Erasmus School of Economics

AllOptions

Distinction between Intra-day and Overnight Stock Return Distribution and its Influence on Option Pricing

Author: Katarina Bojc
Student ID: 355224
Supervisor: Prof. dr. Dick van Dijk
Co-reader: Asst. prof. dr. Marcin Jaskowski
Externl Supervisors: Prof. dr. ir. Michel Vellekoop, Geeske Vlaming and Dr. Sing Kwong Cheung

Econometrics Institute
Erasmus School of Economics
Erasmus University Rotterdam

ROTTERDAM, 2012

" If we knew what it was we were doing, it would not be called research, would it?"

- Albert Einstein

Acknowledgments

First and foremost I would like to thank my parents for giving me the opportunity to study in the Netherlands as well as for their support and encouragement through out my studies.
I wish to express my gratitude to Geeske Vlaming and prof. Michel Vellekoop for giving me an opportunity to do my master thesis at All Options. Their guidance and valuable suggestions were an essential part of this project.
My thanks goes out to Sing Kwong Cheung for his very helpful advices and programming tips as well as to Rob Willems for assisting me with all the practical issues.
I would also like to thank to my supervisor at Erasmus University Rotterdam, Dick van Dijk for his help and valuable remarks.

Abstract

This research investigates the importance of including the overnight behavior of the stock price in an option pricing model. During the trading hours stock prices process follows a stochastic model, which allows for stochastic volatility and random jumps. We model the overnight behavior by a single jump, which is independent of the intra-day component. In our research we consider Heston, Bates and Variance Gamma models, extend them by introducing an overnight jump and compare their performance. We find that introducing an overnight component in already existing option pricing models leads to option prices closer to the ones observed on the market. However the overnight jump alone is not able to capture the stock price behavior, random jumps should also be included. We find that the most successful of the considered option pricing model uses both a random jump during the day and scaled-t distributed overnight jumps, where the improvement from a normally distributed overnight jumps is significant.

KEYWORDS: Option pricing, Overnight jump, Stochastic volatility, Scaled-t distribution, Variance Gamma

JEL CLASSIFICATION: G13, C580

Contents

List of Figures . . . vi
List of Tables . . . vii

1 Introduction **1**

2 Theoretical Overview **3**
2.1 Pricing Options . . . 3
2.1.1 Black-Scholes Model . . . 4
2.2 The Stochastic Volatility Model with Random and Overnight Jumps 5

3 Distribution of Stock Returns **8**
3.1 Data . . . 9
3.2 Alternative distributions . . . 10
3.3 Parameter Estimation and Testing . . . 13
3.3.1 Goodness of Fit Test . . . 15
3.3.2 QQ-plots . . . 16
3.4 Independence of Intra-day and Overnight Returns . . . 19
3.4.1 Bartlett-Box Test . . . 19
3.4.2 Scatter Plots . . . 20

4 Extending the SVRJOJ Model **21**
4.1 Variance Gamma Process . . . 21
4.2 Variance Gamma Model with Normal Overnight Jump . . . 23
4.3 SVRJOJ Model with Scaled-t Overnight Jump . . . 23
4.4 Variance Gamma Model with Scaled-t Overnight Jump . . . 24

5 Option Pricing **25**
5.1 The SVRJOJ model . . . 26
5.2 The Variance Gamma model with a Normal Overnight Jump . . . 28
5.3 The Variance Gamma Model with the Scaled-t Overnight Jump . . . 29
5.4 The SVRJOJ model with Scaled-t Overnight Jump . . . 30

6 Option Data and Model Estimation **31**
6.1 Data . . . 31
6.2 Estimation . . . 33

7 Empirical Results **34**
7.1 Models Based on Stochastic Volatility . . . 34
7.2 Models Based on the Variance Gamma Process . . . 38
7.3 Alternative Evaluation of Option Pricing Models. . . . 40
7.3.1 Statistical indicators . . . 40
7.3.2 Indicators Based on the Absolute Error . . . 42
7.3.3 Model Evaluation Based on the Bid-Ask Spread . . . 48
7.4 Out of Sample Pricing Errors . . . 50

8 Conclusion and Further Research **54**
8.1 Conclusion . . . 54
8.2 Further Research . . . 55

Bibliography **56**

A Appendix **59**
A.1 QQ-plots of the Allianz stock returns . . . 59
A.2 Fitting Distribution to Alternative Data . . . 60
A.2.1 Distributions of BMW Stock Returns . . . 60
A.2.2 Distribution of DBK Stock Returns . . . 63
A.3 Parameter Estimates when using Absolute Error Objective Function 67
A.3.1 Parameter estimates for data with only one maturity . . . 69

List of Figures

3.1 Intra-day distributions 14
3.2 Overnight distributions 14
3.3 QQ-plots: Intra-day returns 17
3.4 QQ-plots: Overnight returns 18
3.5 Scatter plots of intra-day and overnight correlations 20

7.1 Absolute error for maturity 0.031 years (ALV). 42
7.2 Absolute error for maturity 0.120 years (ALV). 43
7.3 Absolute error for maturity 0.210 years (ALV). 43
7.4 Absolute error for maturity 0.4603 years (ALV). 44
7.5 Absolute pricing errors: Bates model 45
7.6 Absolute pricing errors: SVRJOJ model 46
7.7 Absolute pricing errors: SVRJOJ-SC model 46
7.8 Absolute pricing errors: VG-N model 47
7.9 Option prices given by the models compared to mid-market price . 50

A.1 QQ-plots - alternative distributions: Intra-day (ALV) 59
A.2 QQ-plots - alternative distributions: Overnight (ALV) 60
A.3 QQ-plots: Intra-day (BMW) 62
A.4 QQ-plots: Overnight (BMW) 63
A.5 QQ-plots: Intra-day (DBK) 65
A.6 QQ-plots: Overnight (DBK) 66

List of Tables

3.1 Moments of intra-day and overnight returns 10
3.2 Parameter Estimates 13
3.3 Chi Squared Test 16
3.4 Bartlett-Box Test 19

6.1 Option data overview 32

7.1 Parameter estimates: Models with stochastic volatility 35
7.2 Variance decomposition 36
7.3 Parameter estimates: Models based on variance gamma process . . 39
7.4 Statistical indicators for option pricing models 41
7.5 Best performing option pricing models depending on the maturity . 44
7.6 Average percentage deviations from the mid-market price 49
7.7 Absolute pricing errors: Out of sample 51
7.8 Percentage pricing errors: Out-of-sample 52
7.9 Absolute pricing errors: In-sample 52

A.1 Moments of intra-day and overnight returns: BMW 60
A.2 Distribution estimates: BMW 61
A.3 Chi Squared Test: BMW 61
A.4 Moments of intra-day and overnight returns: DBK 63
A.5 Distribution estimates: DBK 64
A.6 Chi Squared Test: DBK 64
A.7 Parameter estimates - Absolute error objective function (Stochastic volatility models) 67
A.8 Parameter estimates - Absolute error objective function (VG) . . . 68
A.9 Parameter estimates - Absolute error objective function (Stochastic volatility models) 69
A.10 Parameter estimates - Absolute error objective function (VG) . . . 70

Chapter 1

Introduction

Information important for financial markets accumulates globally around the clock. However most of financial markets are only opened from morning until late afternoon. During the time they are closed a lot of relevant information can become available, but can not be incorporated in the securities price. For example a European trader might expect the opening price today to be different from the closing prices yesterday, if there were some public announcements during the time the market was closed. Since financial markets around the world are generally correlated, the good or bad performance on Asian and American markets also influences the opening price on the European market. This overnight information is incorporated in the option price by investors who submit the orders to the exchange before the opening hours and with that influence the opening price.

The goal of this research is to capture this overnight behavior and introduce it into an already existing option pricing model. This idea was first presented in the article of Boes, Drost and Werker [10], who extend the Bates model by adding an overnight jump component. In their research they assume that intra-day and overnight stock price processes follow a normal distribution. This assumption of normality has been rejected several times by the literature (Mandelbrot [22] and Fama [12]), since returns usually exhibit negative skewness, have higher peaks and fatter tails than assumed under normality.

Our research incorporates these characteristics of returns, by considering alternative distributions to describe intra-day and overnight returns. Several distributions that describe stock returns more accurately than the normal distribution have been proposed in the literature. Mandelbrot [22] argues that stock returns are well captured by stable Paretian distributions with a characteristic exponent less than 2. The logistic distribution was proposed by Smith [30], exponential power distribution by Hsu [17]. The distribution that is considered to describe overnight returns rather accurately is the scaled-t distribution proposed by Praetz [25]. Madan and Seneta [21] considered the variance gamma distribution when describing the returns.

We look at all of these distributions and test which one captures intra-day and overnight returns of stocks in DAX index best. Intra-day returns are captured most accurately by the variance gamma distribution and overnight returns by the scaled-t distribution. We include those two distributions in our option pricing models and compare their performance to already existing option pricing models, based on the option prices of Allianz, one of the stocks in the DAX index.

The existing option pricing models that we consider in our research are firstly the Black-Scholes model [8], which despite his wide us has two well know short comings [27], the volatility smiles and the skewness premium. Therefore we also evaluate the Heston model [16] that introduces stochastic volatility in the Black-Scholes model and the Bates [5] model that extends the Heston model by adding a random jump process. As mentioned before we also consider the Stochastic Volatility Model with Random and Overnight Jumps (SVRJOJ) and the Stochastic Volatility Model with an Overnight Jumps (SVOJ) proposed by Boes, Drost and Werker [10] that add a normal overnight jump to the Bates and Heston model respectively. Our research extends these models by adding a scaled-t distributed jump to the SVRJOJ model instead of normally distributed jump.

The distribution that fits the intra-day returns best is the variance gamma distribution. Therefore we consider the variance gamma (VG) option pricing model proposed by Madan, Carr and Cheng [19] and extend it by adding a normally and scaled-t distributed jump.

All the above models are fitted to Allianz option data and evaluated based on how well they minimize the squared percentage error objective function, absolute pricing errors and out of sample pricing performance. We find that based on the above criteria models that include stochastic volatility, random and overnight jumps perform best, especially when the overnight jump is scaled-t distributed. Out of sample performance of all the considered models is poor.

The remainder of the report is organized as follows. In the next section we briefly discuss the option pricing theory and present the SVRJOJ model. Section 3 describes how well the above mentioned distributions fit the intra-day and overnight returns. In Section 4 the extended models that we propose are introduced. Later on in Section 6 we explain how to implement the option pricing models we consider in this research using Fourier transform and characteristic function. Our empirical results are presented in section 7. Finally Section 8 concludes.

Chapter 2

Theoretical Overview

In this chapter we briefly discuss the most important definitions and concepts considered in option pricing and this research.

Definition 2.0.1. An option is a financial derivative that represents a contract sold by one party (option writer) to another party (option holder). The contract offers the buyer the right, but not the obligation, to buy (call) or sell (put) a security or other financial asset at an agreed-upon price (the strike price) during a certain period of time or on a specific date (exercise date) [18].

Definition 2.0.2. European option is an option that can only be exercised at the end of its life, at its maturity. European options tend to sometimes trade at a discount to its comparable American option. This is because American options allow investors more opportunities to exercise the contract. The prices of an European call and put option at time t equal

$$c_t = exp(-r(T-t))E_t^{\mathbb{Q}}[max(S_T - K, 0)], \tag{2.1}$$

$$p_t = exp(-r(T-t))E_t^{\mathbb{Q}}[max(K - S_T, 0)], \tag{2.2}$$

where with c and p we denote the price of the European call and put option respectively. T stands for the maturity and r for the interest rate.

2.1 Pricing Options

When pricing options in complete markets risk neutral pricing can be used to determine the theoretical “fair”price of an option. This is done due to the fact that a risk-less portfolio consisting of the position in the option and a position in the stock can be set up. A risk-less portfolio can be constructed due to the fact that the price of an option and the stock price are both effected by the same source of

uncertainty: stock price movements. Therefore when an appropriate portfolio of the stock and an option is established (replicating portfolio), the gain or loss from the stock position always offsets the gain or loss from the option position.
We show how to price options for the example of the Black-Scholes model. For models considered later on, similar principles apply.

2.1.1 Black-Scholes Model

The Black-Scholes model is the most frequently considered option pricing model in the literature. Under this model the stock price follows the process

$$dS_t = \mu_t S_t dt + \sigma_t S_t dW_t, \tag{2.3}$$

where W_t is a Brownian motion under real world measure $\mathbb{P}$. As we mentioned before fair option prices are obtained under the risk neutral measure, therefore we introduce the equivalent martingale measure $\mathbb{Q}$, such that

$$\tilde{W}_t = W_t + \nu t,$$

where $\nu = (\mu - r)/\sigma$. Under this new measure the stock price process follows

$$dS_t = rS_t dt + \sigma_t S_t d\tilde{W}_t. \tag{2.4}$$

Here we show why the price of a call option with strike K and maturity T equals 2.1. The claim for such an option equals $(S_T - K)^+$. The price of an option at time zero should equal the value of the replicating strategy discounted back to time zero at the risk free rate, which equals

$$\exp(-rT)E^{\mathbb{Q}}((S_T - K)^+), \tag{2.5}$$

where $\mathbb{Q}$ is the martingale measure for $S_t e^{-rt}$. The value $(S_T - K)^+$ only depends on the stock price at the maturity T. In order to find an expectation of this claim we need to find the marginal distribution under $\mathbb{Q}$.

To do that we look at the logarithm of S_t that is obtained by applying Itôs lemma to equation (2.4)

$$d(\log(S_t)) = \sigma d\tilde{W}_t + (r - \frac{1}{2}\sigma^2)dt,$$

hence $\log(S_t) = log(S_0) + \sigma\tilde{W}_t + (r - \frac{1}{2}\sigma^2)t$ and thus $S_t = S_0 \exp(\sigma\tilde{W}_t + (r - \frac{1}{2}\sigma^2)t)$. From that we can see that S_T has a marginal distribution equal to the S_0 multiplied with the exponential of the normal distribution with mean $(r - \frac{1}{2}\sigma^2)T$ and variance $\sigma^2 T$. Therefore if we say that Z is normally distributed $N(-\frac{1}{2}\sigma^2 T, \sigma^2 T)$, we can write $S_T = S_0 \exp(Z + rT)$ and (2.5) equals $\exp(-rT)E((s\exp(Z+rT) - K)^+)$, which can be written as

$$V(S_0,T) = S_0\Phi(\frac{\log(\frac{S_0}{K}) + (r+\frac{1}{2}\sigma^2)T}{\sigma\sqrt{T}}) - K\exp(-rT)\Phi(\frac{\log(\frac{S_0}{K}) + (r-\frac{1}{2}\sigma^2)T}{\sigma\sqrt{T}}), \tag{2.6}$$

where $\Phi(x)$ is the cumulative standard normal distribution function. Formula (2.6) is the Black-Scholes formula for pricing European call options. The price of a call option was obtained under the risk neutral measure, however this price also holds under the real world measure. When we move from risk neutral measure to real world measure the expected growth rate in the stock price changes and the discount rate changes as well. It happens that these two changes always offset each other. For more detail on the topic of risk neutral pricing and Black-Scholes model we refer to [7].

Despite its wide use the Black-Scholes model has several shortcomings. As we have see the Black-Scholes model assumes that volatility is constant and independent of the strike price and maturity, hence the volatility surface should be flat. However in practice we observe skewed volatility surface often refereed to as the volatility simile and can not the be explained by the Black-Scholes model. Therefore alternative models were considered. One of those models is the Stochastic Volatility Model with Random and Overnight Jump, which is a building-block of our research and is described in the next section.

2.2 The Stochastic Volatility Model with Random and Overnight Jumps

The vast majority of exchanges are closed from the late afternoon until the next morning and the information that becomes available during that time can not be incorporated in the stock price immediately. However is included into the option price by investors who submit orders to the exchange before opening and hence influence the opening price.

The Stochastic Volatility Model with Random and Overnight Jumps (SVRJOJ), first introduced by Boes, Drost and Werker [10] as an extension of the Bates model, can capture this behavior of financial markets and therefore has the potential to price options more closely to the ones observed by the market. Here their model is introduced. Later on some possible extensions are considered.

The interest rates influence the option prices, so the SVRJOJ model assumes the risk free rate to follow a process $\frac{dB_t}{B_t} = rdt$ i.e $B_t = \exp(rt)$. The model has the following four risk components: stock price movement, volatility, random jump and overnight jump. Since only the stock process is a traded asset, the market is incomplete and therefore the equivalent martingale measure is no longer unique. The value process of the stock at time t under the risk neutral probability measure $\mathbb{Q}$ is defined as

$$\frac{dS_t}{S_{t-}} = rdt + \sigma_t dW_t^S + \sum_{i=1}^{N_t}(Y_i - 1) - \lambda\mu_{RJ}dt + d\sum_{i=1}^{\lfloor 252t \rfloor}(V_i - 1), \tag{2.7}$$

$$\log Y_i \sim N(\log(1+\mu_{RJ}) - \frac{1}{2}\sigma_{RJ}^2, \sigma_{RJ}^2), \tag{2.8}$$

$$\log V_i \sim N(-\frac{1}{2}\sigma_{OJ}^2, \sigma_{OJ}^2), \tag{2.9}$$

where $\{W_t^S\}$ is a standard Brownian Motion independent of the Poisson Process $\{N_t\}$ where,

$$N_t \sim Poisson(\lambda t).$$

Both $\{W_t^S\}$ and $\{N_t\}$ are assumed to be independent of jumps $\{Y_i\}$ and $\{V_i\}$. It is assumed that the weekend is a single night, therefore there are 252 trading days per year, hence also 252 overnight jumps per year. The distribution of random and overnight returns is chosen in such a way that $E[V_i] = 1$ and $E[Y_i] = 1 + \mu_{RJ}$.

