MASTER-ARBEIT

Pricing and Hedging of Weather Derivatives

Institut für Mathematik B
Technische Universität Graz

unter Anleitung von
Univ. Prof. Dr. Albrecher

durch

Markus Zahrnhofer

0357944

im Sommersemester 2009

Acknowledgment

I would like to thank my supervisor, Prof. Dr. Albrecher for giving me the possibility of writing the thesis about this very interesting and complex topic. Further I thank him for his faithful supervision and his patience. Special thanks go to all people, who helped me with the work on the thesis and in particular to my parents for supporting me during my study time.

All errors are my own.

Graz, December 2009 Markus Zahrnhofer

Contents

Chapter 1

Introduction

Weather has an essential influence on many business activities. Regarding for instance energy production, weather has an enormous impact on profits and losses, because the demand for energy (electricity, gas) is highly correlated with the temperature. Therefore, a market for trading financial contracts based on temperature events has emerged in the last decades. Most of these contracts depend on certain temperature conditions, called weather derivatives and more particularly temperature derivatives. The contracts are a new type of securities and differ from insurance contracts. If an insurance owner claims a loss, he has to prove that a loss has occurred on his insured title. Another characteristic of traditional insurance contracts is that they are not geared to cover monetary losses as a consequence of temperature variation, but rather losses as a consequence of extreme weather conditions e.g. floods and drought.

On the contrary, weather derivatives are a valuable tool for managing risk, because they tend to reduce the risk caused by temperature variations. A typical example, already mentioned above, is an energy producer, for whom warm weather during the winter or low temperature during the summer may incur significant losses in earnings. The enterprise may buy a temperature derivative to compensate the losses if the temperature is too high or too low during the according season. Clearly, the enterprise has to pay for this contract, namely the premium charged by the counterparty of the contract. If the summer is too cold or the winter too warm, a contract could cover all, or a part of the incurred losses caused of temperature variation. If the seasons are not atypical, the enterprise will only lose the premium paid for the contract. There are many different structures of weather derivatives and they often depend on the needs of the investor.

Standardized weather derivatives are traded at the Chicago Mercantile Exchange (CME)[1]. The market for weather derivatives exists since 1999 and there are two different classes of standardized contracts: temperature futures and options on temperature futures. All of them are linked to different temperature indices (see Section 3.1.2). The indices are based on measurement locations in the USA, Canada, Europe and Asia.

Especially in Europe, where the weather market is growing, there already exists another en-

[1] http://www.cmegroup.com

terprise which offers weather derivative contracts, namely Celsius Pro[2]. In addition to the regulated standardizied market, there also exists an unregulated market, the so called over-the-counter (OTC) market. Nevertheless, in Austria temperature derivative contracts are rare. Temperature linked derivatives hardly exist in Austria, and also alternative forms to insurance contracts against extreme weather conditions like CAT-bonds, are not popular.

Therefore, the objective of this thesis is the consideration of different methods for modeling the temperature and pricing methods of weather derivatives. We start with considering different stochastic temperature models and apply them to temperature data from Graz Airport (Thalerhof). Subsequently we pay attention to the derivation of derivatives linked to different temperature indices, which are traded at CME. In Chapter 4 we consider a spatial problem of temperature derivatives. This method is helpful for firms whose losses are influenced by temperature changes in entire areas rather than just in cities. This derivative method could be helpful for energy producers and farmers to hedge their monetary losses caused by temperature influences.

In Chapter 5 and 6 we present other methods for pricing derivatives with non-tradeable underlyings. There we first concentrate on optimal cross hedging, where we use the negative correlation of risk exposure of different agents on the market. An improved approach is investigated in Chapter 6, where the price and the optimal hedging strategy is expressed in terms of forward-backward stochastic differential equations. In both parts we assume an exponential utility framework for an indifference pricing approach.

Finally hybrid CAT-bonds are discussed, which are modifications of simple CAT-bonds and linked to the stock market.

