Abstract

The demand for centralised risk metric services is increasing as the regulations and legislations in the financial market becomes stricter. In cooperation with TriOptima a framework for such service is presented. The framework is constructed such that parallel computation using graphical processing units (GPUs) is possible, with the goal to obtain maximal computational speed.

The focus of this thesis is to introduce the pricing framework as well as the procedure of extending the asset class coverage. Four models for pricing crude oil derivatives are tested, each with different characteristics. We have successfully presented a workflow for developing pricing models in the computational framework introduced, with historical data and asset price dynamics in focus. The possible weaknesses are highlighted and improvements are proposed.

Contents

Chapter 1

Introduction

The financial derivatives market can be divided into two main groups based on how the contracts are traded. The exchange traded derivatives are highly standardised. A central clearing house acts as the counterparty between the buyer and seller, and takes the credit risk of both parties. Calculating risk metrics for exchange traded derivatives is because of this a manageable task, even for large portfolios. On the over the counter (OTC) market derivatives are traded between two parties directly without going through an exchange. The contracts are negotiated to fit both parties and are thus not standardised. The use of central clearing houses to manage counterparty risk is present but not a necessity. Calculating risk metrics for large portfolios of OTC derivatives is in some sense more important since the counterparty risk generally is larger. The lack of standardisation makes this a huge task for small banks and institutions. It is not uncommon these calculations take hours and are done overnight.

There is a demand on the market for centralised risk metric services. When developing centralised computational services performance is crucial. One such service is *triCalculate* which is developed by TriOptima[1], a provider of risk management services and post trade infrastructure for the OTC derivatives market. The foundation of their service is described in Albanese et al. [2011]. By using discrete Markov-chain models and performing highly parallel calculations with the help of graphical processing units (GPUs) a significant speedup versus traditional methods is achieved.

The main objective of this thesis is to derive a workflow for, and expand the asset class coverage of such a centralised service to include commodity contracts. Since different commodities have different price dynamics they will each need different models. A reasonable limitation is therefore to study

[1] `http://www.trioptima.com`

a single commodity, and further, only a few specific contracts. Crude oil futures are the most traded commodity derivatives on the markets today and is therefore a good choice. Options written on these futures will also be considered. Note that the contracts covered are exchange traded and not OTC traded, even though we are interested in OTC derivatives. The reason behind this choice is the difficulty to obtain good market data for the OTC trades.

The goal is to replicate quoted market prices for these contracts by modeling the crude oil spot price in the computational framework used by *triCalculate* and calibrate the models to a calibration portfolio. The spot price is what drives the price for all derivatives, exchange traded as well as OTC, so using exchange data should be a non-issue.

The thesis begins with a brief introduction of the commodities market and some common commodity derivatives. Then some models we will consider are presented followed by a walkthrough of the pricing method. The results are then presented showing the performance of the different models. Finally we conclude the results and suggest further improvements.

Chapter 2

Commodity Markets

2.1 Markets

Crude oil is not a homogeneous product because it varies in e.g. density and sulfur content. The two major grades are West Texas Intermediate (WTI) from the US and Brent crude oil sourced in the North Sea. The largest market places for trading crude oil derivatives are the New York Mercantile Exchange (NYMEX) and the Intercontinental Exchange (ICE). The basic factors which affects the price of a commodity are supply and demand. Storability is also an important factor. A stock for example, can be stored indefinitely while electricity cannot be stored at all. Crude oil is considered a storable commodity.

2.2 Spot price dynamics

2.2.1 Mean reversion

In other commodity markets such as electricity or most grains mean reversion is a fact [Escribano et al., 2011]. The last decade the oil price has been rising without falling back to, what appears to be, a normal level. It has been debated if mean reversion really is a property in crude oil prices, see e.g. Geman [2007]. In the scope of this thesis no models with mean reverting dynamics will be considered but an example of such model can be found in Gibson and Schwartz [1990].

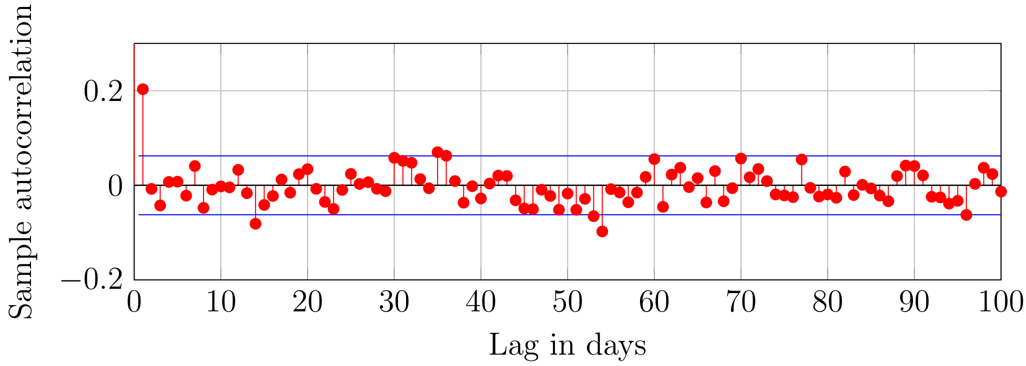

Figure 2.1: Autocorrelation of Brent Index returns from 2010-01-01 to 2013-12-31.

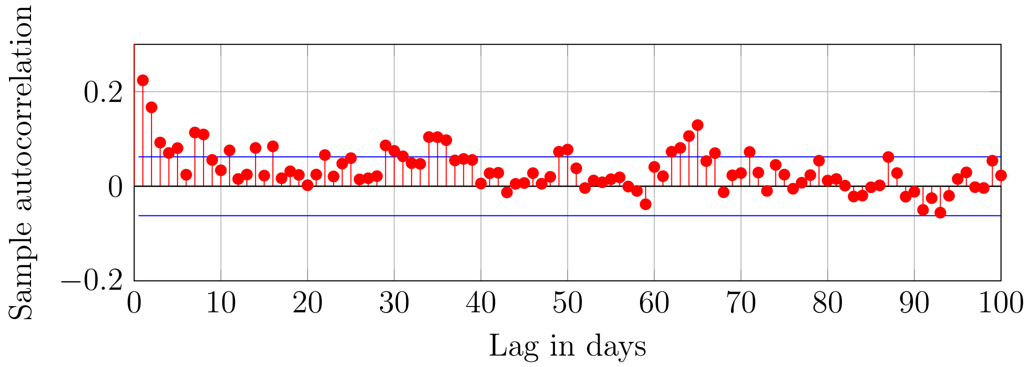

Figure 2.2: Autocorrelation of squared daily returns of the Brent Index from 2010-01-01 to 2013-12-31.

2.2.2 Lack of seasonality

The electricity price definitely has seasonal behaviour. Not only when it comes to the yearly cycles but also week cycles where the demand is higher on working days and lowe in the weekends. Even intraday cycles are present, the demand is higher in the day and lower in the night. These effects are augmented by the lack of storability. In the crude oil market seasonality is not an issue. Oil has obviously already been stored for millions of years, and even when exposed to air it can still be stored for months. This smooths out any seasonal supply problems that might exist. Further argument for the lack of seasonality in crude oil prices can be seen in Figure 2.1. We see that the autocorrelation of daily prices are negligible.

