecast future distributional return properties using the properties implied by the information up to time t.

Strong stationarity is often too restrictive and hard to verify empirically as it requires all moments of the process to be constant over time. Instead, a weaker version of stationarity is often assumed, which relaxes the condition of (2.44) applied to the first two moments. Formally, this states that a stochastic process is considered to be weakly stationary if it satisfies the following properties;

- $\mathrm{E}(X(t)) = \mu$, independent of t,
- $\mathrm{Cov}(X(t), X(t+\ell)) =: \gamma(t, t+\ell) = \gamma(\ell)$, is finite, time-invariant, and only dependent on the time lag ℓ.

For $\ell = 0$, the second property thus also implies that the variance is constant over time as well. Henceforth, when talking about stationary process, it will refer to weakly stationary processes.

Ergodicity

In time series analysis, a stochastic process is said to be ergodic if its statistical properties can be deduced from a single, sufficiently long, sample path. For a weakly stationary process $\{X(t)\}$, with mean and autocovariance function given by μ and γ respectively, X is said to be *wide sense ergodic* if the time series sample mean and autocovariance converges in mean-square to their true values, μ and $\gamma(\cdot)$. That is;

$$\begin{aligned} &\lim_{n\to\infty} \mathrm{E}\left((\hat{\mu} - \mu)^2\right) = 0, \\ &\lim_{n\to\infty} \mathrm{E}\left((\hat{\gamma}(\ell) - \gamma(\ell))^2\right) = 0, \quad 0 \leq \ell < n, \end{aligned}$$

where $\hat{\mu} \coloneqq n^{-1} \sum_{i=1}^{n} X_i$ and $\hat{\gamma}(\ell) \coloneqq n^{-1} \sum_{i=1+\ell}^{n} (X_i - \hat{\mu})(X_{i-\ell} - \hat{\mu})$. Indeed there is no reason to believe that the time series sample mean and variance of a stationary series

necessarily would converge to the true values as the independent assumption of the law of large numbers (LLN) and the central limit theorem (CLT) is violated. However, it turns out that if the time series is ergodic, a counterpart to the CLT and LNN can be established.

Beginning with the sample mean, $\hat{\mu}$ is said to be an unbiased and consistent estimator if $\gamma(n) \to 0$ as $n \to \infty$. Indeed it holds that

$$\begin{aligned} \mathrm{E}\left((\hat{\mu}-\mu)^2\right) = \mathrm{Var}(\hat{\mu}) = \mathrm{Var}\left(\frac{1}{n}\sum_{i,j=1}^{n} X_i\right) = \frac{1}{n^2}\sum_{i,j=1}^{n}\mathrm{Cov}(X_i, X_j) \\ = \frac{1}{n}\sum_{|\ell|<n}\left(1-\frac{|\ell|}{n}\right)\gamma(\ell) \leq \frac{1}{n}\sum_{|\ell|<n}|\gamma(\ell)|. \end{aligned}$$

Thus, if $\gamma(n) \to 0$ as $n \to \infty$, then $\lim_{n\to\infty} n^{-1}\sum_{|\ell|<n}|\gamma(\ell)| = 2\lim_{n\to\infty}|\gamma(n)| = 0$, from where $\mathrm{Var}(\hat{\mu}) \to 0$. Moreover, assuming that the autocovariance is absolutely summable[10] with $\sum_{\ell=-\infty}^{\infty}\gamma(\ell) =: \nu^2 > 0$, it is possible to show that the sample mean of $\{X(t)\}$ is asymptotically Gaussian with mean μ and variance given by $n^{-1}\nu^2$ (e.g., Brockwell and Davis (1991, ch. 7)).

To prove that the sample autocovariance converges in mean square is not as easy. However, under suitable assumptions, as will be presented below, it is possible to show that also the time series sample autocovariance serves as a consistent, although biased, estimation of the second moment of X. A detailed proof with necessary and sufficient condition can be found in Hannan (1970, ch. 4) and Hamilton (1994, ch. 7).

Stationary and Ergodic Moving Average process

A moving average process is a time series of the form

$$X(t) = \mu + \sum_{i=0}^{q}\psi_i\varepsilon(t-i), \tag{2.45}$$

where q defines the moving average order, μ is the mean of $\{X(t)\}$ and $\psi_0 = 1$. $\{\varepsilon(t)\}$ denotes a sequence, commonly referred to as the *innovation*, of independent and identically distributed (IID) random variables with mean zero and variance $\sigma^2 < \infty$. By denoting the moving average process of order q by MA(q), it follows from the representation in (2.45) that the autocovariance function is equal to zero for all lags $\ell > q$.

Assuming that $\{X(t)\}$ is MA(∞) with invariant mean μ, and innovation $\{\varepsilon(t)\} \sim \mathrm{IID}(0, \sigma^2)$, then, by rewriting the process as $\{X(t)-\mu\} = \{Y(t)\}$, it holds that

$$\sqrt{n}\hat{\mu}_Y = \sqrt{n}(\hat{\mu}_X - \mu) \xrightarrow{\mathrm{d.}} \mathsf{N}(0, \nu^2), \tag{2.46}$$

[10] That is, $\sum_{\ell=-\infty}^{\infty}|\gamma(\ell)| < \infty$

as $n \to \infty$, where

$$\nu^2 = \gamma_Y(0) + 2\sum_{\ell=1}^{\infty} \gamma_Y(\ell). \tag{2.47}$$

provided that, $\nu^2 > 0$. Further, if it holds that ψ_i is absolutely summable, then by means of the dominated convergence theorem it holds that

$$\nu^2 = \sigma^2 \left(\sum_{i=0}^{\infty} \psi_i \right)^2. \tag{2.48}$$

Finally, suppose that ε also has finite forth moment, then it holds that the sample autocovariance $\hat{\gamma}_X(\ell)$ is asymptotic normal with mean equal to the true autocovariance function $\gamma_X(\ell)$ (Brockwell and Davis, 1991, ch. 7).

2.4 Risk Management

Risk management within the financial sector is the practice of measuring and managing the risk involved in investing in assets and financial securities. This section involves the procedure of measuring and managing financial risk. In particular the risk of large losses connected to market risk will be discussed. The concept of risk measure is presented after which a series of market relevant risk measures is introduced.

2.4.1 Risk Measures

A risk measure is a mapping $\varrho : \mathcal{L} \to \mathbb{R}$ with the interpretation that $\varrho(L)$ gives the amount of capital needed to cover a loss of magnitude L. In this thesis, $\mathcal{L}$ will denote a linear state space defined as

$$\mathcal{L} \coloneqq \left\{ L : L = c + \boldsymbol{\lambda}^\top \mathbf{X},\ c \in \mathbb{R},\ \lambda \in \mathbb{R}^d \right\}, \tag{2.49}$$

for some risk factors $\mathbf{X}$ and corresponding risk weight vector $\boldsymbol{\lambda}$. The axioms that defines ϱ are as follows;

- *Translation Invariance.* For all $L \in \mathcal{L}$ and every $c \in \mathbb{R}$ we have $\varrho(L + c) = \varrho(L) + c$.
- *Subadditivity.* For all $L_1, L_2 \in \mathcal{L}$ we have $\varrho(L_1 + L_2) \leq \varrho(L_1) + \varrho(L_2)$
- *Positive Homogeneity.* For all $L \in \mathcal{L}$ and $\gamma > 0$ we have $\varrho(\gamma L) = \gamma \varrho(L)$
- *Monotonicity.* For all $L_1, L_2 \in \mathcal{L}$ such that $L_1 \leq L_2$, we have $\varrho(L_1) \leq \varrho(L_2)$
- *Convexity.* For all $L_1, L_2 \in \mathcal{L}$ and $0 \leq \gamma \leq 1$, we have $\varrho\left(\gamma L_1 + (1-\gamma)L_2\right) \leq \gamma\varrho(L_1) + (1-\gamma)\varrho(L_2)$

In particular, a risk measure that satisfies the four axioms; *monotonicity*, *translation invariance*, *subadditivity* and *positive homogeneity* is called a *coherent* risk measure and was proposed in the seminal paper of Artzner et al. (1999). Due to the many desirable features, the coherent risk measures has been widely studied.

Two of the more commonly discussed axioms of the coherent risk measures are subadditivity and positive homegeneity. Intuitively, subadditivity is supposed to represent the idea that diversification of a portfolio lowers overall unsystematic risk and is considered a valid and common market practice in reducing portfolio risk. The idea is that, by adding multiple of assets into a common portfolio, the overall risk of the portfolio can be reduced since assets tends to not move in perfect synchrony with one another. A lack of subadditivity could therefore, as pointed out in Artzner et al. (1999), incline institutions to break up into smaller subsidiaries to lower overall risk and regulatory capital requirement, which can give a false sense of security.

One of the debatable points of subadditivity is its indication that $\varrho(\lambda X) \leq \lambda\varrho(X)$ for $\lambda = 1, 2, \ldots$. Since there is no diversification involved in taking a repeated position in the same asset, positive homogeneity require equality. This, even though there are arguments which support requiring $\varrho(\lambda X) \geq \lambda\varrho(X)$ for large values of λ due to liquidity issues.

Value-at-Risk and Expected Shortfall

Value-at-Risk (VaR) is, to this day one of the most widely used risk measures. It gained large attraction *post* the introduction of J.P. Morgan industry-wide risk model *RiskMetrics* in the beginning of the 1990s. Shortly after, the BCBS followed the adoption of RiskMetrics by including Value-at-Risk in their 1996 Basel I amendment and following as the preferred measure to calculate market risk in the 1999 Basel II Accord.

For a given degree of confidence $\alpha \in [0, 1]$, the α-VaR of an investment is defined as the largest amount l such that, with probability $1-\alpha$ the loss will not exceed l over a specified period of time. Formally, this is defined as

$$\mathrm{VaR}_\alpha(L) = \inf\{l \in \mathbb{R} : \mathrm{P}(L > l) \leq 1 - \alpha\} = \inf\{l \in \mathbb{R} : F_L(l) \geq \alpha\}. \tag{2.50}$$

In probabilistic terms, VaR_α thus correspond to the $100(1-\alpha)$ largest quantile of the loss distribution, that is $\mathrm{VaR}_\alpha = F^-(\alpha)$. Here, F^- denotes the generalised inverse of the distribution function defined as

$$F^-(\alpha) \coloneqq \inf\{x \in \mathbb{R} : F(x) \geq \alpha\}. \tag{2.51}$$

If F is continuous, monotone and strictly increasing, F^- is simply equal to F^{-1} implying that $\mathrm{VaR}_\alpha(L) = F_L^{-1}(\alpha)$.

The drawback and most debatable problem with VaR, besides it not being a coherent risk measure, is its inability to convey the severity of losses that occur with probability less than $1-\alpha$, so-called tail risk. A risk measure that manage to capture this is the

conditional VaR, or more commonly referred to as the Expected Shortfall (ES). For a loss L with $\mathrm{E}(|L|) < \infty$, ES is defined as

$$\mathrm{ES}_\alpha(L) = \frac{1}{1-\alpha}\int_\alpha^1 F_L^-(\gamma)d\gamma, \tag{2.52}$$

where $\mathrm{E}(|L|) < \infty$ ensures that the integral is well defined. As a consequence, ES can be written in terms of VaR as

$$\mathrm{ES}_\alpha(L) = \frac{1}{1-\alpha}\int_\alpha^1 \mathrm{VaR}_\gamma(L)d\gamma. \tag{2.53}$$

The representation in (2.53) therefore means that ES measures the expected loss given that VaR is violated with a given confidence level α. ES thus consider all losses of the loss distribution that occurs with probability less than $1-\alpha$ and therefore convey a more accurate estimate of the magnitude of extreme losses.

The ability to capture tail risk has granted ES large attention within risk management industry, especially after the recent financial crisis. This has lead to the marginalisation of VaR in preference of ES as the preferred approach for calculating market risk in the 2014 Basel III regulatory framework[11].

Diversification Effect

As mentioned above, subadditivity represents the diversification effect of the portfolio in lowering unsystematic risk. Given a loss $L \in \mathcal{L}$ and a risk measure ϱ satisfying translation invariance, subadditivity thus implies

$$\begin{aligned} & \varrho\left(c + \sum_{i=1}^d \lambda_i X_i\right) \leq c + \sum_{i=1}^d \varrho(\lambda_i X_i) \\ \Leftrightarrow \quad & \varrho\left(\sum_{i=1}^d \lambda_i X_i\right) \leq \sum_{i=1}^d \varrho(\lambda_i X_i) \\ \Leftrightarrow \quad & 1 - \frac{\varrho\left(\sum_{i=1}^d \lambda_i X_i\right)}{\sum_{i=1}^d \varrho(\lambda_i X_i)} \geq 0. \end{aligned}$$

This expression will be referred to as the diversification effect, denoted

$$\mathrm{DE}_{\varrho,\lambda}(L) \coloneqq 1 - \frac{\varrho\left(\sum_{i=1}^d \lambda_i X_i\right)}{\sum_{i=1}^d \varrho(\lambda_i X_i)}, \tag{2.54}$$

[11] VaR is however, currently used in the calculation process of solvency capital requirement under Solvency II.

and measures the possible diversification effect implied by position in the risk factors $X_i,\ i = 1, \ldots, d$.

For a coherent risk measure such as ES it is easy to see that the diversification effect is always non-negative and bounded from above by one, whilst for VaR the diversification effect can become negative since VaR does not necessarily satisfy subadditivity. If $\mathrm{DE}_\varrho = 0$, the numerator and the denominator are equal, implying that there is no risk reduction involved in diversification. If instead $\mathrm{DE}_\varrho = 1$, the numerator is equal to zero and thus yielding a total diversification of the risk involved in such a position.

From the expression in (2.54), it is possible to examine the diversification effect by varying the risk factor coefficients λ_i. In the bivariate case, by letting λ_1 and λ_2 equal $\cos\theta$ and $\sin\theta$ respectively it is possible to investigate the behavior of overall diversification effect for different positions,

$$\mathrm{DE}_{\varrho,\theta}(L) = 1 - \frac{\varrho\left(X_1\cos\theta + X_2\sin\theta\right)}{\varrho(X_1\cos\theta) + \varrho(X_2\sin\theta)}. \tag{2.55}$$

The polar coordinate transform consequently implies $|\lambda_i| \leq 1$, with $\lambda < (>)0$ corresponding to a short (long) position of the corresponding risk factor. On the unit circle it is therefore possible to investigate the diversification in different quadrants corresponding to long (L) and short (S) position of the two risk factors. Thus, this allows for investors, given a prior investment preference in the two risk factors, to investigate possible diversification involved in changing the magnitude of their investment in the two risk factors. The position for different values of θ is given by

θ	X_1	X_2
$(0, \frac{\pi}{2})$	L	L
$(\frac{\pi}{2}, \pi)$	S	L
$(\pi, \frac{3\pi}{2})$	S	S
$(\frac{3\pi}{2}, 2\pi)$	L	S

Table 2.1: Position specification of the diversification effect for $\theta \in [0, 2\pi]$.

Here, a short position should be interpreted as possible leverage in the sense that, by going short in one asset allows for a more substantial long position in the other, giving a possible lower capital requirement implied by the risk reduction.

Chapter 3

Methodology

In this chapter, the methodology behind the uncertainty study will be described. In Section 3.1 we introduce the descriptive statistics and risk measures underlying the study along with the estimation procedure that will be used. Section 3.2 continues by considering the simulation methodology of the various models and copula sampling methods used to model their dependence. The chapter concludes by presenting how the method will be used to empirically study the uncertainty of the presented statistics, this in Section 3.3.

3.1 Time Series Analysis

Most financial time series studies involves return series rather than asset prices themselves. The reason for this is twofold. First, in the eyes of an average investor, the return of an asset is a complete characterisation of the performance of an investment. Secondly, from a mathematical point of view, return series tends to have much desirable statistical properties not seen in prices. In this section, statistical analysis related to returns will be presented.

3.1.1 Returns

The definition of asset returns can be defined in two ways; *simple returns* and *continuously compounded returns*. Let $V(t)$ denote the value of a financial asset at time t, then the one-period simple return (*net*) is defined as

$$R(t) \coloneqq \frac{V(t)}{V(t-1)} - 1. \tag{3.1}$$

Rewriting the expression in (3.1) as $1 + R(t)$ (*gross*), the k-period gross return becomes

$$\begin{aligned} 1 + R^{(k)}(t) &= \frac{V(t)}{V(t-1)} \dots \frac{V(t-k+1)}{V(t-k)} \\ &= (1 + R(t)) \dots (1 + R(t-k+1)) \\ &= \prod_{i=0}^{k-1} (1 + R(t-i)), \end{aligned}$$

with the super-script denoting the return period. This means that the k-period simple gross return is simply the product of the k consecutive one-period simple gross returns of the period. The k-period simple net return is given by $R^{(k)}(t) = V(t)/V(t-k) - 1$.

The second definition of returns is that of *continuously compounded returns* or simply log-returns. In terms of the simple gross return, the one period log-return return at time t is given by

$$r(t) \coloneqq \log(1 + R(t)) = \log \frac{V(t)}{V(t-1)} \tag{3.2}$$

and the k-period return by

$$\begin{aligned} r^{(k)}(t) &= \log \left(\prod_{i=0}^{k-1} (1 + R(t-i)) \right) = \sum_{i=0}^{k-1} \log(1 + R(t-i)) \\ &= \sum_{i=0}^{k-1} r(t-i). \end{aligned}$$

In comparison, we see that log-returns are conveniently additive compared to simple returns which are multiplicative. Since, in practice, the actual time interval is often of importance when analysing returns, the way returns are compounded over time is of importance. However, it is not necessarily clear why log-returns are justified in comparison to simple returns. To evaluate this, we consider the k-period simple return once again,

$$\begin{aligned} R^{(k)}(t) &= \left(\prod_{i=0}^{k-1} (1 + R(t-i)) \right) - 1 \\ &= \exp \left(\sum_{i=0}^{k-1} \log(1 + R(t-i)) \right) - 1 \\ &= \exp \left(\sum_{i=0}^{k-1} r(t-i) \right) - 1. \end{aligned}$$

For small values of $r(t)$, it is possible to use a first order Taylor approximation[1] yielding,

$$R^{(k)}(t) \approx \sum_{i=0}^{k-1} r(t-i) = r^{(k)}(t), \tag{3.3}$$

which, indeed is the k-period log-return (Tsay, 2010, ch. 1). This means that in the case of small[2] relative fluctuation in V over a period k, the simple return can preferably be approximated using log returns.

For this thesis, the additive property of returns is a necessary property and for that reason the preferred type of returns to consider are log-returns. Moving forward, if not stated otherwise, log-returns will simply be referred to as returns.

3.1.2 Rolling-Window Analysis

In practice, the amount of data accessible to estimate the stationary distribution might sometimes be inadequate to give a statistically sound estimation of market parameters. To overcome this problem, so-called *rolling analysis* can be implemented. The idea behind rolling analysis is to construct 'new' observations using samples of consecutive observations. For instance, consider a series of returns $\{r(t)\}_{t=1}^{T}$, from which the aim is to estimate an aggregated return over a period ΔT. Then, one can either take the sample $\{r(t)\}_{t=1}^{T}$ and divide it up into $T/\Delta T$ equally sized non-overlapping sub-samples from which aggregated statistical measures of the stationary distribution, e.g., mean and variance, can be calculated. The issue with this procedure is that, if T is not moderately large, the number of ΔT aggregated observation is not sufficiently large to get accurate estimates of the statistics. A preferable approach is instead rolling analysis, where instead of dividing the sample into non-overlapping sub-samples, the aggregated returns is calculated by moving the window forward one observation at a time. The number of aggregated ΔT-returns for this approach then becomes $T - \Delta T + 1$ which is considerably larger than $T/\Delta T$, for $\Delta T < T$.

Assume now that $\{r(t)\}$ is a stationary time series with mean μ_r and autocovariance function $\gamma_r(\cdot)$, then the rolling-returns, denoted by $\{\tilde{r}(t)\}$, is defined as

$$\tilde{r}^{(\Delta T)}(k) := \sum_{t=k}^{k+\Delta T-1} r(t), \tag{3.4}$$

with $k = 1, \ldots, T - \Delta T + 1$. By rewriting the process as

$$\tilde{r}^{(\Delta T)}(k) = \mu_{\tilde{r}} + \sum_{t=k}^{k+\Delta T-1} (r(t) - \mu_r),$$

[1] That is, $\exp(x) \approx 1 + x$.