The time varying stochastic volatility process in (2.7) is taken from Heston [16] and defined as

$$d\sigma_t^2 = -\kappa(\sigma_t^2 - \sigma^2)dt + \sigma_\sigma \sigma_t dW_t^V, \tag{2.10}$$

$$Corr_t(W_t^V, W_t^S) = \rho, \tag{2.11}$$

where κ is the speed of mean reversion, σ^2 is the long run mean of the variance, and σ_σ volatility of volatility. This specification of volatility allows for a negative premium of volatility risk, although we know volatility can not be negative. From equation (2.11) we can see that the Brownian Motions from the stock price process and stochastic volatility process are correlated. With this the fact that the large decline in the stock price is accompanied by a positive shock in volatility levels is incorporated in the model.

Under the risk neutral measure $\mathbb{Q}$ the process $\frac{S_{t+\Delta t}}{S_t}$ is a martingale, therefore the return over Δt should equal

$$E^{\mathbb{Q}}[\frac{S_{t+\Delta t}}{S_t}|S_t] = \exp(r\Delta t).$$

To show that process (2.7) really is a martingale under measure $\mathbb{Q}$ after discounting, we look at expectation of different parts of process (2.7) separately. For process

(2.7) to be a martingale the expectation of the random jump component should equal $E[\sum_{i=1}^{N_t}(Y_i - 1)] = \lambda\mu_{RJ}t$. We can show that is the case, since taking the distribution of a random jump (2.8) into account it follows

$$E[Y_i] = E[\exp(\log(1 + \mu_{RJ}) - \frac{1}{2}\sigma_{RJ}^2 + \sigma_{RJ}Z)],$$

where $Z \sim N(0,1)$. Considering the fact that $E[\exp(-\frac{1}{2}\sigma_{RJ}^2 + \sigma_{RJ}Z)] = 1$, it follows that $E[Y_i] = 1 + \mu_{RJ}$. The number or jump occurrences follows a Poisson process with intensity λ, therefore $E[N_t] = \lambda t$. Using conditioning we see that $E[\sum_{i=1}^{N_t}(Y_i - 1)] = \lambda\mu_{RJ}t$.

For process (2.7) to equal $\exp(rt)$ in the expectation, the overnight jump factor should have a mean of 1. This holds since V_i follows (2.9) and similarly as before we can show that $E[\exp(\sigma_{OJ}Z - \frac{1}{2}\sigma_{OJ}^2)] = 1$, from here it is straightforward to show that $E[\sum_{i=1}^{\lfloor 252t \rfloor}(V_i - 1)] = 0$.

Chapter 3

Distribution of Stock Returns

The assumptions that stock prices and stock returns are log-normally and normally distributed is often used in finance. This is due to the fact that they are rather simple distributions and are therefore it is easy to work with, as well as due to the Central Limit Theorem. They are also used in basic stochastic calculus and Itôs Lemma, that are often applied in derivatives pricing.
Let us first explain why log-normal and normal distribution follow from the Central limit theorem. Using S_i and R_i to represent the stock price and random stock return at time i respectively, we can write

$$S_1 = S_0(1 + R_1)$$

and after n days the return equals to

$$S_n = S_0 \prod_{i=1}^{n} (1 + R_i). \tag{3.1}$$

If we apply the logarithm on both sides of equation (3.1) we get

$$\log(S_n) = \log(S_0) + \sum_{i=1}^{n} \log(1 + R_i).$$

We know that each R_i is random, therefore also each $\log(1 + R_i)$ is random. As long as R_i are independent and identically distributed, with the finite mean and finite standard deviation of $\log(1 + R_i)$, we can apply the Central Limit Theorem and conclude that $\log(S_n)$ are normally distributed. From that it follows that S_n is log-normally distributed and since $\log(\frac{S_n}{S_{n-1}}) \approx \frac{S_n - S_{n-1}}{S_{n-1}}$ (first order Taylor approximation, for small $S_n - S_{n-1}$) we can see that returns are approximately normally distributed.

Even though assuming the returns to be normal is very convenient, several contradictions to it can be found in the literature. First evidences against normality were stated in the empirical research of Mandelbrot [22] and Fama [12], who claimed that price changes follow a stable Pareto distribution, with characteristic exponent less then 2, hence exhibit fat tails and infinite variance and are therefore more risky than assumed under normality. After those influential publications extensive research was done on the topic. The logistic distribution was proposed by Smith [30] due to its fatter tails compared to normal, scaled-t distribution by Praetz [25], since it exhibits fatter tails as well as allows for skewness. Hsu [17] argued that exponential power distribution describes returns well due to its fat tails that shrink at exponential speed. Seneta and Madan [21] proposed a variance gamma distribution that can capture fat tails and skewness. Returns could also be described by the combination of two distributions, Press [26] suggested to use the mixture of two normal distributions, since it would allow for skewness in the model.
In the remainder of this section we first describe the return data used in our research, later on we present all the above mentioned distributions in more detail. Only the pareto distribution was excluded, since its capability to capture returns was rejected by Clark [11]. All the distributions are fitted to the data and the parameter estimates are reported. To conclude which distribution fits the data best, a Chi Squared Goodness of Fit Test is conducted and as a visual tool QQ-plots are presented. We consider intra-day and overnight returns separately, since they possess different characteristics.

3.1 Data

Data used in this research are the stock prices of Allianz SE (ALV). For each trading day between January 2, 2007 to December 30, 2011 the opening and the closing price of the stock was obtained from public data by Bloomberg. The stock prices were corrected for dividends beforehand. The analyzed series returns are defined as

$$OR_t = \frac{OP_{t+1} - CP_t}{CP_t},$$

for overnight returns, where OP denotes opening price of the stock and CP closing price of the stock. Intra-day returns are calculated as

$$DR_t = \frac{CP_t - OP_t}{OP_t},$$

hence we can see the returns expressed as proportion of stock price. Table 1 below reports the empirical statistics of the stock returns under consideration. This information helps to understand the behavior of stock returns and choosing an appropriate distribution to capture them.

Table 3.1: Moments of intra-day and overnight returns

Returns	Mean	St. dev.	Min of series	Max	Skw	Krt
Intra-day	-0.0010	0.0226	-0.1408	0.1729	0.325	9.644
Overnight	0.0009	0.0157	-0.0998	0.1792	1.2339	24.7048

Moments of intra-day and overnight returns.Skw stands for skewness and Krt for kurtosis.

From the table above we can see that the intra-day returns are on average negative, while overnight returns are positive. For a normal distribution the skewness and kurtosis equal zero and three respectively. Since skewness and kurtosis reported in the table are far from those values for both intra-day and overnight returns we can assume that both returns are skewed and leptokurtic. The distribution chosen to fit the returns should be able to mimic this behavior.

3.2 Alternative distributions

As shown in the previous section intra-day, and overnight returns on the Allianz stock exhibit skewness, higher peaks and heavier tails than assumed under normality. In this part we present the above mentioned distributions that have at least one of the described characteristics of returns. The same distributions are considered for intra-day and overnight returns.

Logistic Distribution
This distribution was first proposed by Smith [30] and tested by Gray and French [15]. It is very similar to the normal distribution, however its tails are heavier. The density function of the logistic distribution is

$$f(x) = \frac{\exp(-\frac{x-\mu}{\alpha})}{\alpha[1+\exp(-\frac{x-\mu}{\alpha})]^2},$$

where μ is a location parameter and $\alpha > 0$ is a scale parameter. If a random variable X follows a logistic distribution then $E[X] = \mu$ and $V[X] = (\pi^2/3)\alpha^2$.

Scaled-t Distribution
First this distribution was considered by Praetz [25]. Later on also Blattberg and

Gonedes [9] and Peiro [24] confirmed that scaled-t distributions describes returns better than several other proposed distributions, since it allows for fatter tails when degrees of freedom are small. Its density function has the following form

$$f(x) = \frac{\Gamma(\frac{\nu+1}{2})}{\sigma\sqrt{\nu\pi}\Gamma(\frac{\nu}{2})}[\frac{\nu + (\frac{x-\mu}{\sigma})^2}{\nu}]^{-(\frac{\nu+1}{2})},$$

where μ and σ are location and scale parameter respectively and ν is a degrees of freedom parameter and Γ the gamma function. It holds that $E[X] = \mu$ and $V[X] = \frac{\nu}{\nu-2}\sigma^2$, when $\nu > 2$.

Exponential Power Distribution
The exponential power distribution presented by Hsu [17] allows for fat tails that shrink at an exponential rate and a high peak, therefore it should provide a reasonably good fit for stock returns. The density function is given as

$$f(x) = \frac{\exp[-\frac{1}{2}|\frac{x-\mu}{\alpha}|^{(\frac{2}{1+\beta})}]}{2^{(\frac{3+\beta}{2})}\alpha\Gamma(\frac{3+\beta}{2})},$$

where μ, $\alpha > 0$ and $\beta(-1 < \beta \leq 1)$ are location, scale and shape parameter respectively. The exponential power distribution equals normal distribution when $\beta = 0$. When $0 < \beta \leq 1$ fat tails and high peaks are obtained, with the fatness of tails increasing with β. If X follows an exponential power distribution then $E[X] = \mu$ and $V[X] = 2^{(1+\beta)}\frac{\Gamma[3(1+\beta)/2]}{\Gamma[(1+\beta)/2]}\alpha^2$.

Variance Gamma distribution
This distribution is defined as normal variance mean mixture, where the mixing density is the gamma density. It was first proposed by Madan and Seneta [21] and later on extended by Madan and Milne [20] and Seneta [29]. The variance gamma distribution capturers returns well since it allows for peakedness, fat tails and skewness. Its distribution function is of the form

$$f(x) = \frac{2\exp(\theta x/\sigma^2)}{\sigma\sqrt{2\pi}\nu^{1/\nu}\Gamma(1/\nu)}(\frac{|x|}{\sqrt{(2\sigma^2/\nu + \theta^2)}})^{1/\nu-1/2}K_{1/\nu-1/2}(\frac{|x|\sqrt{2\sigma^2/\nu + \theta^2}}{\sigma^2}), \tag{3.2}$$

where $\nu > 0$ and θ are the shape and asymmetry parameter respectively. Skewness is adjusted by θ, negative values mean negative skewness. If X is a random variable, its mean under variance gamma distributions equals $E[X] = \theta$ and its variance $V[X] = \sigma^2 + \theta^2\nu$. In the above density function $K_{1/\nu-1/2}(.)$ denotes the Bessel function of the third order, also known as Hankel function.

Mixture of two Normal distributions

We can also assume that returns are generated by a mixture of distributions. Press [26] proposed to use a combination of two normal distributions connected by a jump process, to describe the stock returns. This kind of distribution can capture large informational shocks as well as skewness. The density function of the mixture normal distributions is written as

$$f(x) = \frac{1}{\sqrt{2\pi\sigma_1^2}} \exp^{-[\frac{(x-\mu_1)^2}{2\sigma_1^2}]} \lambda + \frac{1}{\sqrt{2\pi\sigma_2^2}} \exp^{-[\frac{(x-\mu_2)^2}{2\sigma_2^2}]} (1-\lambda),$$

where μ_i and σ_i^2 ($\sigma_i > 0$) are the mean and the variance respectively. This mixture can be explained as drawing samples from a normal distribution with mean μ_1 and variance σ_1^2 with probability λ and from an independent normal distribution with mean μ_2 and variance σ_2^2 with probability $1-\lambda$. If a random variable X has a mixture normal distribution, then the mean equals $E[X] = \mu = \lambda\mu_1 + (1-\lambda)\mu_2$ and variance $V[X] = \lambda((\mu_1-\mu)^2 + \sigma_1^2) + (1-\lambda)((\mu_2-\mu)^2 + \sigma_2^2)$ as described in Aparicio and Estrada [1].

3.3 Parameter Estimation and Testing

To conclude which of the above described distributions would fit the intra-day and overnight returns best we perform a goodness of fit test and present QQ-plots. However the theoretical distributions first need to be fitted to the return data. In the table below the maximum likelihood estimations of the parameters are reported.

Table 3.2: Parameter Estimates

Distribution	Parameter	Intra-day	Overnight
N	μ	-0.000998	0.000699
	σ	0.0226	0.0123
S-t	μ	-0.000948	0.000952
	σ	0.0286	0.0064
	ν	2.5185	2.8228
L	μ	-0.0011	0.00091
	α	0.0111	0.05314
EP	μ	-0.0012	0.0012
	α	0.0076	0.0038
	β	1	1
MN	μ_1	-0.0011	0.000528
	σ_1	0.0368	0.0351
	μ_1	-0.0011	0.000843
	σ_2	0.0113	0.0074
	λ	0.3102	0.081
VG	c	-0.000924	0.00115
	σ	0.0217	0.0104
	θ	-0.000072	0.00045
	ν	1.131	1.091

Parameter estimates of the following distributions are reported, N=Normal, S-t = Scaled t, L = Logistic, EP=Exponential power, MN = Mixture Normal, VG=Variance gamma

Looking at the results from the Table 3.2 we can observe that all the estimates for the means of the intra- day returns are negative and positive for the overnight returns, consistent with the results from Table 3.1. It is a well known fact that the student-t distribution converges towards the normal distribution when the number of degrees of freedom is big, however here we can see that the estimate for the degrees of freedom is rather small, indicating fatter tails of the scaled-t distribution in comparison to normal. Similarly the normal distribution equals the exponential power distribution when $\beta = 0$, however Table 3.2 reports very high estimates of β, again implying thick tails. The parameter estimates when considering intra-day returns are rather different to the ones estimated for the overnight returns, showing that the returns exhibit different characteristics and should indeed be treated separately.

For the visual presentation of how well different distributions fit the intra-day and overnight returns, we present the plots of fitted distributions compared to the empirical distribution of returns.

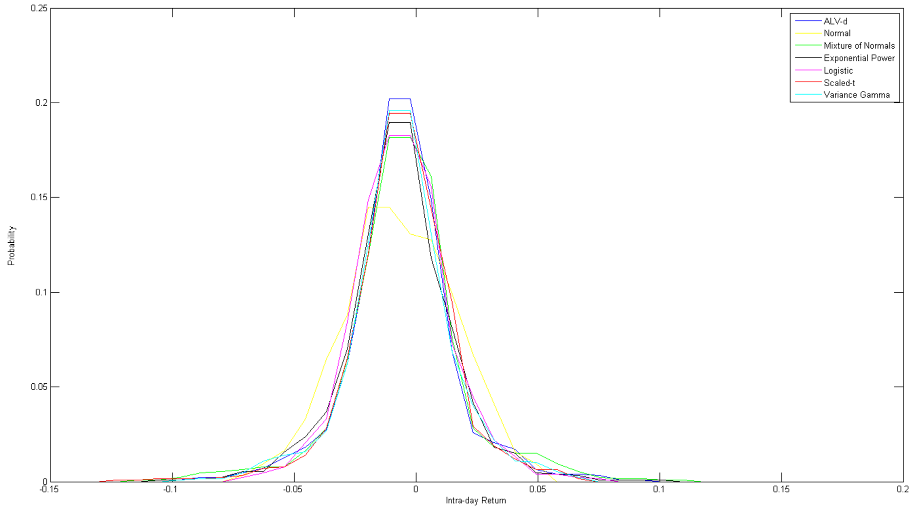


Figure 3.1: Intra-day distributions

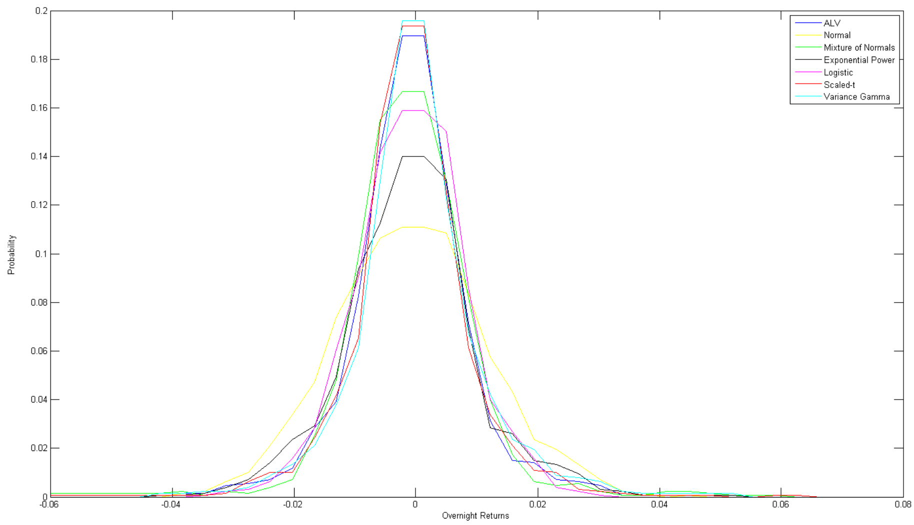


Figure 3.2: Overnight distributions

From at the first plot of intra-day distributions we can see that all distributions, besides normal capture intra-day returns rather well. There are some deviations in description of peaks and tails, where variance gamma and scaled-t distribution present the best fit. As expected the normal distribution does not manage to capture the peak behavior.

In comparison to the previous plot the distributions did not manage to capture the overnight return data well. They had problems especially with description of the peak, where only scaled-t and variance gamma distribution performed well. Again the normal distribution presented the poorest fit.

3.3.1 Goodness of Fit Test

In order to compare the relative fit of the described theoretical distributions, Pearson's Chi Squared Goodness of Fit Test is considered. This test is conducted in the following way. First, observations are divided in k non-overlapping intervals, then the probabilities of the outcomes to fall in intervals by the distribution under H_0 are calculated. The test statistics equals:

$$X^2 = \sum_{i=1}^{k} \frac{(n_i - Np_i)^2}{Np_i},$$

where p_i and n_i stand for the probability of being in partition i and number of observations in partition i respectively, N is the total number of observations ($\sum_{i=1}^{k} n_i = N$). The calculated statistic is compared to the chi-squared distribution with $(k - ep - 1)$ degrees of freedom and chosen significance level, where ep denotes the number of estimated parameters.

