

Analysis of Equity and Interest Rate Returns in South Africa in the Context of Jump Diffusion Processes

Wilson Tsakane Mongwe

A dissertation submitted to the Department of Actuarial Science,
Faculty of Commerce, at the University of the Cape Town, in partial
fulfilment of the requirements for the degree of Master of Philosophy.

September 14, 2015

*Master of Philosophy specializing in Mathematical Finance,
University of the Cape Town,
Cape Town.*



The copyright of this thesis vests in the author. No quotation from it or information derived from it is to be published without full acknowledgement of the source. The thesis is to be used for private study or non-commercial research purposes only.

Published by the University of Cape Town (UCT) in terms of the non-exclusive license granted to UCT by the author.

Declaration

I, Wilson Tsakane Mongwe (MNGWIL004), declare that this dissertation is my own, unaided work. It is being submitted for the Degree of Master of Philosophy in the University of the Cape Town. It has not been submitted before for any degree or examination in any other University.

Signed by candidate

Abstract

Over the last few decades, there has been vast interest in the modelling of asset returns using jump diffusion processes. This was in part as a result of the realisation that the standard diffusion processes, which do not allow for jumps, were not able to capture the stylized facts that return distributions are leptokurtic and have heavy tails. Although jump diffusion models have been identified as being useful to capture these stylized facts, there has not been consensus as to how these jump diffusion models should be calibrated. This dissertation tackles this calibration issue by considering the basic jump diffusion model of Merton (1976) applied to South African equity and interest rate market data. As there is little access to frequently updated volatility surfaces and option price data in South Africa, the calibration methods that are used in this dissertation are those that require historical returns data only. The methods used are the standard Maximum Likelihood Estimation (MLE) approach, the likelihood profiling method of Honore (1998), the Method of Moments Estimation (MME) technique and the Expectation Maximisation (EM) algorithm. The calibration methods are applied to both simulated and empirical returns data. The simulation and empirical studies show that the standard MLE approach sometimes produces estimators which are not reliable as they are biased and have wide confidence intervals. This is because the likelihood function required for the implementation of the MLE method is not bounded. In the simulation studies, the MME approach produces results which do not make statistical sense, such as negative variances, and is thus not used in the empirical analysis. The best method for calibrating the jump diffusion model to the empirical data is chosen by comparing the width of the bootstrap confidence intervals of the estimators produced by the methods. The empirical analysis indicates that the best method for calibrating equity returns is the EM approach and the best method for calibrating interest rate returns is the likelihood profiling method of Honore (1998).

Acknowledgements

First and foremost, I would like to acknowledge the support I received from Rinaye Ravhuhali and Chanty Mathebula during the course of writing this dissertation. Without your valuable input, the completion of this project would not have been possible.

Secondly, I would like to thank the input and help from Melusi Mavuso and Dr Miguel Lacerda. Your advise during the writing of this dissertation was priceless, I will forever remain thankful to both of you. I am also grateful to Hlayisi Baloyi, Delicia Shabangu, Rhulani Ntimbani, Ndaedzo Madume, Gavaza Mukansi, Kendall Carolissen and Cecelia Augustine for all the feedback that you gave me.

I would also like to thank the National Research Fund (NRF) for funding this dissertation. Without your funding, the work presented in this dissertation would not have been completed.

I would also like to thank the support that I have received from my family throughout this degree. Without you, I would not be where I am today.

Last but not least, I would like to thank Mr Dylan Flint for all the advice and valuable input that you gave me during the writing of this dissertation.

Contents

1. Introduction	1
2. Methodology	3
2.1 Data and Calibration	3
2.2 Model Setting	3
2.3 MLE Approach	4
2.4 Likelihood Profiling Approach	7
2.5 EM Approach	8
2.6 MME Approach	9
2.7 Jump Detection Test	10
3. Monte-Carlo Simulations	13
3.1 Simulation background	13
3.2 Simulation Results	14
3.2.1 Large Jumps	14
3.2.2 Small Jumps	20
3.2.3 Varying the Number of Simulations and Years	22
3.2.4 Run Time and Code Considerations	23
4. Empirical Implementation	25
4.1 Jump Test Results	25
4.2 Model Calibration Results	26
4.2.1 Equity Returns	29
4.2.2 Interest Rate Returns	32
4.2.3 Comparing the Equity and Interest Rate Returns	33
5. Conclusions and Recommendations	34
Bibliography	36
A. Proving that the Likelihood function is not bounded	38
A.1 Transition Density of the Jump Diffusion Model	38
A.2 Deriving an Expression for $h(\mu, \sigma)$	39
A.3 Taking the Limit into the Infinite Sum	39

B. Simulation Results	41
B.0.1 Large Jumps: $\{\mu = 0.05, \sigma = 0.2, \lambda\tau = 0.3, \mu_{jump} = 0.05, \sigma_{jump} = 0.07\}$	41
B.0.2 Small Jumps: $\{\mu = 0.05, \sigma = 0.2, \lambda\tau = 0.07, \mu_{jump} = 0.005, \sigma_{jump} = 0.03\}$	41

List of Figures

3.1	Histograms of the estimators of μ using different methods: (a) Estimation using MLEs (b) Estimation using method of profiling the likelihood (c) Estimation using the EM algorithm and (d) Estimation using the MME method. The blue lines indicate the true value of μ . In this case, the true value of μ was set to be 0.05.	15
3.2	Histograms of the estimators of σ using different methods: (a) Estimation using MLEs (b) Estimation using method of profiling the likelihood (c) Estimation using the EM algorithm and (d) Estimation using the MME method. The blue lines indicate the true value of σ . In this case, the true value of σ was set to be 0.2.	16
3.3	Histograms of the estimators of $\lambda\tau$ using different methods: (a) Estimation using MLEs (b) Estimation using method of profiling the likelihood (c) Estimation using the EM algorithm and (d) Estimation using the MME method. The blue lines indicate the true value of $\lambda\tau$. In this case, the true value of $\lambda\tau$ was set to be 0.3.	17
3.4	Histograms of the estimators of μ_{jump} using different methods: (a) Estimation using MLEs (b) Estimation using method of profiling the likelihood and (c) Estimation using the EM algorithm. The blue lines indicate the true value of μ_{jump} . In this case, the true value of μ_{jump} was set to be 0.05.	18
3.5	Histograms of the estimators of σ_{jump} using different methods: (a) Estimation using MLEs (b) Estimation using method of profiling the likelihood (c) Estimation using the EM algorithm and (d) Estimation using the MME method. The blue lines indicate the true value of σ_{jump} . In this case, the true value of σ_{jump} was set to be 0.07. . . .	19
4.1	Equity log return series with the detected jumps indicated by squares: (a) MTN (b) ANG (c) Top40.	27
4.2	Interest rate log return series with the detected jumps indicated by squares: (a) 6 month Jibar (b) 5 year swap (c) 15 year swap.	28

List of Tables

3.1	Simulation results for the estimation of μ (true value = 0.05) using different estimation methods.	14
3.2	Simulation results for the estimation of σ (true value = 0.2) using different estimation methods. For the MME method, σ^2 is shown in the table instead of σ because σ^2 can take on negative values.	14
3.3	Simulation results for the estimation of $\lambda\tau$ (true value = 0.3) using different estimation methods.	20
3.4	Simulation results for the estimation of μ_{jump} (true value = 0.05) using different estimation methods.	20
3.5	Simulation results for the estimation of σ_{jump} (true value = 0.07) using different estimation methods.	21
3.6	Simulation results for the estimation of μ (true value = 0.05) using different estimation methods.	21
3.7	Simulation results for the estimation of σ (true value = 0.2) using different estimation methods.	22
3.8	Simulation results for the estimation of $\lambda\tau$ (true value = 0.07) using different estimation methods.	22
3.9	Simulation results for the estimation of μ_{jump} (true value = 0.005) using different estimation methods.	23
3.10	Simulation results for the estimation of σ_{jump} (true value = 0.03) using different estimation methods.	23
4.1	Jump diffusion parameter estimates for the MTN log returns using different estimation methods. The 95% bootstrap interval is shown in the parenthesis.	29
4.2	Jump diffusion parameter estimates for the ANG log returns using different estimation methods. The 95% bootstrap interval is shown in the parenthesis.	29
4.3	Jump diffusion parameter estimates for the Top40 log returns using different estimation methods. The 95% bootstrap interval is shown in the parenthesis.	30
4.4	Jump diffusion parameter estimates for the 6 month Jibar log returns using different estimation methods. The 95% bootstrap interval is shown in the parenthesis.	30
4.5	Jump diffusion parameter estimates for the 5 year swap log returns using the different estimation methods. The 95% bootstrap interval is shown in the parenthesis.	31

4.6	Jump diffusion parameter estimates for the 15 year swap log returns using the different estimation methods. The 95% bootstrap interval is shown in the parenthesis.	31
B.1	Simulation results for the estimation of μ (true value = 0.05) using different estimation methods.	41
B.2	Simulation results for the estimation of σ (true value = 0.2) using different estimation methods.	41
B.3	Simulation results for the estimation of $\lambda\tau$ (true value = 0.3) using different estimation methods.	42
B.4	Simulation results for the estimation of μ_{jump} (true value = 0.05) using different estimation methods.	42
B.5	Simulation results for the estimation of σ_{jump} (true value = 0.07) using different estimation methods.	42
B.6	Simulation results for the estimation of μ (true value = 0.05) using different estimation methods.	42
B.7	Simulation results for the estimation of σ (true value = 0.2) using different estimation methods.	43
B.8	Simulation results for the estimation of $\lambda\tau$ (true value = 0.07) using different estimation methods.	43
B.9	Simulation results for the estimation of μ_{jump} (true value = 0.005) using different estimation methods.	43
B.10	Simulation results for the estimation of σ_{jump} (true value = 0.03) using different estimation methods.	43

Chapter 1

Introduction

Since the seminal work of Press (1967) and Merton (1976), there has been vast interest in modelling returns using jump diffusion models. Jump diffusion processes are appealing to use to model asset returns, instead of the traditional diffusion models such as Geometric Brownian Motion, because they produce returns which are leptokurtic. The leptokurtic nature of returns is a stylized fact which has been noted by Cont (2001). The more recent jump diffusion models that improve on the model of Merton (1976) include that of Kou and Wang (2004), amongst others.

Other models that capture this distributional characteristic of returns include stochastic volatility models, such as the Heston (1993) model, and pure jump models such as the generalized Hyperbolic motion of Eberlein and Keller (1995) which assumes that log returns follow a Hyperbolic distribution. Although these models capture the leptokurtic nature of returns, they often present computational problems as they are sometimes not able to produce analytic solutions for situations of interest in practice, such as pricing options (Kou and Wang, 2004).

Although many jump diffusion models have been developed, there has not been consensus as to how such models should be calibrated or estimated. The most preferred way of estimating parameters is the Maximum Likelihood Estimation (MLE) approach because of the asymptotic characteristics of the estimator. However, in the jump diffusion case, the MLE approach produces unsatisfactory results because the likelihood is not bounded (Honore, 1998). Some of the other calibration techniques that have been considered in practice, which use historical returns data only, are the Method of Moments Estimation (MME) technique of Press (1967), simulated method of moments approach of Duffie and Singleton (1993), likelihood profiling method of Honore (1998) and the Expectation Maximisation (EM) algorithm. Other calibration methods use option prices and volatility surfaces, and include methods that obtain model parameters that minimise the sum of squared errors between the market and model option prices (West, 2005).

The methods that only use historical returns data are simpler to implement than

those that need option prices and volatility surfaces because historical returns data is easily accessible. The methods that require option prices and volatility surfaces are more difficult to implement in a South African context as access to frequently updated data on volatility surfaces and option prices is limited (West, 2005). Given the data limitations in the South African financial market, this dissertation focuses on calibration methods that only require historical price data. For this reason, the methods that are considered in this dissertation are the standard MLE approach, the likelihood profiling method of Honore (1998), the EM algorithm, as well as the MME technique of Press (1967).

Prior to these estimation techniques being applied, one needs to establish whether the premise that there are jumps in the data is valid. Examples of tests for the presence of jumps that have been developed in practice include the bi-power variations test of Barndorff-Nielsen and Shepard (2006). This jump test compares the estimate of variance that is not robust to the presence of jumps, called *realised variance*, with an estimate of variance that is robust to the presence of jumps, called *bi-power variation*. This test was improved by Lee and Mykland (2008) who tested for the presence of jumps at each observed value of the process while taking into account the volatility of the process at the time the observation was made. The test of Lee and Mykland (2008) has the added advantage that it not only indicates whether or not jumps have occurred, but also gives information as to what time the jumps occurred and their size.

The model that is used in this dissertation is a one-dimensional jump diffusion process with jumps only in returns and without stochastic volatility. For studies which incorporate stochastic volatility and jumps in volatility, refer to the paper by Andersen (2000).

The rest of this dissertation is structured as follows: Chapter 2 looks into the methodology that has been followed in this dissertation, Chapter 3 provides the Monte-Carlo simulation methodology as well as the results of the simulations, Chapter 4 focuses on the empirical implementation of the different calibration methods and finally Chapter 5 concludes the dissertation and recommendations for future work are made.

Chapter 2

Methodology

2.1 Data and Calibration

As mentioned in Chapter 1, there is limited access to frequently published data on option and volatility surfaces in South Africa. For example, the volatility surface produced by the South African Futures Exchange (SAFEX) is updated twice a month and the data available to reproduce this surface is illiquid (West, 2005). For this reason, this dissertation focuses on calibration methods which do not rely on option price data and volatility surfaces. Specifically, the calibration is done by using the underlyings historical price data only.

The required data was collected from the Bloomberg terminal at the University of Cape Town Oppenheimer library. The data consists of share price data and interest rate data. The share price data is composed of MTN, Anglo (ANG) and the Top40 index returns. The interest rate data is made up of the 6 month Jibar rate, 5 year and 15 year swap rates. The span of all the data is a period of 14 years from January 2000 to January 2014. It is important to note that the returns for all the instruments used in this dissertation were calculated as log returns as per the model specification in Section 2.2.