[2] Indeed, the accuracy of the approximation in (3.3) may not be sufficient in some application, especially for highly dispersive time series, where for large market movements log-return can be misleading in yielding percentage losses greater than 1.

where $\mu_{\tilde{r}} \coloneqq \Delta T \mu_r$, it is easy to see that $\tilde{r}^{(\Delta T)}$ represents a moving average process of order ΔT with mean and variance given by

$$\mathrm{E}\left(\tilde{r}^{(\Delta T)}(k)\right) = \mathrm{E}\left(\sum_{t=k}^{k+\Delta T-1} r(t)\right) = \sum_{t=k}^{k+\Delta T-1} \mathrm{E}(r(t)) = \mu_{\tilde{r}}, \tag{3.5}$$

$$\begin{aligned}\mathrm{Var}\left(\tilde{r}^{(\Delta T)}(k)\right) &= \mathrm{Var}\left(\sum_{t=k}^{k+\Delta T-1} r(t)\right) = \sum_{t=k}^{k+\Delta T-1} \mathrm{Var}(r(t)) + \sum_{s\neq t} \mathrm{Cov}(r_t, r_s) \\ &= \Delta T \gamma_r(0) + 2\sum_{\ell=1}^{\Delta T-1} \gamma_r(\ell),\end{aligned} \tag{3.6}$$

and innovation by $\{r(t) - \mu_r\}$. Assuming further that $\{r(t) - \mu_r\} \sim \mathrm{IID}(0, \sigma_r^2)$, it follows that $\{\tilde{r}(t)\}$ is indeed stationary with mean and variance given by

$$\mathrm{E}\left(\tilde{r}^{(\Delta T)}(k)\right) = \mu_{\tilde{r}} \tag{3.7}$$

$$\mathrm{Var}\left(\tilde{r}^{(\Delta T)}(k)\right) = \Delta T \sigma_r^2 =: \sigma_{\tilde{r}}^2, \tag{3.8}$$

and autocovariance function $\gamma_{\tilde{r}}(\cdot)$, with $\gamma_{\tilde{r}}(\ell)$ equal to zero for every lag $\ell > \Delta T$. The impact of the rolling analysis to the autocovariance can be observed[3] in Figure 3.1. Moreover, given that $\{\tilde{r}(t)\}$ is stationary it is easy to see that the autocovariance is absolutely summable, implying that $\{\tilde{r}(t)\}$ is ergodic with its limiting stationary distribution being equal to the corresponding ΔT-stationary distribution.

The fact that the aggregated rolling-return distribution can be seen as the corresponding stationary distribution is one of the central results of this thesis. As will be described in Section 3.3 it allows the empirically study of the accuracy of estimated statistical quantities implied by the distributional properties of the ΔT-stationary distribution.

Despite the benefit of additional observations and maintained stationarity of the return distribution, the rolling-window procedure do experience some problems. In the case of extreme market observations, for instance that of jumps in the modelling of asset prices, the impact of the observation will have a lasting effect in the construction of new observation as long as it is contained within the width of the window. In the case of monthly frequency series and an annual window (twelve observations), the impact of an extreme value would then propagate over a twelve year period using this procedure. This can thus lead to spurious effect in the newly created return series and its estimated statistical quantity. The severeness of the impact of such an observation is of course problematic to forsee as it depends on both the level of "extremeness" of the observation, as well as the width of

[3] Note that it is the autocorrelation function (ACF) that is depicted, i.e., normalize with the variance, see (3.11).

the rolling-window and the accessible number of observations in the series. In addition to sensitivity to outliers, the rolling-window procedure is affected by the introduction of autocorrelation between observation it introduces. As stated in Section 2.3.3, this is indeed not of any relevance in the limit as the number of observation tends towards infinity. This is, however, never the case meaning that the width of ΔT should be chosen accordingly to the amount of data at hand given that, as demonstrated in Figure 3.1, the endurance of the autocorrelation is proportional to ΔT.

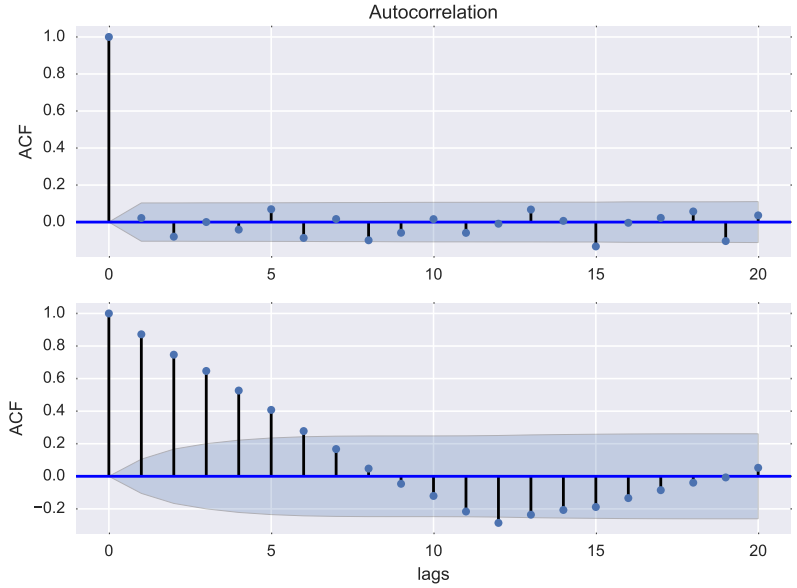


Figure 3.1: Top: Sample autocorrelation function of $r(t)$ simulated using the Black-Scholes model with 95% confidence bounds (shaded area). Bottom: Sample autocorrelation function of $\tilde{r}^{(\Delta T)}(t)$ with $\Delta T = 12$ and 95% confidence bounds (shaded area). Confidence bounds is given by the Ljung-Box test presented in Section 3.1.3.

3.1.3 Moments and Quantiles

For different time horizons ΔT the sample mean and standard deviation will be estimated. Let $\{x_i\}_{i=1}^{N}$ denote a random sample of a stochastic process X, then an estimate of the

first two moments of X is given by

$$\hat{\mu} = \frac{1}{N}\sum_{i=1}^{N} x_i, \tag{3.9}$$

$$\hat{\sigma}^2 = \frac{1}{N-1}\sum_{i=1}^{N}(x_i - \hat{\mu})^2, \tag{3.10}$$

where the factor $N-1$ in the sample variance ensures an unbiased estimator of the variance. Given the sample variance, the sample standard deviation of the distribution is easily attainable by taking the square-root of the sample variance.

As presented in Section 3.1.2, the proposed method relies on the fact that the return series attained using rolling analysis is stationary and ergodic with an autocovariance tending towards zero as the number of lags increases. To test for the significance of the autocorrelation over multiple lags a so-called Ljung-Box test can be performed. By considering the statistic

$$Q(n) = N(N+2)\sum_{\ell=1}^{n}\frac{\hat{\rho}^2(\ell)}{N-\ell},$$

where $\hat{\rho}(\ell)$ is given by

$$\hat{\rho}(\ell) = \frac{\sum_{i=\ell+1}^{N}(x_i - \hat{\mu})(x_{i-\ell} - \hat{\mu})}{\sum_{i=1}^{N}(x_i - \hat{\mu})^2} = \frac{\hat{\gamma}(\ell)}{\hat{\gamma}(0)}, \quad 0 \leq \ell < N-1 \tag{3.11}$$

with $\hat{\mu}$ given by (3.9), the hypothesis test follows by testing the null hypothesis $\mathcal{H}_0 : \rho_1 = \ldots = \rho_n = 0$ against the alternative hypothesis $\mathcal{H}_a : \rho_i \neq 0$ for some $i \in \{1, \ldots, n\}$. Under the null hypothesis the statistic $Q \sim \chi^2_n$, that is chi-squared distributed with n degrees of freedom. For a specified confidence level α, the null-hypothesis is rejected in preference for the alternative hypothesis if

$$Q(n) > \chi^2_n(\alpha).$$

In addition to moments and autocorrelation, sample quantiles are needed to estimate VaR and ES. In doing so, consider the ordered statistic $\{x_{(i)}\}_{i=1}^{N}$ of $\{x_i\}_{i=1}^{N}$, given by $x_{(1)} \leq \ldots \leq x_{(N)}$, then the empirical quantile function is defined as

$$\hat{F}^{-1}_{N,X}(\alpha) = x_{(\lfloor (N-1)\alpha+1 \rfloor)}, \quad \alpha \in (0,1), \tag{3.12}$$

where[4] $\lfloor u \rfloor \coloneqq \sup\{n \in \mathbb{N} : n \leq u\}$, that is, the largest integer less then or equal to u.

[4] This definition is thus in accordance with the quantile function of an empirical distribution in contrast to computer software which tends to use linear interpolation between observation.

With the empirical quantile function defined, the sample VaR and ES of $\{x_{(i)}\}_{i=1}^N$ is easily computed by considering

$$\widehat{\mathrm{VaR}}_\alpha(X) = \hat{F}^{-1}_{N,X}(\alpha), \quad \alpha \in (0,1), \tag{3.13}$$

$$\widehat{\mathrm{ES}}_\alpha(X) = \frac{\sum_{i=\lfloor (N-1)\alpha+1 \rfloor}^N x_{(i)}}{N - \lfloor (N-1)\alpha + 1 \rfloor}, \quad \alpha \in (0,1), \tag{3.14}$$

(Hult et al., 2012, ch. 7).

3.1.4 Correlation and Diversification

In addition to the statistical quantities and the risk measures linked to univariate time series, some bivariate statistics will be considered. In a bivariate setup the correlation and the diversification effect are of interest. Starting with the correlation coefficient which, much equal to the sample autocorrelation in (3.11), for a bivariate sample $\{(x_i, y_i)\}_{i=1}^N$ of (X, Y) is given by

$$\hat{\rho}_{x,y} = \frac{\sum_{i=1}^N (x_i - \hat{\mu}_x)(y_i - \hat{\mu}_y)}{\sqrt{\sum_{i=1}^N (x_i - \hat{\mu}_x)^2}\sqrt{\sum_{i=1}^N (y_i - \hat{\mu}_y)^2}}, \tag{3.15}$$

where, once again, $\hat{\mu}$ is given by (3.9).

For the diversification effect, recall from Chapter 2 that by letting the weights of the risk factors equal $\cos\theta$ and $\sin\theta$, it is possible to evaluate the diversification for a bivariate sample. For a given risk measure ϱ and a loss L given by the two risk factors $\tilde{X}_\theta := X\cos\theta$, $\tilde{Y}_\theta := Y\sin\theta$ the sample diversification is, for some θ, given by

$$\widehat{\mathrm{DE}}_{\hat{\varrho},\theta}(L) = 1 - \frac{\hat{\varrho}(\tilde{X}_\theta + \tilde{Y}_\theta)}{\hat{\varrho}(\tilde{X}_\theta) + \hat{\varrho}(\tilde{Y}_\theta)}. \tag{3.16}$$

For corresponding ordered statistics $\{l_{(i)}\}_{i=1}^N$, $\{\tilde{x}_{(i)}\}_{i=1}^N$, and $\{\tilde{y}_{(i)}\}_{i=1}^N$, the diversification effect can be evaluated by letting $\hat{\varrho}$ equal $\widehat{\mathrm{VaR}}_\alpha$ or $\widehat{\mathrm{ES}}_\alpha$ given by (3.13) and (3.14) respectively.

3.2 Model Simulation

In this section the simulation schemes and discretisation methods used to simulate the various stochastic process will be presented. The simulation methods involved includes the exact scheme where an analytic expression of the solution to the governing dynamics can be derived, as well as a more sophisticated simulation scheme used to simulate the more complex equity models; Heston and Bates. We also present two methods used for copula sampling, which will be used to model the dependence between the two asset classes. A brief discussion about modelling under the objective measure will also be presented.

3.2.1 Risk Premium

As presented in Chapter 2, the equity models, in comparison with the interest rate, is directly modeled under the objective probability measure P. One of the issues with this method is that the drift of the process is not analytically defined without information about the market price of risk, γ. Intuitively, the market price of risk represents the premium an investor should get in compensation for taking on a larger amount of risk. Since investing in equities usually carries more risk than investing in the bank, a risk premia to the drift, under P, is added. Mathematically this is justified through Girsanov's theorem (see Appendix B) giving the relation between the Brownian motion under Q and P as $W^*(t) = \int_0^t \gamma(s)ds + W(t)$. Given the dynamics of a regular geometric Brownian motion, then the relationship between the risk neutral and the objective drift[5] is given by $\mu^{\mathrm{P}} = \mu^{\mathrm{Q}} + \gamma(t)\sigma$. Now, since $\sigma > 0$ by definition, γ represents the premium received for taking a higher risk. For more complex models incorporating stochastic volatility and jumps, γ should, theoretically be greater to reflect the additional risk such additions add to the model. However, as this is not of significant interest to the thesis, the risk premia will be set uniformly across all considered equity models.

3.2.2 The Exact Scheme

The exact scheme utilises the fact that for certain models it is possible to access the distribution of the solution to the SDE governing the dynamics of the model. If so, it is possible to simulate the model exactly without the need of using any discretisation methods such as the Euler-Maruyama and Milstein scheme (e.g., Jäckel (2002)). This ensures accurate simulation algorithms without approximation errors and bias that tends to propagate over time.

Of the models introduced in Chapter 2, both the Black-Scholes and the Merton model are analytically tractable as well as the Hull-White model[6]. Below we present the closed form expressions for the models along with their simulation procedures.

Black-Scholes and Merton

Starting with the Black-Scholes model given in (2.25). By means of stochastic calculus, the explicit solution to the dynamics in (2.25) is given by

$$S(t) = S(0)\exp\left\{\left(\mu - \frac{1}{2}\sigma^2\right)t + \sigma W(t)\right\}, \tag{3.17}$$

[5] In the case of jumps this is of course not the case as the equivalent martingale measure Q is no longer unique. This is however not within the scope of this thesis and the reader is referred to, e.g., Broadie et al. (2007) for a review of risk premia under the addition of jumps.

[6] Technically there also exist a distributional solution to the CIR SDE, however as the computational power of accessing such a distribution has been shown to be rather cumbersome from a simulation point of view, we will not consider such a simulation scheme.

for some initial value $S(0)$.

From a simulation point of view, it is often convenient to work with the logarithm of the solution, and so rewriting the solution in terms of its exponent yields,

$$X(t) = X(0) + \left(\mu - \frac{1}{2}\sigma^2\right) t + \sigma W(t), \tag{3.18}$$

where $X \coloneqq \log S$. For a fixed simulation time step Δ and arbitrary point in time, the simulation recursion becomes

$$X(t+\Delta) = X(t) + \left(\mu - \frac{1}{2}\sigma^2\right) \Delta + \sigma\sqrt{\Delta} z,$$

where z denotes a standard Gaussian variable. The value of S is now easily attainable by exponentiating the simulated paths of X.

The addition of jumps to the Black-Scholes model in the Merton model generalises the solution to the geometric Brownian motion with the addition of a product sum. The solution of the Merton model then becomes

$$S(t) = S_0 \exp\left\{\left(\mu - \frac{1}{2}\sigma^2\right) t + \sigma W(t)\right\} \prod_{j=1}^{N(t)} Y_j, \tag{3.19}$$

and in terms of its exponent, X,

$$X(t) = X(0) + \left(\mu - \frac{1}{2}\sigma^2\right) t + \sigma W(t) + \sum_{j=1}^{N(t)} \log Y_j. \tag{3.20}$$

The time discretisation, for an arbitrary point in time, now follows as

$$X(t+\Delta) = X(t) + \left(\mu - \frac{1}{2}\sigma^2\right) \Delta + \sigma\sqrt{\Delta} z + \sum_{j=N(t)+1}^{N(t+\Delta)} \log Y_j.$$

To discretise the sum we utilise the fact that Poisson increments $N(t+\Delta) - N(t)$ are independent and Poisson distributed with mean $\lambda(t+\Delta-t) = \lambda\Delta$.

Assuming that $N(t+\Delta) - N(t) = \xi$, the exact simulation recursion of the Merton model reads

$$X(t+\Delta) = X(t) + \left(\mu - \frac{1}{2}\sigma^2\right) \Delta + \sigma\sqrt{\Delta} z_1 + \xi\mu_J + \sigma_J\sqrt{\xi} z_2. \tag{3.21}$$

Here, z_1 and z_2 are both standard Gaussian and independent of each other. The entire simulation algorithm of the Black-Scholes and Merton model is outlined in Algorithm 1.

Algorithm 1: Exact Scheme for the Black-Scholes and Merton jump-diffusion model

Data: $\Delta, \mu, \sigma, \lambda, \mu_J, \sigma_J, x_0, mod$
Result: $\{X(t_i)\}_{i=1}^N$

initialization $X(t_0) = x_0$;
for $i \leftarrow 1$ **to** N **do**
  if *mod=Merton* **then**
    draw $\xi \sim \mathsf{Po}(\lambda\Delta)$;
    if $\xi = 0$ **then**
      $J \leftarrow 0$;
    else
      draw $z_1 \sim \mathsf{N}(0,1)$;
      $J \leftarrow \mu_J \xi + \sigma_J \sqrt{\xi} z_1$;
    end
  else
    $J \leftarrow 0$;
  end
  draw $z_2 \sim \mathsf{N}(0,1)$;
  $X(t_i) \leftarrow X(t_{i-1}) + \left(\mu - \frac{1}{2}\sigma^2\right)\Delta t_i + \sigma\sqrt{\Delta t_i} z_2 + J$;
end

Hull-White

To simulate the Hull-White model, recall the analytical decomposition of the model given by

$$r(t) = \tilde{r}(t) + \psi(t). \tag{3.22}$$

As shown in Chapter 2, ψ is a completely deterministic function given by the initial forward curve observed in the market and $\tilde{r}$ is a regular zero mean Vasicek model. As mentioned before, the advantage of this decomposition is the ability of simulating the model by separately simulating ψ and $\tilde{r}$ independent of each other. Due to the deterministic nature of ψ, the entire simulation procedure of the Hull-White model follows from the simulation of the regular Vasicek model, $\tilde{r}$.

Using the definition of $\tilde{r}$ presented in (2.21) it is easy to see that, for any $s < t$, it holds that

$$\tilde{r}(t) = \tilde{r}(s)e^{-\lambda(t-s)} + \lambda \int_s^t e^{-\lambda(t-\tau)} dW^*(\tau),$$

(Brigo and Mercurio, 2012).

This means that $\tilde{r}$ is conditionally Gaussian with mean and variance given by

$$m(t) \coloneqq \mathrm{E}(\tilde{r}(t)|\tilde{r}(s)) = \tilde{r}(s)e^{-\lambda(t-s)},$$
$$s^2(t) \coloneqq \mathrm{Var}(\tilde{r}(t)|\tilde{r}(s)) = \frac{\eta^2}{2\lambda}\left(1 - e^{-2\lambda(t-s)}\right),$$

and so from the Hull-White decomposition the simulation procedure of r can be written as

$$r(t+\Delta) = \psi(t+\Delta) + m(t) + s(t)z, \tag{3.23}$$

with z being a standard Gaussian random variables. The exact simulation scheme of the Hull-White model is outlined in Algorithm 2 below.

Algorithm 2: Exact Scheme for the Hull-White short-rate model.

Data: $\Delta, \{f(0,t_i)\}_{i=1}^N, \lambda, \eta, r_0$
Result: $\{r(t_i)\}_{i=1}^N$
initialization $r(t_0) = \psi(t_0) = r_0$;
for $i \leftarrow 1$ **to** N **do**
  $m \leftarrow \tilde{r}(t_{i-1})\exp(-\lambda\Delta)$;
  $s^2 \leftarrow \frac{\eta^2}{2\lambda}(1 - \exp(-2\lambda\Delta))$;
  $\psi(t_i) \leftarrow f(0,t_{i-1}) + \frac{\eta^2}{2\lambda^2}(1 - \exp(-\lambda t_i))^2$;
  draw $z \sim \mathsf{N}(0,1)$;
  $\tilde{r}(t_i) \leftarrow m + sz$;
  $r(t_i) \leftarrow \psi(t_i) + \tilde{r}(t_i)$;
end

For the bond price, the simulation procedure follows naturally through Algorithm 2 and the analytical expression of the bond price in (2.6). From this setup it is possible to construct a so-called *constant maturity rebalancing scheme*. The idea is that, rather than investing in a bond and holding it until maturity, an investor should always hold a bond with a constant maturity T'. This is ensured by rebalancing. By assuming a finite number of rebalancing times over the investment horizon $[0,T]$ then, at time of rebalancing, t_r, the bond held will be sold whereupon a new, with the same time to maturity, will be bought.

Let $\mathcal{W}(t_r)$ denote the wealth invested in the bond at time of rebalancing and $w(t_r)$ the corresponding number of bonds held at that time, then

$$\begin{aligned} \mathcal{W}(t_r + \Delta) &= w(t_r)P(t_r + \Delta, t_r + T') \\ w(t_r) &= \frac{\mathcal{W}(t_r)}{P(t_r, t_r + T')} \end{aligned} \tag{3.24}$$

with $w(0) = \mathcal{W}(0)/P(0,T')$, i.e. the initial wealth divided by the price of a zero-coupon bond at time $t = 0$ maturing at T'.

3.2.3 The Quadratic Exponential Scheme

As pointed out in Chapter 2, it is possible to show that the conditional distribution of the CIR process in (2.23) includes a non-central chi-square distribution with degrees of freedom equal to $n = 4\kappa\theta/\nu^2$ (e.g., Cox et al. (1985), Glasserman (2004, ch. 3)). This was adapted in Broadie and Kaya (2006), in their development of an exact simulate scheme of the CIR process.

As much as such a simulation scheme is theoretically appealing, studies have shown that simulation through such a procedure becomes practically limited due to its complex nature involving Fourier inversion methods and lack of computational speed (Andersen, 2007).