Despite of the wide use of Chi Squared Goodnes of Fit test, it is still not clear how many partition points should be used and whether the partitions should be equiprobable or not. Based on Stuart, Ord and Arnold [32] we decided to use equiprobable partitions. Literature recommends to use a rather large number of partitions (Mann and Wald [23]), however this drastically reduces the power of the test, therefore we use 8 partitions.

In the table below we report the results of the test, run on all considered distributions. The zero hypothesis of the test is that the returns are governed by the assumed distribution, against the alternative (returns are not governed by the assumed distribution). The significance level used in the test is 0.01.

Table 3.3: Chi Squared Test

	Intra-day return		Overnight return	
Distribution	Test	P-value	Test	P-value
N	1	0.0000	1	0.0000
S-t	0	0.233	0	0.6829
L	1	0.0000	1	0.0000
EP	1	0.0000	1	0.0000
MN	1	0.0002	1	0.0053
VG	0	0.013	1	0.0072

Results of the Chi Squared Test, if 1 is written for the test, that means that the null hypothesis was rejected, if there is 0 the hypothesis was not rejected.

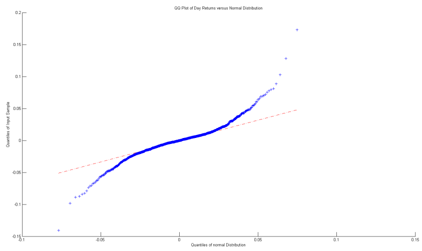
As expected the goodness of fit test rejects the normal distribution, for both intra-day and overnight returns. Despite the fact that the logistic, exponential power and mixture normal distributions are capable of capturing some of the characteristics of returns, the test still rejects the hypothesis that returns are generated by any of those distributions. The results show support for the variance gamma distribution that was not rejected for intra-day returns. However the best fit for both returns is the Scaled-t distribution.

Since Chi Squared Goodness of Fit tests can be biased due to the number of chosen partitions, we also report the QQ-plots before concluding which distributions describe intra-day and overnight returns best.

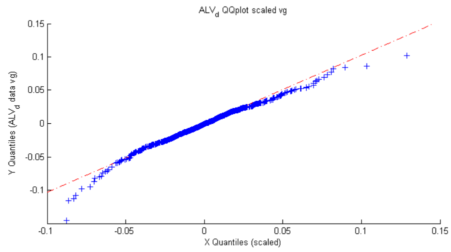
3.3.2 QQ-plots

The quantile-quantile (Q-Q) plot is a graphical technique for determining if two data sets come from the same family of distributions. A QQ-plot is a plot of quantiles of the first data set against the second data set, where quantile is a point below which a given fraction of points lies. Since we are interested whether the returns come from any of the above described distributions, we plot quantiles of returns against quantiles obtained from the random sample from each of the proposed distributions. Here we only report the QQ-plots of normal, scaled-t and variance gamma distributions, since other distributions were clearly rejected by the goodness of fit test, they are however presented in the Appendix. QQ-plots for intra-day and overnight returns are presented separately.

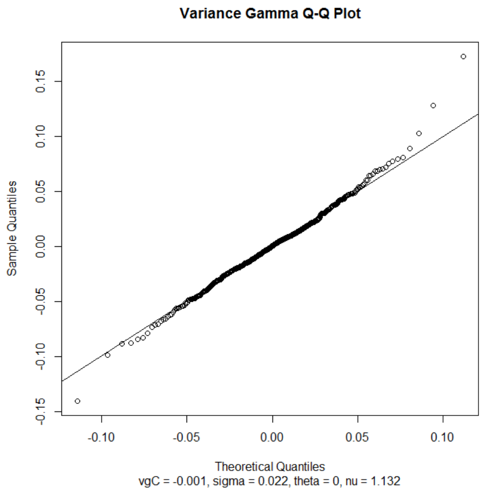
Figure 3.3: QQ-plots: Intra-day returns



(a) Normal Distribution



(b) Scaled-t Distribution

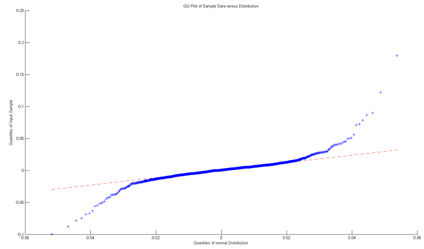


(c) Variance Gamma Distribution

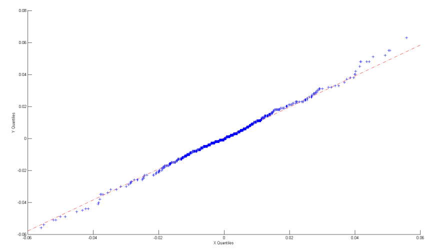
From the QQ-plots we can see that scaled-t and variance gamma distribution present a much better fit to the intra-day returns than the normal distribution. As expected normal distribution fits poorly in the tails. Both the scaled-t and the variance gamma distribution describe tail behavior rather well, however the second captures the left tail behavior better. In the right tail there is some imprecision in both distributions, but the variance gamma distribution fits to the red line longer. From the QQ-plots we can conclude that the variance gamma distribution fits intra-day returns best.

In the plots below, similarly as before the performance of the normal distribution is very poor in comparison to the alternative. The scaled-t and the variance gamma distribution both describe overnight returns well, however the variance gamma does not manage to capture the left tail behavior, therefore we conclude that the scaled-t distribution fits overnight returns best.

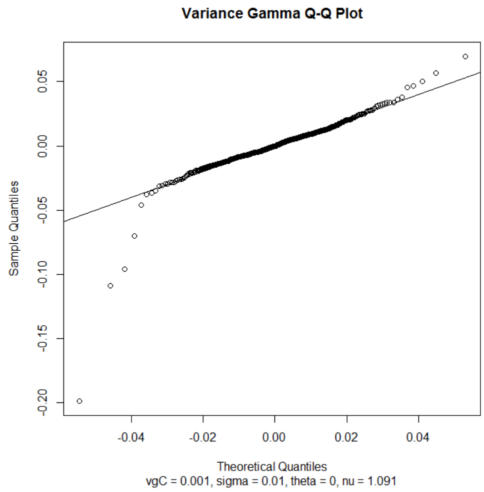
Figure 3.4: QQ-plots: Overnight returns



(a) Normal Distribution



(b) Scaled-t Distribution



(c) Variance Gamma Distribution

The Chi Squared Goodness of Fit Test reports different results from QQ-plots regarding which distribution fits intra-day and overnight returns best. We decide to follow the QQ-plots, since the Goodness of Fit Test is dependent on the number of partitions we choose. Therefore we conclude that the variance gamma distribution describes intra-day returns best, while scaled-t distribution captures overnight jumps most accurately.

Similar research as reported here was done on all stock in DAX index for the same period of time as described for Allianz stock. The results were more or less consistent with results obtained for Allianze stock. For most intra-day returns variance gamma was found to capture them best, while the overnight returns were best described with scaled-t distribution. In the appendix, the similar results as for the Allianz stock are reported for BMW and Deutsche Bank stock.

3.4 Independence of Intra-day and Overnight Returns

The research studies done in this area indicate that small negative correlation between intra-day and overnight returns exists (Wang, Shieh, Havlin and Stanley [33]). For example when overnight returns are small the intra-day returns are likely to be bigger and the other way around. When the overnight returns are high the reason of low intra-day returns is usually the inflation in the stock price at the opening of the exchange.

The SVRJOJ model assumes that intra-day and overnight returns are uncorrelated by implementing independence between the Brownian Motion W_t^S, the Poisson process N_t, the random jumps Y_i and the overnight jumps V_i. In this part we examine whether this assumption holds for intra-day and overnight returns of Allianz stock and whether introducing dependence would lead to a more accurate model. The independence between overnight (OR_{t-1}) and intra-day (DR_t) as well as intra-day (DR_t) and overnight (OR_t) returns is tested. To evaluate whether the returns are correlated we test if the correlation matrix of overnight and intra-day returns is an identity matrix, the test used is a Chi Squared Barlett and Box test. For a visual presentation scatter plots are reported.

3.4.1 Bartlett-Box Test

With Bartlett-Box test we test whether a correlation matrix R is an identity matrix. The test is based on the proposition made by Bartlett [4], that $Y = -\log(\det(R)(N-1-(2p+5)/6))$ is chi squared distributed if R is an identity matrix, where N equals the number of rows or columns in a correlation matrix and p is a p-value. The H_0 of the test: R is an identity matrix. If the test statistics Y exceeds the critical value $\chi^2_{1-p,N-1}$, then the H_0 is rejected at significance level p.

Table 3.4: Bartlett-Box Test

Returns	Correlation	Test Stat.	P-value
$OR_{t-1} - DR_t$	0.0477	2.889	0.0891
$DR_t - OR_t$	0.0466	2.758	0.097

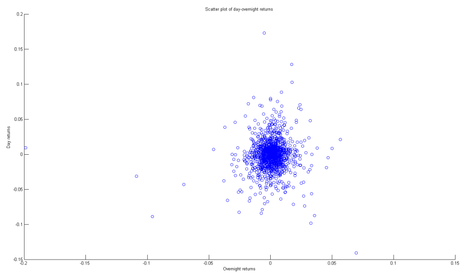
Summery of the Bartlett-Box test. With $OR_{t-1} - DR_t$ we denote the correlation between overnight returns and the intra-day returns next day, similarly $DR_t - OR_t$ denotes the correlation between returns on the same day.

The results of the test do not reject the null hypothesis for the significance level $p = 0.05$ that we choose. Based on that result we will assume that the returns are independent of one another, although the hypothesis not being rejected does not mean the hypothesis is being correct. The table reports very low positive correlation between returns, contradicting the negative values usually denoted in the literature.

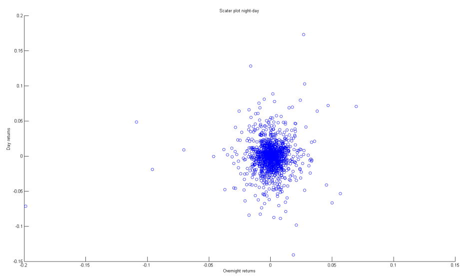
3.4.2 Scatter Plots

Scatter plots show the relationship between two pairs of variables, by plotting them against each other on a two dimensional graph. They are often used for investigating whether two variables are correlated. The variables are positively correlated when the values increase together from left to right and are negatively correlated when they decrease from left to right. However the positive correlations in our case are so small that we can not see the patterns on the scatter plots.

Figure 3.5: Scatter plots of intra-day and overnight correlations



(a) $DR_t - OR_t$



(b) $OR_{t-1} - DR_t$

Scatter plots do not exhibit a clear pattern, therefore we assume they are not correlated and confirm the results we obtained from the Barlett-Box test. This result was expected, due to very low correlation values.

Chapter 4

Extending the SVRJOJ Model

The SVRJOJ model assumes that intra-day and overnight returns follow a normal distribution. However as shown above the normal distribution does not manage to capture the returns well, since it suggests that extreme returns occur less often than they do in the real world. In order to describe the stock price behavior better, other distributions should be considered. From the previous section we can conclude that the variance gamma distribution describes intra-day returns best, while the scaled-t distribution captures the overnight behavior most accurately.

In this part we present three extensions of the SVRJOJ model. The first extension is the model that combines the variance gamma process for intra-day returns with the normal overnight jumps, second is the Bates model with the scaled-t overnight jump and in the last we combine the variance gamma process with the scaled-t overnight jumps.

4.1 Variance Gamma Process

The variance gamma process is a Lévy process obtained by evaluating Brownian Motion at a random time change given by a gamma process (see [19]). It was first introduced by Madan and Seneta [21] and later on extended by Madan, Carr and Chang [19] and Seneta [29]. The model presents two additional parameters, compared to the geometrical Brownian Motion, which control the asymmetry (θ) and shape (ν) of the distribution of returns.

Under the variance gamma process the stock prices evolve over time $t \geq 0$ as

$$S_t = S_0 \exp(rt + \theta T_t + \sigma W(T_t)) = S_0 \exp(rt + X_t), \tag{4.1}$$

where θ and $\sigma \geq 0$ are constants. The time at which the Brownian motion is evaluated is a gamma process and increments $g = T_{t+h} - T_t$ have the following distribution, for $g > 0$

$$f_h(g) = (\frac{\theta}{\nu})^{(\frac{\theta^2 h}{\nu})} \frac{g^{\frac{\theta^2 h}{\nu}-1}\exp(-\frac{\theta}{\nu}g)}{\Gamma(\frac{\theta^2 h}{\nu})},$$

where $\Gamma(.)$ denotes the gamma function. The log returns corresponding to the stock price process can be written as

$$X_t = \theta T_t + \sigma W(T_t), \tag{4.2}$$

with a density function defined as (3.2). Since a density function of X_t is rather complex, its characteristic function is used instead whenever possible, it has the following form

$$\phi_X(u) = E[\exp(iuX_t)](1 - i\theta\nu u + (\sigma^2\nu u^2)/2)^{-t/\nu}, \tag{4.3}$$

for $-\infty < u < \infty$.
To show how parameters of the variance gamma process influence its behavior we present its first four centralized moments. The calculations are omitted here, however they can be found in Madan, Carr and Chang [19]. The moments are equal to

$$E[X_t] = \theta t,$$
$$V[X_t] = (\sigma^2 + \theta^2\nu)t,$$
$$E[(X_t - E[X_t])^3] = (2\theta^3\nu^2 + 3\sigma^2\theta\nu)t$$
$$E[(X_t - E[X_t])^4] = (3\sigma^4\nu + 12\sigma^2\theta^2\nu^2 + 6\theta^4\nu^3)t + (3\sigma^4 + 6\sigma^2\theta^2\nu + 3\theta^4\nu^2)t^2.$$

From the calculated moments we can see that θ and ν are not themselves the skewness and kurtosis of the variance gamma process, however have a big influence on those moments. The equation for skewness shows that skewness has the same sign as θ. When $\theta = 0$ there is no skewness and the variance of the process equals $\sigma^2 t$.

Since options are priced under the risk neutral measure we are interested in the stock price which follows a variance gamma process under the risk neutral measure. As mentioned before stock prices discounted at the risk free rate are martingales under the risk free measure. Therefore the stock price process follows

$$S_t = S_0 \exp((rt + X_t(\sigma, \theta, \nu) + \omega t), \quad t > 0, \tag{4.4}$$

where by setting $\omega = (1/\nu)\log(1 - \theta\nu - \frac{1}{2}\sigma^2\nu)$, the mean rate of return on the stock equals the interest rate r, hence the stock process is a martingale after discounting. We calculate ω by setting $wt = -\log(\phi_{X_t(-i)})$.

4.2 Variance Gamma Model with Normal Overnight Jump

In this part we add an overnight jump to the variance gamma process in order to capture the behavior of stock returns as close as possible. The stock price of this model under measure $\mathbb{Q}$ is written as

$$S_t = S_0 \exp(rt + X_t(\sigma, \theta, \nu) + \omega t) \prod_{i=1}^{\lfloor 252t \rfloor} (V_i), \quad t > 0, \tag{4.5}$$

where V_i follows (2.9). In order for the above equation to be a martingale we set $\omega = (1/\nu) \log(1 - \theta\nu - \frac{1}{2}\sigma^2\nu)$. Here ω is the same as in the variance gamma process without the overnight jump, this follows since the expected value of the overnight jump with distribution (2.9) is equal to 1, as we showed in the Section 2.2.

4.3 SVRJOJ Model with Scaled-t Overnight Jump

The stock price in this model has the same process for intra-day returns as the SVRJOJ model, however overnight returns are no longer normally distributed. We exchange the normal distribution with the student-t distribution, since it offers a better fit to our overnight returns data. The model under the $\mathbb{Q}$ measure is written as

$$\frac{dS_t}{S_{t-}} = r dt + \sigma_t dW_t^S + \sum_{i=1}^{N_t} (Y_i - 1) - \lambda \mu_{RJ} dt + d \sum_{i=1}^{\lfloor 252t \rfloor} (V_i - 1), \quad t > 0. \tag{4.6}$$

where the distribution of $\log V_i$ equals

$$\log(\hat{V}_i) \sim SC(\mu, \sigma^2), \tag{4.7}$$

$$\log(V_i) = \log(\hat{V}_i) - \log(\Phi_{\log \hat{V}_i}(-i)), \tag{4.8}$$

where SC is denoted as a scaled-t distribution. In comparison to the SVRJOJ model where we only simulate from the normal distribution to obtain the samples for the $\log V_i$, here also a part $\log(\phi_{\hat{V}_i}(-i))$ is added. This is the case, since for the scaled-t distribution it no longer holds that the expected value of the exponentiated draw from $St(-\frac{1}{2}\sigma_{OJ}^2, \sigma_{OJ}^2)$ equals to 1. In order for (4.6) to be a martingale $\log(\phi_{\hat{V}_i}(-i))$ is added. The intra-day part of the process (4.6) is a martingale as shown for the SVRJOJ model.

4.4 Variance Gamma Model with Scaled-t Overnight Jump

This model should capture the behavior of the stock price best, since it includes the variance gamma distribution and the scaled-t distribution, which were the best fit for intra-day and overnight returns respectively. In this case the stock price process under measure $\mathbb{Q}$ is equal to equation (4.5), however the overnight jump no longer follows (2.9), but is written as (4.7). The ω remains the same as in VG-N, since the intra-day component remains the same and the overnight jump component is already written in such a form that the VG-SC is a martingale under measure $\mathbb{Q}$.

Chapter 5

Option Pricing

As mentioned in Chapter 2 options are priced in a risk-neutral framework. The price for the European call option at time is calculated as

$$C_t = \exp(-r(T-t))E_t^{\mathbb{Q}}[\max(S_T - K, 0)],$$

where T is the maturity of an option and K is its strike price. For some stock pricing models (for example Black-Scholes) there exists a closed form solution for the option price, although for several other models this is no longer the case. For those models option price can be estimated with Monte Carlo simulation, however this approach is very computationally demanding.