[2]http://www.celsiuspro.com

Chapter 2

Stochastic Modeling of Temperature

In this chapter we analyse stochastic processes for modeling the dynamics of daily mean temperature. We consider a mean-reverting Ornstein-Uhlenbeck process with different volatility functions. In addition to the analysis of the models we apply them to temperature data form Graz Airport (Thalerhof). We simulate the different models and compare the simulated paths with the observed temperature. The chapter ends with some possible improvements of the models. We mainly follow Alaton et al. [3], Bhowan [17] and Benth and Šaltytė-Benth [13].

2.1 Weather data

Let us first consider the temperature data, which we got from ZAMG (Zentralanstalt für Meteorologie und Geodynamik) for the weather station located at the airport of Graz (Thalerhof). The data set consists of daily mean temperature[1] in the period January 1, 1961 until February 16, 2009, resulting in 17579 observations (the data set is complete). The data set includes leap days entries, which we will not consider in the later explanations.

To construct the different models we only use the data until December 31, 2007, because from January 1, 2008 until February 16, 2009 can be used to the compare observed with the modeled temperature. If we take a look at the histogram of the daily average temperature (Figure 2.1), it indicates, that the data are not normally distributed. The Shapiro Wilk test also rejects normality. The figure shows a negative skewness and a negative kurtosis, which is confirmed in Table 2.1. This result is caused by the cold and warm seasons in Austria.
On the other hand it will be reasonable to assume a normal distribution, for the daily temperature differences (cf. Figure 2.2). Though, small differences in the daily mean temperature will be underestimated. This indicates that a Brownian Motion model for the time evolution of daily temperature differences could be useful. (cf. Section 2.5.4).

[1] usually defined to be the average of the maximum and minimum temperature over a 24h-time horizon for the specific date

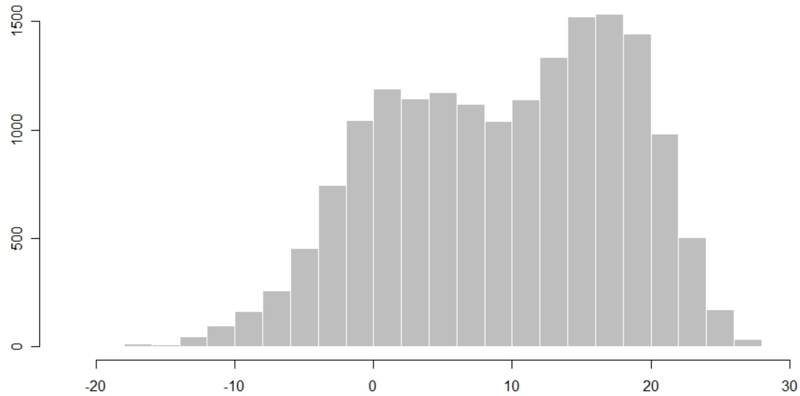

Figure 2.1: Histogram of daily mean temperature from Graz in the period from January 1, 1961 until December 31, 2007

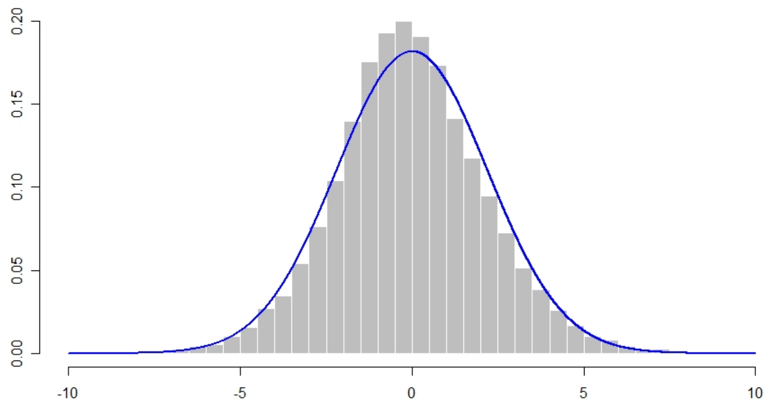

Figure 2.2: Histogram of the daily temperature differences from Graz in the period from January 1, 1961 until December 31, 2007