2.2 Model Setting

The model that is analysed in this dissertation is a one-dimensional Markov process $\{S_t, t \geq 0\}$. This Markov process is characterised by the following Stochastic Differential Equation (SDE):

$$d \ln S_t = \left(\mu - \frac{1}{2} \sigma^2 \right) dt + \sigma dB_t + d \left(\sum_{i=1}^{N_t} Y_i \right) \quad (2.1)$$

where μ is the drift coefficient, σ is the diffusion coefficient, $\{B_t, t \geq 0\}$ is a standard Brownian motion process, Y_i is the random size of the i th jump and $\{N_t, t \geq 0\}$ is a

Poisson process with intensity λ . Furthermore, the jump sizes Y_i are assumed to be independent and identically distributed random variables, and independent of both $\{N_t, t \geq 0\}$ and $\{B_t, t \geq 0\}$. In addition, it is assumed that $Y_i \sim N(\mu_{jump}, \sigma_{jump}^2)$. This model is the jump diffusion model used by Press (1967) in the context of finding a model that accurately describes the underlying return process and Merton (1976) in the context of pricing options when the returns are discontinuous.

The parameter vector that is going to be estimated is $(\mu, \sigma, \lambda, \mu_{jump}, \sigma_{jump})$. In this dissertation, the parameters will be estimated using the Maximum Likelihood Estimation (MLE) method, the likelihood profiling method of Honore (1998), the Expectation Maximization (EM) algorithm and the Method of Moments Estimation (MME) method. The remaining sections of this chapter look at these different estimation methods.

2.3 MLE Approach

When it comes to the estimation of parameters of statistical models, the MLE approach is often preferred. This approach is preferred because it produces estimators that are asymptotically the best in that they attain the Cramèr Rao lower bound, are unbiased and are asymptotically normally distributed (Serlin, 2007). However, these properties are only guaranteed in the limit, thus the finite sample behaviour of this estimator may not have the above desired properties.

The SDE in equation (2.1) implies a particular transition density. In order to use the MLE approach, the transition density of the jump diffusion process has to be determined. As shown in Appendix A.1, the transition density of the returns of the jump diffusion process in equation (2.1) is given as:

$$\mathbb{P}(\ln S(t + \tau) = w | \ln S(t) = x) = \sum_{n=0}^{\infty} \frac{e^{-\lambda\tau} (\lambda\tau)^n}{n!} \frac{\phi\left(\frac{w-x-(\mu\tau+n\mu_{jump})}{\sqrt{\sigma^2\tau+n\sigma_{jump}^2}}\right)}{\sqrt{\sigma^2\tau+n\sigma_{jump}^2}} \quad (2.2)$$

where w and x are real numbers, τ is the time difference between $S(t + \tau)$ and $S(t)$ and ϕ is the probability density function(pdf) of a standard normal random variable. Thus the likelihood function that needs to be maximised is as follows:

$$L(\mu, \sigma, \lambda, \mu_{jump}, \sigma_{jump}) = \prod_{k=1}^N \mathbb{P}(\ln S(t + \tau) = w_k | \ln S(t) = x_k) \quad (2.3)$$

where N is the sample size. As mentioned in Section 2.2, the parameters of the jump diffusion model in equation (2.1) that need to be estimated are $\mu, \sigma, \lambda\tau, \mu_{jump}$ and σ_{jump} .

A careful examination of the transition density in equation (2.2) shows that it is a mixture of normally distributed random variables with the mixing weights being probabilities from a Poisson distribution. Given that the transition density in equation (2.2) is an infinite mixture of normal distributions, it can be shown that the likelihood function in equation (2.3) has many points where it is not defined. The proof that a product of a finite mixture of normals has an infinite number of points where it is not defined can be found in the paper by Kiefer (1978). Honore (1998) develops the idea in Kiefer (1978) and argues that the likelihood of the jump diffusion model of Merton (1976), which is an infinite mixture of normals, also has many singularities. In this dissertation, a proof of why the model by Merton (1976) in equation (2.1) is not bounded is provided.

This means that if the MLE approach is used to estimate the parameters of the model in equation (2.1), the results may be unreliable. These unreliable results have been observed by several authors (see for example, Ball and Torous, 1983, and Bakers, 1981).

To see why the likelihood in equation (2.3) is not bounded, consider re-writing the likelihood in the following form:

$$\begin{aligned} L(\mu, \sigma, \lambda, \mu_{jump}, \sigma_{jump}) &= \prod_{k=1}^N \mathbb{P}(\ln S(t + \tau) = w_k | \ln S(t) = x_k) \\ &= \prod_{k=1}^N \sum_{n=0}^{\infty} \frac{e^{-\lambda\tau} (\lambda\tau)^n}{n!} \frac{\phi\left(\frac{w-x-(\mu\tau+n\mu_{jump})}{\sqrt{\sigma^2\tau+n\sigma_{jump}^2}}\right)}{\sqrt{\sigma^2\tau+n\sigma_{jump}^2}} \\ &= \prod_{k=1}^N \sum_{n=0}^{\infty} \frac{g_n}{\sqrt{\sigma^2\tau+n\sigma_{jump}^2}} \phi\left(\frac{y_k - (\mu\tau + n\mu_{jump})}{\sqrt{\sigma^2\tau + n\sigma_{jump}^2}}\right) \end{aligned}$$

where $g_n = \frac{e^{-\lambda\tau} (\lambda\tau)^n}{n!}$ and $y_k = w_k - x_k$.

The aim is to show that one can create as many spikes as one likes in the above likelihood function. Note that the likelihood function is now seen as a function of $\mu, \sigma, \lambda, \mu_{jump}$ and σ_{jump} . Thus all that needs to be done is to show that this function is unbounded as the parameters tend to certain values.

Now, consider taking the limit as $\mu \rightarrow \frac{y_k}{\tau}$ and $\sigma \rightarrow 0$. Without loss of generality, set $k = 1$ such that $\mu \rightarrow \frac{y_1}{\tau}$. In addition, let the other parameters be $\lambda = a_2, \mu_{jump} = a_3$ and $\sigma_{jump} = a_4$, where a_2 and a_4 are real constants greater than zero and a_3 is real constant.

This means that:

$$\begin{aligned} & \lim_{\sigma \rightarrow 0} \lim_{\mu \rightarrow \frac{y_1}{\tau}} L(\mu, \sigma, \lambda = a_2, \mu_{jump} = a_3, \sigma_{jump} = a_4) \\ &= \lim_{\sigma \rightarrow 0} \lim_{\mu \rightarrow \frac{y_1}{\tau}} \prod_{k=1}^N \sum_{n=0}^{\infty} \frac{g_n}{\sqrt{\sigma^2 \tau + na_4^2}} \phi \left(\frac{y_k - (\mu\tau + na_3)}{\sqrt{\sigma^2 \tau + na_4^2}} \right). \end{aligned}$$

To avoid taking the limit into the product, because the product limit law does not directly apply in this case, the proof follows by first defining a new function, say $h(\mu, \sigma)$, which is less than or equal to $L(\mu, \sigma)$. The proof will then proceed to show that $\lim_{\sigma \rightarrow 0} \lim_{\mu \rightarrow \frac{y_1}{\tau}} h(\mu, \sigma) = \infty$. This would then imply that $\lim_{\sigma \rightarrow 0} \lim_{\mu \rightarrow \frac{y_1}{\tau}} L(\mu, \sigma) = \infty$ as $L(\mu, \sigma) \geq h(\mu, \sigma)$. As shown in Appendix A.2, a possible expression for $h(\mu, \sigma)$ is given as:

$$h(\mu, \sigma) = \prod_{k=1}^N \sum_{n=0}^{\infty} \frac{g_n}{\sqrt{\sigma^2 \tau + na_4^2}} \phi \left(\frac{y_k - (\mu\tau + na_3)}{\sqrt{na_4^2}} \right).$$

Now the limit $\lim_{\sigma \rightarrow 0} \lim_{\mu \rightarrow \frac{y_1}{\tau}} h(\mu, \sigma)$ is then evaluated as:

$$\begin{aligned} &= \lim_{\sigma \rightarrow 0} \lim_{\mu \rightarrow \frac{y_1}{\tau}} h(\mu, \sigma) \\ &= \lim_{\sigma \rightarrow 0} \lim_{\mu \rightarrow \frac{y_1}{\tau}} \prod_{k=1}^N \sum_{n=0}^{\infty} \frac{g_n}{\sqrt{\sigma^2 \tau + na_4^2}} \phi \left(\frac{y_k - (\mu\tau + na_3)}{\sqrt{na_4^2}} \right) \\ &= \lim_{\sigma \rightarrow 0} \lim_{\mu \rightarrow \frac{y_1}{\tau}} \prod_{k=1}^N \sum_{n=0}^{\infty} \frac{g_n}{\sqrt{\sigma^2 \tau + na_4^2}} \phi \left(\frac{y_k - \mu\tau}{\sqrt{na_4^2}} - \sqrt{n} \frac{a_3}{a_4} \right) \end{aligned}$$

Consider now the individual terms in the above product:

For $k = 1$

$$\begin{aligned} &= \lim_{\sigma \rightarrow 0} \lim_{\mu \rightarrow \frac{y_1}{\tau}} \sum_{n=0}^{\infty} \frac{g_n}{\sqrt{\sigma^2 \tau + na_4^2}} \phi \left(\frac{y_1 - \mu\tau}{\sqrt{na_4^2}} - \sqrt{n} \frac{a_3}{a_4} \right) \\ &= \lim_{\sigma \rightarrow 0} \sum_{n=0}^{\infty} \frac{g_n}{\sqrt{\sigma^2 \tau + na_4^2}} \phi \left(-\sqrt{n} \frac{a_3}{a_4} \right) \\ &\quad \text{(Taking limit inside the infinite sum [See Appendix A.3])} \\ &= \lim_{\sigma \rightarrow 0} \frac{g_0}{\sqrt{\sigma^2 \tau}} \phi(0) + \lim_{\sigma \rightarrow 0} \sum_{n=1}^{\infty} \frac{g_n}{\sqrt{\sigma^2 \tau + na_4^2}} \phi \left(-\sqrt{n} \frac{a_3}{a_4} \right) \\ &= \infty + \sum_{n=1}^{\infty} \frac{g_n}{\sqrt{na_4^2}} \phi \left(-\sqrt{n} \frac{a_3}{a_4} \right) \quad \text{(Taking the limit inside the infinite sum)} \\ &\geq \infty \quad \text{(Terms in the infinite sum are all positive.)} \end{aligned}$$

For $k \geq 2$

$$\begin{aligned}
&= \lim_{\sigma \rightarrow 0} \lim_{\mu \rightarrow \frac{y_1}{\tau}} \sum_{n=0}^{\infty} \frac{g_n}{\sqrt{\sigma^2 \tau + n a_4^2}} \phi \left(\frac{y_k - \mu \tau}{\sqrt{n a_4^2}} - \sqrt{n} \frac{a_3}{a_4} \right) \\
&= \lim_{\sigma \rightarrow 0} \sum_{n=0}^{\infty} \frac{g_n}{\sqrt{\sigma^2 \tau + n a_4^2}} \phi \left(\frac{c_k}{\sqrt{n a_4^2}} - \sqrt{n} \frac{a_3}{a_4} \right)
\end{aligned}$$

(Taking limit into the infinite sum and where c_k is some real number.)

$$\begin{aligned}
&= \lim_{\sigma \rightarrow 0} \frac{g_0}{\sqrt{\sigma^2 \tau}} \phi(\pm\infty) + \lim_{\sigma \rightarrow 0} \sum_{n=1}^{\infty} \frac{g_n}{\sqrt{\sigma^2 \tau + n a_4^2}} \phi \left(\frac{c_k}{\sqrt{n a_4^2}} - \sqrt{n} \frac{a_3}{a_4} \right) \\
&= 0 + \sum_{n=1}^{\infty} \frac{g_n}{\sqrt{n a_4^2}} \phi(m_k)
\end{aligned}$$

(Taking the limit inside the sum and where m_k is some real number)

> 0 (Terms in the sum are all positive.)

Given that for $k = 1$ the term is infinity and for $k \geq 2$ the terms are all positive, it means that $\lim_{\sigma \rightarrow 0} \lim_{\mu \rightarrow \frac{y_1}{\tau}} h(\mu, \sigma) = \infty$, which implies that

$$\lim_{\sigma \rightarrow 0} \lim_{\mu \rightarrow \frac{y_1}{\tau}} L(\mu, \sigma) = \infty.$$

The above result shows that there exists a singularity at the point $(y_1, \mu = y_1/\tau, \sigma = 0, \lambda = a_1, \mu_{jump} = a_2, \sigma_{jump} = a_3)$. Given that the data point y_1 was chosen arbitrarily, it thus follows that a singularity can be created at every observation y_k , for $1 \leq k \leq N$.

The arguments above suggest that the likelihood is unbounded when (a) one of the variances, either σ^2 or σ_{jump}^2 , is zero, and the other one is not and (b) when $\mu\tau$ equals a data point. Honore (1998) suggested two solutions to this problem. One is through profiling the likelihood function, and the other is through using the EM algorithm. In Section 2.4, the profiling method is examined in more detail and in Section 2.5 the EM approach is discussed.

2.4 Likelihood Profiling Approach

The method of profiling the likelihood removes the singularity problem by avoiding evaluation of the likelihood at points where it is not defined. Honore (1998) does this by setting the jump variance σ_{jump} to be a multiple of the diffusion variance σ . That is, he ensures that condition (a) mentioned in Section 2.3 above does not hold by setting $\sigma_{jump} = \sqrt{m}\sigma$, where m is a positive constant. The range of values for m

is chosen such that it is less than 1 but greater than zero. This is because empirical studies suggest that σ_{jump} is usually a fraction of σ , see for example Honore (1998).

What this restriction does is that if σ equals 0, then σ_{jump} must equal zero, and if σ does not equal 0, then σ_{jump} will not equal zero as well. This means that it will be impossible for the likelihood function to be evaluated at the points where it is not defined.

The likelihood function in equation (2.3) now becomes

$$L_m(\mu, \sigma, \lambda, \mu_{jump}, \sqrt{m}\sigma),$$

for a fixed m . The optimal value for m is the one that produces the largest value for $L_m(\mu, \sigma, \lambda, \mu_{jump}, \sqrt{m}\sigma)$. The estimators are then found by maximising $L_m(\mu, \sigma, \lambda, \mu_{jump}, \sqrt{m}\sigma)$ for this optimal value of m .