The solution to this problem was presented in a paper of Stein and Stein [31], who suggested rewriting the option price in terms of a characteristic function of the logarithm of the terminal stock price, using Fourier transforms. The main advantage of this method is that it is general, it applies to any option pricing model as long as the terminal stock price has an explicitly known characteristic function. This approach is considerably faster than the methods described in Heston [16] and Stein and Stein [31], therefore it is very convenient for model calibration.

Due to its several advantages the Fourier transforms and characteristic functions are used when pricing options in all the models we consider, even when there exists a closed form solution. In our implementation of this approach we follows the article [14]. The following formula proposed by Scott [28] is used for valuation of the call option

$$C(t,T) = S_t P_1 - \exp(-r(T-t))KP_2, \tag{5.1}$$

where

$$\begin{aligned} P_1 &= \int_K^\infty \frac{S_T}{E_t^{\mathbb{Q}}[S_T]} p(S_T) dS_T, \\ P_2 &= P(S_T > K). \end{aligned}$$

For most models the probability density function for the terminal stock price is unavailable, therefore Fourier inversion techniques are used to derive the formulas for P_1 and P_2.
The Fourier transform ($\mathcal{F}\{.\}$) and the inverse Fourier transform ($\mathcal{F}^{-1}\{.\}$) for an integrable function are given as

$$\mathcal{F}\{f(x)\} = \int_{-\infty}^{\infty} \exp(iux) f(x) dx = \phi(u),$$
$$\mathcal{F}^{-1}\{\phi(u)\} = \frac{1}{2\pi} \int_{-\infty}^{\infty} \exp(-iux)\phi(u) du = f(x),$$

where $f(x)$ stands for risk-neutral density of log-returns of the stock price process. With ϕ we denote the characteristic function.

Here we only report the values for P_1 and P_2, however the exact calculations can be found in Bakshi and Madan [2]

$$P_1 = \frac{1}{2} + \frac{1}{\pi} \int_0^{\infty} Re\left(\frac{\exp(-iu\log(K))\phi(u-i)}{iu\phi(-i)}\right) du, \tag{5.2}$$
$$P_2 = \frac{1}{2} + \frac{1}{\pi} \int_0^{\infty} Re\left(\frac{\exp(-iu\log(K))\phi(u)}{iu}\right) du, \tag{5.3}$$

where $\phi(u)$ equals the characteristic function under $\mathbb{Q}$, conditioned on t, of the logarithm of the terminal stock price $\log(S_T)$ and Re indicates that we only integrate over the real part of the complex number. Due to the definition of characteristic functions we can also write $\phi(u) = E_t^{\mathbb{Q}}[\exp(iu\log S_T)]$. When pricing options, we evaluate the equations 5.2 and 5.3 numerically using Gauss Lobatto integration scheme.

5.1 The SVRJOJ model

Here we show how the option price in the SVRJOJ model is obtained. The calculations follow directly from [10].

The Fourier transformation approach to option pricing is based on the logarithm of the terminal stock price and its characteristic function, therefore we first use *Itô*'s Lemma and transform the stock price process (2.7) into

$$\log(S_t) = log(S_0) + \int_{u=0}^{t} (r - \frac{1}{2}\sigma_u^2) du + \int_{u=0}^{t} \sigma_u W_u^S + (\sum_{i=1}^{N_t} log(Y_i)) - \lambda\mu_{RJ} t + (\sum_{i=1}^{\lfloor 252t \rfloor} log(V_i)), \tag{5.4}$$

where $-\frac{1}{2}\sigma_t^2$ in the first part of the equations 5.4 follows from the transformation of the S_t into $\log(S_t)$. When we take a logarithm of an overnight part of the stock price, the term $\sum_{i=1}^{\lfloor 252t \rfloor}(V_i - 1)$ it is transformed to $\sum_{i=1}^{\lfloor 252t \rfloor} log(V_i)$.

The characteristic function of $\log S_T$ can then be written as

$$\begin{aligned}
&\phi(u) = E_t^{\mathbb{Q}}[\exp(iu \log S_T)] = \\
&= E_t^{\mathbb{Q}}[\exp(iu(\log S_t + r\tau - \frac{1}{2}\int_t^T \sigma_u^2 du + \int_t^T \sigma_u dW_u^S) + \\
&+ \sum_{i=N_t+1}^{N_T} \log Y_i - \lambda\mu_{RJ}\tau + \sum_{\lfloor 252t \rfloor+1}^{\lfloor 252T \rfloor} log(V_i))] = \\
&= E_t^{\mathbb{Q}}[\exp(iu(\log S_t + r\tau - \frac{1}{2}\int_t^T \sigma_u^2 du + \int_t^T \sigma_u dW_u^S))] \times \\
&E_t^{\mathbb{Q}}[\exp(iu(\sum_{i=N_t+1}^{N_t} \log Y_i - \lambda\mu_{RJ\tau}))] E_t^{\mathbb{Q}}[\exp(iu \sum_{\lfloor 252t \rfloor+1}^{\lfloor 252T \rfloor} log(V_i)],
\end{aligned}$$

where $\tau = (T - t)$. The characteristic functions for each part can be calculated separately, due to the property of characteristic functions that states $\phi(XY) = \phi(X)\phi(Y)$, when X and Y are independent random variables. Since the model assumes W_t^S, N_t, Y_i and V_i to be independent of one another under $\mathbb{Q}$ this rule also applies here.

Solving the first expectation is equal to calculating the characteristic function of the Heston model. In literature there are several different approaches, which all lead to the same result.

$$\begin{aligned}
&E_t^{\mathbb{Q}}[\exp(iu(\log S_t + r\tau - \frac{1}{2}\int_t^T \sigma_u^2 du + \int_t^T \sigma_u dW_u^S))] = \\
&= \exp(C(\tau, u) + D(\tau, u)\sigma_t^2 + iu \log S_t),
\end{aligned} \tag{5.5}$$

where

$$\begin{aligned}
C(\tau; u) &= riu\tau + \frac{\kappa\sigma^2}{\sigma_\sigma^2}((\kappa - \rho\sigma_\sigma iu + d)\tau - 2\log\frac{1 - ge^{d\tau}}{1 - g}), \\
D(\tau; u) &= \frac{\kappa - \rho\sigma_\sigma iu + d}{\sigma_\sigma^2}\frac{1 - e^{d\tau}}{1 - ge^{d\tau}}
\end{aligned}$$

and g and d are expressed as

$$\begin{aligned}
g &= \frac{\kappa - \rho\sigma_\sigma iu + d}{\kappa - \rho\sigma_\sigma iu - d}, \\
d &= \sqrt{(\rho\sigma_\sigma iu - \kappa)^2 + \sigma_\sigma(iu + u^2)}.
\end{aligned}$$

Now the characteristic function for the random jump is calculated. If we multiply the characteristic function of the Heston model with the characteristic function of the random jump we obtain the characteristic function of the log stock price of the Bates model.

$$E_t^{\mathbb{Q}}[\exp(iu(\sum_{i=N_t+1}^{N_t} \log Y_i - \lambda\mu_{RJ\tau}))] =$$
$$= \exp(\lambda\tau((1+\mu_{RJ})^{iu}\exp((\frac{iu}{2})(iu-1)\sigma_{RJ}^2) - 1) - iu\lambda\mu_{RJ}\tau). \quad (5.6)$$

As last we calculate the characteristic function of the overnight jump, which follows straight from the characteristic function of the normal distribution.

$$E_t^{\mathbb{Q}}[\exp(iu \sum_{\lfloor 252t \rfloor+1}^{\lfloor 252T \rfloor} log(V_i)] = \exp(-\frac{1}{2}u(u+i)n\sigma_{OJ}^2/252), \quad (5.7)$$

where $n = \lfloor 252T \rfloor - \lfloor 252t \rfloor$.
To calculate an option price of the SVRJOJ model all three characteristic functions are multiplied to obtain the characteristic function of the log price of the stock price process. From that the integrals (5.2) (5.3) are calculated and inserted into the equation (5.1) to get the option price.

5.2 The Variance Gamma model with a Normal Overnight Jump

The approach to option price calculation of this model is the same as the described before for the SVRJOJ model. First we calculate the log stock price process under the risk neutral measure, which follows the process (4.5). Since there is no Brownian Motion in this equation we do not have to use *Itô*'s lemma and therefore the logarithm of the stock price equals

$$\log(S_t) = \log(S_0) + (r+w)t + X_t(\sigma,\theta,\nu) + \sum_{i=1}^{\lfloor 252t \rfloor} log(V_i), \quad (5.8)$$

where $V_i - 1$ was exchanged with $\log V_i$ based on the reasoning explained in the previous model. Since overnight and intra-day returns in this model are assumed to be independent, we calculate their characteristic function separately. The characteristic function of intra-day returns follows

$$\phi_{S_T}(u) = \exp(\log(S_0)iu + (r+\omega)iuT)(1 - i\theta\nu u + \frac{1}{2}\sigma^2 u^2 \nu)^{-T/\nu}, \quad (5.9)$$

where u is an argument of the characteristic function and ω is calculated as shown in [19]. The characteristic function of log stock returns in this model is obtained by multiplying (5.7) and (5.9). To calculate the option price the same steps as in the previous part should be taken.

5.3 The Variance Gamma Model with the Scaled-t Overnight Jump

In this model the logarithm of the intra-day stock price is the same as for the variance gamma model with the normal overnight jump and therefore follows equation (5.8). The characteristic function for an intra-day stock prices is still equal to (5.9), however the characteristic function of an overnight part changes, since the logarithm of the jump is scaled-t distributed.

The characteristic function of a scaled-t distribution with p degrees of freedom is calculated from the characteristic function corresponding to the student-t distribution, which equals

$$\phi_Y(u) = \frac{K_{\frac{p}{2}}(\sqrt{p}|u|)(\sqrt{p}|u|)^{p/2}}{\Gamma(\frac{p}{2})2^{p/2-1}},$$

where $K_{\frac{p}{2}}$ is the modified Bessel Function of the second kind. To obtain the characteristic function of Scaled-t distribution we use the following property. If x is scaled-t distributed then $y = \frac{x-\mu}{\sigma}$ is student-t distributed, therefore the characteristic function of scaled-t distribution can be calculated as $\phi_X = E[\exp(iu(Y\sigma+\mu))]$ which equals

$$\phi_X(u) = \exp(iu\mu)\phi_Y(u\sigma).$$

This is however only a characteristic function of $\log(V_i)$ and we would like to find the characteristic function of $\sum_{i=1}^{\lfloor 252t \rfloor} \log V_i$. Due to the property of characteristic function which states that characteristic function of a sum of independent random variables equals the product of characteristic functions

$$\phi_{\sum_{i=1}^{\lfloor 252t \rfloor} \log V_i}(u) = (\phi_{log(V_i)}(u))^{\lfloor 252t \rfloor}.$$

The calculated characteristic function is only a characteristic function of the logarithm of the scaled-t distributed overnight jump, however we still have to correct the jump part in order for the Variance Gamma Model with Scaled-t Overnight Jump (VG-SC) to be a martingale under measure $\mathbb{Q}$ after discounting, hence the overnight jumps should equal to 1 under expectation. This is achieved by setting the characteristic function of the overnight jump to

$$\phi(u) = \phi_{\sum_{i=1}^{\lfloor 252t \rfloor} \log V_i}(u) \exp(-iu \log(\phi_{\sum_{i=1}^{\lfloor 252t \rfloor} \log V_i}(-i))). \tag{5.10}$$

The characteristic function of the log stock prices following a VG-SC model is now obtained by multiplying (5.10) with (5.9), since it is assumed that the intra-day and overnight stock price process are independent.

5.4 The SVRJOJ model with Scaled-t Overnight Jump

The logarithm of stock price process for this model equals equation (5.4), where the overnight jump is distributed as (4.8). The characteristic function for the intra-day returns in this model is obtained by multiplying (5.5) with (5.6) and the characteristic function of overnight returns equals (5.10).

Chapter 6

Option Data and Model Estimation

To estimate the parameters of the option pricing models, option data is needed as well as the objective function based on which we minimize the error of an option pricing model. First the option data is described; later on we elaborate on estimation and on which objective function to use.

6.1 Data

We use American option prices on the Allianz (ALV) stock between October 1, 2010 to October 14, 2010, without the weekends this gives us 10 days of observation. The data was provided by All Options and is based on the internal option pricing model, which calculates the option prices that are close to the mid market option price at the end of each trading day. Each day option prices for the following four maturities are given: October 15, 2010, November 19, 2010, December 17, 2010 and March 18, 2011. Next to option prices and maturities also strike price, risk free rate and stock price at the end of the trading day are reported.

It is never optimal to exercise an American call option on a non-dividend paying stock early, therefore American and European call options have the same price when there are no dividends until the expiration of an option [18]. Since when the call is in the money and investor plans to hold the stock even after the expiration it is better to hold a call until maturity in order to protect himself/herself from the downside risk as well as earn interest on the strike price between now and the maturity. On the other hand when the investor believes that the stock is overpriced and considers exercising the call and selling the stock immediately, he/she should sell the call instead. This holds due to inequality $C \geq S_0 - K\exp(-rT)$ proven by [18], where C denotes price of an American call option.

During October 1, 2010 to March 18, 2011 there were no dividends on the Allianz SE stock, therefore American options can be used to calibrate models with which we price European options.

Following Bakshi, Charles and Zhiwu [3] Table 5 provides descriptive statistics of call options that satisfy the criteria:

- Options with less then six days to expiration are excluded from the sample, since they might have liquidity related biases.

- Price quotes lower then 0.02 € are not included, due to very high relative errors obtained when using a squared percentage error objective function.

- Quotes for which the following arbitrage restriction does not hold are excluded, since our models assume arbitrage is not possible

$$C(t,T) \geq \max(0, S_t \exp(T-t) - K \exp(-r(T-t))).$$

In the literature they usually also exclude options with too big bid-ask spread, however the bid-ask spreads were not available for our research.

Table 6.1: Option data overview

Day return		Overnight return	
S/K	< 60	60 − 180	Subtotal
OTM (< 0.97)	0.378 (79)	1.253 (172)	 (251)
ATM (0.97 − 1.03)	2.277 (47)	4.4256 (57)	 (104)
ITM (> 1.03)	11.227 (118)	13.716 (210)	 (328)

The reported numbers are average call option prices and in the brackets number of options in each category is reported.

The option data in the table is divided based on the moneyness and time to expiration. The call option is said to be out-of-the-money (OTM) if $S/K \leq 0.97$, at-the-money (ATM) if $0.97 \leq S/K \leq 1.03$ and in-the-money if $S/K \geq 1.03$. By maturity options can be classified as short term (< 60 days) and mid-term (60-180 days). Altogether there are 683 call option observations, with OTM and ITM options amounting to 37% and 48% respectively. In our research we only use OTM options for consistency with [10].

6.2 Estimation

When one wishes to price an option using an option pricing model the strike price and the maturity are specified in the contract, the spot stock price and risk free rate can be taken from published data. However spot volatility and structural parameters are unobserved and need to be estimated. These estimates can be obtained by using maximum likelihood or a generalized method of moments, but this approach might be inconvenient due to large historical data set requirements. Therefore practitioners and academics usually use option-implied volatility given by the model. This approach leads to significant improvement of different option pricing model's performances [5] and [6].

Estimation Procedure
We adopt the estimation procedure proposed by [3], where parameters under the risk neutral measure $\mathbb{Q}$ are estimated as:

Step 1: For each day t N options on the same stock from the same point in time are chosen and their closing price is observed. The number of observed prices should be equal or greater to one plus the number of estimated parameters. For every $i = 1, 2, ..., N$ τ_i represents the time to maturity and K_i the strike price. With $\hat{C}_i(t, \tau_i, K_i)$ we denote the observed option price and with $C_i(t, \tau_i, K_i)$ the one calculated by one of the proposed models. For example in the SVRJOJ model the difference between $\hat{C}_i(t, \tau_i, K_i)$ and $C_i(t, \tau_i, K_i)$ depends on implied volatility σ_t and the parameter vector is denoted by $\Phi = (\mu_{RJ}, \sigma_{RJ}, \lambda, \sigma_{OJ}, \kappa, \sigma, \sigma_\sigma, \rho)$. For each i the error of the pricing model is defined as

$$\epsilon_i[V(t), \Phi] = \hat{C}_i(t, \tau_i, K_i) - C_i(t, \tau_i, K_i).$$

Step 2: Different objective functions can be used to find the parameters for which the model has the smallest error. We will use the objective function that was considered in [10]

$$SSE(t) = \min_{\sigma_t, \Phi} \sum_{i=1}^{N} (\frac{\epsilon_i[\sigma_t, \Phi]}{C_i(t, \tau_i, K_i)})^2. \tag{6.1}$$

This is called the squared percentage error objective function and it assigns more weight to the OTM options. In the literature [3] also the squared relative objective function is considered, which gives more weight to relatively expensive options (ITM options and options with long maturity).

With this approach we obtain different set of parameters for each model each day. Later on we only report their average. Since parameters are estimated separately time variation of the parameters is not excluded. However our research only includes 10 days of option data. Therefore time varying parameters are not required.

Chapter 7

Empirical Results

In this Chapter we report the estimation results by applying the described estimation techniques to all of the described models. For an easier comparison the models are divided in two groups. In first are the models based on the stochastic volatility and in the second the ones containing the variance gamma process. The parameter estimates obtained in our research are compared to the existing literature.

7.1 Models Based on Stochastic Volatility

All the models with exception of SVRJOJ-SC considered in this section were already discussed in the literature. The Heston and Bates models were evaluated by [2] and [10], the latter also introduced and evaluated the SVOJ and SVRJOJ model. Their model estimation techniques are very similar to the ones used in this research, however the options they use are written on the $S\&P500$ index, while we use the options based on Allianz to calibrate the models. In the article [2] they estimate the models based on all available option prices and the at the money options, while in [10] parameters are estimated using out of the money options. When estimating the parameters we use squared percentage error objective function, which is also the case for [10], however in [2] they apply absolute squared errors objective function. Since the techniques used in [10] are more in line with our research, we assume parameter estimates form [10] should offer a better comparison for our estimates.