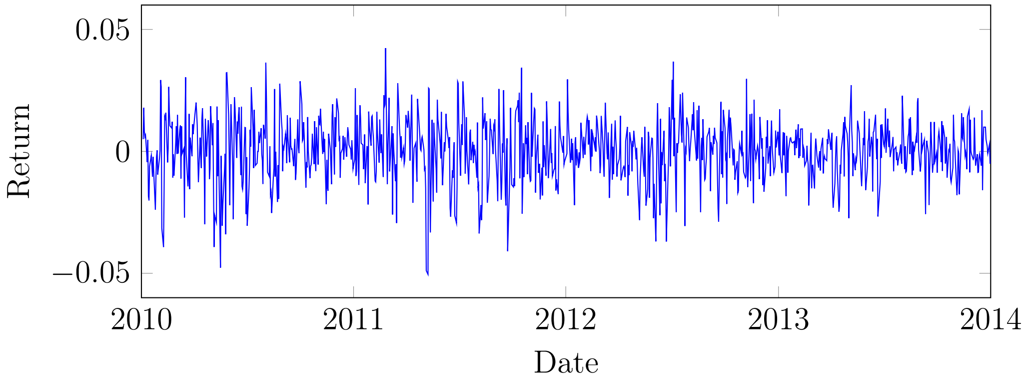

Figure 2.3: Daily returns of the Brent Index from 2010-01-01 to 2013-12-31.

2.2.3 Volatility clustering

In Figure 2.2 the autocorrelation of the second moment of the returns is shown. The figure indicates that the volatility on a specific date is dependent on previous volatilities. The dependence is not very strong but still significant. This effect is called volatility clustering. The result is that there are periods with low volatility and there are stressed periods of high volatility. Just by looking at the returns themselves this property can be hard to observe. Volatility clustering is usually modeled using stochastic volatility, which we will take a look at in Section 4.3.

2.2.4 Excess kurtosis

While returns often are assumed to be normally distributed, it is rarely the case. By observing market data we often find that the return distribution is more fat-tailed. In Figure 2.4 a normal probability plot of the Brent Index returns is shown. We observe that the tails are heavier than for a Gaussian distribution. The way excess kurtosis will be handled in this thesis is by adding jumps to the models. Another common method is to use the student t-distribution in the diffusion term of the models.

2.3 Commodity contracts

2.3.1 Spot Price

The spot market for commodities consist of commercial contracts where the transaction is to be performed immediately or with a minimal lag. This

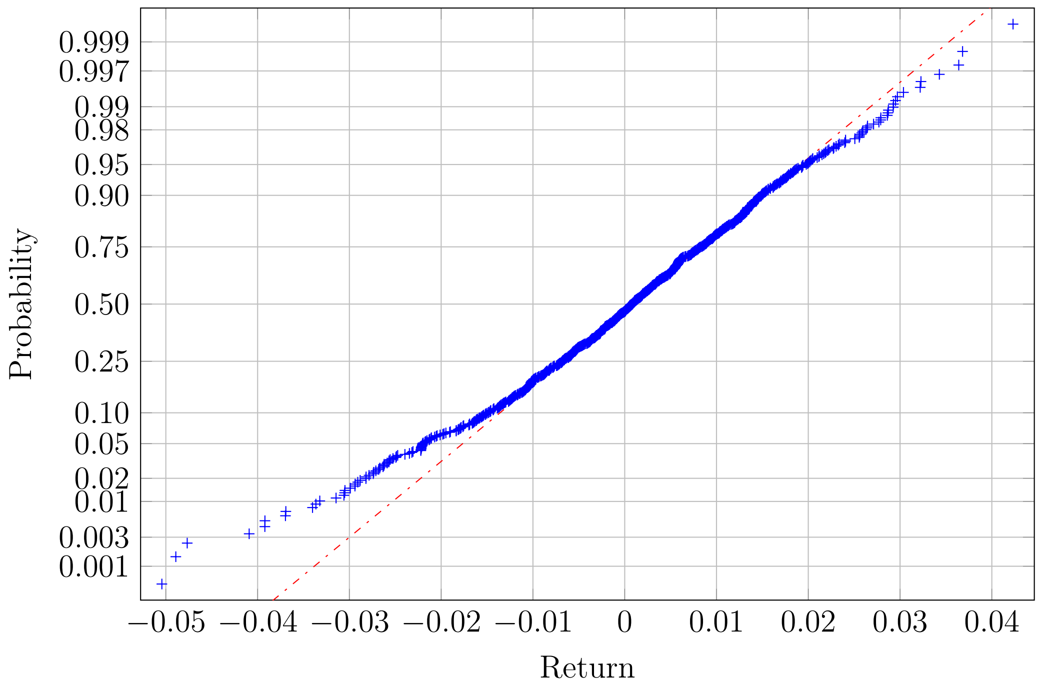

Figure 2.4: Normal probability plot of Brent Index returns from 2010-01-01 to 2013-12-31.

market is not always liquid and usually not transparent but still makes up a large portion of the commodities market.

2.3.2 Forwards

A forward contract is a bilateral agreement between two parties A and B signed at time t_0. Party A is obligated to pay the predetermined price at time t_0 whilst party B is obligated to deliver the underlying asset at a fixed future time T. These contracts are usually traded over the counter (OTC) and thus fully exposed to credit risk. The forward contract can either be connected to delivery of a physical product or be a purely financial contract.

2.3.3 Futures

A futures contract is similar to the forward contract but is highly standardised in terms of maturity, volume, quality and quantity. Futures contracts are traded at exchanges, such as NYMEX and ICE and are cleared at their respective clearing houses. The clearing house act as counterparty to both the seller and the buyer and require them to make an initial margin deposit.

The contracts are marked-to-market at a daily basis, which means the parties have to adjust their deposits according to their profit or loss at the end of each day, therby reducing credit risk. Futures contracts are naturally the most liquid contracts in the commodity markets.

2.3.4 Options

A call option gives the holder the right, but not the obligation, to buy the underlying asset at a specific maturity date to a predetermined strike price. The put option gives the holder the right to sell the underlying asset. The two most common types of options are European style options which only can be exercised at a specified maturity date, and the American style option which can be exercised at any time before or at the maturity date. It can be shown that the price of an American call option is equal to the price of the corresponding European call option [Björk, 2009]. Options are both exchange traded and traded OTC.

2.3.5 Swaps

Commodity swap contracts are agreements where two parties exchange fixed price and floating spot price with each other over a specified period. The majority of swap contracts in the commodities market are based on oil prices. Swap contracts are both exchange and OTC traded.

2.4 Relation between forward, future and spot price

Under the assumption that the market is arbitrage free, the relationship between spot and forward prices is

$$f^T(t) = S(t)e^{(r-\delta)(T-t)}, \tag{2.1}$$

where r is the rate and δ is the convenience yield. In the case of non-constant but deterministic rate and yield the above expression generalises to

$$f^T(t) = S(t)e^{\int_t^T (r_u-\delta_u)\mathrm{d}u}. \tag{2.2}$$

When the rate and convenience yield are deterministic the forward price $f^T(t)$ will coincide with the future price $F^T(t)$ [Geman, 2005, Chapter 2]. We will use this relation and price all future contracts as forwards. Futures can otherwise be tricky to work with since the margining affects the price slightly.

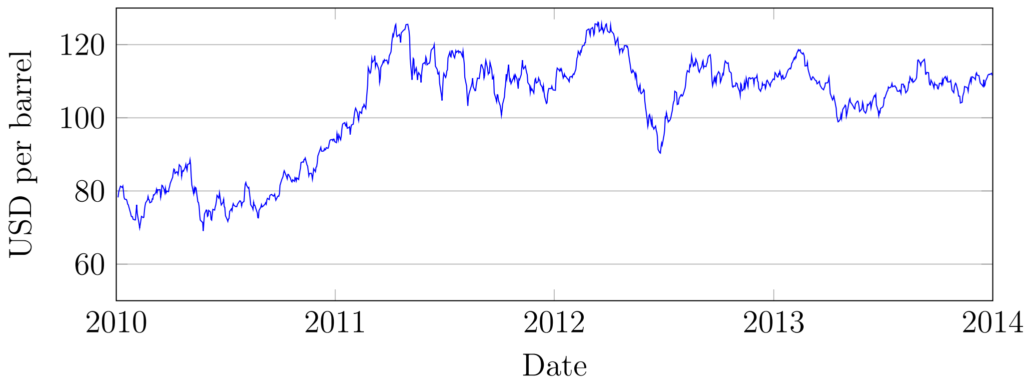

Figure 2.5: ICE Brent Index from 2010-01-01 to 2013-12-31.

