UNIVERSITY OF GOTHENBURG
SCHOOL OF BUSINESS, ECONOMICS AND LAW

Pricing Currency Options with Bates Model: Analytical Tractability versus Empirical Misspecification

Author:
Oscar Thelander

Supervisors:
Stefano Herzel
Alexander Herbertsson

Abstract

In this thesis I complement the results from Bates (1996) wherein a Stochastic Volatility Jump-Diffusion model for pricing foreign currency options is introduced and evaluated against USD/DM foreign exchange options. I complement Bates results with two different calibration methodologies, nonlinear least-squares and the built-in MATLAB function `fmincon`, using the same dataset that was used in Bates (1996). The results shows that the nonlinear least-squares calibration exhibit parameter values closely related to that of Bates (1996) and performs well when testing the pricing performance across moneyness, thus confirming Bates results. For the `fmincon` calibration, certain implicit parameter values are improbable given the model specification. This also corresponds to a comparatively worse pricing performance than that of `lsqnonlin` and an overall inconsistent pricing with respect to theoretical interpretation.

Acknowledgement

I would like to thank my supervisor at University of Rome Tor Vergata, Stefano Herzel, for devoting such a great deal of time, effort, and commitment to guiding and challenging me throughout this process. I also would like to thank my supervisor at University of Gothenburg, Alexander Herbertsson, for helping me with comments and suggestions easing the transition between universities.

List of Figures

List of Tables

Contents

1 Introduction

In this thesis, the focus is on pricing options on foreign currencies using Bates (1996) stochastic volatility jump-diffusion model. For foreign currencies, a number of typical distributional properties has been documented. Firstly, volatility is time-varying as has been evidenced by numerous ARCH/GARCH studies (Bollerslev et al. 1992) or in for example Taylor (1995) where a survey of the exchange rate economics is provided for the 1980s and 1990s. Volatility clustering, the notion that periods of high volatility are contrasted with periods of low volatility, are observed for exchange rate returns (Cheung and Miu 2009). It has also been shown that options on foreign currencies, as is the focus of this thesis, exhibit *volatility smile*-properties (Beneder and Elkenbracht-Huizing 2003).

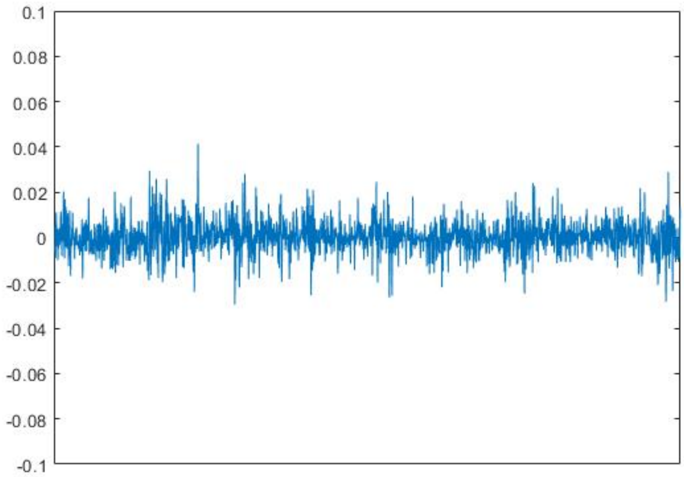

Figure 1: Log-differenced daily returns of USD/DM for Jan 1984-Jun 1991.

Secondly, the unconditional distribution of exchange rate returns exhibits excess kurtosis, i.e. a leptokurtic distribution. This has been shown for example in de Vries and Leuven (1992), Tucker and Pond (1988), and Friedman and Vandersteel (1982). There are however empirical results suggesting that there is an inverse relationship between the length of the holding period and the excess kurtosis, meaning that the return distribution tends towards the normal distribution as the holding period increases. Boothe and Glassman (1987) look at 4 major currencies and find that daily, weekly and monthly data exhibit non-normal distributions while quarterly data seems to be normally distributed.

Thirdly, as discussed in de Vries and Leuven (1992), the distribution of exchange rate returns may become skewed if, for example, there are dissimilar monetary policies between the countries in question. This has been shown in Akgiray et al. (1988) where they observe skewness for exchange-rate return distributions of minor trading currencies.

The seminal paper Black and Scholes (1973) introducing the Black-Scholes model changed the landscape for pricing and modelling options by providing a closed-form analytic pricing formula for European options. The Black-Scholes model however relies on some assumptions, one heavily discussed being the assumptions of a flat term structure of volatilities. This volatility-assumption has been widely challenged and several models have been developed to combat this.
Merton (1976) developed a constant-volatility model which includes a jump-component for the underlying asset. The need for including jumps is motivated by the presence of jumps observed in market prices as well as for risk management purposes (Tankov and Voltchkova 2009). In Dupire (1994) a local volatility model was developed where the volatility is a deterministic function of time and the underlying. Furthermore, there are numerous stochastic volatility models wherein the volatility follows a separate diffusion process and thus allows the volatility to develop stochastically over time. Examples are the Stein and Stein model where the volatility component follows a mean-reverting process (Stein and Stein 1991), the Heston model where the volatility component follows a mean-reverting square root process (Heston 1993), the SABR model which relates a forward under stochastic volatility (Hagan et al. 2002), the Constant Elasticity of Variance model (CEV model) which is based on a relationship between volatility and the underlying using a parameter measuring the leverage effect Cox (1975), and Bates model wherein stochastic volatility and jump-diffusion is combined (Bates 1996).
There are models that combine the local volatility with stochastic volatility and/or jumps, such as the model proposed by Jex et al. (1999) where the Heston model is modelled together with a local volatility correction component, or the JLSV model (Lipton and McGhee 2002) which adds jumps to the model proposed by Jex et al. (1999). Beyond local volatility and stochastic volatility, the model introduced in Hull and White (1990) has a stochastic interest rate-component in order to price interest-rate derivatives, where Amin and Ng (1993) expands this by proposing a model that combine stochastic interest-rate with a stochastic volatility-component.

For pricing options on foreign currencies, an extension of the Black-Scholes model was developed in Garman and Kohlhagen (1983) which incorporates the two interest rates related to currency pairs. Tucker et al. (1988) tests the Garman-Kohlhagen model against the CEV model for prediction purposes looking at options on five major currencies. They find that the CEV model provides better prediction pricing accuracy than the Garman-Kohlhagen model for intervals lower than five days but beyond that, the models' performances is not statistically distinguishable. Shastri and Wethyavivorn (1987) perform empirical tests of the Garman-Kohlhagen model, CEV model, Merton's jump-diffusion model and a pure jump-model on currency options for implied-volatility patterns. They find that the pure jump-model and Merton's jump-diffusion model outperform Garman-Kohlhagen and the CEV. Melino and Turnbull (1990) compare the Garman-Kohlhagen model to a stochastic volatility model for pricing foreign currency options and find that the stochastic volatility model provides a better fit for the USD/CAD exchange rate and in pricing ability. Bates model (Bates 1996) previously mentioned was developed for pricing foreign currency options. In the article the model was tested on USD/DM currency options from 1984 to 1991. The model was contrasted to the Garman-Kohlhagen model, a deterministic volatility/jump-diffusion model and a stochastic volatility model. The pricing results show that allowing for leptokurtic distributions through the non-constant volatility specifications improved the pricing ability, especially on in- and out-of-the-money options with a maturity less than 3 months. Furthermore, it is noted in the article that the stochastic volatility jump-diffusion model presented suffers from parameter instability through two tests that are developed in the article.

Building on this framework, the aim of this thesis is to complement the results in Bates (1996). This is done by implementing two different calibration methodologies, nonlinear ordinary least squares and the built-in MATLAB function `fmincon`, using the same market data. As such, the results can be directly contrasted to that of Bates, where inferences are made for consequences of using different calibration methodologies. The parameters implicit in the option data are also evaluated with respect to the interpretation from the analytical specification of the model, where I attempt to shed some light as to the potential problem of having an analytically pleasing model and whether empirical misspecification might occur as a consequence. This is an issue which has been raised in Bakshi et al. (1997), Bates (2003), and Mills and Patterson (2009) with regards to stochastic volatility models (with or without jump-diffusion) calibrated in empirical settings which pricing performance is better than constant-volatility models but whose resulting parameter values are implausible. Furthermore, an evaluation of the pricing performance under the different calibration methodologies is done for the framework in which Bates model was developed, that of pricing foreign currency option, using the Garman-Kohlhagen model as a benchmark.

The results show that the implicit parameter values of the nonlinear-least squares calibration are largely consistent with those presented in Bates (1996). For `fmincon`, the calibrated parameter values deviate to a greater extent from the Bates (1996) values, most notably the volatility of variance- and correlation-parameters. The pricing performance of the `lsqnonlin` specification is comparatively better than both the `fmincon` specification as well as the Garman-Kohlhagen model. The `fmincon` pricing performance is on an aggregate level the worst, but beats the Garman-Kohlhagen for in-the-money puts and out-of-the-money calls.

The structure of the thesis is the following: Chapter 2 provides a light informal theoretical framework for mathematical concepts used in the thesis. Chapter 3 presents the option pricing models for currency options that are used throughout, namely the Garman-Kohlhagen Model and Bates Stochastic Volatility Jump-Diffusion Model. Chapter 4 presents the calibration and pricing methodology used. Testing of the methodology is done in a theoretical setting using Monte Carlo simulation. The chapter is concluded with a discussion of numerical issues. Chapter 5 presents the foreign currency option transactions data from the Philadelphia Stock Exchange that is used. Chapter 6 presents the results from calibrated parameters implicit in the option prices for the different calibration methodologies as well as pricing evaluation which is contrasted with that of Bates (1996). Chapter 7 concludes the thesis.

2 Theoretical Framework

This chapter is an informal introduction to some of the theoretical framework for this thesis. Firstly, throughout this chapter we will work on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ where Ω denotes the set of all possible outcomes, $\mathcal{F}$ denotes the collection of all subsets and $\mathbb{P}$ denotes the probability measure, i.e. function mapping $\mathcal{F}$ to $[0, 1]$. If we denote X as a random variable, then X is a mapping from Ω to the real numbers $\mathbb{R}$. Together with the probability measure $\mathbb{P}$ on Ω, X determines the probability distribution on $\mathbb{R}$ (Shreve 2004). More than one probability measure is used in this thesis, changing from actual probability measure to risk-neutral probability measure, which is expanded upon in later chapters.

2.1 Stochastic Processes

Definition of Stochastic Process: *A stochastic process X is a family of random variables on a common probability space $(\Omega, \mathcal{F}, \mathbb{P})$ indexed by some set T such that*

$$\{X_t \ ; \ t \in T\} \tag{1}$$

From a financial perspective, we usually work with a finite final time T and then have a filtration $\mathcal{F}$ which is indexed by the time variable t, $\{\mathcal{F}_t; 0 \leq t \leq T\}$. We can interpret $\mathcal{F}_t$ as the information available at time t. Information increases over time such that the information contained in $\mathcal{F}_s$ is contained within $\mathcal{F}_t$ if $0 \leq s \leq t \leq T$. For our purposes, the stochastic process X_t can be thought of as the price of an asset at time t which thus is $\mathcal{F}_t$-measurable, i.e. that the information known at time t incorporates all the past prices of the asset up until time t for $0 \leq t \leq T$.

Two important features of stochastic processes are *martingales* and *Markov processes*. A Markov process says that the estimate of a future value of a stochastic process X only depends on the current value X_s and not its past history. More formally, when we are given a function f and another function g and $0 \leq s \leq t \leq T$, then

$$E[f(X_t)|\mathcal{F}_s] = g(X_s) \tag{2}$$

Martingales suggests that

$$E[X_t|\mathcal{F}_s] = X_s, \quad \text{for all } 0 \leq s \leq t \leq T \tag{3}$$

i.e. that the best guess of a future value is the current value.

2.2 Brownian Motion

A *Brownian motion*, also called *Wiener process*, is a stochastic process which essentially is the continuous-time version of a random walk. Starting with a *symmetric random walk*, where for each step, the path takes either one step up or one step down, with equal size, then it can be modelled as the process of tossing a fair coin a number of times. If we denote p to be the probability of tossing heads (H) and $(1-p)$ to be the probability of tossing tails (T), assume its equal probability of each outcome, and define each outcome ω_j, we have the following

$$X_i = \begin{cases} 1 & \text{if } \omega_j = H \\ -1 & \text{if } \omega_j = T \end{cases} \tag{4}$$

We can construct the symmetric random walk as the successive outcomes of the tosses

$$Y_k = \sum_{i=1}^{k} X_i \ , k = 1, 2, ... \text{ and } Y_0 = 0 \tag{5}$$

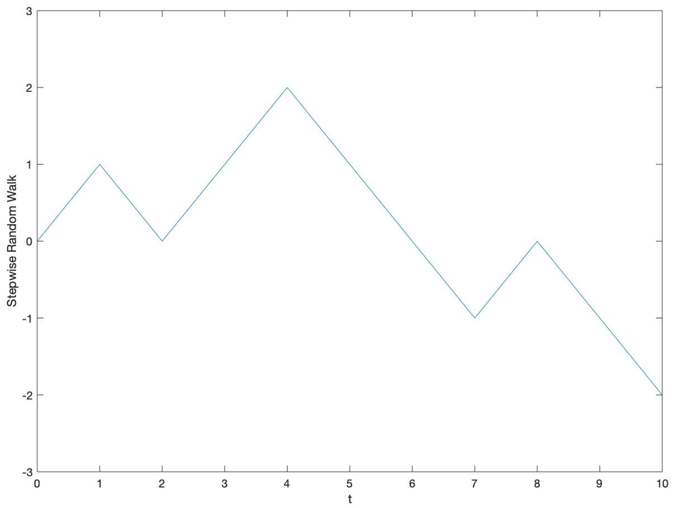

Figure 2: Sample Symmetric Random Walk path.

In order to obtain the Brownian motion we decrease the step sizes and increase the time of the symmetric random walk. We fix a positive integer n and define the *scaled symmetric random walk* as

$$W^{(n)}(t) = \frac{1}{\sqrt{n}} Y_{nt}, \tag{6}$$

provided nt is an integer. We obtain the Brownian motion as $n \to \infty$. A more formal definiton found in Shreve (2004):

Definition of Brownian Motion: Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. For each $\omega \in \Omega$, suppose there is a continuous function W_t of $t \geq 0$ that satisfies $W_0 = 0$ and that depends on ω. Then W_t, $t \geq 0$, is a Brownian motion if for all $0 = t_0 < t_1 < \cdots < t_m$ the increments

$$W_{t_1} = W_{t_1} - W_{t_0}, W_{t_2} - W_{t_1}, \; \cdots, W_{t_m} - W_{t_{m-1}} \tag{7}$$

are independent and each of these increments is normally distributed with

$$\mathbb{E}[W_{t_{i+1}} - W_{t_i}] = 0, \tag{8}$$

$$Var[W_{t_{i+1}} - W_{t_i}] = t_{i+1} - t_i \tag{9}$$

The Brownian motion is both a martingale and a Markov process.