Table 7.1: Parameter estimates: Models with stochastic volatility

	BS	Heston	Bates	SVOJ	SVRJOJ	SVRJOJ-SC
μ_{RJ}			-0.4352		-0.1571	-0.1532
			(21.26%)		(25.31%)	(6.01 %)
σ_{RJ}			0.1906		0.0578	0.1237
			(20.53%)		(6.68%)	(5.81 %)
λ			0.0968		0.1024	0.3976
			(0.15)		(0.13)	(0.11 %)
μ_{OJ}						-0.1499
						(0.122.73 %)
σ_{OJ}				0.1329	0.0951	0.0624
				(5.97%)	(5.71%)	(1.71 %)
κ		1.8909	13.1319	1.5319	3.8419	0.3302
		(1.65)	(2.31)	(3.39)	(5.15)	(0.26)
σ		0.2273	0.2072	0.2801	0.1779	0.2331
		(0.68%)	(1.39%)	(11.92%)	(5.01%)	(1.85%)
σ_σ		1.3855	0.56486	0.4488	0.52	0.5986
		(59.71%)	(40.07%)	(22.56%)	(41.02%)	(15.92 %)
ρ		-0.3567	-0.4946	-0.7256	-0.6569	-0.1444
		(0.08)	(0.44)	(0.25)	(0.31)	(0.15)
σ_t	0.2102	0.1886	0.1953	0.1348	0.1566	0.2200
	(0.48%)	(4.5%)	(3.26%)	(5.21%)	(5.45%)	(0.59 %)
p						3.8810
						(14.52 %)
SSE	0.2376	0.0432	0.0314	0.0383	0.0303	0.0267

The values reported as the parameter estimates are the mean of 10 parameter estimations done for each day separately, in the brackets standard deviations of the parameters are reported. In SSE we report the mean of the objective function over 10 days.

Before commenting on the obtained results it is important to point out that the reported estimates for all of the parameters are actually average values, since we estimate the parameters of each model every day separately. When taking the average of the estimated parameters to calculate the sum of objective functions over 10 days, the result would no longer equal SEE, since the SEE is obtained from the actual parameter estimates. The results are reported in this way, due to easier comparison between parameter estimates in different models and consistency with the existing literature.

Let us first consider the instantaneous volatility (σ_t^2) estimates reported in the Table 7.1. We can see that the highest value is obtained when evaluating the

SVRJOJ-SC model followed by Black Scholes, Bates and Heston model, smaller values are reported for SVOJ and SVRJOJ model. This result is intuitively correct, with the exception of the SVRJOJ-SC and Bates model. Since the total variation in the SVRJOJ model is divided to the variation of the random and overnight jump as well as the variation coming from the stochastic volatility part of the model, therefore the instantaneous volatility can be smaller. These results are to some extent consistent with the literature, however our instantaneous volatility estimates are higher compared to those in [10], which can be explained due to different option data used when estimating the models.

In order to show the complete variance decomposition of the models we calculate the variation of the log-return due to random jump

$$Var(\sum_{i=1}^{N_{t+1}-N_t} \log Y_i) = \lambda\sigma_{RJ}^2 + \lambda(\log(1+\mu_{RJ}) - \frac{1}{2}\sigma_{RJ}^2)^2.$$

The variation of the overnight part is already reported in the σ_{OJ}. In the following table we present the variance decomposition in the Bates, SVOJ, SVRJOJ and SVRJOJ-SC model.

Table 7.2: Variance decomposition

	Bates	SVOJ	SVRJOJ	SVRJOJ-SC
Continuous Part	0.0381	0.0182	0.0245	0.0484
Random Jump Part	0.0378		0.0034	0.0181
Overnight Jump Part		0.0177	0.0090	0.0039
Total	0.0759	0.0357	0.0369	0.0704

Variance decomposition of Bates, SVOJ, SVRJOJ and SVRJOJ-SC model.

From the table we can see that random and overnight jumps can play an important role in the total variance of the model. In the Bates model the random jump variation is as big as the continuous part variation, while in SVRJOJ and SVRJOJ-SC model it accounts for approximately 9% and 26% respectively. The overnight jump variation plays an especially important role in the SVOJ model where it is almost as high as the continuous variation. We assume this is the case due to the absence of the random jump component. In the SVRJOJ model overnight jump accounts for 25%, while in the SVRJOJ-SC model it only represents 5% of the total variation.

Compared to the Bates model the random jump component plays a less important role in SVRJOJ and SVRJOJ-SC model, this can be the case due to introduction of the overnight jump that now accounts for the part of the variation that is considered under the random component in the Bates model. Similarly a decline can also be observed in the mean of the random jump (μ_{RJ}). On the other hand the value of lambda in the SVRJOJ and SVRJOJ-SC model is higher than the one in Bates, therefore we can assume that by introducing the overnight jump component the model can fit more jumps with smaller mean. This result is consistent with [10]. Overall the variation components in table 7 are higher then the ones reported in [10].

Previous research [10], [2] and [5] show that when random or overnight jumps are introduced in the model the values of parameter estimates σ_σ and ρ should decrease, since the on average negative jumps can partly capture the negative skewness that is observed in the distribution of stock returns. The above statement also holds for our research, with the exception of the estimate for the ρ parameter in the Bates model, that has a a higher value compared to the Heston model.

Besides the parameter estimates we also report their standard deviations, this helps us to evaluate how stable different parameter estimates are. However we should consider the fact that standard deviations were only calculated based on 10 estimates for each parameter. Table 7.1 reports relatively small standard deviations for parameters σ_t^2, σ^2 and σ_{OJ}^2. This result is consistent with [10], however they report slightly lower standard deviations. That could be explained by difference in the underlying or the time period. On the other hand the parameter estimates for all the other parameters are relatively unstable. Depending on the model standard deviations of λ, μ_{RJ} and σ_{RJ} lie between 10%-25% . The mean reversion κ has the most extreme standard deviation which goes all the way to 515% in the SVRJOJ model. Big deviations in estimation of κ are also observed in the literature [10]. When comparing different models and their parameter consistency it can be observed that Black Scholes model is by fare the most stable one, however this is not surprising, since it only has one parameter. However the second most stable model is the SVRJOJ-SC model, this result is rather surprising indicating that the scaled-t overnight jump leads to more consistency.

To evaluate the performance of the option pricing models we consider the value of the objective function (SSE). The smaller the objective function, the better the fit of the model. As expected all the described models are a big improvement in comparison to the Black-Scholes model. From the table it can be seen that the best model is the SVRJOJ-SC model, followed by the SVRJOJ model. This result is not surprising, however we find it interesting how much the fit to the objective functions improves when the overnight returns are scaled-t distributed compared to the normal distribution, since changing the distribution leads to a bigger improvement than for example adding an overnight jump to the Bates model. The third best model is the Bates model, from that we can conclude that the random component is more important than the overnight component, since the performance of the Bates model was considerably better then the one of the SVOJ model. Similar conclusions regarding how well different models minimize the value of the objective function were found in [10], however the reported values of SEE in their article are considerably higher. This can be explained with the fact that they use more options to fit the models, as well as options with a maturity of more than 6 months. Such options are on average more expensive compared to the shorter maturities and the squared percentage error objective function is better at fitting relatively cheap options.

7.2 Models Based on the Variance Gamma Process

In this section we consider the Variance Gamma Model (VG), Variance Gamma Model with Normal Overnight Jump (VG-N) and the Variance Gamma Model with Scaled-t Overnight Jump (VG-SC). Only the VG model was already discussed in the literature, first in [19] and later also in [13], however the techniques used in their research differ significantly from our work. First they consider options on the S&P500 index while we use options on Allianz SE stock. Also time period is different. However the biggest change comes from the technique used to estimate the parameters. As mentioned before we minimize the objective function to obtain parameter estimates, on the other hand [19] and [13] use maximum likelihood estimation.

As already discussed σ_ν, θ and ν are not by itself the standard deviation, the skewness and the kurtosis parameter respectively. Therefore it is not possible to separate the impact of a parameter on a moment without affecting the other moments too. Looking at the estimates of parameter θ that describes the mean as well as the skewness, we can see that across all three models the parameter value is negative, however the estimates differ from one another a lot. The estimate for θ reported in [19] is much bigger then in our case and equals -0.1436.

Again big differences can be observed in parameter estimates of ν, where the estimate in VG-N is much higher compared to other models. Partly this can be explained by scaled-t jump in VG-SC model, that can also capture a part of the kurtosis.

When looking at the standard deviation of parameter estimates, reported in the brackets, we can see that VG-SC model has a rather stable parameters, with the exception of parameter p, μ_{OJ} and θ. In model VG estimate of parameter θ is very high and also has a big standard deviation, on the other hand the standard deviation in other two parameters is insignificant.

Table 7.3: Parameter estimates: Models based on variance gamma process

	VG	VG-N	VG-SC
θ	-8.0389	-0.2192	-0.7589
	(274.84 %)	(32.78 %)	(20.93%)
ν	0.001	3.9137	0.0122
	(0.00731 %)	(279.75 %)	(0.96%)
σ_ν	0.000842	0.0982	0.2156
	(0.0098%)	(6.22 %)	(0.79%)
μ_{OJ}			0.3732
			(40.67%)
σ_{OJ}		0.1635	0.0011
		(8.01 %)	(0.10%)
p			5.6664
			(173.35%)
SSE	0.11516	0.09154	0.0677

Parameter estimates of the VG, VG-N and VG-SC. We report the mean and the standard deviation of the parameters across 10 days.With SSE we denote the average value of the objective function over 10 days.

Similarly as before we also evaluate these models based on how well they minimize the objective function. As expected also models based on the variance gamma processes are a big improvement in comparison to the Black-Scholes model, however the improvement is much smaller than in the previous section. This is mostly the case due to the fact that this model could not fit the options for one or two days well, however the fit for other days were similar to the ones of the stochastic volatility models.

To estimate the parameters of the models we used the squared percentage error objective function, however also some other objective functions could be considered. To conclude how the objective function influences the parameter estimates, we evaluate the parameters when using an absolute error objective function. The estimates are reported in the appendix A.7 and A.8.

When comparing the estimates of models based on stochastic volatility we can see there are especially big differences in the estimation of kappa, on the other hand for other parameters the difference between estimates depend on the model. Also the consistency of the parameters depends on the model, for example when using absolute error objective function the parameters of the SVOJ model are more stable, however the opposite holds for the SVRJOJ model. When looking at the models based on the variance gamma process we can see that the situation is similar. There are big differences in the estimation of θ, while other estimates are relatively close together.

Based on these results we can as expected conclude that the choice of an objective function does influence the parameter estimates. On the other hand the models that performed best under squared percentage error objective function also perform best under absolute error objective function.

7.3 Alternative Evaluation of Option Pricing Models.

In the previous section we estimated all the considered option pricing models and evaluate their performance based on how well they minimize the value of the objective function. However the efficiency of the option pricing model can not be evaluated only based on this criteria. Therefore in this section we discuss alternative indicators that help us conclude which model prices the options most accurately. First we consider statistical indicators, such as the mean of the error and the mean squared error. Later on we compare the models based on the absolute errors for different maturities. In the end we evaluate the models, considering an indicator based on the bid and ask spread, that is used in practice. When evaluating models based on the above criteria, we use the parameter estimates obtained by squared percentage error objective function.

7.3.1 Statistical indicators

In the table below we report the statistical indicators that can help us conclude which model captures the option prices most accurately.

Table 7.4: Statistical indicators for option pricing models

	MAE	St. Dev.	Min	Max	MSE	Mean	MPE
BS	2.3506	46.36%	1.6744	3.167	0.6351	-2.2305	2.0666
H	0.7755	35.85%	0.3300	1.3845	0.0908	-0.5384	0.7773
B	0.5717	28.93%	0.2537	1.0915	0.0441	-0.3016	0.6732
$SVOJ$	0.6632	30.46%	0.2573	1.2815	0.0652	-0.4437	0.7464
$SVRJOJ$	0.5304	**25.33%**	0.2872	1.0534	0.0406	-0.2898	0.6509
$SVRJOJ-SC$	**0.4191**	29.41%	**0.1276**	**0.6184**	**0.0202**	**-0.0697**	**0.6480**
VG	1.2342	81.41%	0.7393	1.7197	0.1998	-0.9345	1.3082
$VG-N$	1.0908	67.22%	0.4184	1.8413	0.1596	-0.8473	1.2042
$VG-SC$	1.0109	38.23%	0.4148	1.6582	0.1364	-0.7822	1.0560

Table summarizes statistical indicators for all the models. MAE denotes the mean over 10 days of the sum of absolute errors. St .Dev. tell us how much the MAE variates. Min and Max denote the minimal and the maximal sum of absolute errors over 10 days. MSE and MPE stand for mean squared error and mean absolute percentage error respectively, mean is again taken over the days. With Mean we denote the mean of the sum of the errors. Best result in each category is highlighted.

First we consider the MAE that contains information on the absolute difference between the option price calculated by the models compared to the market price of an option. As expected the SVRJOJ model has the smallest MAE followed by the SVRJOJ and Bates model, surprisingly however the maximum error of the SVRJOJ-SC is much lower compared to the other models. The Bates model is again more accurate than the SVOJ model, confirming the assumption from the previous section, that adding a random jump is more important than adding an overnight jump. Models based on the variance gamma process perform the worst, they are only successful compared to the Black-Scholes model.

In statistics MSE is often used to evaluate the performance of the estimator, since it presents a trade-off between the increase in the bias for a larger decease in the variance and vice-versa. Also based on this criteria the situation is similar as in MAE, SVRJOJ-SC and SVRJOJ again perform the best, while models based on variance gamma process do not seem to be a good fit.

When looking at MAE and MSE we mostly focus on how well models estimate the prices of relatively expensive options, however for an overall view we also compare the models based on the mean squared error. Also in this case the SVRJOJ-SC model is the best, however the difference between SVRJOJ-SC and SVRJOJ model is much smaller here, indicating that SVRJOJ-SC model might be better in capturing relatively expensive options.

Option pricing model should not always overvalue or undervalue the option price, since this leads to the conclusion that there is a structural error in the model. Optimally the mean error should be close to zero indicating that the over and under pricing in the model is balanced. We check whether this is the case for our models, by looking at the mean reported in the table. Regardless of which model we use the mean error is always negative, indicating that models are under-priced on average.

Based on the statistical indicators we come to the same conclusion as in the previous section. The SVRJOJ-SC model performs best based on MAE, MSE as well as the MPE and is followed by the SVRJOJ model.

7.3.2 Indicators Based on the Absolute Error

In this part we evaluate the models based on the absolute error. With the absolute error we mean the absolute difference between the option price calculated by our model compared to the actual option price reported by the market. First we plot graphs of absolute errors separately for all four maturities. For the most successful models, we also present 3-D plots, where it is shown how the absolute error changes through the strike price and the maturity. The results in this section are based on the parameter estimates from October1, 2020.

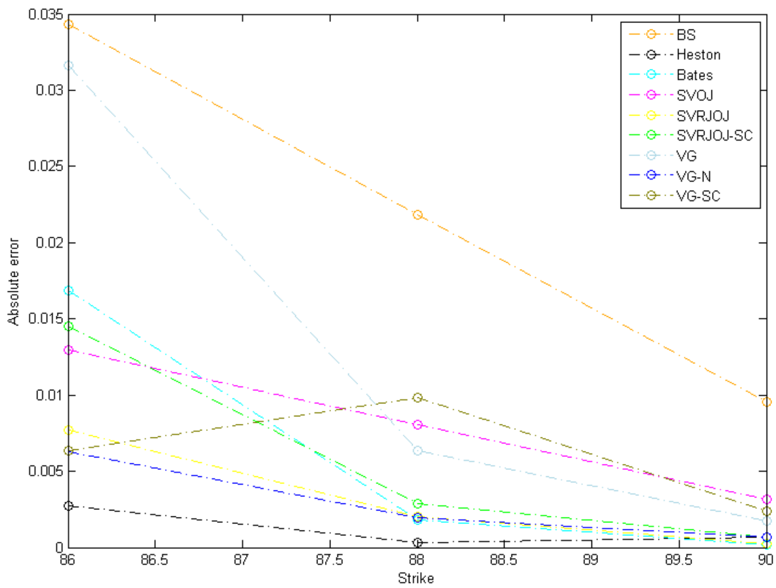


Figure 7.1: Absolute error for maturity 0.031 years (ALV).

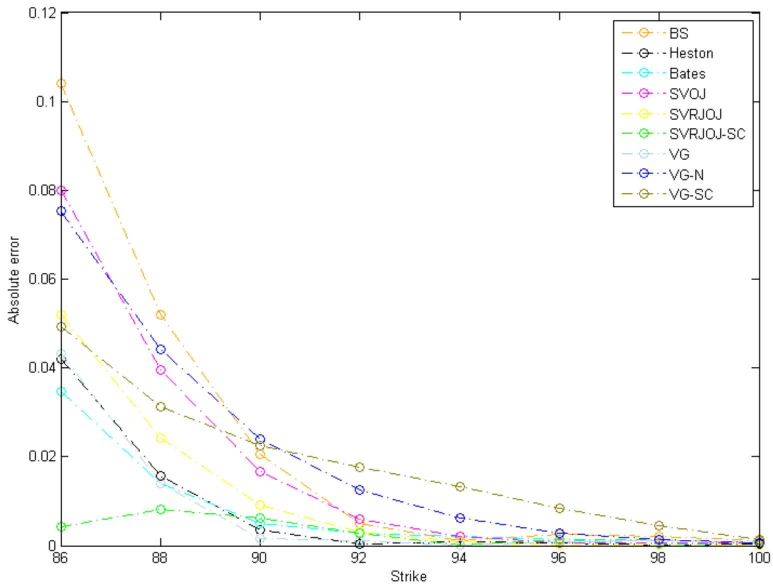


Figure 7.2: Absolute error for maturity 0.120 years (ALV).