2.5 EM Approach

The other method that Honore (1998) mentions in his paper but does not implement is the EM algorithm. The EM algorithm is a method that is often used to work out approximate MLE estimates. It is often used in cases where there is missing data. Serlin (2007) states that finding MLE estimates for a mixture distribution is a case where an EM approach can yield fruitful results.

In this dissertation, the EM algorithm was not applied directly to the jump diffusion model with SDE in equation (2.1), but rather to another process that approximates the process with the SDE in equation (2.1). The approximation was used, instead of the actual process in equation (2.1), as it reduces the computational complexity of the problem.

The jump diffusion model in equation (2.1) was altered by approximating the Poisson probability weights by a Bernoulli distribution. This means that the mixing distribution contains only two normals instead of the infinite mixture in equation (2.2). The transition density of the approximation is:

$$\begin{aligned} \mathbb{P}(\ln S(t + \tau) = w | \ln S(t) = x) \\ = \frac{1 - \lambda}{\sqrt{\sigma^2 \tau}} \phi \left(\frac{w - x - \mu \tau}{\sqrt{\sigma^2 \tau}} \right) + \frac{\lambda}{\sqrt{\sigma^2 \tau + n \sigma_{jump}^2}} \phi \left(\frac{w - x - (\mu \tau + n \mu_{jump})}{\sqrt{\sigma^2 \tau + n \sigma_{jump}^2}} \right) \end{aligned}$$

where now λ is interpreted as the jump probability and not the rate at which the jumps occur. This approximation will only be good if $\lambda \tau$ is small (Honore, 1998).

Honore (1998) finds that this approximation is not good for estimating jump diffusion processes for some stocks listed in the New York Stock Exchange. This

suggests that, for those stocks, either the jumps occur at a higher frequency or that the daily interval chosen was not small enough for his particular case.

The advantage of the EM algorithm is that it can be proven that the algorithm produces consistent and normally distributed estimators (Serlin, 2007). The main disadvantage of the EM algorithm is that it sometimes exhibits slow convergence. For a detailed explanation of the EM algorithm used in this dissertation see Serlin (2007), Cont and Tankov (2003) and Cappe *et al.* (2005).

2.6 MME Approach

Although MLE based approaches described above are preferred, MME based approaches offer an alternative. The alternative MME approaches include equating infinitesimal moments, as in Bandi and Phillips (2003), and equating cumulants, as in Press (1967). In this dissertation, the focus will be on equating population cumulants to sample cumulants. Beckers (1981) provides an improvement to the method of Press (1967) and this is the method that is used in this dissertation.

Beckers (1981) suggests setting the jump mean μ_{jump} of the SDE in equation (2.1) to zero. This reduces the dimensionality of the estimation problem and ensures that the transition density is symmetrical. Setting μ_{jump} to zero means that all the odd cumulants are zero. Beckers (1981) then shows that the population cumulants for the jump diffusion model, with μ_{jump} set to zero, in equation (2.1) are:

$$\begin{aligned} K_1 &= \mu\tau \\ K_2 &= \sigma^2\tau + \lambda\tau\sigma_{jump}^2 \\ K_3 &= K_5 = 0 \\ K_4 &= 3\sigma_{jump}^4\lambda\tau \\ K_6 &= 15\sigma_{jump}^6\lambda\tau \end{aligned}$$

where K_i represents the i th population cumulant. The sample cumulant functions \bar{K}_i , for $i \geq 1$, are then:

$$\begin{aligned} \bar{K}_1 &= m_1 \\ \bar{K}_2 &= m_2 - m_1^2 \\ \bar{K}_4 &= m_4 - 3m_2^2 - 4m_1m_3 + 12m_1^2m_2 - 6m_1^4 \\ \bar{K}_6 &= m_6 - 6m_5m_1 - 15m_4m_2 + 30m_4m_1^2 - 10m_3^2 + 120m_3m_2m_1 \\ &\quad - 120m_3m_1^3 + 30m_2^3 - 270m_2^2m_1^2 + 360m_2m_1^4 - 120m_1^6. \end{aligned}$$

where $m_i = \frac{1}{T} \sum_{k=1}^T (\Delta X_i)^i$ with ΔX_i being the i th return. Equating the population cumulants with the sample cumulants and solving the resulting system of equations

we obtain that:

$$\begin{aligned}\hat{\mu} &= \bar{K}_1/\tau \\ \hat{\lambda} &= \frac{25\bar{K}_4^3}{3\bar{K}_6^2} \\ \hat{\sigma}_{jump}^2 &= \frac{\bar{K}_6}{5\bar{K}_4} \\ \hat{\sigma}^2 &= \left(\bar{K}_2 - \frac{5\bar{K}_4^2}{3\bar{K}_6} \right) / \tau.\end{aligned}$$

As can be seen from the above equations, the MME approach is attractive because it does not have computational difficulties as it is easy to implement. However, MME approaches in general have the disadvantage of producing results which do not make statistical sense. For example, it is possible to obtain negative variances using this approach (Beckers, 1981). The reason why the negative variances can occur is because there is nothing ensuring that $\bar{K}_2 \leq \frac{5\bar{K}_4^2}{3\bar{K}_6}$. Thus a negative value for $\hat{\sigma}^2$ is possible.

2.7 Jump Detection Test

Before the model in equation (2.1) can be calibrated using the above mentioned techniques, a test needs to be conducted as to whether the model is statistically justified for the given data set. The most natural approach would be to perform some sort of likelihood ratio test, but given the deficiencies of the MLE approach mentioned in Section 2.3 above, this approach may not produce satisfactory results.

The method that is used in this dissertation is the jump detection test of Lee and Mykland (2008). This test determines whether there is a jump over a time period, say from t_{i-1} to time t_i , by comparing the absolute return over that time period with the absolute returns in its neighbourhood. Given that there may appear to be jumps in the returns due to high volatility, this test takes volatility into account by standardising the absolute returns with the volatility at that point, thus stripping out the effects of volatility.

The test statistic to test if there was a jump in the return from time t_{i-1} to time t_i is

$$L(i) = \frac{\ln(S_{t_i}/S_{t_{i-1}})}{\hat{\sigma}_{t_i}}$$

where S_{t_i} is the underlings price at time t_i .

Lee and Mykland (2008) suggest using the bi-power variation method of Barndorff-Nielsen and Shepard (2006) to estimate $\hat{\sigma}_{t_i}$ as this method produces a statistically

consistent estimator of $\hat{\sigma}_{t_i}$ which is not affected by the occurrence of jumps at previous time periods. The bi-power variation formula used to calculate $\hat{\sigma}_{t_i}$ is given as

$$\hat{\sigma}_{t_i} = \sqrt{\frac{1}{k-2} \sum_{j=i-k+2}^{i-1} |\log(S_{t_j}/S_{t_{j-1}})| |\log(S_{t_{j-1}}/S_{t_{j-2}})|} \quad \text{for } i > k-2, \quad (2.4)$$

where k represents the k nearest observations to the left of the observation at time t_i . Lee and Mykland (2008) recommend using the optimal window size k for one-week, one-day, one-hour, 30-minute, 15-minute, and 5-minute data to be 7, 16, 78, 110, 156, and 270 respectively.

Before this test can be conducted, one needs to first ensure that the mean return is not statistically different from zero and that the return process is stationary. When performing this test, a further check needs to be conducted to ensure that the drift of the log-returns is zero.

It can be shown that the distribution of $|L(i)|$ under the null hypothesis of no jumps is a Gumbel distribution. Let n be the total number of observations over the time interval $[0, T]$ and \bar{A}_n be the set of all $i \in \{1, 2, 3, \dots, n\}$ where no jump occurred from t_{i-1} to t_i . Lee and Mykland (2008) use extreme value theory to show that as $\Delta t \rightarrow 0$

$$\frac{\max_{i \in \bar{A}_n} (|L(i)| - C_n)}{S_n} \xrightarrow{\text{distribution}} \xi, \quad (2.5)$$

where ξ is a Gumbel random variable with cumulative distribution function $\mathbb{P}(\xi \leq x) = e^{-e^{-x}}$. The constants in equation (2.5) are

$$C_n = \frac{(2 \log n)^{0.5}}{c} - \frac{\log \pi + \log(\log n)}{2c(2 \log(n))^{0.5}} \quad \text{and} \quad S_n = \frac{1}{c(2 \log(n))^{0.5}}$$

where $c = \frac{\sqrt{2}}{\pi}$. Thus the null hypothesis of no jumps from t_{i-1} to t_i would be rejected at a significance level of α if

$$|L(i)| > -\log(-\log(1 - \alpha))S_n + C_n.$$

The significance level that is used in this dissertation is 5% as per convention.

As mentioned in Chapter 1, the advantage of this test is that it not only provides information as to whether jumps have occurred or not, but also gives information as to over which period the jumps have occurred. The disadvantage of this test, for the purposes of this dissertation, is that the authors state that the test is most powerful at the 15 minute time frequency. However, the empirical data used in this dissertation, discussed in Section 2.1, was of a daily frequency. As noted by the simulations performed by Lee and Mykland (2008), the test performs poorly at a

daily frequency with the test only detecting 2.6% of actual jumps. This must be taken into account when interpreting the results in Section 4.1 as the number of jumps are likely to be underestimated.

In order to compare the statistical properties of the estimators discussed above, Monte-Carlo simulations were conducted. The methodology and results of the simulations are presented in Chapter 3.

Chapter 3

Monte-Carlo Simulations

3.1 Simulation background

Monte-Carlo simulations were conducted to get a handle on how the different estimators behave in finite samples. In these simulations, the jump diffusion process in equation (2.1) was simulated for different number of years, and for different number of simulations. The simulation was an Euler simulation. The code that was used to simulate this data can be found on the accompanying cd-rom.

The jump diffusion model in equation (2.1) was simulated with the parameter vector set to the following values: $(\mu = 0.05, \sigma = 0.2, \lambda\tau = 0.3, \mu_{jump} = 0.05, \sigma_{jump} = 0.07)$ and $(\mu = 0.05, \sigma = 0.2, \lambda\tau = 0.07, \mu_{jump} = 0.005, \sigma_{jump} = 0.03)$. These two sets of parameter vectors were chosen as the methods behave differently depending on whether the parameters of the jump component of the model $\{\lambda, \mu_{jump}, \sigma_{jump}\}$ are big or small. Note that when the MME approach was used, the estimator of μ_{jump} was set to zero as per the specification in Section 2.6

All of the simulations were conducted with $\tau = \frac{1}{252}$. This means that the simulations were conducted at a daily frequency. It was assumed that there are 252 business days in a year. This was done so that it is consistent with the frequency of the data that is used in the empirical application in Chapter 4. The `nlm` package in **R** was used for the optimisation. This package uses a Newton type algorithm for the optimisation. When using the MLE and profiling methods, the infinite sum in the probability density function in equation (2.2) was truncated at $N = 10$. This cut-off value was used as suggested by Ball and Torous (1983).

The results for the parameter vector $(\mu = 0.05, \sigma = 0.2, \lambda\tau = 0.3, \mu_{jump} = 0.05, \sigma_{jump} = 0.07)$, for one thousand simulations of one year, are shown in Tables 3.1 to 3.5 and Figures 3.1 to 3.5. The results for the parameter vector $(\mu = 0.05, \sigma = 0.2, \lambda\tau = 0.07, \mu_{jump} = 0.005, \sigma_{jump} = 0.03)$, for one thousand simulations over one year, are shown in Table 3.6 through Table 3.10. The other simulation results can be found on the accompanying cd-rom.

3.2 Simulation Results

Estimates of μ with true value 0.05			
Method	Mode	Lower 95%	Upper 95%
MLE	0.0446	-0.3653	0.4921
Profiling	0.0515	-0.3953	0.4708
EM	0.0234	-0.3832	0.5224
MME	3.2576	2.2446	4.4291

Tab. 3.1: Simulation results for the estimation of μ (true value = 0.05) using different estimation methods.

Estimates of σ with true value 0.2			
Method	Mode	Lower 95%	Upper 95%
MLE	0.1983	0.1765	0.2230
Profiling	0.1964	0.1751	0.2222
EM	0.1980	0.1767	0.2252
MME(σ^2)	0.0001	-0.0612	0.0709

Tab. 3.2: Simulation results for the estimation of σ (true value = 0.2) using different estimation methods. For the MME method, σ^2 is shown in the table instead of σ because σ^2 can take on negative values.

3.2.1 Large Jumps

Tables 3.1 to 3.5 show the results of the simulation studies when the parameter vector was set to $(\mu = 0.05, \sigma = 0.2, \lambda\tau = 0.3, \mu_{jump} = 0.05, \sigma_{jump} = 0.07)$. The estimation method whose results immediately stand out is the MME approach. These results show that this estimation technique produced results that do not make statistical sense such as negative values for σ^2 and very large rate parameter $\lambda\tau$ value. It can be seen from the histograms in Figures 3.1 to 3.5 that there can be some extreme values produced by the MME estimation technique. In addition, this method produced the widest confidence intervals relative to all the methods used in this dissertation.