Instead, Andersen (2007) proposed to approximate the sampling from the CIR distribution using a moment matching technique of the first two (conditional) moments. For sufficiently large values of $V(t)$, Andersen showed that the distribution of $V(t+\Delta)|V(t)$ can be well approximated with a quadratic representation;

$$V(t+\Delta) = (a+bz)^2, \tag{3.25}$$

where z denotes a standard Gaussian random variable and a and b some constants used to match the first two moments of V. Meanwhile, for small values of V, the quadratic approximation falls apart. To address this problem, Andersen proposed a switching rule where, for small value of $V(t)$, the distribution of $V(t+\Delta)|V(t)$ will be approximated by

$$V(T+\Delta) = \begin{cases} 0, & U \leq p \\ \beta^{-1}\log\frac{1-p}{1-U}, & U > p, \end{cases} \tag{3.26}$$

with, U being standard uniform and p and β some appropriately set constants.

The quadratic exponential (QE) scheme now follows by considering the first two moments of $V(t+\Delta)$ given $V(t)$. From the dynamics of the CIR process given in (2.23), the first two conditional moments is given by

$$\begin{aligned} m &\coloneqq \mathrm{E}(V(t+\Delta)|V(t)) = \theta + (V(t)-\theta)e^{-\kappa\Delta}, \\ s^2 &\coloneqq \mathrm{Var}(V(t+\Delta)|V(t)) = \frac{V(t)\nu^2 e^{-\kappa\Delta}}{\kappa}\left(1-e^{-\kappa\Delta}\right) + \frac{\theta\nu^2}{2\kappa}\left(1-e^{-\kappa\Delta}\right)^2. \end{aligned} \tag{3.27}$$

Letting $\psi \coloneqq \frac{s^2}{m^2}$, (3.25) is applicable if $\psi \leq 2$ while (3.26) is used for[7] $\psi \geq 1$.

The involved simulation parameters a, b, p and β is related to m and s through

$$a \coloneqq \frac{m}{1+b^2}, \qquad b^2 \coloneqq 2\psi^{-1} - 1 + \sqrt{2\psi^{-1}(2\psi^{-1}-1)} \tag{3.28}$$

[7] The observant reader realises that the two conditions is indeed overlapping meaning that at least one of the two schemes is always applicable in the sampling procedure.

in (3.25) as well as

$$p \coloneqq \frac{\psi - 1}{\psi + 1}, \qquad \beta \coloneqq \frac{2}{m(\psi + 1)} \tag{3.29}$$

for (3.26).

The QE scheme can now be applied to simulate paths of the Heston and Bates model. Consider the case of the Heston model in (2.26), by rewriting the stock dynamics of (2.26) in terms of its logarithm the dynamics reads

$$\begin{aligned} d\log S(t) &= \left(\mu - \frac{1}{2}V(t)\right)dt + \sqrt{V(t)}dW_S(t), \\ dV(t) &= \kappa(\theta - V(t))dt + \nu\sqrt{V(t)}dW_V(t). \end{aligned} \tag{3.30}$$

Using the fact that $\langle W_S(t)W_V(t)\rangle = \rho t$ it is possible to rewrite the driving process of S in terms of W_V as

$$W_S(t) = \rho W_V(t) + \sqrt{1-\rho^2}W^{\perp}(t), \tag{3.31}$$

where $W^{\perp}$ is a standard Brownian motion independent of W_V. (3.30) now becomes

$$\begin{aligned} d\log S(t) &= \left(\mu - \frac{1}{2}V(t)\right)dt + \rho\sqrt{V(t)}dW_V(t) + \sqrt{1-\rho^2}\sqrt{V(t)}dW^{\perp}(t), \\ dV(t) &= \kappa(\theta - V(t))dt + \nu\sqrt{V(t)}dW_V(t). \end{aligned} \tag{3.32}$$

In integral form,

$$\begin{aligned} \log S(t+\Delta) \;=\; & \log S(t) + \mu\Delta - \frac{1}{2}\int_t^{t+\Delta} V(s)ds \\ & + \rho\int_t^{t+\Delta}\sqrt{V(s)}dW_V(s) + \sqrt{1-\rho^2}\int_t^{t+\Delta}\sqrt{V(s)}dW^{\perp}(s) \end{aligned} \tag{3.33}$$

and the CIR process

$$\begin{aligned} & V(t+\Delta t) = V(t) + \int_t^{t+\Delta}\kappa(\theta - V(s))ds + \nu\int_t^{t+\Delta}\sqrt{V(s)}dW_V(s) \\ \Leftrightarrow \quad & \int_t^{t+\Delta}\sqrt{V(s)}dW_V(s) = \frac{1}{\nu}\left(V(t+\Delta) - V(t) - \kappa\theta\Delta + \kappa\int_t^{t+\Delta}V(s)ds\right). \end{aligned} \tag{3.34}$$

By comparison of (3.33) and (3.34) we see that the second to last integral in (3.33) equals that in (3.34). Using substitution and term rearrangement, (3.33) becomes

$$\begin{aligned} \log S(t+\Delta) \;=\; & \log S(t) + \left(\mu - \frac{\rho\kappa\theta}{\nu}\right)\Delta + \frac{\rho}{\nu}(V(t+\Delta) - V(t)) \\ & + \left(\frac{\rho\kappa}{\nu} - \frac{1}{2}\right)\int_t^{t+\Delta}V(s)ds + \sqrt{1-\rho^2}\int_t^{t+\Delta}\sqrt{V(s)}dW^{\perp}(s), \end{aligned} \tag{3.35}$$

where $V(t+\Delta)$ is simulated through the switching rule of (3.25) and (3.26).

In order to sample from $\log S(t+\Delta t)$ the two integrals in (3.35) needs to be handled. For the first integral Andersen proposed a simple linear discretisation as

$$\int_t^{t+\Delta} V(s)ds = (\gamma_1 V(t+\Delta) + \gamma_2 V(t))\,\Delta \tag{3.36}$$

for some constants γ_1 and γ_2. Here, we will use a regular trapezoidal approximation, letting $\gamma_1 = \gamma_2 = \frac{1}{2}$.

As for the last integral in (3.35), we see, given that $W^{\perp}$ is independent from V, as proposed in (3.31), conditional on $V(t)$, the integral is Gaussian with zero mean and variance given by

$$\int_t^{t+\Delta t} V(s)ds.$$

This integral can, preferably, be approximated using the trapezoidal method once again. The whole discretisation of the Heston model now reads

$$\begin{aligned}\log S(t+\Delta) &= \log S(t) + \left(\mu - \frac{\rho\kappa\theta}{\nu}\right)\Delta + \frac{\rho}{\nu}(V(t+\Delta) - V(t)) \\ &\quad + \left(\frac{\rho\kappa}{\nu} - \frac{1}{2}\right)(\gamma_1 V(t+\Delta) - \gamma_2 V(t))\Delta + \sqrt{1-\rho^2}\sqrt{(\gamma_1 V(t+\Delta) - \gamma_2 V(t))\Delta}z. \\ &= \log S(t) + K_0 + K_1 V(t) + K_2 V(t+\Delta) + \sqrt{K_3 V(t) + K_4 V(t+\Delta)}z,\end{aligned} \tag{3.37}$$

where z denotes a standard Gaussian variable and K_0, K_1, K_2, K_3, K_4 is given below.

$$\begin{aligned} K_0 &\coloneqq \left(\mu - \frac{\rho\kappa\theta}{\nu}\right)\Delta, & K_1 &\coloneqq \gamma_1\left(\frac{\rho\kappa}{\nu} - \frac{1}{2}\right)\Delta - \frac{\rho}{\nu}, \\ K_2 &\coloneqq \gamma_2\left(\frac{\rho\kappa}{\nu} - \frac{1}{2}\right)\Delta + \frac{\rho}{\nu}, & K_3 &\coloneqq \gamma_1(1-\rho^2)\Delta, \\ K_4 &\coloneqq \gamma_2(1-\rho^2)\Delta, & \gamma_1 &= \gamma_2 = \tfrac{1}{2}. \end{aligned}$$

In the case of the Bates model, starting from the Heston discretisation in (3.37), it is possible to use the independent property of the compounded Poisson process to the two Brownian motion in order to extend the simulation. The simulation recursion is then given by

$$\log S(t+\Delta) = \log S(t) + K_0 + K_1 V(t) + K_2 V(t+\Delta) + \sqrt{K_3 V(t) + K_4 V(t+\Delta)}z_1 + J, \tag{3.38}$$

where J, given that $N(t+\Delta) - N(t) = \xi$, is given by

$$J = \xi\mu_J + \sigma_J\sqrt{\xi}z_2.$$

The whole QE algorithm for the Heston and Bates model is outlined in Algorithm 3.

Algorithm 3: QE Scheme for the Heston and Bates model.

Data: $\Delta, \mu, \sigma, \kappa, \theta, \nu, \mu_J, \sigma_J, \lambda, \rho, x_0, v_0, \gamma_1, \gamma_2, mod$
Result: $\{X(t_i)\}_{i=1}^N, \{V(t_i)\}_{i=1}^N$

initialization $X(t_0) = x_0, V(t_0) = v_0$;
$K_0 \leftarrow \left(\mu - \frac{\rho\kappa\theta}{\nu}\right)\Delta$;
$K_1 \leftarrow \gamma_1\left(\frac{\rho\kappa}{\nu} - \frac{1}{2}\right)\Delta - \frac{\rho}{\nu}$;
$K_2 \leftarrow \gamma_2\left(\frac{\rho\kappa}{\nu} - \frac{1}{2}\right)\Delta + \frac{\rho}{\nu}$;
$K_3 \leftarrow \gamma_1(1-\rho^2)\Delta$;
$K_4 \leftarrow \gamma_2(1-\rho^2)\Delta$;
for $i \leftarrow 1$ **to** N **do**
  $m \leftarrow \theta + (V(t_{i-1}) - \theta)\exp(-\kappa\Delta)$;
  $s^2 \leftarrow \frac{V(t_{i-1})\nu^2 e^{-\kappa\Delta}}{\kappa}\left(1 - e^{-\kappa\Delta}\right) + \frac{\theta\nu^2}{2\kappa}\left(1 - e^{-\kappa\Delta}\right)^2$;
  $\psi \leftarrow \frac{s^2}{m^2}$;
  draw $u \sim \mathsf{U}(0,1)$;
  if $\psi \le \psi_C$ **then**
    $b^2 \leftarrow 2\psi^{-1} - 1 + \sqrt{2\psi^{-1}(2\psi^{-1} - 1)}$;
    $a \leftarrow \frac{m}{1+b^2}$;
    draw $z_1 \sim \mathsf{N}(0,1)$;
    $V(t_i) \leftarrow a(b + z_1)^2$;
  else
    $p \leftarrow \frac{\psi - 1}{\psi + 1}$;
    $\beta \leftarrow \frac{2}{m(\psi+1)}$;
    $V(t_i) \leftarrow \begin{cases} 0, & 0 \le u \le p, \\ \beta^{-1}\log\frac{1-p}{1-u}, & p < u \le 1; \end{cases}$
  end
  if *mod=Bates* **then**
    draw $\xi \sim \mathsf{Po}(\lambda\Delta)$;
    if $\xi = 0$ **then**
      $J \leftarrow 0$;
    else
      draw $z_2 \sim \mathsf{N}(0,1)$;
      $J \leftarrow \mu_J\xi + \sigma_J\sqrt{\xi}z_2$;
    end
  else
    $J \leftarrow 0$;
  end
  draw $z_3 \sim \mathsf{N}(0,1)$;
  $X(t_i) \leftarrow X(t_i) + K_0 + K_1 V(t_{i-1}) + K_2 V(t_i) + \sqrt{K_3 V(t_{i-1}) + K_4 V(t_i)}z_3 + J$;
end

3.2.4 Copula Sampling

In order to simulate the dependence between the interest rate and equity process, the Wiener processes needs to be sampled through the copula. In doing so, two methods will be considered. The first one allows to sample from copulas with known distribution such as Gaussian and t copula, whilst the second is used to sample from Archimedean copulas.

Distributional Sampling Approach

The distributional sampling approach (DSA) utilises the one-to-one relationship between copulas and multivariate continuous distribution function. Recall the definition of the d-dimensional copula

$$C(\mathbf{u}) = F\left(F_1^{-1}(u_1), \dots, F_d^{-1}(u_d)\right),$$

with $\mathbf{u} \in [0,1]^d$. Then, if the multivariate distribution function F and the inverse of the marginals $F_i^{-1},\ i = 1, \dots, d$ are known, the DSA algorithm effectively samples from the copula C by drawing an independent d-dimensional sample $\mathbf{X} \sim F$ and inserting each margin X_i into its marginal F_i. Following the probability transformation, the result $(F_1(X_1), \dots, F_d(X_d))^\top \in [0,1]^d$ and will thus be a variate from C. The steps of the DSA algorithm can now be summarized into two steps;

- Sample $\mathbf{X} = (X_1, \dots, X_d)^\top \in \mathbb{R}^d$ with $\mathbf{X} \sim F$,
- Return $(F_1(X_1), \dots, F_d(X_d))^\top$.

As stated above, the DSA can thus only be used to sample from the Gaussian and the t copula with F equal to Φ_R and $t_{\nu,R}$ respectively and with corresponding marginals. Regretfully the DSA can not be used to sample from Archimedean copulas, and in particular the Clayton copula, as neither the joint distribution function nor the marginals are know. To do so, another sampling algorithm will be used considering the inverse of the Laplace-Stiltjes transform presented in Section 2.3.2.

Laplace Transform Simulation

As presented in Section 2.3.2, if ψ is a completely monotone function then the Laplace-Stiltjes transform provides a one-to-one relationship between ψ and some distribution function F as $\psi(t) = \mathcal{L}_t\{F\}$. If F is known and accessible then the inverse of the Laplace-Stiltjes transform provides a sampling algorithm for copulas of the form in (2.36);

- Sample $X \sim F$ with $F = \mathcal{L}_t^{-1}\{\psi\}$,
- Sample independently $U_i \sim \mathsf{U}(0,1)$ and set $Y_i = -\log U_i,\ i = 1, \dots, d$,
- Return $(\psi(Y_1/X), \dots, \psi(Y_d/X))^\top$.

In the case of the Clayton copula, X has a Gamma distribution[8], that is, $F = \mathsf{Ga}(\alpha, \beta)$ with shape parameter, $\alpha = 1/\theta$ and scale parameter, $\beta = 1$. The sampling algorithm for the Clayton copula now follows as;

- Sample $X \sim \mathsf{Ga}(1/\theta, 1)$,
- Sample independently $U_i \sim \mathsf{U}(0,1)$ and set $Y_i = -\log U_i,\ i = 1, \ldots, d$,
- Return $\left(\left(1 + \theta \frac{Y_1}{X}\right)^{-1/\theta}, \ldots, \left(1 + \theta \frac{Y_d}{X}\right)^{-1/\theta} \right)^\top$.

3.3 Empirical Study

To study the accuracy of the statistics using the estimation method presented in Section 3.1.2, the simulation and sampling methods presented in Section 3.2 will be used. By means of Monte Carlo, the copula sampling methods allows to generate trajectories of the joint interest rate-equity framework. Now, consider the case where N trajectories of length T of the joint framework has been generated. Then, using rolling analysis with an overlapping window of length ΔT, sample statistics can be estimate from each trajectory. In the case of the sample mean, this procedure equip us with N estimations of the expected ΔT-mean, $\hat{\mu}_{\tilde{r}}^{(i)} = n^{-1}(\tilde{r}_1^{(i)} + \ldots + \tilde{r}_n^{(i)}),\ n = T - \Delta T + 1,\ i = 1, \ldots, N$. For large N this method gives an empirical distribution of the estimation of the expected return over the investment horizon ΔT. The uncertainty of the estimation follows by constructing confidence intervals of the estimate where the upper and lower bounds are given empirically as the $N\alpha/2$ smallest and $N(1-\alpha)/2$ largest values of the estimates $\{\hat{\mu}_{\tilde{r}}\}_{i=1}^N$, for some confidence level α. The confidence interval of μ is thus given by

$$I_\alpha\left(\mathrm{E}\left(\hat{\mu}_{\tilde{r}}\right)\right) \coloneqq \left(\hat{\mu}_{\tilde{r}}^{(N\alpha/2)}, \hat{\mu}_{\tilde{r}}^{(N(1-\alpha)/2)}\right),$$

and analogously for all other statistical quantities.

By gradually increasing the size of ΔT, this procedure also allows to investigate the accuracy and predictive power implied by a series of historical data of length T under the hypothesis that the model generating the trajectories is the one that generated the historical data. Once again, in the case of the sample mean $\hat{\mu}_{\tilde{r}}$, the width of the empirical confidence level $\Delta I_\alpha(\cdot)$ effectively quantify the predictive power of this procedure. By varying the complexity of the models grants us to study the predictive power of the rolling-window procedure when known market behaviors like stochastic volatility and jumps are introduced in the model.

[8] Derivation can be found in Appendix C

Chapter 4

Results

In this chapter, the results will be presented. The chapter is structured as follows, Section 4.1 and Section 4.2 evaluates the estimation accuracy of the statistical quantities connected to the interest rate and equity models independently. Section 4.4 consider joint framework and the estimation of the multivariate statistics. Comparison studies of the estimation performance between the models and frameworks is discussed in Section 4.3 and Section 4.4.2 respectively.

The study will be based on monthly frequent data simulated over thirty years using the presented framework. Even though it is always preferable to have access to more data, the financial sector is a constantly evolving market with endogenous and exogenous factors affecting its behavior. This means that market behavior seen in the past might not necessarily be relevant and representative in the process of predicting future market movements. In addition, in the cases where longer series of historical data is representative, there are often accessibility issues. The number of index and funds with regular and frequent historical data tracing back further than thirty years are small and when they exist, gaining access is not cheap. Considering these two points a thirty year series can be considered a reasonable choice to base the study on.

4.1 Interest Rate

In this section, the result of the interest rate model will be presented. The section begins by presenting the model parameters used to simulate the model. Section 4.1.2 presents the estimation accuracy of the summary statistics connected to the returns series using rolling analysis. Lastly, the result of the estimation of the risk measures is investigated. This is presented in Section 4.1.3.

4.1.1 Model Parameters

In order to properly investigate the prediction power of the interest rate model, the model parameters of the Hull-White model should be taken to reflect the behavior seen in the market. Since the objective is to investigate the accuracy implied by the model calibrated to historical data, rather than the calibration procedure itself we here use parameter with order of magnitude equal to that seen in, for instance, Kienitz and Wetterau (2012). The initial condition $r(0)$ is, for simplicity, taken equal to the spot rate of the yield curve in Figure 4.1 which corresponds to a three month STIBOR rate[1].

Model	λ	η
Hull-White	0.25	0.02

Table 4.1: Model parameters of the Hull-White model

As presented in Section 2.1.1, the time dependent long term mean $\theta(t)$ is taken to fit the initial term structure as given in (2.15). The initial yield curve is presented in Figure 4.1 and is composed of Swedish annual forward rates as of 2015-12-31 taken from the European Insurance and Occupational Pension Authority (EIOPA). By fitting $\theta(t)$ to the yield curve thus ensures that the interest rate evolves around the yield curve throughout the simulation horizon.

[1] By definition, $r(0) = f(0,0)$, i.e., the instantaneous forward rate seen in the market. As such a rate does not exist in practice, $r(0)$ is approximated with the three month STIBOR rate.

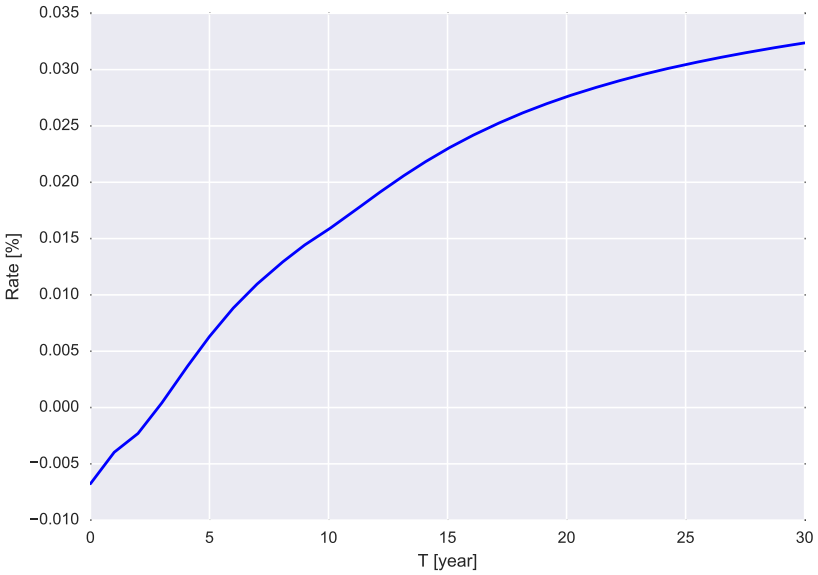


Figure 4.1: Swedish 30 year term structure as of 2015-12-31, Source: EIOPA.

4.1.2 Summary Statistics

In order to investigate the predictability of the interest rate model, the rolling-window procedure presented in Section 3.1.2 was implemented. In doing so, $N = 10\ 000$ sample paths of length $T = 30$ years of the Hull-White constant maturity bond prices were simulated using the simulation algorithm presented in Section 3.2.2 with a time step $\Delta = 1/12$ (monthly frequency) and bond maturity $T' = 10$ years. For each simulated path, the returns were computed after which, for a given window length ΔT, the rolling-return series were calculated. From the $N = 10\ 000$ sample paths, empirical distributions of the sample mean and standard deviation were computed, using, for each path, (3.9) and (3.10). The corresponding empirical confidence interval were computed using a confidence level $\alpha = 0.95$.