Minimum	Maximum	Mean	Skewness	Kurtosis
-20.3	28.5	9.4	-0.26	-0.81

Table 2.1: Summary

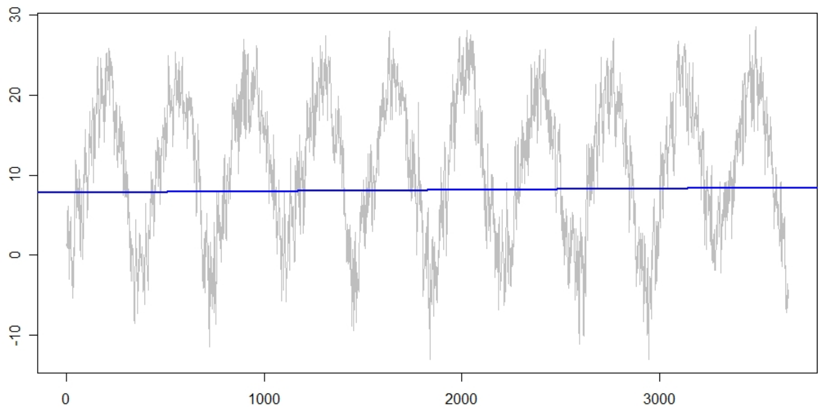

Figure 2.3: Observed daily mean temperature from Graz in the period from January 1, 1998 until December 31, 2007, together with the regression line.

The temperature varies between -7 degrees in the winter and about 23 degrees during the summer, as we can see in Figure 2.3. It is obvious that the temperature process should be a mean-reverting process with seasonality. To shape the seasonal dependence it is possible to use a trigonometric function (e.g. sine or cosine). A simple regression model (Table 2.2) shows that a weak positive, significant, linear trend exists. (in the considered period the daily mean temperature increased by about 3.18 degrees)

Intercept	Slope
7.81	0.000185

Table 2.2: Values of the linear regression model

There are several plausible reasons for the increase, for example global warming or the urban heating[2] or especially at the observation site, the increase of flights at the airport (see Kabas [39]).

2.2 General assumptions for the models

Now we introduce some general assumptions, which are valid for all considered models. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space with a filtration $\{\mathcal{F}_t\}_{t\geq 0}$.

[2]temperature rises in areas nearby big cities

Definition 2.1 (Ornstein-Uhlenbeck-Process) *The unique solution of the equation*

$$dX(t) = -\alpha X(t)dt + \sigma dB(t) \tag{2.1}$$

with $\alpha \in \mathbb{R}$, $\sigma > 0$ *and* $B(t)$ *a Brownian motion, is called Ornstein-Uhlenbeck process.*

Theorem 2.1 *The solution of (2.1) is given by*

$$X(t) = X(0)e^{-\alpha t} + \sigma \int_0^t e^{-\alpha(t-u)} dB(u). \tag{2.2}$$

Proof:
If $\sigma = 0$ then the solution is

$$X(t) = X(0)e^{-\alpha t}.$$

Let us consider

$$Y(t) = X(t)e^{\alpha t}.$$

It follows:

$$\begin{aligned} dY(t) = d(X(t)e^{\alpha t}) &= X(t)d(e^{\alpha t}) + (dX(t))e^{\alpha t} + (dX(t))d(e^{\alpha t}) \\ &= X(t)\alpha e^{\alpha t}dt + e^{\alpha t}\left(-\alpha X(t)dt + \sigma dB(t)\right) \\ &= \sigma e^{\alpha t} dB(t) \end{aligned}$$

