WHOLESALE BANKING

Calibration of Stochastic Convenience Yield Models For Crude Oil Using the Kalman Filter

Master's Thesis

by

Adriaan Krul

born on the 19th of April 1985 in Basingstoke, England

June 25, 2008

Supervisors:
Prof. Dr. Ir. C.W. Oosterlee
Dr. Ir. G.J.M. Pieters

Other thesis committee members:
Prof. Dr. Ir. A.W. Heemink
Dr. Ir. J.A.M. van der Weide

Delft
Faculty of Electrical Engineering, Mathematics and Computer Science
Institute of Applied Mathematics

Acknowledgements

There are numerous persons which I would like to thank. In particular Gert-Jan Pieters, not for only serving as my supervisor but also encouraging and challenging me throughout the last 7 months of my study. He guided me through the dissertation process, never accepting less than my best efforts. I want to thank Kees Oosterlee for initially giving me the opportunity to intend an internship at the ING Bank and for introducing me to financial mathematics. From inside the ING, I would like to name some persons who helped me to understand the economical details of this subject, namely; Antoine van der Ploeg, Vincent Saarloos, Alain Verberkmoes, Luuk Derksen and Per Stahl. Furthermore, I want to thank the other committe members; Hans van der Weide and Arnold Heemink. Last but not least, great love goes out to my parents who supported me. In brief, without the help of these persons, this Master's Thesis would not be a success. Thank you all.

Abstract

Futures contracts depend on, when considering deterministic interest rates, the spot price S_t and the convenience yield δ_t. The former is assumed to follow a Geometric Brownian Motion (GBM) and the latter is usually calibrated via market data every two days, using the futures contracts. The market shows however that the convenience yield behaves stochastically and has a mean-reverting property. In the literature the convenience yield therefore is believed to follow an Ornstein-Uhlenbeck process driving by Brownian motion. Both processes are calculated by the Kalman filter. However, the disadvantage of this process is that the convenience yield can become negative which can result in cost of carry arbitrage possibilities. Therefore we introduce the Cox-Ingersoll-Ross process for the dynamics of the convenience yield. We find strong evidence for the adequacy of this model. Furthermore the extended Kalman filter is explained and tested for the Ornstein-Uhlenbeck process, as well as for pricing put options on these futures contracts.

Summary

In this report we examine the effect of including stochastic dynamics for the two state variables in futures contracts on the commodity, light crude oil. First of all we consider throughout the report deterministic interest rates. This makes calibration and the mathematics considerably easier. Furthermore, the effect of the interest rate turns out to be negligible. We assume a Geometric Brownian Motion (GBM) for the spot price process and an Ornstein-Uhlenbeck (OU) process for the convenience yield as empirical evidence exists for a mean-reverting pattern. Since both these processes are not observable in the market we use a Kalman filter to link these latent state variables with the observable variables, which in this case are the future prices. The Kalman filter turns out to be a very powerful, stable and accurate iterative procedure. The set consisting of the parameters of these stochastic processes is chosen such that the log-likelihood function is maximized and the innovations, which are the errors between the actual future prices and the filtered futures prices, are minimized. For ease of applying the Kalman filter to our model, it is important to have an affine form for the closed form solution of the future prices. In that way, the future prices are linear in the log of the state variable S_t since the requirement for the Kalman filter is linearity of these prices.

The formula for a put price on a futures contract is derived and since this is a nonlinear formula the Kalman filter will not give an optimal result. Consequently, we introduce the extend Kalman filter which deals with nonlinear problems and again, this is a very powerful and accurate method.

At the end of this report we consider a matured contract with different strikes. For a smal data set, it turns out that the parameters (in particular the mean-reverting term of the convenience yield process) will increase to unrealistic values. These values will distort the initial process and it can be that another process for the convenience yield is needed. We tested a Cox-Ingersoll-Ross (CIR) process, but because of the nonnegativity contraint of the convenience yield in this process, the CIR is inadequate. This proves that the period of observations has strong impact on which model to choose.

Calibration of Stochastic Convenience Yield Models For Crude Oil Using the Kalman Filter

Adriaan Krul

Amsterdam 2007-2008

Contents

Chapter 1

Introduction

1.1 Commodity markets, forwards, futures and convenience yield

In the history of commodity trading, buyers and sellers meet at marketplaces and trading of commodities led into immediate delivery. London was the 'main' capital for trading metal as Chicago was for cereals and agricultural products. Since more and more transaction were made, the need of building up central trading points was increasing. In 1842 the New York Cotton Exchange and in 1848 the Chicago Board Of Trade were set up as well as the New York Mercantile Exchange. At these boards, over 600 million contracts are traded every year. Examples of commodities are copper, wheat, cotton, pork belly and light crude oil.

There are two different contracts on commodities, forward and futures contracts. Forward (as well as future) contracts allow parties to buy or sell an asset at a prespecified price on a specific time. Forward contracts are private agreements and because of that, there is always a chance that a party may default on its side of the agreement. futures contracts on the other hand have clearing houses that guarentee the transactions. For forward contracts, settlement of the contract occurs at the end of the contract, but settlement for futures contracts can occur on different dates. Furthermore, futures contracts are so called mark-to-market, i.e. daily changes are settled day by day until the end of the period. There is also a difference in who trades in them. Futures contracts are used by speculators and are usually closed out prior to maturity and delivery usually never happens and forward contracts are used by hedgers. Delivery of the asset will usually take place.

The advantage of holding the physical commodity is that, for example, if there will be a crisis in the world and the demand for that particular commodity is rising, it can be sold with high profit. The convenience yield is the benefit that the owner of the physical commodity makes in this situation and does not accrue to the holder of a futures contract written on it. It is similar to dividend, which is paid to the owner of a stock but not to the holder of a contract written on the stock. Sometimes, due to irregular market movements such as an inverted market (this is when the short-term contract prices are higher than the long-term contracts), the holding of an underlying good or security may become more profitable than owning the contract or derivative instrument, due to its relative scarcity versus high demand. An example would be purchasing physical barrels of oil rather than futures contracts. Suppose there would be a sudden shock-situation wherein the demand for oil increases, the difference between the first purchase price of the oil versus the price after the shock would be the convenience yield.

Another property of futures contract is backwardation and contango. The former means if a far future delivery price is lower than a nearer future delivery price. The latter is the other way around. This implies that futures contracts have a term-structure. One can also have options on future contracts. For example consider a K-strike European call option on a futures contract, that is, the right to buy a futures contract with maturity T_1 for a price K at maturity T_2 of the option. The difference between the two is e.g. the payoff function (it could be negative for futures contracts but strictly positive for options).

1.2 Summary and outline of the thesis

Traders calibrate the convenience yield (CY), δ_t using the futures contract prices. The market shows that the CY behaves stochastically and has a mean-reverting property, see e.g. [5]. The CY may therefore be modelled by e.g. an Ornstein-Uhlenbeck (OU) process driven by a Brownian motion. The disadvantage of this process is that it allows for negative CY which can result in cost of carry arbitrage possibilities. Hence to prevent this from happening we also consider a model in which the CY follows a Cox-Ingersoll-Ross (CIR) process. For both processes, the spot price is assumed to follow a GBM.
For commodities, the spot prices and the CY can not be observed directly in the market. To calibrate the spot and CY model parameters for the stochastic processes OU and CIR, the non-observability of these *state variables* is a challenging difficulty. Since the future prices of commodities are widely observed and traded on the market we use a method that links these actual observations with the latent state variables: The *Kalman filter* (KF) technique. The main idea of the (KF) is to use observable (future) variables to reconstitute the value of the non-obsvervable spot prices and CY. For this purpose, we need a closed form solution for the futures prices. This solution can be easily found by explicitly using the affine structure of the CY model.
The aim of this thesis is to implement the KF for both OU and CIR and compare the results with the market data, i.e. we will price futures contracts on light crude oil and observe whether the price matches the price given in the market. Furthermore we use the results of the Kalman filtering of futures contracts to price a put option on these contracts. The outline of this report is as follows: Chapters 2 and 3 give analytic results, such as first and second other moments of the OU and CIR processes together with the closed form solution of the future prices. Chapter 4 applies the KF and extended Kalman filter (which is an extension of the KF for quasi-linear systems) to the OU process and discusses the numerical results. In Chapter 5 the KF is set up for the CIR process. In Chapter 6 we derive a closed form formula for the price of a put option on a futures contracts and compare it to the market data. We conclude with a discussion and further research in Chapter 7. In Appendix A, a detailed description of the KF is found together with an introductory example. In Appendix B we extend the OU model with a stochastic interest rate and in Appendix C the lengthy proofs of Appendix B are given.

Chapter 2

Stochastic convenience yield model following an Ornstein-Uhlenbeck process

In this chapter, analytic results are presented for the stochastic CY when it follows a time-homogeneous Ornstein-Uhlenbeck process driven by Brownian motion. The assumption is that the spot price S_t is described by a GBM, i.e.

$$dS_t = \mu S_t dt + \sigma_S S_t dW_t, \tag{2.1}$$

where μ is the drift term, σ_S the volatility term, and W_t a standard $\mathbb{P}$-Brownian motion. The form of the convenience yield δ_t is a result of the studies of Gibson and Schwartz (see [5]) where they find empirical evidence that the convenience yield may have a mean-reverting property, i.e.

$$d\delta_t = k(\alpha - \delta_t)dt + \sigma_\delta dZ_t. \tag{2.2}$$

In (2.2), α is the long range mean to which δ_t tends to revert, k the speed of reversion, σ_δ the constant volatility and Z_t a standard $\mathbb{P}-$Brownian motion. Here $dW_t dZ_t = \rho dt$. The solution of (2.1) is given by

$$S_t = S_0 \exp\left\{(\mu - \frac{1}{2}\sigma_S^2)t + \sigma_S \int_0^t dW_s\right\}, \tag{2.3}$$

and the solution of (2.2) is

$$\delta_t = \theta(0,t)\delta_0 + (1 - \theta(0,t))\alpha + \sigma_\delta e^{-kt} \int_0^t e^{ks} dZ_s, \tag{2.4}$$

with $\theta(t,T) = e^{-k(T-t)}$.

2.1 Mean and variance of the Ornstein-Uhlenbeck process

Multiplying (2.4) by e^{kt} gives

$$\delta_t e^{kt} = \delta_0 + k\alpha \int_0^t e^{ks} ds + \sigma_\delta \int_0^t e^{ks} dZ_s.$$

Rearranging and taking expectation yields

$$\mathbb{E}[\delta_t] = \delta_0 e^{-kt} + \alpha(1 - e^{-kt}) = \delta_0\theta(0,t) + \alpha(1 - \theta(0,t)).$$

For the variance we calculate the covariance matrix by applying the Itô isometry, which is defined as follows, [32],

Lemma 2.1. *(The Itô isometry)*
If $\Lambda(t,w)$ is bounded, then

$$\mathbb{E}\left[\left(\int_s^T \Lambda(t,w)dB_t(w)\right)^2\right] = \mathbb{E}\left[\int_s^T \Lambda(t,w)^2 dt\right].$$

In particular, for $\Lambda(t,w) = \Lambda(t)$ we have

$$\mathbb{E}\left[\left(\int_s^T \Lambda(t)dB_t\right)^2\right] = \int_s^T \Lambda(t)^2 dt.$$

Using Lemma 2.1 we have,

$$\begin{aligned}
\mathrm{Cov}(\delta_s,\delta_t) &= \mathbb{E}[(\delta_s - \mathbb{E}[\delta_s])(\delta_t - \mathbb{E}[\delta_t])] \\
&= \mathbb{E}\left[\sigma_\delta \int_0^s e^{-k(u-s)} dZ_u \sigma_\delta \int_0^t e^{-k(v-t)} dZ_v\right] \\
&= \sigma_\delta^2 e^{-k(s+t)} \mathbb{E}\left[\int_0^s e^{ku} dZ_u \int_0^t e^{kv} dZ_v\right] \\
&= \frac{\sigma_\delta^2}{2k} e^{-k(s+t)} (e^{2k(\min(s,t))} - 1).
\end{aligned}$$

For $s = t$ this gives $\mathbb{V}\mathrm{ar}[\delta_t] = \frac{\sigma_\delta^2}{2k}\left(1-\theta(0,t)^2\right)$. We conclude that $\delta_t \sim \mathscr{N}\left(\theta(0,t)\delta_0 + (1-\theta(0,t))\alpha, \frac{\sigma_\delta^2}{2k}(1-\theta(0,t)^2)\right)$.

2.2 From real world to risk neutral world

The unobservable state variable S_t is not a tradable asset since it generates "cash flows" by means of the convenience yield δ_t. A standard strategy would be to reinvest the convenience yield in S_t, i.e. we define the price process for asset $Y_t := e^{\int_0^t \delta_u du} S_t$ made from instantaneously reinvesting the convenience yield δ_t. The asset Y_t is tradable if and only if S_t is a tradable asset. As noticed earlier, this is for the commodity crude oil not the case. Due to this inconvenience Y_t is not tradable itself. However, we assume that S_t is a tradable asset because this will not effect the calibrations throughout this report and thus we continue with this standard approach. Next step is to choose a suitable numeraire B_t for which the relative price process $\frac{Y_t}{B_t}$ is a $\mathbb{Q}$-martingale. The requirement for B_t is that it has to be a tradable asset as well. Since we do not incorporate stochastic interest rates, we may choose the cash bond as numeraire, i.e.

$$dB_t = rB_t dt, \quad \text{or} \quad B_t = e^{\int_0^t r du}. \tag{2.5}$$

The stochastic process for $\frac{Y_t}{B_t}$ is then found by applying Itô's lemma and using (2.1)

$$\begin{aligned}
d\left(\frac{Y_t}{B_t}\right) &= d\left(e^{\int_0^t (\delta_u - r)du} S_t\right) \\
&= (\mu + \delta_t - r)\left(\frac{Y_t}{B_t}\right) dt + \sigma_S \left(\frac{Y_t}{B_t}\right) dW_t.
\end{aligned} \tag{2.6}$$

This price process is a martingal under the $\mathbb{Q}$-measure (risk neutral measure). However, W_t is a $\mathbb{P}$-Brownian motion and therefore it has to be transformed into a $\mathbb{Q}$-Brownian motion. Girsanov's theorem states that there exists a λ_t such that

$$\tilde{W}_t = W_t + \int_0^t \lambda_u du, \tag{2.7}$$

where $\tilde{W}_t$ denotes a $\mathbb{Q}$-Brownian motion. Substitution of (2.7) in (2.6) yields

$$d\left(\frac{Y_t}{B_t}\right) = (\mu + \delta_t - r - \sigma_S \lambda_t)\left(\frac{Y_t}{B_t}\right) dt + \sigma_S \left(\frac{Y_t}{B_t}\right) d\tilde{W}_t. \tag{2.8}$$

This process under the $\mathbb{Q}$-measure is a martingale. Therefore $\mu + \delta_t - r - \sigma_S \lambda_t = 0$ and hence the market price of risk λ_t is given by

$$\lambda_t = \frac{\mu + \delta_t - r}{\sigma_S}. \tag{2.9}$$

This result combined with (2.7) gives for (2.1)

$$\frac{dS_t}{S_t} = (r - \delta_t)dt + \sigma_S d\tilde{W}_t. \tag{2.10}$$

One wants to have a similar result for Z_t, which is again a $\mathbb{P}$-Brownian motion. However, δ_t is not tradable and a construction as outlined above is not possible. We can, of course, still apply Girsanov's theorem, but the market price of risk, λ (now taken as a constant), appears as an extra parameter in the convenience yield process, i.e.

$$d\delta_t = k\left(\alpha - \delta_t - \frac{\sigma_\delta \lambda}{k}\right) dt + \sigma_\delta d\tilde{Z}_t. \tag{2.11}$$

2.3 The value of a future delivery

Define $Y_T = Y(T, S_T, \delta_T)$ to be the contingent claim at time T. The discounted expected value of Y_T at time t is given by, (see Appendix B for details),

$$V_t[Y_T] := B_t \mathbb{E}^{\mathbb{Q}}\left[\frac{Y_T}{B_T}|\mathscr{F}_t\right] = e^{-r(T-t)}\mathbb{E}^{\mathbb{Q}}[Y_T|\mathscr{F}_t], \tag{2.12}$$

where B_t is given by (2.5) and $\mathscr{F}_t$ is the filtration. We consider the case $Y_T := S_T$ and we compute $V_t[S_T]$. The current value of a claim on a future delivery of the commodity on the future date t is

$$\begin{aligned} V_t[S_T] \quad = \quad & S_t \exp\Bigg\{\left[-\alpha + \frac{1}{k}\left(\sigma_\delta\lambda - \sigma_\delta\sigma_S\rho\right) + \frac{1}{2}\left(\frac{1}{k}\right)^2 \sigma_\delta^2\right](T-t) \\ & - \frac{1}{k}\left[\delta_t - \alpha + \frac{1}{k}\left(\sigma_\delta\lambda - \sigma_\delta\sigma_S\rho\right) + \left(\frac{1}{k}\right)^2 \sigma_\delta^2\right](1 - \theta(t,T)) \\ & + \frac{1}{2}\left(\frac{1}{k}\right)^2 \frac{\sigma_\delta^2}{2k}\left(1 - \theta(t,T)^2\right)\Bigg\}. \end{aligned} \tag{2.13}$$

Proof. From (2.3) we have

$$S_T = S_t e^{\int_t^T (r-\delta_s)ds - \frac{1}{2}\sigma_S^2(T-t) + \int_t^T \sigma_s d\tilde{W}_u}. \tag{2.14}$$

Now define

$$\int_t^T \delta_S ds = X_T - X_t, \quad X_t = \int_0^t \delta_s ds. \tag{2.15}$$

Integrating equation (2.11) with respect to t yields

$$\begin{aligned} \delta_T - \delta_t \quad &= \int_t^T k(\alpha - \delta_u - \frac{\sigma_\delta\lambda}{k})dt + \sigma_\delta \int_t^T d\tilde{Z}_u \\ &= k(\alpha - \frac{\sigma_\delta\lambda}{k})(T-t) - k\int_t^T \delta_u du + \sigma_\delta \int_t^T d\tilde{Z}_u \\ &= k(\alpha - \frac{\sigma_\delta\lambda}{k})(T-t) - k(X_T - X_t) + \sigma_\delta \int_t^T d\tilde{Z}_u. \end{aligned} \tag{2.16}$$

Substituting

$$\delta_T = \theta(t,T)\delta_t + (1-\theta(t,T))(\alpha - \frac{\sigma_\delta \lambda}{k}) + \sigma_\delta e^{-kT} \int_t^T e^{ku} d\tilde{Z}_u, \tag{2.17}$$

which follows from (2.4), into (2.16) and rearranging terms gives

$$X_T - X_t = \frac{1}{k}(1-\theta(t,T))\delta_t + \frac{1}{k}(1-\theta(t,T))(\frac{\sigma_\delta \lambda}{k} - \alpha) + (\alpha - \frac{\sigma_\delta \lambda}{k})(T-t) - \frac{\sigma_\delta}{k} e^{-kT} \int_t^T e^{ku} d\tilde{Z}_u + \frac{\sigma_\delta}{k} \int_t^T d\tilde{Z}_u. \tag{2.18}$$

Equation (2.15) together with the result given in (2.18) will give for (2.14)

$$\begin{aligned} S_T &= S_t \exp\Big\{ r(T-t) - \frac{1}{k}(1-\theta(t,T))(\delta_t - \alpha + \frac{\sigma_\delta \lambda}{k}) - (\alpha - \frac{\sigma_\delta \lambda}{k})(T-t) + \frac{\sigma_\delta}{k} e^{-kT} \int_t^T e^{ku} d\tilde{Z}_u \\ &\quad - \frac{\sigma_\delta}{k} \int_t^T d\tilde{Z}_u - \frac{1}{2}\sigma_S^2 (T-t) + \sigma_S \int_t^T d\tilde{W}_u \Big\} \\ &\equiv S_t \exp\{r(T-t) + z\} \end{aligned} \tag{2.19}$$

From basic stochastic calculus it follows

$$\begin{aligned} \tilde{\mu} &= \mathbb{E}[z] \\ &= -(\frac{1}{2}\sigma_S^2 + \alpha - \frac{1}{k}\sigma_\delta \lambda)(T-t) + \frac{1}{k}(\alpha - \delta_t - \frac{1}{k}\sigma_\delta \lambda)(1-\theta(t,T)). \end{aligned} \tag{2.20}$$

Using Lemma 2.1 we can calculate the following expectations

$$\mathbb{E}\Big[\sigma_S^2 \Big(\int_t^T d\tilde{W}_s \int_t^T d\tilde{W}_s\Big)\Big] = \sigma_S^2 \int_t^T dt = \sigma_S^2 (T-t), \tag{2.21}$$

$$\mathbb{E}\Big[\Big(\frac{1}{k}\sigma_\delta \int_t^T d\tilde{Z}_s\Big)^2\Big] = \frac{\sigma_\delta^2}{k^2} \int_t^T dt = \frac{\sigma_\delta^2}{k^2}(T-t), \tag{2.22}$$

$$\begin{aligned} \mathbb{E}\Big[\Big(\frac{1}{k} e^{-kT} \int_t^T e^{ks} d\tilde{Z}_s\Big)^2\Big] &= \Big[\frac{\sigma_\delta^2}{k^2} e^{-2kT} \int_t^T e^{2ks} ds\Big] \\ &= \frac{\sigma_\delta^2}{k^2} e^{-2kT} \Big[\frac{1}{2k} e^{2kT} - \frac{1}{2k} e^{2kt}\Big] \\ &= \frac{\sigma_\delta^2}{k^2} \frac{1}{2k} [1 - \theta(t,T)^2], \end{aligned} \tag{2.23}$$

$$\mathbb{E}\Big[\Big(\sigma_S \int_t^T d\tilde{W}_s\Big)\Big(\frac{1}{k}\sigma_\delta \int_t^T d\tilde{Z}_s\Big)\Big] = \sigma_S \frac{\sigma_\delta}{k} \rho (T-t), \tag{2.24}$$

$$\begin{aligned} \mathbb{E}\Big[\Big(\sigma_S \int_t^T d\tilde{W}_s\Big)\Big(\frac{1}{k}\sigma_\delta e^{-kT} \int_t^T e^{ks} d\tilde{Z}_s\Big)\Big] &= e^{-kT} \frac{\sigma_S \sigma_\delta}{k} \rho \int_t^T e^{ks} ds \\ &= e^{-kT} \frac{\sigma_S \sigma_\delta \rho}{k} \frac{1}{k} [e^{kT} - e^{kt}] \\ &= \frac{\sigma_S \sigma_\delta \rho}{k^2} (1 - \theta(t,T)), \end{aligned} \tag{2.25}$$

$$\begin{aligned} \mathbb{E}\Big[\Big(\frac{1}{k}\sigma_\delta \int_t^T d\tilde{Z}_s\Big)\Big(\frac{1}{k}\sigma_\delta e^{-kT} \int_t^T e^{ks} d\tilde{Z}_s\Big)\Big] &= \frac{1}{k^2}\sigma_\delta^2 e^{-kT} \rho \int_t^T e^{ks} ds \\ &= \frac{1}{k^2}(\sigma_\delta^2 e^{-kT})\rho \frac{1}{k}(e^{kT} - e^{kt}) \\ &= \frac{\sigma_\delta^2 \rho}{k^3}(1 - \theta(t,T)). \end{aligned} \tag{2.26}$$

Finally we get

$$
\begin{aligned}
\tilde{\sigma}^2 &= \mathbb{V}\mathrm{ar}(z) \\
&= \mathbb{E}[(z^2)] - (\mathbb{E}[z])^2 \\
&= (\sigma_S^2 - 2\frac{1}{k}\sigma_\delta\sigma_S\rho + \frac{1}{k^2}\sigma_\delta^2)(T-t) \\
&+ 2(\frac{1}{k^2}\sigma_\delta\sigma_S\rho - \frac{1}{k^3}\sigma_\delta^2)(1-\theta(t,T)) \\
&+ \frac{1}{k^2}\frac{\sigma_\delta^2}{2k}(1-\theta(t,T)^2).
\end{aligned} \tag{2.27}
$$

From the standard formula of the expected value of a lognormal random variable we get

$$
V_t[S_T] = S_t \exp\left\{\tilde{\mu} + \frac{1}{2}\tilde{\sigma}^2\right\}. \tag{2.28}
$$

□

From the absence of risk-free arbitrage opportunities it follows

$$
V_t[S_T - F] = 0. \tag{2.29}
$$

The value F for which (2.29) holds, is called the future price and is

$$
F = F(S,\delta,\tau) := e^{r\tau} V_t[S_T], \tag{2.30}
$$

with $\tau = T - t$ denoting the time to maturity.
In order to find the PDE, which $F(S,\delta,\tau)$ satisfies we use the Feynman- Kac theorem, [1].

Theorem 2.1. *(Feynman-Kac)*
Suppose the underlying processes $y_1(t), y_2(t), .., y_n(t)$ follow the stochastic differential equation

$$
dy_i = \mu_i(y_1, y_2, ..y_n, t)dt + \sigma_i(y_1, y_2, .., y_n, t)dW_i.
$$

Then the function

$$
g(y_1, y_2, .., y_n, t) = \mathbb{E}_{y_1,y_2,..,y_n,t}[f(y_1(T), .., y_n(T))]
$$

is given by the solution of the partial differential equation

$$
\frac{\partial g}{\partial t} + \sum_{i=1}^{n} \mu_i \frac{\partial g}{\partial y_i} + \frac{1}{2}\sum_{i,j=1}^{n} \rho_{ij}\sigma_i\sigma_j \frac{\partial^2 g}{\partial y_i \partial y_j} = 0,
$$

subject to

$$
g(y_1, y_2, .., y_n, T) = f(y_1, y_2, .., y_n),
$$

where $\rho_{ij} = Cov(dW_i, dW_j)/dt$.

$F(S,\delta,\tau)$ satisfies the PDE

$$
\begin{cases}
\frac{1}{2}\sigma_S^2 S^2 F_{SS} + \rho\sigma_S\sigma_\delta S F_{S\delta} + \frac{1}{2}\sigma_\delta^2 F_{\delta\delta} + (r-\delta)SF_S + (k(\alpha-\delta) - \lambda\sigma_\delta)F_\delta - F_\tau = 0, \\
F(S,\delta,0) = S_0.
\end{cases} \tag{2.31}
$$

A similar result exists for $V(S,\delta,t) := V_t[Y_T]$. The $V(S,\delta,t)$ satisfies the PDE

$$\begin{cases} \frac{1}{2}\sigma_S^2 S^2 V_{SS} + \rho\sigma_S\sigma_\delta S V_{S\delta} + \frac{1}{2}\sigma_\delta^2 V_{\delta\delta} + (r-\delta)SV_S + (k(\alpha-\delta)-\lambda\sigma_\delta)V_\delta + V_t - rV = 0, \\ V(S,\delta,T) = Y_T(S,\delta). \end{cases} \tag{2.32}$$

Assume that the solution of (2.32) can be written in the form

$$\begin{cases} F(S,\delta,\tau) = S\exp\{A(\tau)+B(\tau)\delta\}, \\ A(0)=0, \quad B(0)=0. \end{cases} \tag{2.33}$$

Calculating the derivatives and inserting these into (2.31) gives

$$-F\{A'(\tau)+B'(\tau)\delta\} + (r-\delta)F + (k(\alpha-\delta)-\lambda\sigma_S)FB(\tau) + \rho\sigma_S\sigma_\delta FB(\tau) + \frac{1}{2}\sigma_\delta^2 B^2(\tau) = 0, \tag{2.34}$$

or equivalently

$$-A'(\tau) - B'(\tau)\delta + (r-\delta) + \rho\sigma_S\sigma_\delta B(\tau) + (k(\alpha-\delta)-\lambda\sigma_\delta)B(\tau) + \frac{1}{2}\sigma_\delta^2 B^2(\tau) = 0. \tag{2.35}$$

In other words, we have to solve the two differential equations

$$\begin{cases} -B'(\tau) - 1 - kB'(\tau) = 0, \\ -A'(\tau) + r + \rho\sigma_S\sigma_\delta B(\tau) + (k\alpha - \lambda\sigma_\delta)B(\tau) + \frac{1}{2}\sigma_\delta^2 B^2(\tau) = 0. \end{cases} \tag{2.36}$$

Solving (2.36) yields

$$B(\tau) = \frac{1}{k}(e^{-k\tau}-1) = \frac{1}{k}(\theta(\tau)-1); \quad \theta(\tau) = e^{-k\tau}, \tag{2.37}$$

and thus, for $(2.36)_2$,

$$A'(\tau) = r + \frac{\rho\sigma_S\sigma_\delta}{k}(\theta(\tau)-1) + (\alpha - \frac{\lambda\sigma_\delta}{k})(\theta(\tau)-1) + \frac{\sigma_\delta^2}{2k^2}(\theta(\tau)-1)^2. \tag{2.38}$$

Integrating (2.38) in time, and using the fact that $A(0)=0$, gives

$$A(\tau) = \left[\left(r - \tilde{\alpha} + \frac{\sigma_\delta^2}{2k^2} - \frac{\sigma_S\sigma_\delta\rho}{k}\right)\tau\right] + \left[\frac{\sigma_\delta^2}{4}\frac{1-e^{-2k\tau}}{k^3}\right] + \left[\left(\tilde{\alpha}k + \sigma_\delta\sigma_S\rho - \frac{\sigma_\delta^2}{k}\right)\left(\frac{1-e^{-k\tau}}{k^2}\right)\right],$$

with

$$\tilde{\alpha} = \alpha - \frac{\lambda}{k}.$$

2.4 Conclusion

In this chapter we derived analytic results when assuming an OU process for the convenience yield. By means of the mean and variance of S_t and δ_t we are able to price the current value of a claim on a future delivery of the commodity on the future date t. Also, the market price of risk is explained via the Girsanov's theorem. An important remark, as we will notice later on, is the assumption of an affine form for the closed form solution of the future prices. By assuming this form, we can easily solve the PDE given in (2.32).

Chapter 3

Stochastic convenience yield model following a Cox-Ingersoll-Ross process

In this chapter we assume that the convenience yield follows a time-homogeneous Cox-Ingersoll-Ross (CIR) process and the spot price a GBM in which the volatility is proportional to the square root of the instantaneous convenience yield level, i.e.

$$\begin{cases} \dfrac{dS_t}{S_t} = \mu dt + \sigma_S \sqrt{\delta_t} dW_t, \\ d\delta_t = k(\alpha - \delta_t)dt + \sigma_\delta \sqrt{\delta_t} dZ_t. \end{cases} \tag{3.1}$$

In (3.1), S_t is the price of the underlying, δ_t is the CY, μ is the drift term, σ_S is the volatility term of $\frac{dS_t}{S_t}$, W_t a standard $\mathbb{P}$-Brownian motion, α is the long range mean to which δ_t tends to revert, k the speed of adjustment, σ_δ the volatility term of $d\delta_t$, and Z_t a standard $\mathbb{P}$-Brownian motion. Here $dW_t dZ_t = \rho dt$. The reason why we assume that the CY follows a CIR process is the nonnegativity. The CIR process excludes negative CY. A negative CY would make the forward prices go up at more than the interest rate and provide some kind of cost of carry arbitrage through buying the spot commodity and selling a forward.

3.1 Mean and variance of the CIR process

The solution of SDE $(3.1)_2$ is given by

$$\delta_t = \delta_0 + k \int_0^t (\alpha - \delta_u) du + \sigma_\delta \int_0^t \sqrt{\delta_u} dZ_u.$$

Multiplying by e^{kt}, taking expectations on both sides and differentiating with respect to time yields

$$\frac{d}{dt}\mathbb{E}[\delta_t] = k(\alpha - \mathbb{E}[\delta_t]) \quad \Rightarrow \quad \frac{d}{dt} e^{kt}\mathbb{E}[\delta_t] = e^{kt}[k\mathbb{E}[\delta_t] + \frac{d}{dt}\mathbb{E}[\delta_t]] = e^{kt} k\alpha.$$

This leads to

$$e^{kt}\mathbb{E}[\delta_t] - \delta_0 = k\alpha \int_0^t e^{ku} du = \alpha(e^{kt} - 1) \quad \Rightarrow \quad \mathbb{E}[\delta_t] = \alpha + e^{-kt}(\delta_0 - \alpha) = e^{-kt}\delta_0 + (1 - e^{-kt})\alpha. \tag{3.2}$$

Remark. We see that if $\delta_0 = \alpha$ then $\mathbb{E}[\delta_t] = \alpha, \ \forall t$. If $\delta_0 \neq \alpha$, then δ_t exhibits mean reversion, i.e.

$$\lim_{t\to\infty}\mathbb{E}[\delta_t] = \alpha.$$

For the variance we calculate first $d\delta_t^2$ via Itô's lemma. Define $f(x) = x^2$. We have

$$\begin{aligned} d\delta_t^2 &= df(\delta_t) \\ &= f'(\delta_t)d\delta_t + \frac{1}{2}f''(\delta_t)d\delta_t d\delta_t \\ &= 2\delta_t[k(\alpha-\delta_t)dt + \sigma_\delta\sqrt{\delta_t}dZ_t] + [k(\alpha-\delta_t)dt + \sigma_\delta\sqrt{\delta_t}dZ_t]^2 \\ &= 2\alpha k\delta_t dt - 2k\delta_t^2 dt + 2\sigma_\delta\delta_t^{3/2}dZ_t + \sigma_\delta^2\delta_t dt \\ &= (2k\alpha+\sigma_\delta^2)\delta_t dt - 2k\delta_t^2 dt + 2\sigma_\delta\delta_t^{3/2}dZ_t. \end{aligned}$$

This leads to

$$\delta_t^2 = \delta_0^2 + (2k\alpha+\sigma_\delta^2)\int_0^t \delta_u du - 2k\int_0^t \delta_u^2 du + 2\sigma_\delta\int_0^t \delta_u^{3/2}dZ_u.$$

Taking the expectation on both sides and differentiating with respect to t yields

$$\frac{d}{dt}\mathbb{E}[\delta_t^2] = (2k\alpha+\sigma_\delta^2)\mathbb{E}[\delta_t] - 2k\mathbb{E}[\delta_t^2].$$

From this we get

$$\begin{aligned} \frac{d}{dt}e^{2kt}\mathbb{E}[\delta_t^2] &= e^{2kt}[2k\mathbb{E}[\delta_t^2] + \frac{d}{dt}\mathbb{E}[\delta_t^2]] \\ &= e^{2kt}(2k\alpha+\sigma_\delta^2)\mathbb{E}[\delta_t]. \end{aligned} \tag{3.3}$$

Using (3.2) in (3.3) and integrating the equation gives

$$\mathbb{E}[\delta_t^2] = \frac{\alpha\sigma_\delta^2}{2k} + \alpha^2 + (\delta_0-\alpha)(\frac{\sigma_\delta^2}{k} + 2\alpha)e^{-kt} + (\delta_0-\alpha)^2\frac{\sigma_\delta^2}{k}e^{-2kt} + \frac{\sigma_\delta^2}{k}(\frac{\alpha}{2}\delta_0)e^{-2kt}.$$

We finally obtain

$$\begin{aligned} \mathbb{V}\mathrm{ar}[\delta_t] &= \mathbb{E}[\delta_t^2] - (\mathbb{E}[\delta_t])^2 \\ &= \alpha\frac{\sigma_\delta^2}{2k}(1-e^{-kt})^2 + \delta_0\frac{\sigma_\delta^2}{k}(e^{-kt}-e^{-2kt}). \end{aligned} \tag{3.4}$$

3.2 The logprice and inserting the market price of risk

Using similar arguments as explained in Section 2.2, we obtain the price processes under the $\mathbb{Q}$ measure

$$\begin{cases} \dfrac{dS_t}{S_t} = (r-\delta_t)dt + \sigma_S\sqrt{\delta_t}d\tilde{W}_t, \\ d\delta_t = (k(\alpha-\delta_t) - \lambda_{\delta_t})dt + \sigma_\delta\sqrt{\delta_t}d\tilde{Z}_t, \end{cases} \tag{3.5}$$

Remark: Note that λ_{δ_t} is a function of σ_δ and δ_t, however, throughout this report, λ_{δ_t} is considered constant and we shall denote it by λ.