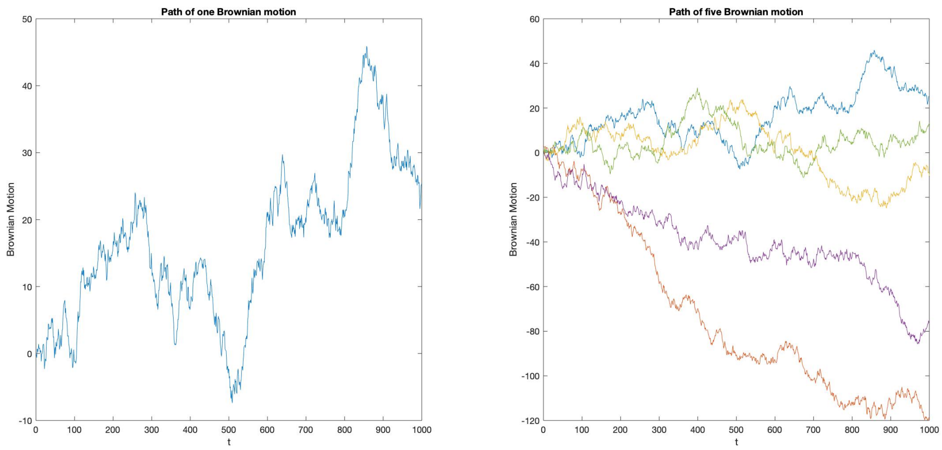

Figure 3: Sample Brownian motion paths.

2.3 Stochastic Differential Equations

In order to introduce stochastic differential equations, lets first consider an ordinary differential equation. The idea is that we are given a functional relationship:

$$f(t, x(t), x'(t), x''(t), ...) = 0, \;\; 0 \leq t \leq T \tag{10}$$

which involves the unknown function $x(t)$, time t and its derivatives $(x'(t), x''(t), ..)$ The solution of the differential equation (10) is to find the function $x(t)$ which describe the dynamics of the process over a given period of time. The simplest version is an ordinary differential equation of order 1:

$$dx(t) = a(t, x(t))dt, \;\; x(0) = x_0 \tag{11}$$

A stochastic differential equation is an extension of the ordinary one, introducing randomness into the equation. This could simply be done by introducing randomness into the initial condition $x(0)$ but for our purposes, this is introduced also via an additional term:

$$dX_t = a(t, X_t)dt + b(t, X_t)dW_t,\ X_0(\omega) = Y(\omega) \tag{12}$$

W_t here denote a Brownian motion, which have the properties explained above. $a(t, X_t)$ and $b(t, X_t)$ are in this example deterministic functions, and the solution, X, is (if it exists) a stochastic process. A popular framework for pricing financial derivatives, where X_t denotes the price of a stock at time t, is one which follows the stochastic differential equation

$$dX_t = \mu X_t dt + \sigma X_t dW_t \tag{13}$$

which is called a Itô drift-diffusion process, where $\mu X_t dt$ denotes the drift-term and $\sigma X_t dW_t$ denotes the diffusion-term. (Klebaner 2005)

2.4 Itô Stochastic Integral and Itô Lemma

Equation (12) can be stated in integral form:

$$X_t = X_0 + \int_0^t a(s, X_s)ds + \int_0^t b(s, X_s)dW_s,\ 0 \leq t \leq T \tag{14}$$

where the first integral $\int_0^t a(s, X_s)ds$ is a Riemann integral, which can be approximated using the Fundamental Theorem of Calculus. The second integral $\int_0^t b(s, X_s)dW_s$ is however an *Itô stochastic integral*, which thus makes Equation (14) into a Itô stochastic differential equation. The Itô stochastc integral is a generalization of the Riemann integral when the integrands and integrators are stochastic processes. The result of the integration is a random variable which is defined as the limit of a sequence of random variables, in general:

$$\int_a^b g(s)dW_s = \sum_{k=0}^{n-1} g(t_k)[W(t_{k+1}) - W(t_k)] \tag{15}$$

(Björk 2009)

Itô lemma is a method for finding the differential of a time-dependent stochastic process, which can be interpreted as the stochastic calculus version of the chain rule from regular calculus. Using the stochastic differential equation:

$$dX_t = \mu dt + \sigma dW_t \tag{16}$$

and a twice-differentiable function $f(t,x)$, Itô's Lemma gives:

$$\begin{aligned} df &= \frac{\partial f}{\partial t}dt + \frac{\partial f}{\partial x}dx + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}dx^2 + ... \\ df &= \frac{\partial f}{\partial t}dt + \frac{\partial f}{\partial x}(\mu dt + \sigma dW_t) + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}(\mu dt + \sigma dW_t)^2 + ... \\ df &= \Big[\frac{\partial f}{\partial t} + \frac{\partial f}{\partial x}\mu + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}\sigma\Big]dt + \frac{\partial f}{\partial x}\sigma dW_t \end{aligned} \tag{17}$$

where in the limit as $dt \to 0$: $dt^2 \approx 0$, $dtdW_t \approx 0$, $dW_t^2 \approx dt$. The partial derivatives beyond $\frac{\partial^2 f}{\partial x^2}$ are thus negligible.

For a popular application to option pricing, let $f(x) = ln(x)$ and consider the following stochastic differential equation:

$$dX_t = \mu X_t dt + \sigma X_t dW_t \tag{18}$$

Itô's Lemma:

$$\begin{aligned} d(ln(X_t)) &= \frac{\partial f(X_t)}{\partial t}dt + \frac{\partial f(X_t)}{\partial x}dx + \frac{1}{2}\frac{\partial^2 f(X_t)}{\partial x^2}dx^2 + ... \\ d(ln(X_t)) &= \frac{\partial f(X_t)}{\partial t}dt + \frac{\partial f(X_t)}{\partial x}(\mu X_t dt + \sigma X_t dW_t) + \frac{1}{2}\frac{\partial^2 f(X_t)}{\partial x^2}(\mu X_t dt + \sigma X_t dW_t)^2 + ... \\ d(ln(X_t)) &= \frac{1}{X_t}(\mu X_t dt + \sigma X_t dW_t) - \frac{1}{2}\frac{1}{X_t^2}(\mu X_t dt + \sigma X_t dW_t)^2 \\ d(ln(X_t)) &= \Big[\mu - \frac{1}{2}\sigma^2\Big]dt + \sigma dW_t \\ \int_0^t d(ln(X_t)) &= \int_0^t \Big[\mu - \frac{1}{2}\sigma^2\Big]dt + \int_0^t \sigma dW_t \\ X_t &= X_0 e^{(\mu - \frac{1}{2}\sigma^2)t + \sigma W_t} \end{aligned} \tag{19}$$

The last equation is called a Geometric Brownian Motion (GBM), which is a stochastic process solution to the stochastic differential equation in Equation (17). The GBM is used to model the underlying asset price in the Black-Scholes model. (Klebaner 2005)

2.5 Poisson Process

The Poisson process is used to model the number of occurrences of certain event in a certain time frame. This could for example be the number of phone calls in a day or the number of visitors in a store in a day. For financial derivatives, the Poisson process is used to model jumps in the asset prices. More formally:

Definition of Poisson process: A Poisson process, N_t, is a stochastic process with intensity $\lambda > 0$ such that:

1. The process starts at $N_0 = 0$.
2. $N_t - N_s$ for $t > s$ are independent, i.e. the increments are independent.
3. The increments $N_t - N_s$, $t > s$ has a Poisson distribution with parameter $\lambda(t-s)$:

$$P(N_t - N_s = k) = e^{-\lambda(t-s)} \frac{\lambda^k (t-s)^k}{k!} \tag{20}$$

(Calin 2015)

In Figure 4, a sample path of the Poisson process is shown where the number of jumps is 15, exactly equal to the expected value $\lambda = 15$.

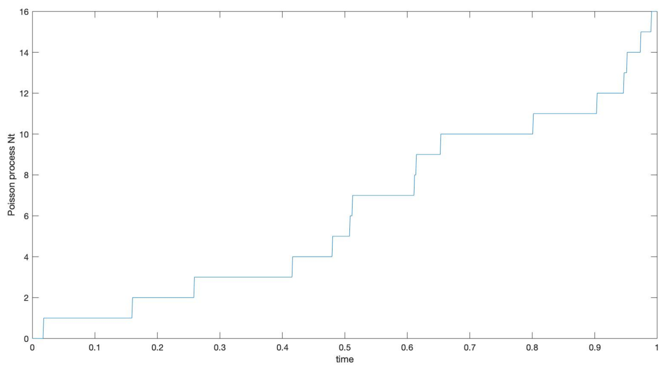

Figure 4: Sample Poisson process.

3 Models for Currency Options

3.1 The Garman-Kohlhagen Model

The Garman-Kohlhagen Model was introduced in 1983 as an application of the original Black-Scholes model (Black and Scholes 1973) for pricing foreign currency options. The model follows a stochastic differential equation (SDE) for the dynamics of the spot price:

$$dS_t = \mu S_t dt + \sigma S_t dZ_t \tag{21}$$

where S_t denotes the spot price of the deliverable currency, μ denotes the instantaneous expected rate of appreciation of the foreign currency, σ denotes the volatility of the spot currency price and Z_t is a standard Brownian motion. Solving the SDE using Itô Calculus amounts to the following analytic solution for S_t:

$$S_t = S_0 e^{(\mu - \frac{1}{2}\sigma^2)t + \sigma W_t} \tag{22}$$

which means that S_t follows a GBM and is log-normally distributed. Utilizing Itô's Lemma for deriving the partial differential equation (PDE) for the value of an option and applying arbitrage arguments[1], Garman and Kohlhagen derives the following closed-form analytic formula for the price of a European call option, C_{GK}:

$$C_{GK} = e^{-r_f(T-t)} S_t N(d_1) - e^{-r_d(T-t)} K N(d_2) \tag{23}$$

where

$$\begin{aligned} d_1 &= \frac{ln(S_t/K) + (r_d - r_f - \frac{\sigma^2}{2})(T-t)}{\sigma\sqrt{T-t}} \\ d_2 &= d_1 - \sigma\sqrt{T-t} \end{aligned} \tag{24}$$

where K denotes the strike price of the option, $(T-t)$ denotes time to maturity, $N(\cdot)$ denotes the cumulative normal distribution function, and r_d, r_f denotes the domestic and foreign interest rate, respectively.

The price of a European put option, P_{GK}, is calculated as:

$$P_{GK} = e^{-r_d(T-t)} K N(-d_2) - e^{-r_f(T-t)} S_t N(-d_1) \tag{25}$$

with the same notational implications as above.

[1]For a full treatment of the derivation for obtaining the pricing formulas, see Garman and Kohlhagen (1983)

3.2 Bates Model

Bates model was introduced in the paper which this thesis largely follow (Bates 1996). The model incorporates both a stochastic volatility-component and a jump diffusion-component. Bates model thus proposes that S_t, the underlying exchange rate, follows a geometric jump diffusion with the variance, V_t, following a mean-reverting square root process:

$$\begin{aligned} dS_t &= (\mu - \lambda\bar{k})S_t dt + \sqrt{V_t}S_t dZ_t + kS_t dP_t \\ dV_t &= \kappa(\eta - V_t)dt + \sigma_v\sqrt{V_t}dW_t \\ E[dZ_t dW_t] &= \rho dt \\ prob(dP_t^* = 1) &= \lambda dt, \quad ln(1+k) \sim N(ln(1+\bar{k}) - \frac{1}{2}\delta^2, \delta^2) \end{aligned} \tag{26}$$

where μ denotes the instantaneous expected rate of appreciation of the foreign currency, λ denotes the annual frequency of jumps, k denotes the size of the random percentage jump conditional on a jump occurring, P_t denotes a Poisson process with intensity λ, κ denotes the mean-reversion speed for the variance, η denotes the long-term variance level, σ_v denotes the volatility of variance, ρ denotes the correlation between the standard Brownian motions Z_t and W_t, and δ denotes the standard deviation of the random jump size.

Volatility, as in the case of the Garman-Kohlhagen model is modelled as constant, here follows the mean-reverting process popularized by Cox-Ingersoll-Ross also known as the CIR-process (Cox et al. 1985). For the stochastic volatility-component, high values of κ smoothes out the process as any deviations from the long-term mean, η, are quickly removed. The kurtosis of the distribution is driven by σ_v while ρ governs the skewness. For the jump diffusion-component, $\bar{k}$ affects the skewness with positive (negative) values implying a positively (negatively) skewed distribution. The standard deviation of the random jump size, δ, affects the kurtosis of the distribution as an increase in δ leads to an increase in kurtosis as the variance of the jumps increases. For λ, the annual frequency of jumps or intensity parameter for the Poisson process P_t, higher values indicate more frequent jumps in the process and as such a higher overall volatility which thus consequently increases the kurtosis of the distribution. As such, Bates proposes that the model having multiple sources governing the different distributional properties of the underlying process allows more channels through which to more accurately fit the market data (Bates 1996).

The risk-neutral measure for Bates model, the measure used for pricing purposes and the parameters that will be referenced to throughout this thesis is defined by:

$$\begin{aligned} dS_t &= (b - \lambda^* \bar{k}^*)S_t dt + \sqrt{V_t} S_t dZ_t^* + k^* S_t dP_t^* \\ dV_t &= \kappa^*(\eta^* - V_t)dt + \sigma_v \sqrt{V_t} dW_t^* \\ E^{\mathbb{Q}}[dZ_t^* dW_t^*] &= \rho dt \\ prob(dP_t = 1) &= \lambda^* dt, \quad ln(1 + k^*) \sim N(ln(1 + \bar{k}^*) - \frac{1}{2}\delta^2, \delta^2) \end{aligned} \tag{27}$$

where $b = r_d - r_f$ denotes the interest-rate differential between r_d and r_f, the domestic and foreign interest rates, as the framework for this thesis is with the pricing of foreign currency options. For stock options this corresponds to $r - q$, where r denote the continuous risk-free rate and q denote the continuous dividend yield. The starred variables represent the risk-neutral versions of the *true* variables presented in Equation (26). The difference as opposed to the parameters in Equation (26) is that the risk-neutral versions take into account the risk premium inherent in the true parameters from jump risk and volatility risk (Bates 1996). The risk-neutral measure will be more thoroughly explained in Chapter 4.

4 Calibration Methodology and Numerical Issues

One of the main motivations for the development of advanced option pricing models with or without stochastic volatility and jump diffusion-components is to better fit market data. The process of fitting theoretical models to market data, referred to as model calibration, entails using optimization techniques for identifying the set of parameters in the model for which the estimated prices from the model well fit the market prices. As such, the better the calibration, the more valuable the model is as a tool for pricing and risk management-purposes. Denoting the set of parameters in a model as θ, the calibration is often conducted by minimizing the value of a loss function conditional on the parameters in the model, $L(\theta)$:

$$\hat{\theta} = \underset{\theta}{\operatorname{argmin}}\, L(\theta) \tag{28}$$

For Bates model, the set of parameters to be estimated are $\theta =< \kappa^*, \eta^*, \sigma_v, \rho, \delta, \lambda^*, \bar{k}^* >$ where the first four parameters refer to the stochastic volatility-component and the latter three refer to the jump diffusion-component as per the description in Chapter 3. As the purpose of this thesis is to complement the results from Bates (1996), I use the methodology therein, where the option pricing residual utilized is:

$$e_i = \frac{O_i^{(M)}}{S_i} - \frac{O_i^{(E)}}{S_i} \tag{29}$$

where $O_i^{(M)}$ denotes the market price of the i'th American option, S_i denotes the underlying exchange rate corresponding to the i'th American option, and $O_i^{(E)}$ denotes the estimated price of the i'th American option from the model given the contractual terms and the parameters. As such, the estimated price of an option using Bates model is a function of $< b, S, T, K, \theta >$. The methodology of using the ratio of option prices to the underlying is motivated by Bates for the problems of cross-sectional heteroskedasticity inherent in using regular dollar and percentage option pricing errors as a metric for in-sample fit. As such, Bates proposes that using a ratio of the option price to the underlying provides a better framework for comparison across different datasets (Bates 2003).