2.5 Data source

In this thesis one month of data from November 2013 will be used. This section presents the sources.

2.5.1 ICE Brent index

The ICE Brent Index is a weighted average of settlement prices for Brent futures. This index will be used as a proxy for the spot price. As said in Section 2.3.1 the spot price is not always a very transparent market and good data can be hard to obtain. The Brent index and the method to calculate it is publicly available. Another reason for modeling the index instead of the spot price is that the futures we will use are settled on the index price. In Figure 2.5 the ICE Brent index is shown.

2.5.2 ICE Brent futures

Daily settlement prices of ICE Brent futures with up to six years to maturity will be used. Contracts with some maturities are not very liquid. We will therefore only consider futures with a daily volume of at least 1000 trades. This will almost certainly include all contracts with maturity in up to one year, and also the 18 month, 24 month and 36 month contracts. An example of a forward curve is shown in Figure 2.6.

2.5.3 ICE Brent american style options

ICE Brent American style options have the ICE Brent future price as underlying asset. Only call options will be considered, and they will be priced as

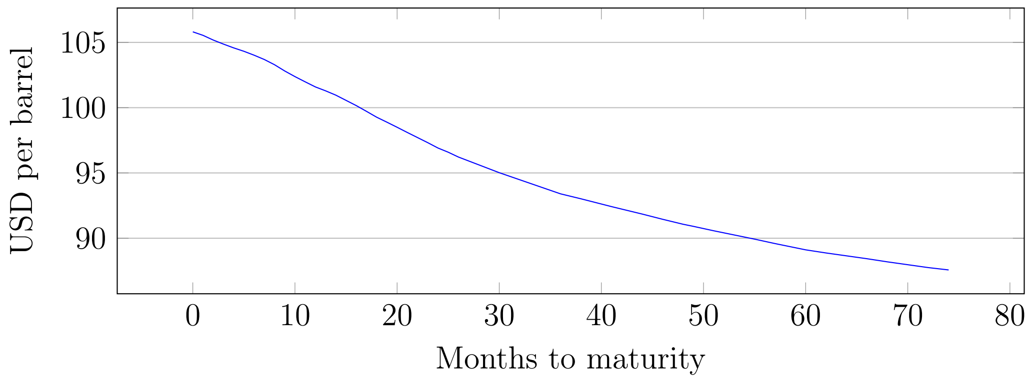

Figure 2.6: The ICE Brent forward curve 2013-11-12 is in contango.

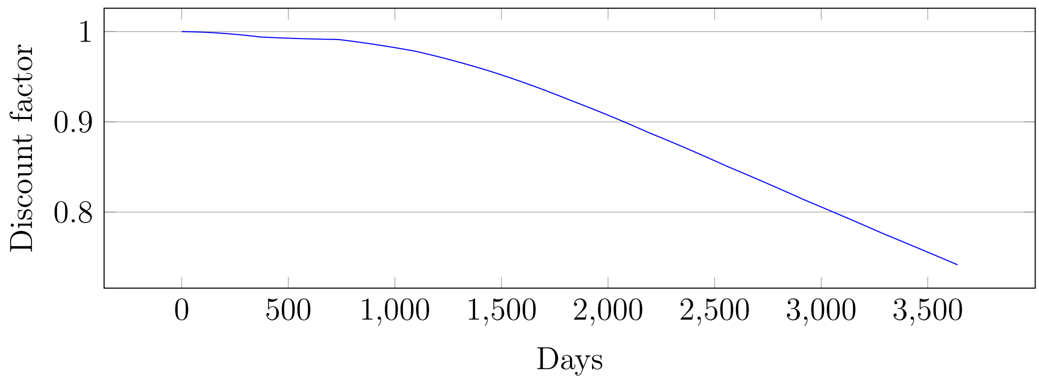

Figure 2.7: Discount curve for 2013-11-12.

if they were of European exercise. As in the case with futures only the most liquid contracts will be considered. This means that only the one month contracts will be considered and only some strike levels close to the money.

2.5.4 Discount curve

We will discount all prices using discount curves collected from the market. An example of one such curve is shown in Figure 2.7. When pricing options we will need the forward rate discount factor used to discount from time T_2 back to T_1. This can be obtained by simply dividing the discount factor for T_1 with the one for T_2 using the discount curve for the day we want to price the option at.

Chapter 3

Markov chain pricing framework

The pricing method described in this chapter is somewhat simplified version of the one used by TriOptima in their risk metrics service. It is closely related to the standard partial differential equation (PDE) methods using finite differences, but not exactly equivalent.

3.1 Expected payoff

The risk neutral valuation formula states that any simple claim can be priced at the current time t as the discounted expected payoff Φ at maturity T,

$$\Pi(t) = P(t,T) \cdot \mathbb{E}^{\mathbb{Q}}\left[\Phi(T)|\mathcal{F}_t\right] \tag{3.1}$$

where $P(t,T)$ is the the price of the zero coupon bond with expiration date T at time t. Essentially the core of derivative pricing methods is to calculate or approximate the expected value in one way or another. In this thesis we will use an approximative method, assuming the underlying asset price S_t only can assume a finite number of values $\Omega = \{z_1, \ldots, z_n\}$. The transition probability matrix P is constructed so that the element at row i and column j is

$$p_{i,j} = \Pr(S_T = z_j | S_t = z_i). \tag{3.2}$$

As usual with transition probability matrices

$$\sum_{k=1}^{n} p_{i,k} = 1 \quad \text{for all } i = 1, \ldots, n. \tag{3.3}$$

If the probability mass function $f_{S_t}(x) = \Pr(S_t = x)$ has the corresponding column vector $\bar{f}_{S_t}$, then

$$\bar{f}_{S_T} = P\bar{f}_{S_t}. \tag{3.4}$$

The expected payoff is then easily calculated as

$$\mathbb{E}^{\mathbb{Q}}\left[\Phi(T)|\mathcal{F}_t\right] = \sum_{i=1}^{n} f_i^{S_T}\Phi(f_i^{S_T}). \tag{3.5}$$

In the following section we will focus on how to obtain the transition probability matrix P given any model described by a stochastic differential equation (SDE).

3.2 Kolmogorov forward equation

Given a general d-dimensional diffusion process with m independent Brownian motion terms

$$\mathrm{d}X_t = a(X_t,t)\mathrm{d}t + b(X,t)\mathrm{d}W_t \tag{3.6}$$

where $a(X_t,t)$ is a $d\times 1$ column vector and $b(X_t,t)$ is a $d\times m$ matrix, the *Kolmogorov forward equation* states that the probability density of the solution to the above SDE satisfies

$$\begin{aligned}\frac{\partial f(x,t)}{\partial t} =& -\sum_{i=1}^{d}\frac{\partial}{\partial x^{(i)}}\left[a_i(x,t)f(x,t)\right]\\ &+\frac{1}{2}\sum_{i=1}^{d}\sum_{j=1}^{d}\frac{\partial^2}{\partial x^{(i)}\partial x^{(j)}}\left[B_{i,j}(x,t)f(x,t)\right]\end{aligned} \tag{3.7}$$

$$f(x,0) =\delta_{X_0}(x) \tag{3.8}$$

where

$$B(x,t) =b(x,t)b(x,t)^{\intercal}. \tag{3.9}$$

This PDE is solved with finite difference methods using forward difference in time and central difference in the price dimensions. The discretisations are given by

$$\frac{\partial f(x,t)}{\partial t}\approx\frac{f(x_i,t+\Delta t)-f(x_i,t)}{\Delta t} \tag{3.10}$$

in time, and

$$\frac{\partial f(x,t)}{\partial x}\approx\frac{f(x_{i+1},t)-f(x_{i-1},t)}{2\Delta x} \tag{3.11}$$

$$\frac{\partial^2 f(x,t)}{\partial x^2}\approx\frac{f(x_{i+1},t)-2f(x_i,t)+f(x_{i-1},t)}{\Delta x^2} \tag{3.12}$$