The logprice process for $(3.5)_1$ is given by

$$dx_t = \left(r - \left(1+\frac{1}{2}\sigma_S^2\right)\delta_t\right)dt + \sigma_S\sqrt{\delta_t}d\tilde{W}_t, \tag{3.6}$$

with $x_t = \log S_t$.

3.3 Partial differential equation for the value of future delivery

Following the approach discussed in Section 2.5, we have the following PDE for the future price $F(S,\delta,\tau)$

$$\begin{cases} \frac{1}{2}\sigma_S^2\delta S^2 F_{SS} + \rho\sigma_S\sigma_\delta\delta S F_{S\delta} + \frac{1}{2}\sigma_\delta^2\delta F_{\delta\delta} + (r-\delta)SF_S + (k(\alpha-\delta)-\lambda)F_\delta - F_\tau = 0, \\ F(S,\delta,0) = S_0. \end{cases} \tag{3.7}$$

We assume again (see e.g. [2]) that this PDE admits the affine form solution

$$\begin{cases} F(S,\delta,\tau) = Se^{A(\tau)-B(\tau)\delta}, \\ A(0) = 0, \quad B(0) = 0. \end{cases} \tag{3.8}$$

To find $B(\tau)$ and $A(\tau)$ we substitute (3.8) into (3.7) to obtain the two equations

$$\begin{aligned} \frac{1}{2}\sigma_\delta^2 B^2 + (k-\rho\sigma_S\sigma_\delta)B - 1 + B'(\tau) &= 0, \\ r + (\lambda - k\alpha)B - A'(\tau) &= 0. \end{aligned}$$

For simplicity first write $a_1 = \frac{1}{2}\sigma_\delta^2$ and $a_2 = k - \rho\sigma_S\sigma_\delta$ which leads to $B'(\tau) = B_\tau = -a_1B^2 - a_2B + 1 =: f(B)$. Now we want to factorize the function f.

$$\frac{f(B)}{-a_1} = B^2 + \frac{a_2}{a_1}B - \frac{1}{a_1} = (B+\epsilon)(B-\gamma).$$

From this we get that $(\epsilon - \gamma) = \frac{a_2}{a_1}$ and $\epsilon\gamma = \frac{1}{a_1}$. From these two relations it follows that we have to solve

$$a_1\gamma^2 + a_2\gamma - 1 = 0.$$

This gives

$$\begin{aligned} \gamma_{1,2} &= -\frac{a_2}{2a_1} \pm \frac{\sqrt{a_2^2 + 4a_1}}{2a_1} \\ &= \frac{-k_2 \pm k_1}{2a_1}, \end{aligned} \tag{3.9}$$

where for simplicity k_1 and k_2 are given by

$$\begin{aligned} k_1 &= \sqrt{k_2^2 + 2\sigma_\delta^2}, \\ k_2 &= (k - \rho\sigma_S\sigma_\delta). \end{aligned}$$

To solve the ODE for $B(\tau)$ we write

$$\frac{1}{-a_1^2B^2 - a_2B + 1}B'(\tau) = 1. \tag{3.10}$$

Using the factorization we can rewrite this in

$$B(\tau) = \frac{(\tilde{c}\epsilon e^{-\frac{\gamma+\epsilon}{a_1}\tau} + \gamma)}{1 - \tilde{c}e^{-\frac{\gamma+\epsilon}{a_1}\tau}},$$

and with initial condition $B(0) = 0$ we get that $\tilde{c} = -\frac{\gamma}{\epsilon}$. Substituting $\tilde{c}$ and calculating for γ_1 (which refers to the plus sign) gives

$$\begin{aligned} B(\tau) &= \frac{\gamma(1 - e^{-\frac{\gamma+\epsilon}{a_1}\tau})}{1 + \frac{\gamma}{\epsilon}e^{-\frac{\gamma+\epsilon}{a_1}\tau}} \\ &= \frac{2(1 - e^{-k_1\tau})}{k_1 + k_2 + (k_1 - k_2)e^{-k_1\tau}}. \end{aligned} \tag{3.11}$$

The same result is found when substituting γ_2 (which refers to the minus sign). Accordingly,

$$A(\tau) = r\tau + (\lambda - k\alpha)\int_t^T B_q dq, \tag{3.12}$$

where

$$\int_t^T B_q dq = \frac{2}{k_1(k_1+k_2)}\ln\left[\frac{(k_1+k_2)e^{k_1\tau}+k_1-k_2}{2k_1}\right] \tag{3.13}$$

$$+ \frac{2}{k_1(k_1-k_2)}\ln\left[\frac{k_1+k_2+(k_1-k_2)e^{-k_1\tau}}{2k_1}\right]. \tag{3.14}$$

3.4 Conclusion

In this chapter derived analytic results when assuming a CIR process for the convenience yield. The reason for this, is the nonnegativity constraint one needs to impose for the convenience yield to prevent some sort of 'cost of carry' arbitrage possibilities. Again, by assuming an affine form of the closed form solution of the future prices we can easily solve the PDE given in (3.7).

Chapter 4

The Kalman filter for the Ornstein-Uhlenbeck process

In 1960 R.E. Kalman introduced the Kalman filter (KF), [34]. This algorithm makes optimal use of imprecise data on a (quasi-) linear system with Gaussian errors (white noise) to continuously update the best estimate of the system current state. The power of KF is; to compute these updates it is only necessary to consider the estimates from the previous time step and the new measurement and not all the previous data. The main idea is that we want to estimate the current state and its uncertainty, but we can not directly observe these states. Instead we observe noisy measurements. Since the Kalman filter deals with uncertain (in our case unobservable) variables, it seems to be the best iterative procedure for calibrating future prices.

To calibrate the state-space model, the state variables (which in this case are the spot prices and the convenience yield) are placed in a state vector. The measurement equation, consisting of this vector and uncorrelated disturbances to account for possible errors in the data, links actual observations (in this case the future prices on several different maturities) with latent variables. These latent variables are assumed to be first-order Markovian processes and are related to systems (2.1) and (2.2). The Kalman filter will give an optimal prediction for the unobserved data by only considering the previously estimated value. For a detailed description and an introductory example, see Appendix A.

4.1 Setup of the Kalman filter applied to the OU process

Consider again

$$F(S,\delta,\tau) = Se^{A(\tau)-B(\tau)\delta}, \tag{4.1}$$

where

$$A(\tau) = \left[\left(r - \tilde{\alpha} + \frac{\sigma_\delta^2}{2k^2} - \frac{\sigma_S\sigma_\delta\rho}{k}\right)\tau\right] + \left[\frac{\sigma_\delta^2}{4}\frac{1-e^{-2k\tau}}{k^3}\right] + \left[\left(\tilde{\alpha}k + \sigma_\delta\sigma_S\rho - \frac{\sigma_\delta^2}{k}\right)\left(\frac{1-e^{-k\tau}}{k^2}\right)\right],$$

and

$$B(\tau) = \frac{1-e^{-k\tau}}{k}, \quad \tilde{\alpha} = \alpha - \frac{\lambda}{k}.$$

In state variable terms, (4.1) can be rewritten in

$$\ln F_t(\tau) = x_t + A(\tau) - B(\tau)\delta_t. \tag{4.2}$$

The measurement equation then reads

$$\mathbf{Y}_{t_i} = \mathbf{d}_{t_i} + \mathbf{Z}_{t_i}\boldsymbol{\alpha}_{t_i} + \boldsymbol{\epsilon}_{t_i}, \quad i = 1, .., N, \tag{4.3}$$

where N denotes the number of observations and

- $\boldsymbol{\alpha}_{t_i} = [x_{t_i} \;\; \delta_{t_i}]'$,
- $\mathbf{Y}_{t_i} = [\ln \mathbf{F}_{t_i}(\tau_j)]$, for $j = 1, ..., n$ is a $n \times 1$ vector for n maturities. τ_j are the maturity dates. $\mathbf{F}_{t_i}(\tau_j)$ are observed from market data,
- $\mathbf{d}_{t_i} = [A(\tau_j)]$ for $i = j, .., n$ is a $n \times 1$ vector,
- $\mathbf{Z}_{t_i} = [1, -B(\tau_j)]$, for $j = 1, ..., n$ is a $n \times 2$ matrix.
- $\boldsymbol{\epsilon}_{t_i}$ is a $n \times 1$ vector of uncorrelated disturbances and is assumed to be normal with zero mean and variance matrix $\mathbf{H}_{t_i}$. $\mathbf{H}_{t_i}$ is a n $\times$ n diagonal matrix with h_j on its diagonal.

The $\boldsymbol{\epsilon}_t$ in (4.3) term is included to account for possible errors in the measurement. These errors especially occur when the state variables are unobservable. To get a feeling for the size of the error suppose that the OU model generates the prices and yields perfectly, and that the state variables can be observed from the market directly. The included error term in the measurement equation can be seen as the uncertainty in bid-ask spreads, errors in the data etc. The error is assumed to be small in comparison to the variation of the yield. The matrix $\mathbf{H}_{t_i}$ is assumed to be diagonal, for convenience, in order to reduce the number of parameters to be estimated.

In combination with the relationship $x_t = \ln S_t$ we have from (2.1) and (2.2)

$$\begin{cases} dx_t = (\mu - \delta_t - \frac{1}{2}\sigma_S^2)dt + \sigma_S dW_t, \\ d\delta_t = k(\alpha - \delta_t)dt + \sigma_\delta dZ_t, \end{cases} \tag{4.4}$$

i.e. we calibrate the model under the $\mathbb{P}$ (real world) measure.
From (4.4) the transition equation follows immediately,

$$\begin{bmatrix} x_{t_i} \\ \delta_{t_i} \end{bmatrix} = \begin{bmatrix} (\mu - \frac{1}{2}\sigma_S^2)\Delta t \\ k\alpha\Delta t \end{bmatrix} + \begin{bmatrix} 1 & -\Delta t \\ 0 & 1 - k\Delta t \end{bmatrix} \begin{bmatrix} x_{t_{i-1}} \\ \delta_{t_{i-1}} \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \xi_{t_i}^1 \\ \xi_{t_i}^2 \end{bmatrix}, \tag{4.5}$$

which for simplicity will be written as

$$\boldsymbol{\alpha}_{t_i} = \mathbf{c}_{t_i} + \mathbf{Q}_{t_i}\boldsymbol{\alpha}_{t_{i-1}} + \mathbf{R}\boldsymbol{\xi}_{t_i}. \tag{4.6}$$

Here $\boldsymbol{\xi}_{t_i}$ takes into account the Brownian motions dW_t and dZ_t. Hence $\boldsymbol{\xi}_{t_i}$ is assumed normal with mean zero and has a covariance-variance matrix given by

$$\mathbf{V}_{t_i} = \begin{bmatrix} \sigma_S^2\Delta t & \rho\sigma_\delta\sigma_S\Delta t \\ \rho\sigma_\delta\sigma_S\Delta t & \sigma_\delta^2\Delta t \end{bmatrix}. \tag{4.7}$$

Note that $\mathbf{V}_{t_i}$ does not depend on the state variables $\{x_t, \delta_t\}$. The matrices **H**, **Z**, **Q**, **c**, **d** given above are parametrized by the unknown parameter set $\boldsymbol{\varphi} := \left\{k, \mu, \alpha, \lambda, \sigma_S, \sigma_\delta, \rho, \{h_j\}_{j=1}^n\right\}$. We dropped the time subscript for each matrix because they are time-independent.

4.1.1 Iterative procedure

The iterative procedure consists of two main parts. Part one is the Kalman filter which is used to update the system matrices. Part two is estimating the parameter set φ. The algorithm of the Kalman filter is given in pseudo code following [30]. Define $\mathbf{a}_{t_i} = \mathbb{E}[\boldsymbol{\alpha}_{t_i}]$ and $\mathbf{P}_{t_i} = \mathbb{E}[(\mathbf{a}_{t_i} - \boldsymbol{\alpha}_{t_i})(\mathbf{a}_{t_i} - \boldsymbol{\alpha}_{t_i})']$.

while (number of iterations has not been reached , optimal φ has not been found) **do**
  (Kalman filter)
  for $i = 1 : N$ **do**
    (Prediction)
    $\tilde{\boldsymbol{a}}_{t_i} = \mathbf{Q} * \mathbf{a}_{t_{i-1}} + \mathbf{c};$[1]
    $\tilde{\mathbf{P}}_{t_i} = \mathbf{Q} * \mathbf{P}_{t_{i-1}} * \mathbf{Q}' + \mathbf{R} * \mathbf{V} * \mathbf{R}';$
    $\tilde{\mathbf{y}}_{t_i} = \mathbf{Z} * \tilde{\mathbf{a}}_{t_i} + \mathbf{d};$
    (Innovations)
    $\mathbf{v}_{t_i} = \mathbf{y}_{t_i} - \tilde{\mathbf{y}}_{t_i};$
    (Updating)
    $\mathbf{G}_{t_i} = \mathbf{Z} * \tilde{\mathbf{P}}_{t_i} * \mathbf{Z}' + \mathbf{H};$
    $\mathbf{a}_{t_i} = \tilde{\mathbf{a}}_{t_i} + \tilde{\mathbf{P}}_{t_i} * \mathbf{Z}' * \mathbf{G}_{t_i}^{-1} * \mathbf{v}_{t_i};$
    $\mathbf{P}_{t_i} = \tilde{\mathbf{P}}_{t_i} - \tilde{\mathbf{P}}_{t_i} * \mathbf{Z}' * \mathbf{G}_{t_i}^{-1} * \mathbf{Z} * \tilde{\mathbf{P}}_{t_i};$
    if $det\mathbf{G}_{t_i} \leqslant 0$ **then**
      $det\mathbf{G}_{t_i}t = 10^{-10};$
    end if
    (Log-likelihood function per iteration)
    $logl(i) = -(n/2) * \ln(2 * \pi) - 0.5 * \ln(det\mathbf{G}_{t_i}) - 0.5 * \mathbf{v}_{t_i}' * \mathbf{G}_{t_i}^{-1} * \mathbf{v}_{t_i};$
  end for
  $Log\mathbf{L} = \sum_i logl(i);$
  (Adjustment for φ via a Matlab optimization routine [2])
end while

Via an optimalisation routine, the vector φ is chosen such that the total sum of the log-likelihood function is maximized and the innovations minimized. After this is done, the matrices in the measurement and transition equation are updated and the Kalman filter algorithm will then be repeated. This iterative procedure is repeated until the optimized φ is found. With this optimized set of parameters the matrices will be updated once more and are used for the last time to generate the paths of the non-observable variables via the Kalman filter. The total sum of the log-likelihood function is calculated as follows

$$\ln\mathbf{L}(\mathbf{Y};\varphi) = -\frac{1}{2}n\ln 2\pi - \frac{1}{2}\sum_t \ln|\mathbf{G}_{t_i}| - \frac{1}{2}\sum_t \mathbf{v}_{t_i}'\mathbf{G}_{t_i}^{-1}\mathbf{v}_{t_i}, \tag{4.8}$$

where $\mathbf{Y}_{t_i}$ is the information vector at time t_i and it is assumed that $\mathbf{Y}_{t_i}$ conditional on $\mathbf{Y}_{t_{i-1}}$ has a normal distribution with mean $\mathbb{E}[\mathbf{Y}_{t_i}|\mathbf{Y}_{t_{i-1}}]$ and covariance matrix $\mathbf{G}_{t_i}$.

4.2 Numerical results for the Ornstein-Uhlenbeck process

We use weekly observations (on every friday to account for possible weekend effects) of light crude oil data from 01-02-2002 until 25-01-2008 (313 observations). At each observation, seven monthly futures

[1] $*$ denotes the multiplication symbol.
[2] The Matlab optimization routine is fmincon.

contracts $\mathbf{F}_t(\tau_1)$,...,$\mathbf{F}_t(\tau_7)$ are observed, where τ_1 denotes contract 1 and so on. Each diagonal element of the error matrix $\mathbf{H}$ corresponds to each maturity, i.e. h_1 corresponds to $\mathbf{F}_t(\tau_1)$, h_2 corresponds to $\mathbf{F}_t(\tau_2)$ and so on. The used market data is such that the time to maturities τ_j are constant for each forward contract j and for each observation data, i.e. for each observation date of contract j there exists a future price with maturity τ_j. Furthermore we assume deterministic interest rates, i.e. for each observation, 7 contracts are observed, meaning 7 different interest rates. r_1 is the interest rate committed to $\mathbf{F}_t(\tau_1)$ etc.

4.2.1 Setting up the calibration

In order to start the iterative process, some difficulties have to be tackled first. Namely, the initial choice of the parameter set φ and the non-observable variables at time zero, i.e. x_0 and δ_0.
We start with $\varphi = \left\{0.3, 0.2, 0.06, 0.01, 0.4, 0.4, 0.8, \left\{\mathbb{V}\mathrm{ar}[\ln \mathbf{F}_t(\tau_j)]^2\right\}_{j=1}^{7}\right\}$. The first seven parameters are chosen arbitrary. However, for the diagonal elements $\{h_j\}_{j=1}^{n}$ of the error variance matrix $\mathbf{H}$ we choose the squares of the measurement variance. Therefore we expect the error to be small in comparison to the variance of the measurements. For the non-observable parameters (cf. [8]), the nearest future price is retained as the spot price S_0 and

$$\delta_{t_i}^{\text{market}}(\tau_1, \tau_2) = \frac{r_1\tau_1 - r_2\tau_2}{\tau_1 - \tau_2} - \frac{\ln(\mathbf{F}_{t_i}(\tau_1)) - \ln(\mathbf{F}_{t_i}(\tau_2))}{\tau_1 - \tau_2}. \tag{4.9}$$

and we use $\delta_0 = \delta_0^{\text{market}}(\tau_1, \tau_2)$.

Implementation issues: There are a few important implementation details worth mentioning. The log-likelihood function estimates the optimal parameter set. For that, it has control of a parameter set bound. We use the following bounds

$$[\quad \mu \quad \sigma_S \quad \alpha \quad k \quad \sigma_\delta \quad \rho \quad \lambda \quad h1 \quad h2 \quad h3 \quad h4 \quad h5 \quad h6\ h7];$$

```
lowerbound=[ -10  0.001   -10   0.001  0.001  -1    -10   -1 -1 -1 -1 -1 -1 -1];
upperbound=[  10     10    10      10     10   1     10    1  1  1  1  1  1  1];
```

Omitting these bounds, the iterative procedure may break down. In case of matrix $\mathbf{H}$, which consists of the variances of the error terms, can become negative. But being negative, this can cause problems in the KF. In the KF algorithm $\mathbf{G}_t$ can become singular and so impossible to inverse. So we square each element of $\mathbf{H}$ before calculating $\mathbf{G}_t$ and hence they effectively enter $\mathbf{G}_t$ as the variance. This is why $\{h_j\}_{j=1}^{n}$ in Table 4.1 and in Table 4.2 are in absolute value.

4.2.2 Results

The parameter set is estimated for different initial sets and given in Table 4.1. To test the robustness of the calibration, we first choose randomly an initial parameter set. This gaves us the optimized set and subsequently we used this optimized set as our initial set. Within a few interation steps the parameters converges to the same values.
Remark: In the following sections we use the third column of Table 4.1 to be the initial parameter set.

Parameters	Ini parset	Opti parset	Ini parset	Opti parset	Ini parset	Opti parset
k	0.3	1.4221 (0.0372)	0.3	1.4221 (0.0380)	2	1.4221 (0.0382)
μ	0.2	0.3733 (0.1471)	0.2	0.3733 (0.1376)	0.2	0.3733 (0.1382)
α	0.06	0.0699 (0.1128)	0.2	0.0699 (0.1025)	0.06	0.0699 (0.1082)
λ	0.01	-0.0183(0.1602)	0.1	-0.0183(0.1459)	0.01	-0.0183(0.1543)
σ_S	0.4	0.3630 (0.0137)	0.4	0.3630 (0.0139)	0.5	0.3630 (0.0153)
σ_δ	0.4	0.4028 (0.0165)	0.4	0.4028 (0.0172)	0.4	0.4028 (0.0181)
ρ	0.8	0.8378 (0.0162)	0.5	0.8378 (0.0164)	0.5	0.8378 (0.0177)
$\|h_1\|$	0.0246	0.0188 (0.0008)	0.0246	0.0188 (0.0007)	0.01	0.0188 (0.0007)
$\|h_2\|$	0.0268	0.0072 (0.0003)	0.0268	0.0072 (0.0003)	0.01	0.0072 (0.0003)
$\|h_3\|$	0.0291	0.0022 (0.0001)	0.0291	0.0022 (0.0001)	0.01	0.0022 (0.0001)
$\|h_4\|$	0.0313	0.0000 (0.0001)	0.0313	0.0000 (0.0001)	0.01	0.0000 (0.0001)
$\|h_5\|$	0.0336	0.0006 (0.0000)	0.0336	0.0006 (0.0000)	0.01	0.0006 (0.0000)
$\|h_6\|$	0.0357	0.0000 (0.0001)	0.0357	0.0000 (0.0001)	0.01	0.0000 (0.0001)
$\|h_7\|$	0.0377	0.0014 (0.0001)	0.0377	0.0014 (0.0001)	0.01	0.0014 (0.0001)
Log-Likelihood		8744.6479		8744.6479		8744.6479

Table 4.1: Optimized parameter set (Opti parset) for different initial parameter sets (Ini parset). Standard errors in parentheses.

With the optimized parameter set given in Table 4.1 we plot, in Figure 4.1, the calibrated future prices versus the one-month maturity of the log of the future prices, i.e. $\ln(\mathbf{F}_{t_i}(\tau_1))$. The calibrated future prices seems to fit the market future prices succesfully. The same results holds for all seven futures contracts.

Nowadays traders determine the market convenience yield via the future prices like in (4.9). In Figure

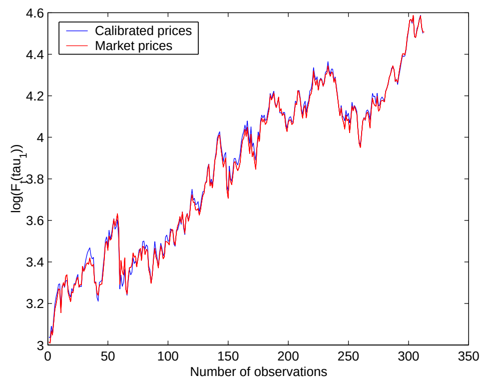


Figure 4.1: Comparison between the calibrated future prices and market future prices for contract τ_1.

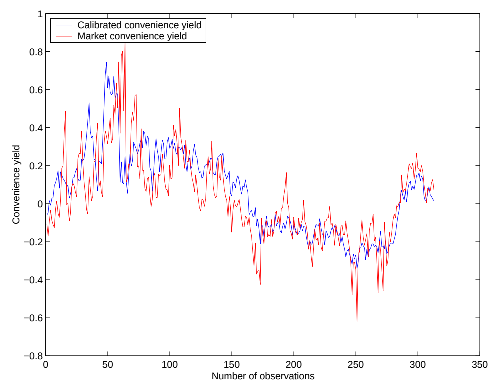


Figure 4.2: Comparison between filtered δ_{t_i} and market convenience yield, $\delta_{t_i}^{\text{market}}(\tau_1, \tau_2)$.

4.2 , the estimated convenience yield is plotted versus the market convenience yield. In Figure 4.2 we see that the calibrated convenience yield does not follow the path of the market convenience yield. This can have several reasons. First of all, we compared it to our calculated convenience yield (4.9) but there are more factors included in the market future prices than the interest rates, spot prices and convenience yield. In fact, the market convenience yield actually consists of numerous terms, e.g. storage costs, security costs, transport costs etc. So the comparison is expected to be not succesfully as we do not compare the pure convenience yield but, what is called, the total cost of carry.

For each futures contract we plot the innovation (ϵ_t in equation (4.3)), i.e. $\mathbf{v}_{t_i} = \mathbf{y}_{t_i} - \tilde{\mathbf{y}}_{t_i}$. We assumed the measurement errors to be normal with mean zero and variance $\mathbf{H}$. Considering the figures, this assumption is acceptable because the mean of the innovations is almost zero for all seven futures contracts. Note that there are some large error 'spikes'. This can be caused by data errors.

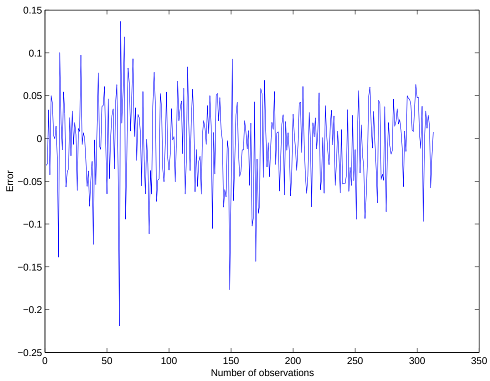


Figure 4.3: Innovation corresponding to $\mathbf{F}_t(\tau_1)$. Mean = -8.1734e-004, Variance = 0.0022.

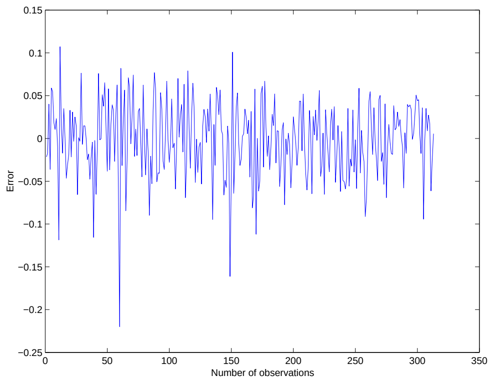


Figure 4.4: Innovation corresponding to $\mathbf{F}_t(\tau_2)$. Mean = 8.6078e-004, Variance = 0.0019.

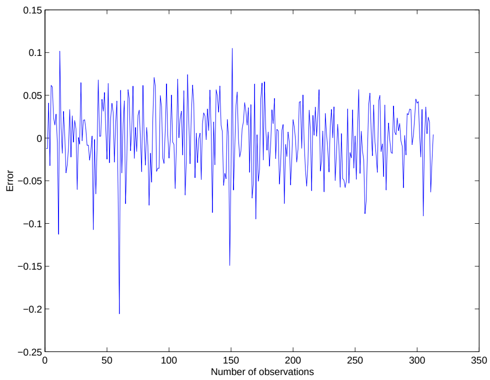


Figure 4.5: Innovation corresponding to $\mathbf{F}_t(\tau_3)$. Mean = 6.4612e-004, Variance = 0.0017.

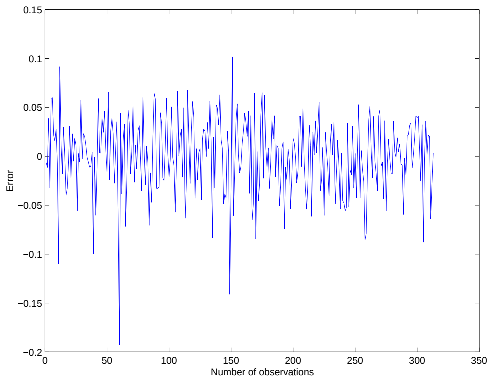


Figure 4.6: Innovation corresponding to $\mathbf{F}_t(\tau_4)$. Mean = -1.0181e-004, Variance = 0.0015.

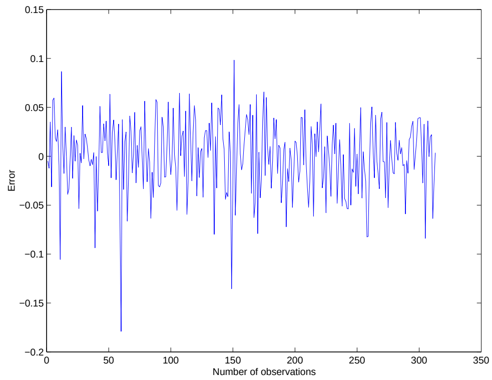


Figure 4.7: Innovation corresponding to $\mathbf{F}_t(\tau_5)$. Mean = -5.5711e-004, Variance = 0.0014.

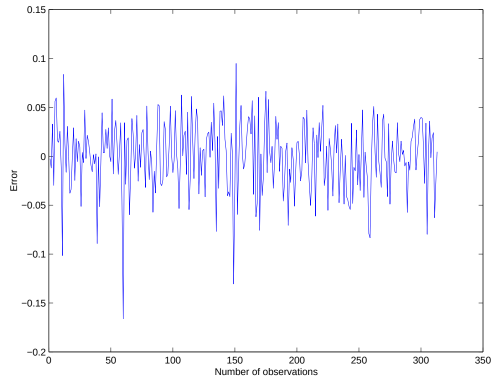


Figure 4.8: Innovation corresponding to $\mathbf{F}_t(\tau_6)$. Mean = -3.4456e-004, Variance = 0.0013.

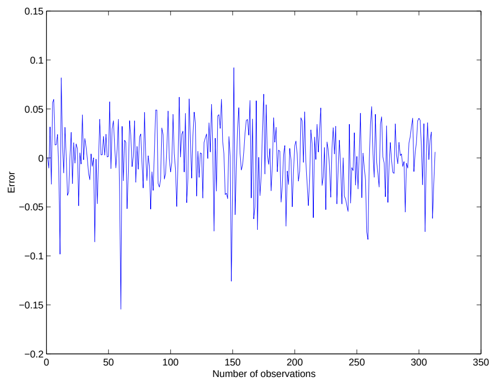


Figure 4.9: Innovation corresponding to $\mathbf{F}_t(\tau_7)$. Mean = 6.1501e-004, Variance = 0.0012.

4.2.3 Results for a different number of contracts

In this section we test the method for different number of contracts. We see in Table 4.2 that the standard error increases if only two contracts are used. This is expected because the Kalman filter now has to estimate two unobservable values using only two data sets.

Contracts	$\mathbf{F}_t(\tau_1),...,\mathbf{F}_t(\tau_7)$	$\mathbf{F}_t(\tau_1),\mathbf{F}_t(\tau_3),\mathbf{F}_t(\tau_5),\mathbf{F}_t(\tau_7)$	$\mathbf{F}_t(\tau_1),\mathbf{F}_t(\tau_7)$
k	1.4221 (0.0372)	1.4591 (0.0602)	1.7628 (0.4897)
μ	0.3733 (0.1471)	0.3654 (0.1354)	0.8134 (0.5488)
α	0.0699 (0.1128)	0.0653 (0.1004)	0.5169 (0.5381)
λ	-0.0183(0.1602)	-0.0331(0.1475)	1.8517 (2.2353)
σ_S	0.3630 (0.0137)	0.3610 (0.0139)	0.3501 (0.0153)
σ_δ	0.4028 (0.0165)	0.3995 (0.0183)	0.4108 (0.0705)
ρ	0.8378 (0.0162)	0.8418 (0.0174)	0.8172 (0.0234)
$\|h_1\|$	0.0188 (0.0008)	0.0161 (0.0007)	0.0188 (0.0012)
$\|h_2\|$	0.0072 (0.0003)	0.0000 (0.0003)	0.0081 (0.0031)
$\|h_3\|$	0.0022 (0.0001)	0.0021 (0.0001)	
$\|h_4\|$	0.0000 (0.0001)	0.0010 (0.0012)	
$\|h_5\|$	0.0006 (0.0000)		
$\|h_6\|$	0.0000 (0.0001)		
$\|h_7\|$	0.0014 (0.0001)		
Log-Likelihood	8744.6479	4001.8997	1440.5857

Table 4.2: Optimized parameter set (Opti parset) for different initial parameter sets (Ini parset). Standard errors in parentheses.

4.2.4 Parameter explanation

If we look at Table 4.1, we see no unrealistic parameter values. The correlation between the two state variables is 0.8378. To find graphical evidence for this positive number we made a scatterplot of the increments of both the state-variables, i.e. $\Delta x_{t_i} := x_{t_i} - x_{t_{i-1}}$ versus $\Delta\delta_{t_i} := \delta_{t_i} - \delta_{t_{i-1}}$.

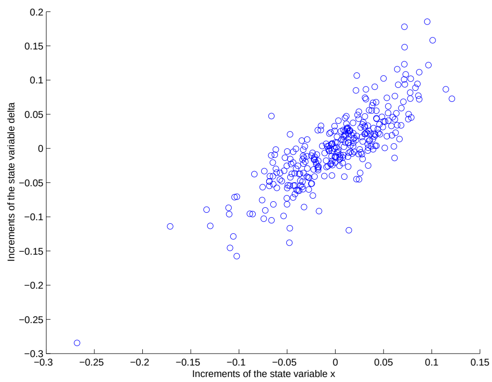


Figure 4.10: Scatterplot of $(\Delta x_{t_i}, \Delta\delta_{t_i})$ for 313 observations.

Since the pattern of dots slopes from lower left to upper right, it suggests a positive correlation between the variables being studied.