In Bates (1996), the parameters implicit in the option prices are calibrated using a nonlinear generalized least-squares methodology. This is motivated in Bates (1996) through which the methodology is not putting equal weight on the heavily traded near-the-money options relative to the less heavily traded in- and out-of-the-money options. In this thesis, the parameters implicit in the market data are calibrated using two other methodologies: nonlinear least-squares and the built-in MATLAB function `fmincon`.

4.1 lsqnonlin & fmincon

The first calibration methodology, nonlinear least-squares, is computed using the built-in MATLAB function `lsqnonlin` which solves nonlinear least-squares problems subject to bounds on the parameter inputs. Mathematically, the algorithm solves the problem:

$$\min_x ||f(x)||_2^2 = \min_x (f_1(x)^2 + f_2(x)^2 + \cdots + f_n(x)^2) \quad \text{s.t. } lb \leq x \leq ub \tag{30}$$

where $f(x)$ is provided as a vector, which for this thesis is a vector of the option pricing residuals in Equation (29), and lb, ub denote the lower and upper bounds of the parameters, respectively. `lsqnonlin` provides the possibility of using two different algorithms: `Levenberg-Marquardt` and `Trust-Region-Reflective`, where in this thesis the latter algorithm is used. The `Trust-Region-Reflective`-algorithm works by approximating the function f with a simpler function q which resembles the behaviour of the function f in a *neighbourhood* N around the current point x. This neighbourhood is what is called the *trust-region* and from there a trial step s is computed by approximately minimizing the step over N which is the *trust-region subproblem*:

$$\min_s q(s), s \in N \tag{31}$$

The current point, x, is updated to $x + s$ if $f(x) > f(x + s)$. If $f(x) \leq f(x + s)$, the current point is not updated and the trial step is repeated. These steps are continued until the algorithm converges, which is evaluated through that the final change in the sum of squares relative to the initial value is less than a set value for the function tolerance (MATLAB 2020b).

The second calibration methodology is computed using the built-in MATLAB function `fmincon`. `fmincon` is a local minimum optimizer that works through a gradient-based method on nonlinear multivariate functions:

$$\min_x f(x) \quad \text{s.t. } lb \leq x \leq ub \tag{32}$$

where x, lb, ub are passed as vectors and $f(x)$ is a function that returns a scalar. Inequalities and equalities can also be added as constraints but in this thesis only bounds for the parameter values are set. For `fmincon` as compared to `lsqnonlin`, a specific loss function to minimize needs to be provided as per the methodology of Equation (28). In this thesis, I opt to use Mean Absolute Error (MAE):

$$MAE = \frac{1}{N} \sum_{i=1}^{N} \Big| \frac{O_i^{(M)}}{S_i} - \frac{O_i^{(E)}}{S_i} \Big| \tag{33}$$

as to complement the nonlinear least-squares calibration rather than using for example a Mean Square Error[2] (MSE) methodology (Brooks 2019).

[2]

$$MSE = \frac{1}{N} \sum_{i=1}^{N} \Big(\frac{O_i^{(M)}}{S_i} - \frac{O_i^{(E)}}{S_i} \Big)^2 \tag{34}$$

For `fmincon`, there are five different algorithms that can be used where in this thesis the `interior-point`-algorithm is used. The algorithm works by solving a sequence of approximate minimization problems by calculating the Hessian using a quasi-Newton approximation. For each iteration, the algorithm takes a step in the parameters. If the iteration does not improve the approximation, the step is rejected and a new step is attempted as to converge to a solution. Convergence is obtained when the objective function is non-decreasing in feasible directions subject to a tolerance level. A strong point for the `interior-point`-algorithm is that it can handle the function to minimize returning infinite or NaN values (MATLAB 2020a).

The nonlinear least squares differs from the generalized version used in Bates (1996) in that it does not include an account for variance-covariance matrix of the residuals. The `fmincon`-methodology with MAE can be contrasted to that of Bates (1996) for penalizing all the pricing residuals equally rather than the squaring of residuals, which consequently penalizes larger errors more heavily.

For both calibration methodologies, the same inputs are used. There are no prescribed constraints but bounds for the parameters to be calibrated are set. In Bates (1996) several calibrations are done, full calibrations for the whole dataset 1984-1991 as well as biyearly subsamples: 1984-1985, 1986-1987, 1988-1989, 1990-1991. In this thesis, the main focus will be on the full calibration over the whole dataset. For the calibrations, the initial values for the parameters are those which were reported in Bates (1996), with lower and upper bounds corresponding to reasonable theoretical boundaries for the parameters in the model. The initial parameter values for the calibration with bounds are shown in Table 1.

Variables	η^*	κ^*	σ_v	ρ	$\bar{k}^*$	δ	λ^*
x_0	0.024	0.78	0.343	0.078	-0.001	0.019	15.01
lb	0.01	0.01	0.01	-1.00	-1	0.01	0.01
ub	10	10	10	1	10	10	20

Table 1: Initial values for main calibration with upper and lower bounds.

Noteworthy is that the annual frequency of jumps, λ^*, is 15.01 while the average random percentage jump, $\bar{k}^*$, is -0.1% with a volatility of 1.9% so although the annual frequency of jumps is high, the estimated average size of the jumps and its volatility are low. The long-term variance level of 0.024 corresponds to a long-term volatility level of 15.49% which contrasted to an average implied volatility for all the options of 13.32% seems reasonable. The calibrated correlation coefficient ρ from Bates (1996) depicts a slight positive relationship. For options on stocks, the relationship between volatility and price has been extensively studied, termed *leverage effect*. The leverage effect occurs as $\rho < 0$ with the motivation that as the price of the underlying stock decreases, the debt-to-equity ratio of the firm increases which thus makes the firm riskier and should increase the future expected volatility (Yu 2005). For foreign currency options the same rationale with regards to the sign is not as theoretically concluded but rather situational as discussed in Chapter 1.

4.2 Pricing Methodology

4.2.1 Numerical Integration

Firstly, as mentioned in Chapter 3, the risk-neutral measure[3] is a probability measure, $\mathbb{Q}$, as opposed to the true probability measure, $\mathbb{P}$. The risk-neutral measure stems from the fundamental theorem of asset pricing which states that a financial market is free of arbitrage if and only if there exists a risk-neutral probability measure $\mathbb{Q}$. Using a risk-neutral measure makes it possible to discard risk-preferences in the pricing framework and price to the fair value as all assets have the same expected rate of return (Yor 2008). As such, the price of European call option can be computed by taking the discounted expected value of the future payoff under the risk neutral measure:

$$\begin{aligned} V(S_t, T) =& e^{-r(T-t)} \mathbb{E}^{\mathbb{Q}}[(S_T - K)^+ | \mathcal{F}_t] \\ =& e^{-r(T-t)} \int_0^\infty (S_T - K)^+ q(S_T|\mathcal{F}_t) dS_T \\ =& e^{-r(T-t)} \int_K^\infty (S_T - K) q(S_T|\mathcal{F}_t) dS_T \end{aligned} \tag{35}$$

where S_t denotes the underlying, K denotes the strike price, r denotes the risk-free discount rate, $(T-t)$ denotes time to maturity, $\mathcal{F}_t$ denotes the filtration at time t or information flow up to and including t, and $q(S_T|\mathcal{F}_t)$ denotes the risk neutral density function. If the probability density function is known in closed form, the option price can be obtained by a single integration. For most extensions of the Black-Scholes framework the closed-form density function is not known which makes this methodology infeasible. Popular pricing methodologies for option pricing models without known density function instead utilize the *characteristic function* (Schmelzle 2010).

The characteristic function of any random variable completely define its probability distribution as there is a one to one relationship between the characteristic function and the probability density function. More formally, the characteristic function $f_X(\phi) = \mathbb{E}[e^{i\phi X}]$ of a $\mathbb{R}$[4] random variable X is defined for real numbers ϕ by taking the expectation of the complex transformation $e^{i\phi X}$, where i denotes the imaginary unit[5]. If $g_X(x)$ denote the probability density function of the random variable

[3] See for example Yor (2008) for reference on risk-neutral measure.
[4] $\mathbb{R}$ denotes real number.
[5] $i = \sqrt{-1}$

then:

$$f_X(\phi) = \mathbb{E}[e^{i\phi X}] = \int_{-\infty}^{\infty} e^{i\phi x} g_X(x) dx \tag{36}$$

is the integral which defines the expected value and is the *Fourier transform* of the density function $g_X(x)$. Linking the characteristic function back to the probability distribution is done using an inverse Fourier transform:

$$G_X(x) = \frac{1}{2} - \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{e^{-i\phi x} f_X(\phi)}{i\phi} d\phi \tag{37}$$

where $G_X(x)$ denotes the cumulative distribution function (Schmelzle 2010).

For option pricing, the methodology is then to evaluate an integral of the payoff function for an option over the probability distribution obtained through the inversion methodology. For Bates model the density function is not known but the characteristic functions is:

$$\begin{aligned}
f_{Bates(\phi)} &= exp(C + DV_0 + i\phi ln(S_t))\cdot \\
&\quad exp(\lambda^*(T-t)(1+\bar{k}^*)[(1+\bar{k}^*)^{i\phi} e^{\delta^2(-\frac{1}{2}i\phi + \frac{(i\phi)^2}{2})} - 1] - \lambda^*(T-t)\bar{k}^* i\phi) \\
C &= (r_d - r_f)i\phi(T-t) + \frac{\kappa^*\eta^*}{\sigma_v^2}\Big[(\kappa^* + \lambda_{Volrisk} - \rho\sigma_v i\phi + d)(T-t) - 2ln\Big(\frac{1-\epsilon e^{d(T-t)}}{1-\epsilon}\Big)\Big] \\
D &= \frac{\kappa^* + \lambda_{Volrisk} - \rho\sigma_v i\phi - d}{\sigma_v^2}\Big(\frac{1-e^{-d(T-t)}}{1-\epsilon e^{-d(T-t)}}\Big) \\
\epsilon &= \frac{\kappa^* + \lambda_{VolRisk} - \rho\sigma_v i\phi - d}{\kappa^* + \lambda_{Volrisk} - \rho\sigma_v i\phi + d} \\
d &= \sqrt{(\kappa^* + \lambda_{Volrisk} - \rho\sigma_v i\phi)^2 + \sigma_v^2(i\phi + \phi^2)}
\end{aligned} \tag{38}$$

where S_t, λ^*, $\bar{k}^*$, δ, r_d, r_f, κ^*, η^*, σ_v, and ρ correspond to the parameters in Chapter 2. $\lambda_{Volrisk}$, the volatility risk premium, is set to zero. The pricing methodology used in this thesis involves pricing the options according to the framework developed in Heston (1993):

$$\begin{aligned}
C &= S_t e^{-r_f(T-t)} P_1 - K e^{-r_d(T-t)} P_2 \\
P &= C + K e^{-r(T-t)} P_2 - S_t e^{-q(T-t)} P_1 \\
P_j &= \frac{1}{\pi} \int_0^{\infty} \mathbb{R}\Big[\frac{e^{-i\phi ln(K)} f_j(\phi)}{i\phi}\Big] d\phi
\end{aligned} \tag{39}$$

where the characteristic function is inverted in order to get the probabilities in the pricing equations C for the call option and P for the put option. The integration in Equation (39) is computed numerically using global adaptive quadrature which is a methodology that works by adaptively refining subintervals of the region that is integrated (Gander and Gautschi 2000). Fortunately for this thesis, the methodology is implemented in the built-in MATLAB function `optByBatesNI`. In

Table 2, the function is tested against the theoretical put option pricing that is provided in Bates (1996):

Option Parameters	Strike	Bates (1996)	European Price NI
$\theta = 0.0225$	38	0.374	0.377
$\sigma_v = 0.15$	39	0.662	0.665
$\rho = 0$	40	1.074	1.077
$\lambda^* = \bar{k}^* = \delta = 0$	41	1.617	1.620
$V_0 = 0.0225$	42	2.283	2.285
$\theta = 0.0225$	38	0.575	0.579
$\sigma_v = 0.15$	39	0.902	0.906
$\rho = 0$	40	1.334	1.338
$\lambda^* = \bar{k}^* = \delta = 0$	41	1.874	1.878
$V_0 = 0.04$	42	2.515	2.518
$\theta = 0.0225$	38	0.369	0.372
$\sigma_v = 0.30$	39	0.648	0.652
$\rho = 0$	40	1.056	1.060
$\lambda^* = \bar{k}^* = \delta = 0$	41	1.601	1.605
$V_0 = 0.0225$	42	2.274	2.276
$\theta = 0.0225$	38	0.369	0.372
$\sigma_v = 0.15$	39	0.658	0.662
$\rho = 0.1$	40	1.074	1.078
$\lambda^* = \bar{k}^* = \delta = 0$	41	1.621	1.624
$V_0 = 0.0225$	42	2.289	2.292
$\theta = 0.0125$	38	0.356	0.360
$\sigma_v = 0.15$	39	0.619	0.626
$\rho = 0$	40	1.018	1.026
$\lambda^* = 2, \bar{k}^* = 0, \delta = 0.07$	41	1.567	1.574
$V_0 = 0.0125$	42	2.252	2.256

Table 2: European put option estimation of theoretical option values, Bates (1996) values in parenthesis. $\mathrm{S}_t = 40, (T - t) = 0.25, b = 0.02$.

The average pricing error and average percentage error over the theoretical put options in Table 2 is -0.004 and -0.48%, respectively, so overall the pricing methodology employed seems to price the theoretical options slightly above the methodology used in Bates (1996). The pricing methodology in Bates (1996) is similar with regards to using numerical integration but the algorithm for numerically computing the integration is different which could influence the consistent overpricing of 0.002-0.008.

4.2.2 Early Exercise Premium Approximation

The options that are used in this thesis are American, that is, they can be exercised prior to expiration. As both of the option pricing models considered in this thesis and the pricing methodology used are derived for pricing European options, where the option can only be exercised at maturity, the early exercise premium included in the American option price is not included in the pricing methodology. In this thesis, the *Quadratic Approximation Method* from Barone-Adesi and Whaley (1987) is implemented to approximate the early exercise premium. The following is the derivation for the American call option value where the notation closely follow that of Barone-Adesi and Whaley (1987).