By integrating we obtain:

$$Y(t) - Y(0) = \sigma \int_0^t e^{\alpha u} dB(u)$$

$$Y(t) = X(0) + \sigma \int_0^t e^{\alpha u} dB(u)$$

using $X(t) = Y(t)e^{-\alpha t}$ and the result follows. □

In the following we want to model the temperature as a stochastic mean-reverting process. Hence we consider the Vasicek mean reverting model, which is a widely used interest rate model. The process is defined as the following stochastic differential equation (SDE):

$$dT(t) = \theta(M(t) - T(t))dt + \sigma(t)dB(t), \tag{2.3}$$

where $T(t)$ is the daily mean temperature, θ the speed of reversion (constant), $\sigma(t)$ the volatility of the process, $B(t)$ a Brownian Motion and $M(t)$ the mean to which the process reverts. The mean reversion model has been suggested by Dornier and Querel in [29]. They showed that (2.3) only reverts to $M(t)$ (we require $E(T(t)) = M(t)$), if the parameter $M(t)$ is constant (see Bhowan [17]).

However, $M = M(t)$ should be a deterministic function for our purposes which models the trend and seasonality of the temperature. To obtain a stochastic process that reverts to $M(t)$ we have to add the term

$$\frac{dM(t)}{dt}.$$

Then we obtain a model for the evolution of temperature as a generalized Ornstein-Uhlenbeck process of the form

$$dT(t) = \left[(\theta(M(t) - T(t)) + \frac{dM(t)}{dt}\right] dt + \sigma(t)dB(t). \tag{2.4}$$

Theorem 2.2 *The solution of (2.4) is*

$$T(t) = (T(0) - M(0))e^{-\theta t} + M(t) + \int_0^t e^{-\theta(t-u)}\sigma(u)dB(u). \tag{2.5}$$

Proof:
Applying partial integration to $dT(t)e^{\theta t}$ we obtain

$$\begin{aligned} dT(t)e^{\theta t} &= T(t)de^{\theta t} + e^{\theta t}dT(t) \\ &= \theta T(t)de^{\theta t} + e^{\theta t}dT(t). \end{aligned}$$

By inserting (2.4) for $dT(t)$ it follows that

$$\begin{aligned} dT(t)e^{\theta t} &= \theta T(t)e^{\theta t}dt + e^{\theta t}dM(t) - \theta e^{\theta t}T(t)dt + \theta e^{\theta t}M(t)dt + e^{\theta t}\sigma(t)dB(t) \\ &= \theta e^{\theta t}M(t)dt + e^{\theta t}dM(t) + e^{\theta t}\sigma(t)dB(t). \end{aligned}$$

We obtain by partial integrating

$$\begin{aligned} T(t)e^{\theta t} &= T(0) + \int_0^t \theta e^{\theta u}M(u)du + \int_0^t e^{\theta u}dM(u) + \int_0^t e^{\theta u}\sigma(u)dB(u) \\ &= T(0) + e^{\theta t}M(t) - M(0) + \int_0^t e^{\theta u}\sigma(u)dB(u). \end{aligned}$$

□

2.3 Model A

After some general assumptions we start with the first stochastic temperature model. It is similar to the model of Alaton et al. [3] and our analysis will be based on the same assumptions. We have to estimate the parameters of equation (2.4). As already mentioned above, $M(t)$ should be a deterministic function which models the trend and the seasonality. If we look at Figure 2.3, the behavior of the temperature suggests (according to Alaton et al. [3]), to choose $M(t)$ as

$$M(t) = a + bt + c\sin(\omega t + \phi) \tag{2.6}$$

with $\omega = \frac{2\pi}{365}$. This gives a good fit of periodic temperature data. To estimate the numerical values of equation (2.6) we use ordinary least squares (OLS).