4.2.5 Kalman forecasting

We test a Kalman forecasting method on half of the observations. The first half, from observation i=1 until i=156 we perform the KF algorithm with the optimized parameter set given in the third column of Table 4.1. From i=157 until i=313 we forecast the state variable δ_t and the future prices by the following recursive equations, keeping the optimized parameter set fixed,

$$\begin{aligned} \mathbf{a}_{t_i} &= \mathbf{Q}\mathbf{a}_{t_i} + \mathbf{c}, \\ \mathbf{P}_{t_i} &= \mathbf{Q}\mathbf{P}_{t_i}\mathbf{Q}' + \mathbf{R}\mathbf{V}_{t_i}\mathbf{R}', \\ \mathbf{y}_{t_i} &= \mathbf{Z}\mathbf{a}_{t_i} + \mathbf{D}. \end{aligned} \tag{4.10}$$

Forecasting for the state variable δ_t is not a succes. This can be explained by the sudden drop at 160.

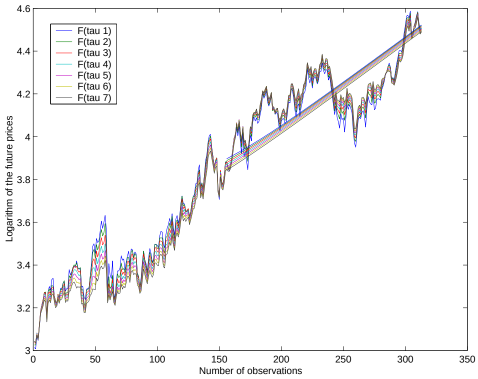


Figure 4.11: Forecast over half a sample of the log of the future prices (F1-F7).

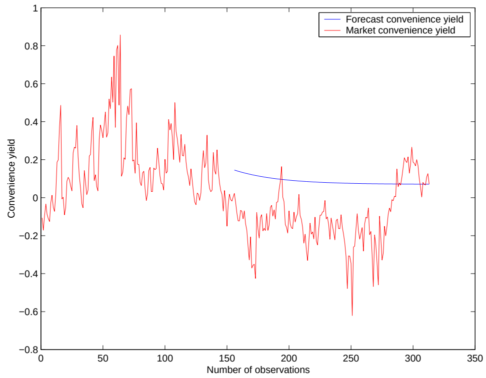


Figure 4.12: Forecast for CY over half a sample.

The Kalman forecasting can not absorb this sudden drop in the convenience yield. For that, it follows the previous estimations. However, as noted earlier, the mean-reverting term is equal to 0.0699, which is approximately equal to the forecasting. Also for the future prices, the sudden drop at 250 can not be absorbed by this forecasting method. This is a real disadvantage of this method. A way to get the method to absorb these drops, jumps can be inserted in the stochastic processes of both the state variables. The smoothness of the Kalman forecasting is due to the fact that the parameters and so the system matrices are not updated anymore. In other words, there is no more stochasticity in the prices. The estimate is based on historical data only. The trend is based on this data and will fit in some optimal sense.

4.3 The extended Kalman filter

The extended Kalman filter (EKF) is an extension of the standard KF. The basic iterative procedure is the same but the main difference between the two is that the latter is limited to linear measurement and transition equations. For the EKF this is not a requisite. Since the OU process is linear in the log of the state variable S_t, see (4.2), applying the EKF to this process does not make much sense. However, we explain in detail how the EKF works and we will apply it to the OU process in which the measurement equation is nonlinear, i.e. we do not use the log transformation. Furthermore we investigate its consistency to the standard KF (i.e. if it converges to the same parameters) and stability of the method.

4.3.1 General setup of the extended Kalman filter

In the EKF the transition and measurement equations consist of nonlinear functions that depend on the state variables. The transition equation is given by

$$\boldsymbol{\alpha}_{t_i} = \mathbf{Q}(\boldsymbol{\alpha}_{t_{i-1}}) + \mathbf{R}(\boldsymbol{\alpha}_{t_i})\boldsymbol{\xi}_{t_i}, \tag{4.11}$$

where $\boldsymbol{\alpha}_{t_i}$ is the state vector at time t_i, $\mathbf{Q}(\boldsymbol{\alpha}_{t_{i-1}})$ and $\mathbf{R}(\boldsymbol{\alpha}_{t_{i-1}})$ are nonlinear differentiable functions depending on the state vector and measured at time t_{i-1} and $\boldsymbol{\xi}_{t_i}$ is the transition error. Again, this error is assumed to be normal with mean zero and covariance matrix $\mathbf{V}_{t_i}$. The measurement equation is given by

$$\mathbf{Y}_{t_i} = \mathbf{Z}(\boldsymbol{\alpha}_{t_i}) + \boldsymbol{\epsilon}_{t_i}, \tag{4.12}$$

where $\mathbf{Y}_{t_i}$ is the informational vector (in this case, consisting of the future prices), $\mathbf{Z}(\boldsymbol{\alpha}_{t_i})$ is a nonlinear function depending on the state vector measured at time t_i and $\boldsymbol{\epsilon}_{t_i}$ the measurement error. Again, this error is assumed to be normal with mean zero and covariance matrix $\mathbf{H}_{t_i}$. Since we assumed that the nonlinear functions, $\mathbf{Q}(\boldsymbol{\alpha}_{t_{i-1}})$ and $\mathbf{R}(\boldsymbol{\alpha}_{t_{i-1}})$, are differentiable, we can linearize them by computing their Jacobian matrices,

$$\hat{\mathbf{Z}} = \frac{\partial \mathbf{Z}(\boldsymbol{\alpha}_{t_i})}{\partial \boldsymbol{\alpha}_{t_i}}, \quad \hat{\mathbf{Q}} = \frac{\partial \mathbf{Q}(\boldsymbol{\alpha}_{t_{i-1}})}{\partial \boldsymbol{\alpha}_{t_{i-1}}}, \quad \hat{\mathbf{R}} \approx \mathbf{R}. \tag{4.13}$$

By using these Jacobians, the linearized state-space model is then given by

$$\begin{cases} \boldsymbol{\alpha}_{t_i} \approx \hat{\mathbf{Q}}\boldsymbol{\alpha}_{t_{i-1}} + \hat{\mathbf{R}}\boldsymbol{\xi}_{t_i}, \\ \mathbf{Y}_{t_i} \approx \hat{\mathbf{Z}}\boldsymbol{\alpha}_{t_i} + \boldsymbol{\epsilon}_{t_i}. \end{cases} \tag{4.14}$$

In other words, the EKF implicitly introduces an extra approximation error in the filtering procedure.

4.3.2 Iterative procedure

The iterative procedure consists of two main parts. Part one is the extended Kalman filter which is used to update the system matrices. Part two is estimating the parameter set φ. The algorithm of the extended Kalman filter is given in pseudo code.

while (number of iterations has not been reached , optimal φ has not been found) **do**
  (Kalman filter)
  for $i = 1 : N$ **do**
    (Prediction)
    $\tilde{\boldsymbol{a}}_{t_i} = \mathbf{Q} * \mathbf{a}_{t_{i-1}} + \mathbf{c}$;
    $\tilde{\mathbf{P}}_{t_i} = \mathbf{Q} * \mathbf{P}_{t_{i-1}} * \mathbf{Q}' + \mathbf{R} * \mathbf{V} * \mathbf{R}'$;
    $\tilde{\mathbf{y}}_{t_i} = \mathbf{Z} * \tilde{\mathbf{a}}_{t_i} + \mathbf{d}$;
    (Innovations)
    $\mathbf{v}_{t_i} = \mathbf{y}_{t_i} - \tilde{\mathbf{y}}_{t_i}$;
    (Updating)
    $\mathbf{G}_{t_i} = \mathbf{Z} * \tilde{\mathbf{P}}_{t_i} * \mathbf{Z}' + \mathbf{H}$;
    $\mathbf{a}_{t_i} = \tilde{\mathbf{a}}_{t_i} + \tilde{\mathbf{P}}_{t_i} * \mathbf{Z}' * \mathbf{G}_{t_i}^{-1} * \mathbf{v}_{t_i}$;
    $\mathbf{P}_{t_i} = \tilde{\mathbf{P}}_{t_i} - \tilde{\mathbf{P}}_{t_i} * \mathbf{Z}' * \mathbf{G}_{t_i}^{-1} * \mathbf{Z} * \tilde{\mathbf{P}}_{t_i}$;
    if $det\mathbf{G}_{t_i} \leqslant 0$ **then**
      $det\mathbf{G}_{t_i}t = 10^{-10}$;
    end if
    (Log-likelihood function per iteration)
    $logl(i) = -(n/2) * \ln(2 * \pi) - 0.5 * \ln(det\mathbf{G}_{t_i}) - 0.5 * \mathbf{v}'_{t_i} * \mathbf{G}_{t_i}^{-1} * \mathbf{v}_{t_i}$;
    (Update system matrices)
  end for

$Log\mathbf{L} = \sum_i logl(i)$;
(Adjustment for φ via a Matlab optimization routine)
end while

The same implementation issues as discussed (in Section 4.2.8) still holds. The EKF works the same as the standard KF described in Section 4.1. However, since the system matrices depend on the state variables, they have to be updated in every timestep i.

4.3.3 Setup of the extended Kalman filter applied to the OU process

The same data is used as for the calibration of the standard KF. The closed form solution of the future prices is given by (4.1). The system matrices are not time-dependent so we drop the time subscript. From (2.1) and (2.2) the transition equation in discrete time is given by

$$\begin{bmatrix} S_{t_i} \\ \delta_{t_i} \end{bmatrix} = \mathbf{Q}(S_{t_{i-1}}, \delta_{t_{i-1}}) + \mathbf{R}(S_{t_{i-1}}, \delta_{t_{i-1}})\boldsymbol{\xi}_{t_i}, \tag{4.15}$$

where

- $\begin{bmatrix} S_{t_i} \\ \delta_{t_i} \end{bmatrix}$ is the (2 $\times$ 1) state vector consisting of the spot price, S_t and the CY δ_t,
- $\mathbf{Q}(S_{t_{i-1}}, \delta_{t_{i-1}})$ is a (2 $\times$ 1) vector: $\mathbf{Q} = \begin{bmatrix} S_{t_{i-1}}(1 + \mu\Delta t - \delta_{t_{i-1}}\Delta t) \\ k\alpha\Delta t + \delta_{t_{i-1}}(1 - k\Delta t) \end{bmatrix}$,
- $\mathbf{R}(S_{t_{i-1}}, \delta_{t_{i-1}})$ is a (2 $\times$ 2) matrix: $\mathbf{R} = \begin{bmatrix} S_{t_{i-1}} & 0 \\ 0 & 1 \end{bmatrix}$,
- $\mathbf{V}$ is a (2 $\times$ 2) matrix: $\mathbf{V} = \begin{bmatrix} \sigma_S^2\Delta t & \rho\sigma_S\sigma_\delta\Delta t \\ \rho\sigma_S\sigma_\delta\Delta t & \sigma_\delta^2\Delta t \end{bmatrix}$.

The measurement equation is defined by

$$\mathbf{Y}_t = \mathbf{Z}(S_{t_i}, \delta_{t_i}) + \boldsymbol{\epsilon}_{t_i}, \tag{4.16}$$

where

- $\mathbf{Y}_{t_i} = \mathbf{F}_{t_i}(\tau_j)$, see Section 4.1,
- $\mathbf{Z}(S_{t_i}, \delta_{t_i})$ is a $(n \times 1)$ matrix. For $j = 1 : n$, $\mathbf{Z} = \left[S_{t_i} \times \exp(-A(\tau_j)\delta_{t_i} + B(\tau_j))\right]$, where $A(\tau_j)$ and $B(\tau_j)$ are given by (4.1),
- $\boldsymbol{\epsilon}_{t_i}$ is a $n \times 1$ vector of uncorrelated disturbances, see Section 4.1.

The Jacobians of the matrices $\hat{\mathbf{Q}}$ and $\hat{\mathbf{Z}}$ are given by

$$\hat{\mathbf{Q}} = \begin{bmatrix} 1 + \mu\Delta t - \delta_{t_i}\Delta t & -S_{t_i}\Delta t \\ 0 & (1 - k\Delta t) \end{bmatrix}, \quad \hat{\mathbf{Z}} = \begin{bmatrix} e^{(-A(\tau_j)\delta_{t_{i-1}} + B(\tau_j))} & -S_{t_{i-1}}A(\tau_j)e^{(-A(\tau_j)\delta_{t_{i-1}} + B(\tau_j))} \end{bmatrix} \tag{4.17}$$

4.3.4 Setting up the calibration

The extended Kalman filter is not an optimal filter. In addition, if the initial estimate of the state is wrong or if the process is modeled incorrectly, the filter may quickly diverge, owing to its linearization. The choice of the initial parameter set should therefore be done carefully. The initial choice of the non-observable variables at time zero, i.e. x_0 and δ_0 remains the same. For the initial parameter set we start with the optimized set which we have found for the standard Kalman filter, that is, $\varphi = \left\{1.422, 0.3733, 0.0699, -0.0183, 0.3630, 0.4028, 0.8378, \frac{1}{1000}\left\{\mathbb{V}\text{ar}[(\ln \mathbf{F}_t(\tau_j))]\right\}_{j=1}^{7}\right\}$. Note that we expect the error to be small in comparison to the variance of the measurements, since we do not consider the logarithm of the futures but the future prices itself, the variance of the measurements is large. Therefore we divide the variances for each contract by 1000.

4.3.5 Results

As noted in Section 4.2.2 we start with an initial parameter set and to test the robustness we differ this set. Results are given in Table 4.3.

Parameters	Ini parset	Opti parset	Ini parset	Opti parset	Ini parset	Opti parset
k	1.4221	1.6159 (0.0423)	1.6	1.6159 (0.0402)	1.3	1.6159 (0.0377)
μ	0.3733	0.4144 (0.0871)	0.2	0.4156 (0.1273)	0.4	0.4156 (0.1498)
α	0.0699	0.1076 (0.0099)	0.1	0.1087 (0.0964)	0.1	0.1087 (0.1121)
λ	-0.0183	0.0589 (0.0198)	-0.0183	0.0607 (0.1561)	0.1	0.0607 (0.11811)
σ_S	0.3630	0.4530 (0.0200)	0.4	0.4530 (0.0200)	0.4	0.4530 (0.0188)
σ_δ	0.4028	0.3641 (0.0137)	0.4	0.3641 (0.0148)	0.4	0.3641 (0.0148)
ρ	0.8378	0.8198 (0.0183)	0.8	0.8198 (0.0164)	0.3	0.8198 (0.0190)
$\|h_1\|$	0.3647	0.6116 (0.0242)	0.0246	0.6116 (0.0227)	0.3	0.6116 (0.0234)
$\|h_2\|$	0.3773	0.1671 (0.0063)	0.0268	0.1671 (0.0066)	0.3	0.1671 (0.0069)
$\|h_3\|$	0.3886	0.0000 (0.0104)	0.0291	0.0000 (0.0106)	0.3	0.0000 (0.0104)
$\|h_4\|$	0.3986	0.0347 (0.0013)	0.0313	0.0347 (0.0014)	0.3	0.0347 (0.0013)
$\|h_5\|$	0.4078	0.0000 (0.0029)	0.0336	0.0000 (0.0028)	0.3	0.0000 (0.0027)
$\|h_6\|$	0.4159	0.0707 (0.0025)	0.0357	0.0707 (0.0026)	0.3	0.0707 (0.0028)
$\|h_7\|$	0.4229	0.1629 (0.0069)	0.0377	0.1629 (0.0061)	0.3	0.1629 (0.0061)
Log-Likelihood		365.6806		365.6806		365.6806

Table 4.3: Optimized parameter set (Opti parset) for different initial parameter sets (Ini parset). Standard errors in parentheses.

Using the optimized parameter set in Table 4.3 we plot, in Figure 4.13, the calibrated future prices versus the one-month maturity of the future prices, i.e. $\mathbf{F}_{t_i}(\tau_1)$. The same results hold for all seven futures contracts. For each futures contract we plot the innovation. From Figures 4.15-4.21 we can see that

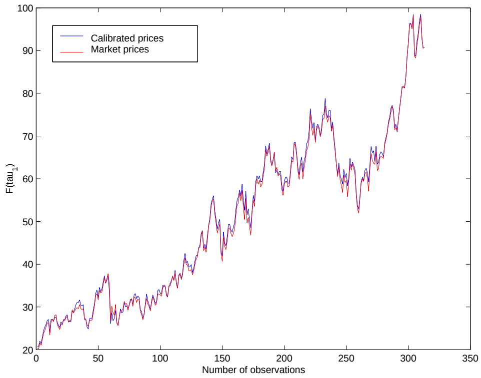


Figure 4.13: Comparison between the calibrated future prices and market future prices for contract τ_1.

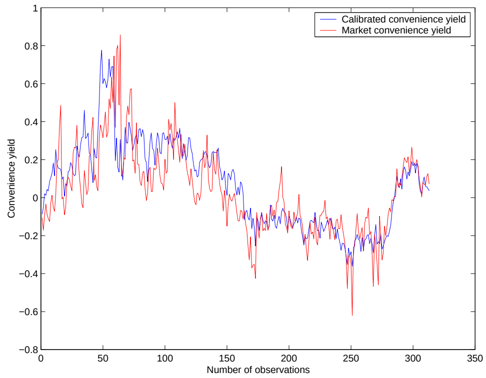


Figure 4.14: Comparison between the calibrated δ_{t_i} and the market convenience yield, $\delta_{t_i}^{\text{market}}(\tau_1, \tau_2)$.

the mean of the innovations differs more from zero than the mean of the innovations shown in Figures 4.3-4.9. This is due to the fact that we used the extended Kalman filter. However, the results are still acceptable.

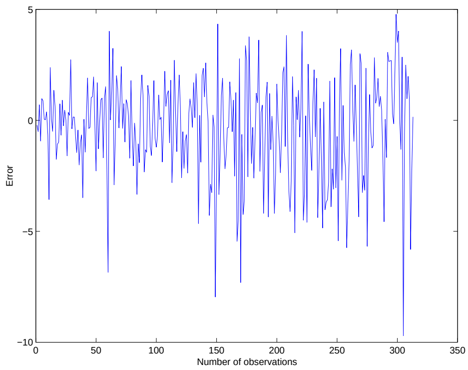


Figure 4.15: Innovation corresponding to $\mathbf{F}_t(\tau_1)$. Mean = -0.3819, Variance = 5.2406.

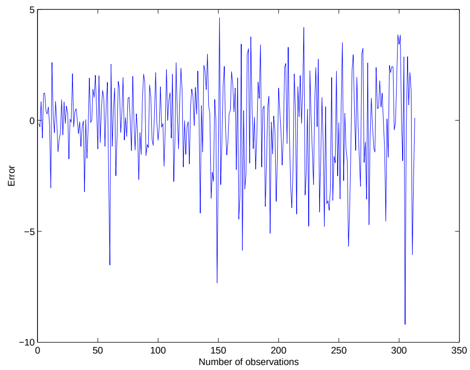


Figure 4.16: Innovation corresponding to $\mathbf{F}_t(\tau_2)$. Mean = -0.1894, Variance = 4.5616.

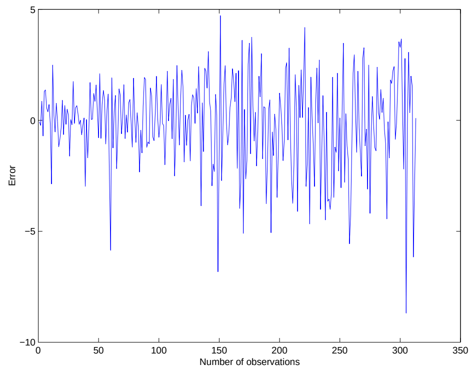


Figure 4.17: Innovation corresponding to $\mathbf{F}_t(\tau_3)$. Mean = -0.1187, Variance = 4.1149.

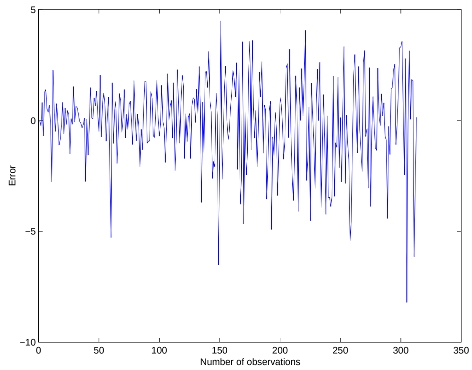


Figure 4.18: Innovation corresponding to $\mathbf{F}_t(\tau_4)$. Mean = -0.1040, Variance = 3.7877.

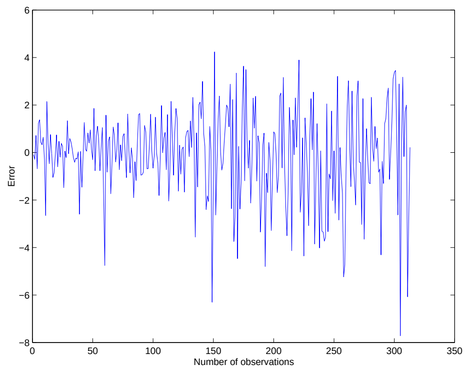


Figure 4.19: Innovation corresponding to $\mathbf{F}_t(\tau_5)$. Mean = -0.1015, Variance = 3.5212.

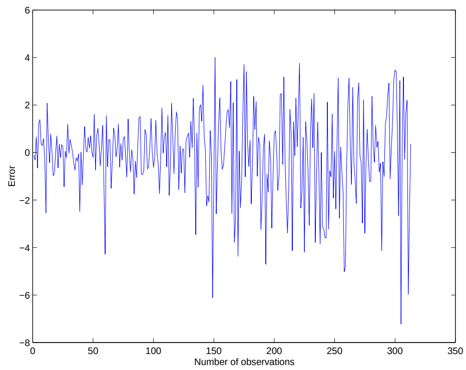


Figure 4.20: Innovation corresponding to $\mathbf{F}_t(\tau_6)$. Mean = -0.0941, Variance = 3.3058.

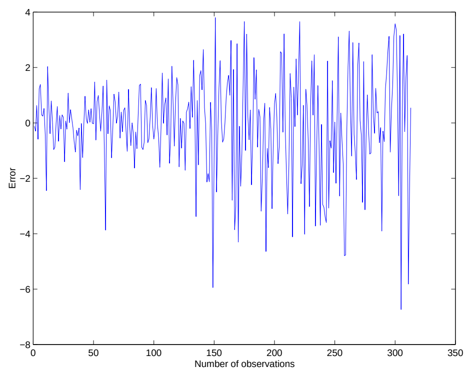


Figure 4.21: Innovation corresponding to $\mathbf{F}_t(\tau_7)$. Mean = -0.0771, Variance = 3.1398.

There are some slight differences in the optimized parameter set. This may be caused by the linearization of the system matrices. The optimized parameters do not differ much from the parameters found for the standard KF. However, as we can see in Table 4.3, the method is stable with respect to different initial values of the parameters.

4.3.6 Results for different number of contracts

As in Section 4.2.3 we see that the standard error increases as soon as two contracts are observed.
The standard errors for two contracts are not as high as those in Table 4.2. This can be explained by

Contracts	$\mathbf{F}_t(\tau_1), ..., \mathbf{F}_t(\tau_7)$	$\mathbf{F}_t(\tau_1), \mathbf{F}_t(\tau_3), \mathbf{F}_t(\tau_5), \mathbf{F}_t(\tau_7)$	$\mathbf{F}_t(\tau_1), \mathbf{F}_t(\tau_7)$
k	1.6159 (0.0423)	1.5799 (0.0655)	1.4987 (0.5099)
μ	0.4144 (0.0871)	0.3979 (0.1461)	0.1322 (0.1468)
α	0.1076 (0.0099)	0.0949 (0.1022)	-0.1621 (0.1034)
λ	0.0589 (0.0198)	0.0346 (0.1614)	-0.9509 (0.2272)
σ_S	0.4530 (0.0200)	0.3965 (0.0214)	0.3409 (0.0597)
σ_δ	0.3641 (0.0137)	0.3559 (0.0143)	0.3327 (0.0152)
ρ	0.8198 (0.0183)	0.8359 (0.0176)	0.7908 (0.0259)
$\|h_1\|$	0.6116 (0.0242)	0.0161 (0.0007)	0.3623 (0.0419)
$\|h_2\|$	0.1671 (0.0063)	0.0000 (0.0003)	0.0000 (0.0620)
$\|h_3\|$	0.0000 (0.0104)	0.0021 (0.0001)	
$\|h_4\|$	0.0347 (0.0013)	0.0010 (0.0012)	
$\|h_5\|$	0.0000 (0.0029)		
$\|h_6\|$	0.0707 (0.0025)		
$\|h_7\|$	0.1629 (0.0069)		
Log-Likelihood	365.6805	-737.1525	1440.5857

Table 4.4: Optimized parameter set (Opti parset) for different initial parameter sets (Ini parset). Standard errors in parentheses.

the initial parameter set. For the EKF this set is almost equal to the optimized set. The initial set for the KF is chosen arbitrary and differs more from the optimized set. This may cause some bigger standard errors for the parameters. Since the prediction errors are quite big, one may have doubts by considering the diagonal elements of the error covariance matrix **H** to be bounded in an interval [-1,1]. In order to let the optimalization routine search for the diagonal elements $\{h_j\}_{j=1}^n$ from a bigger range we took the bounds as follows:

$$[\quad \mu \quad \sigma_S \quad \alpha \quad k \quad \sigma_\delta \quad \rho \quad \lambda \quad h1 \quad h2 \quad h3 \quad h4 \quad h5 \quad h6 \quad h7];$$

```
lowerbound=[ -10  0.001 -10 0.001 0.001  -1 -10  -100 -100 -100 -100 -100 -100 -100];
upperbound=[  10     10  10    10    10   1  10   100  100  100  100  100  100  100];
```

The result are presented in Table 4.5.

As we can see, increasing the lower and upper bounds for the diagonal elements of the covariance matrix **H** does not have any influence on the optimized parameter set.

Parameters	Ini parset	Opti parset
k	1.4221	1.6159 (0.0386)
μ	0.3733	0.4144 (0.1405)
α	0.0699	0.1076 (0.1088)
λ	-0.0183	0.0589 (0.1746)
σ_S	0.3630	0.4530 (0.0198)
σ_δ	0.4028	0.3641 (0.0155)
ρ	0.8378	0.8198 (0.0184)
$\|h_1\|$	36.4707	0.6116 (0.0200)
$\|h_2\|$	37.7297	0.1671 (0.0060)
$\|h_3\|$	38.8572	0.0000 (0.0100)
$\|h_4\|$	39.8636	0.0347 (0.0014)
$\|h_5\|$	40.7809	0.0000 (0.0027)
$\|h_6\|$	41.5923	0.0707 (0.0026)
$\|h_7\|$	42.2907	0.1629 (0.0061)
Log-Likelihood		365.6806

Table 4.5: Optimized parameter set (Opti parset) for different initial parameter sets (Ini parset). Standard errors in parentheses.

4.3.7 Parameter explanation

The correlation between the two state variables is 0.8198. To find graphical evidence for this positive correlation we made a scatterplot of the increments of both the state-variables, i.e. $\Delta S_{t_i} := S_{t_i} - S_{t_{i-1}}$ versus $\Delta\delta_{t_i} := \delta_{t_i} - \delta_{t_{i-1}}$.

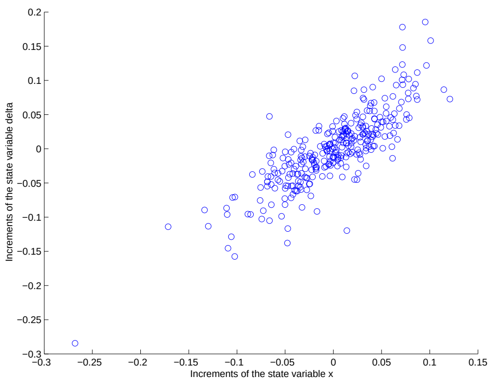


Figure 4.22: Scatterplot of $(\Delta S_{t_i}, \Delta\delta_{t_i})$ for 313 observations.

Since the pattern of dots slopes from lower left to upper right, it suggests a positive correlation between the variables being studied, which is exactly what we expected.

4.3.8 Kalman forecasting

We test a Kalman forecasting method on half of the observations. On the first half, from observation i=1 until i=156 we perform the EKF algorithm with the optimized parameter set given in the third column of Table 4.1. From i=157 until i=313 we forecast the state variable and δ_t and the future prices by the following recursive equations

$$\begin{aligned} \mathbf{a}_{t_i} &= \mathbf{T}_{t_i}, \\ \mathbf{P}_{t_i} &= \hat{\mathbf{Q}}\mathbf{P}_{t_i}\hat{\mathbf{Q}}' + \hat{\mathbf{R}}\mathbf{V}_{t_i}\hat{\mathbf{R}}', \\ \mathbf{y}_{t_i} &= \mathbf{Z}. \end{aligned} \tag{4.18}$$

Note that after each recursion the system matrices $\hat{\mathbf{Q}}$, $\hat{\mathbf{R}}$ and $\mathbf{Z}$ are updated with the state variables.

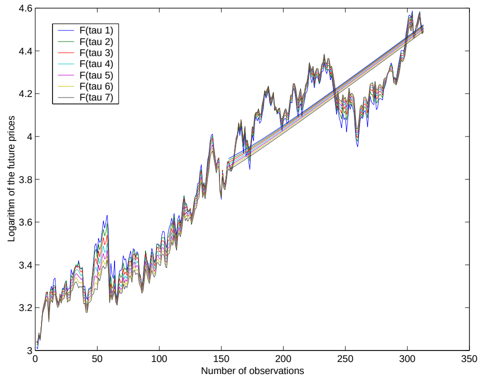


Figure 4.23: Forecast over half a sample of the log of the future prices, $\mathbf{F}_t(\tau_1)$,...,$\mathbf{F}_t(\tau_7)$.

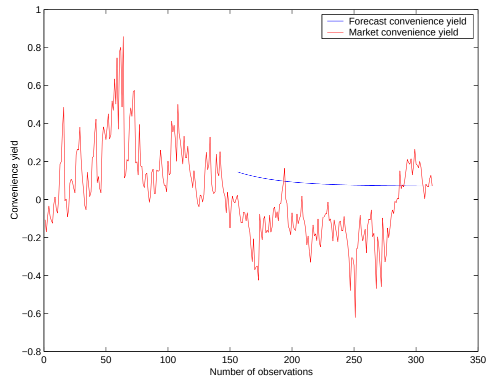


Figure 4.24: Forecast for CY over half a sample.

As can be seen in Figures 4.23 and 4.24, the forecast from the EKF is equal to the forecasts depicted in Figures 4.11 and 4.12.

4.4 Conclusion

In this chapter we apply both the Kalman filter and the extended Kalman filter on the OU process for the convenience yield and a GBM for the spot price. The reason why the Kalman filter is chosen to be the iterative procedure is due to unobservability of the state variables S_t and δ_t. The Kalman filter links these latent variables to the observed variables, in this case the furture prices. As noticed earlier, the affine form of the closed form solution of the future prices gives us immediately the measurement equation (4.3). Also the linearity of this closed form solution in the state variable S_t is an important requirement in order to have an optimal result of the Kalman filter. The covariance-variance matrix given in (4.7) does not depend on the state variables and because of that, we did not need to adjust the Kalman filter algorithm.
Although the convenience yield and spot price are unobservable we need an initial choice of these variables to start the iterative procedure. We assumed the $S_0 = F(0,T)$ and the δ_0 is calculated via (4.9). By doing this, there is an error involved. However, we tested the Kalman filter by setting $\alpha_0 = [10, \quad -10]$, a totally inadequate initial value for both the log of the spot price and the convenience yield. The log-likelihood remained the same, as well as the parameters given in Table 4.1. The only difference is that the value of the log-likelihood function jumped in the beginning. This can be explained by these

incorrect initial values.
The results presented in Table 4.1 show the stability of the procedure (different initial values of the parameter set converges to the same optimized parameter set with the same value of the log-likelihood function). Alhough there are no unrealistic values of these parameters, one can not tell if they are acceptable as they are parameters of unobservable processes. However, looking at Figure 4.1, we see that the calibrated future prices fit the market future prices succesfully. In Figure 4.2 and 4.14, as noted before, we see that the calibrated convenience yield is not following the path of the market convenience yield. This can have several reasons. First of all, we compared it to our calculated convenience yield (4.9) but there are more factors included in the market future prices than the interest rates, spot prices and convenience yield. In fact, the market convenience yield actually consists of numerous terms, e.g. storage costs, security costs, transport costs etc. So the comparison is expected to be not succesfully as we do not compare the pure convenience yield but, what is called, the total cost of carry.
As from Section 4.4 we tested the extend Kalman filter for the OU process. Since the future prices are linear in the state variable S_t, we can expect larger errors than in the previous sections. This is caused by the linearization of the system matrices. The EKF implicitly introduces an extra approximation error in the filtering procedure. However, as we can see from the results (Table 4.3 and Figure 4.13) the method is stable and we are still able to fit the calibrated future prices to the market future prices very succesfully.

Chapter 5

The Kalman filter model for the Cox-Ingersoll-Ross process

In this chapter the KF model for the CIR process is discussed. An essential requirement for the KF to give an optimal result is the Gaussian assumption. Since the CIR process is non-Gaussian we expect quasi-optimal results. Together with the nonnegativity constraint, applying the KF may give some problems as we shall see.