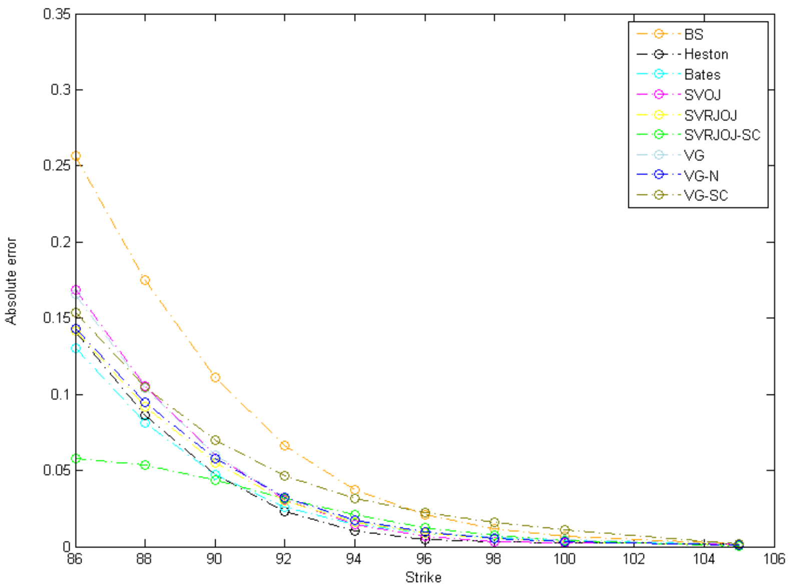


Figure 7.3: Absolute error for maturity 0.210 years (ALV).

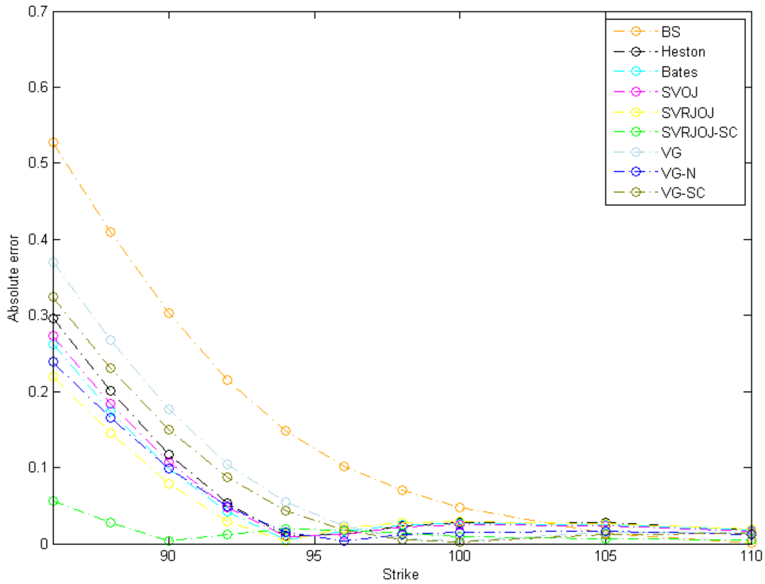


Figure 7.4: Absolute error for maturity 0.4603 years (ALV).

In the table below we summarize which models perform best based on four maturities. Our conclusions are drawn from the graphs reported above.

Table 7.5: Best performing option pricing models depending on the maturity

Maturity	1-mat	2-mat	3-mat	4-mat
	H	SVRJOJ-SC	SVRJOJ-SC	SVRJOJ-SC
	VG-N	VG	B	SVRJOJ
	SVRJOJ	B	H	VG-N

The model reported first for each of the maturities is the model that fitted option prices best.

As we can see from the table and the graphs above, the performance of the option pricing model depends on the maturity. For example SVRJOJ-SC model performs best for all maturities but first, where surprisingly Heston is the most accurate model. Through maturities the accuracy of the models seems to be decreasing, this can be partly explained by the fact that with increasing maturity options become more expensive and are no longer fitted so well, due to squared percentage error objective function.

When we evaluated option pricing models based on statistical indicators and on which model minimizes the objective function best, we concluded that models based on the variance gamma process are only better than the Black-Scholes model. However here we can see this is no longer the case, since especially VG-N model preforms very well for the first and fourth maturity, although the errors at other maturities are then relatively large, leading to the overall poor result.

For the most successful models from the table above we also plot 3-D plots that show the absolute errors across maturities and the strike prices for each model separately. This helps us evaluate at which maturities and which strikes the model performs best. The results presented are again based on the parameter estimates of day one.

When comparing the absolute errors in the Bates and SVRJOJ model we can see that the models perform in a similar way. The errors increase throughout the maturities, they are bigger for higher strikes. The main difference that we notice is that the Bates model captures options with higher strikes better, than the SVRJOJ model. On the other hand SVRJOJ model performs better with relatively more expensive options, since its peaks are colored blue in comparison to the red.

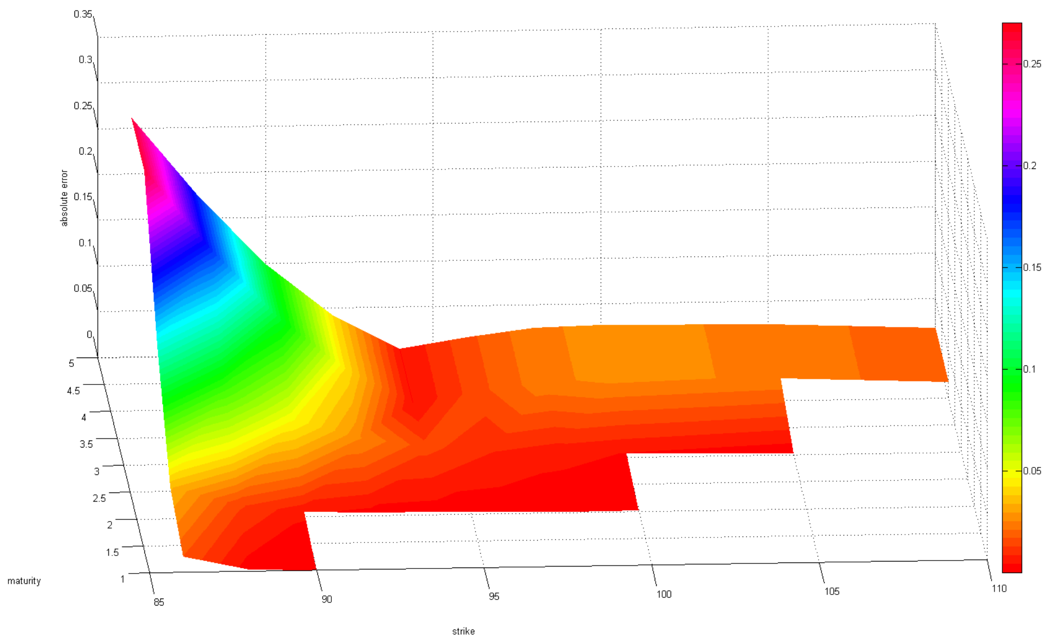


Figure 7.5: Absolute pricing errors: Bates model

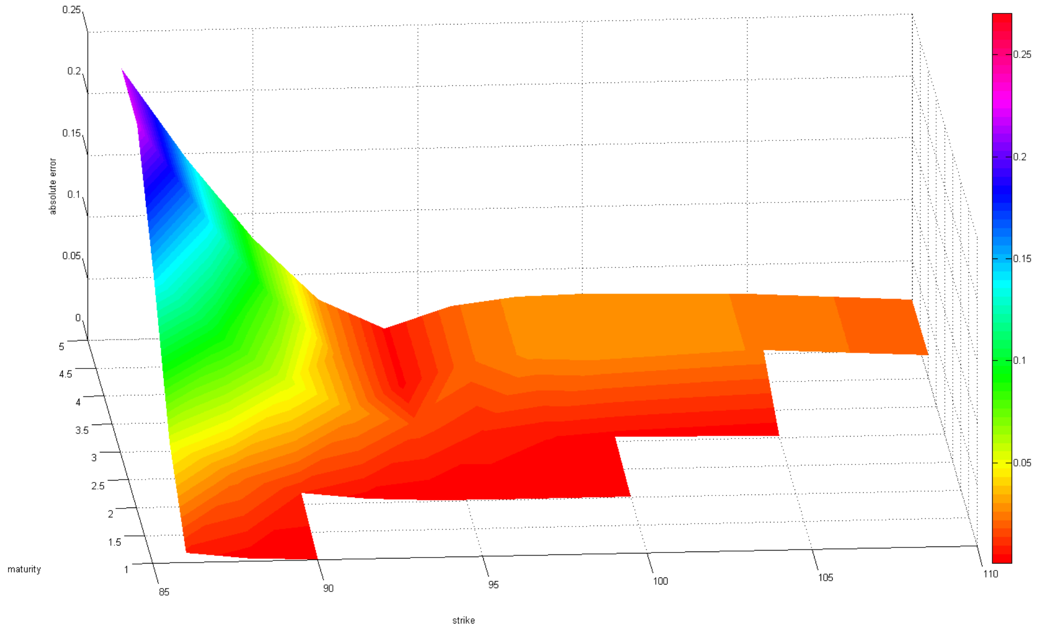


Figure 7.6: Absolute pricing errors: SVRJOJ model

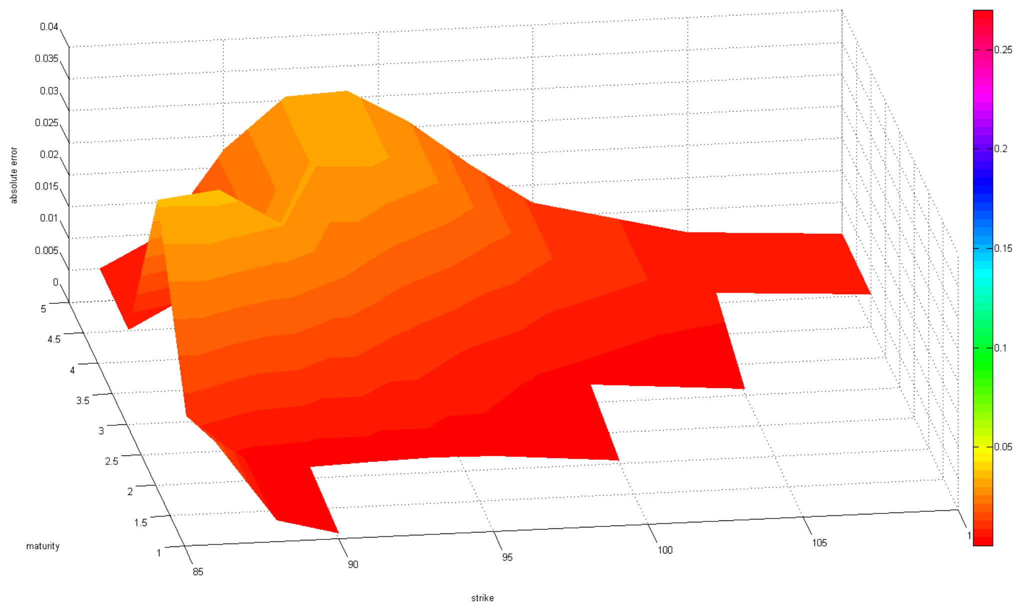


Figure 7.7: Absolute pricing errors: SVRJOJ-SC model

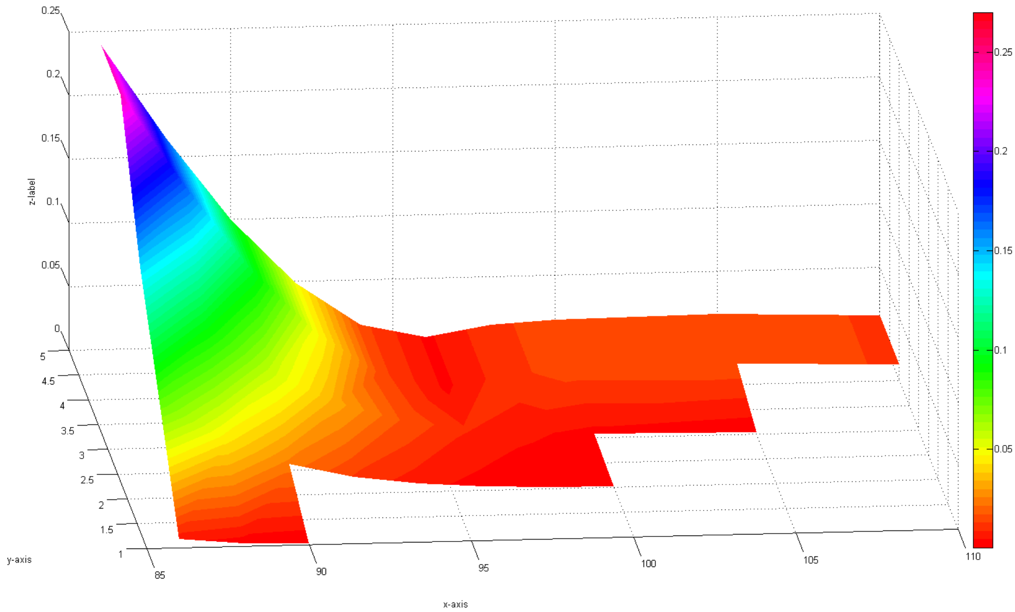


Figure 7.8: Absolute pricing errors: VG-N model

The 3-D plot of absolute errors for VG-N model is very similar to the ones of the Bates and SVRJOJ model. It manages to capture errors at higher strikes well, however at lower strikes it no longer performs so well, for example the errors of value between 0.05-0.1 already start for the first maturity.

On the other hand, the performance of SVRJOJ-SC model is surprisingly good through both strikes and maturities, since the errors never exceed the value of 0.04. In comparison to all other models the errors in SVRJOJ-SC model are no longer increasing with maturity, unexpectedly there is a very small error for the lowest strike and the longest maturity. That can be explained by the fact that SVRJOJ-SC model might capture extremes better due to the scaled-t distribution of the overnight jump.

Considering all the criteria, from the minimization of the objective function to the statistical indicators and plots of absolute errors, we can conclude that SVRJOJ-SC model is the best option pricing model, followed by the SVRJOJ model. This result was to some extend expected, since SVRJOJ-SC model has the most parameters and can therefore capture the option price behavior best. On the other hand we are surprised with the poor performance of the models based on variance gamma process. Even though they manage to capture the option price for certain moments extremely well, they present a poor overall performance.

7.3.3 Model Evaluation Based on the Bid-Ask Spread

Market makers use an internal option pricing model to price the options. They compare the price obtained by their model to the bid and ask price given by the market. Bid is the highest price a buyer is prepared to pay for an option and ask the lowest price for which a a seller is willing to sell an option. The difference between the bid and the ask is called bid-ask spread.

The internal option pricing models should calculate an option price close to the mid market price, which is the mean of bid and ask. From this price given by the internal option pricing model, a market maker can calculate the bid and ask price indicated by his model. For example if the option price given by the model is higher then the mid market price reported, then the bid and ask price implied by the model will also be higher then the ones reported by the market. In the case when the market makers trusts theirr option pricing model, they can quote at the higher bid and ask, since they believes that the option is under-priced.

Depending on the option, the option price given by the model should not be more than 15% of the spread away from the mid market price. Sightly bigger deviations are allowed for options that are far in or far out of the money. The deviation from the mid market indicates on the trading opportunity, however if the model thinks there is a trading opportunity in almost every option it is more likely that the option pricing model does not capture prices well, than that the market is constantly wrong.

To evaluate how well the above described models perform based on the this bid-ask spread criteria we use new date, since no bid and ask prices were available for the options we considered in the previous parts. Here we take the option data on Allianz stock from October 1, 2012 to October 5, 2012, all-together this gives 5 days of observations. Each day we have 18 option prices with different strikes and the maturity December 21, 2012. Compared to the previous parts we use only one maturity when calibrating the models, since this allows for more accurate parameter estimation. This approach is also used in practice.

When calibrating the models we consider the squared percentage error objective function. The estimated parameters are reported in the appendix A.9 and A.10. From those table we can see that the average value of the objective function is smaller in this case, this can be explained by the smaller amount of options we use as well as the fact that we only consider one maturity, which allows us to fit the models more accurately. Big improvement can be noticed especially when looking at the Bates, SVRJOJ, SVRJOJ-SC, VG-N and VG-SC model.

We first report the average absolute deviations of our models from the mid-market price for each day separately. The deviations are expressed in the form of the percentage of the bid-ask spread (absolute error divided by the bid-ask spread).

Table 7.6: Average percentage deviations from the mid-market price

	1-day	2-day	3-day	4-day	5-day
BS	377.15 %	295.82 %	343.39 %	219.40 %	242.11 %
H	54.24	32.55	53.97	27.56	28.74
B	37.35	14.01	50.39	20.88	11.19
SVOJ	45.68	33.89	47.42	22.55	26.11
SVRJOJ	27.60	8.37	49.12	15.51	11.62
SVRJOJ-SC	24.43	7.57	89.44	13.94	10.06
VG	323.45	245.47	307.32	229.89	243.94
VG-N	41.95	29.526	89.83	37.68	12.53
VG-SC	53.29	18.63	101.17	54.20	32.95

The table reports average percentage deviations of the option prices calculated by our models compared to the mid-market price. Results presented in the table equal mean over absolute value of the difference between option prices calculated by our models and the mid-market price.

From the table we can observe that the deviations from the mid-market price are mostly larger than the advised 15% of the bid-ask spread. The Black-Scholes and VG model performance is very poor, exceeding the deviation of 100% each day. On the other hand as expected SVRJOJ and SVRJOJ-SC model perform best and manage to stay in the 15% band for three out of five days. We can see that all models do not manage to capture option price behavior on day 3, where SVRJOJ-SC model performs wost than most of the others.

In order to show how the option price deviations between our models and mid-market price depend on the strike, we plot option prices given by our models, mid-market price, bid and ask price. We do not plot option prices given by Black-Scholes and Variance Gamma model, due to too large deviations. For the better clarity of the results we deduce mid-market price, from option price given by our model, mid-market price, bid and ask price.