Due to the approximation of the transition density in Section 2.5, the EM approach underestimates the rate parameter $\lambda\tau$. In this particular case, it underestimates the rate parameter by 14.6%. This calibration approach produced point estimates of μ that are very different from the other methods. The EM approach also produced estimates with the widest confidence intervals for all the parameters

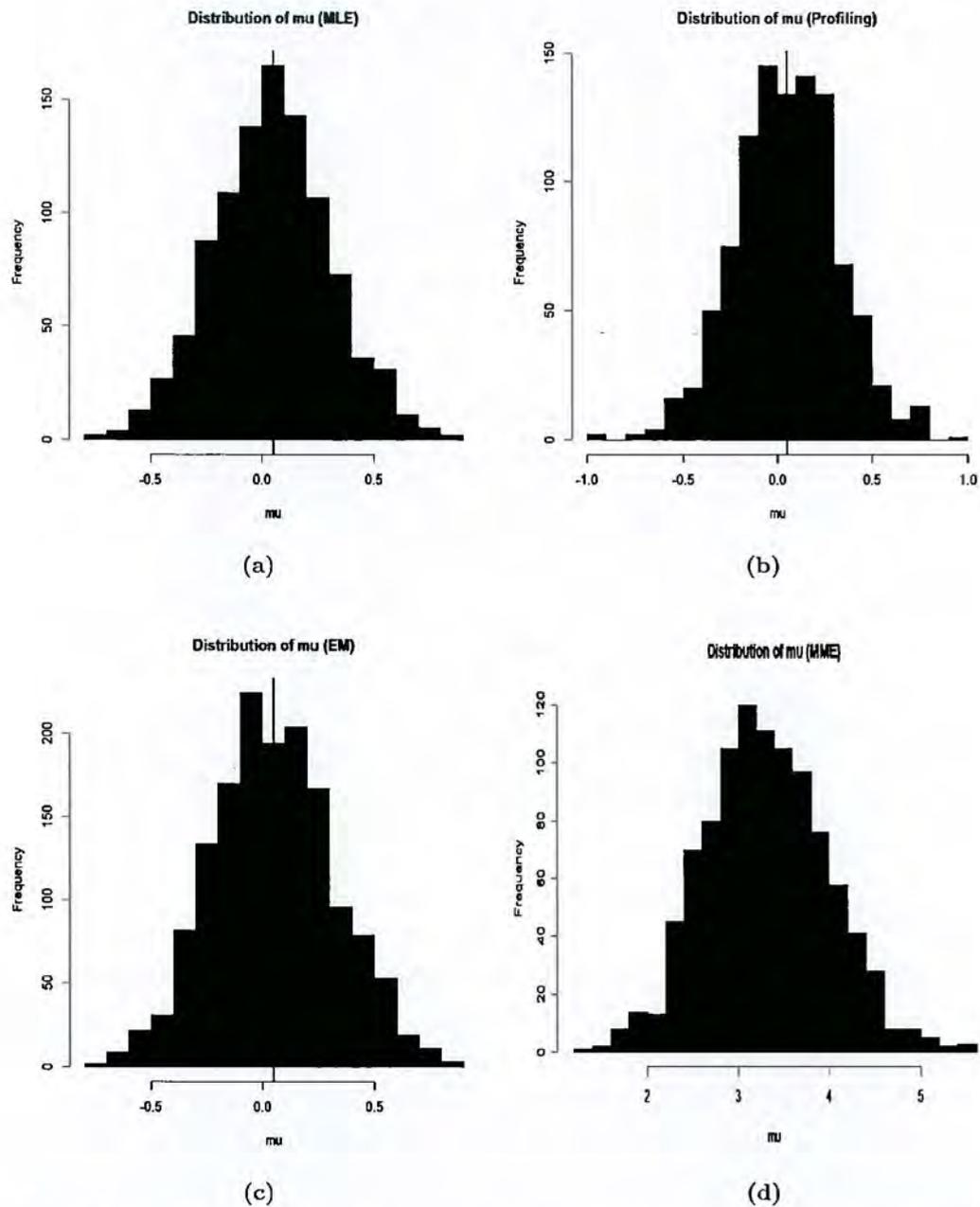


Fig. 3.1: Histograms of the estimators of μ using different methods: (a) Estimation using MLEs (b) Estimation using method of profiling the likelihood (c) Estimation using the EM algorithm and (d) Estimation using the MME method. The blue lines indicate the true value of μ . In this case, the true value of μ was set to be 0.05.

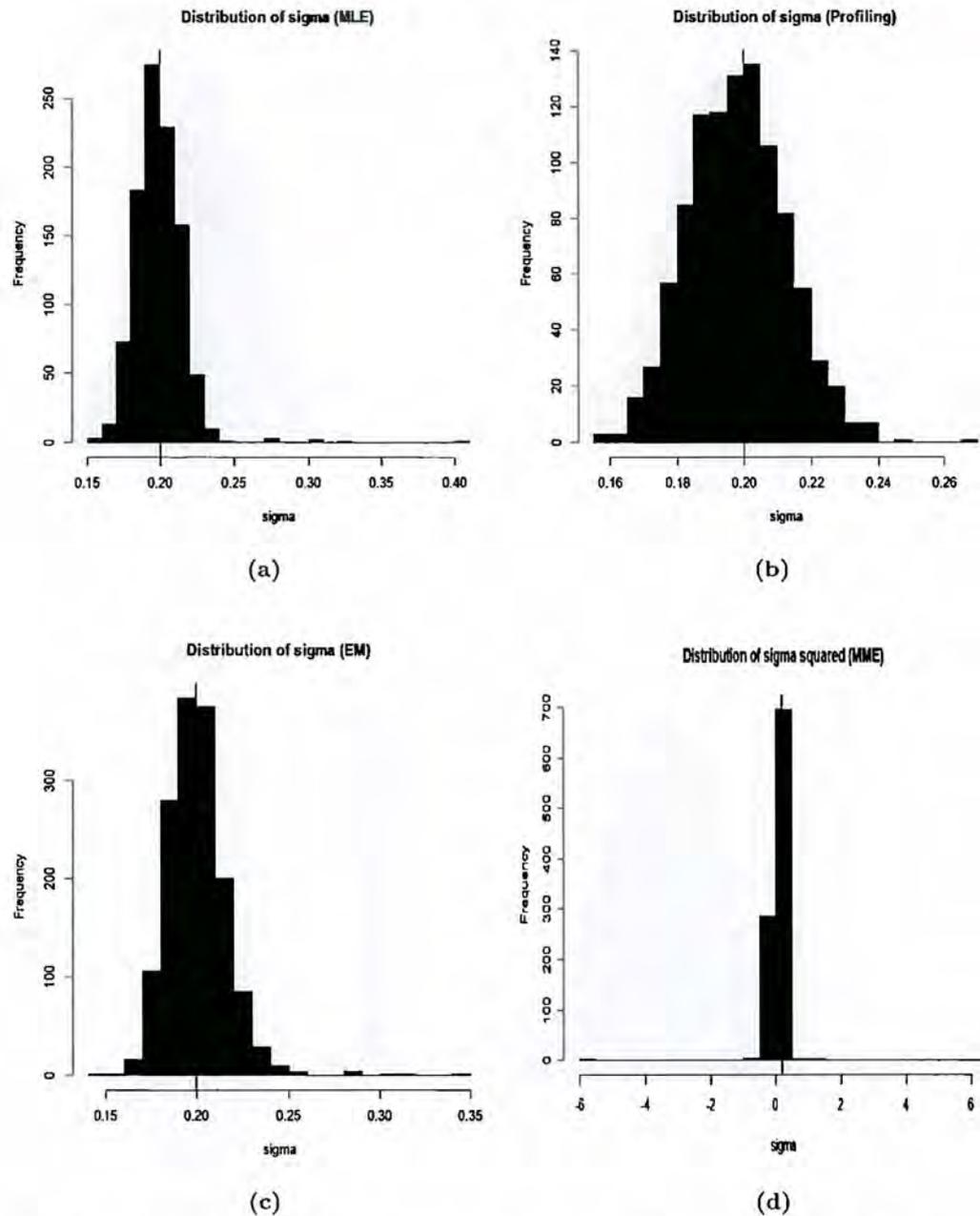


Fig. 3.2: Histograms of the estimators of σ using different methods: (a) Estimation using MLEs (b) Estimation using method of profiling the likelihood (c) Estimation using the EM algorithm and (d) Estimation using the MME method. The blue lines indicate the true value of σ . In this case, the true value of σ was set to be 0.2.

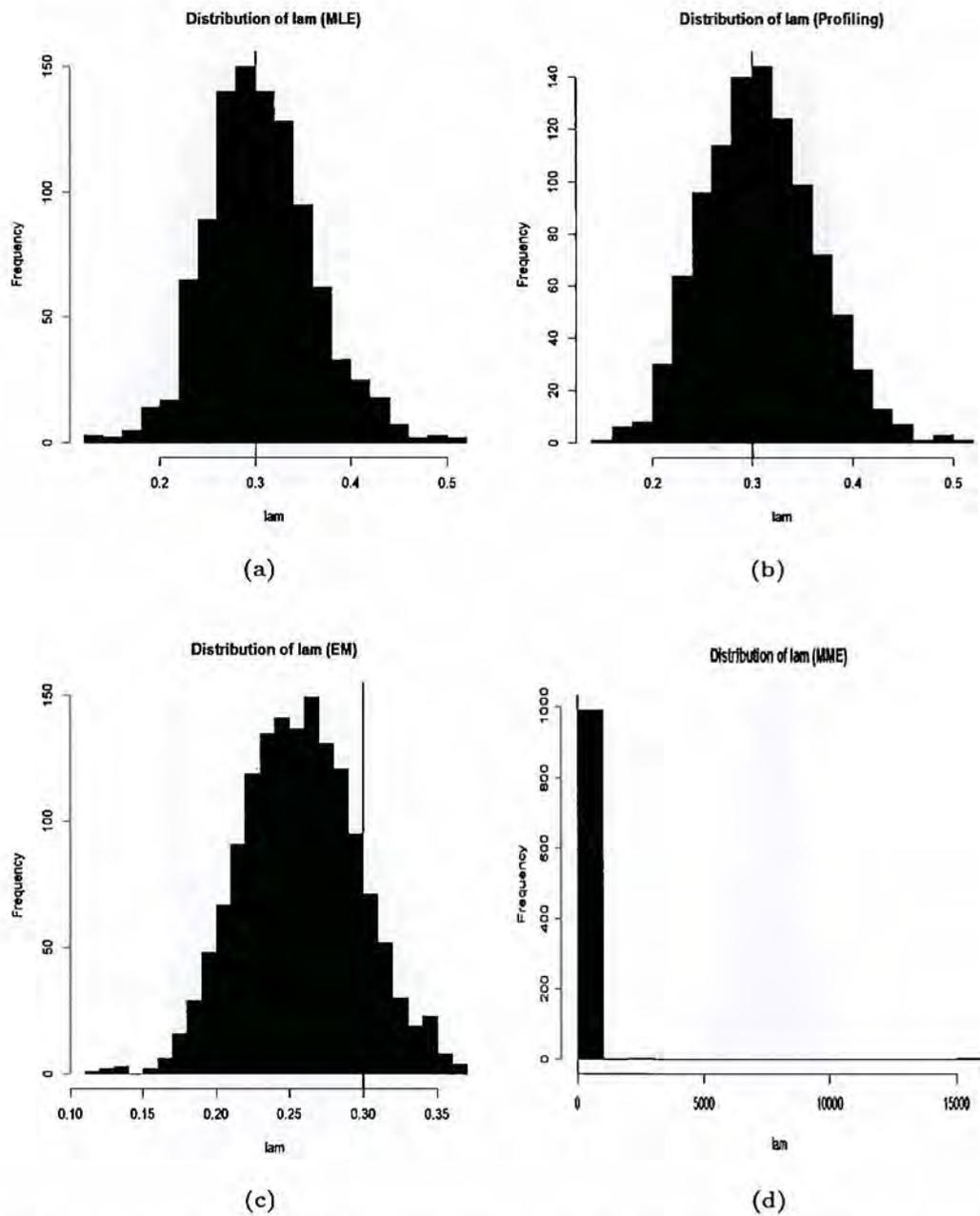


Fig. 3.3: Histograms of the estimators of $\lambda\tau$ using different methods: (a) Estimation using MLEs (b) Estimation using method of profiling the likelihood (c) Estimation using the EM algorithm and (d) Estimation using the MME method. The blue lines indicate the true value of $\lambda\tau$. In this case, the true value of $\lambda\tau$ was set to be 0.3.

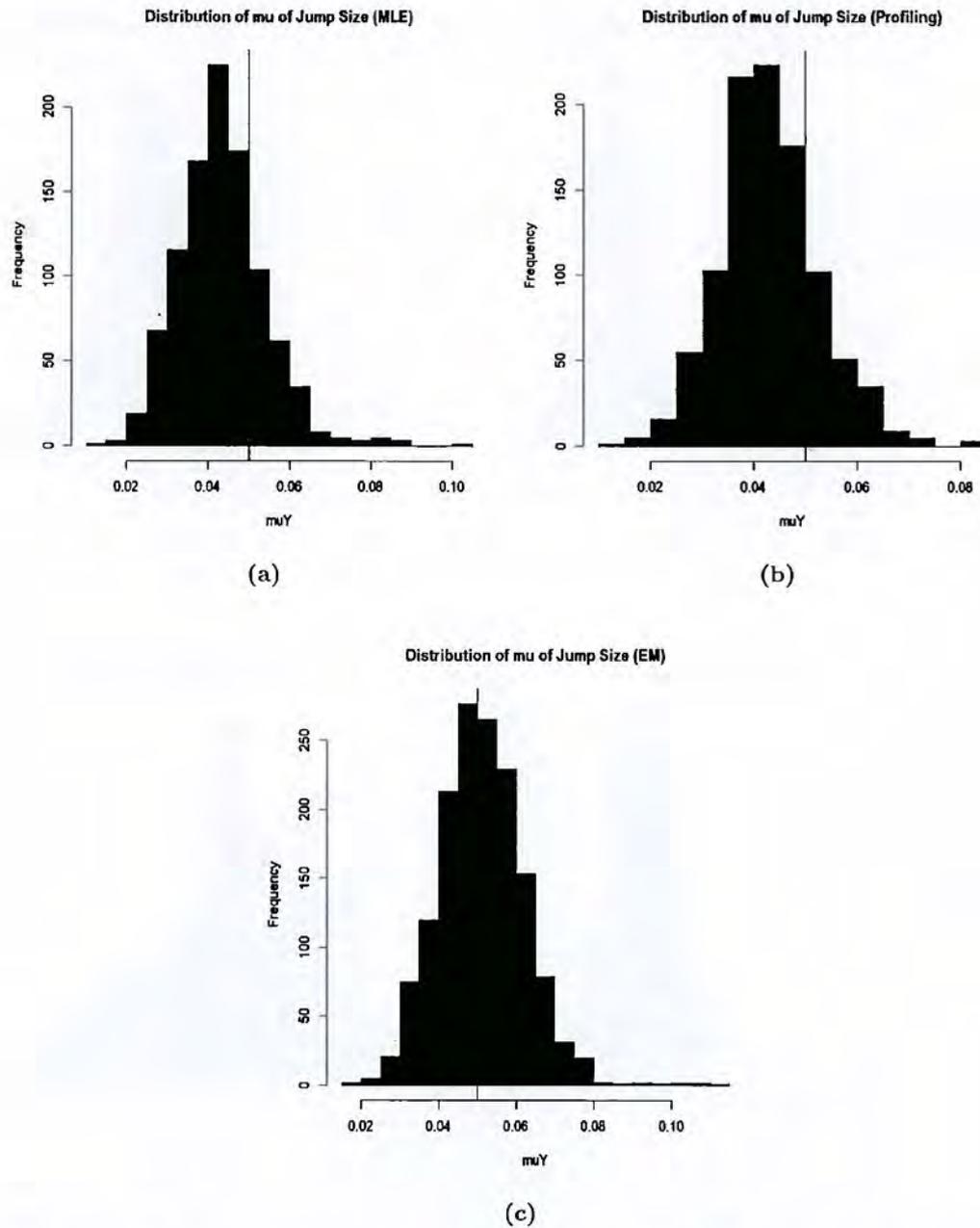


Fig. 3.4: Histograms of the estimators of μ_{jump} using different methods: (a) Estimation using MLEs (b) Estimation using method of profiling the likelihood and (c) Estimation using the EM algorithm. The blue lines indicate the true value of μ_{jump} . In this case, the true value of μ_{jump} was set to be 0.05.