The methodology is based in a framework of the Black-Scholes model wherein the PDE governing the movement of the option value, V, is given by:

$$\frac{1}{2}\sigma^2 S^2 V_{SS} + bSV_S - rV + V_t = 0 \tag{40}$$

where V_S, V_{SS} denotes the first and second partial derivative of V with respect to S, respectively. b denotes the cost of carry, which in the framework of currency options is the interest rate differential. Barone-Adesi and Whaley (1987) use the fact that if Equation (40) applies to both American and European call options, then it should also apply to the early exercise premium. As such, the early exercise premium for a call, $e(S,t)$ is defined by:

$$e(S,t) = C(S,t) - c(S,t) \tag{41}$$

where $C(S,t)$, $c(S,t)$ denotes the American and European call option price, respectively. Equation (40) is thus modified to:

$$\frac{1}{2}\sigma^2 S^2 e_{SS} + bSe_S - re + e_t = 0 \tag{42}$$

At this point, a couple of simplifications are implemented:

1. Denote T as time to maturity in place of $(T-t)$.
2. Equation (42) is multiplied by $2/\sigma^2$
3. Denote $M = 2r/\sigma^2$ and $N = 2b/\sigma^2$

The value of the early exercise premium of a call is assumed to follow $e(S,h) = h(T)f(S,h)$, where $h(T)$ is an arbitrary function of time to maturity and thus expressing the early exercise premium as a function of time to maturity and the underlying asset price. As such, the corresponding partial derivatives follows: $e_{SS} = hf_{SS}$, $e_S = hf_S$, $e_T = h_T f + hh_T f_K$ which yields the PDE

$$S^2 f_{SS} + NSf_S - Mf[1 + (h_T/rh)(1 + hf_h/f)] = 0 \tag{43}$$

By choosing $h = 1 - e^{-rT}$ and substituting into (43):

$$S^2 f_{SS} + NSf_S - (M/h)f - (1-h)Mf_h = 0 \tag{44}$$

Here an approximation is made, the $(1-h)Mf_h$-term is set equal to 0. This is motivated with that for commodity options with very short (long) time to expiration (T approaches 0 (∞)), f_h approaches 0 (h approaches 1), and the $(1-h)Mf_h$-term disappears. This yields:

$$S^2 f_{SS} + NSf_S - (M/h)f = 0 \tag{45}$$

where it can be noted that Equation (45) is a second-order ordinary differential equation (ODE). The ODE has two linearly independent solutions of the form aS^q. Substituting $f = aS^q$ into Equation (45) yields:

$$S^2 aq(q-1)S^{q-2} + NSaqS^{q-1} - (M/h)aS^q = 0 \tag{46}$$

as the partial derivatives $f_{SS} = aq(q-1)S^{q-2}$ and $f_S = aqS^{q-1}$. Rearranging yields:

$$aS^q[q^2 + q(N-1) - M/h] = 0 \tag{47}$$

which has the following two roots:

$$q_1 = [-(N-1) - \sqrt{(N-1)^2 + 4M/h}]/2$$

$$q_2 = [-(N-1) + \sqrt{(N-1)^2 + 4M/h}]/2$$

Note that $q_1 < 0$ and $q_2 > 0$ as $M/h > 0$ which follows because M/h is expressed as follows

$$M/h = \frac{\frac{2r}{\sigma^2}}{1 - e^{-rT}} \tag{48}$$

The general solution to the second-order ODE is

$$f(S) = a_1 S^{q_1} + a_2 S^{q_2} \tag{49}$$

where a_1 and a_2 are left to be determined. Since this is the case for the call option, having $q_1 < 0$ and $a_1 \neq 0$ leads to the fact that f approaches ∞ as S approaches 0. This is an undesirable property for the value of the early exercise premium of an American call option. Therefore, the first constraint to be set is that $a_1 = 0$ for call options, which leads to the following approximate value for the American call:

$$C(S,T) = c(S,T) + ha_2 S^{q_2} \tag{50}$$

At this point, it is important to discuss the appropriate constraints to impose on a_2. As $S = 0$, $C(S,T) = 0$ since both of the terms on the RHS in Equation (50) are equal to zero. As S increases, the value of $C(S,T)$ increases because of both of the RHS terms, assuming $a_2 > 0$. To represent the value of the American call option the function should touch, but not intersect, the boundary imposed by the early exercise premium of the American Call, $S - K$. The critical value of S implied by the point of tangency is denoted S^*, and the American call value should be represented by Equation (50) when $S < S^*$. When $S > S^*$, the American call value should theoretically immediately be exercised to the value $S - K$.

In order to find the critical price S^*, the exercisable value of the American call option is set to the value of $C(S^*, T)$:

$$S^* - K = c(S^*, T) + ha_2 S^{*q_2} \tag{51}$$

and the slope of the exercisable value of the call, that is the delta of the option equal to 1, is set equal to the slope of $C(S^*, T)$:

$$1 = e^{(b-r)T} N[d_1(S^*)] + hq_2 a_2 S^{*q_2 - 1} \tag{52}$$

which follows from the partial derivative of Equation (51) with respect to S^*. $d_1(S^*)$ denotes:

$$d_1(S^*) = \frac{ln(S^*/K) + (b + 0.5\sigma^2)T}{\sigma\sqrt{T}}$$

As such, there are two unknown parameters, a_2 and S^*, and two equations, Equation (51) and (52). Rewriting Equation (52) in order to isolate a_2 yields:

$$a_2 = \frac{1 - e^{(b-r)T} N[d_1(S^*)]}{hq_2 S^{*q_2 - 1}} \tag{53}$$

Substituting into Equation (51) and simplifying yields:

$$S^* - K = c(S^*, T) + \frac{S^*}{q_2}[1 - e^{(b-r)T} N[d_1(S^*)]] \tag{54}$$

where finding S^* needs to be solved iteratively. With S^* known, a_2 can be obtained in Equation (51). Thus, the approximate price of the American call option is given by:

$$\begin{aligned} C(S, T) &= c(S, T) + A_2(S/S^*)^{q_2}, \text{ when } S < S^* \\ C(S, T) &= S - K, \qquad\qquad\qquad \text{when } S \geq S^* \end{aligned} \tag{55}$$

where $A_2 = \frac{S^*}{q_2}[1 - e^{(b-r)T} N[d_1(S^*)]]$. Note that $A_2 > 0$ because q_2, S^*, $1 - e^{(b-r)T} N[d_1(S^*)]$ are positive when $b < r$. When $b \geq r$, the American call will theoretically never be exercised early, and the Garman-Kohlhagen model can be used. For this thesis, r denotes the US interest rate and as such $b \geq r$ only occurs when the foreign interest rate is equal to or below 0.

The approximate value of the American put option is derived with similar arguments, and the value is computed as:

$$\begin{aligned} P(S, T) &= p(S, T) + A_1(S/S^{**})^{q_1}, \text{ when } S > S^* \\ P(S, T) &= K - S, \qquad\qquad\qquad \text{when } S \leq S^* \end{aligned} \tag{56}$$

where $A_1 = -(S^{**}/q_1)[1 - e^{(b-r)T} N[-d_1(S^{**}]]$. Note that $A_1 > 0$ because $S^{**} > 0$, $q_1 < 0$ and $N[-d_1(S^{**})] < e^{-bT}$. The critical price S^{**} for the put is determined by iteratively solving:

$$K - S^{**} = p(S^{**}, T) - \frac{S^{**}}{q_1}\{1 - e^{(b-r)T} N[-d_1(S^{**})]\} \tag{57}$$

To test the implemented analytic approximation I compared it to the theoretical computations in Barone-Adesi and Whaley (1987) which is depicted in Table 3. The pricing error being zero for all the approximations follows from the deterministic nature of the updating scheme for obtaining the critical price (S^*, S^{**}) of the underlying[6].

Option Parameters	S	American Call Approximation	American Put Approximation
$r = 0.08$	80	0.05 (0.05)	20.00 (20.00)
$\sigma = 0.20$	90	0.85 (0.85)	10.18 (10.18)
$T = 0.25$	100	4.44 (4.44)	3.54 (3.54)
	110	11.66 (11.66)	0.80 (0.80)
	120	20.90 (20.90)	0.12 (0.12)
$r = 0.12$	80	0.05 (0.05)	20.00 (20.00)
$\sigma = 0.20$	90	0.84 (0.84)	10.16 (10.16)
$T = 0.25$	100	4.40 (4.40)	3.53 (3.53)
	110	11.55 (11.55)	0.79 (0.79)
	120	20.69 (20.69)	0.12 (0.12)
$r = 0.08$	80	1.29 (1.29)	20.53 (20.53)
$\sigma = 0.40$	90	3.82 (3.82)	12.93 (12.93)
$T = 0.25$	100	8.35 (8.35)	7.46 (7.46)
	110	14.80 (14.80)	3.96 (3.96)
	120	22.72 (22.72)	1.95 (1.95)
$r = 0.08$	80	0.41 (0.41)	20.00 (20.00)
$\sigma = 0.20$	90	2.18 (2.18)	10.71 (10.71)
$T = 0.50$	100	6.50 (6.50)	4.71 (4.71)
	110	13.42 (13.42)	1.76 (1.76)
	120	22.06 (22.06)	0.55 (0.55)

Table 3: American option estimation of theoretical options values using Barone-Adesi and Whaley (1987) analytic approximation, Barone-Adesi and Whaley (1987) values in parenthesis. b = 0.04, K = 100.

[6] For the updating scheme, see Appendix of Barone-Adesi and Whaley (1987).

4.3 Discretization, Monte Carlo Simulation

In order to test the stability of the pricing and calibration methodology used in this thesis, this section will present the tests that was performed prior to working with the market data. Bates model is discretized using an Euler-Maruyama approximation[7], simulated paths for the underlying and the stochastic volatility are computed where theoretical option prices are generated from the simulated paths. Bates model is then calibrated using the theoretical option prices using both of the methodologies described in Section 4.1.

The Euler-Maruyama discretization of Bates model used has the following form:

$$ln\ S_{t+1} = ln\ S_t + (\mu - \lambda\bar{k} - \frac{1}{2}V_t^+)\Delta t + \sqrt{V_t^+}\sqrt{\Delta t}Z + ln(1+k)dP_t$$

$$V_{t+1} = V_t + \kappa(\eta - V_t)\Delta t + \sigma_v\sqrt{V_t}\sqrt{\Delta t}W$$

with the same notational implications as in Chapter 3 for Bates model. This discretization makes it possible to sample paths from the model given parameter inputs.
In this discretization setting, the standard Brownian motions may give rise to negative values for V_t, due to drawing from a standard normal distribution. To deal with this problem, one of two approaches is typically applied. The first one called the *reflecting assumption* uses the absolute value:

$$\text{if } \ V_t < 0 \ \text{ then } \ V_t = -V_t$$

The other solution is using an *absorbing assumption*:

$$\text{if } \ V_t < 0 \ \text{ then } \ V_t = 0$$

which is the methodology applied here as per the V_t^+-notation[8] (Gatheral 2006).

For simulating paths of the underlying as well as the volatility, the same parameter values were used as in Table 1 in order to test the calibration methodology in a similar environment to the purpose of the thesis. Furthermore, the length of the simulated paths corresponds to the time span for the dataset that is used in this thesis, that is from the 4th of Jan 1984 to the 19th of June 1991, corresponding to 2724 days. As such, the idea of the simulated paths is to resemble daily data for the same duration. In Figure 2, a sample path for the two processes is shown:

[7] See Kloeden and Platen (1995) for reference on Euler-Maruyama approximation.

[8] $V_t^+ = \max(V_t, 0)$

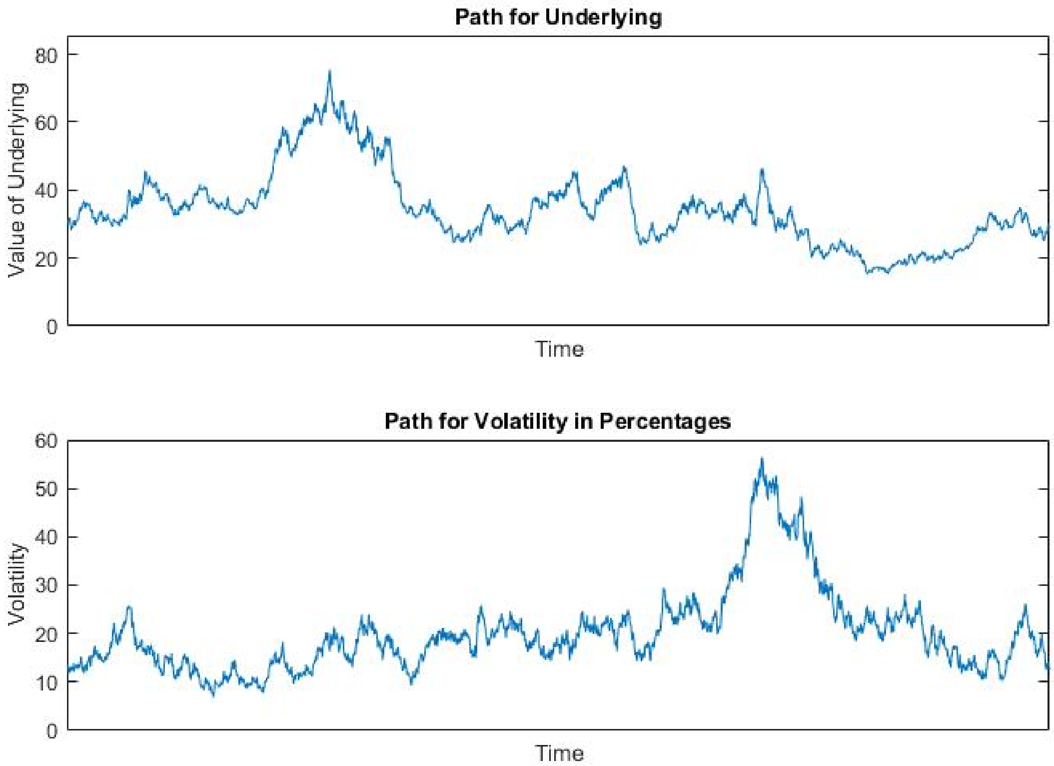

Figure 5: Simulated paths for underlying and volatility, $S_0 = 30$.