2.3.1 Estimation of the mean temperature function

We can write equation (2.6) as follows

$$M(t) = a + bt + c(\sin(\omega t)\cos(\phi)) + c(\sin(\phi)\cos(\omega t)). \tag{2.7}$$

It would be possible to estimate (2.7) using a nonlinear regression, but we transform (2.7) to a linear function.

$$M(t) = a + bt + \alpha_1 t_1 + \alpha_2 t_2 \tag{2.8}$$

with $\alpha_1 = c\cos(\phi)$, $\alpha_2 = c\sin(\phi)$, $t_1 = \sin(\omega t)$ and $t_2 = \cos(\omega t)$. To estimate the parameters of (2.8) we then apply OLS. We get the following numerical values:

$$a = 7.86, \tag{2.9}$$

$$b = 0.00017, \tag{2.10}$$

$$c = \frac{\alpha_1}{\cos(\phi)} = \frac{-2.70}{0.24} = -10.89, \tag{2.11}$$

$$\phi = \tan^{-1}\left(\frac{\alpha_2}{\alpha_1}\right) = \tan^{-1}\left(\frac{-10.55}{-2.70}\right) = 1.32. \tag{2.12}$$

We obtain the following function for the mean temperature:

$$M(t) = 7.86 + 0.00017t - 10.89\sin\left(\frac{2\pi t}{365} + 1.32\right). \tag{2.13}$$

In this formula we see the weak significant trend already mentioned above. Figure 2.4 represents the observed mean temperature with $M(t)$.

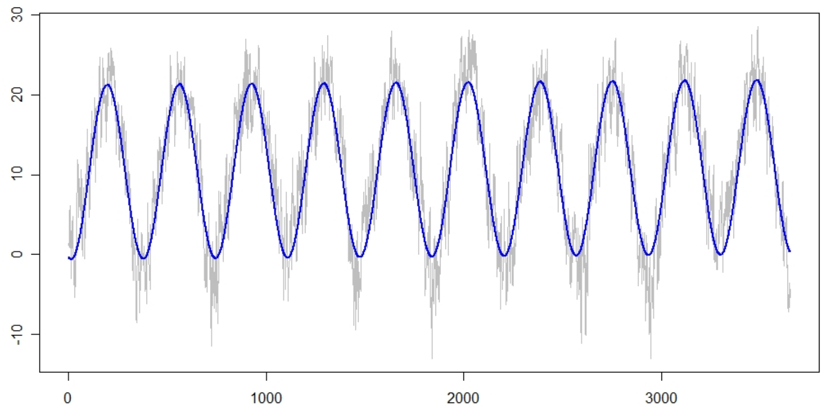

Figure 2.4: Observed daily mean temperature from Graz in the period from January 1, 1998 until December 31, 2007, together with fitted mean temperature function $M(t)$.

2.3.2 Estimation of the volatility σ

This section utilizes ideas of Alaton et al. [3]. To estimate σ from (2.4), we assume that the quadratic variation σ^2 of the temperature is nearly constant during one month, while the quadratic variation varies across the different months of the year. We assume that $\sigma(t)$ is a piecewise constant function during each month. We get 12 different values for $\sigma(t)$, $\sigma(1)$ during January, $\sigma(2)$ during February and so on.
For a specific month η ($\eta = 1, \ldots, 12$), N_η denotes the observed temperatures $T(j)$ $j = 1, \ldots, N_\eta$ during one month η. (i.e. $\eta = 1$: $N_\eta = 31 \cdot 47$ "days of January" $\cdot$ "count of years"). At first we derive one estimator for $\sigma(\eta)$ and later a second one and then we will take the average. The first estimator is based one the quadratic variation of $T(t)$ (see Basawa and Prasaka Rao [12]):

$$\hat{\sigma}^2(\eta) = \frac{1}{N_\eta} \sum_{j=1}^{N_\eta - 1} (T(j+1) - T(j))^2. \tag{2.14}$$