Factor 2	0				1				$\cdots$				$\|\Omega_2\|$			
Factor 1	0	1	$\cdots$	$\|\Omega_1\|$	0	1	$\cdots$	$\|\Omega_1\|$	0	1	$\cdots$	$\|\Omega_1\|$	0	1	$\cdots$	$\|\Omega_1\|$

Figure 3.1: Mapping the state matrix to a vector.

in space. To achieve satisfactory computational performance we limit the number of dimensions in the diffusion process by only considering one-factor and two-factor models. This will also help simplify the solution of the Kolmogorov forward equation significantly. In the case of a two-factor process we will also need the cross derivative

$$\begin{aligned}\frac{\partial^2 f(x^{(1)}, x^{(2)}, t)}{\partial x^{(1)} \partial x^{(2)}} \approx \frac{1}{4\Delta x^{(1)} \Delta x^{(2)}} \Big[& f(x^{(1)}_{i-1}, x^{(2)}_{i-1}, t) - f(x^{(1)}_{i+1}, x^{(2)}_{i-1}, t) \\ & - f(x^{(1)}_{i-1}, x^{(2)}_{i+1}, t) + f(x^{(1)}_{i+1}, x^{(2)}_{i+1}, t) \Big] .\end{aligned} \tag{3.13}$$

By inserting the discretised derivatives into 3.7 we get

$$\begin{aligned} f(x_i, t + \Delta t) = {} & f(x_i, t) \\ & - \frac{\Delta t}{2\Delta x} \Big[a(x_{i+1}, t) f(x_{i+1}, t) - a(x_{i-1}, t) f(x_{i-1}, t) \Big] \\ & + \frac{\Delta t}{\Delta x^2} \Big[B(x_{i+1}, t) f(x_{i+1}, t) - 2B(x_i, t) f(x_i, t) \\ & \qquad + B(x_{i+1}, t) f(x_{i-1}, t) \Big] . \end{aligned} \tag{3.14}$$

The two dimensional case is similar to the above.

3.3 State mapping

When we only have one dimensional models the state corresponding to a row in the transition probability matrix is natural. The ith row represents the ith state in Ω. When there are two dimensions we have to choose two state spaces Ω_1 and Ω_2, one for each dimension. Each row in the matrix now represents a pair of states. We will map them as Figure 3.1 illustrates.

3.4 Markov chain analogy

In the continous case the initial condition is a Dirac delta function. When we discretise by limiting the possible prices to the state space Ω we have to modify the initial condition as well. The initial state X_0 is observed in the market and is therefore fixed. Instead we utilise the freedom of choosing the state space to fit the calculations. By selecting Ω such that $X_0 = z_k \in \Omega$ the discrete version of the initial condition in (3.8) becomes

$$p(z_j, 0) = \begin{cases} 1 & \text{if } j = k \\ 0 & \text{otherwise} \end{cases}. \tag{3.15}$$

When we insert this into (3.14) and the corresponding two dimensional case we will get the neccessary values $p_{i,j}$ to populate the transition probability matrix P. The one dimensional case gives us

$$\begin{aligned} p_{i,i+1} &= \frac{\Delta t}{2\Delta x^2}\Big[\Delta x \cdot a(x_i,t) + B(x_i,t)\Big], \\ p_{i,i} &= 1 - \frac{\Delta t}{\Delta x^2} B(x_i,t), \\ p_{i,i-1} &= \frac{\Delta t}{2\Delta x^2}\Big[-\Delta x \cdot a(x_i,t) + B(x_i,t)\Big]. \end{aligned} \tag{3.16}$$

The two dimensional caclulations are somewhat tedious and are therefore omitted. The interesting result of using the boundary condition in (3.15) is that the transition probability matrix becomes tridiagonal in the one dimensional case, and block-tridiagonal in the two dimensional case.

3.4.1 Boundary conditions

All rows in the transition probability matrix now has a unit sum, except for at the boundary states. It is natural to assume that

$$\Pr(S_T = z_j | S_t = z_i) = 0 \quad \text{if } z_j \notin \Omega. \tag{3.17}$$

The state space Ω is chosen such that the boundary states are far enough from the current asset price that the probability of reaching the boundaries during the contract period is small. One type of boundary condition that is very simple to implement is to use absorbing boundary condition. The transition probability matrix is easily modified to achieve this and is now ready to use.

3.4.2 Jumps

Jump diffusion models does not add much complexity to the creation of the probability matrix. The jump probabilities are simply added to the corresponding element of the matrix. The diagonal is then adjusted so that each row still has a unit sum.

3.5 Time step selection

We cannot generally create the transition probability matrix corresponding to prices at the current time t and at some future time T. The problem is that we have to ensure the calculations are numerically stable. The final transition probability matrix is created as the product of matrixes corresponding to shorter time steps Δt,

$$P = \prod_{i=0}^{k-1} P_{\Delta t} \tag{3.18}$$

for some integer k. We note that the expressions in (3.16) implies that the matrix $P_{\Delta t}$ can be expressed as

$$P_{\Delta t} = G + \Delta t \cdot I \tag{3.19}$$

where I is the identity matrix and G is a matrix that does not depend on Δt. The time step Δt can then be selected small enough, such that it satisfies the Courant condition

$$\Delta t < \frac{1}{\max(|\operatorname{diag}(G)|)}. \tag{3.20}$$

Finally, the time time step is selected such that $\Delta t \cdot 2^n = T - t$ for some integer n. This makes it possible to use consecutive squaring to achieve maximal computational performance without compromising on stability. In the case where it normally would take k full matrix multiplications to get to the final P we only have to perform $n = \log_2 k$ full matrix multiplications if we utilise consecutive squaring. The price we have to pay is that our models will have to be time homogeneous. In section 4.4 however, we will present a workaround for this limitation. For more information on this subject see Albanese et al. [2011, Section 5].

Chapter 4

Spot price models

In this section, the models we will test are presented. Four models will be considered, each one with different characteristics.

4.1 Geometric brownian motion

Geometric Brownian motion described by the SDE

$$\mathrm{d}S_t = \mu S_t \mathrm{d}t + \sigma S_t \mathrm{d}W \tag{4.1}$$

is the fundamental model in asset pricing. It is used in the famous Black-Scholes model introduced in Black and Scholes [1973]. This is the most simple model we will test and is used as a building block in the other models. The model only has one-factor and one Brownian motion term. Only two parameters will need to be calibrated, so the risk for over fitting is small. It should be easy to calibrate this model to market data but the dynamics might not be captured fully with a simple model like this.

4.2 Merton

The Merton model proposed in Merton [1976] is an extension of geometric Brownian motion. The model is described by

$$\mathrm{d}S_t = \mu S_t \mathrm{d}t + \sigma S_t \mathrm{d}W_t + S_t \mathrm{d}Z_t \tag{4.2}$$

where the additional term $\mathrm{d}Z_t$ is a compound Poisson process. The jumps occur with intensity λ and has log-normally distributed amplitude k according to

$$\ln(1+k) \sim \mathcal{N}(\ln(1+\chi) - \frac{1}{2}\alpha^2, \alpha^2). \tag{4.3}$$

Note that the support of the jump amplitude is $[-1,\infty)$. The reason for adding jumps to the model is to capture the fat tails of the return distribution. Three parameters are used to describe the jumps, resulting in a total of five parameters we need to calibrate to market data. The calibration should still be fairly easy but the risk for over fitting is greater. There is also a greater risk for ending up in local minimas during calibration.