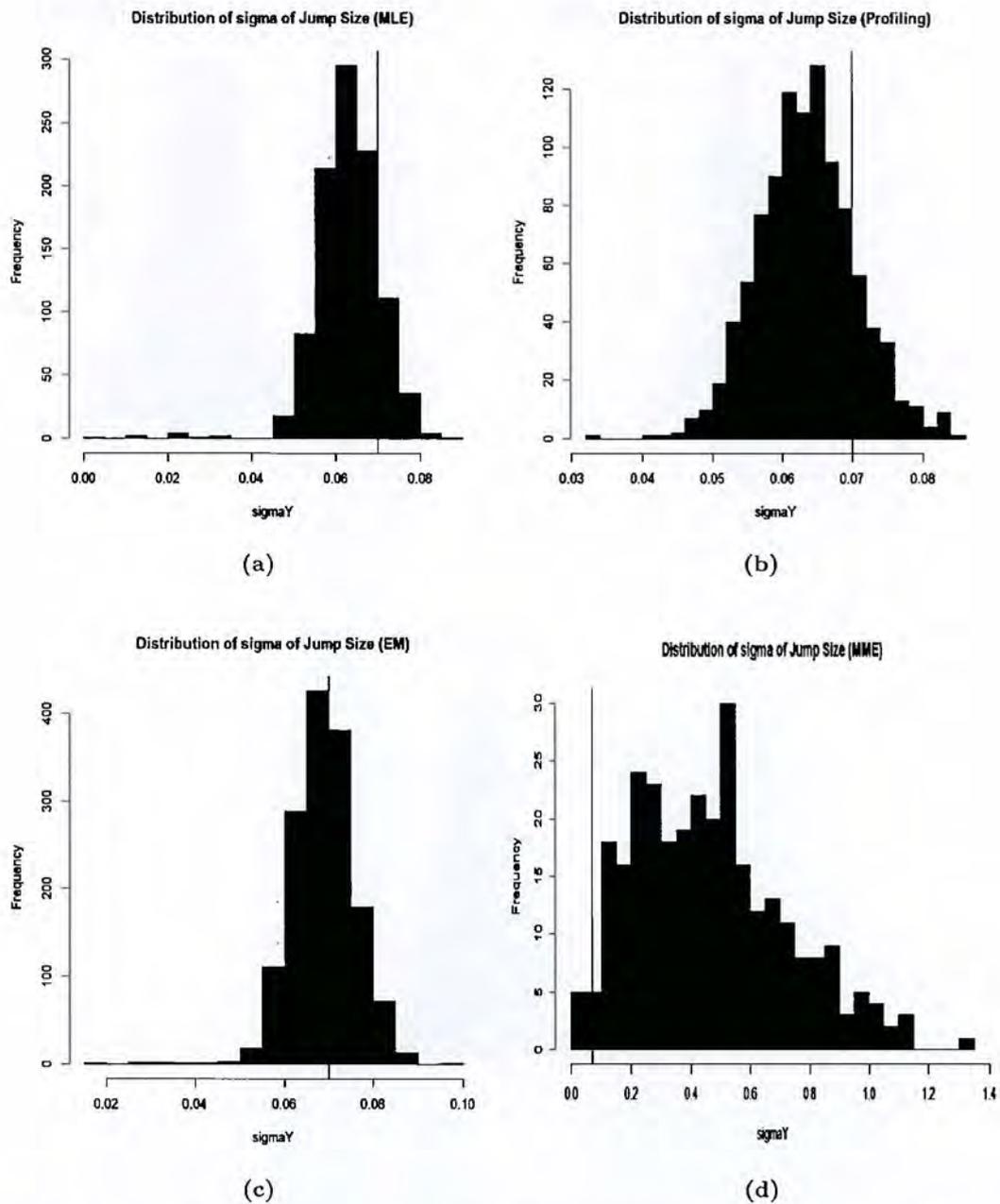


Fig. 3.5: Histograms of the estimators of σ_{jump} using different methods: (a) Estimation using MLEs (b) Estimation using method of profiling the likelihood (c) Estimation using the EM algorithm and (d) Estimation using the MME method. The blue lines indicate the true value of σ_{jump} . In this case, the true value of σ_{jump} was set to be 0.07.

Estimates of $\lambda\tau$ with true value 0.3			
Method	Mode	Lower 95%	Upper 95%
MLE	0.2954	0.2240	0.4047
Profiling	0.3063	0.2207	0.4005
EM	0.2561	0.1961	0.3232
MME	0.0164	0.0085	6.1602

Tab. 3.3: Simulation results for the estimation of $\lambda\tau$ (true value = 0.3) using different estimation methods.

Estimates of μ_{jump} with true value 0.05			
Method	Mode	Lower 95%	Upper 95%
MLE	0.0424	0.0277	0.0605
Profiling	0.0423	0.0276	0.0603
EM	0.0510	0.0334	0.0686

Tab. 3.4: Simulation results for the estimation of μ_{jump} (true value = 0.05) using different estimation methods.

compared with the MLE and profiling methods. This suggests that the EM approach does not perform well when the rate parameter is large.

For the diffusion component $\{\mu, \sigma\}$, the MLE approach has the narrowest confidence interval, while for the jump parameters $\{\lambda\tau, \mu_{jump}, \sigma_{jump}\}$ the profiling method has the narrowest confidence intervals.

The histograms in Figures 3.1 through 3.5 show that the MLE approach produced skewed empirical distributions for the parameters. In addition, these histograms show some extreme tail behaviour, showing that the MLE approach can produce estimates which are far from the expected values.

The results discussed above suggest that when the rate parameter $\lambda\tau$ value is large, the MLE and profiling method produce results that are better than the EM approach, with the MME approach being the poorest amongst the four.

3.2.2 Small Jumps

Tables 3.6 to 3.10 show the results of the simulation studies when the parameter values are set to $(\mu = 0.05, \sigma = 0.2, \lambda\tau = 0.07, \mu_{jump} = 0.005, \sigma_{jump} = 0.03)$. When the mean jump size μ_{jump} and the jump variance σ_{jump} are small (implying that the jumps are small) and the rate parameter $\lambda\tau$ is small, the MLE and profiling methods produce estimates that are not reliable over one year's worth of data.

Estimates of σ_{jump} with true value 0.07			
Method	Mode	Lower 95%	Upper 95%
MLE	0.0631	0.0518	0.0743
Profiling	0.0634	0.0526	0.0751
EM	0.0686	0.0577	0.0804
MME	0.3525	0.1168	0.9466

Tab. 3.5: Simulation results for the estimation of σ_{jump} (true value = 0.07) using different estimation methods.

Estimates of μ with true value 0.05			
Method	Mode	Lower 95%	Upper 95%
MLE	0.0871	-0.3401	0.4709
Profiling	0.0719	-0.3279	0.4346
EM	0.0767	-0.3125	0.4272
MME	0.0968	-0.2945	0.4873

Tab. 3.6: Simulation results for the estimation of μ (true value = 0.05) using different estimation methods.

The MLE approach implies that the rate parameter equals zero, while the profiling method implies that it is 77.0% smaller than the true value. However, both methods still estimate the σ well as the estimated value is close to the true value, and the confidence intervals are narrow.

The other jump parameters, being $\{\mu_{jump}, \sigma_{jump}\}$, are also poorly estimated by these methods. The MLE underestimated μ_{jump} and σ_{jump} by 39.7% and 46.3% respectively. The profiling method underestimated the same parameters by 33.6% and 16.7% respectively.

The EM approach produced the more reliable estimates as the estimates are closer to the true known values and the estimator has the narrowest confidence interval compared to all the methods. The simulation studies also indicated that the EM approach produced more reliable rate parameter estimators when $\lambda\tau$ is less than 0.08. This explains why the EM approach produced better results than the other methods in this case as $\lambda\tau$ is set to 0.07.

The MME approach once again produced estimates that did not have any statistical justification, such as negative values for the σ^2 parameter. It should however be noted that this is possible, as mentioned in Section 2.6. However, the MME approach produced faster run times compared to the other methods.

Estimates of σ with true value 0.2			
Method	Mode	Lower 95%	Upper 95%
MLE	0.2086	0.1749	0.2354
Profiling	0.0719	-0.3279	0.4346
EM	0.0767	-0.3125	0.4272
MME	0.0968	-0.2945	0.4873

Tab. 3.7: Simulation results for the estimation of σ (true value = 0.2) using different estimation methods.

Estimates of $\lambda\tau$ with true value 0.07			
Method	Mode	Lower 95%	Upper 95%
MLE	0.0000	0.0000	0.1913
Profiling	0.0161	0.0040	0.1755
EM	0.0574	0.0138	0.1786
MME	0.0020	0.0011	6.5414

Tab. 3.8: Simulation results for the estimation of $\lambda\tau$ (true value = 0.07) using different estimation methods.

The results discussed above suggest that when the rate parameter $\lambda\tau$ is small and the jump sizes are small, the EM approach performs better than the other approaches for small samples.

3.2.3 Varying the Number of Simulations and Years

For the fixed number of simulations of one thousand, increasing the number of years improves the estimation results. As shown in Tables B.1 to B.5 and Tables B.6 to B.10 in Appendix B, both the MLE and profiling methods estimate the smaller jumps better (as they are narrower confidence intervals) when the sample size is larger. Given that the data that is used in this dissertation was collected over a period spanning 14 years, it is good news that the estimation methods are behaving well at 14 years. However, the EM approach still produces estimates that are not reliable when $\lambda\tau$ is large. The MME still produces results that do not make any statistical sense, such as negative variances even when the number of years is set at 14. For this reason, it was decided that the MME approach would not be applied to the empirical data.

The MLE and profiling method estimators become more precise, as the number of years increases, in the sense that they exhibit narrower confidence intervals.

Estimates of μ_{jump} with true value 0.005			
Method	Mode	Lower 95%	Upper 95%
MLE	0.0030	-0.0661	0.0706
Profiling	0.0033	-0.0414	0.0655
EM	0.0030	-0.0088	0.0528

Tab. 3.9: Simulation results for the estimation of μ_{jump} (true value = 0.005) using different estimation methods.

Estimates of σ_{jump} with true value 0.03			
Method	Mode	Lower 95%	Upper 95%
MLE	0.0161	0.0000	0.5618
Profiling	0.0250	0.0065	0.1003
EM	0.0277	0.0104	0.0465
MME	0.1733	0.0528	0.4734

Tab. 3.10: Simulation results for the estimation of σ_{jump} (true value = 0.03) using different estimation methods.

However, the estimators for μ_{jump} and σ_{jump} become more underestimated as the number of years increases. The EM approach does not show much improvement when the number of years increases as it was already producing estimators with small confidence intervals when the number of years was fixed at one year.

Note that studies were also conducted to examine how the parameters behave when the number of simulations were changed, and the number of years were held constant at 3 years. The number of simulations was varied between 500, 1000, 2000 and 4000. The results indicated that the estimation methods produced better estimates (in terms of smaller confidence intervals) as the number of simulations increased. It was decided that using a 1000 simulations offered the right balance between run time and accuracy.

3.2.4 Run Time and Code Considerations

The run time of the code was a key consideration in this dissertation. The code initially took roughly 24 hours to run one thousand simulations of 14 years. However, the run time was improved by a factor of three when parallel computing was used on a machine with four cores. The parallel computing was facilitated by the `doParallel` package in the statistical language **R**.

The code for jump test of Lee and Mykland (2008) is an adjusted version of the

MATLAB code supplied by the authors. The code for the EM approach that was used is an adjusted version of the code in the **EMjumpdiffusion R** package. The code used in this dissertation can be found on the accompanying cd-rom.

Chapter 4

Empirical Implementation

4.1 Jump Test Results

As mentioned in Section 2.7, before the jump diffusion model in equation (2.1) can be calibrated using the different calibration techniques, the returns need to be tested for the presence of jumps.

The test of Lee and Mykland (2008) was applied to the data specified in Section 2.1. The test was performed at a 5% significance level and the value of k was set to 252. In addition, the returns from the instruments considered in this dissertation were found to be stationary using the Augmented Dickey-Fuller test. This means that the jump test is applicable to the empirical data. The output from the Lee and Mykland (2008) test is shown in Figures 4.1 and 4.2.

For the equity returns data, four jumps were detected in the MTN returns, five were detected for the ANG returns data and two jumps were detected in the Top40 index returns. For MTN and ANG returns, most of the jumps were up jumps. This may indicate that the true mean jump size μ_{jump} parameter is positive for these two sets of returns. This was confirmed by the fact that the empirical average jump size for MTN was 0.066 and that of ANG was 0.079, which are both positive.