By discretizing (2.4) we can derive a second estimator of $\sigma(\eta)$. The discretized equation has the following form during a given month η :

$$T(j) = M(j) - M(j-1)\theta M(j-1) + (1-\theta)T(j-1) + \sigma(\eta)\epsilon(j-1) \quad j = 1, \ldots, N_\eta \tag{2.15}$$

with $\{\epsilon(j)\}_{j=1}^{N_\eta}$ independent standard normally distributed random variables. We can write (2.15) as follows

$$\hat{T}(j) = \theta M(j-1) + (1-\theta)T(j-1) + \sigma(\eta)\epsilon(j-1), \tag{2.16}$$

with $\hat{T}(j) := T(j) - (M(j) - M(j-1))$. According to Brockwell [20], an efficient estimator is

$$\hat{\sigma}(\eta)^2 = \frac{1}{N_\eta - 2} \sum_{j=0}^{N_\eta} (\hat{T}(j) - \hat{\theta} M(j-1) - (1-\hat{\theta})T(j-1))^2. \tag{2.17}$$

To derive the second estimator of $\sigma(\eta)$, we need the estimator of θ, which is the objective of the following section.

2.3.3 Estimation of the mean-reverting parameter θ

According to Bibby und Sørensen [18], an unbiased estimator of θ is the root of the equation:

$$G_n(\theta) = \sum_{i=1}^{n} \frac{\dot{b}(T(i-1);\theta)}{\sigma^2(i-1)} \left(T(i) - E\left[T(i)|T(i-1)\right]\right) \tag{2.18}$$

where n is the number of observations and $\dot{b}(T(i-1);\theta)$ denotes $\frac{\partial b}{\partial \theta}$. To solve (2.18) we have to determine $E\left[T(i)|T(i-1)\right]$. Equation(2.5), for $s \leq t$;

$$T(t) = (T(s) - M(s))e^{\theta t} + M(t) + \int_s^t e^{-\theta(t-u)}\sigma(u)dB(u), \tag{2.19}$$

yields

$$E\left[T(i)|T(i-1)\right] = (T(i-1) - M(i-1))e^{\theta} + M(i). \tag{2.20}$$

By substituting in (2.18), we get

$$G_n(\theta) = \sum_{i=1}^{n} \frac{M(i-1) - T(i-1)}{\sigma^2(i-1)} \left[T(i) - (T(i-1) - M(i-1))e^{-\theta} - M(i)\right]. \tag{2.21}$$

The unique solution of (2.21) is

$$\hat{\theta} = -\log\left(\frac{\sum_{i=1}^{n} \frac{M(i-1)-T(i-1)}{\sigma^2(i-1)}[T(i) - M(i)]}{\sum_{i=1}^{n} \frac{M(i-1)-T(i-1)}{\sigma^2(i-1)}[T(i-1) - M(i-1)]}\right) \tag{2.22}$$

for $i = 1, \ldots, n$. Here $\sigma^2(i-1)$ are the associated $\hat{\sigma}^2(\eta)$ estimated in (2.14).
Now we are able to derive the last unknown parameters of (2.4). The numerical value of the estimator of θ is

$$\hat{\theta} = 0.26. \tag{2.23}$$

The estimators of σ are listed in Table 2.3, and we see that the volatility of the temperature during the summer months is lower than in the winter. It is also obvious that estimator 1 and 2 are of the same magnitude for almost all months.