4.3 Heston

Another way to extend the geometric Brownian motion is to add stochastic volatility. The Heston model described by

$$\begin{aligned} \mathrm{d}S_t &= \mu S_t \mathrm{d}t + \sqrt{V_t} S_t \mathrm{d}W_t^S \\ \mathrm{d}V_t &= \kappa(\theta - V_t)\mathrm{d}t + \gamma\sqrt{V_t}\mathrm{d}W_t^V \\ \mathrm{d}W_t^S \mathrm{d}W_t^V &= \rho \mathrm{d}t \end{aligned} \tag{4.4}$$

was first proposed in Heston [1993]. A second factor modeling the volatility of the price process as a Cox–Ingersoll–Ross (CIR) model is added. By doing this the volatility clustering effect could be captured. The cost of adding a second factor to the model is large though. Just by adding eight possible volatility levels we have to reduce the number of possible price states with a factor eight, see Figure 3.1 for an illustration. Additionally, the initial volatility V_0 is an unobserved state and will have to be calibrated together with the rest of the model parameters. This results in a total of six parameters, making the PDE more complicated and the calibration problem harder.

4.4 Fitted term structure

The last model differs from the previous ones in the sense that it allows for capturing the shape of the forward curve term structure. The idea is that we no longer will look at the distribution of underlying prices but instead a transformation of the underlying prices [Albanese et al., 2011, Section 6],

$$\bar{S}_t = \lambda(t) S_t. \tag{4.5}$$

The transformation captures all time variation so that an arbitrary time homogeneous model can be used to model the transformed prices

$$\mathrm{d}\bar{S}_t = a(\bar{S}_t)\mathrm{d}t + b(\bar{S}_t)\mathrm{d}W_t. \tag{4.6}$$

The model we will try is the simplest possible. Assume that the spot price S_t is modelled using geometric Brownian motion with time dependent drift

$$\mathrm{d}S_t = (r_t - \delta_t)S_t\mathrm{d}t + \sigma S_t \mathrm{d}W_t. \tag{4.7}$$

The transformation $\lambda(t) = \exp(-\int_0^t r_u - \delta_u \mathrm{d}u)$ captures the forward curve precisely. Itō's lemma then gives us

$$\begin{aligned}\mathrm{d}\bar{S}_t &= \left(-S_t(r_t - \delta_t)\mathrm{e}^{-\int_0^t r_u - \delta_u \mathrm{d}u} + S_t(r_t - \delta_t)\mathrm{e}^{-\int_0^t r_u - \delta_u \mathrm{d}u}\right)\mathrm{d}t \\ &\quad + \sigma S_t \mathrm{e}^{-\int_0^t r_u - \delta_u \mathrm{d}u}\mathrm{d}W_t = \sigma \bar{S}_t \mathrm{d}W_t.\end{aligned} \tag{4.8}$$

This model has no drift and there is only one single parameter to calibrate.

Chapter 5

Method

The method we use to evaluate and compare models is divided in two parts. The first part is the calibration of model parameters using the market data from the sources presented in Section 2.5. The next step is to perform an out of sample goodness of fit test to assess which model we consider the best.

5.1 Parallel computing

The last decade the semiconductor industry have focused heavily on increasing computational power by making it possible to perfor multiple operations simultaneously. Parallel computing has gained immense popularity in a vast range of industries, especially in finance. While normal computational processing units (CPUs) can perform 2, 4, 8 or even 16 tasks in parallel modern graphical processing units (GPUs) can perform thousands. Some operations are particularly suitable for paralell computing, such as matrix multiplications, fourier transforms and Monte Carlo simulations. Programming frameworks, such as CUDA from Nvidia or OpenCL, together with basic linear algebra subprograms (BLAS) implementations makes it easy to utilise this raw computational power. Recent versions of Matlab also has parallel computation capabilities. GPU computing is not essential for the method itself, however it has been very important in the sence that the method intentionally is developed to suit such parallel operations. By implementing the pricing framework in this thesis with GPU capabilities we would notice a significant performance boost.

5.2 Calibration

The calibration is done once for each day in the dataset. We will only use the data corresponding to that particular day. The initial guess of parameters is the same for each calibration and is choosen arbitrarily but reasonable. One could of course use the previous days' calibration as initial guess but to minimise the risk of bad parameter fits to propagate, we choose not to. The loss function is very simple,

$$L(t;\theta) = \sum_k \left(f^{T_k}(t) - \tilde{f}^{T_k}(t;\theta) \right)^2 + \sum_l \left(c^{K_l}(t) - \tilde{c}^{K_l}(t;\theta) \right)^2 . \qquad (5.1)$$

It is simply the sum of squared errors between the observed future and option prices and the modelled prices. Weights could be assigned to each term, but then the problem of how such weights should be selected arises. The objective function is minimised using the damped least-squares (DLS) method.

The price states are selected uniform ranging from the lower price of 50 USD to the upper price of 150 USD. The prices are then shifted such that the state closest to the current spot price is moved to the exact current spot price. The volatility states in the Heston model are selected uniform from 0.005 to 0.06. No shifting is done to the volatility states since we do not know the exact initial volatility. All models have a total of 64 states. The Heston model have 8 states for each factor, all other models only have one factor and thus all states correspond to the spot price. If the implementation would utilise GPUs the total number of states could be increased to at least 1024, the limiting factor being the amount of memory available. Without GPUs the limiting factor is not so much the memory available but the patience to wait for the calibrations to finish. In Appendix A all the calibrated parameters are shown but the results are summarised in Chapter 6.

5.3 Validation

The model validation is performed by evaluating the loss function (5.1) each day, predicting the price of the contracts at the following day. The idea is that the model with the lowest average error also is the model we prefer. As we will see though, other factors will be taken into account. Along with the calibrations all validation results is presented in detail in Appendix A and summarised in Chapter 6

Chapter 6

Results

In this chapter all the results from the model calibration and validations are described and summarised. All results in entirety are located in Appendix A and should be simple to understand. Each row in the tables represent one day of data. The last four columns in each table are root mean squared error values. We separate the errors in one column dedicated to the forward price error and one column for option prices.

6.1 Models

We will start with an example of how the results are interpreted before we dive into the results for each model. In Figure 6.1 the calibrated variables when using the geometric Brownian motion model are shown. We see that the value for μ is fairly constant while the value for σ fluctuates somewhat more. If the shape of the forward curve don't change much during the calibration period and the calibration is good, the desired effect is an approximately constant value of μ, which we have in this case. The value of σ should also be constant but this is not what we observe. In this case the reason can be seen in Figure 6.2 and Figure 6.3. The dynamics of the geometric Brownian motion model cannot capture the exact shape of the forward curve. Since the options we use have the shortest forward contract as underlying, the mispricing observed propagates to the calibration of the option prices, causing the inability to calibrate σ properly. This reason could by itself motivate why we should transform prices before applying a model, like in the fitted term structure model. The same type of analysis is performed for all models and the observations are summarised in the rest of this section.

The results for the geometric Brownian motion model are shown in Table A.1. We see that the parameters are stable and fairly constant. The out

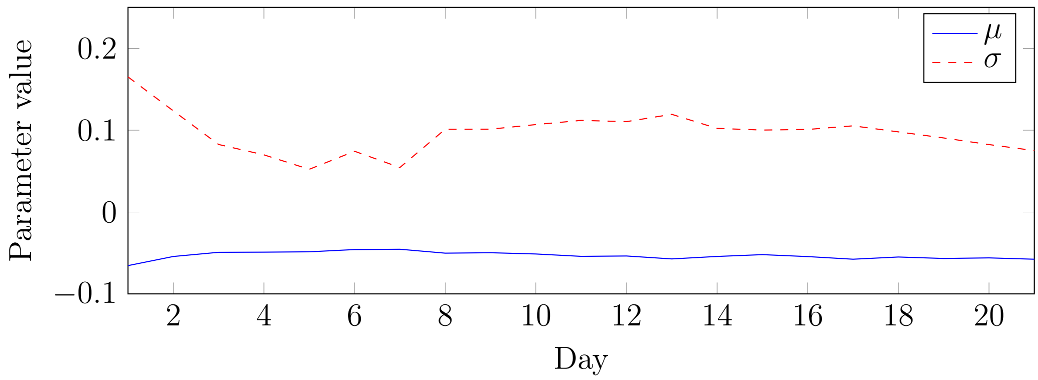

Figure 6.1: The calibrated parameter value for each day in the data set using the geometric Brownian motion model.