5.1 Setup of the Kalman filter applied to the CIR process

In state variable terms, (3.8) can be written in

$$\ln F_t(\tau) = x_t + A(\tau) - B(\tau)\delta_t, \tag{5.1}$$

with $x_t = \log S_t$ and where $A(\tau)$ and $B(\tau)$ are given by (3.12) and (3.11). Again, the vector $\boldsymbol{\alpha}_{t_i}$ contains the state variables we want to estimate. From (3.1) and (3.6) the transition equation immediately follows

$$\begin{bmatrix} x_{t_i} \\ \delta_{t_i} \end{bmatrix} = \begin{bmatrix} \mu\Delta t \\ k\alpha\Delta t \end{bmatrix} + \begin{bmatrix} 1 & -(1+\frac{1}{2}\sigma_S^2)\Delta t \\ 0 & 1-k\Delta t \end{bmatrix} \begin{bmatrix} x_{t_{i-1}} \\ \delta_{t_{i-1}} \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \xi_{t_i}^1 \\ \xi_{t_i}^2 \end{bmatrix},$$

which for simplicity will be again written as

$$\boldsymbol{\alpha}_{t_i} = \mathbf{c}_{t_i} + \mathbf{Q}_{t_i}\boldsymbol{\alpha}_{t_{i-1}} + \mathbf{R}\boldsymbol{\xi}_{t_i}, \quad i = 1, ..N. \tag{5.2}$$

Again $\boldsymbol{\xi}_{t_i}$ has expectation zero and its covariance matrix $\mathbf{V}_{t_i}$ is given by

$$\mathbf{V}_{t_i} = \begin{bmatrix} \mathbb{E}[(x_{t_i} - \mathbb{E}[x_{t_i}])^2 | x_{t_{i-1}}] & \mathbb{E}[(x_{t_i} - \mathbb{E}[x_{t_i}])(\delta_{t_i} - \mathbb{E}[\delta_{t_i}]) | x_{t_{i-1}}, \delta_{t_{i-1}}] \\ \mathbb{E}[(x_{t_i} - \mathbb{E}[x_{t_i}])(\delta_{t_i} - \mathbb{E}[\delta_{t_i}]) | x_{t_{i-1}}, \delta_{t_{i-1}}] & \mathbb{E}[(\delta_{t_i} - \mathbb{E}[\delta_{t_i}])^2 | \delta_{t_{i-1}}] \end{bmatrix}$$

$$= \begin{bmatrix} \mathbb{V}\mathrm{ar}[x_{t_i} | x_{t_{i-1}}] & \sqrt{\mathbb{V}\mathrm{ar}[x_{t_i} | x_{t_{i-1}}]}\sqrt{\mathbb{V}\mathrm{ar}[\delta_{t_i} | \delta_{t_{i-1}}]} \\ \sqrt{\mathbb{V}\mathrm{ar}[x_{t_i} | x_{t_{i-1}}]}\sqrt{\mathbb{V}\mathrm{ar}[\delta_{t_i} | \delta_{t_{i-1}}]} & \mathbb{V}\mathrm{ar}[\delta_{t_i} | \delta_{t_{i-1}}] \end{bmatrix}$$

$$= \begin{bmatrix} \sigma_S^2 \Delta t \delta_{t_{i-1}} & \rho\sigma_S\sqrt{\Delta t}\sqrt{\delta_{t_{i-1}}}\sqrt{\mathbb{V}\mathrm{ar}[\delta_{t_i} | \delta_{t_{i-1}}]} \\ \rho\sigma_S\sqrt{\Delta t}\sqrt{\delta_{t_{i-1}}}\sqrt{\mathbb{V}\mathrm{ar}[\delta_{t_i} | \delta_{t_{i-1}}]} & \mathbb{V}\mathrm{ar}[\delta_{t_i} | \delta_{t_{i-1}}] \end{bmatrix},$$

where

$$\mathbb{V}\mathrm{ar}[\delta_{t_i} | \delta_{t_{i-1}}] = \alpha(\frac{\sigma_\delta^2}{2k})(1 - e^{-k\Delta t})^2 + \delta_{t_{i-1}}(\frac{\sigma_\delta^2}{k})(e^{-k\Delta t} - e^{-2k\Delta t}). \tag{5.3}$$

The difference between $\mathbf{V}_{t_i}$ for the OU process and the CIR process is that in the latter the $\mathbf{V}_{t_i}$ depends on the state variable $\delta_{t_{i-1}}$. This error matrix should be updated every timestep. In the following section the pseudo code for the iterative procedure is given.

5.1.1 Iterative procedure

The iterative procedure consists of two parts. Part one is the Kalman filter which is used to update the system matrices. Part two is estimating the parameter set φ. The algorithm of the Kalman filter is given in pseudo code. The CIR model does not allow negative convenience yields. Therefore we also have to extend the Kalman filter in order to exclude negative convenience yields. For this, we simply replace any negative element of the estimate δ_{t_i} by zero. The reason why this is done is just a matter of convenience and it is rather difficult to impose the positivity constraint in the estimation procedure. However, this solution will turn out to have several consequences as we shall see.

while (number of iterations has not been reached , optimal φ has not been found) **do**
  (Kalman filter)
  for $i = 1 : N$ **do**
    (Prediction)
    $\tilde{\boldsymbol{a}}_{t_i} = \mathbf{Q} * \mathbf{a}_{t_{i-1}} + \mathbf{c}$;
    $\tilde{\mathbf{P}}_{t_i} = \mathbf{Q} * \mathbf{P}_{t_{i-1}} * \mathbf{Q}' + \mathbf{R} * \mathbf{V} * \mathbf{R}'$;
    $\tilde{\mathbf{y}}_{t_i} = \mathbf{Z} * \tilde{\mathbf{a}}_{t_i} + \mathbf{d}$;
    (Innovations)
    $\mathbf{v}_{t_i} = \mathbf{y}_{t_i} - \tilde{\mathbf{y}}_{t_i}$;
    (Updating)
    $\mathbf{G}_{t_i} = \mathbf{Z} * \tilde{\mathbf{P}}_{t_i} * \mathbf{Z}' + \mathbf{H}$;
    $\mathbf{a}_{t_i} = \tilde{\mathbf{a}}_{t_i} + \tilde{\mathbf{P}}_{t_i} * \mathbf{Z}' * \mathbf{G}_{t_i}^{-1} * \mathbf{v}_{t_i}$;
    $\mathbf{P}_{t_i} = \tilde{\mathbf{P}}_{t_i} - \tilde{\mathbf{P}}_{t_i} * \mathbf{Z}' * \mathbf{G}_{t_i}^{-1} * \mathbf{Z} * \tilde{\mathbf{P}}_{t_i}$;
    if $det\mathbf{G}_{t_i} \leqslant 0$ **then**
      $det\mathbf{G}_{t_i} t = 10^{-10}$;
    end if
    (Log-likelihood function per iteration)
    $logl(i) = -(n/2) * \ln(2 * \pi) - 0.5 * \ln(det\mathbf{G}_{t_i}) - 0.5 * \mathbf{v}_{t_i}' * \mathbf{G}_{t_i}^{-1} * \mathbf{v}_{t_i}$;
    $Log\mathbf{L} = \sum_i logl(i)$;
    if $\tilde{\mathbf{a}}_{t_i}(2)<0$ **then**
      $\tilde{\mathbf{a}}_{t_i}(2)=0$; ($\tilde{\mathbf{a}}_{t_i}(2)$ corresponds to δ_{t_i})
    end if
    (update matrix $\mathbf{V}_{t_i}$)
  end for
  (Adjustment for φ via a Matlab optimization routine)
end while

The same procedure to optimize the parameter set φ holds and is explained in Section 4.1. The pseudo code given above differs from the standard KF as discussed in Section 4.6 in two parts: A negativity constraint for the convenience yield is added and the matrix $\mathbf{V}_{t_i}$ is updated in each step because it depends on the state variable $\delta_{t_{i-1}}$.

5.2 Numerical results for the Cox-Ingersoll-Ross process

5.2.1 Choosing the right start value for the convenience yield

We use the same data as in Chapter 4. The initial value of $\delta_0^{\text{market}} = -0.1063$. Such a start value, however, is not allowed in the CIR framework. Therefore we try three approaches to tackle the problem of choosing the correct starting value for the convenience yield.

Case one. We set the starting value for δ_0^{market} equal to zero. This implies that $\mathbf{V}_{t_i}$ contains zero elements, i.e.

$$\mathbf{V}_{t_i} = \begin{bmatrix} 0 & 0 \\ 0 & 0.7784e^{-6} \end{bmatrix}. \tag{5.4}$$

These initial values are used in every iteration. We choose the initial parameter set equal to the optimized parameter set given in Table 4.1. The diagonal elements $\{h_j\}_{j=1}^{n}$ are initially set equal to the variance of the data. The standard errors are extremely high. Since $\mathbf{V}_{t_i}$ depends strongly on the parameters k, σ_δ and σ_S we can expect that the optimalization routine can not optimize the parameters because $\mathbf{V}_{t_i}$ is almost zero for every σ_δ, σ_S and k. As a consequence the parameters σ_δ, λ and k are retained on the upper bound 10. This is certainly not an optimal result.

Parameters	Ini parset	Opti parset
k	1.4221	10.0000 (48.4424)
μ	0.3733	0.8616 (3.9825)
α	0.0699	1.1192 (3.1169)
λ	-0.0183	10.0000 (82.4541)
σ_S	0.3630	0.4219 (9.7882)
σ_δ	0.4028	10.0000 (7.8778)
ρ	0.8378	0.2196 (3.3102)
$\|h_1\|$	0.0246	0.0609 (0.0217)
$\|h_2\|$	0.0268	0.0382 (0.0108)
$\|h_3\|$	0.0291	0.0231 (0.0013)
$\|h_4\|$	0.0313	0.0109 (0.0031)
$\|h_5\|$	0.0336	0.0000 (0.0015)
$\|h_6\|$	0.0357	0.0096 (0.0016)
$\|h_7\|$	0.0377	0.0181 (0.0031)
Log-Likelihood		5104.3040

Table 5.1: Optimized parameter set (Opti parset) for different initial parameter sets (Ini parset). Standard errors in parentheses.

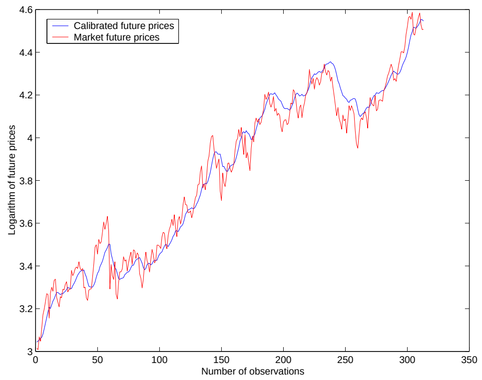


Figure 5.1: Comparison between the calibrated future prices and market future prices for contract τ_1.

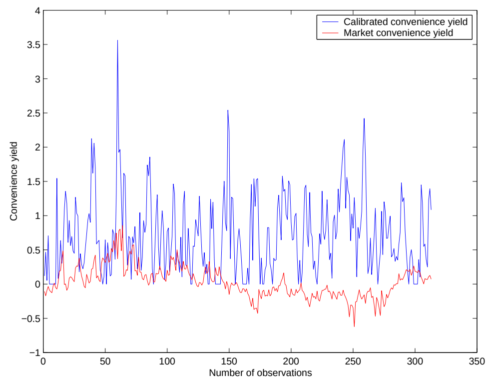


Figure 5.2: Comparison between the calibrated δ_t and the market implied convenience yield.

We can see from Figure 5.1 that the estimated future prices are quite satisfactory. However, if we look at Figure 5.2 we see that the calibrated convenience yield differs significantly from the market convenience yield. This can be easily explained by the fact that the convenience yield volatility stays at $\sigma_\delta = 10$ during the filtering.

Case two. As an alternative approach, we can shift δ_0^{market} by a constant. Table 5.2 shows the calibrated parameters. For this calibration we set $\delta_0^{\text{market}} = 0.3$. As we can see, the three parameters are now optimized.

Parameters	Ini parset	Opti parset
k	1.4221	1.6608 (0.0455)
μ	0.3733	0.3254 (0.1325)
α	0.0699	0.0730 (0.0841)
λ	-0.0183	-0.0754(0.1405)
σ_S	0.3630	0.6574 (0.0255)
σ_δ	0.4028	0.7250 (0.0293)
ρ	0.8378	0.8352 (0.0166)
$\lvert h_1 \rvert$	0.0246	0.0609 (0.0217)
$\lvert h_2 \rvert$	0.0268	0.0382 (0.0108)
$\lvert h_3 \rvert$	0.0291	0.0231 (0.0013)
$\lvert h_4 \rvert$	0.0313	0.0109 (0.0031)
$\lvert h_5 \rvert$	0.0336	0.0000 (0.0015)
$\lvert h_6 \rvert$	0.0357	0.0096 (0.0016)
$\lvert h_7 \rvert$	0.0377	0.0181 (0.0031)
Log-Likelihood		8742.6128

Table 5.2: Optimized parameter set (Opti parset) for different initial parameter sets (Ini parset). Standard errors in parentheses.

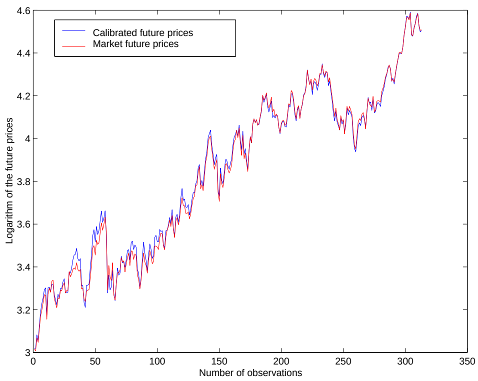


Figure 5.3: Comparison between the calibrated future prices and market future prices for contract τ_1.

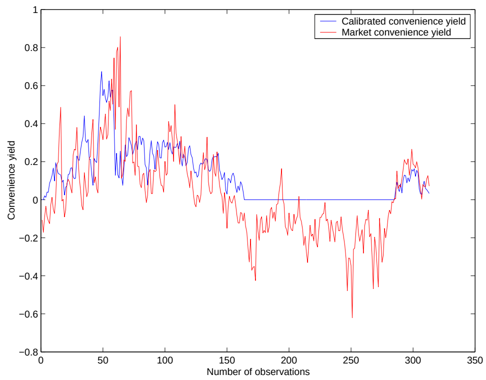


Figure 5.4: Comparison between the calibrated δ_{t_i} and the market convenience yield, $\delta_{t_i}^{\text{market}}(\tau_1, \tau_2)$.

From Figure 5.4 it is clear that the KF can not generate the convenience yield because of the nonnegativity constraint. This constraint implies that the CY equals zero for a long period of time. Note, however, that the future prices (of contract τ_1) are filtered very well. The fact that $\delta_{t_i} = 0$ for extensive periods makes this appoach basically useless.

Case three. Finally, one can also shift the data in time such that $\delta_0^{\text{market}} > 0$. We shift the data 14 weeks ahead (see Figure 5.2). Results are given in Table 5.3. We will test its robustness by changing the initial parameter set. If we compare the optimized parameter set of Table 5.2 with that of Table 5.3 we do

Parameters	Ini parset	Opti parset	Ini parset	Opti parset	Ini parset	Opti parset
k	1.4221	1.6118 (0.0389)	2	1.6118 (0.0485)	1.4	1.6118 (0.0285)
μ	0.3733	0.2932 (0.1442)	0.5	0.2932 (0.1321)	0.2	0.2932 (0.1221)
α	0.0699	0.0590 (0.0973)	0.1	0.0590 (0.0843)	0.1	0.0590 (0.0880)
λ	-0.15	-0.1029(0.1571)	-0.1	-0.1029(0.1360)	0.2	-0.1029(0.1369)
σ_S	0.3630	0.5864 (0.0247)	0.3	0.5864 (0.0249)	0.5	0.5864 (0.0239)
σ_δ	0.4028	0.6598 (0.0274)	0.3	0.6598 (0.0249)	0.3	0.6598 (0.0262)
ρ	0.8378	0.8357 (0.1442)	0.4	0.8357 (0.0179)	0.6	0.8357 (0.0190)
$\|h_1\|$	0.0246	0.0189 (0.0008)	0.0246	0.0189 (0.0008)	0.1	0.0189 (0.0008)
$\|h_2\|$	0.0268	0.0073 (0.0003)	0.0268	0.0073 (0.0003)	0.1	0.0073 (0.0003)
$\|h_3\|$	0.0291	0.0022 (0.0001)	0.0291	0.0022 (0.0001)	0.1	0.0022 (0.0001)
$\|h_4\|$	0.0313	0.0000 (0.0001)	0.0313	0.0000 (0.0001)	0.1	0.0000 (0.0001)
$\|h_5\|$	0.0336	0.0006 (0.0000)	0.0336	0.0006 (0.0000)	0.1	0.0006 (0.0000)
$\|h_6\|$	0.0357	0.0000 (0.0001)	0.0357	0.0000 (0.0001)	0.1	0.0000 (0.0001)
$\|h_7\|$	0.0377	0.0015 (0.0001)	0.0377	0.0015 (0.0001)	0.1	0.0015 (0.0001)
Log-Likelihood		8341.0028		8341.0028		8341.0028

Table 5.3: Optimized parameter set (Opti parset) for different initial parameter sets (Ini parset). Standard errors in parentheses.

not see any significant differences in the parameters. This can also be seen when comparing Figures 5.3 and 5.4 with Figures 5.5 and 5.6 respectively (there is hardly any difference). Despite the fact that the CY equals zero for a large period of time, the Kalman filter is stable. Different initial parameter sets will converge to the same optimized parameter set (see Table 5.3). But again, the fact that $\delta_{t_i} = 0$ for extensive periods makes this approach useless. If we compare Figure 4.2 and 5.6, we see that both processes are not able to fit the market convenience yield very well. As noted earlier, we can conclude that the market convenience yield is not a pure convenience yield, but can be interpreted as the total cost of carry. For each futures contract we plot the innovation. Again, since the mean of the innovations

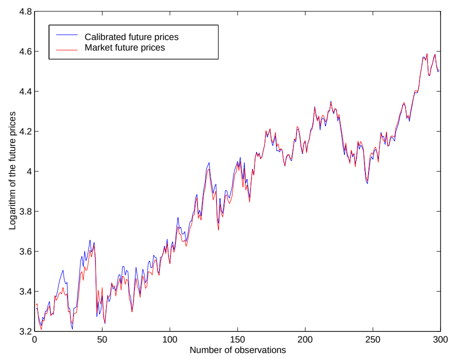


Figure 5.5: Comparison between the calibrated future prices and market future prices for contract τ_1.

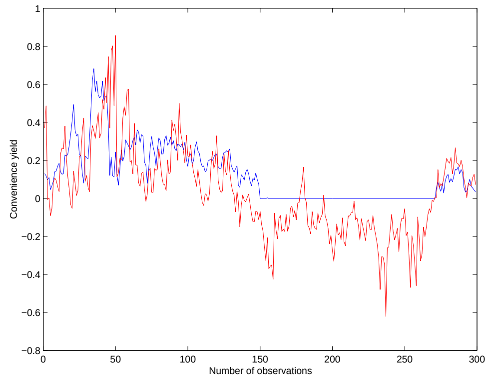


Figure 5.6: Comparison between the calibrated δ_{t_i} and the market convenience yield, $\delta_{t_i}^{\text{market}}(\tau_1, \tau_2)$.

is nearly zero, the assumption of zero-mean normal distributed error terms is acceptable.

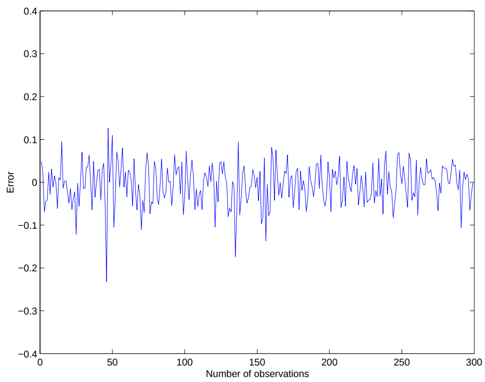


Figure 5.7: Innovation corresponding to $\mathbf{F}_t(\tau_1)$. Mean = -8.2412e-004, Variance = 0.0003.

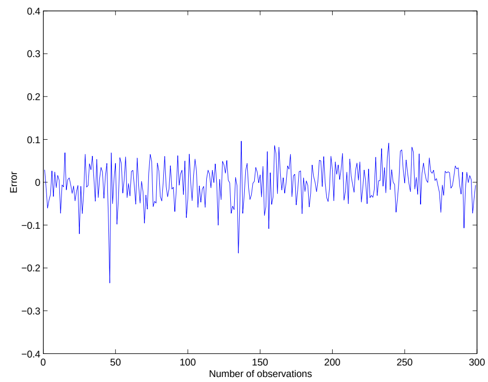


Figure 5.8: Innovation corresponding to $\mathbf{F}_t(\tau_2)$. Mean = 7.6238e-004, Variance = 0.0009.

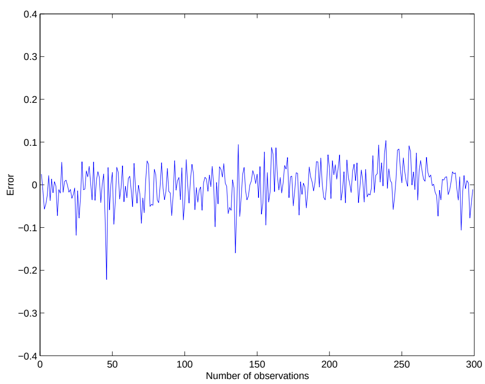


Figure 5.9: Innovation corresponding to $\mathbf{F}_t(\tau_3)$. Mean = 6.1235e-004, Variance = 0.0007.

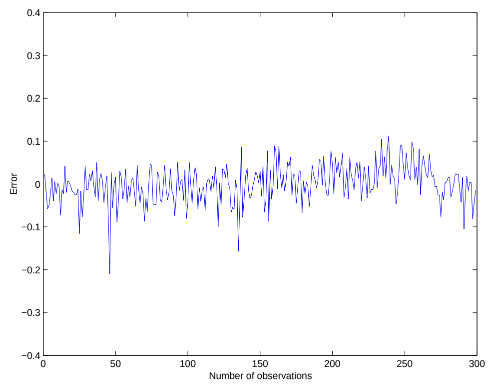


Figure 5.10: Innovation corresponding to $\mathbf{F}_t(\tau_4)$. Mean = -1.0181e-004, Variance = 0.0011.

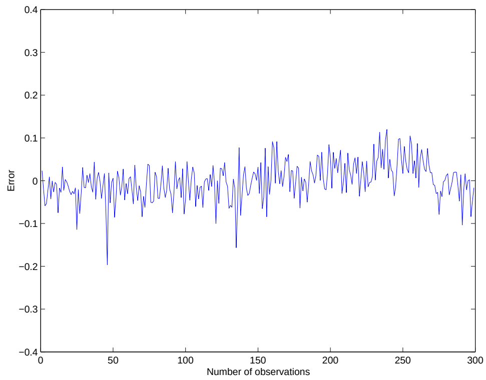


Figure 5.11: Innovation corresponding to $\mathbf{F}_t(\tau_5)$. Mean = -5.5711e-004, Variance = 0.0020.

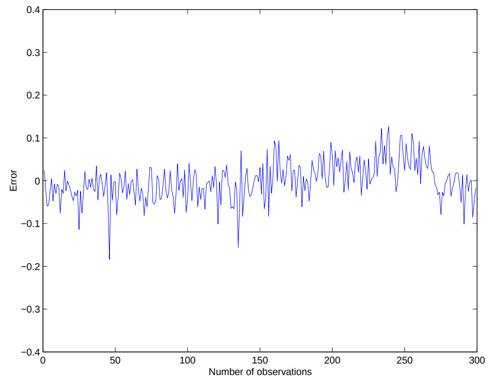


Figure 5.12: Innovation corresponding to $\mathbf{F}_t(\tau_6)$. Mean = -3.4434e-004, Variance = 0.0010.

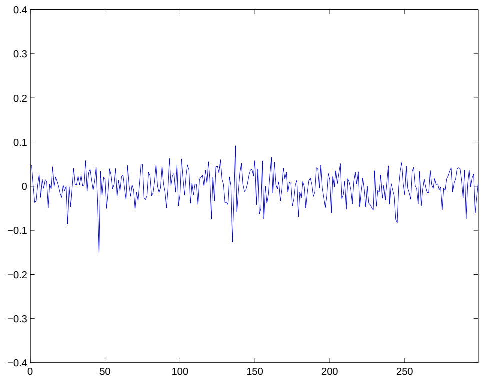


Figure 5.13: Innovation corresponding to $\mathbf{F}_t(\tau_7)$. Mean = 6.1123e-004, Variance = 0.0012.

5.2.2 Parameter explanation

If we look at Table 5.3, we see a correlation factor of 0.8357. This is the correlation between the 'true' state variables (i.e. including the artificial nonnegativity constraint of the convenience yield). Since we only test if the estimate of the convenience yield is zero and not the update (see pseudo code), we can measure the correlation of the 'actual' state variables (i.e. including negative values of the convenience yield). In Figure 5.14 we see the scatterplot of the increments of the 'actual' state variables and in Figure 5.15 we see the scatterplot between the increments of the 'true' state variables presented in Figures 5.5 and 5.6. The dots with correlation zero can be explained by the fact that if the convenience yield is set to zero the correlation between $\Delta\delta_t$ and Δx_t is therefore zero. The 'actual' correlation is 0.4995.

5.3 Conclusion

We expected, due to the non-Gaussianity of the CIR process, quasi-optimal results for the Kalman filter. The covariance-variance matrix given in (5.2) depends on the state variable δ_t. Since δ_t can not be negative, we simply adjust the Kalman filter by setting δ_{t_i} zero if it is negative. However, if $\delta_0 = 0$, the matrix $\mathbf{V}_{t_i}$ will be almost zero as well and from Table 5.1 we can see that this will not give an optimal result. Shifting δ_0 such that it is positive gives an acceptable result. In Figure 5.3 we can see that the calibrated future prices fit the market future prices succesfully, but the fact that $\delta_{t_i} = 0$ for extensive

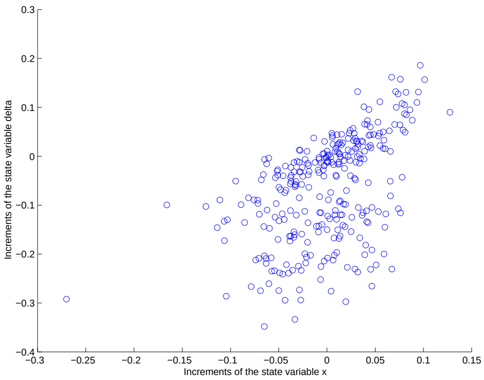


Figure 5.14: Scatterplot of the actual state variables (including negative convenience yield) $(\Delta x_t, \Delta\delta_t)$ for 299 observations.

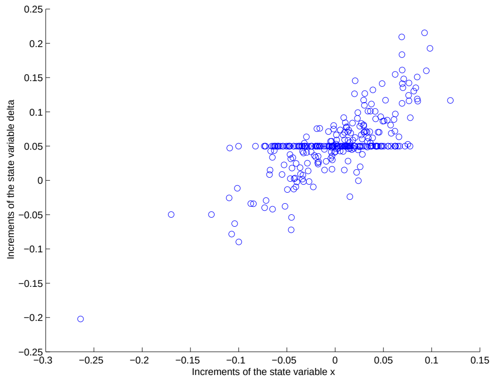


Figure 5.15: Scatterplot of the true state variables $(\Delta x_t, \Delta\delta_t)$ for 299 observations.

periods makes this approach basically useless. The same can be said about case 3, where we shift the data set such that $\delta_0^{\text{market}} > 0$. From Table 5.3 we see that the procedure is stable but the same arguments hold for the convenience yield. In short, the choice of process for the convenience yield strongly depends on the data period of the commodity. In the following chapter we price put options on future prices. Because of the negative results of Chapter 5 we assume the convenience yield to follow an OU process as given in Chapter 2 and 4.

Chapter 6

Pricing options on futures

In this chapter a closed form solution for a European call option on a futures contract with strike price K and maturity T will be derived via the PDE given in equation (2.32). More generally, when considering stochastic interest rate, Appendix B gives a detailed description of the derivation of the same form of the closed form solution via a change of measure etc. Furthermore, with the use of this closed form solution and the calibrated future prices, put prices are calibrated to the market put prices via the extended Kalman filter. It turns out that parametrization of the volatility of the spot processes, S_t and δ_t is a challenging difficulty and we conclude that one needs a matured contract to do so. At the end of this chapter, both initially assumed stochastic processes for the spot price and convenience yield are retrieved by starting from the future prices process.

6.1 Deriving the closed form solution of a K-strike European call option on a futures contract

Consider again the PDE given in (2.32) and assume

$$V(S,\delta,t) = \bar{V}(F,t), \tag{6.1}$$

calculating all derivatives and inserting these in (2.32) gives

$$\begin{aligned}&\bar{V}_t + \{-A'(\tau) - B'(\tau)\} F\bar{V}_F + \{k(\alpha-\delta) - \lambda\sigma_\delta\} B(\tau)F\bar{V}_F + \{r-\delta\} F\bar{V}_F + \rho\sigma_S\sigma_\delta F^2 B(\tau)\bar{V}_{FF} \\ &+\rho\sigma_S\sigma_\delta B(\tau)F\bar{V}_F + \frac{1}{2}\sigma_S^2 F^2\bar{V}_{FF} + \frac{1}{2}\sigma_\delta^2 B^2(\tau)F^2\bar{V}_{FF} - r\bar{V} = 0.\end{aligned} \tag{6.2}$$

Together with the former differential equation presented in (2.36) it holds that

$$\bar{V}_t + \left\{ \underbrace{r - A'(\tau) + k\alpha B(\tau) - \lambda\sigma_S B(\tau) + \rho\sigma_S\sigma_\delta B(\tau)}_{=0,\ \text{cf.}(2.36)_2} \right\} F\bar{V}_F + \frac{1}{2}\sigma_F^2(t,T)F^2\bar{V}_{FF} - r\bar{V} = 0, \tag{6.3}$$

with

$$\sigma_F^2(t,T) = \sigma_S^2 + \sigma_\delta^2 B^2(t,T) + 2\rho\sigma_S\sigma_\delta B(t,T). \tag{6.4}$$

Equation (6.3) is recognized as the Black equation. Suppose now that the payoff is given by a call, i.e. $\bar{V}(F,T) = \max(F-K)^+$. Then the solution is given by

$$\bar{V}(F,t) = e^{-r(T-t)}F\mathcal{N}(d_1) - e^{-r(T-t)}K\mathcal{N}(d_2), \tag{6.5}$$

with

$$d_1 = \frac{\ln(F/K) + \frac{1}{2}\hat{\sigma}_F^2(T-t)}{\hat{\sigma}_F\sqrt{T-t}}, \quad d_2 = \frac{\ln(F/K) - \frac{1}{2}\hat{\sigma}_F^2(T-t)}{\hat{\sigma}_F\sqrt{T-t}}$$

and

$$\begin{aligned} \hat{\sigma}_F^2(t,T) &= \frac{1}{T-t}\int_t^T \sigma_F^2(t,u)du \\ &= \frac{1}{T-t}\left\{\frac{2\rho\sigma_S\sigma_\delta}{k^2}(1-\theta) - \frac{2\rho\sigma_S\sigma_\delta}{k}(T-t) + \sigma_S^2(T-t) + \frac{\sigma_\delta^2}{2k^3}(1-\theta^2) - \frac{2\sigma_\delta^2}{k^3}(1-\theta) + \frac{\sigma_\delta^2}{k^2}(T-t)\right\}. \end{aligned} \tag{6.6}$$

Alternatively, we can derive the same formula as given in (6.5) by considering (2.12):

$$V_t[Y_T] := e^{-r(T-t)}\mathbb{E}[Y_T|\mathscr{F}_t]. \tag{6.7}$$

The value of a European call option at time t is given by

$$V_t[(S_T-K)^+] = e^{-r(T-t)}\mathbb{E}[(S_T-K)^+] \tag{6.8}$$

Combining (2.19) and (6.8) we obtain

$$V_t[(S_T-K)^+] = \mathbb{E}[(S_t e^z - e^{-r(T-t)}K)^+]. \tag{6.9}$$

Since e^z is normally distributed with mean $\hat{\mu}$ and variance $\hat{\sigma}$, we can compute the expectation in (6.9) explicitly.

$$V_t[(S_T-K)^+] = \exp\left\{\hat{\mu} + \frac{1}{2}\tilde{\sigma}\right\} S_t\mathscr{N}(\hat{d}_1) - \exp\{-r(T-t)\}K\mathscr{N}(\hat{d}_2), \tag{6.10}$$

with

$$\hat{d}_1 = \frac{\ln(S_t/K) + r(T-t) + \hat{\mu} + \tilde{\sigma}^2}{\tilde{\sigma}}, \quad \hat{d}_2 = \frac{\ln(S_t/K) + r(T-t) + \hat{\mu}}{\tilde{\sigma}}. \tag{6.11}$$

The parameters $\hat{\mu}$ and $\tilde{\sigma}^2$ are given by (2.20) and (2.27) respectively. Since

$$V_t[S_T] = S_t \exp\left\{\hat{\mu} + \frac{1}{2}\tilde{\sigma}^2\right\}, \tag{6.12}$$

we can rewrite (6.10) in

$$V_t[(S_T-K)^+] = V_t[S_T]\mathscr{N}(d1) - e^{-r(T-t)}K\mathscr{N}(d_2), \tag{6.13}$$

and

$$d_1 = \frac{\ln(V_t[S_T]/K) + r(T-t) + \frac{1}{2}\tilde{\sigma}^2}{\tilde{\sigma}}, \quad d_2 = \frac{\ln(V_t[S_T]/K) + r(T-t) - \frac{1}{2}\tilde{\sigma}^2}{\tilde{\sigma}}.$$

Obviously $\tilde{\sigma}^2$ is given by $\hat{\sigma}_F^2(T-t)$. Using definition (2.30) and $V(S,\delta,t) = V_t[Y_T]$ in (6.13) we arrive at expression (6.5).

For every observed future option price, we can solve formula (6.5) for $\hat{\sigma}_F$. This specific value for the volatility multiplied by $\sqrt{T-t}$ will from this point on be referred to as the real dynamic volatility and it will be denoted by $\sigma_{\mathrm{RD}}(t,T)$.

6.2 Calibrating the model to market put prices

Since for the data used in Chapter 4,5 and 6 we do not have corresponding put prices, we use 3 futures contracts on the same commodity (light crude oil) with different maturities. They are denoted by M8, Z8 and M9. Instead of weekly observations we now have daily observations. The number of observations is 79. One important remark is that the maturity of the option is the same as the maturity of the futures contract!

Contract	Date at t=0	Date at t=T	Last observation day	Observed trading days	Total trading days
M8	20/11/2007	20/05/2008	17/03/2008	79	125
Z8	20/11/2007	20/11/2008	17/03/2008	79	258
M9	20/11/2007	19/05/2009	17/03/2008	79	386

Table 6.1: Contract details.

Puts on contracts M9 and Z8 with three different strikes (70,75 and 80) will generate three different put options prices. Contract M8 only has two put options with strikes 80 and 85. In this section we investigate how well our optimized parameter set, which we derived from the Kalman filter, fits the put option market prices. We insert the parameters in the put option formula and compare these prices to the market prices. The Kalman filter will calibrate the parameters committed to these contracts. The calibrated values of the future prices and the volatilities of the two stochastic processes (spot price and the convenience yield) are then used in formula (6.5). From the put-call parity it follows

$$\mathcal{P} = \mathcal{C} + Ke^{-r(T-t)} - F(0,T)e^{-r(T-t)}, \tag{6.14}$$

with obvious notation. Note that $\mathcal{P}$ is the price of a put option with maturity time T and strike price K written on a futures contract which expires at time T.