From the simulated path of the underlying, I generated theoretical option prices. This was done by constructing strike prices for each integer within ±5% of the price of the underlying for each point in the time-series corresponding to a Wednesday, as per the structure of the market data described in Chapter 5. This means that for an asset price of 50, the highest and lowest strike price is 53 and 47 (as $5\% * Price$ is rounded up), respectively. From there, a put and a call option is generated for each integer between 53 and 47, that is: 53,52,51,50,49,48,47. This resulted for the simulation presented here in 3102 options (1551 calls, 1551 puts). In Figure 6 the max and min theoretical strike prices for each "Wednesday" are plotted together with the simulated path of the underlying:

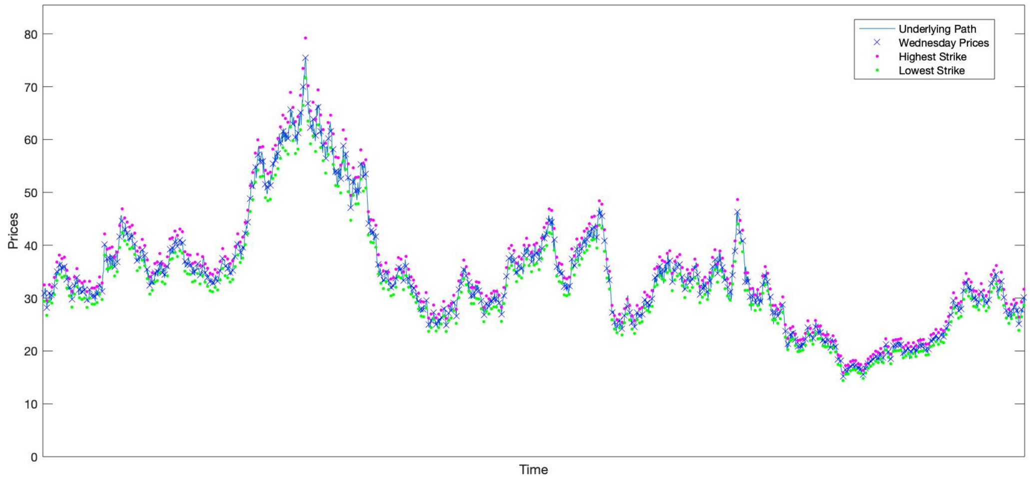

Figure 6: Simulated path for underlying with min and max of strike prices for each Wednesday.

In Figure 4, the theoretical option prices are plotted in histograms, separated into call and put options. As can be observed, the call and put options have a somewhat deviating distribution for the prices. The average price for the theoretical call and put options are 2.9078 and 2.4389, respectively.

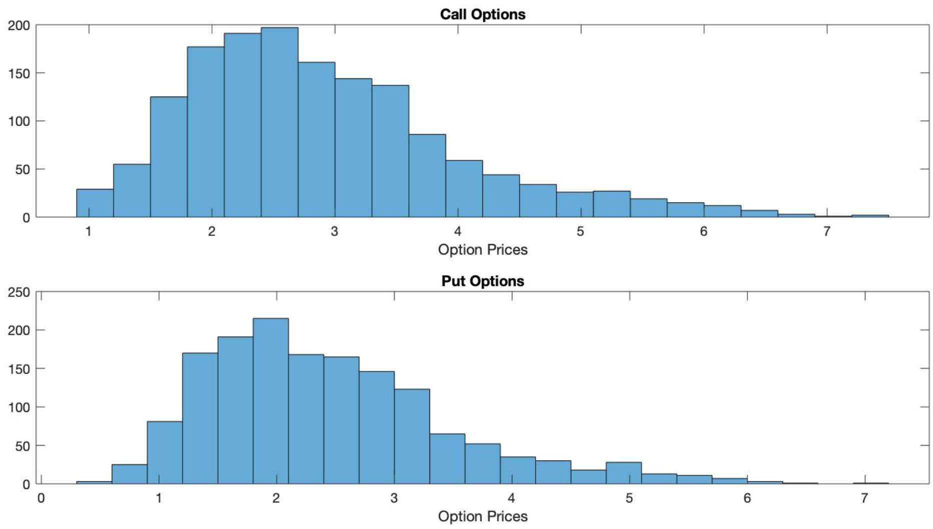

Figure 7: Histogram of theoretical option prices.

At this point, the calibration methodology explained in Chapter 4 is employed as to calibrate the parameter implicit in the theoretical options from the simulated path. The *correct answer* for the calibration would be to obtain the parameter values used for the simulation, that is, Table 1. For this calibration, I used arbitrary inputs for all the parameters as starting values in the calibration:

Variables	η^*	κ^*	σ_v	ρ	$\bar{k}^*$	δ	λ^*
x_0	0.09	1.2	0.3	-0.5	-0.04	0.1	2
lb	0.01	0.01	0.01	-1.00	-0.90	0.01	0.01
ub	10	10	10	1	10	10	20

Table 4: Starting values for the simulation calibration.

In Table 5 the resulting parameters calibrated from nonlinear least-squares and `fmincon` are presented together with the parameters used for the simulation. As can be observed, there is a quite wide margin to the parameter values that was used for the simulation. Most notably, looking at σ_v, the `fmincon` optimization volatility of variance-parameter has a value of 668.92%. This is one of the main issues raised by Bates (1994) and Bakshi et al. (1997) in that an option pricing model with a stochastic volatility-component might produce implicit parameter values for the volatility of variance-parameter that are too high. Furthermore, the correlation coefficient for `fmincon` suggests a very strong negative skew, which is contrasted by an expected percentage jump size of 12.55%.

Variables	η^*	κ^*	σ_v	ρ	$\bar{k}^*$	δ	λ^*
x_0	0.09	1.2	0.3	-0.5	-0.04	0.1	2
`lsqnonlin`	0.0288	2.4558	0.3094	-0.0984	0.0732	0.0640	1.9036
`fmincon`	0.2671	2.8938	6.6892	-0.7926	0.1255	0.0518	4.0024
Simulation Parameters	0.024	0.78	0.343	0.078	-0.001	0.019	15.01

Table 5: Calibrated parameter values from the theoretical options.

For the overall pricing performance of the two calibration methodologies, this is provided in Table 6. AAPPE denotes Absolute Average Percentage Pricing Error[9], AAPE denotes Absolute Average Pricing Error[10], APPE denotes Average Percentage Pricing Error[11], and APE denotes Average Pricing Error[12]:

[9] $AAPPE = \frac{1}{N}\sum_{i=1}^{N}\left|\frac{O_i^{(M)} - O_i^{(E)}}{O_i^{(M)}}\right|$

[10] $AAPE = \frac{1}{N}\sum_{i=1}^{N}|O_i^{(M)} - O_i^{(E)}|$

[11] $APPE = \frac{1}{N}\sum_{i=1}^{N}\left(\frac{O_i^{(M)} - O_i^{(E)}}{O_i^{(M)}}\right)$

[12] $APE = \frac{1}{N}\sum_{i=1}^{N}\left(O_i^{(M)} - O_i^{(E)}\right)$

	lsqnonlin	fmincon
AAPPE	$1.9809 \cdot 10^{-4}\%$	$3.3005 \cdot 10^{-2}\%$
AAPE	$4.8374 \cdot 10^{-6}$	$7.1668 \cdot 10^{-4}$
APPE	$4.0031 \cdot 10^{-6}\%$	$-1.8602 \cdot 10^{-2}\%$
APE	$1.5892 \cdot 10^{-7}$	$-3.5361 \cdot 10^{-4}$

Table 6: Pricing errors for theoretical options.

The pricing residuals from the calibration methodologies are small looking from an average perspective with well below 1% for both `lsqnonlin` and `fmincon` considering both regular and absolute terms. Looking at all the pricing measurements used in this section, the `lsqnonlin` methodology performs better than the `fmincon` specification. The maximum pricing error is $1.5261 \cdot 10^{-5}$ for the `lsqnonlin` specification and 0.0013 for `fmincon`. In Figure 5, scatterplots of the theoretical call option prices against the estimated prices from `lsqnonlin` and `fmincon` are shown. As is visible, the prices are indistinguishable from each other. See Appendix for corresponding graph for the puts.

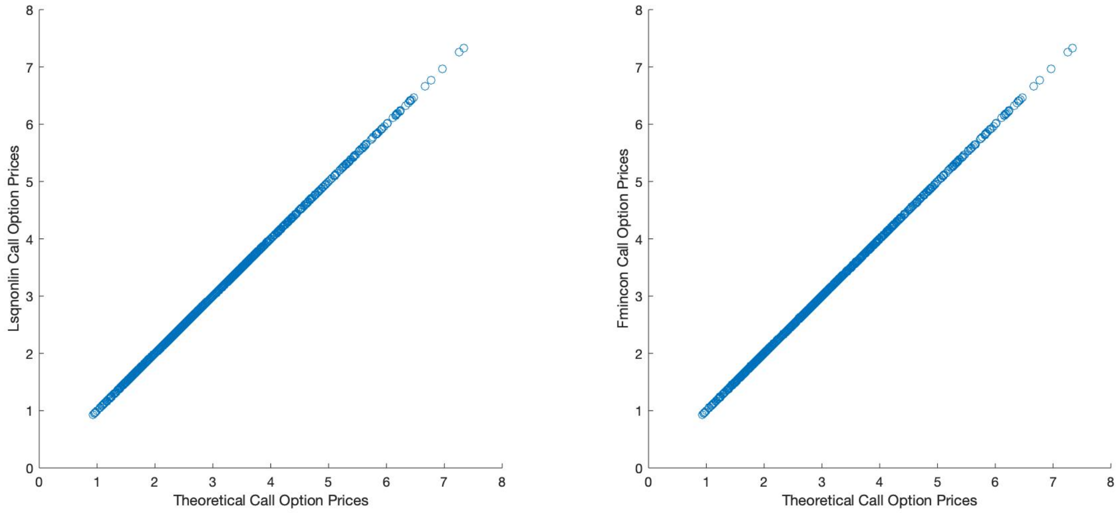

Figure 8: Scatterplot of theoretical call options prices against `lsqnonlin` and `fmincon` estimated prices.

It is important to note the limitations of the testing which influence the performance of the calibration methodology. Firstly, the same time to maturity and interest rate differential is used for all the options. The time to maturity is set to 3 months and the interest rate differential used is the average of the dataset presented in Chapter 5. Moreover, the moneyness does not exceed 5% as per the design of the theoretical options. As such, a restricted theoretical setting such as this naturally simplifies the model calibration to obtain parameter values that well fit the theoretical options.

4.4 Limitations

One of the numerical issues with respect to the calibration methodology of this thesis is the sole use of local optimization techniques. This is because there are usually numerous local optima in a multivariate scenario. The motivation behind only using local optimization was partly due to the extensive calibration time for the main calibrations as well as the fact that the purpose of the thesis is to evaluate with respect to Bates (1996). However, not using a global optimization technique does not ensure obtaining a global optima. As noted in Mikhailov and Nögel (2004), when only using a local optimization technique, choosing the initial guess for the parameter vector is important. Thus, for the purposes of testing with respect to Bates (1996) it was deemed reasonable to solely use local optimization techniques with the calibrated values in Bates (1996) as starting point.

The analytic approximation of the early exercise premium that is used in Bates (1996) is the approximation developed in Bates (1991) which is an extension of the methodology by Barone-Adesi & Whaley used in this thesis. The extension introduced in Bates (1991) is for Merton's jump-diffusion model and utilizes the same principles with deriving the early exercise premium from the PDE and obtaining the critical value of the underlying above or below which the option is exercised immediately. I implemented the analytic approximation but due to computational issues, the methodology was inconsistent when applied to real data. It should be noted that this naturally lessens the comparison of the results in this thesis to that of Bates (1996) but not to a substantial extent. From a theoretical standpoint, both methodologies are proposed from the same principles. Furthermore, as noted in Barone-Adesi (2005), the analytic approximation that is used in this thesis works very well for maturities less than one year. This is no problem as all the option transactions in the dataset have a maturity of less than 6 months.

5 Market Data

The same market data that was used in Bates (1996) is used in this thesis. The dataset consists of transaction data for Deutsche Mark (DM) foreign currency options from the Philadelphia Stock Exchange (PHLX) for the period January 1984 to June 1991. The option transactions are all on Wednesdays, yielding a weekly panel data set. The transactions were recorded in the morning, between 9AM-12AM, and only includes options with a maturity of 6 months or less. The options matured on the third Wednesday of March, June, September or December as per the standard of PHLX during that period.

The dataset consists of 19,689 option transactions (11,952 calls, 7,737 puts) collected over 372 Wednesdays with an average of 53 trades per Wednesday. Other than option transactions, the dataset includes the spot exchange rate obtained from the Chicago Mercantile Exchange (CME), a risk-free rate which was obtained from daily 3-month Treasury bills, domestic/foreign interest rate differential which was obtained from interpolating spot rates of 1- and 3-month forward rates of the spot exchange rate using covered interest rate parity[13].

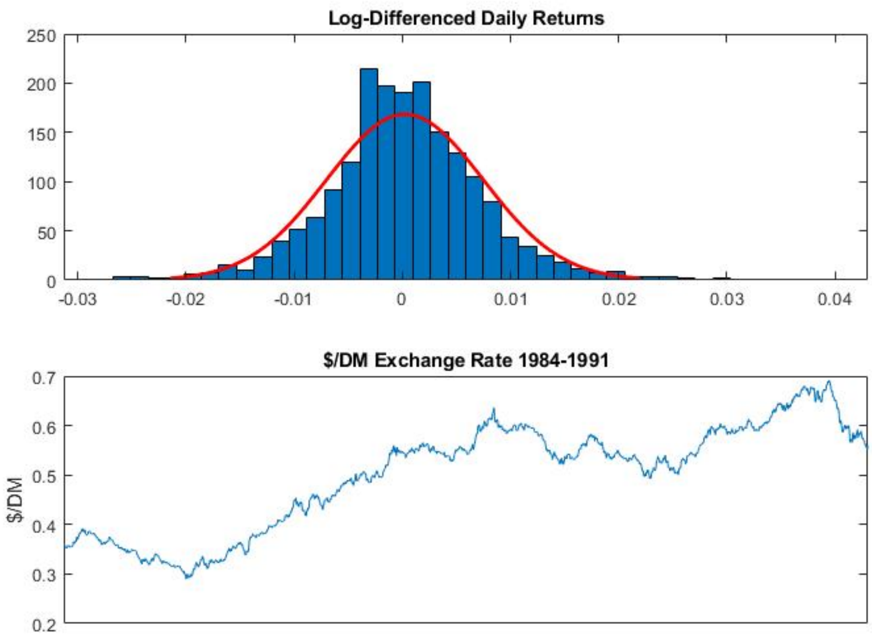

Figure 9: Log-differenced daily returns of USD/DM and USD/DM exchange rate for Jan 1984-Jun 1991.

As was discussed in Chapter 1, one of the typical properties for exchange rate returns is a leptokurtic distribution. This can be observed in Figure 9 where log-differenced daily returns for the exchange

[13]For a full treatment of the filtration of the dataset, see Bates (1996).

rate underlying the options is shown for the period 1984-1991[14], depicting a more peaked center of the distribution with fatter tails than the normal distribution. A two-sample Kolmogorov-Smirnov test further supports this notion by rejecting the null hypothesis that the log-differenced returns are normally distributed.

Looking at the first histogram in Figure 10 showing the price of the options, 67.65% of the options have a price less than \$1, 91.41% of the options have a price less than \$2, and the largest option price is \$11.2. The second histogram in Figure 10 depicts the implied volatility calculated from the Garman-Kohlhagen model for all the option transactions. Here, as with the option prices, it can be observed that most of the implied volatilities are low, 95.78% of implied volatilities are less than 20%. This can be contrasted to the overall volatility of the USD/DM exchange rate in Figure 6 which is 10.58%.