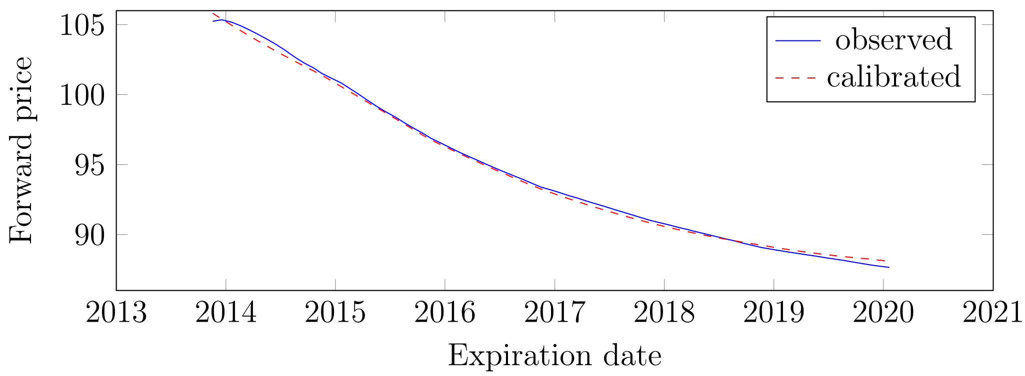

Figure 6.2: The calibrated forward curve with geometric Brownian motion compared to the actual forward curve at 2013-11-06.

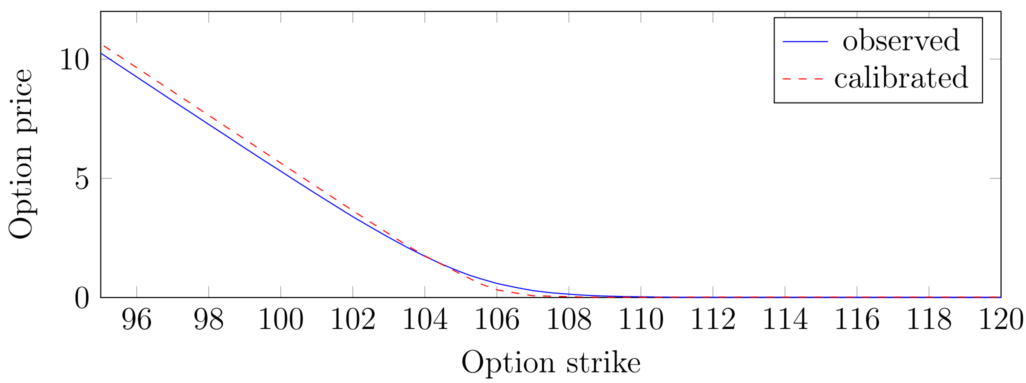

Figure 6.3: The calibrated option prices with geometric Brownian motion compared to the actual prices at 2013-11-06.

of sample fit is not much worse than calibration fit. This model is used as a baseline model which the other models are compared to.

In Table A.2 the results for the Merton model are shown. We note that the parameters for this model is stable as well. The fit is just barely better than the geometric Brownian motion but not good enough to motivate the calibration overhead introduced by the increased model complexity.

The results for the Heston model are shown in table A.3. We notice some possible issues with the calibration. Some of the parameters varies a lot while some are constant. It seems that the initial volatility parameter cannot be calibrated correctly with the methods used in this thesis. The correlation parameter assumes unreasonable values. A perfect negative correlation invalidates the arguments for choosing a two-factor model to begin with. Possible reasons why the calibration is particularly hard for this model are discussed in Chapter 7.

Last model to analyse is the fitted term structure model. The results are presented in Table A.4 We note that the parameter calibration is stable but is probably underestimated at the days where the options are close to expiry and the option curve has a sharper bend. The most notable result is that there are no errors in the forward curve calibration, which is a feature of this specific model. The model is not only stable and simple to calibrate but also has the best out of sample performance of the considered models.

6.2 Method

Besides comparing the models, we can conclude some valuable insights regarding the method itself and models in general. The first insight is that 64 states in total probably is lower than what should have been used to obtain sufficient accuracy. This especially means that we cannot sacrifice the size of state dimensions when using multi factor models. By increasing the distance between grid points the objective function seems to increase in complexity and the time it takes to perform a calibration increases. Since some model calibrations shows signs of beeing local minimisations, it is also likely that the DLS method for optimisation is not the best choice. During the work of this thesis other methods were tried, such as simulated annealing and multistart methods. Despite the fact that these methods generall outperformed the DLS method for the more complex models, the faster DLS method was chosen. The decision later turned out to be suboptimal.

Chapter 7

Discussion

None of the models we have tested stands out as particularly good. What we can observe though, is that some models show desirable features while some shows undesirable features. In this chapter we will try to identify why the problems occur and how to deal with them.

7.1 Models

The geometric Brownian motion is a simple model with few parameters to calibrate. The calibration itself is most probably as good as possible, we have probably not ended up in any local minimas. The price we pay for a simple model is that we don't get a good fit to the market data.

The reason for adding jumps to the geometric Brownian motion is to incorporate the thicker tails of the price distribution observed in the market data. The results for the Merton model does not show the significant performance gain we had hoped for. There are a numerous possible reasons for this but to identify which, we would have to spend some more time analysing the results. One possible reason is that the price distribution doesn't have the significant wide tails. While the tails in Figure 2.4 appear to be wider than a normal distribution it might not be enough. A further step would be to perform some statistical significance test to possibly reject the use of more heavy tailed distributions or jump models. The calibration of the Merton model did also take more than twice as long as the calibration of any of the other models. The reason of this is that by adding jumps the jump probability matrix representing the model no longer is tridiagonal or block-tridiagonal but full instead, increasing the objective function complexity and possibly adding multiple local minimas.

By adding a second factor to the geometric Brownian motion model describing the volatility we get the Heston model. The idea behind adding another factor is to model time dependent volatility. In Figure 2.2 the effect of heteroscedasticity is observed, but we have not performed any significance tests which would be the next step. We observe that of all models this has the worst calibration. The reason for this is most certainly that by adding en extra factor we limit the number of states in each factor drastically to achieve the same computational performance as the one-factor models. This will expand the distance between grid points and the pricing will get less accurate. Another problem with the two factor model is that the second factor is modelling a hidden state. We cannot observe the volatility of the price process and we do not know the initial volatility. The initial volatility does not seem to calibrate correctly with the methods we have used. To investigate why this is, some more time will have to be spent on researching. During the writing of this thesis two more two-factor models were tested. The first of them beeing the Bates model which is a combination of Heston and Merton. The second model we considered was a model introduced by Gibson and Schwartz [1990] where the second factor models the convenience yield. Both of these models were left out because of the problems with two factor models we have observed in the Heston model.

Last we will take a look at the fitted term structure model. This is not only the model with the best fit but also the simplest model to calibrate since it only have one single parameter. The reason this model outperforms the other is mainly because of it's ability to capture the shape of the forward curve, also leading to more accurate option prices. A problem that might occur if another data set is used though is that by assuming the forward curve is similar to the previous day's, fast changes will not be possible to model.