On October 2, 2012 Allianz stock was worth 94.24€, therefore the options on the left from the strike 94.24€ are out of the money puts and on the right we have out of the money calls. From the picture we can see that the bid-ask spread changes through different strikes. When the options are further out of the money the spread is bigger. This is the case since options that are further out of the money are usually less liquid. Similar behavior can be observed when looking at the difference between the option price calculated by our model and the mid-market price. All the models give an option price that is relatively close to the mid-market price for puts that are not that far out of the money, on the other hand only SVRJOJ, SVRJOJ-SC, Bates and VG-N stay close to the mid market price when looking at the out of the money calls.

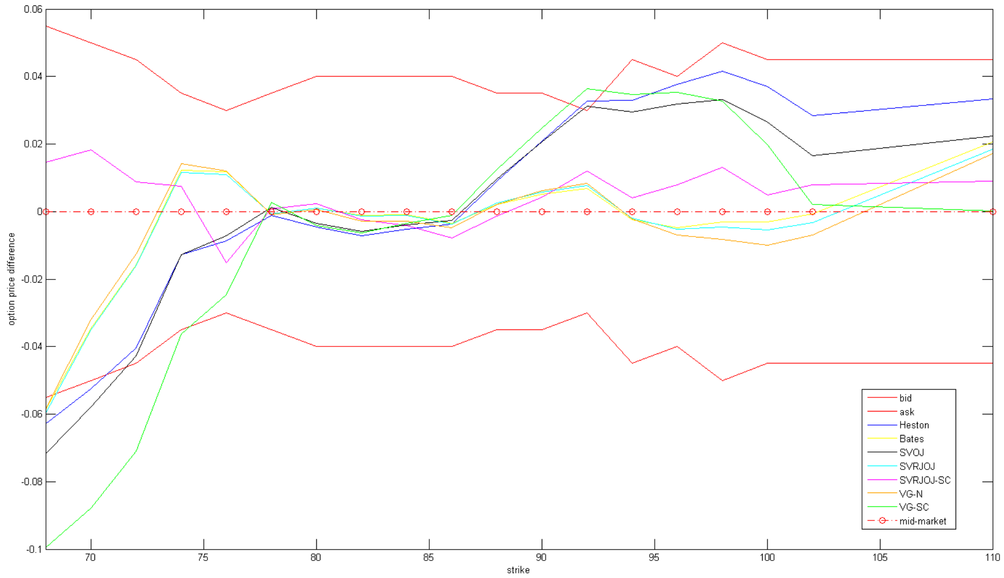


Figure 7.9: Option prices given by the models compared to mid-market price

From the Table 7.6 and the picture above we can see that SVRJOJ-SC, SVRJOJ and Bates model manage to price options closest to the mid-market price and stay between the bid and the ask price for most of the strikes. Based on this we can assume that those option pricing models are more reliable. However we can not come to solid conclusions since the check for which deviations from the mid-market price were actually trading opportunities and which were model errors exceeds the scope of this research.

7.4 Out of Sample Pricing Errors

In the previous parts we have shown that in-sample fit of option prices is the best for the SVRJOJ-SC model. However one could argue that this is the case due to a large number of structural parameters compared to the other models. To lower the impact of the number of parameters on the model performance, we examine the out-of-sample pricing accuracy of the models. This approach might favor models with less parameters, since the presence of extra parameters might cause over-fitting.
To evaluate the out-of-sample pricing performance of all the considered models, we use the parameter estimates of each of the models calculated for every day separately. When calculating the option price using the proposed model, we use

the model parameters estimated the previous day. We report the average absolute and absolute percentage pricing errors for each model, every day separately. In the end we also show the average errors of the option prices calculated by parameters estimates for that day, to show the difference in option price accuracy in sample and out of sample.

Table 7.7: Absolute pricing errors: Out of sample

	BS	H	B	SVOJ	SVRJOJ	SVRJOJ-SC	VG	VG-N	VG-SC
2	0.135	0.081	0.080	0.087	0.080	**0.065**	0.097	0.087	0.085
3	0.066	0.068	0.093	0.080	0.095	0.095	0.071	**0.062**	0.064
4	0.082	0.028	0.026	0.031	0.031	**0.025**	0.053	0.045	0.037
5	0.116	0.040	0.036	0.037	**0.028**	0.041	0.081	0.047	0.062
6	0.116	**0.013**	0.018	0.016	0.016	0.019	0.038	0.056	0.049
7	0.110	0.015	0.014	**0.015**	0.015	0.029	0.053	0.075	0.022
8	0.099	0.044	0.029	0.038	0.030	**0.018**	0.055	0.038	0.022
9	0.067	0.038	0.038	0.037	**0.042**	0.068	0.045	0.059	0.047
10	0.103	0.033	0.026	0.037	0.030	0.038	0.030	**0.022**	0.024
Mean	0.089	**0.036**	**0.036**	0.038	0.037	0.039	0.052	0.049	0.041
Max Err.	0.134	**0.081**	0.093	0.087	0.095	0.095	0.097	0.087	0.085

The table reports average absolute pricing errors for each day separately. When calculating the option price given by the model, parameters estimates from the previous day are used. With Mean we denote the mean percentage absolute error and with Max Err. we denote maximum average absolute error. Errors are denoted in Euros.

The highlighted numbers denote the model that has the smallest average absolute error for each day separately. From the above table we can see that non of the models significantly outperforms the others. Based on the overall mean and maximum average absolute error we can see Heston, Bates, SVOJ, SVRJOJ and SVRJOJ-SC behave very similarly, the differences in their performance are so small that we can not conclude which one is the best. Also the best model based on variance gamma process is not to fare behind, although based on the low number of parameters we could expect that VG and VG-N model would perform better. They seem to me more sensitive to parameter change than some other models.

From the below table that summarizes the mean absolute percentage errors we can see that the Heston and Bates model are again very close. We can learn that the performance between models is very similar, for example on day 2 all models are extremely inaccurate, indicating on a significant change in the option prices from the previous day.

Table 7.8: Percentage pricing errors: Out-of-sample

	BS	H	B	SVOJ	SVRJOJ	SVRJOJ-SC	VG	VG-N	VG-SC
2-day	16.68 %	**12.99** %	13.75 %	13.91 %	13.80 %	13.09%	15.09%	14.77 %	14.27 %
3-day	19.07	19.01	20.43	21.66	20.95	21.17	17.98	**17.28**	18.17
4-day	9.11	5.41	5.43	6.43	5.73	**5.04**	6.89	6.14	6.21
5-day	10.42	3.99	3.97	**3.76**	4.37	4.76	8.013	6.05	5.92
6-day	9.67	**2.05**	2.58	2.65	2.74	2.87	7.22	4.50	4.82
7-day	9.68	**3.37**	3.52	3.44	3.53	4.21	5.28	5.81	3.75
8-day	10.57	5.66	4.89	6.75	**4.87**	6.10	6.28	6.35	5.77
9-day	11.18	9.22	9.13	**8.62**	9.05	11.58	9.40	9.24	11.12
10-day	7.99	5.02	5.00	4.92	5.16	4.47	4.77	**3.82**	3.91
Mean	10.44	**6.67**	6.68	7.21	7.02	7.33	8.09	7.40	7.39
Max Err.	19.07	19.01	20.43	21.66	20.95	21.17	17.98	**17.28**	18.17

The table reports average absolute percentage pricing errors for each day separately. When calculating the option price given by the model, parameters estimates from the previous day are used. With Mean we denote the mean percentage absolute error and with Max Err. we denote maximum absolute percentage error. Results are given in percentages

When comparing the mean absolute errors between models with parameter estimates from that day to the ones that use the parameter estimates from the previous day, we can observe large differences. First as expected, the mean errors in Table 7.9 are much smaller, compared to Table 7.7 with the biggest differences on day 2 and 3. Also maximum mean absolute errors are two or three times higher when using the parameters estimates from the previous day. Surprisingly we can see that in some cases the models with parameter estimates from the previous day even perform better. This can be explained with the fact that the models were fitted to the squared error objective function and here we are comparing the absolute errors.

Table 7.9: Absolute pricing errors: In-sample

	BS	H	B	SVOJ	SVRJOJ	SVRJOJ-SC	VG	VG-N	VG-SC
2-day	0.060	0.046	0.012	0.026	0.010	**0.008**	0.046	0.061	0.055
3-day	0.085	0.019	0.019	0.021	0.022	**0.012**	0.051	0.041	0.032
4-day	0.101	0.033	0.029	0.031	**0.016**	0.025	0.066	0.038	0.049
5-day	0.122	0.014	0.014	0.010	0.011	**0.010**	0.045	0.064	0.054
6-day	0.107	0.014	0.011	0.011	0.012	**0.010**	0.054	0.076	0.017
7-day	0.099	0.039	0.024	0.028	0.024	**0.018**	0.052	0.027	0.022
8-day	0.089	0.038	0.036	0.034	0.031	**0.027**	0.049	0.032	0.033
9-day	0.107	0.032	0.021	0.035	**0.024**	0.028	0.037	0.020	0.025
10-day	0.080	0.032	0.026	0.023	0.029	**0.026**	0.035	0.026	0.028
Mean	0.094	0.029	0.022	0.025	0.020	0.017	0.0484	0.043	0.035
Max Err.	0.122	0.046	0.035	0.035	0.031	0.028	0.066	0.076	0.055

The table reports average absolute pricing errors of option calculated with parameter estimates for the that day. Highlighted numbers denote the model with the smallest average absolute error.

From the above results we can conclude that parameter estimates from the previous day should not be used when modeling option prices, since the errors on average deviate from 6% - 10 % and can in the worst case even reach 20% or more. Such errors are much to high, therefore we would advice to reestimate the parameters each day separately. For models to still be able to calculate the option prices on the day of interest, parameters could be estimated in the first few trading hours and later on used for pricing options. The effectiveness of models with such parameters estimates exceeds the scope of this project and will be left for future research.

As expected, models with more parameters that performed best in-sample are no longer as successful, due to the over-fitting effect.

The fact that distributions are very similar intuitively makes sense since the models were calibrated using the same data.

Chapter 8

Conclusion and Further Research

8.1 Conclusion

In our research we show that incorporating an overnight jump into already existing option pricing models leads to more accurate option prices. The existing research assumed that intra-day as well as overnight returns followed a normal distribution. However by looking at the intra-day and overnight returns of stocks in the DAX index we conclude that the variance gamma distribution describes intra-day returns best, while overnight returns are captured most accurately by the scaled-t distribution.
We introduce those two distributions into the option pricing modes and propose the SVRJOJ-SC, VG-N and VG-SC model. Comparing the performance of the models based on the minimization of the objective function, absolute errors and statistical indicators we conclude that SVRJOJ-SC model captures the option price behavior best, followed by the SVRJOJ model. The good performance of the above mentioned models can be partly explained by a large number of parameters that enables them to fit the option data better. Models based on the variance gamma process only outperform the Black-Scholes model. By comparing the performance of SVOJ and Bates model, we conclude that a random jump is more important than the overnight jump, since Bates model outperforms the SVOJ model.
When looking at the out of sample performance of the option pricing models, there is no clear winner. Models that captured option prices best in sample, no longer do so well due to the over-fitting effect. Therefore we advice to reestimate the parameters each day separately.
This research concludes that total jump risk should be separated in the random and jump risk component. Where the overnight jump component should follow the scaled-t distribution, since this leads to improvement in pricing options.

8.2 Further Research

In this section we propose some possible extensions of our research.
First, we suggest looking at the bigger date set when estimating the parameters of the model. In our research we took 10 days of data, however a longer time period would allow to compare the standard deviations of the parameters in different models more accurately. This research shows that the parameters are more stable in SVRJOJ-SC model compared to SVRJOJ, Bates, SVOJ and Heston model, however the date set is too short to come to any reliable conclusion.
Second, out of sample performance of all the models we consider was relatively poor. Therefore we suggest reestimating the parameters more often. For example use the option data from the previous hour to fit the parameters of the models for the next hour.
Finally, hedging performance under the considered models should be evaluated. Similar research as in [10] for Heston and Bates model could be also conducted for the new models we proposed, especially for SVRJOJ and SVRJOJ-SC model, due to their good in sample performance.

Bibliography

[1] Felipe Aparicio and Javier Estrada. Empirical Distributions of Stock Returns: European Securities Markets. *European Journal of Finance*, 7:1–21, 1991.

[2] Gurdip Bakshi, Charles Cao, and Zhiwu Chen. Empirical Performance of Alternative Option Pricing Models. *Journal of Finance*, 52(5):2003–2049, 1997.

[3] Gurdip Bakshi and Dilip Madan. Spanning and Derivative-Security Valuation. *Journal of Financial Economics*, 55(1):205–238, 2000.

[4] Maurice S. Barlett. The Effect of Standardization on a Chi Square Approximation in Factor Analysis. *Biometrika*, 38:337–344, 1951.

[5] David S. Bates. Jumps and Stochastic Volatility: Exchange Rate Processes Implicit in Deutsche Mark Options. *Review of Financial Studies*, 9(1):69–107, 1996.

[6] David S. Bates. Post-87 Crash Fears in S&P 500 Futures Option, 1997.

[7] Martin Baxter and Andrew Rennie. *Financial Calculus: An Introduction to Derivative Pricing*. Cambridge University Press, 1996.

[8] Fischer Black and Myron Scholes. The Pricing of Options and Corporate Liabilities. *The Journal of Political Economy*, 81(3):637–654, 1973.

[9] Robert C. Blattberg and Nicholas J. Gonedes. A Comparison of the Stable and Student Distributions as Statistical Models for Stock Prices. *Journal of Business*, 47(2):244–280, 1974.

[10] Mark-Jan Boes, Feike C. Drost, and Bas J.M. Werker. The Impact of Overnight Periods on Option Pricing. *Journal of Financial and Quantitative Analysis*, 42:517–533, 2007.

[11] Peter K. Clark. A Subordinated Stochastic Process Model With Finite Variance for Speculative Prices. *Econometrica*, 41(1):135–155, 1973.

[12] Eugene F. Fama. The Behavior of Stock Market Prices. *Journal of Business*, 38(1):34–105, 1965.

[13] Fiorani Filo. Option Pricing Under the Variance Gamma Process, 2004.

[14] Manfred Gilli and Enrico Schumann. Calibrating Option Pricing Models with Heuristics. . *COMISEF Working Papers Series WPS-030*, 2010.

[15] Brian D. Gray and Dan W. French. Empirical Comparisons of Distributional Models for Stock Index Returns. *Journal of Business*, 17(3):451–459, 1990.

[16] Steven L. Heston. A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options. *The Review of Financial Studies*, 6:327–343, 1993.

[17] D A Hsu. A Bayesian Robust Detection of Shift in the Risk Structure of Stock Market Returns. *Journal of the American Statistical Association*, 77(2):29–39, 1982.

[18] John C. Hull. *Options, Futures, and other Derivative Securities.* Prentice Hall, 2000.

[19] Dilip B. Madan, Peret P. Carr, and Eric C. Chang. The Variance Gamma Process and Option Pricing. *European Finance Review*, 2:97–105, 1998.

[20] Dilip B. Madan and Frank Milne. Option Pricing with VG Martingale Components. *Mathematical Finance*, 4(1):39–55, 1991.

[21] Dilip B. Madan and Eugene Seneta. The Variance Gamma (V.G.) Model for Share Market Returns. *Journal of Business*, 63(4):511–524, 1990.

[22] Benoit Mandelbrot. The Variation of Certain Speculative Prices Pricing. *Journal of Business*, 36(4):394–419, 1963.

[23] H. B. Mann and A. Wald. On the Stochastic Treatment of Linear Stochastic Difference Equations. *Econometrica*, 3/4:173–200, 1943.

[24] Amando Peiro. The Distribution of Stock Returns: International Evidence. *Applied Financial Economics*, 4(6):431–439, 1994.

[25] Peter D. Praetz. The Distribution of Share Price Changes. *Journal of Business*, 45(1):49–55, 1972.

[26] James Press. A Compound Events Model for Security Prices. *Journal of Business*, 40(3):317–335, 1999.

[27] Mark Rubinstein. Implied Binomial Trees. *Journal of Finance*, 49(3):771–818, 1994.

[28] Luis O. Scott. Pricing Stock Options in a Jump-Diffusion Model with Stochastic Volatility and Interest Rates: Applications of Fourier Inversion Methods. *Mathematical Finance*, 7:413–426, 1987.

[29] Eugen Seneta. Fitting the Variance-Gamma Model to Financial Data. *Stochastic Methods and Their Applications*, 41:177–187, 1999.

[30] Jaye B. Smith. *The Probability Distribution of Market Returns: A Logistic Hypothesis*. PhD thesis, 1981.

[31] Elias M. Stein and Jeremy C. Stein. Stock Price Distributions with Stochastic Volatility: An Analytic Approach. *Review of Financial Studies*, 4(4):727–752, 1991.

[32] Alan Stuart, Keith Ord, and Steven Arnold. *Kendall's Advanced Theory of Statistics, Classical Inference and the Linear Model*. John Wiley & Sons, 2009.

[33] Fengzhong Wang, Shwu-Jane Shieh, Shlomo Havlin, and H. Eugene Stanley. Statistical Analysis of the Overnight and Daytime Return. *Physical Review*, 79, 2009.

Appendix A

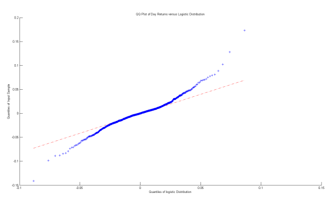
Appendix

In the appendix we first present results regarding the distribution of intra-day and overnight returns, that were mentioned but not shown in the report until now. Later on we give parameter estimates of the considered models when using absolute error objective function as well as parameter estimates for the models used in section 7.3.3.