The distributions of the sample mean and standard deviation of the rolling-return procedure along with 95% confidence interval is presented in Figure 4.2a and Figure 4.2b respectively using a window length $\Delta T \in \{12, 24, 48, 96\}$ months. The corresponding values of the sample mean and standard deviation and their confidence intervals are given in Table 4.2.

The evaluation of the predictability of the model relies on the fact that the rolling-return series for a given ΔT can be considered stationary and ergodic. This means that the estimated sample statistics are consistent estimates of the invariant quantities implied by the information up to time T. The predictive power of the model can therefore be evaluated

by considering the width of the confidence intervals of the estimations, $\Delta I_\alpha(\cdot)$. The width of the confidence interval as a function of ΔT of the sample mean and standard deviation is depicted in the lower right plot of Figure 4.2a and Figure 4.2b respectively. From the two figures and the values presented in Table 4.2, it is possible to conclude that the Hull-White model seems to perform reasonable well over a 12 to 48 months prediction horizon, where, for the expected return the uncertainty is contained within a 0.5% dispersion from the expected mean return for the 12 months horizon up to 2.5% for the 48 months. For the expected standard deviation, the corresponding uncertainty is ranging between 1.3% for 12 months and 4.5%. For 96 months, the uncertainty has grown significantly with a total width just shy of 13% for the mean and 15% for the standard deviation.

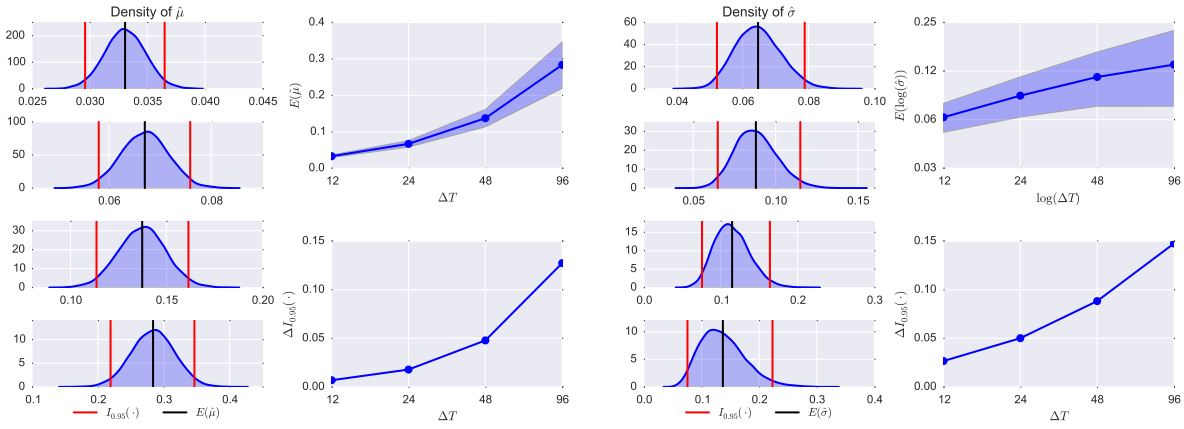
The expected standard deviation plot presented in the top right plot of Figure 4.2b is plotted in loglog-scale. Under the assumption that the innovation of the series is $\text{IID}(0, \sigma_r^2)$, the logarithm of the rolling-window standard deviation should demonstrate a gradient equal to $1/2$ as implied by (3.8),

$$\begin{aligned} \sigma_{\tilde{r}} &= \sigma_r (\Delta T)^{1/2} \\ \Rightarrow \quad \log \sigma_{\tilde{r}} &= \log \sigma_r + \frac{1}{2} \log \Delta T. \end{aligned}$$

From the result in Figure 4.2b we see that the slope is approximately $1/2$ with a slight flattening in the gradient between 48 and 96 months. This is probably a combination of the increased persistency of the autocorrelation and the reduced number of observations generated when considering a $\Delta T = 96$ months. The result is a slower ergodic convergence to the stationary statistics.

	Hull-White			
ΔT	$\mathrm{E}(\hat{\mu})$	$I_{0.95}(\cdot)$	$\mathrm{E}(\hat{\sigma})$	$I_{0.95}(\cdot)$
12	0.033	(0.030, 0.037)	0.065	(0.052, 0.079)
24	0.067	(0.058, 0.076)	0.088	(0.064, 0.115)
48	0.138	(0.113, 0.162)	0.115	(0.075, 0.164)
96	0.285	(0.220, 0.349)	0.137	(0.075, 0.224)

Table 4.2: Table of expected ΔT rolling-mean and standard deviation with corresponding 95% confidence interval for the Hull-White model and $\Delta T \in \{12, 24, 48, 96\}$ months.



a: $\hat{\mu}$ b: $\hat{\sigma}$

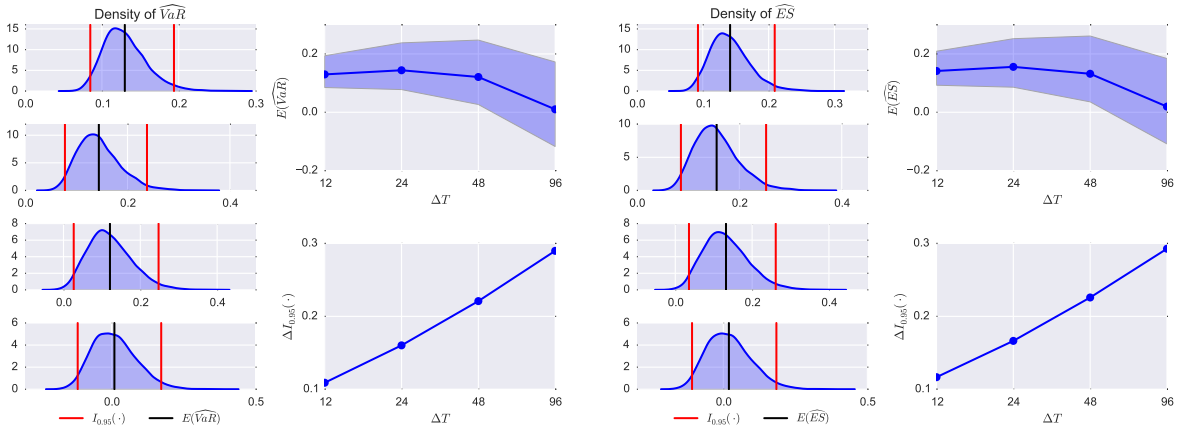
Figure 4.2: Top to Bottom Left: Monte Carlo estimated empirical densities of summary statistics for $\Delta T \in \{12, 24, 48, 96\}$ months. Expected ΔT estimate (black vertical line) with corresponding 95% empirical confidence interval $I_{0.95}(\mathrm{E}(\cdot))$ (red vertical lines). Top Right: Expected summary statistic with corresponding 95% confidence interval as a function of $\Delta T \in \{12, 24, 48, 96\}$ months. Bottom Right: Width of 95% confidence interval as a function of $\Delta T \in \{12, 24, 48, 96\}$ months. All plots generated using $N = 10\ 000$ simulated paths of the Hull-White model.

4.1.3 Risk Measures

For the risk measures the same procedure was implemented as for the sample statistics. The empirical distributions over the $N = 10\ 000$ estimated expected Value-at-Risk and Expected Shortfall of the 99.5th quantile is presented in Figure 4.3a and Figure 4.3b respectively and for the different window lengths. The values of the empirical expected values with corresponding confidence interval is found in Table 4.3. For the risk measures, we consider losses as the negative of returns. This means that a positive value of VaR and ES in Table 4.3 corresponds to a percentage loss of the invested capital. The result show the difficulty to estimate risk measures where, for a 12 months horizon, it is not possible to predict the 99.5%-VaR and ES with a greater certainty than $\pm$ 6%. For longer horizon the uncertainty increase further which is expected given that the autocorrelation between observations generated within the same window prevail longer whilst the overall number of observations decreases. We also see that the magnitude of the risk over longer horizons decreases as depicted in the top right plot of Figure 4.3a and Figure 4.3b. This is a result of the magnitude of the drift of the portfolio in comparison to its volatility. On longer investment horizons bonds are, due to their 'less' volatile nature, expected to generate profit. This behavior is what is reflected in the drop of expected VaR and ES for $\Delta T \in \{48, 96\}$

months.

Despite the poor precision in the estimation of the risk measures, it is possible to conclude that the difference in estimation uncertainty between VaR and ES seems to be insignificant. This, in light of the that that Expected Shortfall is superseding Value-at-Risk as the preferred risk measure used to calculate market risk under Basel III, is an interesting result as it would indicate that the difference in the amount of data needed to obtain a given level of accuracy of them two seem to be small, c.f. Yamai and Yoshiba (2002). This behavior will be further discussed in Section 4.3.2.



a b

Figure 4.3: Top to Bottom Left: Monte Carlo estimated empirical densities of risk measures for $\Delta T \in \{12, 24, 48, 96\}$ months. Expected ΔT estimate (black vertical line) with corresponding 95% empirical confidence interval $I_{0.95}(\mathrm{E}(\cdot))$ (red vertical lines). Top Right: Expected risk measure with corresponding 95% confidence interval as a function of $\Delta T \in \{12, 24, 48, 96\}$ months. Bottom Right: Width of 95% confidence interval as a function of $\Delta T \in \{12, 24, 48, 96\}$ months. All plots generated using $N = 10\ 000$ simulated paths of the Hull-White model.

	Hull-White			
ΔT	$\mathrm{E}(\widehat{\mathrm{VaR}}_\alpha)$	$I_{0.95}(\cdot)$	$\mathrm{E}(\widehat{\mathrm{ES}}_\alpha)$	$I_{0.95}(\cdot)$
12	0.129	(0.085, 0.193)	0.141	(0.093, 0.207)
24	0.144	(0.076, 0.238)	0.155	(0.084, 0.252)
48	0.121	(0.027, 0.251)	0.132	(0.036, 0.263)
96	0.009	(-0.118, 0.174)	0.018	(-0.109, 0.186)

Table 4.3: Table of expected $\Delta T-$ VaR_α and ES_α for $\alpha = 0.995$ with corresponding 95% confidence interval for the Hull-White model and $\Delta T \in \{12, 24, 48, 96\}$ months.

4.2 Equity

With the result connected to the interest rate presented, we move on to the equity processes. The section is structure as in Section 4.1. Firstly a review of the utilised model parameter is presented as well as a short discussion regarding the simulation of the models. With the model parameters presented, Section 4.2.2 moves on to study the performance of the different models in estimating statistics concerning equity returns. A dependence study of the return's sensitivity to changes in certain model parameters is presented in Section 4.2.3. The section is concluded with the presentation of the predictability of the risk measures in Section 4.2.4.

4.2.1 Model Parameters

Just as for the Hull-White model, the choice of model parameters is of significance to evaluate the performance of the rolling analysis estimation procedure. Due to the greater range of complexity within the equity models, the choice of model parameters is key in order to properly evaluate the estimation procedure.

Firstly, we consider the drift of the equity models. As discussed in Section 3.2.1, the drift should reflect the investment in the risk-free rate with an additional risk premia. In doing so, the objective drift μ will be set equal to the 30 year forward rate $f(0, 30)$ given by the initial term structure in Figure 4.1 with and additional annual 1% reflecting the market price of risk.

In the case of models with jumps, the jump frequency is set to $\lambda = 0.2$. This corresponds to an expected number of jumps equal to 6 over a 30 year simulation horizon. The jump sizes is set deterministic with mean $\mu_J = -0.1$ and standard deviation $\sigma_J = 0$, meaning that for each point in time a jump occur, the equity process will experience a fall in its price corresponding to 10% of the price prior to the jump. This is justified considering that over longer time horizon, it is primarily the significant drawdowns that stands out. In accordans with the discussion in Section 3.2.1, stocks drifts upward proportional to the risk-free rate disrupted by unexpected falls during extreme market conditions.

For the stochastic volatility models, the correlation is set equal to $\rho = -0.1$. The negative nature of the correlation between the two Brownian motions is known as the *leverage effect* and represents the fact that in a bearish market where prices decreases, one tends to see increasing levels of volatility due to market uncertainty, Cont (2001).

The remaining of the model parameters is set similar to those seen in Kienitz and Wetterau (2012). All parameter values can be found in Table 4.4.

Model	μ	σ	κ	θ	ν	μ_J	σ_J	λ	ρ
BS	0.043	0.04	-	-	-	-	-	-	-
M	0.043	0.04	-	-	-	-0.1	0	0.2	-
H	0.043	-	0.2	0.04	0.1	-	-	-	-0.1
B	0.043	-	0.2	0.04	0.1	-0.1	0	0.2	-0.1

Table 4.4: Model parameters of the Equity models, Black-Scholes (BS), Merton (M), Heston (H) and Bates (B). The drift is taken as the 30 year spot forward rate $f(0, 30)$ with an added 1% premium, corresponding to the market price of risk γ for simulating under the objective measure.

4.2.2 Summary Statistics

To evaluate the predictive power of the equity models the same rolling-window procedure as in Section 4.1.3 were used. The empirical results of the summary statistics for the Black-Scholes model and Merton model is presented in Figure 4.4 and Figure 4.5 respectively and in Figure 4.6 and Figure 4.7 for Heston and Bates. The values of the expected statistical quantities along with their 95% confidence intervals for each window length ΔT is presented in Table 4.5.

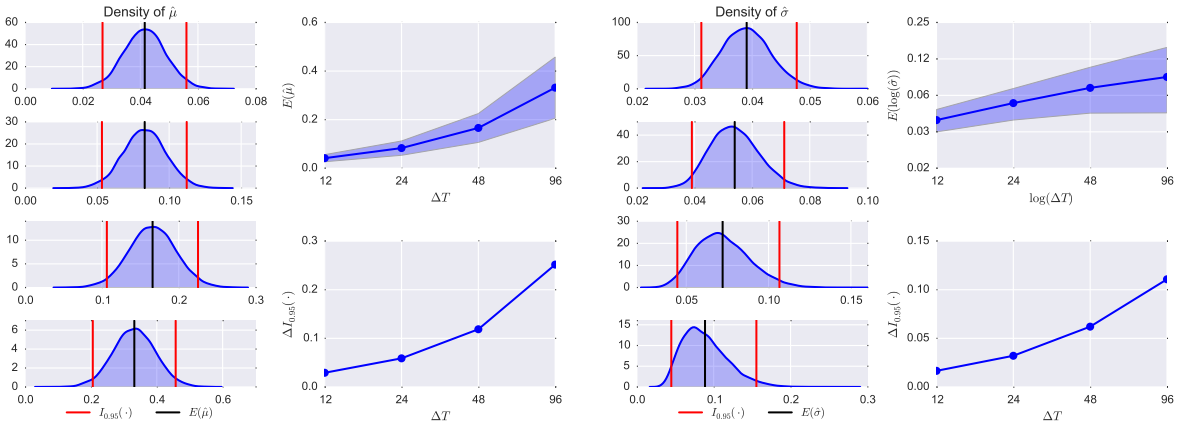
The results in Table 4.5 suggest that that the Black-Scholes model performs best which is expected given its naïve dynamics without the addition of jumps and stochastic volatility. The sole addition of jumps given by the Merton model show an increase in the estimation uncertainty of the statistics, especially for longer horizons $\Delta T \in \{48, 96\}$ months, where the predictive power breaks down and on a 95% statistical level it is not possible to say with certainty whether the expected return will be positive or negative. This behavior is further elevated with the introduction of stochastic volatility (with or without jumps) where even for $\Delta T \in \{12, 24\}$ months it is hard to, statistically, draw any conclusion of the value of future return statistics. Given Bates for instance, under the hypothesis that Bates is the true model, on a one year forecasting horizon, it would not be able to conclude whether an investor would experience an 8% loss or a 7% gain of the invested capital. This, on a 95% confidence level.

The large level of uncertainty exhibited for the equity models on longer forecasting horizon is a result of the rolling-window procedure. As discussed in Section 3.1.2 one of the

problems with the rolling-window method is its sensitivity to extreme outcomes due to the introduction of autocorrelation in the generated sample of observations. This is obvious when comparing the performance between Black-Scholes and Merton. The only difference between these two models is exactly Mertons ability to produce large market movements using jumps. These large outcomes has a strong impact when using the rolling-window procedure as they propagate throughout all observation created withing the width of the window. For longer forecasting horizons, the number of observations affected of course becomes larger as well, resulting in increasing uncertainty. Another contributing factor is the inverse relationship between the number of observation and the length of the forecasting horizon the rolling-window procedure induce. This, together with the introduced autocorrelation has a significant impact in the estimation accuracy and is clearly depicted in Table 4.5.

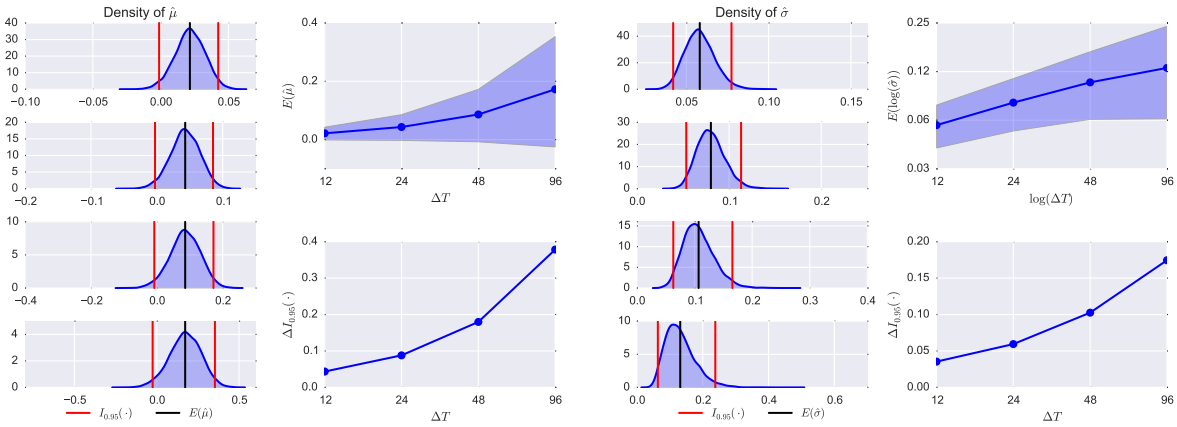
	Black-Scholes			
ΔT	$\mathrm{E}(\hat{\mu})$	$I_{0.95}(\cdot)$	$\mathrm{E}(\hat{\sigma})$	$I_{0.95}(\cdot)$
12	0.042	(0.027, 0.056)	0.039	(0.031, 0.048)
24	0.083	(0.054, 0.113)	0.054	(0.039, 0.072)
48	0.167	(0.106, 0.227)	0.072	(0.044, 0.108)
96	0.333	(0.207, 0.460)	0.088	(0.044, 0.155)
	Merton			
ΔT	$\mathrm{E}(\hat{\mu})$	$I_{0.95}(\cdot)$	$\mathrm{E}(\hat{\sigma})$	$I_{0.95}(\cdot)$
12	0.021	(-0.001, 0.043)	0.058	(0.042, 0.077)
24	0.043	(-0.003, 0.086)	0.081	(0.054, 0.114)
48	0.086	(-0.009, 0.173)	0.108	(0.062, 0.168)
96	0.172	(-0.024, 0.350)	0.133	(0.064, 0.240)
	Heston			
ΔT	$\mathrm{E}(\hat{\mu})$	$I_{0.95}(\cdot)$	$\mathrm{E}(\hat{\sigma})$	$I_{0.95}(\cdot)$
12	0.022	(-0.061, 0.093)	0.192	(0.119, 0.293)
24	0.045	(-0.124, 0.187)	0.266	(0.152, 0.430)
48	0.090	(-0.258, 0.382)	0.357	(0.181, 0.635)
96	0.179	(-0.558, 0.791)	0.440	(0.187, 0.895)
	Bates			
ΔT	$\mathrm{E}(\hat{\mu})$	$I_{0.95}(\cdot)$	$\mathrm{E}(\hat{\sigma})$	$I_{0.95}(\cdot)$
12	0.002	(-0.084, 0.074)	0.197	(0.125, 0.297)
24	0.003	(-0.169, 0.145)	0.272	(0.159, 0.434)
48	0.007	(-0.350, 0.308)	0.365	(0.189, 0.642)
96	0.015	(-0.737, 0.639)	0.449	(0.192, 0.905)

Table 4.5: Table of expected ΔT rolling-mean and standard deviation with corresponding 95% confidence interval for the Black-Scholes, Merton, Heston and Bates model and $\Delta T \in \{12, 24, 48, 96\}$ months.