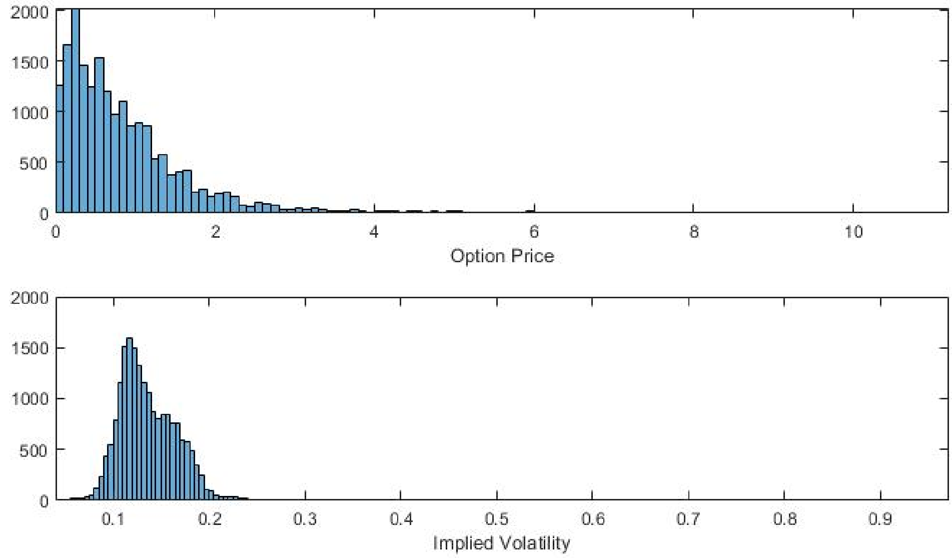

Figure 10: Histogram of option prices and implied volatilities.

[14]Retrieved from FRED, Federal Reserve Bank of St. Louis, https://fred.stlouisfed.org/series/EXGEUS, May 20, 2020.

6 Results

6.1 Model Calibration

In Table 7 the calibrated parameters from the market data are depicted for `lsqnonlin` and `fmincon`.

	η^*	κ^*	σ_v	ρ	$\bar{k}^*$	δ	λ^*
Bates (1996)	0.024	0.78	0.343	0.078	-0.001	0.019	15.01
`lsqnonlin`	0.010	1.600	0.364	-0.100	0.010	0.061	2.635
Error	0.014	-0.82	-0.021	0.178	-0.011	-0.042	12.375
% Error	58.30%	-105.20%	-6.10%	228.80%	1129.6%	-218.90%	82.40%
`fmincon`	0.012	1.357	9.946	-0.998	-0.126	0.012	0.691
Error	0.012	-0.577	-9.603	1.076	0.125	0.007	14.319
% Error	50.31%	-73.99%	2799.67%	1379.38%	12540.55%	37.56%	95.40%

Table 7: Calibrated parameters with error and percentage error compared to Bates (1996). Errors calculated as Bates (1996) less the calibrated parameter.

Firstly, looking at the `lsqnonlin` parameter values overall, they correspond quite closely to the calibrated parameters from Bates (1996). η^*, the long-term variance level, suggests a long-term variance level of 0.010, corresponding to a long-term volatility level of 10%. κ^*, the mean-reversion speed for the variance, is greater than Bates (1996) value, suggesting a faster smoothing of deviations from the long-term mean. σ_v, the volatility of variance parameter also closely follows the initial parameter inputs with a value of 36.40%. ρ on the other hand suggests a negative correlation as opposed to Bates (1996). The negative correlation being modest, only suggests a slight impact on the skewness. For the jump-diffusion parameters, the relationship mirrors that of Bates (1996) with respect to low expected jump size (albeit a positive one), a low volatility of the random jump size, but λ^* depicts a lower expected annual frequency of jumps than that of Bates (1996).

For `fmincon`, the calibrated parameters deviate to a larger extent from the calibrated parameters in Bates (1996) compared to `lsqnonlin`. η^*, the long-term variance level, and κ^*, the mean-reversion speed for the variance, corresponds closely to the values in `lsqnonlin` by a long-term volatility of 10.95% and a mean-reversion speed which faster smoothes out deviations from the long-term mean compared to Bates (1996). σ_v, the volatility of variance parameter, however has a value of 9.946, i.e. a volatility of variance value of 994.6%. This is very close to the upper bound of 1000%. This follows the discussion in Section 3.3 regarding previous studies suggesting that the volatility of variance parameter in stochastic volatility models might imply values that are very high compared to the analytical specification (Bates 2003). ρ has a value of -0.998, almost a perfect negative correlation, suggesting a very strong negative skewness as opposed to Bates (1996) where a slight positive skew is implied. The skewness is further impacted by $\bar{k}^*$, the expected random percentage jump, with a value of -0.126, meaning that the expected percentage jump size conditional on a jump occurring is -12.6%. Looking at the jump-diffusion component altogether: $\bar{k}^*$ together with δ and λ^* depicts an opposite relationship to that of Bates (1996). The jump-frequency in `fmincon` suggests 0.691 expected annual jumps with a comparatively larger expected jump size contrasted to the parameter values in Bates (1996) where a higher frequency of jumps with smaller expected jump size is implied.

Which of the methodologies' resulting implicit parameter values that are deviating the most from Bates (1996) is not surprising. This is because the `lsqnonlin` specification much more closely mirrors the NL-GLS-methodology in Bates (1996) and as such the initial parameter values for the calibration, that is Table 1, should be a better fit for the `lsqnonlin` methodology. This was also evident with respect to calibration times where the `fmincon` calibration exceeded `lsqnonlin`. For `fmincon`, the behaviour of σ_v, ρ are improbable as to the specification of the model given that the parameter values are close to the boundaries set. This behaviour was however also documented in Section 4.3 wherein large parameter values were documented for σ_v and ρ.

6.2 Pricing Performance

For the pricing performance, I recreated the pricing structure of Bates (1996) wherein the pricing performance is filtered according to moneyness and time to maturity. Note that the column showing the number of observations depicts a deviation between this thesis and those of Bates (1996), which unfortunately has not been resolved. The difference could at least partially be attributed to spill over between the closely defined moneyness categories.

In Table 7 and Table 8, the residuals from the `lsqnonlin` calibrated parameter call and put values are shown, respectively.

Moneyness (K/S - 1)	Number of Observations	Average Errors	Standard Deviations
Short-Term (0-3 month) calls			
<-6%	445 (646)	0.053% (0.027%)	0.231% (0.048%)
[-6%,-4%]	697 (1140)	-0.015% (0.016%)	0.322% (0.055%)
[-4%,-2%]	1715 (1051)	0.041% (0.011%)	0.363% (0.070%)
[-2%,-1%]	1179 (1151)	0.012% (0.011%)	0.307% (0.076%)
[-1%,0%]	1331 (1360)	-0.039% (0.011%)	0.384% (0.085%)
[0%,1%]	1049 (1053)	0.141% (0.006%)	0.428% (0.100%)
[1%,2%]	495 (499)	0.038% (-0.030%)	0.340% (0.104%)
[2%,4%]	691 (334)	0.152% (-0.003%)	0.382% (0.113%)
[4%,6%]	313 (576)	-0.029% (-0.039%)	0.244% (0.116%)
>6%	422 (534)	-0.045% (-0.070%)	0.185% (0.115%)
Medium-Term (3-6 month) calls			
<-6%	342 (492)	0.036% (-0.027%)	0.429% (0.076%)
[-6%,-4%]	358 (434)	-0.039% (-0.009%)	0.540% (0.088%)
[-4%,-2%]	599 (333)	0.028% (-0.009%)	0.552% (0.083%)
[-2%,-1%]	414 (434)	0.104% (0.013%)	0.548% (0.086%)
[-1%,0%]	479 (476)	-0.113% (-0.001%)	0.549% (0.094%)
[0%,1%]	419 (414)	0.010% (-0.009%)	0.580% (0.094%)
[1%,2%]	311 (322)	0.108% (-0.015%)	0.587% (0.116%)
[2%,4%]	293 (190)	0.056% (-0.046%)	0.504% (0.125%)
[4%,6%]	190 (237)	-0.121% (-0.062%)	0.498% (0.125%)
>6%	210 (276)	-0.130% (-0.125%)	0.339% (0.160%)

Table 8: Residuals from `lsqnonlin` calibration, call options. Bates values in parenthesis. Average Errors calculated as in Equation (29).

Moneyness (K/S - 1)	Number of Observations	Average Errors	Standard Deviations
Short-Term (0-3 month) puts			
<-6%	56 (93)	-0.082% (-0.002%)	0.206% (0.135%)
[-6%,-4%]	96 (132)	-0.053% (-0.001%)	0.222% (0.091%)
[-4%,-2%]	218 (151)	-0.022% (0.023%)	0.327% (0.110%)
[-2%,-1%]	309 (283)	-0.026% (0.003%)	0.259% (0.090%)
[-1%,0%]	566 (585)	-0.093% (0.002%)	0.328% (0.097%)
[0%,1%]	933 (920)	-0.007% (0.034%)	0.333% (0.093%)
[1%,2%]	905 (911)	-0.032% (0.027%)	0.303% (0.080%)
[2%,4%]	1280 (788)	-0.001% (0.022%)	0.292% (0.072%)
[4%,6%]	630 (939)	-0.011% (0.024%)	0.220% (0.064%)
>6%	610 (828)	0.078% (0.033%)	0.193% (0.047%)
Medium-Term (3-6 month) puts			
<-6%	24 (38)	-0.149% (-0.095%)	0.464% (0.124%)
[-6%,-4%]	25 (29)	-0.277% (-0.053%)	0.331% (0.100%)
[-4%,-2%]	63 (41)	-0.268% (-0.097%)	0.496% (0.115%)
[-2%,-1%]	78 (76)	-0.003% (-0.019%)	0.582% (0.132%)
[-1%,0%]	184 (175)	-0.235% (-0.021%)	0.479% (0.120%)
[0%,1%]	209 (208)	-0.169% (-0.009%)	0.523% (0.096%)
[1%,2%]	276 (271)	0.082% (0.001%)	0.647% (0.095%)
[2%,4%]	419 (197)	-0.129% (0.000%)	0.447% (0.087%)
[4%,6%]	316 (410)	-0.063% (-0.011%)	0.453% (0.066%)
>6%	540 (663)	-0.063% (-0.003%)	0.274% (0.050%)

Table 9: Residuals from `lsqnonlin` calibration, put options. Bates values in parenthesis. Average Errors calculated as in Equation (29).

Looking across moneyness, the pricing performance of `lsqnonlin` both under- and overprices the short- and medium-term call options, thus depicting no apparent pricing bias. The average errors for the calls never exceed 0.2%. For the put options, the `lsqnonlin` specification consistently overprices for both short-term and medium-term options, with the exception of short-term options with moneyness of $[> 6\%]$ and medium-term options with moneyness of $[1\%, 2\%]$. The highest average errors across call and put options can be observed for the medium-term puts where average error exceeds 0.2% for three out of five negative moneyness categories, i.e. for out-of-the money puts. The `lsqnonlin` fits the 0-3 month call and put options better than the 3-6 month options. This can visually be observed in Figure 15 in the Appendix where the average errors for `lsqnonlin` are plotted against moneyness.

The pricing performance of the `lsqnonlin` calibration fare quite well to the errors from Bates (1996). The average errors for `lsqnonlin` are smaller than Bates (1996) for 27.5% of the pricing categories, although sporadically throughout the tables so no pattern of overperformance can be observed. The magnitude of the standard deviations are however quite large with respect to Bates (1996), depicting a wider spread for the pricing errors.

In Table 10 and Table 11, the corresponding pricing performance for `fmincon` is depicted.

Moneyness (K/S - 1)	Number of Observations	Average Errors	Standard Deviations
Short-Term (0-3 month) calls			
<-6%	445 (646)	0.398% (0.027%)	0.329% (0.048%)
[-6%,-4%]	697 (1140)	0.549% (0.016%)	0.404% (0.055%)
[-4%,-2%]	1715 (1051)	0.853% (0.011%)	0.590% (0.070%)
[-2%,-1%]	1179 (1151)	1.004% (0.011%)	0.497% (0.076%)
[-1%,0%]	1331 (1360)	0.748% (0.011%)	0.411% (0.085%)
[0%,1%]	1049 (1053)	0.558% (0.006%)	0.464% (0.100%)
[1%,2%]	495 (499)	0.159% (-0.030%)	0.373% (0.104%)
[2%,4%]	691 (334)	0.042% (-0.003%)	0.377% (0.113%)
[4%,6%]	313 (576)	-0.275% (-0.039%)	0.250% (0.116%)
>6%	422 (534)	-0.225% (-0.070%)	0.201% (0.115%)
Medium-Term (3-6 month) calls			
<-6%	342 (492)	0.805% (-0.027%)	0.504% (0.076%)
[-6%,-4%]	358 (434)	1.319% (-0.009%)	0.588% (0.088%)
[-4%,-2%]	599 (333)	1.569% (-0.009%)	0.528% (0.083%)
[-2%,-1%]	414 (434)	1.221% (0.013%)	0.587% (0.086%)
[-1%,0%]	479 (476)	0.635% (-0.001%)	0.573% (0.094%)
[0%,1%]	419 (414)	0.442% (-0.009%)	0.578% (0.094%)
[1%,2%]	311 (322)	0.294% (-0.015%)	0.605% (0.116%)
[2%,4%]	293 (190)	0.018% (-0.046%)	0.512% (0.125%)
[4%,6%]	190 (237)	-0.332% (-0.062%)	0.520% (0.125%)
>6%	210 (276)	-0.261% (-0.125%)	0.357% (0.160%)

Table 10: Residuals from `fmincon` calibration, call options. Bates values in parenthesis. Average Errors calculated as in Equation (9).

Moneyness (K/S - 1)	Number of Observations	Average Errors	Standard Deviations
Short-Term (0-3 month) puts			
<-6%	56 (93)	0.188% (-0.002%)	0.260% (0.135%)
[-6%,-4%]	96 (132)	0.339% (-0.001%)	0.409% (0.091%)
[-4%,-2%]	218 (151)	0.626% (0.023%)	0.609% (0.110%)
[-2%,-1%]	309 (283)	0.878% (0.003%)	0.499% (0.090%)
[-1%,0%]	566 (585)	0.573% (0.002%)	0.364% (0.097%)
[0%,1%]	933 (920)	0.353% (0.034%)	0.353% (0.093%)
[1%,2%]	905 (911)	0.074% (0.027%)	0.296% (0.080%)
[2%,4%]	1280 (788)	-0.100% (0.022%)	0.287% (0.072%)
[4%,6%]	630 (939)	-0.271% (0.024%)	0.226% (0.064%)
>6%	610 (828)	-0.183% (0.033%)	0.202% (0.047%)
Medium-Term (3-6 month) puts			
<-6%	24 (38)	0.554% (-0.095%)	0.508% (0.124%)
[-6%,-4%]	25 (29)	0.998% (-0.053%)	0.253% (0.100%)
[-4%,-2%]	63 (41)	1.365% (-0.097%)	0.480% (0.115%)
[-2%,-1%]	78 (76)	1.077% (-0.019%)	0.644% (0.132%)
[-1%,0%]	184 (175)	0.529% (-0.021%)	0.515% (0.120%)
[0%,1%]	209 (208)	0.260% (-0.009%)	0.508% (0.096%)
[1%,2%]	276 (271)	0.265% (0.001%)	0.669% (0.095%)
[2%,4%]	419 (197)	-0.191% (0.000%)	0.455% (0.087%)
[4%,6%]	316 (410)	-0.256% (-0.011%)	0.453% (0.066%)
>6%	540 (663)	-0.206% (-0.003%)	0.292% (0.050%)

Table 11: Residuals from `fmincon` calibration, put options. Bates values in parenthesis. Average Errors calculated as in Equation (9).