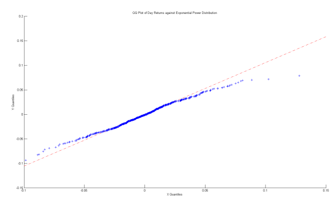
A.1 QQ-plots of the Allianz stock returns

QQ-plots of the Allianz intra-day and overnight stock returns for logistic, exponential power and mixture normal distribution are presented in this part.

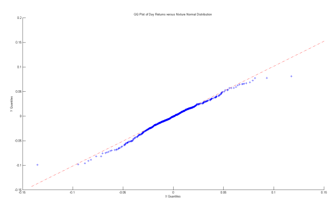
Figure A.1: QQ-plots - alternative distributions: Intra-day (ALV)



Logistic Distribution

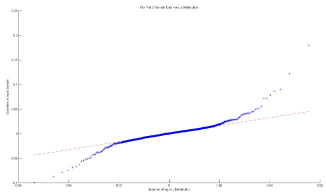


Exp. Power Distribution

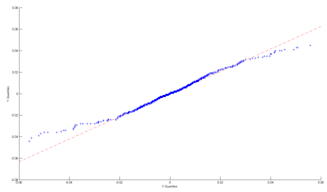


Mixture Normal Distribution

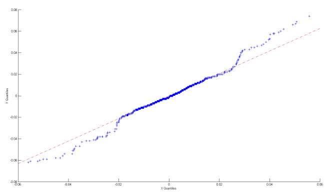
Figure A.2: QQ-plots - alternative distributions: Overnight (ALV)



Logistic Distribution



Exp. Power Distribution



Mixture Normal Distribution

A.2 Fitting Distribution to Alternative Data

We determine which distributions describe intra-day and overnight returns on BMW and Deutsche Bank stock best. This stocks are considered to show that distributions obtained to fit intra-day and overnight returns on Allanz stock best also manage to capture returns of other stocks most accuratelly. Similar results were obtained for all the stock in German DAX.

A.2.1 Distributions of BMW Stock Returns

First we report the empirical statistics of BMW stock returns.

Table A.1: Moments of intra-day and overnight returns: BMW

Returns	Mean	St. dev.	Min of series	Max	Skw	Kr
Intra-day	0.0001	0.0226	-0.1319	0.166	0.308	7.5103
Overnight	0.0004	0.0127	-0.0974	0.0857	-0.7093	13.6367

Moments of intra-day and overnight returns on BMW stock. Sskw = Standerdized skewness ans Skrt = standardized kurtosis

The parameter estimates of the considered distributions are summarized in the following table:

Table A.2: Distribution estimates: BMW

Distribution	Parameter	Intra-day	Overnight
N	μ	0.000094223	0.00042888
	σ	0.0226	0.0127
S-t	μ	-0.0002488	0.0006657
	σ	0.0168	0.0075
	ν	4.3469	2.9753
L	μ	-0.0001666	0.0006009
	α	0.0119	0.061
EP	μ	-0.0002849	0.0007273
	α	0.0115	0.0042
	β	1	0.673
MN	μ_1	-0.00027337	0.0001899
	σ_1	0.0403	0.0296
	μ_1	-0.0001394	0.0006147
	σ_2	0.0163	0.008
	λ	0.1775	0.1204
VG	c	-0.001918	0.001475
	σ	0.022078	0.0117
	θ	-0.002012	0.001044
	ν	0.59489	0.87696

Distributions are fitted to the BMW stock.Parameter estimates of the following distributions are reported, N=Normal, S-t = Scaled t, L = Logistic, EP=Exponential power, MN = Mixture Normal, VG=Variance gamma.

The Goodness of fit test results show similar results as when considering ALV stock

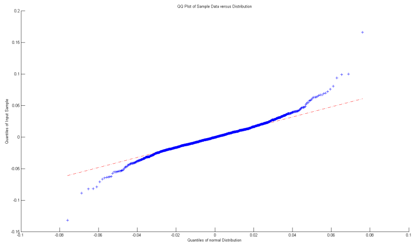
Table A.3: Chi Squared Test: BMW

	Intra-day return		Overnight return	
Distribution	Test	P-value	Test	P-value
N	1	0.0000	1	0.0000
S-t	0	0.5132	0	0.1905
L	1	0.0000	1	0.0000
EP	1	0.0023	1	0.0000
MN	1	0.0097	1	0.0093
VG	0	0.0713	1	0.0083

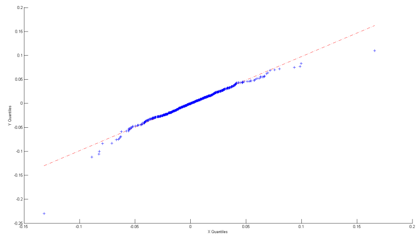
Results of the Chi Squared Test, if 1 is written under the test, that means that the null hypothesis was rejected, if there is 0 the hypothesis was not rejected. Test is based on the BMW data.

QQ-plots for intra-day and overnight returns, for normal, variance gamma and scaled-t distribution are reported on the next page.

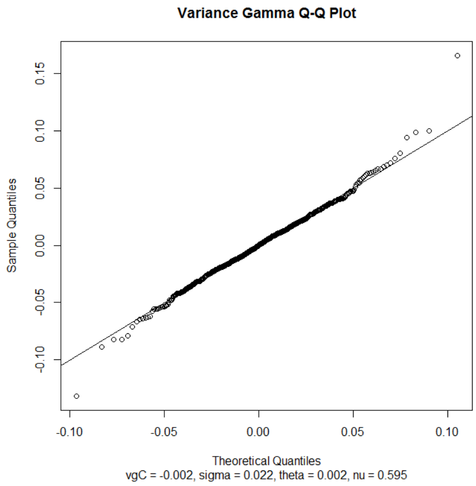
Figure A.3: QQ-plots: Intra-day (BMW)



BMW Normal Distribution

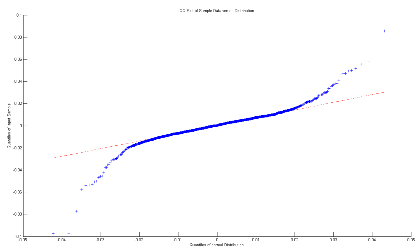


BMW Scaled-t Distribution

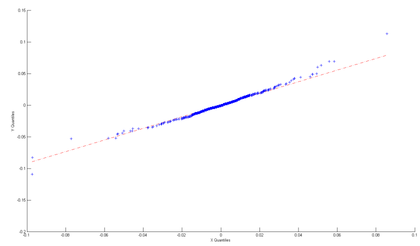


BMW Variance Gamma Distribution

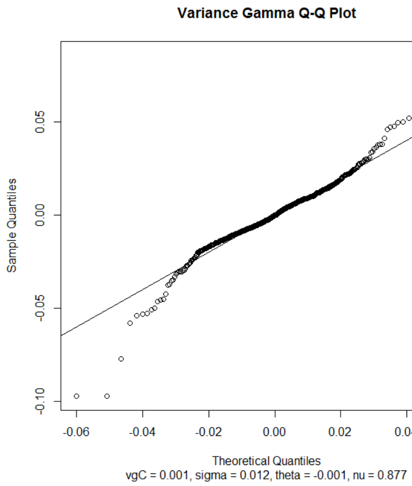
Figure A.4: QQ-plots: Overnight (BMW)



BMW Normal Distribution



BMW Scaled-t Distribution



BMW Variance Gamma Distribution

A.2.2 Distribution of DBK Stock Returns

First we report the empirical statistics of DBK stock returns.

Table A.4: Moments of intra-day and overnight returns: DBK

Returns	Mean	St. dev.	Min of series	Max	Skw	Krt
Intra-day	-0.0014	0.0236	-0.0931	0.1027	0.1431	5.1643
Overnight	0.0015	0.0141	-0.0987	0.1082	0.374	12.0142

Moments of intra-day and overnight returns on DBK stock.

Table A.5: Distribution estimates: DBK

Distribution	Parameter	Intra-day	Overnight
N	μ	-0.0014	0.0015
	σ	0.0236	0.0141
S-t	μ	-0.0015	0.0015
	σ	0.0169	0.0077
	ν	3.7053	2.4272
L	μ	-0.0015	0.0014
	α	0.0124	0.068
EP	μ	-0.0012	0.0018
	α	0.0104	0.0048
	β	0.8024	1
MN	μ_1	-0.0011	0.0019
	σ_1	0.0341	0.0257
	μ_1	-0.0017	0.0015
	σ_2	0.0138	0.0075
	λ	0.1317	0.2377
VG	c	-0.004028	0.001506
	σ	0.022831	0.013269
	θ	0.002657	-0.000035
	ν	0.118402	1.15778

Distributions are fitted to the DBK stock.Parameter estimates of the following distributions are reported, N=Normal, S-t = Scaled t, L = Logistic, EP=Exponential power, MN = Mixture Normal, VG=Variance gamma.

The Goodness of fit test results

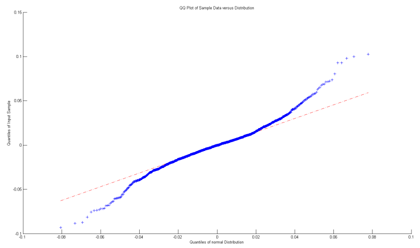
Table A.6: Chi Squared Test: DBK

	Intra-day return		Overnight return	
Distribution	Test	P-value	Test	P-value
N	1	0.0000	1	0.0000
S-t	0	0.4832	0	0.3435
L	1	0.0000	1	0.0000
EP	1	0.0031	1	0.0000
MN	1	0.0000	1	0.0000
VG	0	0.0045	1	0.0013

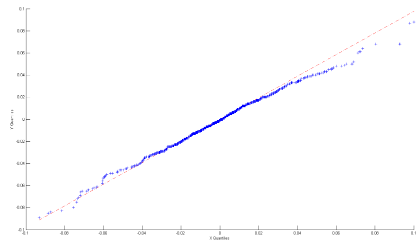
Results of the Chi Squared Test, if 1 is written under the test, that means that the null hypothesis was rejected, if there is 0 the hypothesis was not rejected. Test is based on the DBK data.

QQ-plots for intra-day and overnight returns, for normal, variance gamma and scaled-t distribution are reported bellow.

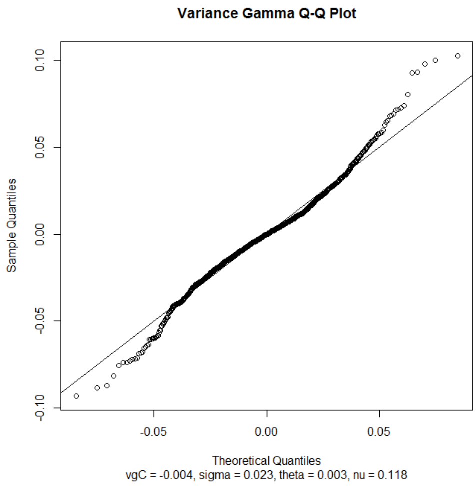
Figure A.5: QQ-plots: Intra-day (DBK)



DBK Normal Distribution

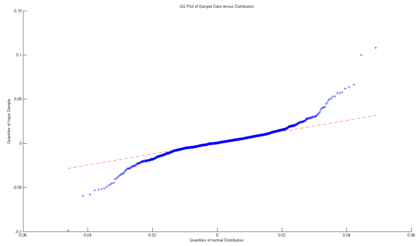


DBK Scaled-t Distribution

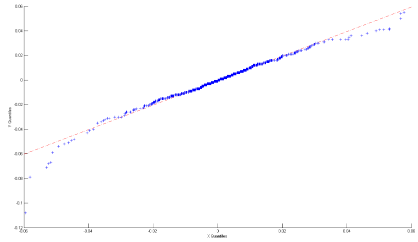


DBK Variance Gamma Distribution

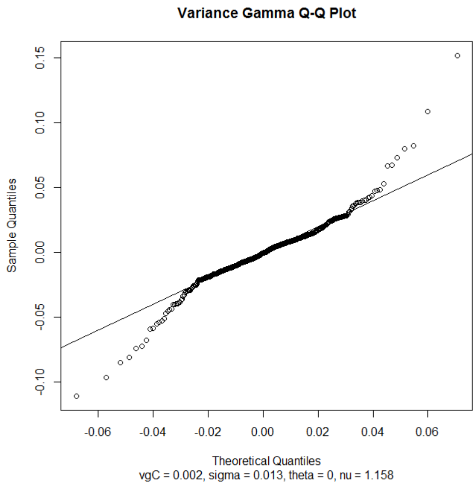
Figure A.6: QQ-plots: Overnight (DBK)



DBK Normal Distribution



DBK Scaled-t Distribution



DBK Variance Gamma Distribution

A.3 Parameter Estimates when using Absolute Error Objective Function

Table A.7: Parameter estimates - Absolute error objective function (Stochastic volatility models)

	BS	Heston	Bates	SVOJ	SVRJOJ	SVRJOJ-SC
μ_{RJ}			-0.6945		-0.3500	-0.1896
			(16.96%)		(73.82%)	(4.51 %)
σ_{RJ}			0.7176		0.0578	0.1758
			(63.98%)		(54.41%)	(5.55 %)
λ			0.0675		0.0538	0.6351
			(1.47%)		(2.64%)	(33.32 %)
μ_{OJ}						0.0038
						(0.71%)
σ_{OJ}				0.050	0.1106	0.0591
				(3.67%)	(5.67%)	(3.25 %)
κ		0.1203	13.1319	0.0517	1.1576	0.366
		(7.91%)	(231.54%)	(6.19%)	(162.23%)	(69.80 %)
σ		0.2530	0.1966	0.763	0.7704	0.3512
		(0.89%)	(0.81%)	(40.7%)	(77.72%)	(38.80%)
σ_σ		1.060	0.3209	0.835	0.52	0.1539
		(33.63%)	(120.47%)	(13.91%)	(41.02%)	(4.73%)
ρ		-0.2695	-0.4946	-0.509	-0.4775	0.3660
		(17.91%)	(44.02%)	(3.49%)	(47.35%)	(51.26%)
σ_t	0.2191	0.2139	0.1953	0.2197	0.1802	0.2033
	(0.37%)	(1.5%)	(1.50%)	(1.74%)	(4.51%)	(1.11 %)
p						4.7028
						(39.23 %)
SEE	1.3798	0.33264	0.29535	0.3231	0.2833	0.1946

Estimates are based on the absolute error objective function. The values reported as the parameter estimates are the mean of 10 parameter estimations done for each day separately and in brackets standard deviations of the parameters are reported. With SSE we denote the average value of the objective function over 10 days.

Table A.8: Parameter estimates - Absolute error objective function (VG)

	VG	VG-N	VG-SC
θ	-2.6961	-0.1625	-0.5257
	(59.92%)	(24.60%)	(1.46%)
ν	0.0081	4.3038	0.0366
	(0.25 %)	(303.86 %)	(1.46%)
σ_ν	0.0021	0.0713	0.2278
	(0.28%)	(5.70%)	(0.63%)
μ_{OJ}			-0.1518
			(34.25%)
σ_{OJ}		0.1766	0.0022
		(5.89 %)	(0.15%)
p			4.4763
			(176.80%)
SEE	0.7691	0.5227	0.5096

Estimates are based on the absolute error objective function. The values reported as the parameter estimates are the mean of 10 parameter estimations done for each day separately and in brackets standard deviations of the parameters are reported. With SSE we denote the average value of the objective function over 10 days.

A.3.1 Parameter estimates for data with only one maturity

Table A.9: Parameter estimates - Absolute error objective function (Stochastic volatility models)

	BS	Heston	Bates	SVOJ	SVRJOJ	SVRJOJ-SC
μ_{RJ}			-0.1781		-0.1869	-0.0773
			(11.90%)		(11.57%)	(5.18 %)
σ_{RJ}			0.1353		0.3897	0.23.04
			(12.52%)		(27.00%)	(13.31 %)
λ			0.1973		0.2978	0.2978
			(5.51%)		(13.25%)	(13.25 %)
μ_{OJ}						0.2839
						(80.79%)
σ_{OJ}				0.0573	0.0634	0.0457
				(1.46%)	(1.62%)	(1.48%)
κ		0.01	1.8234	1.0569	0.0887	0.5281
		(0.01%)	(200.93%)	(3.39%)	(11.34%)	(24.56%)
σ		1.2550	1.4657	0.763	0.9099	0.5905
		(2.71%)	(50.19%)	(40.7%)	(20.69%)	(7.91%)
σ_σ		0.362	0.2276	0.6961	0.2695	0.2257
		(1.37%)	(0.71%)	(22.42%)	(3.52%)	(1.16%)
ρ		-0.4184	-0.4535	-0.4456	-0.4510	-0.3849
		(1.50%)	(31.74%)	(1.42%)	(8.33%)	(13.81%)
σ_t	0.2219	0.032	0.1089	0.011	0.1596	0.2267
	(0.29%)	(0.1%)	(7.58%)	(0.22%)	(7.76%)	(1.39 %)
p						3.8896
						(11.75 %)
SEE	4.7029	0.0349	0.0157	0.03198	0.0121	0.0117

Estimates are based on the absolute error objective function. The values reported as the parameter estimates are the mean of 5 parameter estimations done for each day separately and in brackets standard deviations of the parameters are reported. With SSE we denote the average value of the objective function over 5 days.

Table A.10: Parameter estimates - Absolute error objective function (VG)

	VG	VG-N	VG-SC
θ	-1.5891	-0.1706	-0.2212
	(11.67%)	(2.37%)	(5.57%)
ν	0.0273	0.6694	0.1444
	(0.45%)	(21.67 %)	(3.57%)
σ_ν	0.0001	0.0803	0.2605
	(0.01%)	(5.62%)	(2.62%)
μ_{OJ}			0.0151
			(61.23%)
σ_{OJ}		0.0333	0.0076
		(0.77 %)	(0.57%)
p			4.4143
			(48.23%)
SEE	0.5824	0.0306	0.0351

Estimates are based on the absolute error objective function. The values reported as the parameter estimates are the mean of 10 parameter estimations done for each day separately and in brackets standard deviations of the parameters are reported. With SSE we denote the average value of the objective function over 10 days.