Month	1st Estimation	2nd Estimation	Average
January	2.74	2.74	2.74
February	2.58	2.58	2.58
March	2.70	2.70	2.70
April	2.37	2.38	2.37
May	2.31	2.32	2.32
June	2.16	2.17	2.17
July	2.04	2.05	2.05
August	1.96	1.96	1.96
September	2.12	2.12	2.12
October	2.46	2.46	2.46
November	2.71	2.71	2.71
December	2.74	2.74	2.74

Table 2.3: Estimators of σ

2.4 Model B

A modification of Model A and was developed by Bhowan [17]. The volatility is estimated differently. In Model A we argued that the volatility of the temperature is nearly constant during a month, but varies across the year. This means that σ is a piecewise constant function,

changing monthly. Bowhan [17] proposes to apply a stochastic process for the volatility. The volatility changes randomly on a monthly basis, but is still constant during one month.
To estimate the mean temperature function, we use the same idea as in Section 2.3.1. The parameters are then again:

$$a = 7.86 \quad b = 0.00017 \quad c = -10.89 \quad \phi = 1.32.$$

2.4.1 Estimation of the volatility process σ

Before we consider a stochastic mean reverting process for the volatility, let us consider Figure 2.5 representing the observed monthly volatility. To calculate the monthly volatility, take equation (2.14) and replace N_η by the number of days of the according month (i.e. 31 for January 1961 and so on), which yields 564 different volatilities, for each month the according value. The volatility should revert to a long term trend, that is why the stochastic differential equation has the form,

$$d\sigma(\tau) = -\theta_\sigma(\sigma(\tau) - M_\sigma)d\tau + \sigma_\sigma dB_\sigma(\tau). \tag{2.24}$$

M_σ is constant and the estimated value is

$$\hat{M}_\sigma = 2.14. \tag{2.25}$$

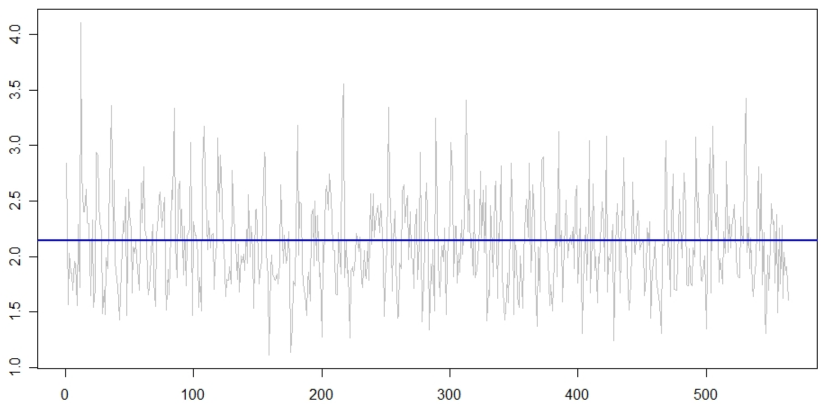

Figure 2.5: Observed monthly volatility of temperature from Graz in the period from January 1, 1998 until December 31, 2007, together with $\hat{M}_\sigma$.

Two parameters remain to be estimated, namely σ_σ and θ_σ. In the case of σ_σ we use the estimator of the quadratic variation

$$\hat{\sigma}_\sigma^2 = \frac{1}{n}\sum_{j=0}^{n-1}(\sigma(j+1) - \sigma(j))^2 \tag{2.26}$$

with $\sigma(j)$ the quadratic variation of the temperature during the month j. We obtain

$$\hat{\sigma}_\sigma = 0.51. \tag{2.27}$$

θ_σ is estimated by a modification of (2.22)

$$\hat{\theta}_\sigma = -\log\left(\frac{\sum_{i=1}^{n}\frac{M_\sigma-\sigma(i-1)}{\sigma^2(i-1)}[\sigma(i)-M_\sigma]}{\sum_{i=1}^{n}\frac{M_\sigma-\sigma(i-1)}{\sigma^2(i-1)}[\sigma(i-1)-M_\sigma]}\right), \tag{2.28}$$

then

$$\hat{\theta}_\sigma = 1.39 \tag{2.29}$$

2.4.2 Estimation of the mean-reverting parameter θ

In this model the estimation of the mean-reverting parameter is similar to Section 2.3.3. Only the $\sigma^2(i-1)$ must be modified. Performing the modified calculation in (2.22) gives