As mentioned above, there were only two jumps detected from the Top40 returns over this period. This suggests that the frequency of jumps parameter $\lambda\tau$ must be lower than that of MTN and ANG. This is also confirmed by the model calibration results in Section 4.2. A possible reason why the Top40 index did not have as many jumps as the single stock returns is that the Top40 is a weighted index, so the returns are in effect smoothed out amongst the member stocks. This reduces the likelihood of the presence of jumps.

For the interest rate data, the test of Lee and Mykland (2008) detected 81 jumps in the 6 month Jibar log returns, 21 jumps were detected for the 5 year swap returns and 20 for the 15 year swap returns. The high number of jumps in the 6 month Jibar log returns was expected as it is in the short end of the interest rate curve and

is directly affected by changes of the repo rate by the South African Reserve Bank. The empirical average jump size for 6 month Jibar was found to be -0.008 . This suggests that the calibration results from all the methods should have a negative μ_{jump} parameter for the 6 month Jibar returns. This is confirmed by the results in Section 4.2.

The empirical average jump size μ_{jump} of the 5 year and 15 year returns was 0.013 and 0.004 respectively. The sign of this parameter for both sets of data is positive, suggesting that the true parameter value for μ_{jump} is positive. This is also supported by the sign of the μ_{jump} estimates in the calibration results in Section 4.2.

Through examining Figures 4.1 and 4.2, it can be noted that the jumps for the interest rate data are smaller, in absolute terms, than the equity jumps that were detected. In addition, the interest rate returns seem to have jumps at the same time as each other. It is suspected that this is because there are dependencies between the swaps returns of different maturities.

It is also worth remembering that the authors of the jump detection test recommend using the test at the 15 minute frequency as this is the frequency at which the test is most powerful. Given that the data used in this dissertation was at a daily frequency, the results could have been heavily influenced by the frequency of the data used. In particular, the total number of jumps may be underestimated as simulations by Lee and Mykland (2008) show that the jump test only detects 2.6% of the actual jumps when applied to data sampled at a daily frequency.

4.2 Model Calibration Results

The results shown in Figures 4.1 to 4.2 and the comments in Section 4.1 strongly suggest that the returns from all of the instruments have jumps. This means that the model in equation (2.1) can be calibrated using the different calibration techniques.

As in Chapter 3, when using the MLE and profiling methods, the infinite sum in the probability density function in equation (2.2) was truncated at $N = 10$. This cut-off value was used as per the suggestion of Ball and Torous (1983). The MLE approach was applied with 100 different starting values and the profiling method with 10 different starting values to improve the precision of the methods.

In order to assess the uncertainty in the parameter estimates, a bootstrapping technique was used. This involved sampling with replacement from the returns data and applying the methods repeatedly to the re-sampled data. This technique assumes that the returns are independent. The re-sampling was repeated one thousand times. This resulted in a histogram of parameter estimates that could then be used

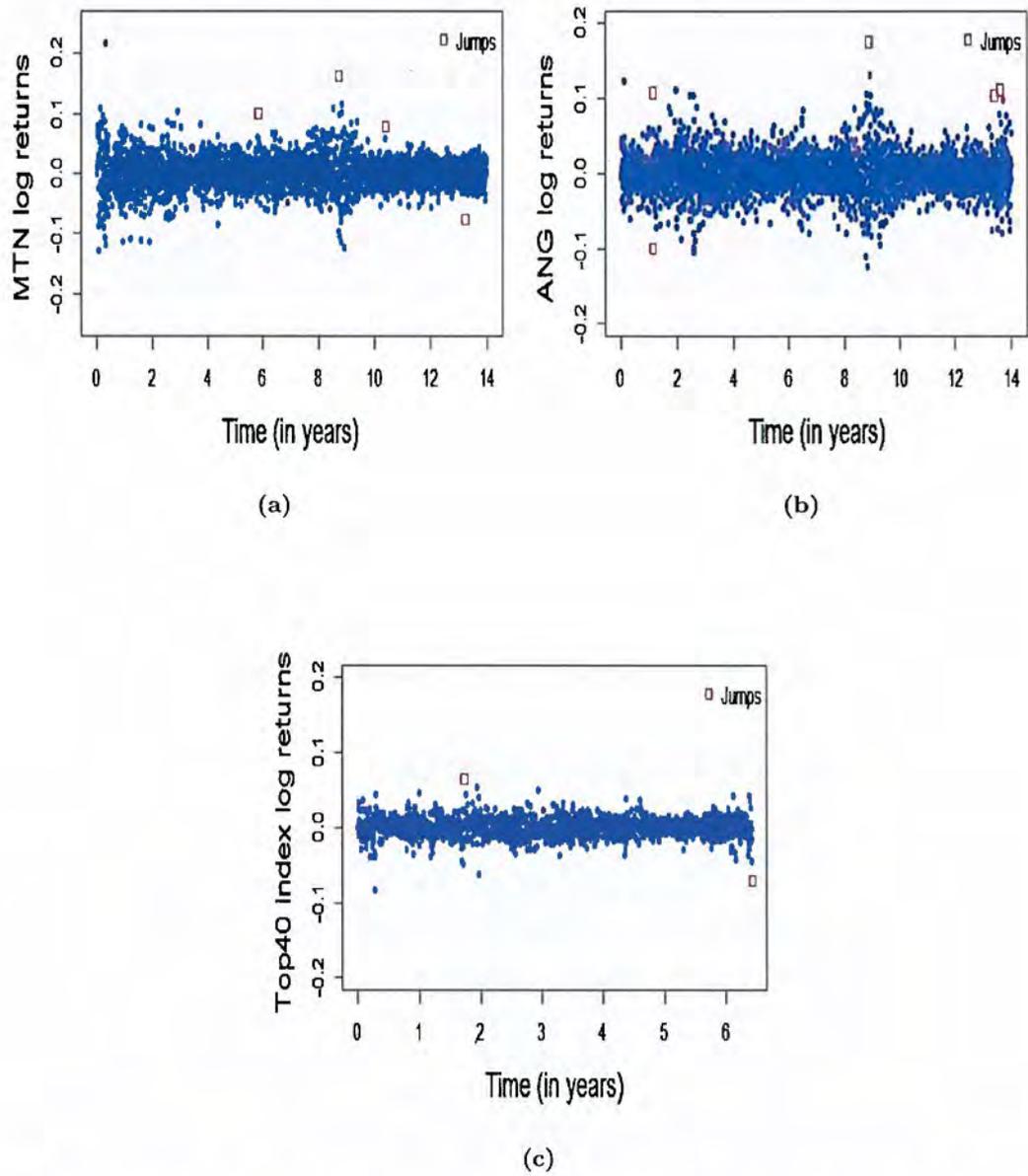


Fig. 4.1: Equity log return series with the detected jumps indicated by squares:
(a) MTN (b) ANG (c) Top40.

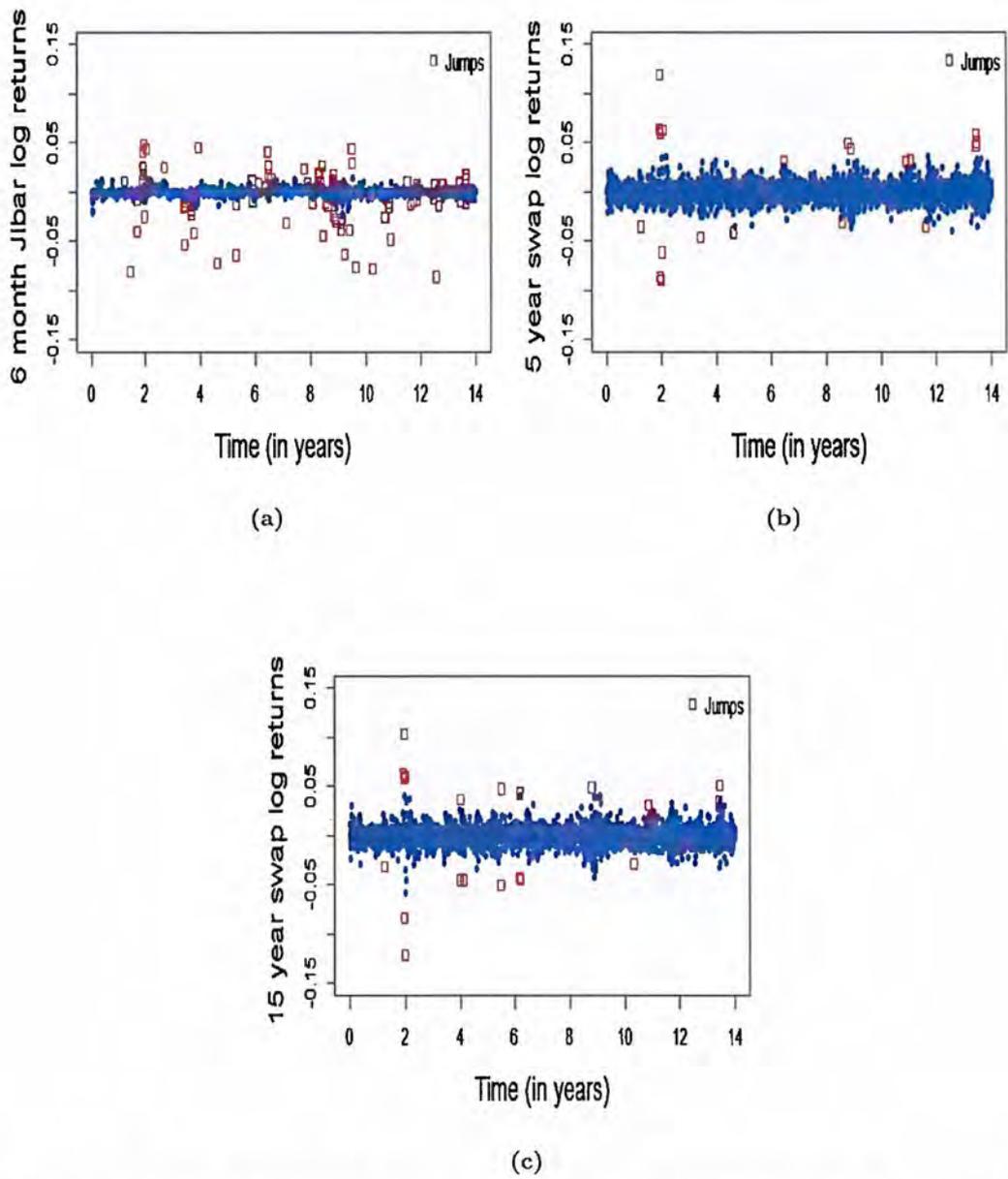


Fig. 4.2: Interest rate log return series with the detected jumps indicated by squares: (a) 6 month Jibar (b) 5 year swap (c) 15 year swap.

to form 95% bootstrap confidence intervals for the parameters. The 95% bootstrap intervals were formed by calculating the 97.5th percentile and the 2.5th percentile of the parameter estimate data.

The results of the model calibration are shown in Tables 4.1 to 4.6.

MTN results			
	MLE Estimate	Profiling Estimate	EM Estimate
μ	0.0352 (-0.118; 0.528)	0.0392 (-0.081; 0.201)	0.0421 (-0.149; 0.221)
σ	0.2744 (0.192; 0.234)	0.2778 (0.234; 0.306)	0.2861 (0.264; 0.305)
$\lambda\tau$	0.3157 (0.191; 0.330)	0.2810 (0.122; 0.409)	0.2121 (0.135; 0.296)
μ_{jump}	0.0021 (-0.001; 0.058)	0.0023 (0.000; 0.005)	0.0030 (0.000; 0.007)
σ_{jump}	0.0326 (0.000; 0.038)	0.0345 (0.029; 0.053)	0.0386 (0.033; 0.045)

Tab. 4.1: Jump diffusion parameter estimates for the MTN log returns using different estimation methods. The 95% bootstrap interval is shown in the parenthesis.

ANG results			
	MLE Estimate	Profiling Estimate	EM Estimate
μ	-0.1052 (-0.320; 0.108)	-0.0868 (-0.250; 0.049)	-0.0833 (-0.278; 0.101)
σ	0.2708 (0.228; 0.294)	0.2810 (0.252; 0.316)	0.2843 (0.260; 0.304)
$\lambda\tau$	0.3631 (0.224; 0.695)	0.2679 (0.095; 0.396)	0.2337 (0.161; 0.326)
μ_{jump}	0.0014 (0.000; 0.004)	0.0017 (0.000; 0.006)	0.0019 (-0.001; 0.006)
σ_{jump}	0.0303 (0.024; 0.037)	0.0350 (0.006; 0.055)	0.0364 (0.032; 0.041)

Tab. 4.2: Jump diffusion parameter estimates for the ANG log returns using different estimation methods. The 95% bootstrap interval is shown in the parenthesis.

4.2.1 Equity Returns

The equity returns data consisted of MTN, ANG and the Top40 index. From the jump test analysis that was performed, it was noted that MTN (5 jumps) and ANG (4 jumps) has the largest number of jumps and the Top40 index (2 jumps) has the least. The results of the calibrations, from all the methods, confirm this result as $\lambda\tau$ for the MTN and ANG returns was higher than that of the Top40 index data.

With regards to the estimated value of μ , the three methods agree as to the sign as well as the approximate magnitude of the drift parameter μ . The MTN

Top40 results			
	MLE Estimate	Profiling Estimate	EM Estimate
μ	0.2509 (0.150; 0.358)	0.2469 (0.117; 0.341)	0.2396 (0.147; 0.343)
σ	0.15507 (0.142; 0.166)	0.1561 (0.146; 0.182)	0.1600 (0.150; 0.169)
$\lambda\tau$	0.27240 (0.167; 0.429)	0.2545 (0.054; 0.330)	0.1940 (0.140; 0.267)
μ_{jump}	-0.0017 (-0.003; 0.000)	-0.0018 (-0.005; 0.000)	-0.0022 (-0.004; 0.000)
σ_{jump}	0.0188 (0.016; 0.023)	0.0194 (0.018; 0.043)	0.0217 (0.019; 0.025)

Tab. 4.3: Jump diffusion parameter estimates for the Top40 log returns using different estimation methods. The 95% bootstrap interval is shown in the parenthesis.