6.2.1 Setting up the calibration for the put option prices

To price the put options we use again the extended Kalman filter, which is explained in Section 4.3. More specifically, we have

- $\mathbf{Q}(\boldsymbol{\alpha}_{t_{i-1}}) = \mathbf{F}_{t_i}(\tau_j) \quad j = 1,..,n, \quad i = 1,..,N,$
- $\mathbf{R}(\boldsymbol{\alpha}_{t_{i-1}}) = 1,$
- $\mathbf{Z}(\boldsymbol{\alpha}_{t_i}) = \mathcal{P},$
- $\hat{\mathbf{Z}}(\boldsymbol{\alpha}_{t_i}) = \frac{\partial \mathcal{P}}{\partial F_{t_i}(\tau_j)},$
- $\hat{\mathbf{T}}(\boldsymbol{\alpha}_{t_i}) = 1,$
- $\hat{\mathbf{R}}(\boldsymbol{\alpha}_{t_i}) = 0,$

- $\mathbf{H} = h_1$.

The first contract expires after 182 calender days (125 trading days) on the first observation. τ is set to 182/360 at this day, since the system matrices are calculated annually. After day two, τ is set to 181/360 etc. Weekends are excluded. So for each observation we have a $\hat{\sigma}_F$.

6.2.2 Results following a naïve approach

The KF is set up in such a way that it incorporates all market future prices in the calibration of all model parameters in the set φ. In this approach, the parameter dependency on the strikes K and maturities T is fully disregarded. In Figures 6.1-6.3 the calibrated future prices against the market future prices are presented. In Table 6.1 the optimized parameter set is shown. The initial values of the parameters are the optimized parameter set from Table 4.1. For the $\{h_j\}_{j=1}^n$ we used the variance of the logarithm of the future prices. We continue with the valuation of put options using the optimized parameter set

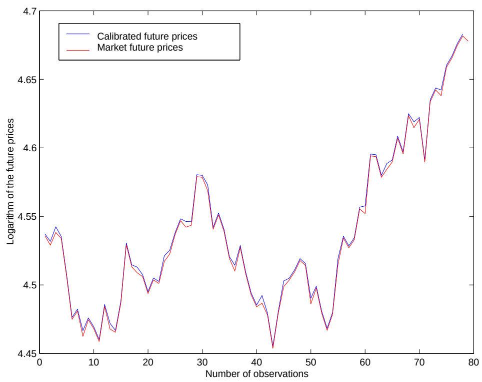


Figure 6.1: Comparison between the calibrated future prices and market future prices for contract M8.

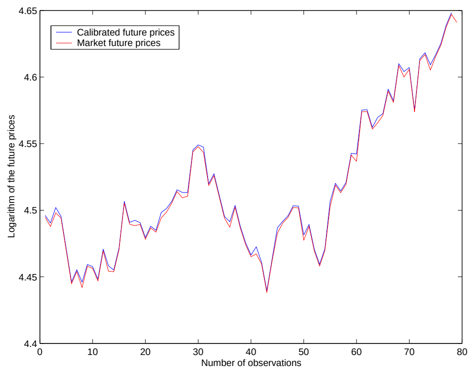


Figure 6.2: Comparison between the calibrated future prices and market future prices for contract Z8.

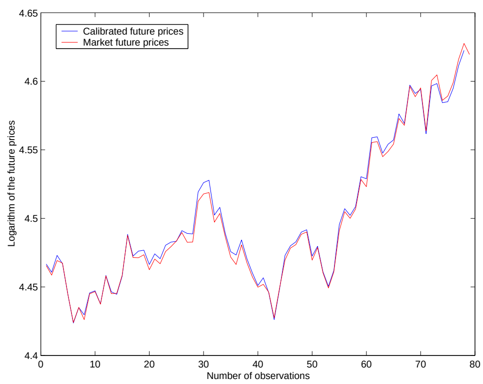


Figure 6.3: Comparison between the calibrated future prices and market future prices for contract M9.

presented in Table 6.2. The put prices are computed by substituting the calibrated future prices, together with all other parameters in Table 6.2, in the put formula given by (6.14).

Parameters	Ini parset	Opti parset
k	1.4221	0.8046 (0.1012)
μ	0.3733	0.5374 (0.4453)
α	0.0699	0.0851 (0.0169)
λ	-0.0183	-0.0754(0.1405)
σ_S	0.3630	0.3493 (0.0305)
σ_δ	0.4028	0.1981 (0.0234)
ρ	0.8378	0.7171 (0.0593)
$\lvert h_1 \rvert$	0.0341	0.0000 (0.0005)
$\lvert h_2 \rvert$	0.0300	0.0000 (0.0005)
$\lvert h_3 \rvert$	0.0291	0.0027 (0.0002)
Log-Likelihood		902.7301

Table 6.2: Optimized parameter set (Opti parset). Standard errors in parentheses.

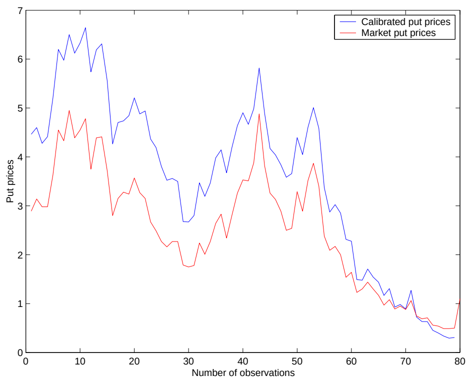


Figure 6.4: Comparison between the calibrated put prices and market put prices for contract M8 with K=85.

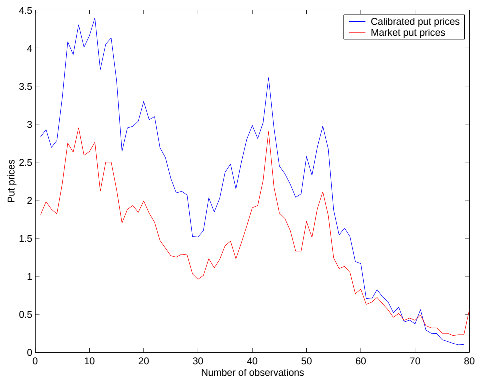


Figure 6.5: Comparison between the calibrated put prices and market put prices for contract M8 with K=80.

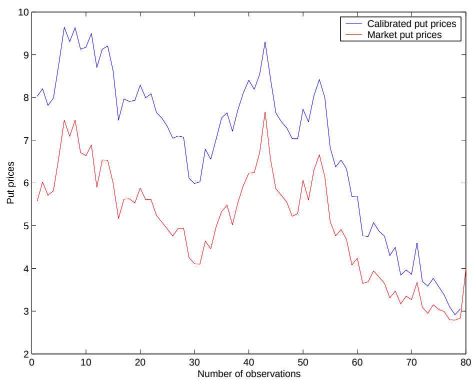


Figure 6.6: Comparison between the calibrated put prices and market put prices for contract Z8 with K=85.

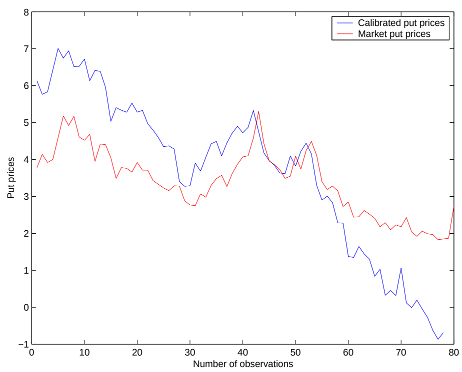


Figure 6.7: Comparison between the calibrated put prices and market put prices for contract Z8 with K=80.

From these figures, it can be concluded that the calibrated put prices are not acceptable. As to be expected, the spot volatilities depend on both the strike prices and the time to maturity.

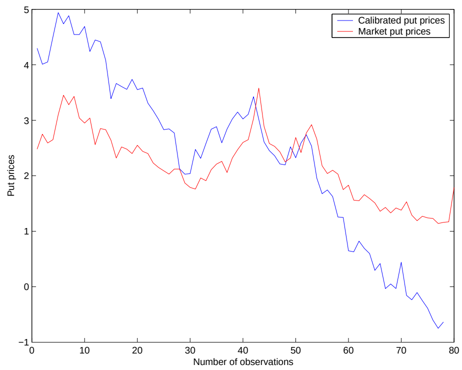


Figure 6.8: Comparison between the calibrated put prices and market put prices for contract Z8 with K=75.

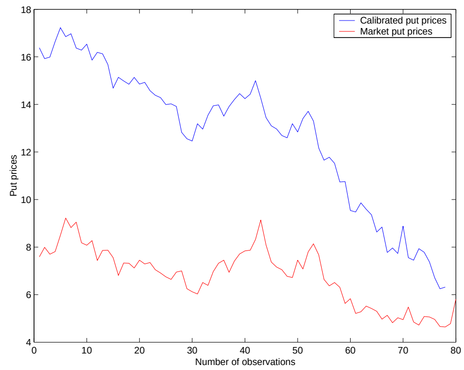


Figure 6.9: Comparison between the calibrated put prices and market put prices for contract M9 with K=85.

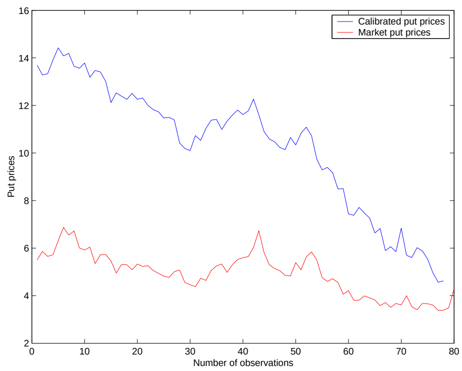


Figure 6.10: Comparison between the calibrated put prices and market put prices for contract M9 with K=80.

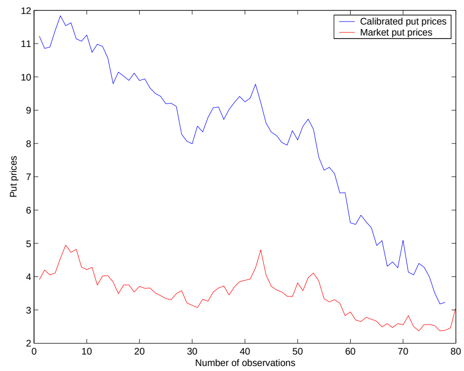


Figure 6.11: Comparison between the calibrated put prices and market put prices for contract M9 with K=75.

6.3 Results which incorporate explicit strike dependency

In this experiment we use the market data of expired put options on future M8 for four different strikes. See Table 6.3 for a quick overview of the details for contract M8. The procedure is as follows: For each

Contract	K	Date at t=0	Date at t=T	Observed trading days	Total trading days
M8	K= 85	20/11/2007	20/05/2008	125	125
M8	K=110	05/12/2007	20/05/2008	115	115
M8	K=115	10/03/2008	20/05/2008	51	51
M8	K=125	13/03/2007	20/05/2008	48	48

Table 6.3: Contract details.

strike we run the Kalman filter over the future prices. Parameters are given in Table 6.4 for four different strikes. The initial parameter set is given by the one in Table 6.2.

In Figure 6.12 we see that the calibrated future prices are fitted succesfully to the market future prices. This holds for all four different strike prices despite the very high values of some parameters. To calibrate the model to the market put prices as well, we keep the optimized parameter set constant for each

Parameters	K=85	K=110	K=115	K=125
k	0.2089	0.2144	0.0014	0.0630
μ	1.4354	1.3117	6.5556	3.9980
α	11.2299	8.0303	0.2753	109.2796
λ	3.1413	1.9467	7.5792	18.3986
σ_S	0.2613	0.2587	0.3013	0.1649
σ_δ	0.0303	0.0455	3.0048	1.6310
ρ	0.9999	0.9999	0.6635	-0.9999
$\lvert h_1 \rvert$	0.0341	0.0231	0.0013	0.0320

Table 6.4: Optimized parameter set for different strike prices for the contract M8.

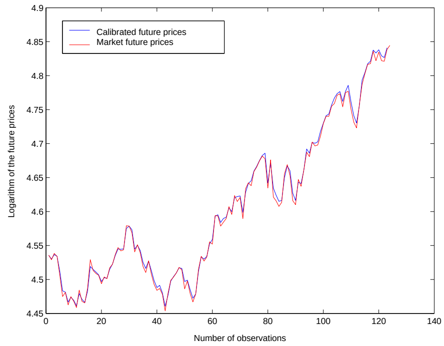


Figure 6.12: Comparison between the calibrated future prices and market future prices for contract M8 K=85.

strike price, and only σ_S and σ_δ are incorporated in the EKF. Results are given in Table 6.5 Figures 6.13-

Parameters	K=85	K=110	K=115	K=125
σ_S	0.3737	0.3756	0.3090	0.2415
σ_δ	0.6701	0.4923	3.8011	1.7195
$\lvert h_1 \rvert$	0.0145	0.0159	0.0056	0.0265

Table 6.5: Optimized parameter set for different strike prices for the contract M8.

6.16 depict calibrated put prices for four strikes, where the σ_S and σ_δ in Table 6.4 are replaced by the ones in Table 6.5. The σ_S and σ_δ in Table 6.5 are calibrated exclusively to the market put prices, but from Figures 6.13-6.16 we can conclude that for these parameters the put prices still fit the market data exceptionally well. Figures 6.17-6.20 represent the volatility $\tilde{\sigma}_F(t,T)$ compared to real dynamic volatility. In Figure 6.17 we see that the real dynamic volatility differs from $\tilde{\sigma}_F(t,T)$ in the end of the observation period. However, this is due to the fact that the put prices, with strike 85, will become worthless. Trying to get data-fitted volatility for that period is inadequate. Also the bumb around observation 100 can be confirmed in the Figure 6.13, where the market put prices differ from the calibrated ones.

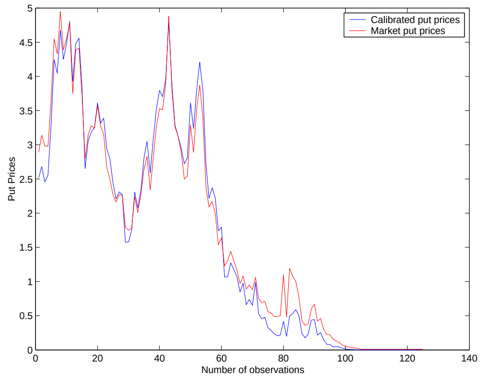


Figure 6.13: Comparison between the calibrated put prices and market put prices for contract M8 K=85.

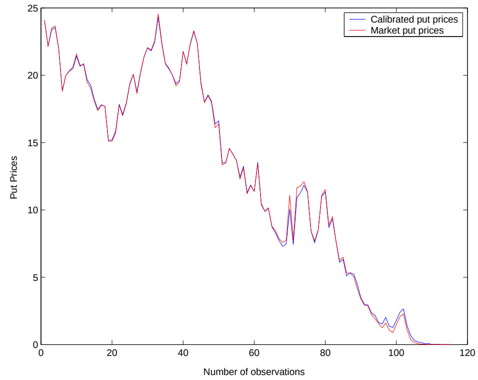


Figure 6.14: Comparison between the calibrated put prices and market put prices for contract M8 K=110.

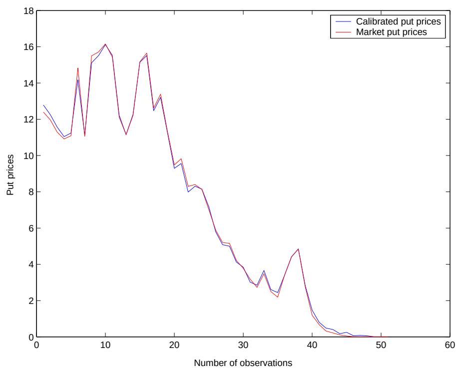


Figure 6.15: Comparison between the calibrated put prices and market future put for contract M8 K=115.

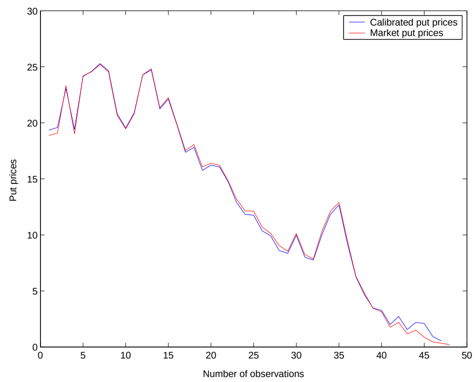


Figure 6.16: Comparison between the calibrated put prices and market future put for contract M8 K=125.

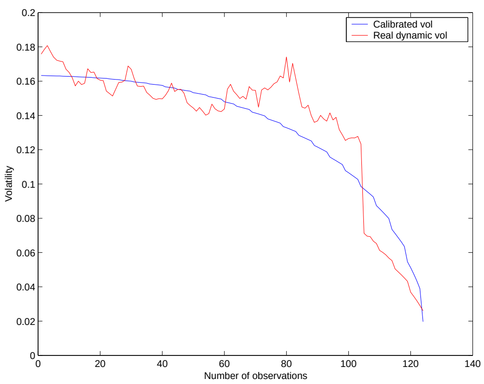


Figure 6.17: Comparison between $\tilde{\sigma}_F(t,T)$ and $\sigma_{\mathrm{RD}}(t,T)$ for contract M8 with K=85.

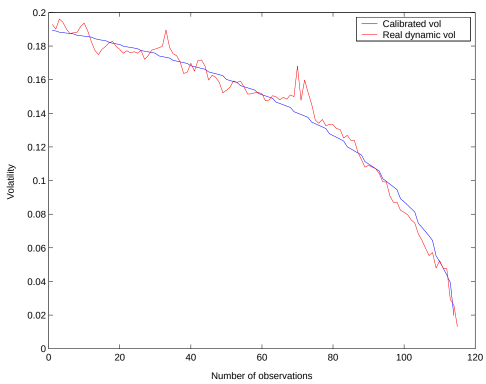


Figure 6.18: Comparison between $\tilde{\sigma}_F(t,T)$ and $\sigma_{\mathrm{RD}}(t,T)$ for contract M8 with K=110.

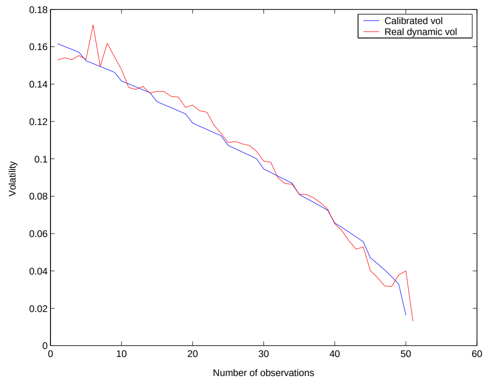


Figure 6.19: Comparison between $\tilde{\sigma}_F(t,T)$ and $\sigma_{\mathrm{RD}}(t,T)$ for contract M8 with K=115.

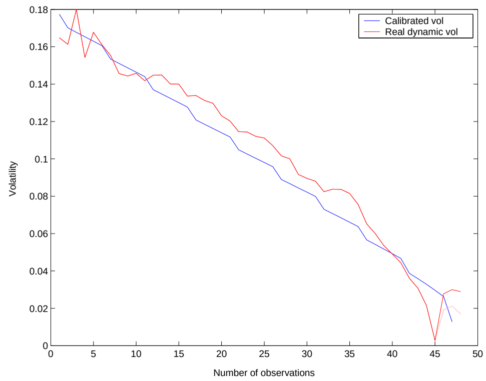


Figure 6.20: Comparison between $\tilde{\sigma}_F(t,T)$ and $\sigma_{\mathrm{RD}}(t,T)$ for contract M8 with K=125.

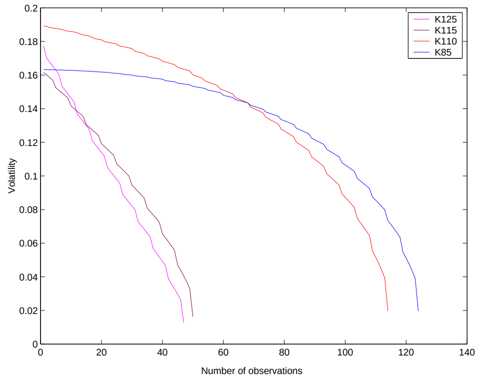


Figure 6.21: Calibrated volatility contract M8 with different strike prices

6.4 Results which incorporates both strike and maturity dependency

In Section 6.3 we again assume incomplete data sets for all three contracts, see Table 6.1. If we look at equation (6.5), the only parameter which is very sensitive is the $\hat{\sigma}_F^2$ given in (6.6). This volatility contains the components σ_S and σ_δ and a term which dictates the dependency on $\tau = t - T$. The idea is to parametrize $\sigma_S = \bar{\sigma}_S \tau^\beta$ and $\sigma_\delta = \bar{\sigma}_\delta \tau^\gamma$ to include more (time-dependent) structure into (6.6), keeping all the other parameters constant, i.e. we only calibrate $\bar{\sigma}_S$, β, $\bar{\sigma}_\delta$ and γ via the EKF.
For contracts with a maturity less than a year (i.e. $T < 1$) the $-2 \leq \beta \leq 0$ and for contracts with a maturity with one year or more, $-10 \leq \beta \leq 0$. This is done because otherwise σ_S will diverge in the calibration process.

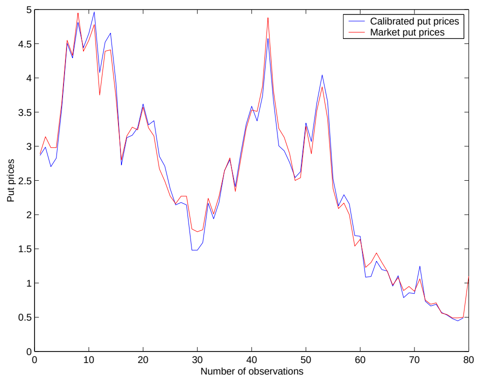


Figure 6.22: Comparison between the calibrated put prices and market put prices for contract M8 with K=85.

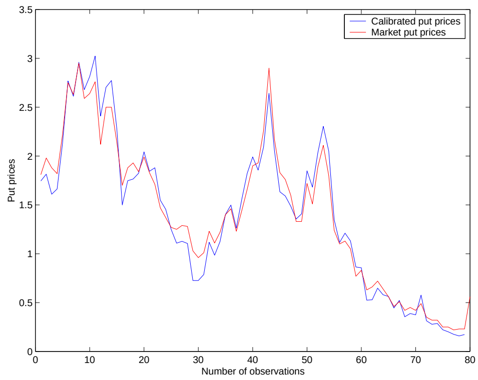


Figure 6.23: Comparison between the calibrated put prices and market put prices for contract M8 with K=80.

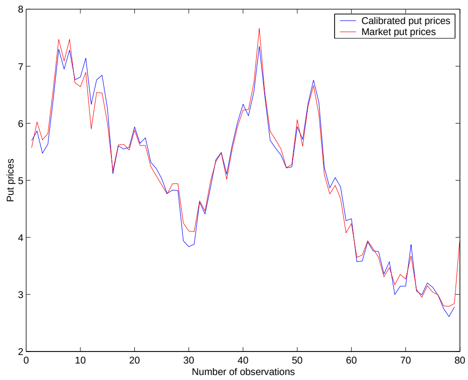


Figure 6.24: Comparison between the calibrated put prices and market put prices for contract Z8 with K=85.

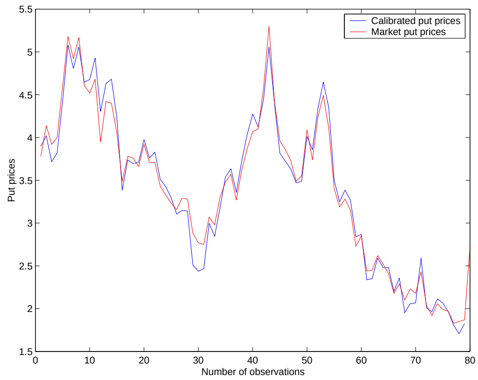


Figure 6.25: Comparison between the calibrated put prices and market put prices for contract Z8 with K=80.

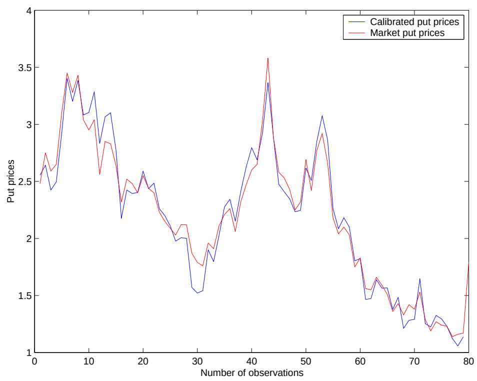


Figure 6.26: Comparison between the calibrated put prices and market put prices for contract Z8 with K=75.

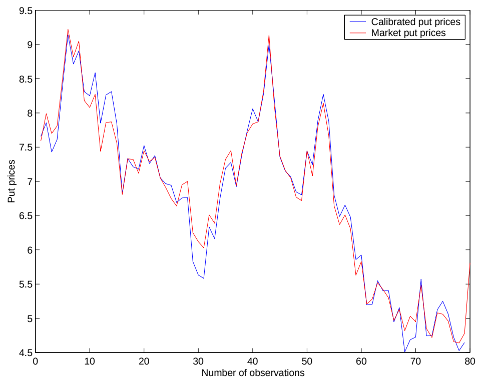


Figure 6.27: Comparison between the calibrated put prices and market put prices for contract M9 with K=85.

Note that $\bar{\sigma}_S$ and $\bar{\sigma_\delta}$ both are positive, guaranteeing positiveness for σ_S and σ_δ. However, this is accomplished by artificially setting zero-underbounds for $\bar{\sigma}_S$ and $\bar{\sigma_\delta}$. In Figures 6.30-6.37 $\tilde{\sigma}_F(t,T)$ and $\sigma_{\text{RD}}(t,T)$ are plotted for each contract and for each strike.

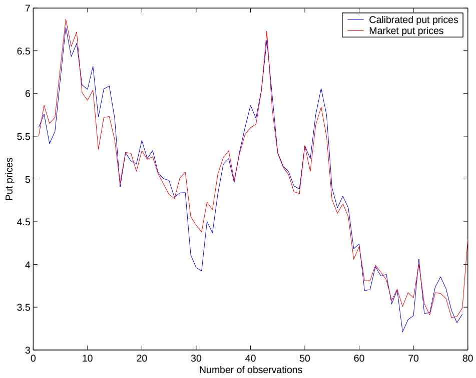


Figure 6.28: Comparison between the calibrated put prices and market put prices for contract M9 with K=80.

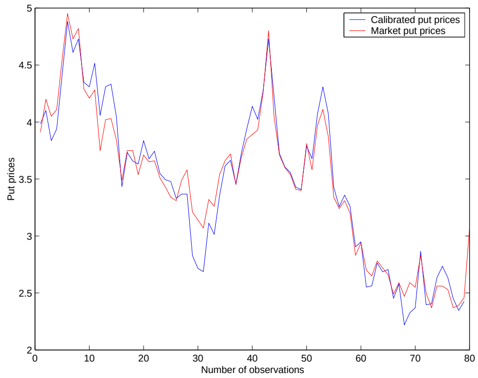


Figure 6.29: Comparison between the calibrated put prices and market put prices for contract M9 with K=75.

Contract names	β	$\bar{\sigma}_S$	$\bar{\sigma}_\delta$	γ	$\bar{v}$	$\lvert h_1 \rvert$
M8 K=85	-0.8563	0.1130	0.7012	-0.8653	0.003	0.003
M8 K=80	-0.3622	0.2333	1.1734	-0.0367	0.002	0.002
Z8 K=85	-1.4589	0.2290	0.6718	-0.2131	0.005	0.003
Z8 K=80	0.8784	0.2941	0.5632	-2.0777	0.008	0.006
Z8 K=75	0.7970	0.3013	0.5824	-2.0423	0.008	0.007
M9 K=85	-1.1096	0.4073	2.2486	-6.7991	0.011	0.008
M9 K=80	-0.9248	0.3944	2.1129	-6.2171	0.013	0.005
M9 K=75	-0.7054	0.3795	1.9293	-5.5284	0.094	0.003

Table 6.6: Estimated parameters.

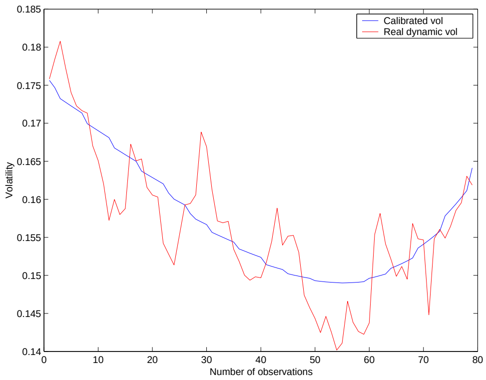


Figure 6.30: Comparison between $\tilde{\sigma}_F(t,T)$ and $\sigma_{\mathrm{RD}}(t,T)$ for contract M8 with K=85.

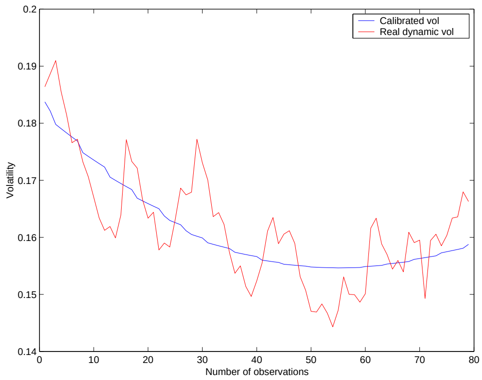


Figure 6.31: Comparison between $\tilde{\sigma}_F(t,T)$ and $\sigma_{\mathrm{RD}}(t,T)$ for contract M8 with K=80.

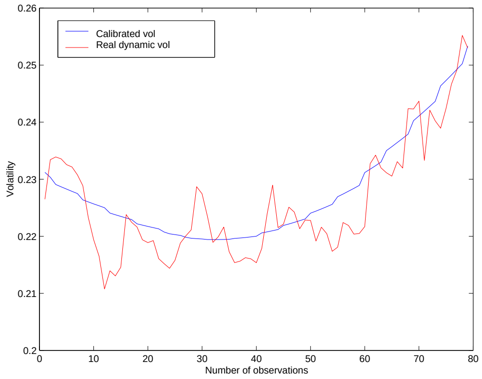


Figure 6.32: Comparison between $\tilde{\sigma}_F(t,T)$ and $\sigma_{\text{RD}}(t,T)$ contract Z8 with K=85.

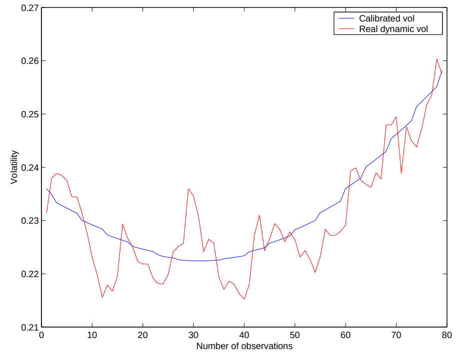


Figure 6.33: Comparison between $\tilde{\sigma}_F(t,T)$ and $\sigma_{\text{RD}}(t,T)$ contract Z8 with K=80.

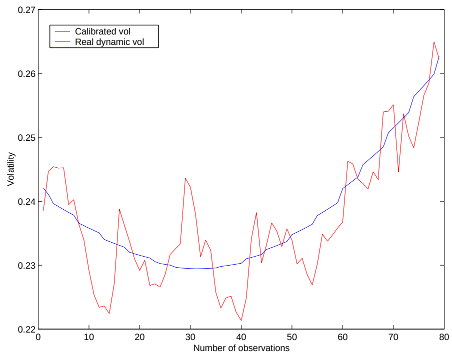


Figure 6.34: Comparison between $\tilde{\sigma}_F(t,T)$ and $\sigma_{\text{RD}}(t,T)$ for contract Z8 with K=75.

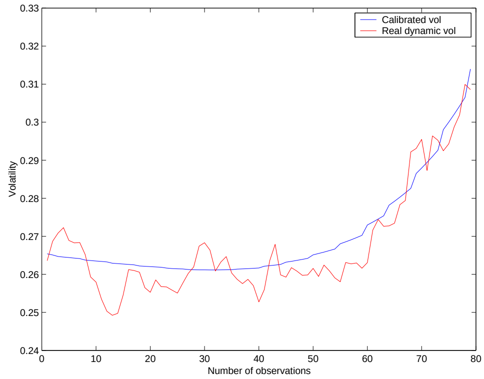


Figure 6.35: Comparison between $\tilde{\sigma}_F(t,T)$ and $\sigma_{\text{RD}}(t,T)$ for contract M9 with K=85.

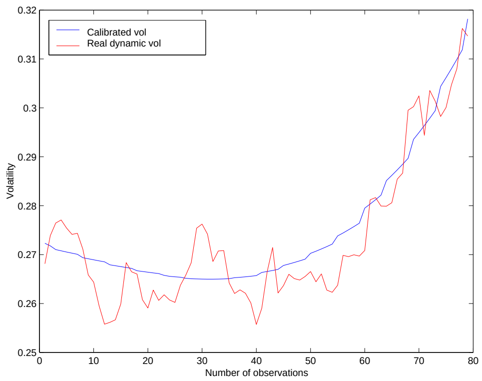


Figure 6.36: Comparison between $\tilde{\sigma}_F(t,T)$ and $\sigma_{\text{RD}}(t,T)$ for contract M9 with K=80.

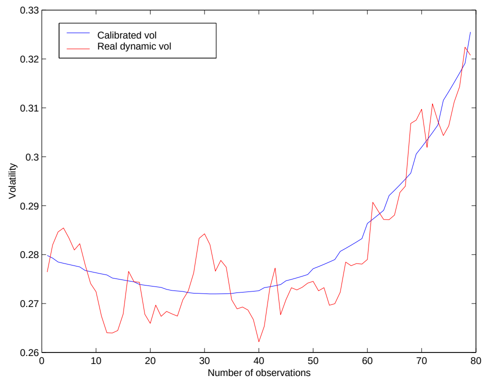


Figure 6.37: Comparison between $\tilde{\sigma}_F(t,T)$ and $\sigma_{\text{RD}}(t,T)$ for contract M8 with K=75.