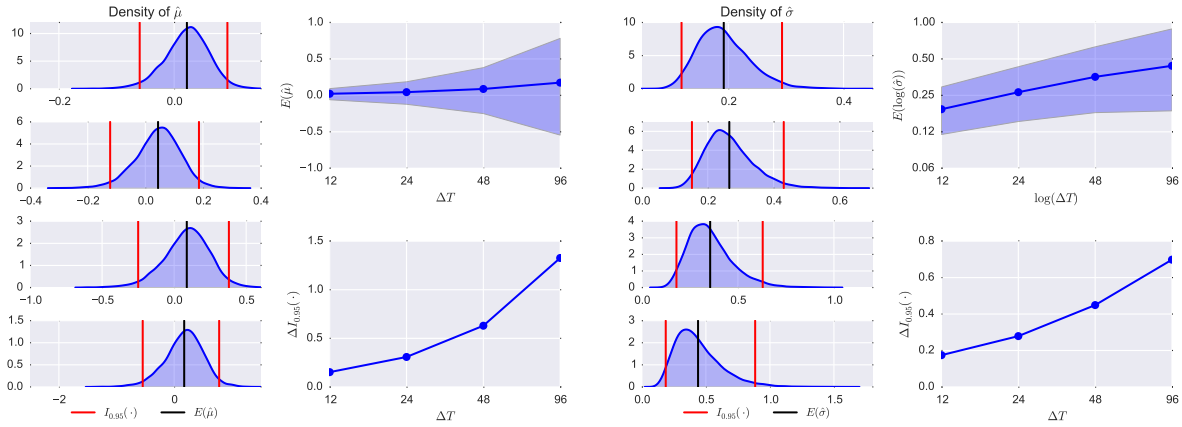
a: $\hat{\mu}$ b: $\hat{\sigma}$

Figure 4.4: Top to Bottom Left: Monte Carlo estimated empirical densities of summary statistics for $\Delta T \in \{12, 24, 48, 96\}$ months. Expected ΔT estimate (black vertical line) with corresponding 95% empirical confidence interval $I_{0.95}(\mathrm{E}(\cdot))$ (red vertical lines). Top Right: Expected summary statistic with corresponding 95% confidence interval as a function of $\Delta T \in \{12, 24, 48, 96\}$ months. Bottom Right: Width of 95% confidence interval as a function of $\Delta T \in \{12, 24, 48, 96\}$ months. All plots generated using $N = 10\ 000$ simulated paths of the Black-Scholes model.



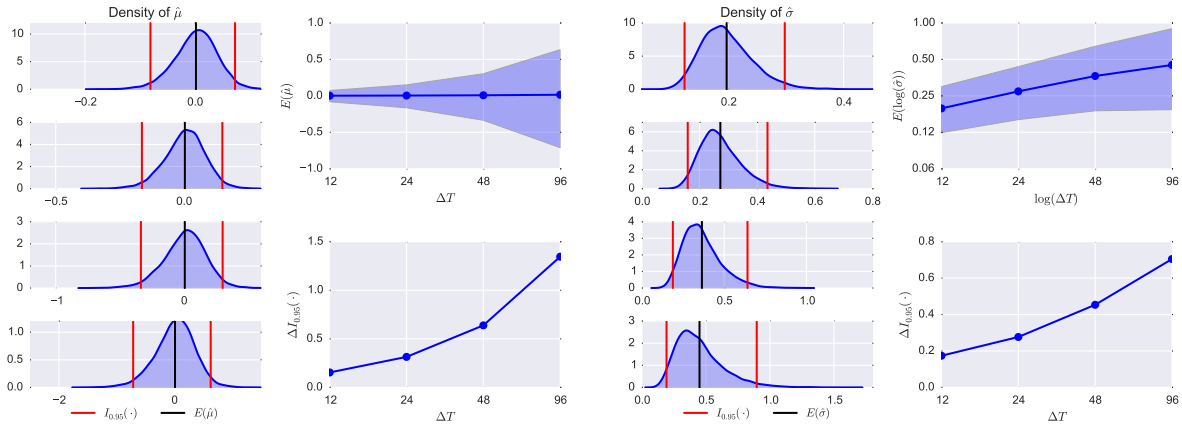
a: $\hat{\mu}$ b: $\hat{\sigma}$

Figure 4.5: Top to Bottom Left: Monte Carlo estimated empirical densities of summary statistics for $\Delta T \in \{12, 24, 48, 96\}$ months. Expected ΔT estimate (black vertical line) with corresponding 95% empirical confidence interval $I_{0.95}(\mathrm{E}(\cdot))$ (red vertical lines). Top Right: Expected summary statistic with corresponding 95% confidence interval as a function of $\Delta T \in \{12, 24, 48, 96\}$ months. Bottom Right: Width of 95% confidence interval as a function of $\Delta T \in \{12, 24, 48, 96\}$ months. All plots generated using $N = 10\ 000$ simulated paths of the Merton model.



a: $\hat{\mu}$ b: $\hat{\sigma}$

Figure 4.6: Top to Bottom Left: Monte Carlo estimated empirical densities of summary statistics for $\Delta T \in \{12, 24, 48, 96\}$ months. Expected ΔT estimate (black vertical line) with corresponding 95% empirical confidence interval $I_{0.95}(\mathrm{E}(\cdot))$ (red vertical lines). Top Right: Expected summary statistic with corresponding 95% confidence interval as a function of $\Delta T \in \{12, 24, 48, 96\}$ months. Bottom Right: Width of 95% confidence interval as a function of $\Delta T \in \{12, 24, 48, 96\}$ months. All plots generated using $N = 10\,000$ simulated paths of the Heston model.



a: $\hat{\mu}$ b: $\hat{\sigma}$

Figure 4.7: Top to Bottom Left: Monte Carlo estimated empirical densities of summary statistics for $\Delta T \in \{12, 24, 48, 96\}$ months. Expected ΔT estimate (black vertical line) with corresponding 95% empirical confidence interval $I_{0.95}(\mathrm{E}(\cdot))$ (red vertical lines). Top Right: Expected summary statistic with corresponding 95% confidence interval as a function of $\Delta T \in \{12, 24, 48, 96\}$ months. Bottom Right: Width of 95% confidence interval as a function of $\Delta T \in \{12, 24, 48, 96\}$ months. All plots generated using $N = 10\,000$ simulated paths of the Bates model.

4.2.3 Return Sensitivity

Here we study the impact of jumps and stochastic volatility to the estimation of the mean and standard deviation. In light of the severe lack of predictive power of the stochastic volatility and jump models for longer horizons presented in Section 4.2.2, the following study will only consider a one year rolling-window procedure with $\Delta T = 12$.

For the jumps we recall the properties of the compounded Poisson process introduced in (2.28). Since the intensity of the jumps is governed by a regular Poisson process, the expected jump frequency can easily be controlled by varying the intensity parameter λ. It is commonly known that the expected value of a regular Poisson process N is given by

$$\mathrm{E}(N(\Delta)) = \lambda\Delta, \tag{4.1}$$

where Δ denotes the simulation step of the model. In our case, all models are simulated using a monthly frequency which means that $\Delta = \frac{1}{12}$, and so by ranging $\lambda \in [0, 12]$ the expected number of jump in each time step will vary between zero and one.

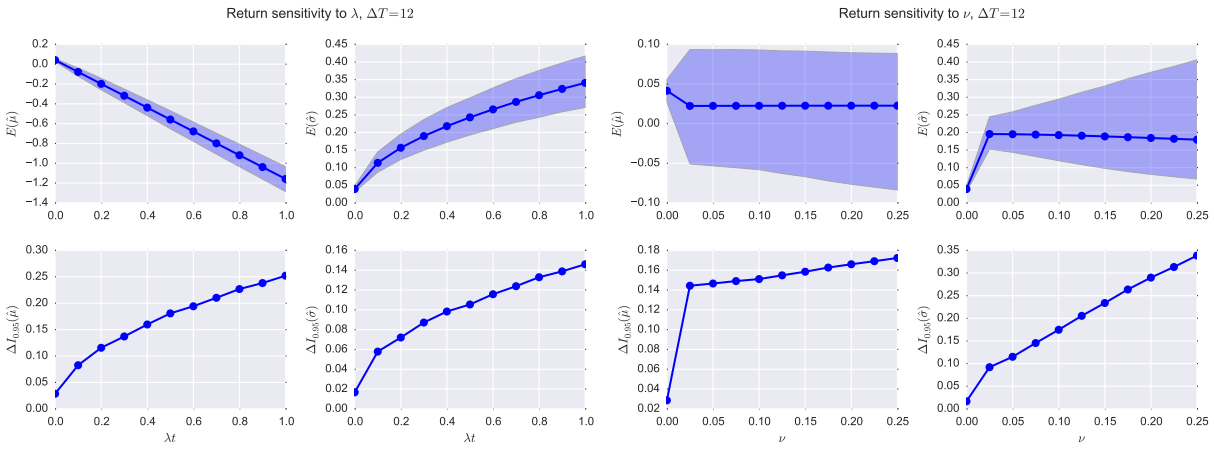
To accurately study the impact of increasing jumps, the study was performed using $N = 10\ 000$ simulated scenarios of the Merton model with the jump frequency $\lambda \in [0, 12]$. Here a $\lambda = 0$ corresponds to no jumps which means that the trajectories is given by the Black-Scholes model.

The result is depicted in Figure 4.8a and shows the expected mean and standard deviation with corresponding confidence interval (top) and the width of the confidence interval (bottom). As can be seen in Figure 4.8a, the width of the confidence interval increases with the jump frequency for both the mean and standard deviation. This is an expected result considering the main reasons for the usage of jumps in equity modelling is the fact that it allows to capture larger variation in price movements then allowed by a diffusion process driven by a Brownian motion solely.

For the stochastic volatility, the impact study is composed of the change in the volatility of variance ν. The volatility is varied for $\nu \in [0, 0.25]$, quoted monthly, using $N = 10\ 000$ simulated scenarios of the Heston model. For $\nu = 0$, κ is also set equal to zero giving a constant volatility corresponding to that in the Black-Scholes model.

The result of the volatility study is presented in Figure 4.8b with the mean and variance plotted as a function of ν and corresponding confidence interval (top), and the width of the confidence interval (bottom). For the mean, we see a significant increase in the in uncertainty between $\nu = 0$ and $\nu > 0$ after which the uncertainty slowly increase as ν grows larger. However, this is not the case for the standard deviation, where the uncertainty seems to depend linearly with respect to ν. This can be explain by the fact that since ν corresponds to the volatility of the variance process, an increase in ν effectively makes the variance more volatile. Since the variance process determines the overall volatility of the return series, an increase in ν also corresponds to a higher uncertainty in the standard deviation of the returns and consequently therefore also the uncertainty in the estimation of the standard deviation.

To conclude, we see that both the introduction of jumps and stochastic volatility has a substantial impact on the uncertainty of the parameter estimations. However, given the nature of the two extensions, their effects are different for the estimation of the statistics. Given that for both studies the benchmark is given by the Black-Scholes model meaning that for $\lambda = 0$ and $\nu = 0$, $I_{0.95}(\mathrm{E}(\hat{\mu}))$ and $I_{0.95}(\mathrm{E}(\hat{\sigma}))$ are equal to each other, it is possible to conclude that the addition of jumps has a considerably greater impact on the estimation uncertainty of the expected one year return in comparison to the stochastic volatility. This, in comparison to the standard deviation where the introduction of stochastic volatility carries a far greater uncertainty on the estimation of the one year expected standard deviation than seen in jumps.



a b

Figure 4.8a Top Left: Sensitivity of expected one year mean as a function of λ with corresponding 95% confidence interval. Top Right: Sensitivity of expected one year standard deviation as a function of λ with corresponding 95% confidence interval. Bottom Left: Width of confidence interval of the expected one year mean as a function of λ. Bottom Right: Width of confidence interval of the expected one year standard deviation as a function of λ. $\lambda \in [0, 12]$, $\Delta T = 12$ months.

Figure 4.8b Top Left: Sensitivity of expected one year mean as a function of ν with corresponding 95% confidence interval. Top Right: Sensitivity of expected one year standard deviation as a function of ν with corresponding 95% confidence interval. Bottom Left: Width of confidence interval of the expected one year mean as a function of ν. Bottom Right: Width of confidence interval of the expected one year standard deviation as a function of ν. $\nu \in [0, 0.25]$ quoted monthly, $\Delta T = 12$ months.

4.2.4 Risk Measures

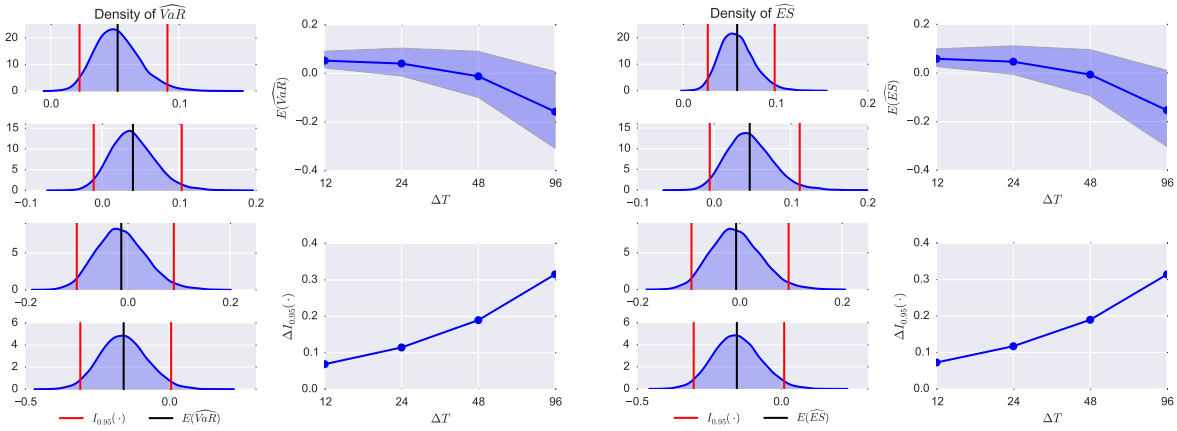
For the risk measures connected to the equity models, the same rolling-window procedure as the one presented in Section 4.2.2 were used. The overall observations regarding the uncertainty in the models seen in Section 4.2.2 is also visible in the estimation of the risk measures where the uncertainty is to extensive to, on a 95% level, draw any conclusion about future risk exposures of equities. The result of the estimation of the risk measures is presented in Figure 4.9 (Black-Scholes), Figure 4.10 (Merton), Figure 4.11 (Heston), Figure 4.12 (Bates) and Table 4.6.

For the Heston and Bates model, we also notice that the confidence levels includes VaR and ES values greater than 1. This is a drawback of using log-returns to approximate simple returns. As discussed in Section 3.1.1, the approximation of simple returns with log-returns in (3.3) is only justifiable for small return values. Since the application of both VaR and ES is to measure the magnitude of extreme losses (99.5th quantile), by default the returns involved in such estimation tends to be large. This is why the same behavior is not observed for the return statistics presented in Table 4.5. Indeed, the choice of ΔT is also a factor in this problem since the dispersion of the model is directly proportional to time and which would increase the probability in seeing large fluctuations in the price process, at least larger than for shorter horizons.

Nevertheless, the substantial increase in the width between the results connected to Black-Scholes or Merton in comparison with those of Heston and Bates demonstrate the uncertainty stochastic volatility readily introduce in the estimation of Value-at-Risk and Expected Shortfall, especially for $\Delta T = 12$ months.

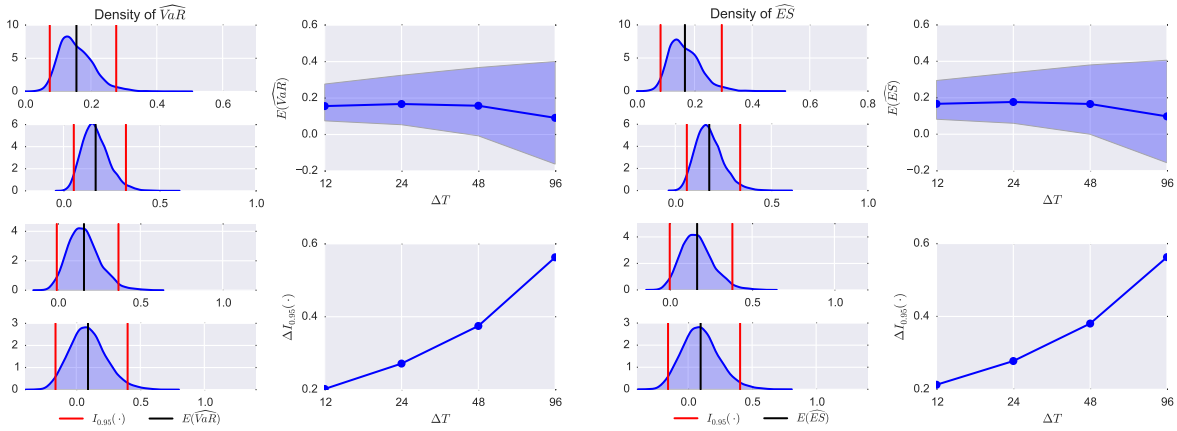
	Black-Scholes			
ΔT	$\mathrm{E}(\widehat{\mathrm{VaR}}_\alpha)$	$I_{0.95}(\cdot)$	$\mathrm{E}(\widehat{\mathrm{ES}}_\alpha)$	$I_{0.95}(\cdot)$
12	0.053	(0.022, 0.092)	0.059	(0.028, 0.101)
24	0.040	(-0.01, 0.103)	0.047	(-0.005, 0.111)
48	-0.012	(-0.097, 0.090)	-0.006	(-0.092, 0.096)
96	-0.157	(-0.306, 0.009)	-0.152	(-0.301, 0.014)
	Merton			
ΔT	$\mathrm{E}(\widehat{\mathrm{VaR}}_\alpha)$	$I_{0.95}(\cdot)$	$\mathrm{E}(\widehat{\mathrm{ES}}_\alpha)$	$I_{0.95}(\cdot)$
12	0.156	(0.076, 0.278)	0.167	(0.083, 0.293)
24	0.168	(0.055, 0.325)	0.177	(0.062, 0.337)
48	0.158	(-0.004, 0.370)	0.166	(0.002, 0.378)
96	0.089	(-0.158, 0.395)	0.095	(-0.151, 0.400)
	Heston			
ΔT	$\mathrm{E}(\widehat{\mathrm{VaR}}_\alpha)$	$I_{0.95}(\cdot)$	$\mathrm{E}(\widehat{\mathrm{ES}}_\alpha)$	$I_{0.95}(\cdot)$
12	0.496	(0.235, 0.899)	0.538	(0.260, 0.964)
24	0.621	(0.233, 1.232)	0.659	(0.256, 1.293)
48	0.713	(0.156, 1.599)	0.747	(0.176, 1.650)
96	0.700	(-0.111, 1.951)	0.726	(-0.093, 1.989)
	Bates			
ΔT	$\mathrm{E}(\widehat{\mathrm{VaR}}_\alpha)$	$I_{0.95}(\cdot)$	$\mathrm{E}(\widehat{\mathrm{ES}}_\alpha)$	$I_{0.95}(\cdot)$
12	0.527	(0.269, 0.925)	0.568	(0.294, 0.998)
24	0.676	(0.291, 1.276)	0.715	(0.315, 1.342)
48	0.810	(0.250, 1.722)	0.843	(0.272, 1.766)
96	0.876	(0.045, 2.174)	0.903	(0.069, 2.204)

Table 4.6: Table of expected $\Delta T-$ VaR_α and ES_α for $\alpha = 0.995$ with corresponding 95% confidence interval for the Black-Scholes, Merton, Heston and Bates model and $\Delta T \in \{12, 24, 48, 96\}$ months.



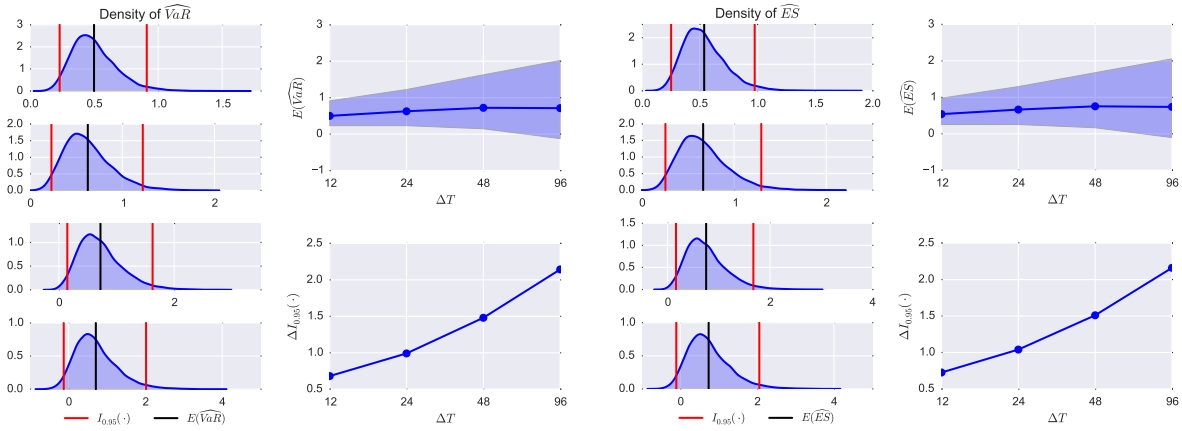
a: $\widehat{\text{VaR}}_{0.995}$ b: $\widehat{\text{ES}}_{0.995}$

Figure 4.9: Top to Bottom Left: Monte Carlo estimated empirical densities of risk measures for $\Delta T \in \{12, 24, 48, 96\}$ months. Expected ΔT estimate (black vertical line) with corresponding 95% empirical confidence interval $I_{0.95}(\text{E}(\cdot))$ (red vertical lines). Top Right: Expected risk measure with corresponding 95% confidence interval as a function of $\Delta T \in \{12, 24, 48, 96\}$ months. Bottom Right: Width of 95% confidence interval as a function of $\Delta T \in \{12, 24, 48, 96\}$ months. All plots generated using $N = 10\,000$ simulated paths of the Black-Scholes model.