6 month Jibar results			
	MLE Estimate	Profiling Estimate	EM Estimate
μ	0.0000 (-0.000; 0.000)	0.0016 (-0.002; 0.005)	0.0169 (0.010; 0.023)
σ	0.0000 (0.000; 0.000)	0.0032 (0.003; 0.004)	0.0109 (0.009; 0.012)
$\lambda\tau$	0.9545 (0.444; 1.346)	0.8710 (0.778; 0.961)	0.259 (0.224; 0.310)
μ_{jump}	-0.0005 (-0.004; 0.003)	-0.0002 (-0.001; 0.000)	-0.0010 (-0.002; 0.000)
σ_{jump}	0.0031 (0.003; 0.010)	0.0031 (0.003; 0.004)	0.0106 (0.009; 0.013)

Tab. 4.4: Jump diffusion parameter estimates for the 6 month Jibar log returns using different estimation methods. The 95% bootstrap interval is shown in the parenthesis.

and Top40 instruments have a positive drift, suggesting that on average they were trending upwards over the period of study. The Top40 index drift parameter μ was 82% larger than that of MTN, suggesting that the Top40 index has a stronger tendency of drifting upwards when compared with MTN. The drift parameter for ANG was negative, suggesting that it had the tendency of drifting downwards over the period of study.

The diffusion volatility parameter σ for the MTN and ANG returns were higher, at around 28% for both of them, than that of the Top40 index, which was around 15%. A possible explanation of this is that the Top40 index is a weighted sum of the returns from the Top40 stocks on the JSE by market capitalisation, meaning that it is possible for the volatility to be lower than for the individual stocks due to the possibility of negatively correlated components, and thus diversification.

The results suggest that the average jump sizes are positive for MTN and ANG, as represented by μ_{jump} , while being negative for the Top40 index. In addition,

5 year swap results			
	MLE Estimate	Profiling Estimate	EM Estimate
μ	-0.0837 (-0.147; 0.000)	-0.0808 (-0.146; -0.013)	-0.0951 (-0.132; -0.031)
σ	0.0878 (0.060; 0.103)	0.0852 (0.071; 0.110)	0.1034 (0.094; 0.112)
$\lambda\tau$	0.3230 (0.134; 0.832)	0.4306 (0.079; 0.686)	0.1204 (0.068; 0.206)
μ_{jump}	0.0006 (0.000; 0.002)	0.0004 (0.000; 0.004)	0.0020 (0.001; 0.005)
σ_{jump}	0.0125 (0.007; 0.019)	0.0106 (0.009; 0.023)	0.0193 (0.015; 0.026)

Tab. 4.5: Jump diffusion parameter estimates for the 5 year swap log returns using the different estimation methods. The 95% bootstrap interval is shown in the parenthesis.

15 year swap results			
	MLE Estimate	Profiling Estimate	EM Estimate
μ	-0.0768 (-0.131; -0.015)	-0.0758 (-0.133; -0.012)	-0.0727 (-0.116; -0.013)
σ	0.1016 (0.090; 0.111)	0.1037 (0.091; 0.114)	0.1087 (0.100; 0.116)
$\lambda\tau$	0.1812 (0.101; 0.337)	0.1498 (0.071; 0.376)	0.1030 (0.059; 0.171)
μ_{jump}	0.0011 (0.000; 0.003)	0.0014 (-0.001; 0.004)	0.0019 (0.000; 0.004)
σ_{jump}	0.0161 (0.012; 0.021)	0.0180 (0.011; 0.027)	0.0211 (0.016; 0.028)

Tab. 4.6: Jump diffusion parameter estimates for the 15 year swap log returns using the different estimation methods. The 95% bootstrap interval is shown in the parenthesis.

the jump size volatility parameter σ_{jump} is also smaller than the diffusion volatility parameter σ , suggesting that most of the volatility in the instruments is from the diffusion component.

For the equity returns data, the three methods were more or less in agreement as to the approximate value of the parameter vector $(\mu, \sigma, \lambda\tau, \mu_{jump}, \sigma_{jump})$. As was observed in the simulation study, the EM algorithm consistently underestimates the value of $\lambda\tau$ when the true value of $\lambda\tau$ is small. However, the other parameters $\{\mu, \sigma, \mu_{jump}, \sigma_{jump}\}$ have been estimated with narrower bootstrap confidence intervals compared to the MLE and profiling methods.

The MLE approach produced estimates with the larger 95% bootstrap confidence intervals, while the profiling method produced estimates that are in-between the MLE and EM approach with regards to bootstrap confidence width. Thus, given that the point estimates produced by the three methods are within the same approximate value, the best method can be judged based on the width of the bootstrap

confidence interval. Based on this criterion, the EM approach was the best approach followed by the profiling method.

4.2.2 Interest Rate Returns

The interest rate returns data consisted of 6 month Jibar, 5 year swap and the 15 year swap. From the jump test analysis that was performed in Section 4.1, it was established that 6 month Jibar has the largest number of jumps with 81 jumps, while the 5 year and 15 year swap returns have 21 and 20 jumps respectively. Thus one would expect that the rate parameter $\lambda\tau$ for 6 month Jibar will be higher than both the 5 year and the 15 year swap. This was confirmed by the calibration returns in Tables 4.4 to 4.6 as $\lambda\tau$ was higher for 6 month Jibar than for the 5 year and the 15 year returns data.

One would also expect that the jump parameters for the 5 year and 15 year swap will be similar, as they have a similar number of jumps as per the Lee and Mykland (2008) test. However, for the MLE and profiling methods, the results appear to be different between the 5 year and the 15 year returns. On the other hand, the EM approach seems to estimate the parameters for the 5 year and 15 year swap in a consistent manner, as the parameter estimates are very similar. However, across all the methods, μ and σ appear to be estimated well as they have similar values for the three methods.

In an analysis of the actual jump sizes detected by the Lee and Mykland (2008) test, it was established that the jump sizes of the 5 year returns are larger in absolute terms than that of the 15 year swap returns. This means that the different parameter estimates obtained by the MLE and profiling methods for the parameters $\{\lambda\tau, \mu_{jump}, \sigma_{jump}\}$ between the 5 and 15 year swap rates was possible as the jump characteristics are different, although the number of jumps is almost the same.

The calibration methods produced varying results for the 6 month Jibar data. The results from the MLE approach suggest that there is no diffusion component to the model as both the drift μ and diffusion coefficient σ are estimated to be zero. However, the rate parameter was estimated to be 0.953 which turned out to be the highest for all the instruments that are under study. This was as expected as a large number of jumps (81) were detected by the Lee and Mykland (2008) jump test.

The profiling method also produced a large value for the rate parameter $\lambda\tau$. However, unlike the MLE approach, it suggests that there is a diffusion component to the model, although the estimated values for μ and σ are small, as can be seen in Table 4.4.

As mentioned in Chapter 3, when $\lambda\tau$ is large, the estimation results for the EM approach are not reliable. Given that $\lambda\tau$ for 6 month Jibar returns is large, the

results of the EM method are considered to be unreliable.

It is important to note that for the interest rate data it is **not** the short rate that is being modelled. Instead, returns of a point on the yield curve (e.g. 6 month Jibar) is being modelled. Thus the point on the yield curve is treated as an asset, and its returns are modelled. This approach was taken as the focus is on modelling the jump risk of points on the yield curve, and not a model of the entire term structure. This is important as models with mean reversion are usually used to model the short rate. In this dissertation, the Merton (1976) model was used and it does not allow for mean reversion. This should be taken into account when analysing the above calibration results.

For the estimation of the interest returns parameters, the MLE and profiling methods are more reliable than the EM approach as the EM approach is not reliable when there are many jumps, which is the case for the interest rate data. The best method for estimating the interest rate parameters, based on the width of the bootstrap intervals, is the profiling method followed by the MLE approach.

4.2.3 Comparing the Equity and Interest Rate Returns

Given that the equity jump sizes are larger than the interest rate jump sizes, and that the number of jumps for equity was smaller than the interest rate jumps, it may imply that the two sets of data need different models. Specifically, the model in equation (2.1) assumes that the jumps are of finite activity, that is, finite jumps occur in finite time. Given that the number of jumps for the equity returns are small, one would argue that the Merton (1976) model is well specified for the South African equity market.

There are, however, models that allow for an infinite number of jumps to occur in finite time, such as the variance gamma model of Madan and Seneta (1990) and the generalized Hyperbolic motion of Eberlein and Keller (1995). A test for the activity of jumps, such as that of Ait-Sahalia and Jacod (2011), may be applied to the equity and interest rate data. It may well be discovered that the model needed for the interest rate is an infinite jump activity model, such as the variance gamma model. This may then mean that the model used in this dissertation would be misspecified for the interest rate data. No jump activity test was performed in this dissertation. It is thus an area open for future work.

In addition, the empirical drift sizes μ for the interest rate data are small relative to the equity drift sizes. This may be because the Merton (1976) model is not a suitable model for the interest rate data as it does not allow for mean reversion, resulting in small drift sizes. It would be useful to compare the fit of a jump diffusion model with mean reversion with the Merton (1976) model for the interest rate data.

Chapter 5

Conclusions and Recommendations

This dissertation studied the application of jump diffusion processes in the South African context. This study included examining the jump diffusion properties of the South African equity market, as well as the South African interest rate market. The test for the presence of jumps of Lee and Mykland (2008) was applied to market data and jumps were detected. This suggested that fitting a jump diffusion process to the data over the period concerned was statistically justified. However, the test was applied to data observed at a daily frequency. This may have produced inaccurate results as Lee and Mykland (2008) recommend applying the test at the 15 minute frequency as this is the frequency at which the test has the most power.

Different methods of calibrating the jump diffusion model were used in this dissertation. These methods were the standard MLE approach, the likelihood profiling approach of Honore (1998), the EM algorithm and the MME approach. The methods were first applied to simulated data, and the simulation results indicated that the MME approach was not an appropriate method to apply to empirical data as it produced results which did not make statistical sense, such as negative variances. For the case when the rate parameter $\lambda\tau$ is small, the simulation results indicated that the EM approach was the most precise of the other three methods, followed by the likelihood profiling approach of Honore (1998). The MLE approach gave unsatisfactory results for the simulated data, the reason being that the likelihood function is not bounded. However, when applying to empirical data, the MLE and the profiling methods were improved by using many different starting values.

For empirical data, the results were varied between the equity and interest rate data. It was found that for the equity returns data the EM approach produced the most satisfactory results as its estimators had the narrowest 95% bootstrap confidence intervals. Although the EM approach produced the smallest confidence intervals for the interest rate data as well, the results were not reliable as the method

is not reliable (due to the pdf approximation in Section 2.5) when $\lambda\tau$ is large. It was then found that the profiling method produced the parameter estimates with a smaller bootstrap confidence interval than the MLE approach, and was thus the best method to use for the interest rate data.

A possible extension to this dissertation would be to expand the analysis for interest rate returns by using an Ornstein-Uhlenbeck process and compare the fit to the Geometric jump diffusion model that was used in this dissertation. This is because some of the results may have been unsatisfactory due to the mis-specification of the asset return model as interest rate usually have mean reversion properties. In addition a model which allows for the infinite activity of jumps can be tried as well, as the interest rate data seems to suggest that such a model may be useful. To check for the mis-specification of the models, the non-parametric specification test of Aït-Sahalia *et al.* (2009) could be used.

It may be worth while to use a better approximation for the probability density function in Section 2.5 on which the EM approach was used. Although this may lead to an increase in run time, it can possibly increase the accuracy of the results. Furthermore, this may lead to it being the overall best method to use for both the equity and interest rate returns over the period of study as it will no longer be affected by large values of $\lambda\tau$.

Another aspect that could be looked at is modelling the occurrence of co-jumps in the equity and interest rate market. This is because the jump test results suggest that the equity jump times are a subset of the jump times for the interest rate return jump times. That is, one could specify a model that allows for the possibility of jumps occurring in both markets simultaneously. The method of Lahaye *et al.* (2011) could be used to assist with analysing the co-jumps. The the non-parametric specification test of Aït-Sahalia *et al.* (2009) could be used to check if the model is misspecified.

In many financial institutions, first order models such as the Black and Scholes (1973) model, for equity returns, and the Black (1976) model, for interest rates, are used in the calculation of economic and regulatory capital. As outlined in this dissertation, these first order models do not explicitly allow for the jump risk which is present in these markets. It may be worth the regulators of financial institutions considering making the jump diffusion models as a basic model requirement when calculating regulatory capital. This is because losses from jumps can be significant and can even result in the insolvency of a company, such as those witnessed during the 2008 financial crisis.

Bibliography

- Aït-Sahalia, Y., Fan, J. and Peng, H. (2009). Non-parametric transition-based tests for jump diffusions, *American Statistical Association* **104**(487): 1102–1116.
- Aït-Sahalia, Y. and Jacod, J. (2011). Testing whether jumps have finite or infinite activity, *The Annals of Statistics* **39**(3): 1689–1719.
- Andersen, L. (2000). Jump diffusion processes: Volatility smile fitting and numerical methods for option pricing, *Review of Derivatives Research* **4**(3): 231.
- Ball, C. A. and Torous, W. N. (1983). A simplified jump process for common stock returns, *Journal of Financial and Quantitative Analysis* **18**(1): 53–65.
- Bandi, F. M. and Phillips, P. C. B. (2003). Fully non-parametric estimation of scalar diffusion models, *Econometrica* **71**(1): 241–283.
- Barndorff-Nielsen, O. E. and Shepard, N. (2006). Econometrics of testing for jumps in financial economics using bipower variation, *Journal of Financial Econometrics* **4**(1): 1.
- Beckers, S. (1981). A note on estimating the parameters of the diffusion-jump model of stock returns, *Journal of Financial and Quantitative Analysis* **16**(1): 127–140.
- Black, F. (1976). The pricing of commodity contracts, *Journal of Financial Economics* **3**: 167–179.
- Black, F. and Scholes, M. (1973). The pricing of options and corporate liabilities, *Journal of Political Economy* **81**(3): 637–654.
- Cappe, O., Moulines, E. and Ryden, T. (2005). Interference in hidden markov models, Springer Series in Statistics.
- Cont, R. (2001). Empirical properties of asset returns: stylized facts and statistical issues, *Quantitative Finance* **1**: 223 – 236.
- Cont, R. and Tankov, P. (2003). Financial modelling with jump processes, Chapman and Hall/CRC Press.
- Duffie, D. and Singleton, K. (1993). Simulated moments estimation of markov models of asset prices, *Econometrica* **61**(4): 929–952.
- Eberlein, E. and Keller, U. (1995). Hyperbolic distributions in finance, *Bernoulli* **1**(3): 281–299.