7.2 Method

Many of the reasons none of the models has great performance probably is due to the small scale of the computations in this thesis, not the framework itself. All calculations in this thesis was performed in a significantly smaller scale than what TriOptima is using in production. To be able to really draw any conclusion all computations would have to be redone in a larger scale. By adding prices for a larger range of contracts each day, the calibrations would be more accurate. Since local optimisation routines seem to fail some global or at least semi global optimisations will have to be performed. The goal of the thesis however, to develop a workflow for model creation in the computational framework, can be considered fulfilled.

Chapter 8

Conclusion

There are two main conclusions we can draw from the work performed in this thesis. The first one is that model calibration is very hard and the second one is that we should try to use as simple models as possible as long as the performance is adequate. If one of the models in this thesis has to be selected as the best it would have to be the fitted term structure model. An improvment to this particular model could be not to use the empirical forward curve. Instead, something like Nelson-Siegel functions could be used to create a parametrisation. While the empirical curve gives a perfect calibration fit the risk of over fitting is higher.

While none of the models are perfect in it's current forms there are a lot of possible improvements and possible further research projects. An interesting project is to research how we can combine the technique of transforming the prices before modelling with more complex models. Another project would be to investigate if it is possible to choose the state space in a better way. By modelling prices in a Markov chain we could have more freedom to create models than by using plain SDEs, something we cold have taken advantage of better. By using the theory stated in this thesis and using the workflow presented further research is facilitated.

Bibliography

Claudio Albanese, Toufik Bellaj, Guillaume Gimonet, and Giacomo Pietronero. Coherent global market simulations and securitization measures for counterparty credit risk. *Quantitative Finance*, 11(1):1–20, 2011. doi: 10.1080/14697688.2010.542633. URL `http://www.tandfonline.com/doi/abs/10.1080/14697688.2010.542633`.

David S. Bates. Jumps and stochastic volatility: Exchange rate processes implicit in deutsche mark options. *Review of Financial Studies*, 9(1):69–107, 1996.

Thomas Björk. *Arbitrage Theory in Continuous Time*. Oxford Finance Series. Oxford University Press, 2009. ISBN 9780191610295. URL `http://books.google.se/books?id=YoMr5Mbo6coC`.

Fischer Black and Myron Scholes. The pricing of options and corporate liabilities. *The Journal of Political Economy*, pages 637–654, 1973.

John C. Cox, Stephen A. Ross, and Mark Rubinstein. Option pricing: A simplified approach. *Journal of Financial Economics*, 7(3):229 – 263, 1979. ISSN 0304-405X. doi: http://dx.doi.org/10.1016/0304-405X(79)90015-1. URL `http://www.sciencedirect.com/science/article/pii/0304405X79900151`.

John C Cox, Jonathan E Ingersoll Jr, and Stephen A Ross. A theory of the term structure of interest rates. *Econometrica: Journal of the Econometric Society*, pages 385–407, 1985.

Alvaro Escribano, J Ignacio Peña, and Pablo Villaplana. Modelling electricity prices: International evidence. *Oxford Bulletin of Economics and Statistics*, 73(5):622–650, 2011.

Hélyette Geman. *Commodities and Commodity Derivatives: Modeling and Pricing for Agriculturals, Metals and Energy*. The Wiley Finance Series. Wiley, 2005. ISBN 9780470012185. URL http://books.google.se/books?id=X9C3jwEACAAJ.

Hélyette Geman. Mean reversion versus random walk in oil and natural gas prices. In *Advances in Mathematical Finance*, pages 219–228. Springer, 2007.

Rajna Gibson and Eduardo S. Schwartz. Stochastic convenience yield and the pricing of oil contingent claims. *The Journal of Finance*, 45(3):959–976, 1990.

Steven L. Heston. A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Review of Financial Studies*, 6(2):327–343, 1993.

Intercontinental Exchange. Brent crude futures, a. URL https://www.theice.com/productguide/ProductSpec.shtml?pdf=&specId=219.

Intercontinental Exchange. The brent index, b. URL https://www.theice.com/publicdocs/futures/ICE_Futures_Europe_Brent_Index.pdf.

Intercontinental Exchange. Brent crude american-style options, c. URL https://www.theice.com/productguide/ProductSpec.shtml?pdf=&specId=218.

Peter E. Kloeden and Eckhard Platen. *Numerical Solution of Stochastic Differential Equations*. Applications of mathematics : stochastic modelling and applied probability. Springer, 1992. ISBN 9783540540625. URL http://books.google.se/books?id=BCvtssom1CMC.

Robert C. Merton. Option pricing when underlying stock returns are discontinuous. *Journal of Financial Economics*, 3(1–2):125 – 144, 1976. ISSN 0304-405X. doi: http://dx.doi.org/10.1016/0304-405X(76)90022-2. URL http://www.sciencedirect.com/science/article/pii/0304405X76900222.

Charles R Nelson and Andrew F Siegel. Parsimonious modeling of yield curves. *Journal of Business*, 60(4):473–489, 1987.

George E Uhlenbeck and Leonard Salomon Ornstein. On the theory of the brownian motion. *Physical Review*, 36(5):823, 1930.

Appendix A

Tables

Table A.1: Calibrations and validation results for the geometric Brownian motion model in Section 4.1.

	μ	σ	rmse futures	rmse options	validation futures	validation options
2013-11-01	−0.0655	0.1651	0.9054	1.8065		
2013-11-04	−0.0543	0.1236	0.2784	0.3421	0.8664	1.7574
2013-11-05	−0.0492	0.0825	0.1878	0.1827	0.5503	1.0802
2013-11-06	−0.0490	0.0699	0.2364	0.2629	0.2711	0.2272
2013-11-07	−0.0486	0.0522	0.5012	1.6046	0.5376	1.6877
2013-11-08	−0.0458	0.0743	1.0465	1.0967	0.5744	0.3812
2013-11-11	−0.0455	0.0544	1.3580	2.1211	1.4819	2.2797
2013-11-12	−0.0502	0.1012	0.4313	0.6930	1.1012	1.3762
2013-11-13	−0.0497	0.1012	0.6700	0.9510	1.1453	1.1563
2013-11-14	−0.0512	0.1068	0.9063	1.3745	1.1295	1.5871
2013-11-15	−0.0541	0.1119	0.4033	0.8210	1.0839	1.4310
2013-11-18	−0.0537	0.1104	0.3843	0.8504	0.3236	0.7596
2013-11-19	−0.0573	0.1194	0.2886	0.1769	1.0200	0.2183
2013-11-20	−0.0543	0.1022	0.3941	0.7711	0.6003	0.5812
2013-11-21	−0.0521	0.1001	1.4754	2.0862	1.1279	1.7617
2013-11-22	−0.0544	0.1009	1.3009	1.9726	2.0352	2.6426
2013-11-25	−0.0576	0.1053	0.4790	0.9200	1.3853	1.9190
2013-11-26	−0.0549	0.0979	0.8293	1.3449	0.4489	0.7929
2013-11-27	−0.0567	0.0905	0.5462	0.8268	1.2801	1.6097
2013-11-28	−0.0560	0.0824	0.5017	0.5344	0.5651	0.5152
2013-11-29	−0.0576	0.0751	0.3268	0.3614	0.7242	0.4159
mean	−0.0532	0.0965	0.6405	1.0048	0.9126	1.2090
std	0.0046	0.0250	0.3881	0.6266	0.4462	0.7053

Table A.2: Calibrations and validation results for the Merton model in Section 4.2.