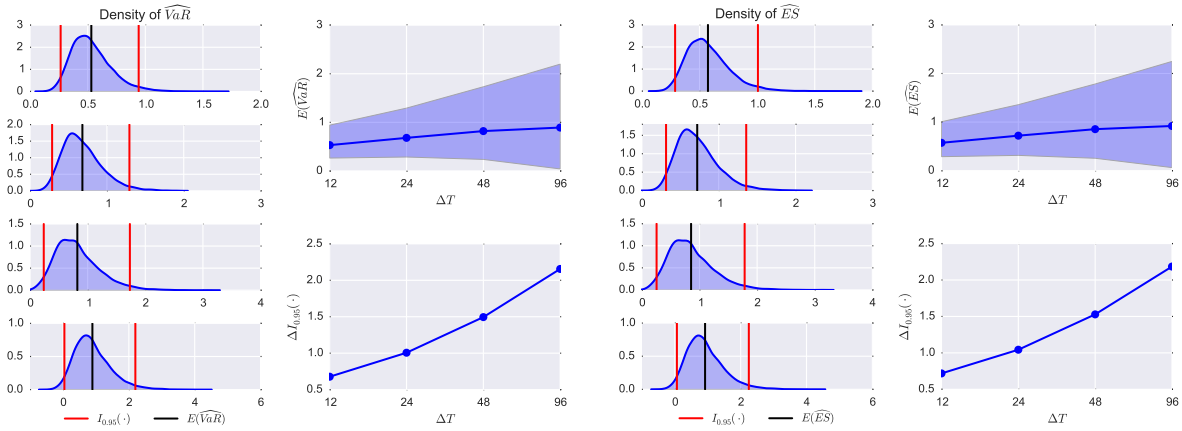
a: $\widehat{\text{VaR}}_{0.995}$ b: $\widehat{\text{ES}}_{0.995}$

Figure 4.10: Top to Bottom Left: Monte Carlo estimated empirical densities of risk measures for $\Delta T \in \{12, 24, 48, 96\}$ months. Expected ΔT estimate (black vertical line) with corresponding 95% empirical confidence interval $I_{0.95}(\text{E}(\cdot))$ (red vertical lines). Top Right: Expected risk measure with corresponding 95% confidence interval as a function of $\Delta T \in \{12, 24, 48, 96\}$ months. Bottom Right: Width of 95% confidence interval as a function of $\Delta T \in \{12, 24, 48, 96\}$ months. All plots generated using $N = 10\,000$ simulated paths of the Merton model.



a: $\widehat{\text{VaR}}_{0.995}$ b: $\widehat{\text{ES}}_{0.995}$

Figure 4.11: Top to Bottom Left: Monte Carlo estimated empirical densities of risk measures for $\Delta T \in \{12, 24, 48, 96\}$ months. Expected ΔT estimate (black vertical line) with corresponding 95% empirical confidence interval $I_{0.95}(\mathrm{E}(\cdot))$ (red vertical lines). Top Right: Expected risk measure with corresponding 95% confidence interval as a function of $\Delta T \in \{12, 24, 48, 96\}$ months. Bottom Right: Width of 95% confidence interval as a function of $\Delta T \in \{12, 24, 48, 96\}$ months. All plots generated using $N = 10\ 000$ simulated paths of the Heston model.



a: $\widehat{\text{VaR}}_{0.995}$ b: $\widehat{\text{ES}}_{0.995}$

Figure 4.12: Top to Bottom Left: Monte Carlo estimated empirical densities of risk measures for $\Delta T \in \{12, 24, 48, 96\}$ months. Expected ΔT estimate (black vertical line) with corresponding 95% empirical confidence interval $I_{0.95}(\mathrm{E}(\cdot))$ (red vertical lines). Top Right: Expected risk measure with corresponding 95% confidence interval as a function of $\Delta T \in \{12, 24, 48, 96\}$ months. Bottom Right: Width of 95% confidence interval as a function of $\Delta T \in \{12, 24, 48, 96\}$ months. All plots generated using $N = 10\ 000$ simulated paths of the Bates model.

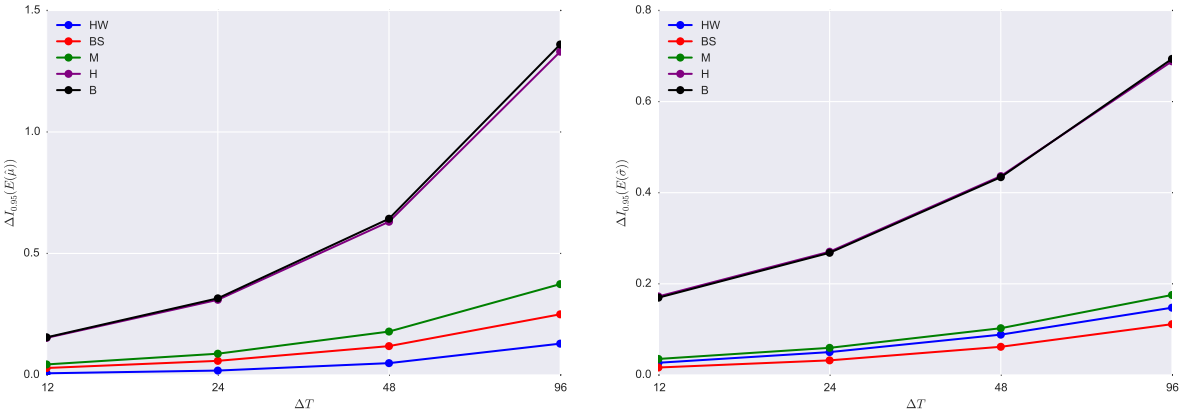
4.3 Model Comparison

To better evaluate the difference in predictive power of the models, this section will study the uncertainty of the considered statistical quantities and compare them between the different models. The result is based on $N = 10\ 000$, $T = 30$ year simulated trajectories of the different models using the rolling-window procedure for $\Delta T = \{12, 24, 48, 96\}$ months. The outline of the section is equal to that seen in Section 4.1 and Section 4.2.

4.3.1 Summary Statistics

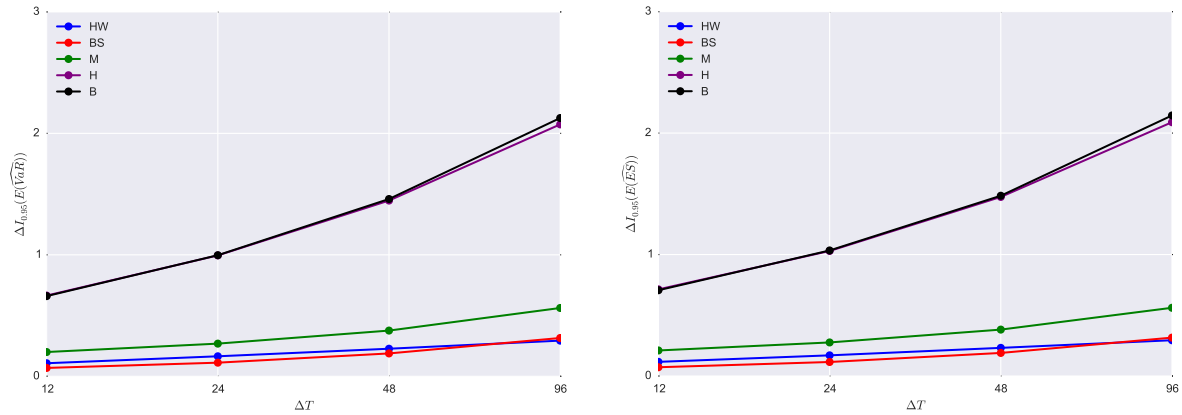
The result of the model comparison is given in Figure 4.13 where the width of the 95% confidence interval of the sample mean (Figure 4.13a) and standard deviation (Figure 4.13b), $\Delta I_{0.95}(\cdot)$ is plotted as a function of ΔT. For both quantities we see a clear division in the predictive power of the models. The Hull-White, Black-Scholes and Merton model perform similar to each other with a considerably narrower confidence interval than for Bates and Heston for all window lengths.

Considering the fact that Heston and Bates both perform similar over all window length, compared to the Merton model, the result in Figure 4.13 would strongly indicate that the introduction of stochastic volatility in the model directly determines the accuracy of which the rolling-window procedure manage to predict future returns statistics. The addition of jumps, as seen if comparing Black-Scholes and Merton do effect the performance, albeit more so for longer prediction horizon such as $\Delta T \in \{48, 96\}$ months. This is, however not of any great significance in comparison to stochastic volatility, and if stochastic volatility is already incorporated as in the Bates model, the addition of jumps carries little additional uncertainty.



a: $\hat{\mu}$ b: $\hat{\sigma}$

Figure 4.13: Left: Width of 95% confidence interval of the expected $\Delta T - \hat{\mu}$ as a function of $\Delta T \in \{12, 24, 48, 96\}$ months. Right: Width of 95% confidence interval of the expected $\Delta T - \hat{\sigma}$ as a function of $\Delta T \in \{12, 24, 48, 96\}$ months. Hull-White (HW), Black-Scholes (BS), Merton (M), Heston (H) and Bates (B)



a: $\widehat{\mathrm{VaR}}_{0.995}$ b: $\widehat{\mathrm{ES}}_{0.995}$

Figure 4.14: Left: Width of 95% confidence interval of the expected $\Delta T - \widehat{\mathrm{VaR}}_{\alpha}$, $\alpha = 0.995$ as a function of $\Delta T \in \{12, 24, 48, 96\}$ months. Right: Width of 95% confidence interval of the expected $\Delta T - \widehat{\mathrm{ES}}_{\alpha}$, $\alpha = 0.995$ as a function of $\Delta T \in \{12, 24, 48, 96\}$ months. Hull-White (HW), Black-Scholes (BS), Merton (M), Heston (H) and Bates (B)

4.3.2 Risk Measures

For the result of the model comparison of the risk measures given in Figure 4.14 we see a similar behavior as in Figure 4.13, where there is an evident division in the performance of the models including stochastic volatility to those without. This would strongly support the discussion in Section 4.3.1 where the major increase in uncertainty lie in whether or not stochastic volatility is incorporate into the model.

Value-at-Risk versus Expected Shortfall

Given the recent adoption of the Expected Shortfall being the new preferred risk measure to use in the evaluation of market risk under Basel III, it would be of interest to study whether ES poses a more complex estimation procedure, requiring a more substantial amount of historical data than VaR in order to achieve the same level of accuracy. To further study this, we will consider the quotient

$$\tau := \frac{\Delta I_{0.95}\left(\mathrm{E}\left(\widehat{\mathrm{ES}}_{0.995}\right)\right)}{\Delta I_{0.95}\left(\mathrm{E}\left(\widehat{\mathrm{VaR}}_{0.995}\right)\right)}.$$

If $\tau > 1$ this means that the width of the confidence interval is smaller for VaR than ES and thus a more accurate estimate given the amount of historical data accessible. If however, $\tau < 1$ the opposite is true and ES is a more accurate estimate. If τ is equal to one, the estimation accuracy is the same for ES as for VaR and the rolling-window procedure allows for future market risk prediction with the same level of accuracy for both risk measures.

From the result depicted in Figure 4.15, we see that all models manage to estimate the Value-at-Risk better than the Expected Shortfall with $\tau > 1$ for all models and $\Delta T = 12$ months. However, for larger ΔT, this effect becomes smaller and for $\Delta T = 96$ months, the difference between the estimation of VaR and ES is small with a quotient less than 1.01% for all models. This is presumably a consequence due to the overall lack of predictive power of the models for longer horizons in general, as seen in Section 4.1.3 and Section 4.2.4. When the estimation uncertainty becomes larger, the difference between the accuracy of ES and VaR is expected to become less pronounced due to the overall lack of accuracy.

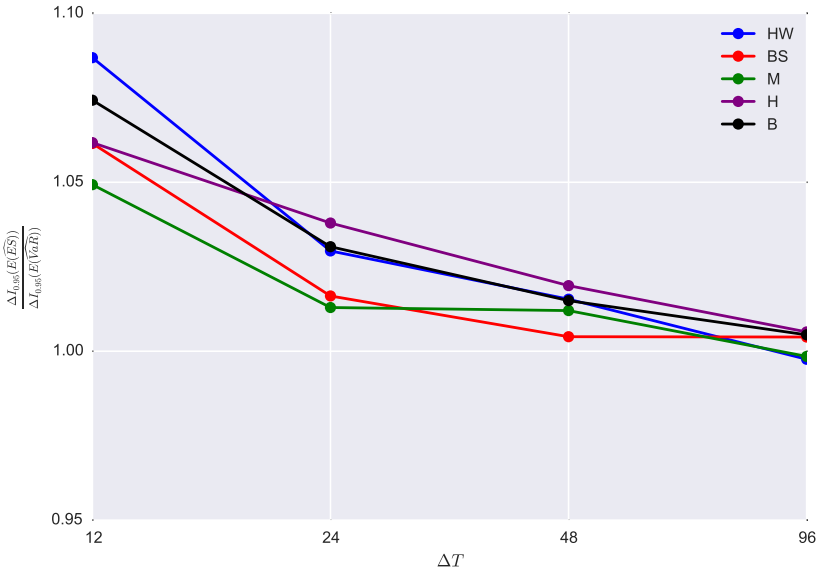


Figure 4.15: Uncertainty comparison of $\widehat{\text{VaR}}_\alpha$ and $\widehat{\text{ES}}_\alpha$ for $\alpha = 0.995$ and $\Delta T \in \{12, 24, 48, 96\}$ months. Hull-White (HW), Black-Scholes (BS), Merton (M), Heston (H) and Bates (B)

4.4 Joint Framework

In this section, the result of the estimation of the bivariate statistical quantities will be presented. The section begins with the estimation of the correlation between the two asset classes and continues with the evaluation of the accuracy of the diversification effect. The section concludes with Section 4.4.2 where a performance comparison of the different frameworks, like the one seen in Section 4.3, is presented.

4.4.1 Risk Measure and Dependence

In order to evaluate the estimation accuracy of the Pearson rho correlation coefficient, the two asset classes needs to be simulated dependently to one another. In order to do so, the sampling of the joint framework was done using a Gaussian copula with zero mean and correlation matrix R equal to

$$R = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}, \tag{4.2}$$

with $\rho = 0.05$ using the distributional sampling approach presented in Section 3.2.4.

The estimation of the correlation between the two[2] asset classes was then performed using the rolling-window procedure and (3.15) for the simulated trajectories of the four frameworks[3].

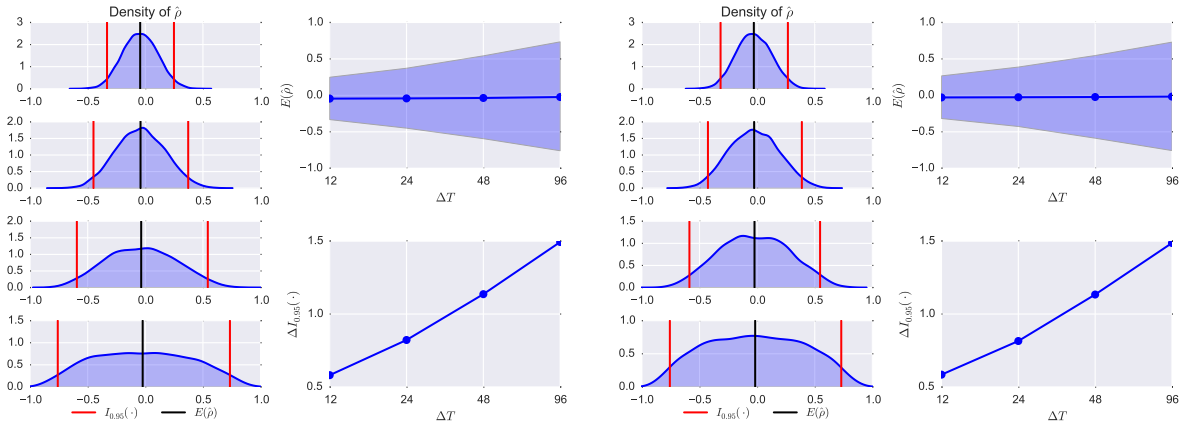
From the result presented in Figure 4.16 and Table 4.7 it is possible to conclude that all frameworks fail to, over all horizons, accurately predict future correlations and would strongly indicate that the estimation accuracy of correlation is not dependent on the complexity of the model.

	BS-HW		M-HW	
ΔT	$\mathrm{E}(\hat{\rho})$	$I_{0.95}(\cdot)$	$\mathrm{E}(\hat{\rho})$	$I_{0.95}(\cdot)$
12	-0.043	(-0.333, 0.254)	-0.030	(-0.315, 0.261)
24	-0.041	(-0.450, 0.371)	-0.028	(-0.426, 0.378)
48	-0.038	(-0.596, 0.532)	-0.026	(-0.581, 0.545)
96	-0.033	(-0.760, 0.729)	-0.022	(-0.751, 0.728)
	H-HW		B-HW	
ΔT	$\mathrm{E}(\hat{\rho})$	$I_{0.95}(\cdot)$	$\mathrm{E}(\hat{\rho})$	$I_{0.95}(\cdot)$
12	-0.040	(-0.325, 0.253)	-0.041	(-0.335, 0.248)
24	-0.036	(-0.435, 0.370)	-0.039	(-0.442, 0.374)
48	-0.029	(-0.582, 0.544)	-0.034	(-0.589, 0.529)
96	-0.020	(-0.745, 0.738)	-0.024	(-0.751, 0.725)

Table 4.7: Expected ΔT correlation with corresponding 95% confidence interval for $\Delta T \in \{12, 24, 48, 96\}$ months and the four different frameworks; Black-Scholes-Hull-White (BS-HW), Merton-Hull-white (M-HW), Heston-Hull-White (H-HW) and Bates-Hull-White (B-HW).

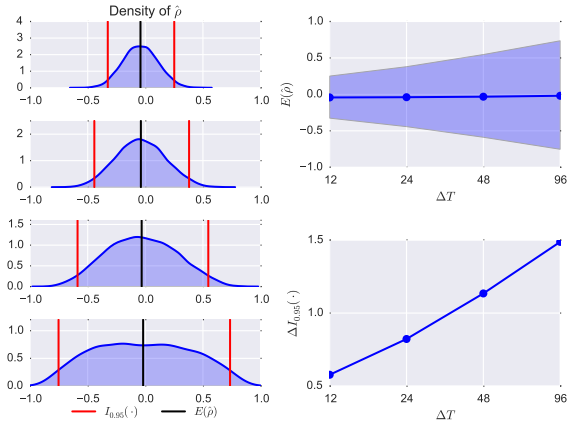
[2] Here we want to point out the fact that bonds and interest rates are negatively correlated and so, by imposing a positive correlation between the interest rate and equity process, we would expect a negative correlation between bonds and equities.