For `fmincon`, the call option errors depicts an underpricing for the majority of moneyness categories across both short- and medium-term. The exceptions are the out-of-the-money calls when moneyness is greater than 4% where the call options are on average overvalued with respect to the market price. For put options, the same pattern emerges wherein put options with negative moneyness, i.e. the out-of-the-money puts, are consistenly undervalued. Both short- and medium-term put options with a moneyness greater than 2%, i.e. the in-the-money puts, are overvalued.
As is evident, the performance of the `fmincon` calibration is inferior to that of Bates (1996) for all moneyness categories except the 3-6 month calls with a moneyness of $[2\%, 4\%]$. The average errors for the calls rarely exceed 1%, except for the in-the-money short-term calls with a moneyness of $[-2\%, -1\%]$ and in-the-money medium-term calls with a moneyness between -6% and -1%. For the puts, nearly the same pattern emerges with exceeding 1% average pricing error for the 3-6 month puts with moneyness $[-4\%, -2\%]$ and $[-2\%, -1\%]$. Compared to the `lsqnonlin` calibration, the `fmincon` calibration has a lower average pricing error for 2 of the pricing categories, thus performing overall worse than the `lsqnonlin` calibration. As was the case for `lsqnonlin`, the magnitude of the standard deviations for `fmincon` are also larger with respect to Bates (1996), thus suggesting wider

spread among pricing errors.

The performance of the Garman-Kohlhagen model is found in Table 12 and 13 in the Appendix. It is important to note that given that the analytic approximation is derived in the structure of the Garman-Kohlhagen model, it should perform well in this setting as using the same volatility as was done for approximating the early exercise premium would yield a pricing error equal to zero for all options. To avoid this problem, the implicit volatility was used to estimate the early exercise premium and the overall volatility for the exchange rate in Figure 9 was used in the pricing performance calculations.
The results over moneyness and maturity for the Garman-Kohlhagen model is quite consistent, valuing the options on average below the market price across all categories. The `lsqnonlin` specification performs comparatively better across all categories than the Garman-Kohlhagen model. The same is not observed for the `fmincon` specification where the Garman-Kohlhagen model has a lower average error for 70% of the moneyness-categories looking across both short- and medium-term maturities. To compare the models at an aggregate level I calculated a modified version of Theil's U-statistic for the different specifications:

$$U = \frac{\sqrt{\sum_{i=1}^{N}\left(\frac{O_i^{(M)}-O_i^{(B)}}{O_i^{(M)}}\right)^2}}{\sqrt{\sum_{i=1}^{N}\left(\frac{O_i^{(M)}-O_i^{(GK)}}{O_i^{(M)}}\right)^2}} \tag{58}$$

where as in Equation (29), $O_i^{(M)}$ denotes the market option price, $O_i^{(B)}$ denotes the estimated option price from Bates model, and $O_i^{(GK)}$ denotes the estimated option price from the Garman-Kohlhagen model. Note the different specifications in the denominator as compared to the pricing residual in Equation (29) and consequently the tables provided in this chapter. For interpretation of the U-statistic a value of $U = 1$ implies that the models under consideration are equal performance-wise, a value of $U < 1$ implies that the model in the numerator is superior, and a value of $U > 1$ implies that the model in the denominator is superior (Brooks 2019).
For the different parameter specifications, the U-statistic for the `lsqnonlin` is 0.0015 and for `fmincon` 1.5017 using the Garman-Kohlhagen model as the benchmark model in the denominator. As such, the parameter specification in `lsqnonlin` is comparatively better than the `fmincon` specifications relative to the Garman-Kohlhagen model. This also means that the `fmincon` performs worse than the Garman-Kohlhagen overall in this framework.

To visualize the performances over moneyness, Figure 11 depicts the average pricing errors for puts over moneyness separated into 0-3 months and 3-6 months for `lsqnonlin`, `fmincon`, Bates (1996), and the Garman-Kohlhagen model. The corresponding figure for call options can be found in Figure 14 in the Appendix.

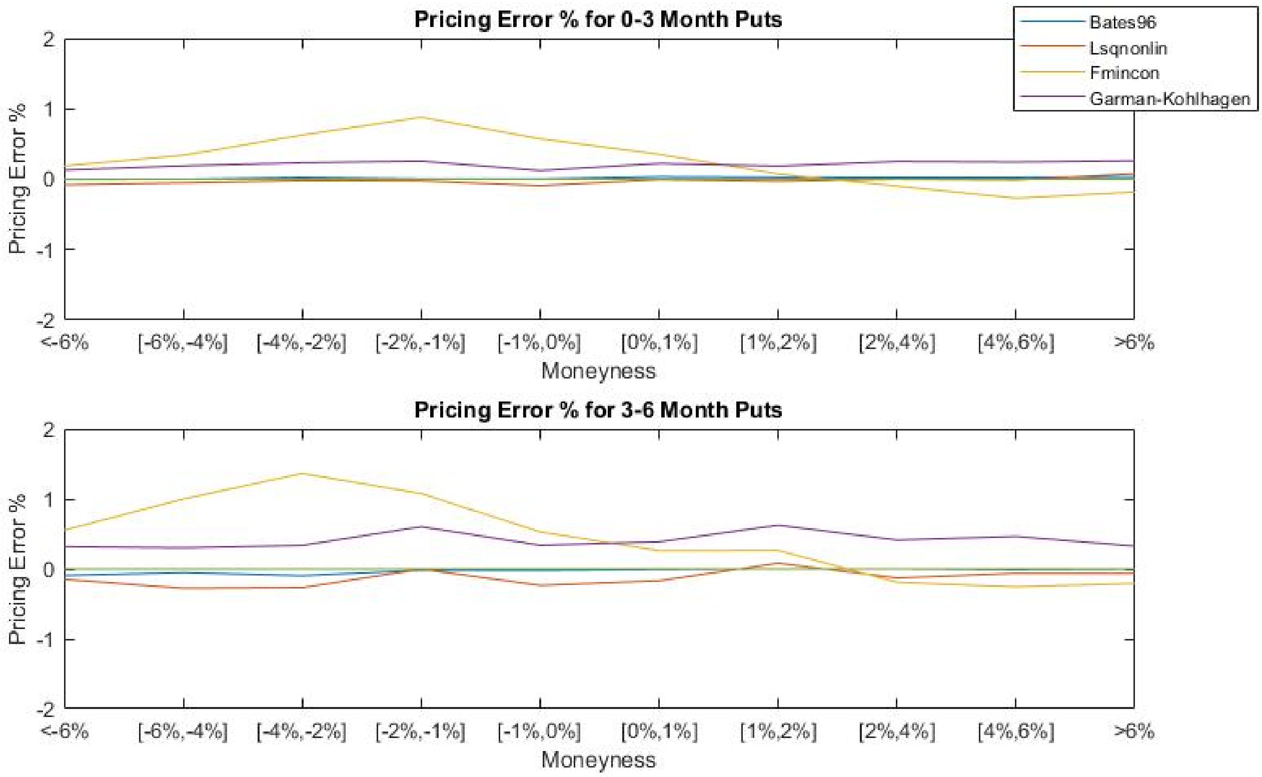

Figure 11: Pricing error over moneyness and maturity.

As can be observed, Bates (1996), `lsqnonlin`, and Garman-Kohlhagen show a less volatile pricing error across moneyness than the `fmincon` specification. `fmincon` has the largest average errors for moneyness less than 0, i.e. out-of-the-money puts by on average estimating below the market price. This pattern can also be observed in Figure 14, but then for in-the-money calls where `fmincon` also undervalues the market prices. Given the parameter values, there is a mixed relationship between σ_v contributing heavily to the excess kurtosis and ρ to the skewness which could explain the inconsistent theoretical interpretation of the pricing performance. For the pricing errors of the options with a positive moneyness, the `fmincon` specifications however performs better. By limiting the pricing to options with a positive moneyness, the U-statistic for `fmincon` is 0.0981, thus performing better than the Garman-Kohlhagen model overall for that subsection for both calls and puts.

7 Conclusion

The purpose of this thesis was to complement the results from Bates (1996) by performing calibration over the same dataset using two different methodologies. This was successfully implemented using the nonlinear least squares methodogy and the built-in MATLAB function `fmincon`.

The calibration using `lsqnonlin` performs well over the pricing framework, surpassing both the Garman-Kohlhagen model and the `fmincon` specification. The calibrated parameter values correspond closely to those in Bates (1996) and given the market data exhibit reasonable values to the analytical specification of the model.
For `fmincon`, the calibrated parameters are deviating widely, providing extreme parameters values for certain variabels that are improbable given the analytical specification of Bates model. This is an issue discussed in this thesis, that of having a stochastic volatility-model that is analytically pleasing but might cause empirical misspecification with respect to the theoretical specification of the model (Mills and Patterson 2009). This behaviour can be observed for the `fmincon` calibration, most notably for the volatility of variance- and correlation-parameter. However, since the pricing performance of `fmincon` exhibit inconsistent behaviour across moneyness as well as performing overall worse than the Garman-Kohlhagen model, this is not an issue. Furthermore, for both the calibration results in Section 4.3 and Section 6.1, the `fmincon` methodology exhibits extreme values for both σ_v and ρ. The results from Section 6.1 could be attributed to the initial starting values for the parameters which should favour the `lsqnonlin` methodology as it more closely follows that of Bates (1996). However, given that the calibrated parameter values in Section 4.3 exhibit similar behaviour with respect to σ_v and ρ, wherein arbitrary starting points where used as starting values, it suggests that the `fmincon` calibration methodology with MAE tends towards extreme values for σ_v and ρ.

For future research, it would be interesting to further carry out this methodology on other option pricing models, replicating and extending with different calibration methodologies. Furthermore, from the discussion of local optimization techniques, it would be useful to test several starting points for the calibrations, or global optimization techniques. This was not done in this thesis due to the already extensive calibration times. Lastly, as noted in the end of Chapter 4, it would be interesting to test the impact of using different analytic approximation techniques.

References

Akgiray, V., Booth, G. G., and Seifert, B. (1988). Distribution properties of latin american black market exchange rates. *Journal of International Money and Finance*, pages 37–48.

Amin, K. and Ng, V. (1993). Options valuation with systematic stochastic volatility. *Journal of Finance*, pages 881–910.

Bakshi, G., Cao, C., and Chen, Z. (1997). Empirical performance of alternative option pricing models. *The Journal of Finance*, pages 2003–2049.

Barone-Adesi, G. (2005). The saga of the american put. *Journal of Banking Finance*, pages 2909–2918.

Barone-Adesi, G. and Whaley, R. E. (1987). Efficient analytic approximation of american option values. *The Journal of Finance*, pages 301–320.

Bates, D. S. (1991). The crash of 87': Was it expected? the evidence from options markets. *The Journal of Finance*, pages 1009–1044.

Bates, D. S. (1994). Post-'87 crash fears in the sp 500 futures option market. *Journal of Econometrics*, pages 181–238.

Bates, D. S. (1996). Jumps and stochastic volatility: Exchange rate processes implicit in deutsche mark options. *The Review of Financial Studies*, pages 69–107.

Bates, D. S. (2003). Empirical option pricing: a retrospection. *Journal of Econometrics*, pages 387–404.

Beneder, R. and Elkenbracht-Huizing, M. (2003). Foreign exchange options and the volatility smile. *Medium Econometrische Toepassingen.*

Björk, T. (2009). *Arbitrage Theory in Continuous Time.* Oxford University Press, Great Clarendon Street, Oxford OX2 6DP, third edition edition.

Black, F. and Scholes, M. (1973). The pricing of options and corporate liabilities. *The Journal of Political Economy*, pages 637–654.

Bollerslev, T., Chou, R. Y., and Kroner, K. F. (1992). Arch modelling in finance. *Journal of Econometrics*, pages 5–59.

Boothe, P. and Glassman, D. (1987). The statistical distribution of exchange rates: Empirical evidence and economic implications. *Journal of International Economics*, pages 297–319.

Brooks, C. (2019). *Introductory Econometrics for Finance*. Cambridge University Press, One Liberty Plaza, 20th Floor, New York, NY 10006, USA, fourth edition edition.

Calin, O. (2015). *An Informal Introduction to Stochastic Calculus with Applications*. World Scientific Publishing Co. Pte. Ltd., first edition edition.

Cheung, C. S. and Miu, P. (2009). Currency instability: Regime switching versus volatility clustering. *Quarterly Journal of Finance and Accounting*.

Cox, J. (1975). Notes on option pricing: Constant elasticity of variance diffusions.

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

de Vries, C. G. and Leuven, K. U. (1992). Stylized facts of nominal exchange rate returns. *Handbook of International Macroeconomics*, pages Oxford and Cambridge, US: Blackwell.

Dupire, B. (1994). Pricing with a smile.

Friedman, D. and Vandersteel, S. (1982). Short-run fluctuations in foreign exchange rates. *Journal of International Economics*, pages 171–186.

Gander, W. and Gautschi, W. (2000). Adaptive quadrature - revisited. *BIT Numerical Mathematics*.

Garman, M. and Kohlhagen, S. (1983). Foreign currency option values. *Journal of International Money and Finance*, pages 231–237.

Gatheral, J. (2006). *The Volatility Surface*. John Wiley Sons, Inc., Hoboken, New Jersey.

Hagan, P. S., Kumar, D., Lesniewski, A. S., and Woodward, D. E. (2002). Managing smile risk.

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

Hull, J. and White, A. (1990). Pricing interest-rate derivative securities. *The Review of Financial Studies*, pages 573–592.

Jex, M., Henderson, R., and Wang, D. (1999). Pricing exotics under the smile. *Risk*.

Klebaner, F. C. (2005). *Introduction to Stochastic Calculus with Applications*. Imperial College Press, 57 Shelton Street, Covent Garden, London WC2H 9HE, second edition edition.

Kloeden, P. E. and Platen, E. (1995). *Numerical Solution of Stochastic Differential Equations*. Springer, second edition edition.

Lipton, A. and McGhee, W. (2002). Universal barriers. *Risk*.

MATLAB (2020a). Constrained nonlinear optimization algorithms. MathWorks Help Center.

MATLAB (2020b). Least-squares (model fitting) algorithms. MathWorks Help Center.

Melino, A. and Turnbull, S. M. (1990). Pricing foreign currency options with stochastic volatility. *Journal of Econometrics*, pages 239–265.

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

Mikhailov, S. and Nögel, U. (2004). Heston's stochastic volatility model implementation, calibration and some extensions. pages 401–407.

Mills, T. C. and Patterson, K. (2009). *Handbook of Econometrics, Applied Econometrics*. Palgrave Macmillan, Saffron House, 6–10 Kirby Street, London EC1N 8TS, second edition edition.

Schmelzle, M. (2010). Option pricing formulae using fourier transform: Theory and application.