6.4.1 Time-local volatility parametrization for the spot process dynamics

In the previous section we introduced two parametrisations $\sigma_S = \bar{\sigma}_S \tau^{\beta}$ and $\sigma_\delta = \bar{\sigma}_\delta \tau^{\gamma}$. These parametrisations bring more structure in $\hat{\sigma}_F$, resulting in better option estimations, but it does not bring in additional dynamics in the spot processes for S_t and δ_t since we assumed σ_S and σ_δ to be constant. Nevertheless, if we can find a parametrization of the spot volatilities and inserting these into

$$\sigma_F^2 = \sigma_S^2 + \sigma_\delta^2 B^2(\tau) + 2\rho\sigma_S\sigma_\delta B(\tau), \tag{6.15}$$

such that one can still explicitly solve the integral given in the relation (6.6), then the problem is solved. In general, for arbitrary parametrizations $\sigma_S(t)$ and $\sigma_\delta(t)$, the explicit computation of integral (6.6) is basically impossible.
Assume the following form for $\hat{\sigma}_F^2$, (see e.g. (6.6)):

$$\hat{\sigma}_F^2 = \frac{1}{\tau}\int_0^\tau D(\tau)^2 - C(\tau)d\tau, \quad C(\tau) \le 0, \quad D(\tau) \ge 0, \tag{6.16}$$

or, equivalently,

$$\tilde{\sigma}_F^2 = \int_0^\tau D(\tau)^2 - C(\tau)d\tau, \tag{6.17}$$

and set

$$\sigma_S = D(\tau). \tag{6.18}$$

Substitution of (6.18) into (6.15) yields

$$\sigma_\delta^2 B(\tau)^2 + 2\rho D(\tau)\sigma_\delta B(\tau) = -C(\tau), \tag{6.19}$$

or equivalently

$$\sigma_\delta = \frac{-2\rho D(\tau)B(\tau) \pm \sqrt{4\rho^2 D(\tau)^2 B(\tau)^2 - 4B(\tau)^2 C(\tau)}}{2B(\tau)^2}. \tag{6.20}$$

Since $\rho = 0.7171$ we see that $-2\rho D(\tau)B(\tau) \ge 0$ because $B(\tau) \le 0$. For the square root to be positive, it is required to have

$$C(\tau) \le \rho^2 D(\tau)^2. \tag{6.21}$$

Looking at Figures 6.20-6.27 we assume

$$\tilde{\sigma}_F = a^2\tau^b + c^2\tau^\gamma + d^2, \quad \gamma \le 0, \tag{6.22}$$

or, equivalently,

$$\tilde{\sigma}_F^2 = a^4\tau^{2b} + c^4\tau^{2\gamma} + d^4 + 2a^2c^2\tau^{b+\gamma} + 2a^2d^2\tau^b + 2c^2d^2\tau^\gamma. \tag{6.23}$$

Now

$$\frac{\partial\tilde{\sigma}_F^2}{\partial\tau} = 2ba^4\tau^{2b-1} + 2\gamma c^4\tau^{2\gamma-1} + (b+\gamma)2a^2c^2\tau^{b+\gamma-1} + 2ba^2d^2\tau^{b-1} + \gamma 2c^2d^2\tau^{\gamma-1}. \tag{6.24}$$

Take

$$\begin{aligned} D^2(\tau) &= 2ba^4\tau^{2b-1} + 2ba^2d^2\tau^{b-1} + \oplus(b+\gamma)2a^2c^2\tau^{b+\gamma-1} \\ C(\tau) &= 2\gamma c^4\tau^{2\gamma-1} + (1-\oplus)(b+\gamma)2a^2c^2\tau^{b+\gamma-1} + \gamma 2c^2d^2\tau^{\gamma-1}, \end{aligned} \tag{6.25}$$

where

$$\oplus = \begin{cases} 1, & \text{if} \quad (b+\gamma)2a^2c^2\tau^{b+\gamma-1} \ge 0 \\ 0, & \text{if} \quad (b+\gamma)2a^2c^2\tau^{b+\gamma-1} \le 0, \end{cases} \tag{6.26}$$

In doing so, we force $D(\tau)$ to be positive and $C(\tau)$ to be negative and thus satisfying the contraint given in (6.21). Substituting (6.25) into (6.16) we get

$$\begin{aligned} \tau\hat{\sigma}_F^2 &= \int_0^\tau D(\tau)^2 - C(\tau)d\tau \\ &= a^4\tau^{2b} + c^4\tau^{2\gamma} + d^4 + a^2c^2\tau^{b+\gamma} + a^2d^2\tau^b + 2c^2d^2\tau^\gamma. \end{aligned} \tag{6.27}$$

returning to the original form (6.22). So the parametrisation of σ_S and σ_δ is given by

$$\begin{aligned} \sigma_S &= \sqrt{2ba^4\tau^{2b-1} + 2ba^2d^2\tau^{b-1} + \oplus(b+\gamma)2a^2c^2\tau^{b+\gamma-1}} \\ \sigma_\delta &= \frac{-2\rho\sigma_S}{B(\tau)} + \frac{\sqrt{4\rho^2\sigma_S^2B(\tau)^2 - 4B(\tau)^2(2\gamma c^4\tau^{2\gamma-1} + (1-\oplus)(b+\gamma)2a^2c^2\tau^{b+\gamma-1} + \gamma 2c^2d^2\tau^{\gamma-1})}}{2B(\tau)^2}, \end{aligned} \tag{6.28}$$

where we take the plus sign which ensures positiveness of σ_δ. This parametrisation gives the same results (i.e. figures) as in Sections 6.2 and 6.3. The parameters are given in Table 6.7.

Contact names	a	b	c	d	γ	h
M8 K=85	-0.7083	4.0083	-0.0002	0.3816	-7.6388	0.009
M8 K=80	-1.5430	6.2778	0.0000	0.3924	-12.1297	0.022
Z8 K=85	0.1053	31.6287	-0.0202	0.4672	-12.3967	0.012
Z8 K=80	0.1136	29.0214	-0.0279	0.4693	-10.8131	0.009
Z8 K=75	0.1076	31.2684	-0.0239	0.4775	-11.4608	0.007
M9 K=85	-0.4076	0.4654	-1.1737	0.2502	-17.1171	0.012
M9 K=80	-0.3081	1.0447	-0.9980	0.3500	-14.2833	0.002
M9 K=75	-0.2808	1.4806	-0.8593	0.3587	-11.8414	0.003

Table 6.7: Estimated parameters.

If we look at equation (6.28), σ_δ is strictly positive $\forall -1 \le \rho \le 1$. In fact, we must have, [cf. (6.20)],

$$\begin{aligned} -2\rho D(\tau)B(\tau) + \sqrt{4\rho^2D(\tau)^2B(\tau)^2 - 4B(\tau)^2C(\tau)} &\geq 0 \\ -2\rho D(\tau)B(\tau) &\geq -\sqrt{4\rho^2D(\tau)^2B(\tau)^2 - 4B(\tau)^2C(\tau)} \\ 2\rho D(\tau)B(\tau) &\leq \sqrt{4\rho^2D(\tau)^2B(\tau)^2 - 4B(\tau)^2C(\tau)} \\ 4\rho^2D(\tau)^2B(\tau)^2 &\leq 4\rho^2D(\tau)^2B(\tau)^2 - 4B(\tau)^2C(\tau) \\ 0 &\leq -4B(\tau)^2C(\tau), \end{aligned} \tag{6.29}$$

which is always true thanks to the initial assumption $C(\tau) \le 0$.

Since we set $\gamma \le 0$ as a constraint during the calibration, $\tilde{\sigma}_F$ blows up for $\tau \leftarrow 0$. Hence, if $\tau \leftarrow 0$ (and this will be the case if one tries to calibrate a matured option), the KF will try to neutralize the $c^2\tau^\gamma$ term. See also Table 6.7 for contract M8, which is close to maturity. The form of $\hat{\sigma}_F$ is only known for the first 79 observation days. Since we are considering put options on these futures, and since the futures on oil are rising, one can imagine that the put prices are deeply out of the money and are worth nothing (looking at the strikes of these options). In fact, $\tilde{\sigma}_F$ will be very small in that case (when close to maturity). Finally, assuming a form for the volatility curve, as we did in Section 6.4.1 is not reliable due to the fact that one does not know how the curve looks in the future. Compare e.g. Figure 6.17 and 6.30. We can conclude that trying to fit the curve of $\tilde{\sigma}_F$, we first need all the historical data.

6.5 From future price process to the spot processes

In the previous chapters and sections we initially defined S_t and δ_t by some kind of stochastic dynamics resulting in

$$\bar{V}_t + \frac{1}{2}\sigma_F^2(t,T)F^2\bar{V}_{FF} - r\bar{V} = 0, \tag{6.30}$$

where $\sigma_F^2(t,T)$ in given in (6.4). However, in the commodity market, and in particular the commodity light crude oil, the spot price and convenience yield process are not observable. In practice, it is therefore preferable to work with the price processes of futures. Reisman uses "as the inputs the parameters of the process specifying the local second moments of future prices of all maturities and their current prices" and showed "that any arbitrary specification of these inputs implies an arbitrage free furtures markets, derives the spot price and the convenience yield processes implied by such specification", ([33]: p.2). An particular example, which we discuss, is the spot price process and convenience yield process given in Chapter 4. Looking at the driftless term structure of the future price, we have

$$\begin{aligned}
\frac{dF(t,T)}{F(t,T)} &= (\sigma_S^2 + \sigma_\delta^2 B^2(t,T) + \rho\sigma_S\sigma_\delta B(t,T))dW_t^F \\
&= \sqrt{1-\rho^2}\sigma_S dW_t^1 + (\sigma_\delta B(t,T) + \rho\sigma_\delta)dW_t^2 \\
&= \sigma_1(t,T)dW_t^1 + \sigma_2(t,T)dW_t^2,
\end{aligned} \tag{6.31}$$

where W_t^1 and W_t^2 are assumed to be two independent Brownian motions. We define

$$\begin{cases} W_t^1 = \dfrac{W_t^S}{\sqrt{1-\rho^2}} - \dfrac{\rho W_t^\delta}{\sqrt{1-\rho^2}}, \\ W_t^2 = W_t^\delta, \end{cases} \tag{6.32}$$

where W_t^1 is orthogonal to W_t^2. Since in the limit of $T \to t$ we have

$$\sigma_1(t) = \sqrt{1-\rho^2}\sigma_S \quad \text{and} \quad \sigma_2(t) = \rho\sigma_S. \tag{6.33}$$

and

$$\begin{aligned}
\frac{dS_t}{S_t} &= (r-\delta_t)dt + \sqrt{1-\rho^2}\sigma_S dW_t^1 + \rho\sigma_S dW_t^2 \\
&= (r-\delta_t)dt + \sqrt{1-\rho^2}\sigma_S \left\{ \frac{dW_t^S - \rho dW_t^\delta}{\sqrt{1-\rho^2}} \right\} + \sigma_S\rho dW_t^S \\
&= (r-\delta_t)dt + \sigma_S dW_t^S.
\end{aligned} \tag{6.34}$$

Note that, together with

$$\delta_t = r - \frac{\partial}{\partial t}\ln F(0,t) + k\int_0^t e^{-k(t-u)}\left\{\frac{-\sigma_\delta}{k^2}\left[e^{-2kt}-1\right] - \frac{\rho\sigma_S\sigma_\delta}{k}\right\}du - \sigma_\delta\int_0^t e^{-k(t-u)}dW_u^2, \tag{6.35}$$

(6.34) is our initial process for the spot price. (6.35) is found by applying Itô's lemma to the logarithm of the future price $\ln F(t,T)$, substituting $S(t) = F(t,t)$ and finally take an Itô differential of that equation.

Applying Itô's lemma on (6.35) yields

$$\begin{aligned} d(e^{kt}\delta_t) &= \left[e^{kt}\frac{\partial}{\partial t}\ln F(0,t)\right]' dt - ke^{kt}\left\{\frac{-\sigma_\delta}{k^2}\left[e^{-2kt}-1\right] - \frac{\rho\sigma_S\sigma_\delta}{k}\right\}dt + \sigma_\delta e^{kt}dW_t^2 \\ &= \left[ke^{kt}\frac{\partial}{\partial t}\ln F(0,t) + e^{kt}\frac{\partial^2 \ln F(0,t)}{\partial t^2}\right]dt - ke^{kt}\left\{\frac{-\sigma_\delta}{k^2}\left[e^{-2kt}-1\right] - \frac{\rho\sigma_S\sigma_\delta}{k}\right\}dt + \sigma_\delta e^{kt}dW_t^2 \\ &= -ke^{kt}\left\{-\frac{\partial}{\partial t}\ln F(0,t) - \frac{1}{k}\frac{\partial^2 \ln F(0,t)}{\partial t^2} - \frac{\rho\sigma_S\sigma_\delta}{k} - \frac{\sigma_\delta^2}{2k^2}(e^{-2kt}-1)\right\}dt + \sigma_\delta e^{kt}dW_t^2. \quad (6.36) \end{aligned}$$

Define $\alpha(t)$, the mean-reversion term of $d\delta_t$ which is now assumed to be time-dependent, by

$$\alpha(t) = \frac{1}{k}\frac{\partial^2}{\partial t^2}\ln F(0,t) + \frac{\partial}{\partial t}\ln F(0,t) + \frac{\sigma_\delta^2}{2k^2}(e^{-2kt}-1) + \frac{\rho\sigma_S\sigma_\delta}{k}. \quad (6.37)$$

(6.37) follows from rewriting our initial spot convenience yield process, i.e. consider again

$$d\delta_t = k(\alpha(t) - \delta_t)dt + \sigma_\delta dW_t^\delta. \quad (6.38)$$

We have

$$\begin{aligned} d(e^{kt}\delta_t) &= ke^{kt}\delta_t dt + e^{kt}d\delta_t \\ &= ke^{kt}\delta_t dt + e^{kt}(k(\alpha(t)-\delta_t)dt + \sigma_\delta dW_t^\delta) \\ &= k\alpha(t)e^{kt}dt + \sigma_\delta e^{kt}dW_t^\delta. \quad (6.39) \end{aligned}$$

By comparing (6.36) by (6.39) the result in (6.37) follows immediately.

6.6 Conclusion

Pricing put options on futures contracts seems to work quite well. Matching the model's implied volatility with the real dynamic volatility seems rather difficult if one does not have a matured contract. However, as we can conclude from Section 6.3, pricing of options on a matured contract using the optimized parameter set φ and only differ ρ, σ_S and σ_δ, is succesful. In Table 6.4 we can see high values of α. An explanation for this can be that our initially assumed stochastic processes for the state variables are not sufficiently structured. Extending this stochastic system with e.g. stochastic convenience yield, stochastic interest rates, seasonality factors, jump processes etc, can be an improvement. Furthermore, starting from the future price process given in (6.31), we see that α_t is time dependent and depends on $F(0,t)\ \forall\ t$. Since the future prices are not observable for an arbitrary point in time, one should find a way to estimate them. This can be done by extrapolation, but, since we must calculate first and second order derivatives of $\ln F(0,t)$, this can cause serious problems. Therefore, we took α constant.

Chapter 7

Discussion and further research

7.1 Discussion

In this Master's Thesis we calibrate future prices and we calibrate put options on these future prices as well. By means of the Kalman filter we are able to fit the future prices very succesfully to the market. The choice of the KF is made because of the fact that the spot processes of the future prices are unobservable. For the spot price we assumed throughout this thesis a GBM and for the convenience yield either an OU process or a CIR process. We find inadequacy of the CIR process due to the nonnegativity constraint for the convenience yield. Hence we continued with the OU process to calibrate the put options. Again, using the extended Kalman filter we can fit the put option prices succesfully to the market. However, due to the unrealistic values of some parameters in Table 6.4, the stochastic process assumed initially for the spot price and the convenience are not satisfactory. In fact, adding extensions such as stochastic volatility and interest rates appear to be necessary. Further research and remark points are given below

- Although the EKF is not optimal for the Ornstein-Uhlenbeck process, it is still acceptable and therefore we use this technique to price put options on futures contracts.

- Another important issue to be considered is the strong evidence against the adequacy of the CIR model for the future prices. This has several reasons. The first one is that the CIR process is not Gaussian which is essential to obtain an optimal result from the KF. Secondly, the nonnegativity constraint gives highly inaccurate results for the path of the convenience yield. And last, if the calibration starts with data such that the convenience yield is negative, the iterative procedure is useless, causing very large standard errors and not realistic parameters. However, if we shift the data (see case two of Section 5.2) we find that the robustness of the method is satisfactory. Changing the initial parameter set will give the same optimized parameter set and the same value of the log-likelihood function. Chapter 5 was inspired by [4]. They considered data of light crude oil ranging for a period from 17^{th} of March 1999 to the 24^{th} of December 2003. In that period they find approximately 20 observations for which the convenience yield is negative. If a replacement of the negative convenience yield by zero is done for these few observations, it will not destroy the total path. In our period however, as we can see in the Figure 5.2, the convenience yield is negative for almost half of our observations. Replacement by zero will therefore cause great disturbances in the path.

- The initial values of the parameters and the state variables are not known. The theory, however, states that, when considering the stochastic system

$$\begin{cases} \mathbf{Y}_t = \mathbf{d} + \mathbf{Z}[x_t, \delta_t]' + \boldsymbol{\epsilon}, \\ \boldsymbol{\alpha}_t = \mathbf{c} + \mathbf{Q}\boldsymbol{\alpha}_{t-1} + \mathbf{R}\boldsymbol{\xi}, \end{cases} \tag{7.1}$$

the state variable $\boldsymbol{\alpha}_t = [x_t \quad \delta_t]'$ in the KF algorithm should start with

$$\boldsymbol{\alpha}_0 = \mathbb{E}[\mathbf{c} + \mathbf{Q}\boldsymbol{\alpha}_{t-1} + \mathbf{R}\boldsymbol{\xi}] = (\mathbf{I} - \mathbf{Q})^{-1}\mathbf{c}, \tag{7.2}$$

$$\tag{7.3}$$

that is, the conditional expectation of the starting variables. The matrix $\mathbf{H}$ consists of the the variance of the state variables, so

$$\mathbb{V}\mathrm{ar}[\boldsymbol{\alpha}_0] = \mathbf{Q}^2\mathbb{V}\mathrm{ar}[\boldsymbol{\alpha}_0] + \sigma^2_{\boldsymbol{\xi}}, \quad \text{or equivalently,} \quad \mathbb{V}\mathrm{ar}[\boldsymbol{\alpha}_0] = \frac{\sigma^2_{\boldsymbol{\xi}}}{\mathbf{I} - \mathbf{Q}^2}. \tag{7.4}$$

Since the KF is a stable algorithm, the effect of initial conditions will vanish almost immediately in our test situation. We tested this for the OU process presented in Chapter 4 by setting $\alpha_0 = [10, \quad -10]$, a totally inadequate initial value for both the log of the spot price and the convenience yield. The log-likelihood remained the same, as well as the parameters given in Table 4.1. The only difference is that the value of the log-likelihood function jumped in the beginning. This can be explained by this incorrect initial value.

- The market price of risk is sometimes negative and sometimes positive. The precise explanation of this behaviour is difficult and is still an important subject of study. The Black-Scholes equation is an example of modelling the perfect world in which hedging a portfolio can be done continuously and perfectly. So the portfolio is riskless and that is why the market price of risk does not appear in the model. In our situation, considering non-observable quantities (spot price and convenience yield) it is not possible to perfectly hedge the risk-free portfolio. So one trades assets that depend on these quantities (in our case the futures contracts). It is therefore impossible to build a completely risk-free portfolio and therefore an investor will ask a premium for investing. In fact, the market price of risk is the difference between the price a trader would buy the asset for and the theoretical price.

- When pricing a put option on a futures contract, there are some important details to be considered. First of all, when deriving the closed form solution of a future put option, one should be careful whether or not introducing stochastic interest rates. In Appendix B we show the type of problems that may arise when doing this. Appendix B is not meant to be a fully detailed description of stochastic theory but it shows some important problems. Although we assumed deterministic (constant) interest rates we are able to price the future put option succesfully. Our model implied volatility follows exactly the pattern of the real dynamic volatility from the Black-Scholes framework. In fact, it seems as if the former is a perfect fit for the latter. The errors in our calibrated put prices can be explained by the use of the EKF. When linearizing the system matrices, we introduce an approximation in the estimation. Since the closed form solution of a future put option is not linear, the EKF still seems to be the best iterative procedure. However, both the resulting parametrizations given in Section 6.4 are not acceptable for actual trading.

- In Table 6.4 we see unrealistic high values of some parameters, in particular the mean-reverting term α. This points out that the stochastic processes of the spot price and the convenience yield are not sufficient to price the futures contracts. In fact, it is possible to price the futures and put options on these futures very accurately, but the parameters are chosen such that they do not obey the initial assumptions of the stochastic processes. Model sensitivity analyses based on these parameters may then be misleading. It clearly proves that the model should be improved.

7.2 Further research

After reading the points of discussion one may wonder why we initially modeled unobserved variables stochastic processes. According to Section 6.4 we can find the stochastic processes back by assuming some kind of term structure (parametrization) of the volatility of the future prices. The fact that this volatility is, in this case, known, makes it some kind of reverse engeneering. However, starting from the future process, one does not know how to choose this volatility. In fact, the future price process may be given by

$$\frac{dF(t,T)}{F(t,T)} = \sum_{i=1}^{n} \sigma_i dW_t^i,$$

with i-independent Brownian motions. Another reason for initially assuming stochastic processes for the state variables is because of the Greeks. In the PDE (2.31) traders would actually need the first and second order derivatives with respect to S and δ to hedge their exposure caused by joining in on the future market.

Although we can calibrate both future prices and put prices very accurately, the parameters presented in Table 6.4 are unrealistic. To improve this, the following research directions could be recommended. First of all we may just accept these parameters as they are parameters of unobserved processes. Secondly, we can introduce stochastic volatility to create a three factor joint stochastic process. The problem with including stochastic volatility is that the affine form solution of the closed form solution of the futures is no longer applicable. Another way is including a stochastic interest rate, which has been explained in Appendix B. A closed form solution for the K-strike European put option can still be derived but since our results are succesful with constant interest rates of 4 percent. Including stochastic interest rates would not make a significant difference. A further research topic could be including time and state dependent volatility in the initial processes for S_t and δ_t, i.e.

$$\begin{cases} dS_t = \mu S_t dt + \sigma_S(t, S_t) S_t dW, \\ d\delta_t = k(\alpha - \delta_t)dt + \sigma_\delta(t, \delta_t) dZ_t. \end{cases} \tag{7.5}$$

But instead of making the volatilities state-dependent, one could also insert jumps into the processes. The simplest mean-reversion jump-diffusion model [31] for spot prices is,

$$dS_t = \alpha(\mu - \psi K_m - \ln S_t)S_t dt + \sigma_S S_t dW + K S_t dq, \tag{7.6}$$

where K is the jump with log-normal distribution, i.e. $\ln(1 + K) \sim \mathcal{N}\left(\ln(1 + K_m) - \gamma^2/2, \gamma^2\right)$ with K_m mean jump size and γ the jump-volatility. ψ is the average number of jumps per year and dq is a Poisson process.

To improve modelling of the spot price model of future prices on commodities like light crude oil even further, one can insert a seasonality factor, time-varying mean-reverting, etc. The problem with forecasting or calibrating oil is that oil is available globally. There are numerous factors that can affect the price of oil. Seasonality, for example, can influence the demand for oil. During cold periods, more oil is used which can make the price of oil rise and the other way around when it is warm. Oil prices also depend on the economic environment. When the economy grows, oil demand in our daily lives will increase, causing the prices of oil to rise. On the other hand, the probability of war in the top oil producing countries is a serious factor. In short, the analysis of future oil price is complicated and one can make as many adjustments to the model as one wishes.

It is a real challenge to end up with a realistic model for oil prices which, at the same time, can be dealt with in a fast and efficient way. The models presented in this work are a first step in that direction.

Appendix A

Derivations of the Kalman filter equations

A.1 Kalman Algorithm

In general the Kalman filter tries to estimate a state $x \in \Re^n$ where x_t is given by the stochastic differential equation

$$x_t = A_t x_{t-1} + w_{t-1}, \tag{A.1}$$

with a measurement $z \in \Re^n$ given by

$$z_t = H_t x_t + v_t, \tag{A.2}$$

where

- A_t is a $n \times n$ matrix,

- H_t is a $m \times n$ matrix,

- w_t and v_t are the process and measurement noise (resp.) with mean zero and covariance matrices Q and R (resp.).

The error term $\tilde{x}_{t-1} = \hat{x}_{t-1} - x_{t-1}$ is assumed to have mean zero and covariance P_{t-1}. In other words, we assume the estimator $\hat{x}_{t-1}$ to be unbiased. The goal of the KF is to find an unbiased minimum variance estimator of the state at time t, of the form

$$\hat{x}_t = K'_t \hat{x}_{t-1} + K_t z_t, \tag{A.3}$$

where K_t is called the *Kalman Gain*. Note that this estimator is linear in x_t and in z_t. Now substract x_t from both the left and right side to get

$$\hat{x}_t - x_t = K'_t \hat{x}_{t-1} + K_t z_t - x_t. \tag{A.4}$$

Substituting (A.2) we get

$$\hat{x}_t - x_t = K'_t \hat{x}_t + K_t(H_t x_t + v_t) - x_t - K'_t x_{t-1} + K'_t x_{t-1}, \tag{A.5}$$

where we substract and add the term $K_t' x_{t-1}$. Substituting (A.1) and rearranging terms finally results in

$$\hat{x}_t - x_t = K_t'[\hat{x}_{t-1} - x_{t-1}] + [K_t H_t A_t - A_t + K_t']x_{t-1} + (K_t H_t - I)w_t + K_t v_{t-1}. \tag{A.6}$$

Since we assumed $\hat{x}_t$ to be unbiased, i.e.

$$\begin{aligned}\mathbb{E}[\hat{x}_t - x_t] &= \mathbb{E}[K_t'[\hat{x}_{t-1} - x_{t-1}] + [K_t H_t A_t - A_t + K_t']x_{t-1} + (K_t H_t - I)w_t + K_t v_{t-1}] \\ &= 0 + [K_t H_t A_t - A_t + K_t']\mathbb{E}[x_{t-1}] + 0 + 0 = 0,\end{aligned} \tag{A.7}$$

where $[K_t H_t A_t - A_t + K_t']$ is just a number and so it can be taken out of the expectation. This implies that $[K_t H_t A_t - A_t + K_t'] = 0$ or equivalently

$$K_t' = (I - K_t H_t)A_t. \tag{A.8}$$

Substituting this result into (A.3) gives

$$\begin{aligned}\hat{x}_t &= (I - K_t H_t)A_t \hat{x}_{t-1} + K_t z_t \\ &= A_t \hat{x}_{t-1} + K_t(z_t - H_t A_t \hat{x}_{t-1}).\end{aligned} \tag{A.9}$$

Now we want to find K_t such that the covariance of the estimation error is minimized. The covariance matrix of the error, P_t is (for an unbiased error) a diagonal matrix with the variances on the diagonal, i.e. we want to find K_t such that $trace(P_t)$ is minimized. We have

$$P_t = \mathbb{E}[(\hat{x}_t - x_t)(\hat{x}_t - x_t)'] = \mathbb{E}[\tilde{x}_t \tilde{x}_t']. \tag{A.10}$$

The extrapolated estimate is defined as

$$\hat{x}_t^- = A_t \hat{x}_{t-1}, \tag{A.11}$$

and the covariance of this error is noted as P_t^-. The extrapolated estimate error is then given by

$$\tilde{x}_t^- = \hat{x}_t^- - x_t = A_t[\hat{x}_{t-1} - x_{t-1}] - w_t. \tag{A.12}$$

We have

$$\begin{aligned}P_t^- &= \mathbb{E}[\tilde{x}_t(\tilde{x}_t^-)'] \\ &= \mathbb{E}[(\hat{x}_t - x_t)(\hat{x}_t^- - x_t)'] \\ &= \mathbb{E}[(\hat{x}_t - (A_t x_{t-1} + w_t))(A_t(\hat{x}_{t-1} - x_{t-1}) - w_{t-1})'] \\ &= \mathbb{E}[(\hat{x}_t - A_t x_{t-1} - w_t)((\hat{x}_{t-1} - x_{t-1})' A_t' - w_{t-1}')] \\ &= \mathbb{E}[(A_t(\hat{x}_{t-1} - x_{t-1}) - w_{t-1} + K_t(z_t - H_t A_t \hat{x}_t))((\hat{x}_t - x_t)' A_t' - w_{t-1}')] \\ &= A_t \mathbb{E}[(\hat{x}_{t-1} - x_{t-1})(\hat{x}_{t-1} - x_{t-1})']A_t' - 2A_t \mathbb{E}[(\hat{x}_{t-1} - x_{t-1})w_t'] \\ &+ \mathbb{E}[w_{t-1} w_{t-1}'] + \mathbb{E}[K_t(z_t - H_t A_t \hat{x}_{t-1})(\hat{x}_{t-1} - x_{t-1})'] - \mathbb{E}[K_t(z_t - H_t A_t \hat{x}_{t-1})w_{t-1}'] \\ &= A_t P_t A_t' + Q_t.\end{aligned} \tag{A.13}$$

Note that $\mathbb{E}[w_t] = \mathbb{E}[(\hat{x}_t - x_{t-1})] = 0$. In the same we way we can derive P_t, noting that from (A.9) and (A.11) it follows that

$$\hat{x}_t = \hat{x}_t^- - K_t(z_t - H_t \hat{x}_t^-), \tag{A.14}$$

holds. Using this result we can calculate the error term $\tilde{x}_t$,

$$\begin{aligned}\tilde{x}_t &= \hat{x}_t - x_t \\ &= [I - K_t H_t]\hat{x}_t^- + K_t z_t - x_t \\ &= [I - K_t H_t]\tilde{x}_t^- - K_t v_t, \end{aligned} \tag{A.15}$$

where we substituted (A.2) and (A.12) and rearranged terms. And so

$$\begin{aligned} P_t &= \mathbb{E}[\tilde{x}_t \tilde{x}_t'] \\ &= \mathbb{E}[(\hat{x}_t - x_t)(\hat{x}_t - x_t)'] \\ &= \mathbb{E}[([I - K_t H_t]\tilde{x}_t^- - K_t v_t)([I - K_t H_t]\tilde{x}_t^- - K_t v_t)'] \\ &= \mathbb{E}[([I - K_t H_t]\tilde{x}_t^- - K_t v_t)((\tilde{x}_t^-)'[I - K_t H_t]' - v_t' K_t')] \\ &= \mathbb{E}[([I - K_t H_t]\tilde{x}_t^-(\tilde{x}_t^-)'[I - K_t H_t]')] - \mathbb{E}[([I - K_t H_t]\tilde{x}_t^- v_t' K_t')] - \mathbb{E}[(K_t v_t (\tilde{x}_t^-)'[I - K_t H_t]')] + \mathbb{E}[K_t v_t v_t' K_t'] \\ &= [I - K_t H_t] P_t^- [I - K_t H_t]' + K_t R_t K_t', \end{aligned} \tag{A.16}$$

where we made use of the fact that $\mathbb{E}[v_t] = 0$ and $\mathbb{E}[x_t v_t] = \mathbb{E}[x_t] + \mathbb{E}[v_t]$. We finally have to find K_t such that $trace(P_t)$ is minimized, we have

$$\begin{aligned} P_t &= [I - K_t H_t] P_t^- [I - K_t H_t]' + K_t R_t K_t' \\ &= P_t^- - K_t H_t P_t^- - P_t^- H_t' K_t' + K_t H_t P_t^- H_t' K_t' + K_t R_t K_t'. \end{aligned} \tag{A.17}$$

Accordingly[1]

$$trace(P_t) = trace(P_t^-) - 2trace(K_t H_t P_t^-) + trace(K_t H_t P_t^- H_t' K_t') + trace(K_t R_t K_t'). \tag{A.18}$$

By using the matrix identities

$$\frac{\partial trace(ABA')}{\partial A} = 2AB, \text{where B is symmetric and} \quad \frac{\partial trace(AC)}{\partial A} = C', \tag{A.19}$$

we can differentiate P_t with respect to K_t,

$$\frac{\partial trace(P_t^-)}{\partial K_t} = -2P_t^- H_t' + 2K_t H_t P_t^- H_t' + 2K_t. \tag{A.20}$$

This results in

$$K_t = P_t^- H_t' (H_t P_t^- H_t' + R)^{-1}, \tag{A.21}$$

which is, as noted earlier, the Kalman Gain. Subsitute this into (B.49) we get

$$P_t = (I - K_t H_t) P_t^-. \tag{A.22}$$

The KF algorithm is a predictor-corrector algorithm and uses the time updates for the prediction and the measurement updates as the corrector.

[1] Note that $(P_t^- H_t' K_t')' = K_t H_t P_t^-$ and that $trace(A) = trace(A')$, so $trace(P_t^- H_t' K_t') = trace(K_t H_t P_t^-)$.

The KF algorithm is given by

Predictor (Time updates)
$\hat{x}_t^- = A_t\hat{x}_{t-1}$(A.11)
$P_t^- = A_tP_{t-1}A_t' + Q_t$(A.13)

Table A.1:

Corrector (Measurement updates)
$K_t = P_k^- H_t'(H_tP_t^- H_t' + R_t)^{-1}$(A.21)
$\hat{x}_t = \hat{x}_t^- + K_t(z_t - H_t\hat{x}_t^-)$(A.14)
$P_t = (I - K_tH_t)P_t^-$(A.22)

Table A.2:

The system matrices consists of unknown parameters. If φ is the vector of unknown parameters and x_t is the vector of observed values, in order to find the appropriate parameters we maximize the loglikelihood function. This function estimation is a well-known method for calculating the parameters such that the mathematical model fits best to the data. It finds the most likely values of distribution parameters for a set of data by maximizing the value of the log-likelihood function. Considering a multivariate Gaussian distribution, the log-likelihood function is given by

$$\ln L(Y;\varphi) = -\frac{1}{2}n\ln 2\pi - \frac{1}{2}\sum_t \ln|F_t| - \frac{1}{2}\sum_t v_t'F_t^{-1}v_t, \tag{A.23}$$

where v_t are the innovations (error between the observed and numeric value) and F_t is the covariance matrix of the innovations at time t. In maximizing $\ln L$ and minimizing the innovations, an optimalization routine finds the most likely parameters. In the following section we introduce a simple example. Note that the log-likelihood function is not used here, because the system matrices are taken constant and they thus not constists of unknown parameters.