[3] Black-Scholes-Hull-White (BS-HW), Merton-Hull-white (M-HW), Heston-Hull-White (H-HW) and Bates-Hull-White (B-HW)

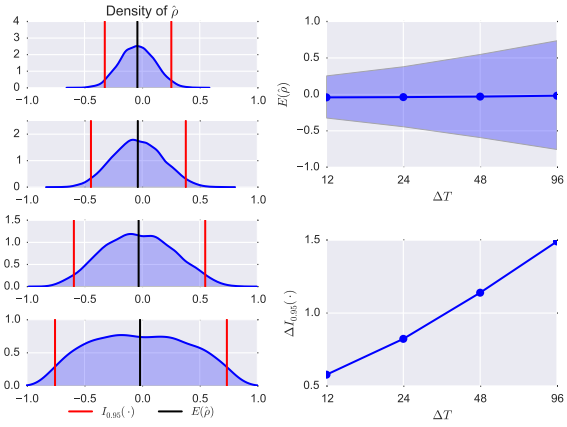


a: BS-HW

b: M-HW



c: H-HW



d: B-HW

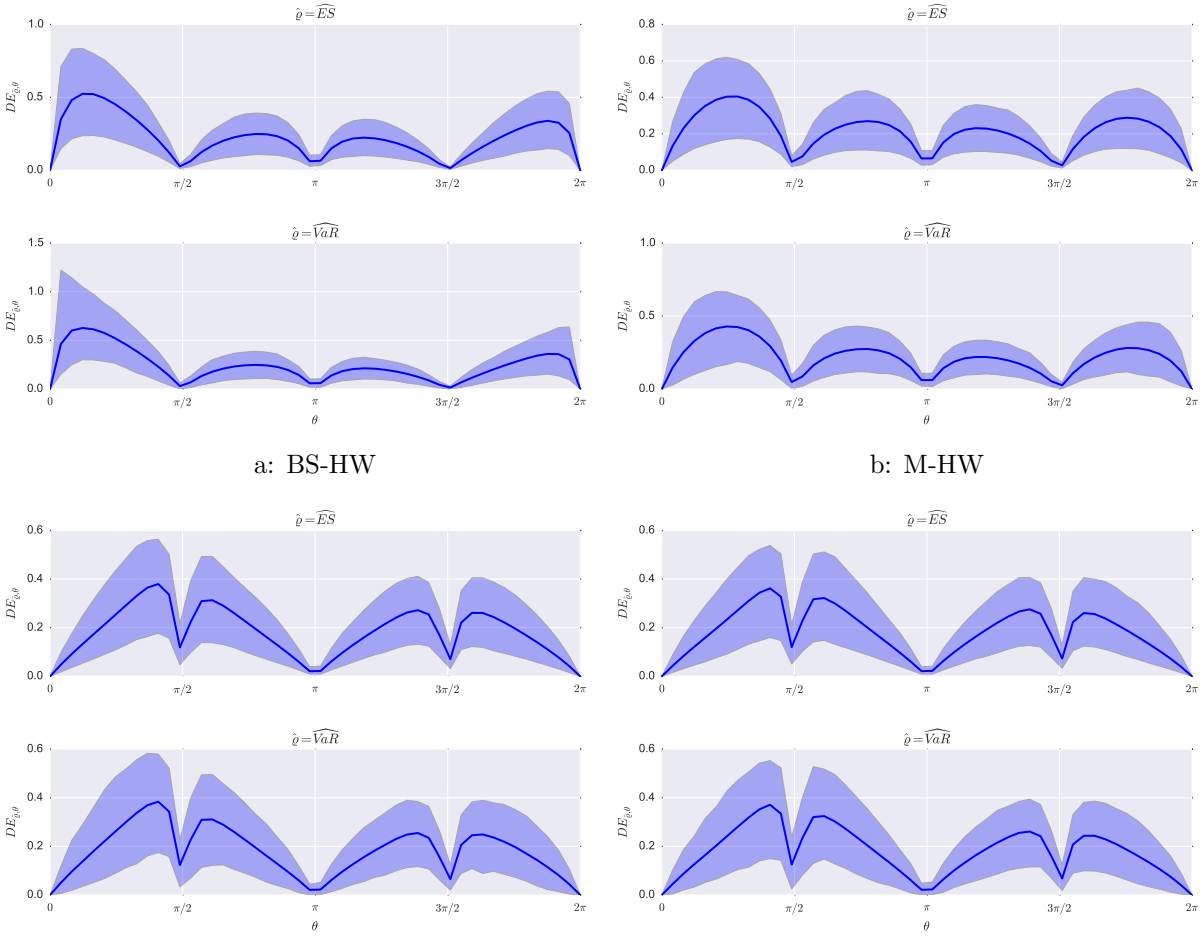
Figure 4.16: Top to Bottom Left: Monte Carlo estimated empirical densities of $\hat{\rho}$ for $\Delta T \in \{12, 24, 48, 96\}$ months. Expected ΔT-correlation (black vertical line) with corresponding 95% empirical confidence interval $I_{0.95}(\mathrm{E}(\hat{\rho}))$ (red vertical lines). Top Right: Expected correlation with corresponding 95% confidence interval as a function of $\Delta T \in \{12, 24, 48, 96\}$ months. Bottom Right: Width of the 95% confidence interval as a function of $\Delta T \in \{12, 24, 48, 96\}$ months and the four different frameworks; Black-Scholes-Hull-White (BS-HW), Merton-Hull-white (M-HW), Heston-Hull-White (H-HW) and Bates-Hull-White (B-HW).

Diversification Effect

To evaluate the diversification we recall from its definition given in (2.55) that it requires the existence of a loss distribution with mass on both side of the origin. For longer horizons, $\Delta T \in \{24, 48, 96\}$ months, the rolling-window procedure tends to generate a majority of negative losses due to the positive drift of the processes. For that reason, we will only consider the evaluation of the diversification effect for $\Delta T = 12$ months. Using (3.16), the diversification effect was evaluated for $\hat{\varrho} \in \{\widehat{\mathrm{VaR}}_\alpha, \widehat{\mathrm{ES}}_\alpha\}$, $\alpha = 0.95$ and $\theta \in [0, 2\pi]$, using 50 different portfolio combinations, equally spaced over the unit circle. The position of the two asset classes is equal to that presented in Table 2.1, where the weight of equities and interest rates is given by $\cos\theta$ and $\sin\theta$ respectively. The result, presented in Figure 4.17; BS-HW (Figure 4.17a), M-HW (Figure 4.17b), H-HW (Figure 4.17c) and B-HW (Figure 4.17d), would indicate, for $\rho = 0.05$ that the optimal diversification effect between the two risk factors is obtained for $\theta \in (0, \frac{\pi}{2})$ which corresponds to a long position in both equities and interest rates (i.e., short in bonds) where, for Heston and Bates, the optimal position is shifted towards a more substation long position in interest rates. For Black-Scholes and Merton we see a more equally weighted portfolio.

For each quadrant, the corresponding 95% confidence interval of the estimates is depicted (shaded blue). Here, we see large uncertainty in the estimates, especially for $\theta \in (0, \frac{\pi}{2})$ where, on a 95% statistical level, it is not possible to evaluate the diversification effect more accurately than $\pm 20\%$.

However, worth pointing out is the fact that the diversification effect appears to be a stable estimate, where the choice of risk measure, Value-at-Risk or Expected Shortfall, appears to have no significant impact. This is surprising given that the base assumption of the diversification effect is that the risk measure is coherent which, indeed is not the case for VaR.



a: BS-HW

b: M-HW

c: H-HW

d: B-HW

Figure 4.17: Monte Carlo estimated Diversification Effect for $\hat{\varrho} \in \left\{\widehat{\mathrm{VaR}}_\alpha, \widehat{\mathrm{ES}}_\alpha\right\}$, $\alpha = 0.95$ with 95% confidence interval as a function of $\theta \in [0, 2\pi]$ for $\Delta T = 12$ months of the four frameworks; Black-Scholes-Hull-White (BS-HW), Merton-Hull-white (M-HW), Heston-Hull-White (H-HW) and Bates-Hull-White (B-HW).

4.4.2 Framework Comparison

To compare the accuracy of the four frameworks we will consider the correlation coefficient and the diversification effect.

Starting with the correlation coefficient, the estimation of the coefficient using the four frameworks were performed. The resulting width of the confidence interval $\Delta I_{0.95}(\mathrm{E}(\hat{\rho}))$ was then plotted as a function of ΔT. From the result, depicted in Figure 4.18, it is clear

that all frameworks perform equally with the width of the confidence interval being equal over all models. This result is thus in accordance with the discussion of the result seen in Figure 4.16 and Table 4.7 and strongly demonstrate the difficulty in estimating future correlation between asset classes.

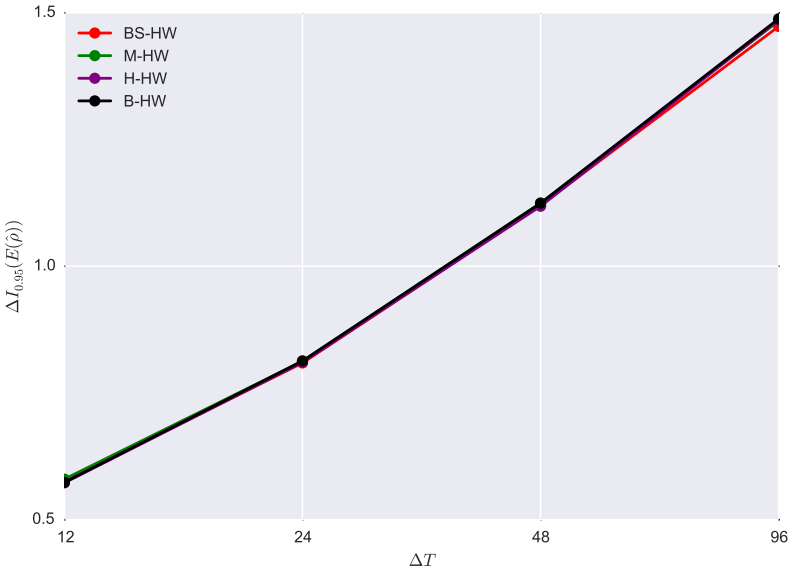


Figure 4.18: Width of 95% confidence interval of the expected $\Delta T - \hat{\rho}$ as a function of $\Delta T \in \{12, 24, 48, 96\}$ months for Black-Scholes-Hull-White (BS-HW), Merton-Hull-white (M-HW), Heston-Hull-White (H-HW) and Bates-Hull-White (B-HW).

Diversification Effect and Copulas

To evaluate the behavior of the diversification effect between the four frameworks, we will study the dependence of the uncertainty in its estimation as a function of the choice of copula through the coefficient of lower tail dependence λ_l. Due to the to lack of effect in the choice of risk measure seen in Figure 4.17, we here only present the result connected to $\hat{\varrho} = \widehat{\mathrm{ES}}_\alpha,\ \alpha = 0.95$.

In order to evaluate the copula dependence of the diversification effect we will study the uncertainty of the diversification effect $\Delta I_{0.95}(\mathrm{E}(\widehat{\mathrm{DE}}_{\hat{\varrho},\theta}))$ as a function of the coefficient of lower tail dependence λ_l. Given that the analytical expression of λ_l is known for both the t copula and the Clayton copula (see (2.42) and (2.43)) this allows to investigate the dependence of the estimation accuracy of the diversification effect in the choice of copula. λ_l is calculated using a t copula with varying degrees of freedom $\nu \in [2, 30]$ as well as a Clayton copula. The overall correlation coefficient for this analysis is set to $\rho = 0.8$

and $\theta \in \frac{1}{4}\{\pi, 3\pi, 5\pi, 7\pi\}$. The result of the study presented in Figure 4.19 is based on $N = 10\ 000$ simulated trajectories of the frameworks for $\Delta T = 12$ months.

From the result in Figure 4.19, we see that the uncertainty level fail to demonstrate any clear dependence towards the choice of copula, in any of the four quadrants. Based on this observation, we therefore conclude that the choice of copula does not seem to have any impact in the estimation uncertainty of the diversification effect.

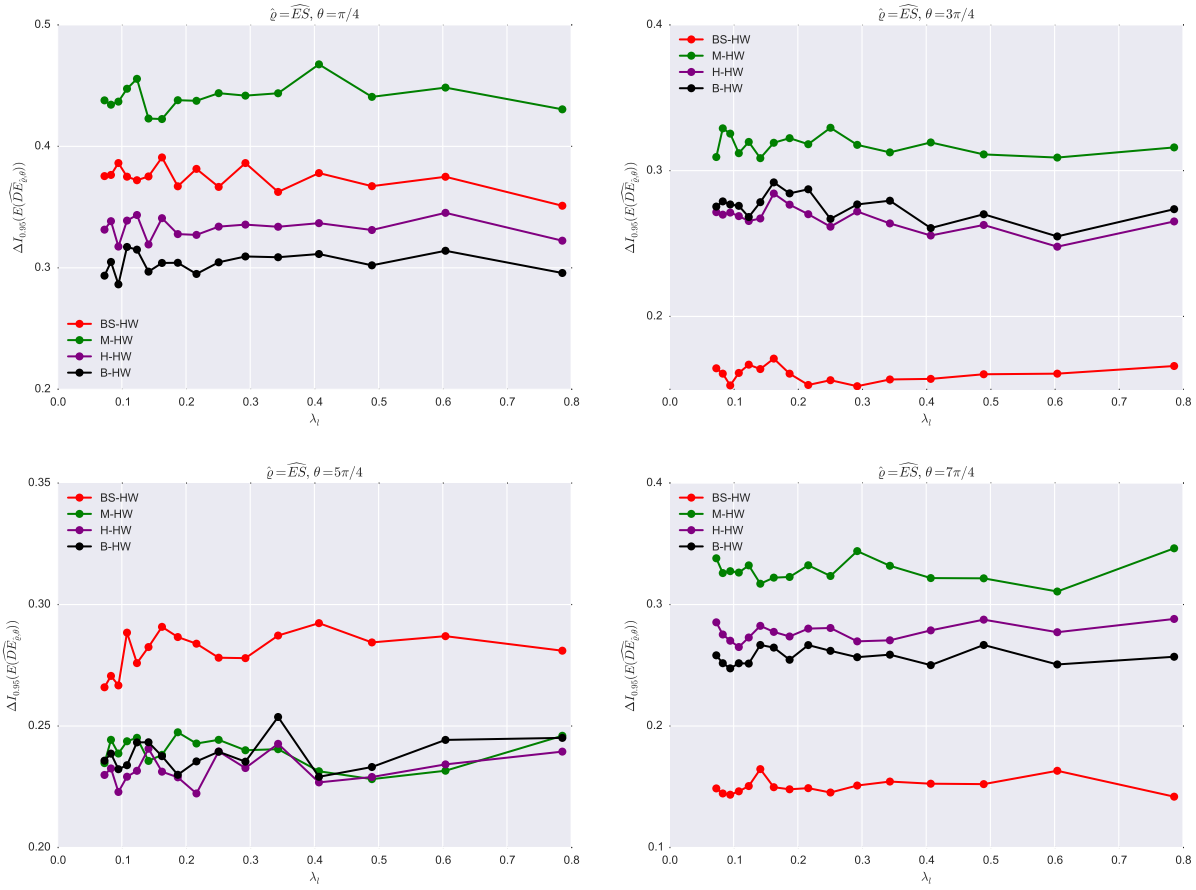


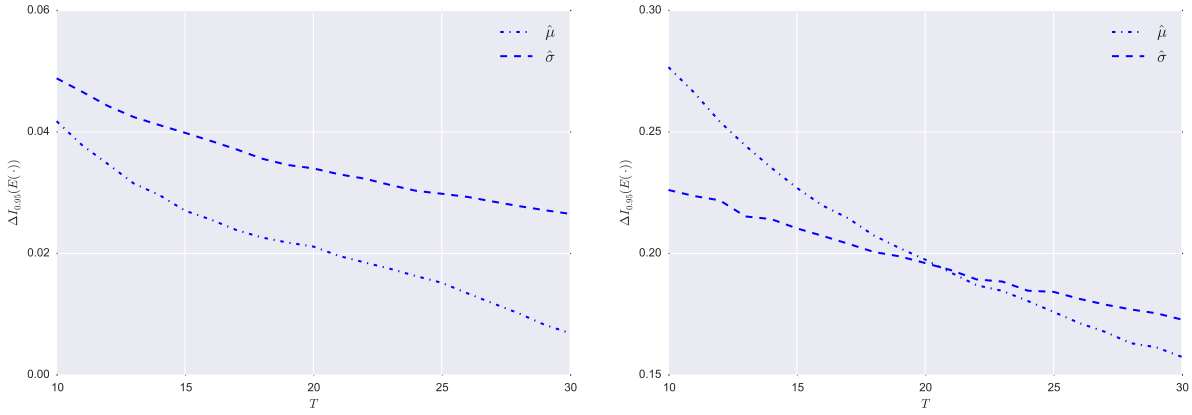
Figure 4.19: Estimation sensitivity to the choice of copula in the estimation of the diversification effect. Top Left: Estimation sensitivity of diversification effect for $\theta = \pi/4$. Top Right: Estimation sensitivity of diversification effect for $\theta = 3\pi/4$. Bottom Left: Estimation sensitivity of diversification effect for $\theta = 5\pi/4$. Bottom Right: Estimation sensitivity of diversification effect for $\theta = 7\pi/4$. $\Delta T = 12$ months and $\hat{\varrho} = \widehat{\mathrm{ES}}_{0.95}$.

4.5 Data Comparative Study

To further visualise the necessity of sufficient amount of data available we, in this section compare the prediction accuracy of the statistics based on data ranging between $T = 10$ and $T = 30$ years using monthly frequent historical data and a prediction horizon of $\Delta T = 12$ months. For visualisation purposes, we here only consider the estimation accuracy of the summary statistics for Hull-White and Bates. As demonstrated in Section 4.1.2 and Section 4.2.2 these two models constitute the two endpoints in the complexity spectra of the models being considered in this thesis with Bates being the most volatile and Hull-White its adverse. The convergence rate for the rest of the models will, trivially, range between that of Bates and Hull-White dependent on their complexity.

In order to study the dependency of the precision and the amount of available data available, increasingly longer trajectories between $T = 10$ and $T = 30$ were simulated. Here, with a one year increase between each set of simulation. For each T, $N = 10\ 000$ sample paths were generated from which a ΔT-statistic was calculated for each trajectory. The result is a series of confidence widths for each statistic and each trajectory length. The result for the two models and their summary statistics can be seen in Figure 4.20a (Hull-White) and Figure 4.20b (Bates). For both models we see a significant improvement between using $T = 10$ and $T = 30$ years of data. This is of course a reflection of the difference in the number of observations available yielding a more accurate ergodic convergence to their stationary statistics. Moreover, we see that the convergence rate of the standard deviation seems to be slower than for the mean.

The corresponding study of the risk measures is omitted here because of the overall lack of accuracy in the result presented in Section 4.1.3 and Section 4.2.4 making the data comparison inconclusive.



a: Hull-White

b: Bates

Figure 4.20: Left: Width of 95% confidence interval of the expected $\Delta T - \hat{\mu}$ and $\Delta T - \hat{\sigma}$ as a function of T ranging between 10 and 30 years for Hull-White. Right: Width of 95% confidence interval of the expected $\Delta T - \hat{\mu}$ and $\Delta T - \hat{\sigma}$ as a function of T ranging between 10 and 30 years for Bates. $\Delta T = 12$ months.

Chapter 5

Discussion and Conclusion

In this final chapter, we summarises the thesis and its findings. Section 5.1 discuss the results presented in Chapter 4 and the performance of the considered method. This is followed by a conclusion summarising the findings in Section 5.2. Section 5.3 concludes this thesis by outlining some available extensions to the model and possible future work to considers.

5.1 Discussion

By inspection of the results presented in Chapter 4, the Hull-White and Black-Scholes can be considered the best performing models. Given the less volatile behavior of these two models without the addition of stochastic volatility or jumps, this can be considered an expected result. The addition of stochastic volatility was found to be the most significant contributing factor in the prediction uncertainty, as demonstrated by the poor descriptive statistics and risk measures produces by the Heston and Bates models. In this case both models performed considerably worse than by solely introducing jumps as in the Merton model.

In Section 4.2.3 the sensitivity of the return statistics of jumps and stochastic volatility was studied. By varying the jump intensity and the level of volatility of variance it was found that the sensitivity of estimating the mean, the jump intensity had a greater impact than the stochastic volatility. The converse was found for the standard deviation.

For the bivariate statstics, both the correlation coefficient and diversification effect were found to be hard to estimate. For the correlation, this was found to be exceedingly hard to estimate. This can be explained by the sensitivity correlation has to varying dispersion, which makes the linear relationship between the two models hard to estimate. The found result is uniform across the framework with an uncertainty of $\pm 0.25\%$ from the true value on a one year prediction horizon. For the diversification effect, on the other hand, the estimation uncertainty is a result of the difficulty of accurately predict future risk measures.

As found in Section 4.1.3 and Section 4.2.4, estimation of risk measures is not trivial and carry a great amount of uncertainty. Given the definition of the diversification effect, the uncertainty of the estimated risk measure trivially translate over into the accuracy of the diversification effect.

For the considered models, the rolling-window method was found to be most applicable in the forecasting of statistical quantities related to portfolio returns up to one year. However, the estimation of risk measures carried greater uncertainty especially for the equity models. For longer prediction horizons, the method lack significantly in estimation accuracy. One of the reason for this is that the introduced autocorrelation between observations within a window becomes more persistent. At the same time, the number of observation reduces, making the rate at which the statistics convergence to their stationary values to become weaker. Another reason for this is the model specific level of dispersion in the sampling of the trajectories. For some of the models the dispersion becomes too extensive, leading to problems in the validation of approximating simple returns with log-returns. This problem is especially prominent in the forecasting of risk measures corresponding to high loss quantiles, $\geq$ 95th. The result should therefore be interpreted accordingly and not as simple returns.

Albeit the overall performance of the rolling-window method with thirty year might not be 'satisfactory', its performance should be set into comparable context. As seen in Section 4.5 a thirty year rolling-window procedure yields a significantly more accurate estimation than when using ten or twenty years of monthly data.

5.2 Conclusion

This thesis has studied the accuracy implied by a series of thirty year monthly historical data in the forecasting of portfolio and risk measure statistics. In doing so, the rolling-window method has been used in order to retrieve stationary statistical quantities implied by the historical data. To empirically measure the accuracy of the estimated future statistics, Monte Carlo methods were used to calculate relevant confidence intervals for each respective model. The considered models constitutes some of the most widely studied and used financial models in the industry and recent literature for their ability to capture known market behaviors.

The rolling-window method was found to be valid in predicting future statistical quantities related to portfolio return of up to one year maximum. For longer prediction horizons, the uncertainty of the estimates becomes too significant to statistically predict future outcomes. In the case of risk measure, multiple of the models lacked in the accuracy of estimating future market risk on even one year horizon and highlights the difficulty in foreseeing extremities of future market movements. Despite some of its shortcomings, the rolling-window procedure constitutes an easily implemented and versatile predictions methods as it allows for easy extensions in the incorporation of multiple asset classes and

corresponding types of risk.

5.3 Future Work

In this final section, we consider some different ways to move forward in order to obtain more accurate future estimation of statistical quantities using the rolling-window procedure as well as possible extension to the current framework in order to capture risk outside of the market risk umbrella.