Shastri, K. and Wethyavivorn, K. (1987). The valuation of currency options for alternate stochastic processes. *Journal of Financial Research*.

Shreve, S. E. (2004). *Stochastic Calculus for Finance II : Continuous-Time Models*. Springer, New York ; London.

Stein, E. M. and Stein, J. C. (1991). Stock price distributions with stochastic volatility: An analytic approach. *Review of Financial Studies*, pages 727–752.

Tankov, P. and Voltchkova, E. (2009). Jump-diffusion models: a practicioner's guide.

Taylor, M. (1995). The economics of exchange rates. *Journal of Economic Literature*, pages 13–47.

Tucker, A. L., Peterson, D. R., and Scott, E. (1988). Tests of the black-scholes and constant elasticity of variance currency call option valuation models. *The Journal of Financial Research*, pages 201–.

Tucker, A. L. and Pond, L. (1988). The probability distribution of foreign exchange price changes: Tests of candidate processes. *The Review of Economics and Statistics*, pages 638–647.

Yor, M. (2008). *Aspects of Mathematical Finance*. Springer, Berlin Heidelberg.

Yu, J. (2005). On leverage in a stochastic volatility model. *Journal of Econometrics*, pages 165–178.

8 Appendix

8.1 Tables

Moneyness (K/S - 1)	Number of Observations	Average Errors	Standard Deviations
Short-Term (0-3 month) calls			
<-6%	445 (646)	0.315% (0.027%)	0.277% (0.048%)
[-6%,-4%]	697 (1140)	0.314% (0.016%)	0.336% (0.055%)
[-4%,-2%]	1715 (1051)	0.366% (0.011%)	0.400% (0.070%)
[-2%,-1%]	1179 (1151)	0.309% (0.011%)	0.331% (0.076%)
[-1%,0%]	1331 (1360)	0.227% (0.011%)	0.399% (0.085%)
[0%,1%]	1049 (1053)	0.410% (0.006%)	0.465% (0.100%)
[1%,2%]	495 (499)	0.268% (-0.030%)	0.376% (0.104%)
[2%,4%]	691 (334)	0.413% (-0.003%)	0.446% (0.113%)
[4%,6%]	313 (576)	0.187% (-0.039%)	0.288% (0.116%)
>6%	422 (534)	0.053% (-0.070%)	0.235% (0.115%)
Medium-Term (3-6 month) calls			
<-6%	342 (492)	0.515% (-0.027%)	0.438% (0.076%)
[-6%,-4%]	358 (434)	0.568% (-0.009%)	0.537% (0.088%)
[-4%,-2%]	599 (333)	0.650% (-0.009%)	0.571% (0.083%)
[-2%,-1%]	414 (434)	0.710% (0.013%)	0.555% (0.086%)
[-1%,0%]	479 (476)	0.479% (-0.001%)	0.556% (0.094%)
[0%,1%]	419 (414)	0.593% (-0.009%)	0.583% (0.094%)
[1%,2%]	311 (322)	0.675% (-0.015%)	0.610% (0.116%)
[2%,4%]	293 (190)	0.597% (-0.046%)	0.513% (0.125%)
[4%,6%]	190 (237)	0.381% (-0.062%)	0.528% (0.125%)
>6%	210 (276)	0.213% (-0.125%)	0.412% (0.160%)

Table 12: Residuals from Garman-Kohlhagen model, call options. Bates values in parenthesis. Average Errors calculated as in Equation (9).

Moneyness (K/S - 1)	Number of Observations	Average Errors	Standard Deviations
Short-Term (0-3 month) puts			
<-6%	56 (93)	0.129% (-0.002%)	0.227% (0.135%)
[-6%,-4%]	96 (132)	0.189% (-0.001%)	0.293% (0.091%)
[-4%,-2%]	218 (151)	0.233% (0.023%)	0.370% (0.110%)
[-2%,-1%]	309 (283)	0.256% (0.003%)	0.292% (0.090%)
[-1%,0%]	566 (585)	0.122% (0.002%)	0.316% (0.097%)
[0%,1%]	933 (920)	0.223% (0.034%)	0.343% (0.093%)
[1%,2%]	905 (911)	0.186% (0.027%)	0.297% (0.080%)
[2%,4%]	1280 (788)	0.251% (0.022%)	0.318% (0.072%)
[4%,6%]	630 (939)	0.243% (0.024%)	0.236% (0.064%)
>6%	610 (828)	0.261% (0.033%)	0.234% (0.047%)
Medium-Term (3-6 month) puts			
<-6%	24 (38)	0.317% (-0.095%)	0.480% (0.124%)
[-6%,-4%]	25 (29)	0.301% (-0.053%)	0.290% (0.100%)
[-4%,-2%]	63 (41)	0.335% (-0.097%)	0.511% (0.115%)
[-2%,-1%]	78 (76)	0.603% (-0.019%)	0.544% (0.132%)
[-1%,0%]	184 (175)	0.337% (-0.021%)	0.466% (0.120%)
[0%,1%]	209 (208)	0.386% (-0.009%)	0.538% (0.096%)
[1%,2%]	276 (271)	0.624% (0.001%)	0.634% (0.095%)
[2%,4%]	419 (197)	0.412% (0.000%)	0.444% (0.087%)
[4%,6%]	316 (410)	0.463% (-0.011%)	0.456% (0.066%)
>6%	540 (663)	0.324% (-0.003%)	0.291% (0.050%)

Table 13: Residuals from Garman-Kohlhagen model, put options. Bates values in parenthesis. Average Errors calculated as in Equation (9).

8.2 Figures

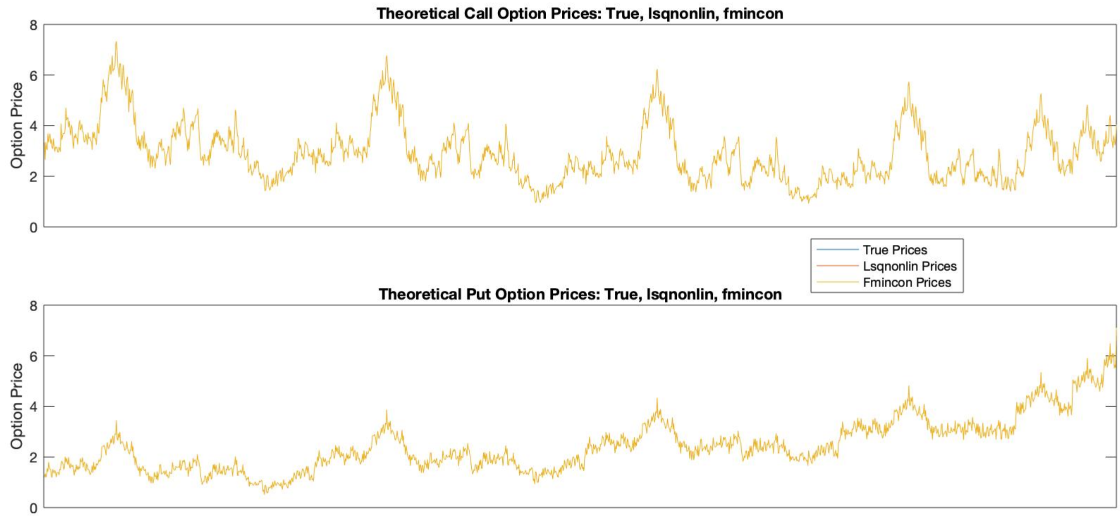

Figure 12: Chronological order of theoretical option prices separated into calls and puts.

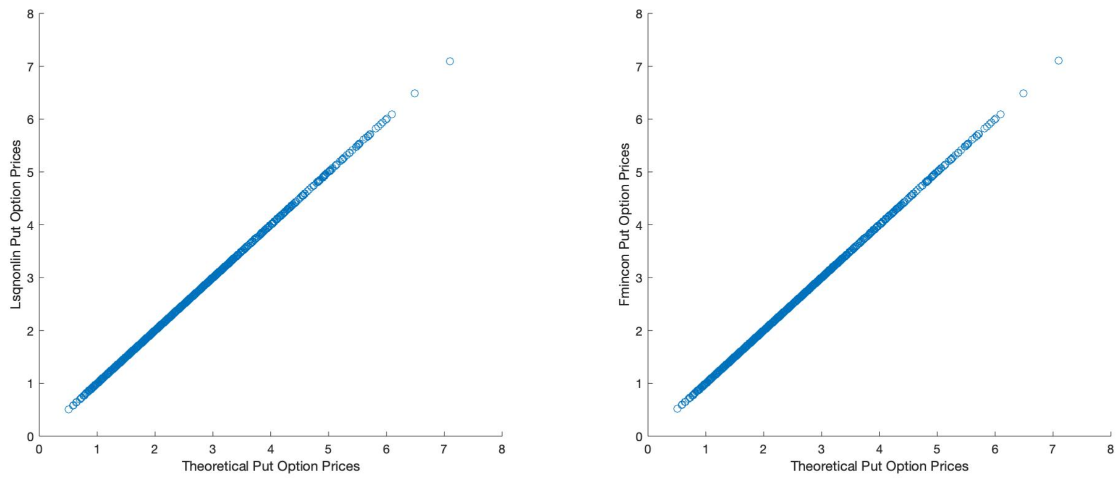

Figure 13: Scatterplot of theoretical put options prices against `lsqnonlin` and `fmincon` estimated prices.

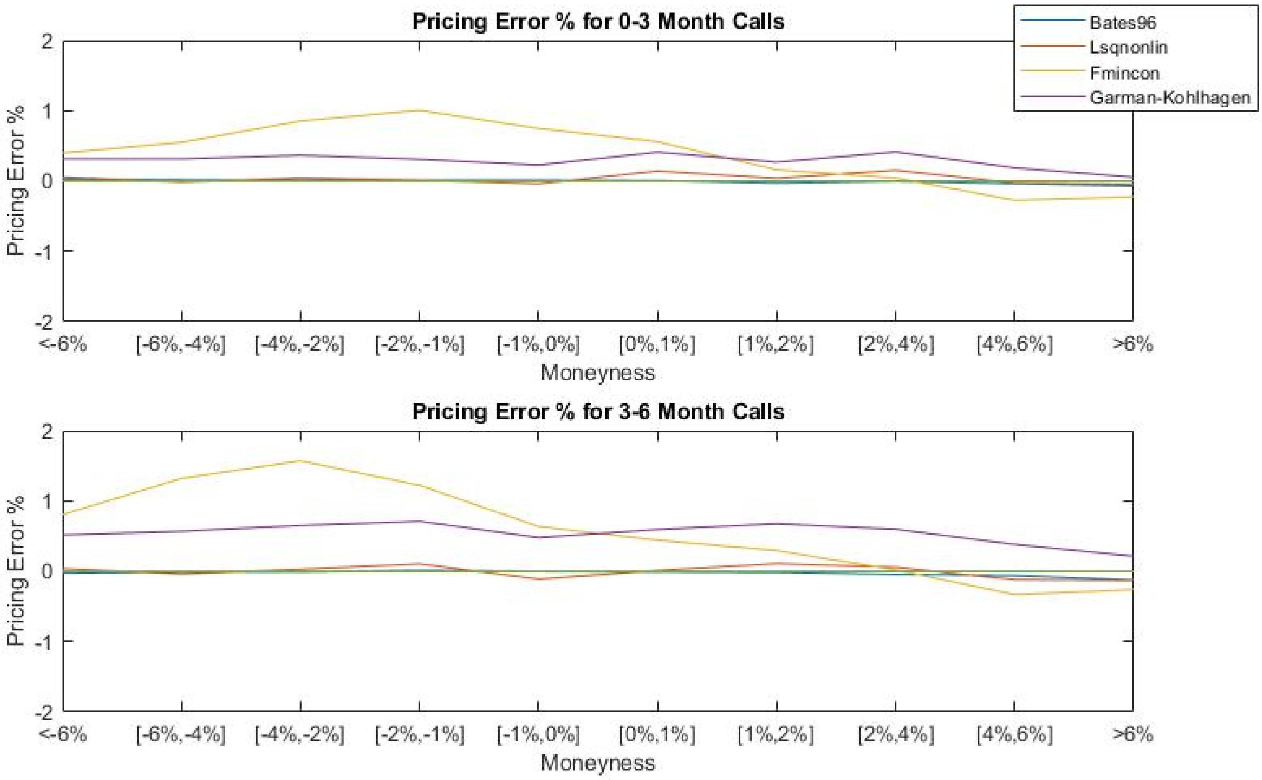

Figure 14: Pricing error over moneyness and maturity.

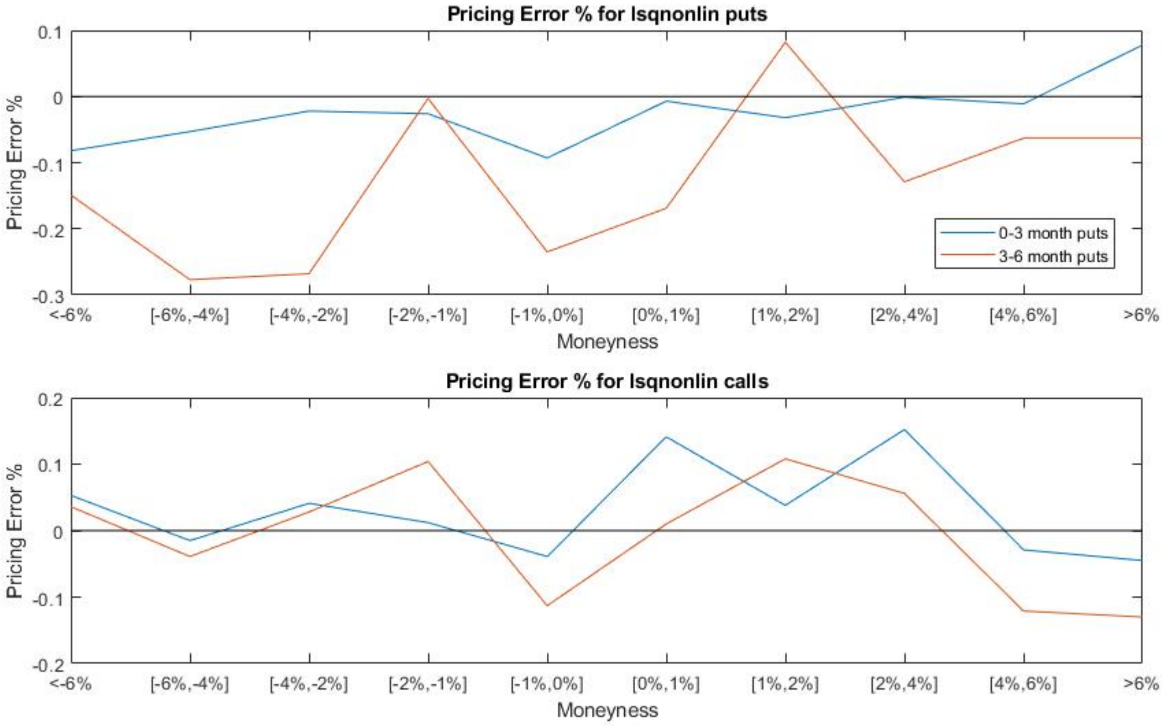

Figure 15: `lsqnonlin` average error over moneyness