A.2 Introductory example

Following [6],we use the KF to estimate a random constant. We assume all matrices in Table B.1 and B.2 to be constant. By setting $A = 1$, we obtain $\hat{x}_k^- = \hat{x}_{k-1}$, i.e. we skip the updating step. By setting $H = 1$, we get $z_k = x_k + v_k$, i.e. the measurements come directly from the state x_k. We set $Q = 1E - 05$. Now we have to choose an initial state to begin with. Since a random variable is normally distributed with mean zero, we take $x_0 = 0$ to be the initial state. Accordingly we must start with an initial state for P_k, P_0. It turns out that we can arbitrary choose $P_0 \neq 0$ and the filter will eventually converge. Take $P_0 = 1$, by taking P_0 large enough, the choice of x_0 does not influence the Kalman filter [7]. The crosses in the following figures are generated by the matlab function

```
y=-0.37727+normrnd2(0,0.025^2,samples,1);
```

A.2.1 Simulation results

By randomly choosing $R = 0.0238$ we get,

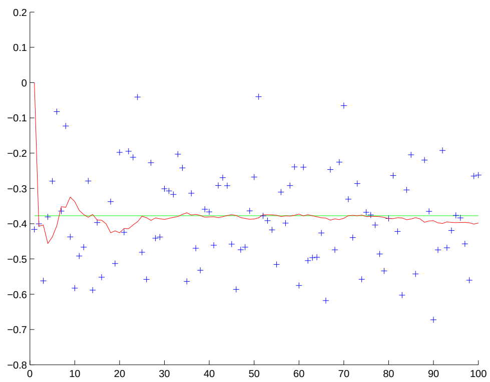


Figure A.1: Simulation of a random constant, with $R = 0.0238$. The true value of $x = -0.37727$ is given by the green line, the Kalman filter by the red line and the crosses are the noisy measurements.

Changing the choice of R is made clear by Figures A.2 and A.3. In Figure A.2 we take $R = 1$, this will cause a much slower converging behaviour than in Figure A.1. This is because the filter responds slower to the noisy measurements. However the filter will eventually converge to the true value of x. In Figure A.3 we take $R = 0.0001$, as we can see the filter quickly responds to the measurements and tries to fit it. With the choice of this R we can expect that it will take very long for the filter to converge to the green line.

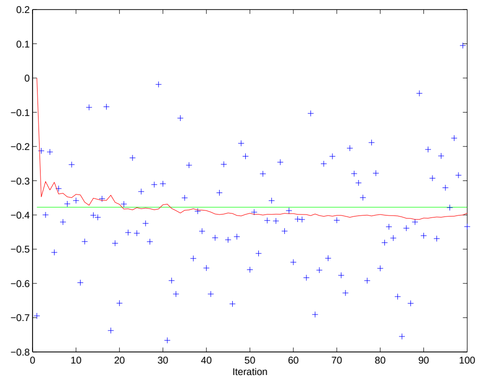


Figure A.2: Simulation of a random constant, with $R = 1$. The true value of $x = -0.37727$ is given by the green line, the Kalman filter by the red line and the crosses are the noisy measurements.

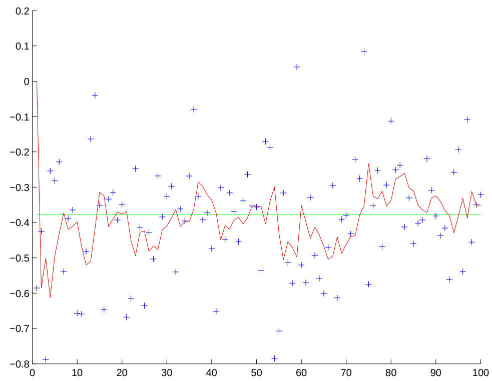


Figure A.3: Simulation of a random constant, with $R = 0.0001$. The true value of $x = -0.37727$ is given by the green line, the Kalman filter by the red line and the crosses are the noisy measurements.

A.2.2 How to choose R

Consider for simplicity $Q = 0$, so (A.2) can be written as $z_t = H_t x_t$ or $I_t = z_t - H_t x_t$, where I_t is usually referred to as the innovations. If we plot I_t in Figure A.4 we see that the crosses are distributed around zero. This is expected because R_{I_t} is statistically equal to the variance of the error term of v_t, R. We assumed v_t is $\mathcal{N}(0, R)$. But from the literature [7] it is also known that $R_{I_k} = H_t P_t^- H_t' + R$. We first choose R randomly, to say $R = 0.10$. If we plot both the statistically innovations and the theoretical ones we should have that 68 percent of the crosses lies inside the theoretical boundary. In Figure A.4 we can see that only 54 percent of the crosses lies inside the boundary.

It is therefore necessary to change the value of R, say $R = 0.16$. We see that 71 percent of the red crosses lies between the blue lines. If we now plot the Kalman filter using this R we see that this is one of the best choices for R.

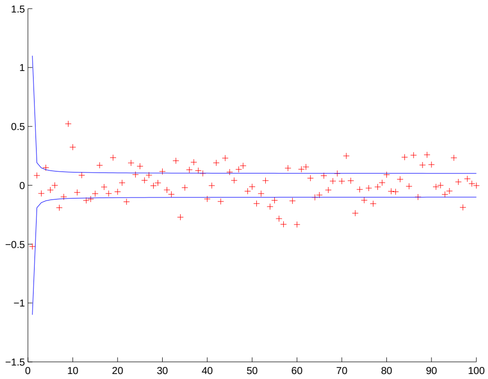


Figure A.4: Comparison between the theoretical innovations (blue lines) and the statistical innovations (red crosses). The number of red crosses between the blue lines is 54 percent. $R = 0.10$.

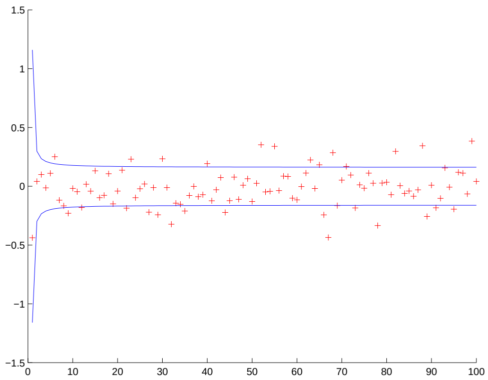


Figure A.5: Comparison between the theoretical innovations (blue lines) and the statistical innovations (red crosses). The number of red crosses between the blue lines is 71 percent. $R = 0.16$.

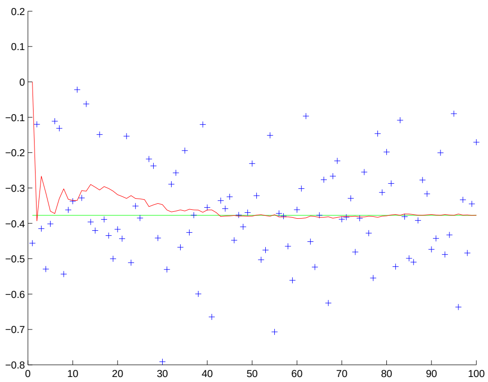


Figure A.6: Simulation of a random constant, with $R = 0.16$. The true value of $x = -0.37727$ is given by the green line, the Kalman filter by the red line and the crosses are the noisy measurements.

Appendix B

Pricing of options on commodity futures

In this appendix, the closed form solution of a European option with strike price K and maturity time t written on a futures contract with maturity time T is derived. Although we considered deterministic interest rates throughout the previous chapters, we include a stochastic interest rate in this chapter. The purpose of doing this is to show what kind of problems arise when dealing with stochastic interest rates. The derivation of the formula follows the work by Schwartz [11]. However, we give a more detailed derivation, where lengthy proofs are moved to the appendix C.
In the introduction the difference between forward and futures contracts have been made. Schwartz, in [11], showed that, considering this distinction will lead to differences in the option prices. Since we are only interested in the case where the interest rates are deterministic we give the fully derivation of the call option on a futures contract and highlight the difference between the call option on a forward contract. Definitions like change of numeraire, Radon-Nikodym derivative and the equivalent martingale measure are briefly explained. However, these definitions are only used and are not proven because this would deviate from the main goal of this thesis.

B.1 Back to basics

Consider a probability space $(\Omega, \mathscr{F}, \mathbb{P})$, where $\mathscr{F}_t$ is the filtration (i.e. all available information up to time t). We consider three adapted (i.e. filtered) stochastic processes: the spot prices S_t, the convenience yield, δ_t and the interest rate, r_t and we assume that the probability space is chosen such that the first and second moments of these processes are well defined. From (4.4) we see that the risk adjusted drift of the commodity price process should be $r - \delta_t$. Since, as noted earlier, the CY can not be hedged, the risk-adjusted CY process will have a market price of risk associated with it. For the derivation of the closed form solution, we assume for now that the interest rate is stochastic and follows, for simplicity a mean-reverting pattern. We have the joint-stochastic process

$$\begin{cases} dS_t = (r_t - \delta_t)dt + \sigma_S(t)dW_t, \\ d\delta_t = k_\epsilon(\alpha - \delta_t)dt + \sigma_\epsilon(t,t)dW_t, \\ dr_t = k_f(m - r_t)dt + \sigma_f(t,t)dW_t. \end{cases} \tag{B.1}$$

In (B.1), α and m are the mean-reverting parameters of the CY and interest rate, respectively. After the derivation of the closed form solution, we set r constant ($\sigma_f(t,t) = 0$).

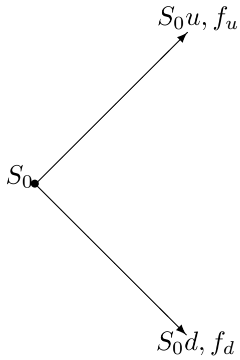
The future price, $G(t,T)$, for delivery at date $T > t$ is given by

$$G(t,T) = \mathbb{E}[S_T|\mathscr{F}_t], \tag{B.2}$$

that is, given that the spot price at time T is a random quantity S_T viewed from date t, $G(t,T)$ is the expectation taken under an equivalent martingale measure[1], conditional on the information up to time t ($\mathscr{F}_t$). The non-arbitrage conditions assumed in this thesis imply the existence of a probability measure $\mathbb{Q}$, such that the price of any basic security at time t is equal to, what is called, the $\mathbb{Q}$-expectation of its discounted future payments. Futhermore, the discounted price of these securities is a $\mathbb{Q}$-martingale between two payments dates. The mathematical explanation is given in the following sections.

B.1.1 Basic principles of change of numeraire

In order to find the formula for the forward prices, we must first explain what a change of numeraire technique is. Consider a tree



We start with S_0, the price of the commodity at $t = 0$. In one step the price can go up, denoted as S_0u with $u > 1$ or can go down with $d < 1$ denoted as, S_0d. After n steps we got our price at $t = T$, S_T. If the price goes up, the payoff is given by f_u and visa versa as f_d. Given a portfolio consisting of Π shares in the long position and one derivative in the short position. If the price goes up, the value of the portfolio is $S_u\Pi - f_u$ and if it goes down the price will be $S_d\Pi - f_d$. This portfolio is riskless if and only if the prices are equal, that is if,

$$\Pi = \frac{f_u - f_d}{S_0u - S_0d}. \tag{B.3}$$

The present value of the portfolio is $(S_0u\Pi - f_u)e^{-rT}$, because since it is riskless, it must earn the risk-free interest rate r (note that, in contrast to the beginning of this chapter, r is now assumed constant). Since setting up the portfolio costs $S_0\Pi - f$ we have

$$(S_0u\Pi - f_u)e^{-rT} = S_0\Pi - f, \quad \text{or equivalently,} \quad f = e^{-rT}(pf_u + (1-p)f_d), \tag{B.4}$$

where

$$p = \frac{e^{rT} - d}{u - d}. \tag{B.5}$$

Under the quivalent martingale measure we see that, together with (B.5)

$$\mathbb{E}[S_T] = pS_0u + (1-p)S_0d = S_0e^{rT}, \quad \text{or equivalently,} \quad S_0 = e^{-rT}\mathbb{E}[S_T]. \tag{B.6}$$

[1]The term, an equivalent martingale measure, is made clear in the following section.

In general, it holds that, $S_t = \phi^u S^u_{t+dt} + \phi^d S^d_{t+dt}$, where for simplicity we write S^u_{t+dt} for $S_0 u$ and S^d_{t+dt} for $S_0 d$. This also holds for a bond, $B_t = \phi^u B_{t+dt} + \phi^d B_{t+dt}$. For a portfolio v_t consisting of x shares of S_t and y bonds, we have

$$\begin{aligned} v_t &= x(\phi^u S^u_{t+dt} + \phi^d S^d_{t+dt}) + y(\phi^u B_{t+dt} + \phi^d B_{t+dt}) \\ &= \phi^u v^u_{t+dt} + \phi^d v^d_{t+dt}. \end{aligned} \tag{B.7}$$

So we can interpret ϕ^u and ϕ^d as the prices in the up and down states of the tree, respectively. Let L be a linear price functional and v_{t+dt} the payoff, we have

$$v_t = L(v_{t+dt}) = \phi^u v^u_{t+dt} + \phi^d v^d_{t+dt} = \langle v_{t+dt}, \phi \rangle, \tag{B.8}$$

where $\phi = [\phi^u \quad \phi^d]$. For a stock and a bond, this holds as well, i.e. let

$$L : X \to \Re \quad \text{or equivalently} \quad L : \text{payoff} \to \text{price} \tag{B.9}$$

be the pricing functional. For a stock and bond we therefore have

$$L(S_T) = S_0 \quad \text{and} \quad L(B_T) = B_0. \tag{B.10}$$

Using the *Riesz representation formula*, i.e.

Theorem B.1. *Riesz representation formula [cf. [15]]*
Let $L(\cdot) : L^2(\mathbb{P}) \to \Re$ be a continuous linear functional, then there exists a unique $\phi \in L^2(\mathbb{P})$ such that

$$L(f) = \int f \phi d\mathbb{P}$$

ϕ is known as the pricing Kernel or the stochastic discount factor.

Using this we have

$$L(S_T) = \int_0^T S_T \phi d\mathbb{P} = S_0, \tag{B.11}$$

$$L(B_T) = \int_0^T B_T \phi d\mathbb{P} = B_0. \tag{B.12}$$

Since pricing options using a tree, it is common to start with S_T and calculate back to S_0. So for T tree steps we have, in the continuous limit,

$$S_0 = \sum_0^T S_T \phi \approx \int_0^T S_T \phi d\mathbb{P}. \tag{B.13}$$

B.1.2 An equivalent martingale measure

To show what is meant the term 'equivalent martingale measure' we note that

$$1 = \frac{B_0}{B_0} = \int_0^T \frac{B_T}{B_0} \phi d\mathbb{P}. \tag{B.14}$$

So it seems that $\frac{B_T}{B_0}\phi d\mathbb{P}$ acts like a probability density, since it integrates to one. We define

$$d\mathbb{Q}^B = \frac{B_T}{B_0} \phi d\mathbb{P}. \tag{B.15}$$

Q is referred to as the forward risk neutral measure. Also

$$\frac{S_0}{B_0} = \int_0^T \frac{S_T}{B_0}\phi d\mathbb{P} = \int_0^T \frac{S_T}{B_T}\frac{B_T}{B_0}\phi d\mathbb{P} = \int_0^T \frac{S_T}{B_T} d\mathbb{Q}^B, \tag{B.16}$$

or equivalently,

$$\frac{S_0}{B_0} = \mathbb{E}^{\mathbb{Q}^B}[\frac{S_T}{B_T}]. \tag{B.17}$$

Now assume constant interest rate, r, and B_0=1 we have

$$\frac{S_0}{1} = \int_0^T \frac{S_T}{e^{rT}} d\mathbb{Q}^B = e^{-rT}\mathbb{E}^{\mathbb{Q}^B}[S_T], \tag{B.18}$$

which is the same as (B.6). This same trick holds for the stock price

$$1 = \frac{S_0}{S_0} = \int_0^T \frac{S_T}{S_0}\phi d\mathbb{P} \Rightarrow d\mathbb{Q}^S = \frac{S_T}{S_0}\phi d\mathbb{P}, \tag{B.19}$$

and so,

$$\frac{B_0}{S_0} = \int_0^T B_T S_0 \phi d\mathbb{P} = \int_0^T \frac{B_T}{S_T}\frac{S_T}{S_0}\phi d\mathbb{P} = \int_0^T \frac{B_T}{S_T} d\mathbb{Q}^S, \tag{B.20}$$

which implies,

$$\frac{B_0}{S_0} = \mathbb{E}^{\mathbb{Q}^S}[\frac{B_T}{S_T}]. \tag{B.21}$$

In brief, normalizing the bond means that everything is measured in units of the bond and normalizing in the stock means that everything is measured in units of the stock. This is called *change of numeraire*. In general the pricing formulas for S_t and P_t are given by

$$\frac{S_t}{B_t} = \mathbb{E}^{\mathbb{Q}^B}[\frac{S_T}{B_T}|\mathscr{F}_t] \quad \text{and} \quad \frac{B_t}{S_t} = \mathbb{E}^{\mathbb{Q}^S}[\frac{B_T}{S_T}|\mathscr{F}_t] \quad \text{respectively.} \tag{B.22}$$

It is known that: all assets normalized by the numeraire are martingales under the forward risk neutral probabilities with respect to the numeraire, [12]. The forward risk neutral probability is also referred to as the *equivalent martingale measure*. From this it follows that the martingales of the pricing formulas for S_t and P_t are respectively

$$\frac{S_t}{B_t} \quad \text{and} \quad \frac{B_t}{S_t}. \tag{B.23}$$

Using the result in Subsection B.1.1,

$$\frac{S_0}{B_0} = \int_0^T \frac{S_T}{B_T} d\mathbb{Q}^B = \mathbb{E}^{\mathbb{Q}^B}[\frac{S_T}{B_T}], \tag{B.24}$$

accordingly

$$S_0 = \mathbb{E}^{\mathbb{Q}^B}[\exp\left\{-\int_0^T r_s ds\right\} S_T], \tag{B.25}$$

which implies

$$S_t = \mathbb{E}^{\mathbb{Q}^B}[\exp\left\{-\int_t^T r_s ds\right\} S_T|\mathscr{F}_t]. \tag{B.26}$$

B.2 Back to pricing an option

Consider a market value M_t, where

$$dM_t = r_t M_t dt, \quad \text{with} \quad C_0 = 1 \quad \Rightarrow \quad M_t = e^{\int_0^t r_s ds}. \tag{B.27}$$

We see that M_t is not riskless since r is evolving over time. The stochastic discount factor is given by

$$\phi = \frac{M_t}{M_T} = \exp\left\{-\int_t^T r_s ds\right\}. \tag{B.28}$$

The price of an option, $V_t[\nu]$, at time t under the equivalent martingale measure, can be generally written as

$$V_t[\nu] = \mathbb{E}^{\mathbb{P}}[\nu \exp\left\{-\int_t^T r_s ds\right\}|\mathscr{F}_t], \tag{B.29}$$

where ν is a random variable (in this case it is the payoff). Note: if ν and r are independent we have

$$\begin{aligned} V_t[\nu] &= \mathbb{E}^{\mathbb{P}}[\nu \exp\left\{-\int_t^T r_s ds\right\}|\mathscr{F}_t] \\ &= \mathbb{E}^{\mathbb{P}}[\exp\left\{-\int_t^T r_s ds\right\}|\mathscr{F}_t]\mathbb{E}^{\mathbb{Q}}[\nu|\mathscr{F}_t] \\ &= P(t,T)\mathbb{E}^{\mathbb{P}}[\nu|\mathscr{F}_t], \end{aligned} \tag{B.30}$$

which is a good results because $P(t,T)$ (zero coupon bond) can be observerd directly from the markt. However the two random variables, ν and r are in general not independent, that is why the change of numeraire is important, in order to calculate the expectation in (B.29).

B.2.1 Change of numeraire and the Radon-Nikodym derivative

From [13] it follows that the change of numeraire works as follows

- choose as numeraire $P(t,T)$, that is the zero coupon price,
- choose as probability measure the T-forward risk neutral probability measure, $\mathbb{Q}^T$, defined by its *Radon-Nikodym derivative* with respect to $\mathbb{P}$

$$\frac{d\mathbb{Q}^T}{d\mathbb{P}} = \frac{\exp\left\{-\int_t^T r_s ds\right\}}{P(t,T)}. \tag{B.31}$$

To understand what is meant by the Radon-Nikodym derivative we give the theorem below and the proof of (B.31).

Theorem B.2. *Radon-Nikodym derivative [cf. [14]]*
Let $\mathbb{Q}$ and $\mathbb{P}$ be two probability measures on a space $(\Omega, \mathscr{F})$. Assume that for every $A \in \mathscr{F}$ satisfying $\mathbb{P}(A) = 0$, we also have $\mathbb{Q}(A) = 0$. Then we say that $\mathbb{P}$ is absolutely continuous with respect to $\mathbb{Q}$. Under this assumption, there is a nonnegative random variable Z such that

$$\mathbb{Q}(A) = \int_A Z d\mathbb{P}, \quad \forall A \in \mathscr{F}, \tag{B.32}$$

and Z is called the Radon-Nikodym derivative of $\mathbb{P}$ with respect to $\mathbb{Q}$.

Looking at (B.15) we see, that over a period $[t, T]$ it holds, using (B.28)

$$\begin{aligned}\frac{d\mathbb{Q}^P}{d\mathbb{P}} &= \frac{P(T,T)}{P(t,T)}\frac{M_t}{M_T} \\ &= \frac{1}{P(t,T)}\exp\left\{-\int_t^T r_s ds\right\}. \end{aligned} \tag{B.33}$$

An important difference between the discount factor ϕ from the market account M_t and the zero-copuon bond, $P(t,T)$, resides that $P(t,T)$ must be known at time t while ϕ depends on the evolution of r_t. For deterministic interest rates, ϕ=$P(t,T)$, but since our interest rate is expected to be stochastic we have

$$P(t,T) = \mathbb{E}^{\mathbb{P}}[\exp\left\{-\int_t^T r_s ds\right\}|\mathscr{F}_t], \tag{B.34}$$

which is consonant with (B.30).

B.2.2 Relation between the future and forward price

Next we substitute (B.33) into (B.29) and we define $\nu = S_T$, i.e. the price of an asset at time T, making $V_t[\nu] = V_t[S_T]$ being its purchase price today (S_t). This results in

$$\begin{aligned} V_t[\nu] &= \mathbb{E}^{\mathbb{P}}[\nu\exp\left\{-\int_t^T r_s ds\right\}|\mathscr{F}_t] \\ &= \mathbb{E}^{\mathbb{P}}[\nu P(t,T)\frac{d\mathbb{Q}^T}{d\mathbb{P}}|\mathscr{F}_t] \\ &= P(t,T)\mathbb{E}^{\mathbb{P}}[\nu\frac{d\mathbb{Q}^T}{d\mathbb{P}}|\mathscr{F}_t] \\ &= P(t,T)\mathbb{E}^{\mathbb{Q}^T}[\nu|\mathscr{F}_t] \\ V_t[S_T] &= P(t,T)\mathbb{E}^{\mathbb{Q}^T}[S_T|\mathscr{F}_t] \\ S_t &= P(t,T)\mathbb{E}^{\mathbb{Q}^T}[S_T|\mathscr{F}_t] \\ \frac{S_t}{P(t,T)} &= \mathbb{E}^{\mathbb{Q}^T}[\frac{S_T}{P(T,T)}|\mathscr{F}_t], \end{aligned} \tag{B.35}$$

since $P(T,T) = 1$. Next we note that $\frac{S_t}{P(t,T)}$ is the forward price of a non-dividend-paying stock that does not pay convenience yield. Since we do have convenience yield, specifically stochastic convenience yield we have[2],

$$S_t = \mathbb{E}^{\mathbb{P}}[S_T\exp\left\{-\int_t^T \delta_s ds\right\}\exp\left\{-\int_t^T r_s ds\right\}|\mathscr{F}_t]. \tag{B.36}$$

Next we proof that the forward price of the commodity can be written, under the equivalent martingale measure as

$$\boxed{F(t,T) = \frac{S_t - \mathrm{cov}\left(\exp\left\{\int_t^T \delta_s ds\right\}, \exp\left\{-\int_t^T r_s ds\right\} S_T|\mathscr{F}_t\right)}{P(t,T)\mathbb{E}[\exp\left\{\int_t^T \delta_s ds\right\}|\mathscr{F}_t]}} \tag{B.37}$$

and that the future price of the commodity can be written as

$$\boxed{G(t,T) = F(t,T) - \frac{S_t}{P(t,T)}\mathrm{cov}\left(\exp\left\{-\int_t^T r_s ds\right\}, \frac{S_T}{S_t}|\mathscr{F}_t\right)}. \tag{B.38}$$

[2]This results, i.e. bringing in the stochastic convenience yield into the expectation is not proven here. However, it can be found in [12]

Remark: From here on, we assume that everything is calculated under the equivalent martingale measure and we will therefore skip the measure subscript on the expectation.

Proof. We summarize the results from Appendix B. We have

$$\boxed{P(t,T) = \mathbb{E}[\exp\left\{-\int_t^T r_s ds\right\}|\mathscr{F}_t]}, \tag{B.39}$$

$$\boxed{S_t = \mathbb{E}[S_T \exp\left\{-\int_t^T \delta_s ds\right\}\exp\left\{-\int_t^T r_s ds\right\}|\mathscr{F}_t]}, \tag{B.40}$$

$$\boxed{G(t,T) = \mathbb{E}[S_T|\mathscr{F}_t]}, \tag{B.41}$$

$$\boxed{F(t,T) = \frac{\mathbb{E}[\exp\left\{-\int_t^T r_s ds\right\} S_T|\mathscr{F}_t]}{P(t,T)}}. \tag{B.42}$$

Rewriting the second equility gives

$$\begin{aligned} S_t &= \mathbb{E}[S_T \exp\left\{-\int_t^T \delta_s ds\right\}\exp\left\{-\int_t^T r_s ds\right\}|\mathscr{F}_t] \\ &= \mathbb{E}[\exp\left\{\int_t^T \delta_s ds\right\}|\mathscr{F}_t]\mathbb{E}[\exp\left\{-\int_t^T r_s ds\right\} S_T|\mathscr{F}_t] \\ &\quad + \text{cov}\left(\exp\left\{\int_t^T \delta_s ds\right\}, \exp\left\{-\int_t^T r_s ds\right\} S_T|\mathscr{F}_t\right) \\ &= \mathbb{E}[\exp\left\{\int_t^T \delta_s ds\right\}|\mathscr{F}_t]P(t,T)F(t,T) + \text{cov}\left(\exp\left\{\int_t^T \delta_s ds\right\}, \exp\left\{-\int_t^T r_s ds\right\} S_T|\mathscr{F}_t\right). \end{aligned} \tag{B.43}$$

Rewriting (B.42) results in

$$\begin{aligned} F(t,T) &= \frac{\mathbb{E}[\exp\left\{-\int_t^T r_s ds\right\} S_T|\mathscr{F}_t]}{P(t,T)} \\ &= \frac{1}{P(t,T)}\left(\text{cov}\left(\exp\left\{-\int_t^T r_s ds\right\}, S_T|\mathscr{F}_t\right)\right) + \mathbb{E}[\exp\left\{-\int_t^T r_s ds\right\}]\mathbb{E}[S_T|\mathscr{F}_t] \\ &= \frac{1}{P(t,T)}\left(\text{cov}\left(\exp\left\{-\int_t^T r_s ds\right\}, S_T|\mathscr{F}_t\right)\right) + G(t,T) \end{aligned} \tag{B.44}$$

Now the substitution of (B.44) into (B.42) results in

$$\begin{aligned} S_t &= \mathbb{E}[\exp\left\{\int_t^T \delta_s ds\right\}|\mathscr{F}_t]P(t,T)G(t,T) + \mathbb{E}[\exp\left\{\int_t^T \delta_s ds\right\}|\mathscr{F}_t]\cdot\text{cov}\left(\exp\left\{-\int_t^T r_s ds\right\}, S_T|\mathscr{F}_t\right) \\ &\quad + \text{cov}\left(\exp\left\{\int_t^T \delta_s ds\right\}, \exp\left\{-\int_t^T r_s ds\right\} S_T|\mathscr{F}_t\right). \end{aligned} \tag{B.45}$$

(B.43) and(B.45) implies that (B.37) and (B.38) holds.

Remark: Note that for the derivation of this proof we use a standard results, namely, for two dependent random variables, X and Y it holds that

$$\mathbb{E}[XY] = \mathbb{E}[X]\mathbb{E}[Y] + \text{cov}(X, Y). \tag{B.46}$$

The dependency of the random variables $e^{\int_t^T \delta_s ds}$ and $e^{-\int_t^T r_s ds} S_T$ (used in (B.43)) and between the random variables S_T and $e^{-\int_t^T r_s ds}$ (used in (B.44)) clearly holds. □

B.3 Instantaneous forward interest rate and the Heath-Jarrow-Morton framework

The Heath-Jarrow-Morton (HJM) framework is a general framework to model the evolution of interest rates. The main idea of this framework is that the drift of the instantanuous interest rate depends on the volatility, i.e. no drift estimation is needed. The HJM is applicable to long rate models because it captures to whole dynamics of the interest rate curve.

Theorem B.3. *HJM-framework, [cf. [27]]*
The instantanuous forward interest rate prevailing at time t for the maturity $T > t$ is denoted by $f(t, T)$ and is defined by

$$f(t, T) := \lim_{S \to T} F(t; T, S) = -\frac{\partial lnP(t, T)}{\partial T}, \tag{B.47}$$

which gives the well-known result

$$P(t, T) = e^{-\int_t^T f(t,s)ds}. \tag{B.48}$$

Intuitively we can say that the instantanuous forward interest rate $f(t, T) \approx F(t; T, T + \Delta T)$, that is, its maturity is very close to its expiry T (considering ΔT small). For pricing the forward prices we would like the same approach for the forward prices of commodities (see (B.37)). So

$$F(t, T) = \frac{S(t)}{P(t, T)} e^{-\int_t^T \delta(t,s)ds} = S_t e^{\int_t^T (f(t,s) - \delta(t,s))ds}, \tag{B.49}$$

where $\delta(t, s)$ is the continuously compounded forward convenience yield which, as noted in the introduction, only accrues to the holder of the physical commodity and not to the holder of the contract on the commodity. From (B.37) and (B.49) it follows that we have

$$e^{-\int_t^T \delta(t,s)ds} = \frac{1 - \text{cov}\left(\exp\left\{\int_t^T \delta_s ds\right\}, \exp\left\{-\int_t^T r_s ds\right\} S_T | \mathscr{F}_t\right)}{\mathbb{E}[\exp\left\{\int_t^T \delta_s ds\right\} | \mathscr{F}_t]} \tag{B.50}$$

To get a feeling of the economical interpretation of $\delta(t, s)$ consider buying one forward contract at date t for future delivery of the commmodity at maturity T and selling one forward contract at date t for future delivery of the commodity at date $T + \Delta$. So in this period you own the physical commodity and so $\delta(t, s)$ is the premium of the difference between the value of the commodity at date T to $T + \Delta$ per unit time per unit of the commodity. To see this mathematically, we have

$$\begin{aligned} F(t, T) - F(t, T + \Delta) &= \frac{S_t e^{-\int_t^T \delta(t,s)ds}}{P(t, T)} - \frac{S_t e^{-\int_t^{T+\Delta} \delta(t,s)ds}}{P(t, T + \Delta)} \\ F(t, T)P(t, T) - F(t, T + \Delta)P(t, T + \Delta) &= S_t e^{-\int_t^T \delta(t,s)ds} - S_t e^{-\int_t^T \delta(t,s)ds} e^{-\int_T^{T+\Delta} \delta(t,s)ds} \\ &= F(t, T)P(t, T)(1 - e^{-\int_T^{T+\Delta} \delta(t,s)ds}), \end{aligned} \tag{B.51}$$

dividing LHS by Δ and taking limits ($\Delta \downarrow 0$) we have

$$\begin{aligned}
\lim_{\Delta \downarrow 0} \frac{F(t,T)P(t,T) - F(t,T+\Delta)P(t,T+\Delta)}{\Delta} &= -\frac{\partial}{\partial T}(F(t,T)P(t,T)) \\
&= -\frac{\partial}{\partial T}(S_t e^{-\int_t^T \delta(t,s)ds}) \\
&= \delta(t,T) S_T e^{-\int_t^T \delta(t,s)ds} \\
&= P(t,T)F(t,T)\delta(t,T). \qquad \text{(B.52)}
\end{aligned}$$

In other words we have

$$\lim_{\Delta \downarrow 0} \frac{F(t,T)P(t,T) - F(t,T+\Delta)P(t,T+\Delta)}{P(t,T)F(t,T)\Delta} = \delta(t,T), \qquad \text{(B.53)}$$

which corresponds to the economic interpretation.
Relation (B.48) is performed under no-arbitrage conditions. We therefore should have the relation

$$f(t,t) = r_t. \qquad \text{(B.54)}$$

We can see this by (step 1)[3] differentiating the LHS and RHS of (B.48) with respect to T, (step 2) dividing by $P(t,T)$ and (step 3) taking the limit ($T \downarrow t$).

$$\begin{aligned}
&\boxed{\text{Step 1}} \\
\frac{\partial P(t,T)}{\partial T} &= \frac{\partial}{\partial T}\mathbb{E}[e^{-\int_t^T r_s ds}|\mathscr{F}_t] = \frac{\partial}{\partial T} e^{-\int_t^T f(t,s)ds} \\
&= \mathbb{E}[\frac{\partial}{\partial T} e^{-\int_t^T r_s ds}|\mathscr{F}_t] = -f(t,T)P(t,T) \\
&= r_T P(t,T) = f(t,T)P(t,T) \\
&\boxed{\text{Step 2}} \\
r_T &= f(t,T) \\
&\boxed{\text{Step 3}} \\
\lim_{T \downarrow t}(r_T) &= \lim_{T \downarrow t} f(t,T) \\
r_t &= f(t,t). \qquad \text{(B.55)}
\end{aligned}$$

In order words, considering an infinitesimal interval ($T \downarrow t$) we should have that the instantanuous interest rate is equal to the forward interest rate at time t. If this was not true, arbitrage possibilities can arise by lending money with one interest rate and putting money on the bank with the other one.

[3] In step 1 we assume all usual rules of interchange derivative operator and expectation.