5.3.1 Model Extensions

As expressed in Section 5.1 one of the problem with the selected class of models used in this thesis is that many of them simply demonstrate too volatile behavior. This is, in particular, a problem for long-term simulation where the dispersion of the generated trajectory simply becomes too extensive. A solution to this problem could be a ”regime switch”-like model where, for long horizons, the model incorporating a mean-reverting trend in the modelling of equities to compress the dispersion of the models. This would probably improve the result for the more complex equity models including stochastic volatility and jumps.

In the case of interest rates, there are multiple ways forward. Given the vast number of fixed income products relying on the modelling of forward rates rather than short-rates, a natural extension would be to consider other types of interest rate models. Two possible extensions could be, the Heat-Jarrow-Morton framework, where forward rates are directly modelled through diffusion processes. Another would be to use so-called market models which aim to model the dynamics of LIBOR/STIBOR-forward rates (e.g., Brigo and Mercurio (2012), Filipović (2009)). The implementation of such models would make it possible to consider risk connected to positions in numerous of fixed income products, not just bonds.

5.3.2 Alternative Model Selection

An alternative direction would be to model returns of asset classes directly instead of through prices. This approach opens up to a large class of discrete-time models such as autoregressive and moving average models, or a combination of them two. Stochastic volatility can be incorporated by considering a GARCH-like or autoregressive models of some sort and jumps can, for instance, be introduced by considering a Bernoulli random variable as seen in Chib et al. (2002).

5.3.3 Credit Risk Modelling

The fact that bond prices are modelled without the possibility of default in this thesis is of course a simplification. Independent on the issuer of the bond, there is always a probability that the company cannot comply with the obligation to pay the bondholder at maturity.

The fact that firms are susceptible to defaults is reflected in so-called credit risk, which can be accounted for through the modelling of defaults. Denote by τ the time of default of the reference company of a bond. In case of default a so-called recovery value, $R \leq 1$, is usually paid. This is a pre-determined value the company has agreed to pay the bondholder in case of such an events. Given this, the price of the defaultable zero-coupon bonds is given by

$$P^d(t,T) = \mathrm{E}^{\mathrm{Q}}\left(\exp\left\{-\int_t^T r(s)ds\right\}(1(\tau > T) + R1(\tau \leq T)) \;\middle|\; \mathcal{G}(t)\right), \tag{5.1}$$

where $\mathcal{G}(t) \coloneqq \mathcal{F}(t) \vee \sigma\left(\{\tau < u\}, u \leq t\right)$ denotes the filtration with information on whether a default has occurred before t and if so at what time (Brigo and Mercurio, 2012). The super-scripted d in $P^d(\cdot,\cdot)$ reference to the fact that the bond is susceptible to default.

In case $\tau > T$, the price of the bond at maturity is simply equal to one as in the non-defaultable case. If, however, the company defaults before or at maturity, $\tau \leq T$, the value of the bond at maturity is equal to $RP^d(t,\tau) \leq 1$.

Whether or not a bond defaults is often modelled by a so-called *hazard rate* model driven by an inhomogeneous Poisson process with intensity determining the probability of defaults. In the case of the rolling-window procedure this would most likely increase the level of uncertainty in the estimation of future statistics. However, in case of predicting future risk involved in bond investment, this is a necessary characteristic that needs to be incorporated to fully reflect the risk of bond investment.

5.3.4 Historical Validation Procedure

Given that the rolling-window procedure rely on the fact that the empirically estimated return distribution is indeed the stationary counterpart, it would be of interest to compare the results found in Chapter 4 with some historically estimated counterparts. Under the hypothesis that the framework generating the trajectories is the one that generated the historical data, the historically estimated statistics should, in theory, lie within the empirical confidence interval, and preferably as close to the expected statistics as possible. If this is not the case, the considered model is not fully capable of regenerating the stationary distributional properties implied by the historical data and could thus be considered invalid.

Now, this clearly requires some historical calibration procedure of the framework and the included models. As this was not within the scope of this thesis, no relevant conclusions can be drawn in the validation of the result generated by the rolling-window procedure compared to that implied by historical data. Nevertheless, validation procedures are still of great importance and should therefore be furthered studied.

Bibliography

L. Andersen. Efficient simulation of the heston stochastic volatility model. *Technical Report*, 2007.

P. Artzner, F. Delbaen, J.-M. Eber, and D. Heath. Coherent meausures of risk. *Mathematical Finance, Vol. 9, No. 3*, page 203–228, July 1999.

S. D. Bates. Jumps and stochastic volatility: Exchange rate processes implicit in deutsche mark options. *The Review of Financial Studies Springer Vol. 9 No. 1*, pages 69–107, 1996.

T. Björk. *Interest rate Theory*. Springer, 1997.

T. Björk. *Arbitrage Theory in Continuous Time*. Oxford University Press Inc., New York, 3rd edition, 2009.

F. Black and S. Myron. The pricing of options and corporate liabilities. *The Journal of Political Economy, Vol. 81, No 3*, pages 637–654, May - Jun. 1973.

D. Brigo and F. Mercurio. *Interest Rate Models - Theory and Practice*. Springer Finance, 2nd edition, 2012.

M. Broadie and O. Kaya. Exact simulation of stochastic volatility and other affine jump diffusion processes. *Operations Research, Vol. 54, No. 2*, page 217–231, 2006.

M. Broadie, M. Chernov, and M. Johannes. Model specification and risk premia: Evidence from futures options. *The Journal of Finance, Vol. 62, No. 3*, pages 1453–1490, June 2007.

J. P. Brockwell and A. R. Davis. *Time Series: Theory and Methods*. Springer Science + Business Media, LLC, 2nd edition, 1991.

S. Cambanis, S. Huang, and G. Simons. On the theory of elliptically contoured distributions. *Journal of Multivariate Analysis 11*, page 368–385, 1981.

P. Carr and L. Wu. Time-changed lévy processes and option pricing. *Journal of Financial Economics 71*, page 113–141, 2004.

S. Chib, F. Nardari, and N. Shephard. Markov chain monte carlo methods for stochastic volatility models. *Journal of Econometrics 108*, page 281–316, 2002.

R. Cont. Empirical properties of asset returns: stylized facts and statistical issues. *Quantitative Finance Volume 1*, pages 223–236, 2001.

R. Cont and P. Tankov. *Financial Modelling With Jump Processes*. Chapman & Hall/CRC Financial Mathematics Series, 2004.

C. J. Cox, E. J. Ingersoll, and A. S. Ross. A theory of the term structure of interest rates. *Econometrica 53*, pages 385–407, 1985.

D. Duffie. *Dynamic Asset Pricing Theory*. Princeton University Press, 3rd edition, 2001.

B. Eraker. Do stock prices and volatility jump? reconciling evidence from spot and option prices. *The Journal of Finance, Vol. 59, No. 3*, pages 1367–1403, June 2004.

B. Eraker, M. Johannes, and N. Polson. The impact of jumps in volatility and returns. *The Journal of Finance, Vol. 58, No. 3*, pages 1269–1300, June 2003.

W. Feller. Two singular diffusion problems. *Annals of Mathematics, Vol. 54*, pages 173–182, 1951.

W. Feller. *An Introduction to Probability Theory and Its Application Volume II*. John Wiley & Sons Inc., New York, N.Y., 2nd edition, 1971.

D. Filipović. *Dynamic Asset Pricing Theory*. Springer Finance, 2009.

J. Gatheral. *The Volatility Surface: A Practitioner's Guide*. John Wiley & Sons, Inc., 2006.

C. Genest and J. MacKay. The joy of copulas: bivariate distributions with uniform marginals. *The American Statistician, Volume 40, Issue 4*, page 280–283, November 1986.

P. Glasserman. *Monte Carlo Methods in Financial Engineering*. Springer-Verlag, 2004.

D. J. Hamilton. *Time Series Analysis*. Princeton University Press, 1994.

J. E. Hannan. *Multiple Time Series*. John Wiley and Sons, Inc., 1970.

L. S. Heston. A closed-form solution for options with stochastic volatility with applications to bond and currency options. *The Review of Financial Studies Volume 6, number 2*, pages 327–343, 1993.

C. J. Hull and A. White. Branching out. *Risk 7*, pages 34–37, 1994.

H. Hult, F. Lindskog, O. Hammarlid, and C. J. Rhen. *Risk and Portfolio Analysis: Principles and Methods*. Springer, 2012.

P. Jäckel. *Monte Carlo Methods in Finance*. John Wiley & Sons, Inc., Wiley Finance, 2002.

I. Karatzas and S. Shreve, E. *Browian Motion and Stochastic Calculus*. Springer, 2nd edition, 1991.

J. Kienitz and D. Wetterau. *Financial Modelling: Theory, Implementation and Practice with MATLAB Source*. John Wiley & Sons, Inc., Wiley Finance, 2012.

H. C. Kimberling. A probabilistie interpretation of complete monotonicity. *Aequationes Mathematicae 10*, page 152–164, 1974.

G. S. Kou. A jump-diffusion model for option pricing. *Management Science Vol. 48, No 8*, pages 1086–1101, August 2002.

J. A. McNeil and J. Nešlehová. Multivariate archimedean copulas, d-monotone functions and ℓ_1-norm symmetric distributions. *The Annals of Statistic, Vol. 37, No. 5B*, page 3059–3097, 2009.

J. A. McNeil, F. Rüdiger, and P. Embrechts. *Quantitative Risk Management: Concepts, Techniques and Tools*. Princeton University Press, Princeton Series in Finance, revised edition, 2015.

C. R. Merton. Option pricing when underlying stock returns are discontinuous. *Journal of Financial Economics 3*, pages 125–144, 1976.

B. Øksendal. *Stochastic Differential Equations: An Introduction with Applications*. Springer, 6th edition, 2003.

E. P. Protter. *Stochastic Integration and Differential Equations*. Springer-Verlag Berlin Heidelberg, 2nd edition, 2004.

R. Rebonato. *Volatility and Correlation: The Perfect Hedger and the Fox*. John Wiley & Sons, Ltd, 2nd edition, 2004.

S. R. Tsay. *Analysis of Financial Time Series*. John Wiley & Sons, Inc., 3rd edition, 2010.

Y. Yamai and T. Yoshiba. Comparative analyses of expected shortfall and value-at-risk: Their estimation error, decomposition, and optimization. *Monetary and Economic Studies, Vol. 20, Issue 1*, pages 87–121, January 2002.

Appendix A

Hull-White

A.1 Derivation of $\theta(t)$

Following the expression of $\psi(t)$ in (2.16),

$$\psi(t) = e^{-\lambda t} r(0) + \lambda \int_0^t e^{-\lambda(t-s)} \theta(s) dW(s), \tag{A.1}$$

its derivative yields

$$\begin{aligned}
\frac{\partial \psi(t)}{\partial t} &= -\lambda e^{-\lambda t} r(0) + \lambda e^{-\lambda(t-t)} \theta(t) + \lambda \int_0^t \frac{\partial}{\partial t} e^{-\lambda(t-s)} dW(s) \\
&= -\lambda e^{-\lambda t} r(0) + \lambda \theta(t) - \lambda^2 \int_0^t e^{-\lambda(t-s)} dW(s) \\
&= \lambda \theta(t) - \lambda \left(e^{-\lambda t} r(0) - \lambda \int_0^t e^{-\lambda(t-s)} dW(s) \right) \\
&= \lambda \theta(t) - \lambda \psi(t).
\end{aligned}$$

By term rearranging, $\theta(t)$ can be expressed in terms of $\psi(t)$ as

$$\theta(t) = \frac{1}{\lambda} \frac{\partial \psi(t)}{\partial t} + \psi(t). \tag{A.2}$$

To express $\theta(t)$ in terms of the initial term structure we utilis (2.6) and the fact that $f(0,t) = -\frac{\partial}{\partial t} \log P(0,t)$. The forward rate can now be expressed as

$$f(0,t) = \frac{\partial}{\partial t} A(0,t) + r(0) \frac{\partial}{\partial t} B(0,t). \tag{A.3}$$

By solving (2.10) for A and B,

$$A(t,T) = -\frac{\eta^2}{2}\int_t^T B^2(s,T)ds + \lambda\int_t^T \theta(s)B(s,T)ds,$$
$$B(t,T) = \frac{1}{\lambda}\left(1 - e^{-\lambda(T-t)}\right).$$

and substituting it into (A.3) yields

$$\begin{aligned} f(0,t) &= \frac{\eta^2}{2}\int_0^t \frac{\partial}{\partial s}B^2(s,t)ds + \lambda\int_0^t \theta(s)\frac{\partial}{\partial s}B(s,t)ds + r(0)\frac{\partial}{\partial t}B(0,t) \\ &= e^{-\lambda t}r(0) + \lambda\int_0^t e^{-\lambda(t-s)}\theta(s)ds - \frac{\eta^2}{2\lambda^2}\left(1 - e^{-\lambda t}\right)^2 \\ &= \psi(t) - \frac{\eta^2}{2\lambda^2}\left(1 - e^{-\lambda t}\right)^2. \end{aligned}$$

Consequently we get that

$$\psi(t) = f(0,t) + \frac{\eta^2}{2\lambda^2}\left(1 - e^{-\lambda t}\right)^2 \tag{A.4}$$

and thus from (A.2), $\theta(t)$ is given by

$$\theta(t) = f(0,t) + \frac{1}{\lambda}\frac{\partial}{\partial t}f(0,t) + \frac{\eta^2}{2\lambda^2}\left(1 - e^{-2\lambda t}\right). \tag{A.5}$$

A.2 Derivation of the Zero-Coupon Bond Price

To derive the zero-coupon bond price given in (2.22) we recall the Hull-White decomposition given in (2.20) and the affine bond price

$$P(t,T) = e^{-A(t,T)-B(t,T)r(t)}. \tag{A.6}$$

By rewriting $r(t)$ in terms of $\psi(t)$ and $\tilde{r}(t)$, the bond price becomes

$$P(t,T) = \exp\{-A(t,T) - B(t,T)(\tilde{r}(t) + \psi(t))\}. \tag{A.7}$$

In terms of the ATS functions of $\tilde{r}$, A and B are given by;

$$A(t,T) = \tilde{A}(t,T) + \lambda\int_t^T \theta(s)B(s,T)ds, \tag{A.8}$$
$$B(t,T) = \tilde{B}(t,T). \tag{A.9}$$

This means that (A.7) can be rewritten as

$$\begin{aligned} P(t,T) &= \exp\left\{-\tilde{A}(t,T) - \tilde{B}(t,T)\tilde{r}(t) - \lambda\int_t^T \theta(s)B(s,T)ds - \psi(t)B(t,T)\right\} \\ &= \tilde{P}(t,T)\exp\left\{-\lambda\int_t^T \theta(s)B(s,T)ds - \psi(t)B(t,T)\right\}. \end{aligned}$$

Substituting the expression for $\theta(t)$ in (A.2) and integration by parts now yields the bond price expression given in (2.22),

$$\begin{aligned} P(t,T) &= \tilde{P}(t,T)\exp\left\{-\lambda\int_t^T \left\{\frac{1}{\lambda}\frac{\partial\psi(s)}{\partial s} + \psi(s)\right\}B(s,T)ds - \psi(t)B(t,T)\right\} \\ &= \tilde{P}(t,T)\exp\left\{-\left(\int_t^T \frac{\partial\psi(s)}{\partial s}B(s,T)ds + \lambda\int_t^T \psi(s)B(s,T)ds\right) - \psi(t)B(t,T)\right\} \\ &= \tilde{P}(t,T)\exp\left\{-\left(\psi(s)B(s,T)\Big|_t^T - \int_t^T \psi(s)\frac{\partial B(s,T)}{\partial s}ds + \lambda\int_t^T \psi(s)B(s,T)ds\right) - \psi(t)B(t,T)\right\} \\ &= \tilde{P}(t,T)\exp\left\{-\left(-\psi(t)B(t,T) - \int_t^T \psi(s)\frac{\partial B(s,T)}{\partial s}ds + \lambda\int_t^T \psi(s)B(s,T)ds\right) - \psi(t)B(t,T)\right\} \\ &= \tilde{P}(t,T)\exp\left\{\int_t^T \psi(s)\frac{\partial B(s,T)}{\partial s}ds - \lambda\int_t^T \psi(s)B(s,T)ds\right\} \\ &= \tilde{P}(t,T)\exp\left\{\int_t^T \psi(s)\left\{\frac{\partial B(s,T)}{\partial s} - \lambda B(s,T)\right\}ds\right\} \\ &= \tilde{P}(t,T)\exp\left\{\int_t^T \psi(s)\left\{-e^{-\lambda(T-s)} - (1-e^{-\lambda(T-s)})\right\}ds\right\} \\ &= \tilde{P}(t,T)\exp\left\{-\int_t^T \psi(s)ds\right\} \\ &\overset{(A.4)}{=} \tilde{P}(t,T)\exp\left\{-\int_t^T \left\{f(0,s) + \frac{\eta^2}{2\lambda^2}\left(1-e^{-\lambda s}\right)^2\right\}ds\right\} \\ &= \tilde{P}(t,T)\frac{P(0,T)}{P(0,t)}\exp\left\{-\frac{\eta^2}{2\lambda^2}\int_t^T \left(1-e^{-\lambda s}\right)^2 ds\right\}. \end{aligned}$$

Appendix B

Stochastic Calculus

B.1 Equivalent Martingale Measures

The notion of equivalent martingale measures (EMM) is a central part in the pricing of financial derivatives. The existence of an EMM allows to express the price of financial derivative at time t as the discounted expected price at time T of the very same derivative.

The existence of an EMM $\mathrm{Q} \sim \mathrm{P}$ is given by the Radon-Nikodym derivative

$$\left.\frac{d\mathrm{Q}}{d\mathrm{P}}\right|_{\mathcal{F}(t)} = \mathcal{E}(-\gamma(t) \cdot W(t)). \tag{B.1}$$

Here, $\mathcal{E}$ denotes the Doléans-Dade exponential (also called the stochastic exponential) defined as

$$\mathcal{E}(X(t)) = \exp\left\{X(t) - \frac{1}{2}\langle X(t)\rangle\right\}$$

and γ is the market price of risk. As seen from (B.1), γ is the main component in the connection between the two probability spaces.

B.2 Girsanov's Theorem

Girsanov's theorem states that, given an EMM, $\mathrm{Q} \sim \mathrm{P}$ such that $\mathcal{E}(-\gamma(t) \cdot W(t))$ is a martingale and W a P−Brownian motion, then it holds that

$$W^*(t) = W(t) + \int_0^t \gamma(s)ds \tag{B.2}$$

is a Brownian motion under Q.

In the case of the short-rate model defined in Chapter 2, the transformation between the dynamics under P and Q is easily derived considering the general dynamics

$$dr(t) = \mu(t, r(t))dt + \sigma(t, r(t))dW(t). \tag{B.3}$$

By substituting $dW(t)$ for $dW^*(t) - \gamma(t)dt$ given in (B.2), the dynamics in (B.3) becomes

$$dr(t) = (\mu(t, r(t)) - \gamma(t)\sigma(t, r(t)))\, dt + \sigma(t, r(t))dW^*(t). \tag{B.4}$$

The relationship between the dynamics of r under P and Q therefore directly rely on the knowledge of γ. As long as we are only concerned with the pricing of derivatives, the market price of risk is given endogenously and we can directly model r under Q.

Appendix C

Clayton

Let $X \sim \mathsf{Ga}(\alpha, \beta)$ with corresponding density function $f(x) = \frac{1}{\Gamma(\alpha)\beta^\alpha} x^{\alpha-1} e^{-x/\beta}$, where $\Gamma(z)$ is the gamma function given by $\Gamma(z) = \int_0^\infty x^{z-1} e^{-x} dx$. Using Laplace-Stiltjes transform on $X \sim \mathsf{Ga}(1/\theta, 1)$ yields

$$\psi(t) = \int_0^\infty \frac{e^{-tx} x^{1/\theta-1} e^{-x}}{\Gamma(1/\theta)} dx = \frac{1}{\Gamma(1/\theta)} \int_0^\infty e^{-(1+t)x} x^{1/\theta-1} dx. \tag{C.1}$$

Using the variable substitution $s \coloneqq 1 + t$ and $p \coloneqq 1/\theta - 1$, the integral in (C.1) becomes

$$\int_0^\infty e^{-sx} x^p dx.$$

As it turns out, this is nothing else than the Laplace transform of $f(x) = x^p$ given by

$$\mathcal{L}_s\{x^p\} = \frac{\Gamma(p+1)}{s^{p+1}}, \quad \Re(p) > -1.$$

Substituting[1] the transform back into (C.1) yields

$$\psi(t) = \frac{1}{\Gamma(1/\theta)} \frac{\Gamma(p+1)}{s^{p+1}} \bigg|_{\substack{s=1+t \\ p=1/\theta-1}} = (1+t)^{-1/\theta}.$$

Indeed this is not exactly the Archimedean copula generator of the Clayton copula introduced in Chapter 2, however it is easy to see that $\psi(t)$ and $\psi(\theta t)$ generates the same Archimedean copula.

[1] Here, $\Re(\cdot)$ denotes the real part operator such that for a $p \in \mathbb{C}$, $\Re(p) \in \mathbb{R}$.

TRITA -MAT-E 2017:12
ISRN -KTH/MAT/E--17/12--SE