	μ	σ	λ	χ	α	rmse futures	rmse options	validation futures	validation options
2013-11-01	−0.1888	0.0361	1.2450	0.4354	0.1165	0.7941	1.5641		
2013-11-04	−0.1008	0.0455	1.2673	0.3210	0.1106	0.2717	0.3583	0.7912	1.5474
2013-11-05	−0.0800	0.0805	1.2761	0.3127	0.0001	0.1438	0.1717	0.5420	1.0520
2013-11-06	−0.1455	0.0297	1.2121	0.3197	0.0276	0.2027	0.2809	0.2303	0.2132
2013-11-07	−0.0748	0.0289	1.4024	0.4365	0.0353	0.5013	1.6027	0.5737	1.6936
2013-11-08	−0.1931	0.0353	1.0150	0.1549	0.0354	0.9074	1.0384	0.5860	0.3810
2013-11-11	−0.2180	0.0448	0.9993	0.1625	0.0035	1.1867	2.0051	1.3203	2.2254
2013-11-12	−0.1786	0.0420	0.9972	0.1154	0.0124	0.3447	0.6118	0.9731	0.9076
2013-11-13	−0.1902	0.0443	1.0922	0.2348	0.0010	0.5220	0.8379	1.0354	1.0140
2013-11-14	−0.2108	0.0452	1.0891	0.2536	0.0085	0.7679	1.1783	0.9651	1.4218
2013-11-15	−0.1871	0.0616	1.1236	0.2628	0.0096	0.3477	0.7614	0.9410	1.2458
2013-11-18	−0.1839	0.0646	1.1243	0.2602	0.0130	0.3596	0.7966	0.3023	0.7267
2013-11-19	−0.0637	0.1189	1.2824	0.2904	0.0105	0.2888	0.1752	1.0610	0.3202
2013-11-20	−0.0510	0.0999	1.2352	0.2303	0.0242	0.3978	0.7682	0.6008	0.5799
2013-11-21	−0.2128	0.0412	1.0646	0.2273	0.0158	1.3625	1.8941	1.1304	1.7591
2013-11-22	−0.0507	0.0985	1.2169	0.2115	0.0253	1.3034	1.9693	1.9136	2.4333
2013-11-25	−0.0004	0.0878	1.3301	0.2494	0.0125	0.4776	0.9182	1.3878	1.9160
2013-11-26	0.0065	0.0769	1.3496	0.2614	0.0035	0.8228	1.3324	0.4477	0.7905
2013-11-27	0.0023	0.0690	1.2999	0.2183	0.0148	0.5491	0.8130	1.2748	1.5970
2013-11-28	−0.0226	0.0751	1.1743	0.1325	0.0005	0.3639	0.5247	0.5698	0.5067
2013-11-29	−0.1299	0.0366	1.2105	0.2920	0.0208	0.3201	0.3547	0.6993	0.4227
mean	−0.1178	0.0601	1.1908	0.2563	0.0239	0.5826	0.9503	0.8673	1.1377
std	0.0792	0.0260	0.1182	0.0831	0.0316	0.3598	0.5844	0.4145	0.6615

Table A.3: Calibrations and validation results for the Heston model in Section 4.3.

	μ	κ	θ	γ	ρ	V_0	rmse futures	rmse options	validation futures	validation options
2013-11-01	−0.0693	16.5730	0.0303	0.3484	−0.3517	0.0286	0.8719	1.9104		
2013-11-04	−0.0545	17.2495	0.0305	0.3473	−0.3195	0.0286	0.2589	0.5046	0.8451	1.8482
2013-11-05	−0.0461	17.2590	0.0315	0.3018	−0.3184	0.0286	0.1686	0.3141	0.5345	1.2247
2013-11-06	−0.0455	16.5892	0.0302	0.3399	−0.3562	0.0286	0.1356	0.4282	0.1955	0.3986
2013-11-07	−0.0478	18.0696	0.0296	0.3620	−0.2429	0.0286	0.5290	1.7760	0.6058	1.8645
2013-11-08	−0.0204	104.2848	0.0292	0.5935	−1.0000	0.0286	0.3716	0.8830	0.6131	0.5462
2013-11-11	−0.0165	118.3777	0.0294	0.6427	−1.0000	0.0286	0.6180	1.9230	0.7836	2.1154
2013-11-12	−0.0385	40.9774	0.0308	0.4177	−0.7618	0.0286	0.1139	0.4790	0.5031	0.8589
2013-11-13	−0.0307	40.7870	0.0306	0.4238	−0.9998	0.0286	0.2708	0.6588	0.8804	0.9221
2013-11-14	−0.0234	106.5053	0.0297	0.6233	−1.0000	0.0286	0.4511	1.0173	0.6874	1.2709
2013-11-15	−0.0378	46.2406	0.0304	0.4369	−1.0000	0.0286	0.1643	0.5504	0.5669	1.0823
2013-11-18	−0.0380	54.4207	0.0301	0.4588	−1.0000	0.0286	0.1743	0.5969	0.2152	0.5111
2013-11-19	−0.0556	18.0115	0.0315	0.3039	−0.3646	0.0286	0.2639	0.3482	1.1188	0.4159
2013-11-20	−0.0423	22.8584	0.0307	0.3201	−0.9992	0.0286	0.1595	0.5269	0.5775	0.4935
2013-11-21	−0.0213	45.9239	0.0305	0.4769	−1.0000	0.0286	0.8629	1.7135	0.8517	1.4774
2013-11-22	−0.0214	98.8979	0.0299	0.6147	−1.0000	0.0286	0.6671	1.5179	1.2540	2.2775
2013-11-25	−0.0359	70.4038	0.0301	0.5252	−1.0000	0.0286	0.1809	0.5631	0.7985	1.5139
2013-11-26	−0.0267	105.4928	0.0295	0.6119	−1.0000	0.0286	0.3563	0.9330	0.1788	0.4590
2013-11-27	−0.0335	81.3855	0.0298	0.5499	−1.0000	0.0286	0.1320	0.4165	0.7295	1.2026
2013-11-28	−0.0391	33.4659	0.0306	0.3875	−1.0000	0.0286	0.1002	0.2824	0.2566	0.2834
2013-11-29	−0.0543	16.0020	0.0274	0.3408	−0.3079	0.0286	0.2378	0.6433	0.7709	0.8174
mean	−0.0380	51.8941	0.0301	0.4489	−0.7629	0.0286	0.3375	0.8565	0.6483	1.0792
std	0.0138	36.3423	0.0009	0.1175	0.3237	0.0000	0.2405	0.5596	0.2903	0.6135

Table A.4: Calibrations and validation results for the forward curve term structure fitted model in Section 4.4.

	σ	rmse futures	rmse options	validation futures	validation options
2013-11-01	0.0000	0.0000	2.1565		
2013-11-04	0.0726	0.0000	0.6221	0.1973	0.5666
2013-11-05	0.0914	0.0000	0.3696	0.3434	0.2135
2013-11-06	0.0675	0.0000	0.5006	0.2300	0.4770
2013-11-07	0.0000	0.0000	1.8579	0.0946	1.9503
2013-11-08	0.2325	0.0000	0.7156	0.3400	1.0115
2013-11-11	0.5436	0.0000	1.2270	0.2282	1.8582
2013-11-12	0.1495	0.0000	0.0410	0.2581	3.9401
2013-11-13	0.1770	0.0000	0.1561	0.9989	0.5143
2013-11-14	0.2192	0.0000	0.3841	0.6479	0.7501
2013-11-15	0.1588	0.0000	0.0605	0.1960	0.6059
2013-11-18	0.1649	0.0000	0.1005	0.1311	0.0483
2013-11-19	0.0983	0.0000	0.4198	0.1580	0.6658
2013-11-20	0.1576	0.0000	0.0594	0.2835	0.3222
2013-11-21	0.2966	0.0000	0.8602	0.1716	1.4906
2013-11-22	0.2758	0.0000	0.8061	0.1091	0.8411
2013-11-25	0.1730	0.0000	0.1794	0.4414	0.7010
2013-11-26	0.2202	0.0000	0.4993	0.9316	0.1605
2013-11-27	0.1556	0.0000	0.1166	0.7897	0.8136
2013-11-28	0.1232	0.0000	0.1405	0.0195	0.2393
2013-11-29	0.0714	0.0000	0.7310	0.1631	0.8684
mean	0.1642	0.0000	0.5716	0.3366	0.9019
std	0.1177	0.0000	0.5761	0.2828	0.8824