A similar task can be done for (B.50) and (B.60) (note that step 2[4] is not needed). For the former we have

$$\boxed{\text{Step 1}}$$

$$\begin{aligned}
\frac{\partial}{\partial T}e^{-\int_t^T \delta(t,s)ds} &= \frac{\partial}{\partial T}\left\{\frac{1-\mathrm{cov}\left(\exp\left\{\int_t^T \delta_s ds\right\},\exp\left\{-\int_t^T r_s ds\right\}S_T|\mathscr{F}_t\right)}{\mathbb{E}[\exp\left\{\int_t^T \delta_s ds\right\}|\mathscr{F}_t]}\right\} \\
-\delta(t,T)e^{-\int_t^T \delta(t,s)ds} &= \frac{\mathbb{E}[e^{-\int_t^T \delta_s ds}|\mathscr{F}_t]\left(-\mathrm{cov}\left(\frac{\partial}{\partial T}e^{-\int_t^T \delta_s ds},\frac{\partial}{\partial T}e^{-\int_t^T r_s ds}\frac{S_T}{S_t}|\mathscr{F}_t\right)\right)}{\mathbb{E}[e^{\int_t^T \delta_s ds}|\mathscr{F}_t]^2} \\
&\quad -\frac{\left(1-\mathrm{cov}\left(e^{-\int_t^T \delta_s ds},e^{-\int_t^T r_s ds}\frac{S_T}{S_t}|\mathscr{F}_t\right)\right)\mathbb{E}[\frac{\partial}{\partial T}e^{\int_t^T \delta_s ds}|\mathscr{F}_t]}{\mathbb{E}[e^{\int_t^T \delta_s ds}|\mathscr{F}_t]^2} \\
&= \frac{\delta_T r_T \mathrm{cov}\left(e^{\int_t^T \delta_s ds},e^{-\int_t^T r_s ds}\frac{S_T}{S_t}|\mathscr{F}_t\right)-\delta_T(1-\mathrm{cov}\left(e^{-\int_t^T \delta_s ds},e^{-\int_t^T r_s ds}\frac{S_T}{S_t}|\mathscr{F}_t\right))}{\mathbb{E}[e^{\int_t^T \delta_s ds}|\mathscr{F}_t]}
\end{aligned}$$

$$\boxed{\text{Step 3}}$$

$$\delta(t,t) = \delta_t, \tag{B.56}$$

where is step 3 we made use of the fact that $\mathrm{cov}(1,1)=0$. For the latter we have the same procedure which together results into

$$\delta(t,t)=\epsilon(t,t)=\delta_t. \tag{B.57}$$

And so we have, consonant with the no-arbitrage principle

$$F(t,t)=G(t,t)=S_t. \tag{B.58}$$

A similar task, which resulted in (B.50) can be done for the future prices

$$G(t,T)=\frac{S_t}{P(t,T)}e^{-\int_t^T \epsilon(t,s)ds}=S_t e^{\int_t^T (f(t,s)-\epsilon(t,s))ds}, \tag{B.59}$$

accordingly (from (B.38))

$$\begin{aligned}
e^{\int_t^T \epsilon(t,s)ds} &= \frac{1-\mathrm{cov}\left(\exp\left\{\int_t^T \delta_s ds\right\},\exp\left\{-\int_t^T r_s ds\right\}\frac{S_T}{S_t}|\mathscr{F}_t\right)}{\mathbb{E}[\exp\left\{\int_t^T \delta_s ds\right\}|\mathscr{F}_t]} \\
&- \mathrm{cov}\left(\exp\left\{-\int_t^T r_s ds\right\},\frac{S_T}{S_t}|\mathscr{F}_t\right),
\end{aligned} \tag{B.60}$$

where $\epsilon(t,s)$ is the future convenience yield.[5]

[4] In step 2 we assume all usual rules of interchange derivative operator with the covariance.

[5] The future convenience yield is just a definition. The economical interpretation is difficult to give. Since we are only interested in deterministic interest rates, we do not even use this definition and so we do not devote too much attention to all the economical interpretations of these definitions.

B.4 Drift calculation for the forward interest rate, future convenience yield and the spot price

For the three stochastic processes we have (cf. e.g. [28], [29] or [11])

$$\begin{aligned} f(t,s) &= f(0,s) + \int_0^t \mu_f(u,s)ds + \int_0^t \sigma_f(u,s)\cdot dW_u, && \text{(B.61)}\\ \epsilon(t,s) &= \epsilon(0,s) + \int_0^t \mu_\epsilon(u,s)du + \int_0^t \sigma_\epsilon(u,s)\cdot dW_u, && \text{(B.62)}\\ S_t &= S_0 + \int_0^t S_u\mu_S(u)du + \int_0^t S_u\sigma_S(u)\cdot dW_u, && \text{(B.63)} \end{aligned}$$

where we used for all three processes the same k-dimensional Brownian motions, which explains the possible correlations among the processes.[6] This notation will be used further on in this chapter.
It follows that working under an equavalent martingale measure and assuming no arbitrage possibilities that we can write an explicit form of the drift terms of the above processes. From Section 2.2 and (B.57) it follows that

$$\mu_S(t) = r_t - \delta_t = f(t,t) - \epsilon(t,t) \tag{B.64}$$

Furthermore

$$\mu_f(t,s) = \sigma_f(t,s)\cdot\left(\int_s^t \sigma_f(t,v)dv\right). \tag{B.65}$$

Before we give a mathematical proof, we first give some facts about the forward interest rate. From Theorem 8.3

$$f(t,T) = -\frac{\partial \ln \mathbb{E}[e^{-\int_t^T r_s ds}]}{\partial T}. \tag{B.66}$$

Since $f(t,T)$ depends on an expectation, it implies some kind of dynamics for the forward interest rate. Heath, Jarrow and Morton [28] showed that this dynamics is given by

$$d(f(t,s)) = \mu_f(t,s)dt + \sigma_f(t,s)\cdot dW_t, \tag{B.67}$$

where $\mu_f(t,s)$ can not be chosen arbitrarily but must equal a some kind of relationship with $\sigma_f(t,s)$. That is, there is no need to estimate the drift terms and one can only pay attention to how to calibrate the volatility.

Proof. The proof can be found in Appendix C.1. □

For $\mu_\epsilon(t,s)$ we have the relationship

$$\mu_\epsilon(t,s) = \sigma_f(t,T)\cdot\left(\int_t^T \sigma_f(t,s)ds\right) + (\sigma_f(t,T)-\sigma_\epsilon(t,T))\cdot\left(\sigma_S(t) + \int_t^T(\sigma_f(t,s)-\sigma_\epsilon(t,s))ds\right) \tag{B.68}$$

Furthermore, from (C.6) it follows (see e.g. [28] or [11])

$$\begin{aligned} P(t,T) &= P(0,T) + \int_0^t P(u,T)\left(f(u,u) - \int_u^T u_f(u,s)ds + \frac{1}{2}\left\|\int_u^T \sigma_f(u,s)ds\right\|^2\right)du \\ &\quad - \int_0^t P(u,T)\left(\int_u^T \sigma_f(u,s)ds\right)\cdot dW_u. \end{aligned} \tag{B.69}$$

[6] Here (,·,) is the inner product, i.e. $\int_0^t \sum_{i=1}^n \sigma_f^i(u,s)dW_u^i$.

Proof. The proof can be found in Appendix C.2. □

Now we have derived explicit formulas for every drift term of the three stochastic processes given in the beginning of Section B.2.2. We can now price a K-strike call option on a futures contract using the formula

$$C^G = \mathbb{E}[e^{-\int_0^t f(s,s)ds}(G(t,s) - K)^+]. \tag{B.70}$$

In the following section, a closed form solution of this expression is derived.

B.5 Closed form solution of a call option on a futures contract

This section is again based on the calculations in [11]. However a detailed derivation is reproduced and given below. To stay in the same notation as in [11] define

$$\sigma_{P_t}(u) = -\int_u^t \sigma_f(u,s)ds, \tag{B.71}$$

$$\sigma_{G_T}(u) = \sigma_S(u) + \int_u^T (\sigma_f(u,s) - \sigma_\epsilon(u,s))ds. \tag{B.72}$$

We can write

$$e^{-\int_0^t f(s,s)ds} = Ae^{-X}, \tag{B.73}$$

if we define[7], with (B.61),

$$X = \int_0^t \int_0^t \sigma_f(u,s) \cdot dW_u ds = -\int_0^t \sigma_{P_t}(u) \cdot dW_u, \tag{B.74}$$

and in A the non-stochastic terms are hidden. The exact form is not important. Similar, we can write

$$G(t,T) = Be^Z, \tag{B.75}$$

where, with (C.11),

$$Z = \int_0^t \left(\sigma_S(u) + \int_u^T (\sigma_f(u,s) - \sigma_\epsilon(u,s))ds \right) \cdot dW_u = \int_0^t \sigma_{G_T}(u) \cdot dW_u \tag{B.76}$$

and in B the non-stochastic terms are hidden. Again the exact form of B does not come up. Since X and Z are both desribed using the same Brownian motion and since σ_{G_T} and σ_{P_t} are deterministic, we can conclude that (X, Z) is jointly normally distributed with mean zero and variances and covariance given by

$$\sigma_X^2 = \int_0^t \|\sigma_{P_t}(u)\|^2 du, \tag{B.77}$$

$$\sigma_Z^2 = \int_0^t \|\sigma_{G_T}(u)\|^2 du, \tag{B.78}$$

$$\sigma_{XZ} = -\int_0^t \sigma_{P_t}(u)\sigma_{G_T}(u)du. \tag{B.79}$$

[7]Where the approbation of interchanging stochastic with integral operator is assumed.

Via the *tower rule*[8] and (B.70) we get

$$\begin{aligned} C^G &= A\mathbb{E}[e^{-X}(Be^Z - K)^+] \\ &= A\mathbb{E}[\mathbb{E}[e^{-X}|Z](Be^Z - K)^+]. \end{aligned} \tag{B.80}$$

The closed-from solution for the price of a European call option with maturity t and exercise price K written on the commodity future prices with maturity T, and is given by

$$\boxed{C^G = P(0,t)\left(G(0,T)e^{-\sigma_{xz}}\mathscr{N}\left(\frac{\log(\frac{G(0,T)}{K}) - \sigma_{xz} + \frac{1}{2}\sigma_z^2}{\sigma_z}\right) - K\mathscr{N}\left(\frac{\log(\frac{G(0,T)}{K}) - \sigma_{xz} - \frac{1}{2}\sigma_z^2}{\sigma_z}\right)\right)}. \tag{B.81}$$

Proof. The proof is given in Appendix C.3. □

B.6 Relation between forward and future prices

This section will briefly highlight the relation between the forward and future prices, but we will not derivate the formula. However, every detail of the derivation can be found in [11].

Lemma B.1. *Relation between forward and future prices*

$$F(t,T) = G(t,T)H(t,T), \tag{B.82}$$

where

$$H(t,T) = \exp\left\{-\int_t^T \left(\int_u^T \sigma_f(u,s)ds\right)\cdot\left(\sigma_S(u) + \int_u^T (\sigma_f(u,s) - \sigma_\epsilon(u,s))ds\right)du\right\}. \tag{B.83}$$

In (B.1), $\sigma_f(t,t) = 0$ if we consider deterministic interest rates resulting in $H(t,T) = 1$ and thus $F(t,T) = G(t,T)$. In the next chapter, (B.81) will be used to calibrate the prices.

[8] Tower rule: let X be an integrable random variable (i.e. $\mathbb{E}[|X|] < \infty$) and Y is any random variable living on the same probability space, then $\mathbb{E}[X] = \mathbb{E}[\mathbb{E}[X|Y]]$.

Appendix C

Proofs

C.1 Proof of equation (B.65)

Define

$$\varpi(t,s) := -\int_0^s f(t,u)du, \tag{C.1}$$

and so via Itô's lemma

$$\begin{aligned} d\varpi(t,s) &= f(t,t)dt - \int_t^s df(t,u)du \\ &= r_t dt - \int_t^s [\mu_f(t,u)dt + \sigma_f(t,u)\cdot dW_t]du. \end{aligned} \tag{C.2}$$

Assume that the Stochastic Fubini Theorem is fullfilled, i.e. we can interchange stochastic and the integral operator and define

$$\begin{aligned} \mu_f^*(t,s) &= \int_t^s \mu_f(t,u)du \\ \sigma_f^*(t,s) &= \int_t^s \sigma_f(t,u)du \end{aligned} \tag{C.3}$$

and get

$$d\varpi(t,s) = r_t dt - \mu_f^*(t,s)dt - \sigma_f^*(t,s)\cdot dW_t. \tag{C.4}$$

Furthermore

$$P(t,s) = e^{\varpi(t,s)}. \tag{C.5}$$

Applying Itô's lemma we have

$$\begin{aligned} dP(t,s) &= e^{\varpi(t,s)}d\varpi(t,s) + \frac{1}{2}e^{\varpi(t,s)}(d\varpi(t,s))^2 \\ &= P(t,s)(r_t - \mu_f^*(t,s) + \frac{1}{2}\sigma_f^*(t,s)\cdot\sigma_f^*(t,s))dt - P(t,s)\sigma_f^*(t,s)\cdot dW_t \end{aligned} \tag{C.6}$$

Implying no arbitrage conditions we must have $dP(t,s) = r_t dt$ and so

$$\mu_f^*(t,s) = \frac{1}{2}\sigma_f^*(t,s)\cdot\sigma_f^*(t,s) = \frac{1}{2}\left\|\int_t^s \sigma_f(t,u)du\right\|^2. \tag{C.7}$$

Differentiating LHS and RHS of (C.7) with respect to s gives

$$\mu_f(t,s) = \sigma_f(t,s)\cdot\left(\int_s^t \sigma_f(t,u)du\right), \tag{C.8}$$

which is the desired result.

C.2 Proof of equation (B.68)

From equation (C.6) we have in other words, $P(t,T)$ is defined by $P(0,T)$ + P times the interest rate drift term + P times the interest rate volatility[1]. This also holds for $Y(t,T)$, defined by $e^{-\int_t^T (f(t,s)-\epsilon(t,s))ds}$, using the dynamics of the forward interest rates and the future convenience yields, so

$$\begin{aligned} Y(t,T) &= Y(0,T) + \int_0^t Y(u,T)\Big(-f(u,u) + \epsilon(u,u) + \int_u^T \mu_f(u,s)ds - \int_u^T \mu_\epsilon(u,s)ds \\ &\quad + \frac{1}{2}\left\|\int_u^T \sigma_f(u,s)ds\right\|^2 + \frac{1}{2}\left\|\int_u^T \sigma_\epsilon(u,s)ds\right\|^2 - \left(\int_u^T \sigma_f(u,s)ds\right)\cdot\left(\int_u^T \sigma_\epsilon(u,s)ds\right)\Big)du \\ &\quad + \int_0^t Y(u,T)\left(\int_u^T \sigma_f(u,s)ds - \int_u^T \sigma_\epsilon(u,s)ds\right)\cdot dW_u. \end{aligned} \tag{C.9}$$

In defining $Y(t,T)$ in this way, the future price of the commodity can be written as

$$G(t,T) = S_t Y(t,T). \tag{C.10}$$

Using the two-dimensional Itô lemma, it yields[2]

$$\begin{aligned} G(t,T) &= S_0 Y(0,T) + \int_0^t S_u Y(u,T)\Big(-f(u,u) + \epsilon(u,u) + \int_u^T \mu_f(u,s)ds - \int_u^T \mu_\epsilon(u,s)ds \\ &\quad + \frac{1}{2}\left\|\int_u^T \sigma_f(u,s)ds\right\|^2 + \frac{1}{2}\left\|\int_u^T \sigma_\epsilon(u,s)ds\right\|^2 - \left(\int_u^T \sigma_f(u,s)ds\right)\cdot\left(\int_u^T \sigma_\epsilon(u,s)ds\right)\Big)du \\ &\quad + \int_0^t S_u Y(u,T)\left(\int_u^T \sigma_f(u,s)ds - \int_u^T \sigma_\epsilon(u,s)ds\right)\cdot dW_u \\ &\quad + \int_0^t Y(u,T) S_u \mu_S(u)du + \int_0^t Y(u,T) S_u \sigma_S(u)\cdot dW_u \\ &\quad + \underbrace{\int_0^t Y(u,T) S_u \left(\sigma_S(u)\cdot\left(\int_u^T (\sigma_f(u,s) - \sigma_\epsilon(u,s))ds\right)\right)du}_{\text{correction term due to Itô's lemma}} \\ &= G(0,T) + \int_0^t G(u,T)\Big(-\int_u^T \mu_\epsilon(u,s)ds + \left\|\int_u^T \sigma_f(u,s)ds\right\|^2 + \frac{1}{2}\left\|\int_u^T \sigma_\epsilon(u,s)ds\right\|^2 \\ &\quad - \left(\int_u^T \sigma_f(u,s)ds\right)\cdot\left(\int_u^T \sigma_\epsilon(u,s)ds\right) + \sigma_S(u)\cdot\left(\int_u^T (\sigma_f(u,s) - \sigma_\epsilon(u,s)ds)\right)\Big)du \\ &\quad + \int_0^t G(u,T)\left(\sigma_S(u) + \int_u^T (\sigma_f(u,s)\sigma_\epsilon(u,s))ds\right)\cdot dW_u. \end{aligned} \tag{C.11}$$

Since we are working under an equivalent martingale measure, the future price process is a martingale, that is, the drift term should equal zero ($\mathbb{E}[G(\cdot,T)] = G(0,T)$). In fact, any zero-drift process is a

[1]Note that the same explanation holds for the dynamics for the forward interest rates, future convenience yield and the spot prices, see e.g. (B.61).

[2]Note that in the second equality we substituted (C.10), (B.65) and (B.64).

martingale process.

$$
\begin{aligned}
&-\int_u^T \mu_\epsilon(u,s)ds + \left\| \int_u^T \sigma_f(u,s)ds \right\|^2 + \frac{1}{2}\left\| \int_u^T \sigma_\epsilon(u,s)ds \right\|^2 \\
&-\left(\int_u^T \sigma_f(u,s)ds\right)\cdot\left(\int_u^T \sigma_\epsilon(u,s)ds\right) + \sigma_S(u)\cdot\left(\int_u^T(\sigma_f(u,s)-\sigma_\epsilon(u,s)ds)\right) = 0 \quad \text{(C.12)}
\end{aligned}
$$

Differentiating with respect to T, substituting $u = t$ and gathering terms it follows

$$
\begin{aligned}
\mu_\epsilon(t,T) &= 2\sigma_f(t,T)\int_t^T \sigma_f(t,s)ds + \sigma_\epsilon(t,T)\int_t^T \sigma_\epsilon(t,s)ds - \sigma_f(t,T)\int_t^T \sigma_\epsilon(t,s)ds \\
&\quad -\sigma_\epsilon(u,s)\int_t^T \sigma_f(u,s)ds + \sigma_S(t)\cdot(\sigma_f(t,T)-\sigma_\epsilon(t,T)) \\
&= \sigma_f(t,T)\cdot\left(\int_t^T \sigma_f(t,s)ds\right) + (\sigma_f(t,T)-\sigma_\epsilon(t,T))\cdot\left(\sigma_S(t) + \int_t^T(\sigma_f(t,s)-\sigma_\epsilon(t,s))ds\right). \quad \text{(C.13)}
\end{aligned}
$$

C.3 Proof of equation (B.81)

We want to calculate $\mathbb{E}[\mathbb{E}[e^{-X}|Z]]$ and in doing so we must know what the conditional distribution of X given Z is, i.e. we proof

$$X|Z = z \sim \mathcal{N}\left(z\frac{\sigma_{xz}}{\sigma_z^2}, \sigma_x^2(1-\frac{\sigma_{xz}^2}{\sigma_x^2\sigma_z^2})\right). \quad \text{(C.14)}$$

Proof. Firstly, the joint density function of X and Z is defined by

$$f(x,z) = \frac{1}{2\pi\sigma_x\sigma_z\sqrt{1-\rho^2}}\exp\left\{-\frac{1}{2(1-\rho^2)}\left[\left(\frac{x-\mu_x}{\sigma_x}\right)^2 + \left(\frac{z-\mu_z}{\sigma_z}\right)^2 - 2\rho\frac{(x-\mu_x)(z-\mu_z)}{\sigma_x\sigma_z}\right]\right\}. \quad \text{(C.15)}$$

Since (X,Z) are two zero-mean random variables having a bivariate normal distribution we must calculate

$$\mathbb{E}[X|Z=z] = \int_{-\infty}^{\infty} x f_{X|Z}(x|z)dx, \quad \text{(C.16)}$$

i.e. we have to determine the conditional density of $X|Z = z$ and in doing so we collect all factors that do not depend on x, which be denoted as the constant C_i.

$$
\begin{aligned}
f_{X|Z}(x|z) &= \frac{f(x,z)}{f_z(z)} \\
&= C_1 f(x,z) \\
&= C_2\exp\left\{-\frac{1}{2(1-\rho^2)}\left[(\frac{x}{\sigma_x^2})^2 - 2\rho\frac{x(z)}{\sigma_x\sigma_z}\right]\right\} \\
&= C_3\exp\left\{-\frac{1}{2\sigma_x^2(1-\rho^2)}\left[x^2 - 2x(\rho\frac{\sigma_x}{\sigma_z}(z))\right]\right\} \\
&= C_4\exp\left\{-\frac{1}{2\sigma_x^2(1-\rho^2)}\left[x-\rho\frac{\sigma_x}{\sigma_z}(z)\right]^2\right\}, \quad \text{(C.17)}
\end{aligned}
$$

since (C.17) is proportional[3] to a

$$\mathcal{N} \sim \left(z\rho\frac{\sigma_x}{\sigma_z}, \sigma_x^2(1-\rho^2) \right), \tag{C.18}$$

we have that, together with

$$\rho = \frac{\text{Cov(X,Z)}}{\sigma_x\sigma_z} = \frac{\sigma_{xz}}{\sigma_x\sigma_z}, \tag{C.19}$$

the final result

$$X|Z = z \sim \mathcal{N}\left(\frac{\sigma_{xz}}{\sigma_z^2}z, \sigma_x^2(1-\frac{\sigma_{xz}^2}{\sigma_x^2\sigma_z^2}) \right). \tag{C.20}$$

□

From this we can calculate the conditional expectation as

$$\mathbb{E}[e^{-X}|Z=z] = e^{-\frac{\sigma_{xz}}{\sigma_z^2}z + \frac{1}{2}\sigma_x^2(1-\frac{\sigma_{xz}^2}{\sigma_x^2\sigma_z^2})}. \tag{C.21}$$

Hence, equation (B.70) can be written as

$$C^G = Ae^{\frac{1}{2}\sigma_x^2(1-\frac{\sigma_{xz}^2}{\sigma_x^2\sigma_z^2})}\mathbb{E}[e^{-Z\frac{\sigma_{xz}}{\sigma_z^2}}(Be^Z - K)^+]. \tag{C.22}$$

Note[4]

$$\begin{aligned} C^G &= Ae^{\frac{1}{2}\sigma_x^2(1-\frac{\sigma_{xz}^2}{\sigma_x^2\sigma_z^2})}\mathbb{E}[e^{-Z\frac{\sigma_{xz}}{\sigma_z^2}}(Be^Z - K)^+] \\ &= Ae^{\frac{1}{2}\sigma_x^2(1-\frac{\sigma_{xz}^2}{\sigma_x^2\sigma_z^2})}\mathbb{E}[\mathcal{I}_{Z>\log(\frac{K}{B})}e^{-Z\frac{\sigma_{xz}}{\sigma_z^2}}(Be^Z - K)] \\ &= ABe^{\frac{1}{2}\sigma_x^2(1-\frac{\sigma_{xz}^2}{\sigma_x^2\sigma_z^2})}\mathbb{E}[\mathcal{I}_{Z>\log(\frac{K}{B})}e^{Z(1-\frac{\sigma_{xz}}{\sigma_z^2})}] \\ &\quad -AKe^{\frac{1}{2}\sigma_x^2(1-\frac{\sigma_{xz}^2}{\sigma_x^2\sigma_z^2})}\mathbb{E}[\mathcal{I}_{Z>\log(\frac{K}{B})}e^{-Z\frac{\sigma_{xz}}{\sigma_z^2}}], \end{aligned} \tag{C.23}$$

where $\mathcal{I}_{Z>\log(\frac{K}{B})}$ is the indicator function.
Now calculate the two expectations. For the former introduce $b = (1 - \frac{\sigma_{xz}}{\sigma_z^2})$ for notational convenience.

[3] Again, μ_z and μ_x are zero.
[4] $(Be^Z - K)^+ = \max(0, (Be^Z - K))$ and so it only counts for $Z > \log(\frac{K}{B})$.

We have

$$\begin{aligned}
\mathbb{E}[\mathcal{I}_{Z>\log(\frac{K}{B})}e^{bZ}] &= \underbrace{\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}\sigma_z}\mathcal{I}_{Z>\log(\frac{K}{B})}\exp\left(-\frac{z^2}{2\sigma_z^2}\right)dz}_{\text{Integrate against the density of a random variable with } \mathcal{N}(0,\sigma_z)} \\
&= \int_{\log(Z>\frac{K}{B})}^{\infty}\frac{1}{\sqrt{2\pi}\sigma_z}e^{bz}\exp\left(-\frac{z^2}{2\sigma_z^2}\right)dz \\
&= \underbrace{\int_{\log(Z>\frac{K}{B})/\sigma_z}^{\infty}\frac{1}{\sqrt{2\pi}}e^{b\sigma_z u-\frac{1}{2}u^2}du}_{\text{Substitute } z=\sigma_z u} \\
&= \underbrace{\int_{\log(Z>\frac{K}{B})/\sigma_z-b\sigma_z}^{\infty}\frac{1}{\sqrt{2\pi}}e^{b\sigma_z(v+b\sigma_z)-\frac{1}{2}(v^2+2vb\sigma_z+b^2\sigma_z^2)}dv}_{\text{Substitute } u=v+b\sigma_z} \\
&= e^{\frac{1}{2}b^2\sigma_z^2}\int_{-\infty}^{b\sigma_z-\log(\frac{K}{B})/\sigma_z}\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}v^2}dv \\
&= e^{\frac{(\sigma_z^2-\sigma_{xz})^2}{\sigma_z^2}}\mathcal{N}\left(\frac{\log(\frac{B}{K})+\sigma_z^2-\sigma_{xz}}{\sigma_z}\right) \qquad \text{(C.24)}
\end{aligned}$$

and via the same way

$$\mathbb{E}[\mathcal{I}_{Z>\log(\frac{k}{b})}e^{-Z\frac{\sigma_{xz}}{\sigma_z^2}}] = e^{\frac{\sigma_{xz}^2}{2\sigma_z^2}}\mathcal{N}\left(\frac{\log(\frac{B}{K})-\sigma_{xz}}{\sigma_z}\right). \qquad \text{(C.25)}$$

Since

$$\begin{aligned}
Ae^{\frac{1}{2}\sigma_x^2(1-\frac{\sigma_{xz}^2}{\sigma_x^2\sigma_z^2})} &= Ae^{\frac{1}{2}\sigma_x^2} = A\mathbb{E}[e^{-X}] \\
&= \mathbb{E}[e^{-\int_0^t f(s,s)ds}] \\
&= P(0,t) \qquad \text{(C.26)}
\end{aligned}$$

and

$$\begin{aligned}
ABe^{\frac{1}{2}\sigma_x^2(1-\frac{\sigma_{xz}^2}{\sigma_x^2\sigma_z^2})}e^{\frac{(\sigma_z^2-\sigma_{xz})^2}{\sigma_z^2}} &= ABe^{\frac{1}{2}(\sigma_x^2+\sigma_z^2-2\sigma_{xz})} \\
&= AB\mathbb{E}[e^{-X+Z}] \\
&= \mathbb{E}[e^{-\int_0^t f(s,s)ds}G(t,T)] \\
&:= G(0,t,T). \qquad \text{(C.27)}
\end{aligned}$$

We have

$$Be^{\frac{1}{2}\sigma_z^2-\sigma_{xz}} = \frac{\mathbb{E}[e^{-\int_0^t f(s,s)ds}G(t,T)]}{P(0,t)}. \qquad \text{(C.28)}$$

Consequently,

$$\log(\frac{B}{K})+\frac{1}{2}\sigma_z^2-\sigma_{xz} = \log\frac{\mathbb{E}[e^{-\int_0^t f(s,s)ds}G(t,T)]}{P(0,t)K}. \qquad \text{(C.29)}$$

From (C.27) we have

$$\begin{aligned} G(0,t,T) &= Be^{\frac{1}{2}(\sigma_x^2+\sigma_z^2-2\sigma_{xz})} = A\mathbb{E}[e^{-X}]B\mathbb{E}[e^{Z}]e^{-\sigma_{xz}} \\ &= P(0,t)\mathbb{E}[G(t,T)]e^{-\sigma_{xz}} = P(0,t)\underbrace{G(0,T)}_{G(t,T)\text{is a martingale}}e^{-\sigma_{xz}}. \end{aligned} \tag{C.30}$$

Inserting (C.28) and (C.30) into (C.23) results into the closed-from solution for the price of a European call option with maturity t and exercise price K written on the commodity future prices with maturity T, and is given by

$$\boxed{C^G = P(0,t)\left(G(0,T)e^{-\sigma_{xz}}\mathcal{N}\left(\frac{\log(\frac{G(0,T)}{K}) - \sigma_{xz} + \frac{1}{2}\sigma_z^2}{\sigma_z}\right) - K\mathcal{N}\left(\frac{\log(\frac{G(0,T)}{K}) - \sigma_{xz} - \frac{1}{2}\sigma_z^2}{\sigma_z}\right)\right)}. \tag{C.31}$$

Bibliography

[1] F.H.C. Naber, *Fast solver for the three-factor Heston-Hull-White problem*, 2006-2007, Wholesale Banking Amsterdam ING.

[2] A. Krul, *L^∞ Error estimates for the solution of the Black-Scholes and Hull-White models on truncated domains with a payoff boundary condition*, 2007, Wholesale Banking Amsterdam ING.

[3] P. Bjerksund, *Contingent Claims Evaluation when the Convenience Yield is stochastic: Analytical Results*, 1991.

[4] D.R. Ribeiro and S. D. Hodges, *A Two-Factor Model for Commodity Prices and Futures Valuation*, working paper, 2004.

[5] R. Gibson and E. S. Schwartz, *Stochastic Convenience Yield and the Pricing of Oil Contingent Claims*, The Journal of Finance, Volume 45, Issue 3(Jul.1990), 959-976.

[6] G. Welch and G. Bishop, *An Introduction to the Kalman filter*, July 24, 2006.

[7] A. Heemink , *http://ta.twi.tudelft.nl/wagm/users/heemink/ch5dataas.doc*, Delft University of Technology.

[8] D. Lautier, *The Kalman Filter In Finance: An Application to Term Structure Models of Commodity Prices and a Comparison between the Simple and the extended Filters*, 06/12/02, working paper.

[9] D. Lautier, *The Informational Value of Crude Oil Future Prices*, working paper, June 2003.

[10] J-C Duan and J-G Simonato, *Estimating and Testing Exponential-Affine Term Structure Models by Kalman Filter*, octobre 1995, working paper.

[11] K.R. Miltersen and E.S. Schwartz, *Pricing of Options on Commodity Futures with Stochastic Term Structure of Convenience Yields and Interest Rates*, Journal of Fin. and Quan. ANA. VOL. 33, No.1, March 1998.

[12] H. Geman, *Commodities and commodity derivatives, Modeling and Pricing for Agriculturals, Metals and Energy.*, Whiley Finance, 2005.

[13] H. Geman, *The importance of the forward neutral probability measure for stochastic interest rates.*, Working paper. ESSEC, 1989.

[14] S. Shreve, *Stochastic Calculus and Finance*, lecture notes, 1997.

[15] J. Primps, *MS and E 142, lecture notes*, 2007.

[16] J. C. Hull, *Options, Futures, and other derivatives*, Third Edition, 1992.

[17] D. Tavella and C. Randall, *Pricing Financial Instruments, The Finite Difference Method*, Whiley Finance, 2000.

[18] S. Barnett, *Matrices, Methods and Applications*, Clarendon Press, Oxfored, 1992.

[19] A. B. Trolle and E.S. Schwarz, *Unspanned stochastic volatility and the pricing of commodity derivatives*, working paper, July 2007.

[20] F. Caumon, *Redefining the Convenience Yield in the North Sea Crude Oil Market*, Oxford Inst. for Energy Studies, July 2004.

[21] M. Ludkovski, *Spot Convenience Yield Models for Energy Assets*, working paper, August 2003.

[22] D.R. Ribeiro and S.D. Hodges, *A Contango-Constrained Model for Storable Commodities*, working paper, October 13, 2004.

[23] G. Cortazar and E. S. Schwartz, *The Valuation of Commodity-Contingent Claims*, The Journal of Derivatives, summer 1994 pag. 27-39.

[24] H. Reisman, *Movements of the term structure of commodity futures and pricing of commodity claims*, working paper, Faculty of I. E. and management, technion-Israel Institute of Technology, 1992.

[25] T. Bjork and C. Landen, *On the term structure of futures and forward prices*, working paper, december 7, 2000.

[26] F. Black and M. Scholes, *The Pricing of Options and Corporate Liabilities*, The Journal of Political Economy, 81 (3,1973), 637-654.

[27] D. Brigo and F. Mercurio, *Interest Rate Models - Theory and Practice*, Second Edition, Springer.

[28] D. Heath, R. Jarrow and A. Morton, *Bond Pricing and the term structure of interest rates: A new methodology for contingent claims valuation*, Econometrica, Vol. 60, No. 1. (Jan., 1992), pp. 77-105.

[29] D. Heath, R. Jarrow and A. Morton, *Bond Pricing and the term structure of interest rates: A discrete time approximation*, The Journal of Fin. and Quan. Ana., Vol. 25, No.4. (Dec., 1990), pp.419-440.

[30] http://www.hec.unil.ch/matlabcodes/.

[31] L. Clewlow and C. Strickland, *Energy derivatives: Pricing and risk management.* London: Lacima Publications, 2000.

[32] B.Oksendal, *Stochastic Differential Equations, Springer Universitext*, 6^{th} Edition, 2007.

[33] H. Reisman, *Movements of the term structure of commodity futures and pricing of commodity claims*, working paper, 1992.

[34] R. E. Kalman, *A new approach to linear filtering and prediction problems* Journal of Basic Engineering 82 (1): 3545, 1960.