- Heston, S. L. (1993). A closed-form solution for options with stochastic volatility with applications to bond and currency options, *The Review of Financial Studies* **6**(2): 327–343.
- Honore, P. (1998). Pitfalls in estimating jump diffusion models.
- Kiefer, N. M. (1978). Discrete parameter variation: efficient estimation of a switching regression model, *Econometrica: Journal of the Econometric Society* **46**(2): 427–434.
- Kou, S. G. and Wang, H. (2004). Option pricing under a double exponential jump diffusion model, *Management Science* **50**(9): 1178–1192.
- Lahaye, J., Laurent, S. and Neely, C. J. (2011). Jumps, co-jumps and macro announcements, *Journal of Applied Econometrics* **26**(6): 893–921.
- Lee, S. S. and Mykland, P. A. (2008). Jumps in financial markets: A new non-parametric test and jump dynamics, *The Review of Financial Studies* **21**(6): 2535–2563.
- Madan, D. and Seneta, E. (1990). The variance gamma(vg) model for share returns, *Journal of business* pp. 511–524.
- Merton, R. C. (1976). Option pricing when underlying stock returns are discontinuous, *Journal of Financial Economics* **3**(12): 125 – 144.
- Press, S. J. (1967). A compound events model for security prices, *Journal of business* **40**(3): 317–335.
- Serlin, R. H. (2007). Towards better estimation of jump diffusion models.
- West, G. (2005). Calibration of the sabr model in illiquid markets, *Applied Mathematical Finance* **12**(4): 371–385.

Appendix A

Proving that the Likelihood function is not bounded

This Appendix provides the additional proofs (to the proof in Chapter 2) that are required to show that the Likelihood function from the Merton (1976) model is not bounded.

A.1 Transition Density of the Jump Diffusion Model

In this section, we derive the transition density of the Merton (1976) model: Consider the jump diffusion model in equation (2.1). Conditioning on n jumps, the increment of $\ln S(t)$ is $\mathbb{N}(\mu\tau + n\mu_{jump}, \sigma^2\tau + n\sigma_{jump}^2)$, where τ is a time increment. It then follows that:

$$\mathbb{P}\left(\ln\left(\frac{S(t+\tau)}{S(t)}\right) = w - x | N(t) = n\right) = \frac{\phi\left(\frac{w-x-(\mu\tau+n\mu_{jump})}{\sqrt{\sigma^2\tau+n\sigma_{jump}^2}}\right)}{\sqrt{\sigma^2\tau+n\sigma_{jump}^2}}$$
$$\mathbb{P}\left(\ln\left(\frac{S(t+\tau)}{S(t)}\right) = w - x, N(t)\right) = \sum_{n=0}^{\infty} \frac{e^{-\lambda\tau}(\lambda\tau)^n \phi\left(\frac{w-x-(\mu\tau+n\mu_{jump})}{\sqrt{\sigma^2\tau+n\sigma_{jump}^2}}\right)}{n! \sqrt{\sigma^2\tau+n\sigma_{jump}^2}}.$$

A.2 Deriving an Expression for $h(\mu, \sigma)$

In this section, the expression for $h(\mu, \sigma)$ used in Section 2.3 is derived. Consider:

$$\begin{aligned}
0 &\leq \sigma^2\tau + na_4^2 \geq na_4^2 \\
\frac{1}{\sigma^2\tau + na_4^2} &\leq \frac{1}{na_4^2} \\
\frac{(y_k - (\mu\tau + na_3))^2}{\sigma^2\tau + na_4^2} &\leq \frac{(y_k - (\mu\tau + na_3))^2}{na_4^2} \\
\frac{1}{2} \frac{(y_k - (\mu\tau + na_3))^2}{\sigma^2\tau + na_4^2} &\geq -\frac{1}{2} \frac{(y_k - (\mu\tau + na_3))^2}{na_4^2} \\
\exp\left(-\frac{1}{2} \frac{(y_k - (\mu\tau + na_3))^2}{\sigma^2\tau + na_4^2}\right) &\geq \exp\left(-\frac{1}{2} \frac{(y_k - (\mu\tau + na_3))^2}{na_4^2}\right) \\
g_n \frac{\exp\left(-\frac{1}{2} \frac{(y_k - (\mu\tau + na_3))^2}{\sigma^2\tau + na_4^2}\right)}{\sqrt{2\pi(\sigma^2\tau + na_4^2)}} &\geq \frac{\exp\left(-\frac{1}{2} \frac{(y_k - (\mu\tau + na_3))^2}{na_4^2}\right)}{\sqrt{2\pi(\sigma^2\tau + na_4^2)}} g_n
\end{aligned}$$

It then follows that:

$$\begin{aligned}
\prod_{k=1}^N \sum_{n=0}^{\infty} \frac{\exp\left(-\frac{1}{2} \frac{(y_k - (\mu\tau + na_3))^2}{\sigma^2\tau + na_4^2}\right)}{\sqrt{2\pi(\sigma^2\tau + na_4^2)}} g_n &\geq \prod_{k=1}^N \sum_{n=0}^{\infty} \frac{\exp\left(-\frac{1}{2} \frac{(y_k - (\mu\tau + na_3))^2}{na_4^2}\right)}{\sqrt{2\pi(\sigma^2\tau + na_4^2)}} g_n \\
L(\mu, \sigma) = \prod_{k=1}^N \sum_{n=0}^{\infty} \frac{\exp\left(-\frac{1}{2} \frac{(y_k - (\mu\tau + na_3))^2}{\sigma^2\tau + na_4^2}\right)}{\sqrt{2\pi(\sigma^2\tau + na_4^2)}} g_n &\geq h(\mu, \sigma) = \prod_{k=1}^N \sum_{n=0}^{\infty} \frac{\exp\left(-\frac{1}{2} \frac{(y_k - (\mu\tau + na_3))^2}{na_4^2}\right)}{\sqrt{2\pi(\sigma^2\tau + na_4^2)}} g_n
\end{aligned}$$

Now set $h(\mu, \sigma)$ to be:

$$h(\mu, \sigma) = \prod_{k=1}^N \sum_{n=0}^{\infty} \frac{\exp\left(-\frac{1}{2} \frac{(y_k - (\mu\tau + na_3))^2}{na_4^2}\right)}{\sqrt{2\pi(\sigma^2\tau + na_4^2)}} g_n.$$

A.3 Taking the Limit into the Infinite Sum

In this section, the Weierstrass M-test was used to show that the double limit can be taken into the infinite sum. Now, let

$$f_n(\mu, \sigma) = \frac{g_n}{\sqrt{\sigma^2\tau + na_4^2}} \phi\left(\frac{y_k - (\mu\tau + na_3)}{\sqrt{na_4^2}}\right).$$

The infinite sum that is of concern is

$$\sum_{n=0}^{\infty} f_n(\mu, \sigma)$$

and the following limit is of interest:

$$\lim_{\sigma \rightarrow 0} \lim_{\mu \rightarrow y_1/\tau} \sum_{n=0}^{\infty} f_n(\mu, \sigma).$$

The Weierstrass-M test says that if there exists a function M_n , independent of μ and σ , such that $|f_n(\mu, \sigma)| \leq M_n$ for all $n \geq 0$ and

$$\sum_{n=0}^{\infty} M_n < \infty,$$

then $\sum_{n=0}^{\infty} f_n(\mu, \sigma)$ converges absolutely and uniformly.

If it can be shown that it converges uniformly, and since $f_n(\mu, \sigma)$ is continuous, then it will mean that

$$\sum_{n=0}^{\infty} f_n(\mu, \sigma)$$

is continuous and

$$\lim_{\sigma \rightarrow 0} \lim_{\mu \rightarrow y_1/\tau} \sum_{n=0}^{\infty} f_n(\mu, \sigma) = \sum_{n=0}^{\infty} \lim_{\sigma \rightarrow 0} \lim_{\mu \rightarrow y_1/\tau} f_n(\mu, \sigma).$$

The following steps show that such an M_n exists.

$$\begin{aligned} \frac{1}{\sqrt{2\pi(\sigma^2\tau + na_4^2)}} &\leq \frac{1}{\sqrt{2\pi a_4^2}} \\ \frac{e^{-\lambda\tau}(\lambda\tau)^n}{n!} \frac{1}{\sqrt{2\pi(\sigma^2\tau + na_4^2)}} &\leq \frac{e^{-\lambda\tau}(\lambda\tau)^n}{n!} \frac{1}{\sqrt{2\pi a_4^2}} \\ \frac{e^{-\lambda\tau}(\lambda\tau)^n}{n!} \frac{e^{\frac{(y_k - (\mu\tau + na_3))^2}{-2na_4^2}}}{\sqrt{2\pi(\sigma^2\tau + na_4^2)}} &\leq \frac{e^{-\lambda\tau}(\lambda\tau)^n}{n!} \frac{1}{\sqrt{2\pi a_4^2}} \\ |f_n(\mu, \sigma)| &\leq \frac{e^{-\lambda\tau}(\lambda\tau)^n}{n!} \frac{1}{\sqrt{2\pi a_4^2}} \\ |f_n(\mu, \sigma)| &\leq M_n \end{aligned}$$

where

$$M_n = \frac{e^{-\lambda\tau}(\lambda\tau)^n}{n!} \frac{1}{\sqrt{2\pi a_4^2}}.$$

It also follows that

$$\sum_{n=0}^{\infty} M_n = \frac{1}{\sqrt{2\pi a_4^2}} < \infty.$$

This means that the limit can then be taken into the infinite sum.

Appendix B

Simulation Results

This Appendix shows the simulation results for large and small jumps, all based on a time horizon of 14 years and one thousand simulations.

B.0.1 Large Jumps:

$$\{\mu = 0.05, \sigma = 0.2, \lambda\tau = 0.3, \mu_{jump} = 0.05, \sigma_{jump} = 0.07\}$$

Estimates of μ with true value 0.05			
Method	Mode	Lower 95%	Upper 95%
MLE	0.0495	-0.0651	0.2238
Profiling	0.0431	-0.0635	0.1662
EM	0.0535	-0.0641	0.1632
MME	3.2798	2.9828	3.5779

Tab. B.1: Simulation results for the estimation of μ (true value = 0.05) using different estimation methods.

Estimates of σ with true value 0.2			
Method	Mode	Lower 95%	Upper 95%
MLE	0.0495	-0.0651	0.2238
Profiling	0.1988	0.1922	0.2051
EM	0.1999	0.1934	0.2062
MME(σ^2)	-0.0142	-0.2048	0.2038

Tab. B.2: Simulation results for the estimation of σ (true value = 0.2) using different estimation methods.

B.0.2 Small Jumps:

$$\{\mu = 0.05, \sigma = 0.2, \lambda\tau = 0.07, \mu_{jump} = 0.005, \sigma_{jump} = 0.03\}$$

Estimates of $\lambda\tau$ with true value 0.3			
Method	Mode	Lower 95%	Upper 95%
MLE	0.3053	0.2640	0.3388
Profiling	0.3049	0.2845	0.3293
EM	0.2597	0.2425	0.2750
MME	0.0829	0.0340	70.7099

Tab. B.3: Simulation results for the estimation of $\lambda\tau$ (true value = 0.3) using different estimation methods.

Estimates of μ_{jump} with true value 0.05			
Method	Mode	Lower 95%	Upper 95%
MLE	0.0423	0.0369	0.0644
Profiling	0.0424	0.0387	0.0464
EM	0.0500	0.0457	0.0543

Tab. B.4: Simulation results for the estimation of μ_{jump} (true value = 0.05) using different estimation methods.

Estimates of σ_{jump} with true value 0.07			
Method	Mode	Lower 95%	Upper 95%
MLE	0.0640	0.0601	0.0761
Profiling	0.0639	0.0609	0.0677
EM	0.0700	0.0672	0.0727
MME	0.3159	0.0832	0.5971

Tab. B.5: Simulation results for the estimation of σ_{jump} (true value = 0.07) using different estimation methods.

Estimates of μ with true value 0.05			
Method	Mode	Lower 95%	Upper 95%
MLE	0.0477	-0.0527	0.1512
Profiling	0.0494	-0.0536	0.1548
EM	0.0509	-0.0457	0.1497
MME	0.1139	0.0036	0.2199

Tab. B.6: Simulation results for the estimation of μ (true value = 0.05) using different estimation methods.

Estimates of σ with true value 0.2			
Method	Mode	Lower 95%	Upper 95%
MLE	0.1996	0.1933	0.2065
Profiling	0.2003	0.1925	0.2063
EM	0.2001	0.1934	0.2059
MME(σ^2)	-0.0001	-0.0003	0.0001

Tab. B.7: Simulation results for the estimation of σ (true value = 0.2) using different estimation methods.

Estimates of $\lambda\tau$ with true value 0.07			
Method	Mode	Lower 95%	Upper 95%
MLE	0.0713	0.0437	0.1067
Profiling	0.0585	0.0412	0.1167
EM	0.0663	0.0460	0.0983
MME	0.0027	0.0009	0.0105

Tab. B.8: Simulation results for the estimation of $\lambda\tau$ (true value = 0.07) using different estimation methods.

Estimates of μ_{jump} with true value 0.005			
Method	Mode	Lower 95%	Upper 95%
MLE	0.0043	0.0006	0.0102
Profiling	0.0047	0.0006	0.0100
EM	0.0048	0.0008	0.0097

Tab. B.9: Simulation results for the estimation of μ_{jump} (true value = 0.005) using different estimation methods.

Estimates of σ_{jump} with true value 0.03			
Method	Mode	Lower 95%	Upper 95%
MLE	0.0289	0.0239	0.0340
Profiling	0.0245	0.0239	0.0358
EM	0.0297	0.0254	0.0346
MME	0.2905	0.2073	0.4254

Tab. B.10: Simulation results for the estimation of σ_{jump} (true value = 0.03) using different estimation methods.