$$\hat{\theta} = 0.21. \tag{2.30}$$

2.5 Model C

The analysis of this last model is based on the model of Benth and Šaltytė-Benth [13]. He proposes to model the volatility as a truncated Fourier series. This choice leads to a seasonal volatility. The procedure of estimating the parameters of equation (2.4) completely differs from Model A and B. At first we reformulate the continuous time dynamics of the solution of the Ornstein-Uhlenbeck process (2.5) to a time series. Recall (2.5), which is equivalent to

$$\begin{aligned}T(t) &= M(t) + (T(0)-M(0))\,e^{-\theta t} + e^{-\theta}\int_0^t \sigma(u)e^{-\theta(t-u)}dB(u)\\ &\quad + (1-e^{-\theta})\int_0^t \sigma(u)e^{-\theta(t-u)}dB(u).\end{aligned}$$

From the solution (2.5) we obtain

$$T(t+1) = M(t+1) + (T(0)-M(0))\,e^{-\theta(t+1)} + e^{-\theta}\int_0^{t+1}\sigma(u)e^{-\theta(t-u)}dB(u)$$

and

$$\begin{aligned}(1-e^{-\theta})\int_0^t \sigma(u)e^{-\theta(t-u)}dB(u) &= (1-e^{-\theta})\,(T(t)-M(t))\\ &\quad - \left(e^{-\theta t}-e^{-\theta(t+1)}\right)(T(0)-M(0)).\end{aligned}$$

We denote $\Delta T(t) = T(t+1) - T(t)$ and get

$$\begin{aligned}
\Delta T(t) &= M(t+1) - M(t) + \left(e^{-\theta(t+1)} - e^{-\theta t}\right)(T(0) - M(0)) \\
&\quad + e^{-\theta} \int_t^{t+1} \sigma(u) e^{-\theta(t-u)} dB(u) - \left(1 - e^{-\theta}\right) \int_0^t \sigma(u) e^{-\theta(t-u)} dB(u) \\
&= \Delta M(t) + \left(e^{-\theta(t+1)} - e^{-\theta t}\right)(T(0) - M(0)) + e^{-\theta} \int_t^{t+1} \sigma(u) e^{-\theta(t-u)} dB(u) \\
&\quad - \left(1 - e^{-\theta}\right)(T(t) - M(t)) + \left(e^{-\theta t} - e^{\theta(t+1)}\right)(T(0) - M(0)) \\
&= \Delta M(t) - (1 - e^{-\theta})(T(t) - M(t)) + e^{-\theta} \int_t^{t+1} \sigma(u) e^{-\theta(t-u)} dB(u).
\end{aligned}$$

Next we approximate the stochastic integral, which yields

$$\Delta T(t) \approx \Delta M(t) - (1 - e^{-\theta})(T(t) - M(t)) + e^{-\theta} \sigma(t) \Delta B(t). \tag{2.31}$$

This equation can be written as a time series model (AR(1)-Model) of the following form:

$$\tilde{T}(t+1) = \rho \tilde{T}(t) + \tilde{\sigma}(t) \epsilon(t) \tag{2.32}$$

with $\tilde{T}(t) = T(t) - M(t)$, $\tilde{\sigma}(t) = \rho\sigma(t)$, $\rho = e^{-\theta}$ and $\epsilon(t)$ i.i.d standard normally distributed. To estimate (2.32) we have to follow several steps. At first we have to remove the seasonality and the linear trend.

2.5.1 Estimation of the mean temperature function

We already mentioned that $M(t)$ should be a deterministic function modeling the trend and the seasonality. In this model we specify $M(t)$ to be of the form

$$M(t) = a + bt + b_1 + b_2 \cos\left(\frac{2\pi(t - b_3)}{365}\right). \tag{2.33}$$

At the beginning we already showed that there exists a weak linear trend (see Figure 2.3 and Table 2.2). After removing the trend we can determine the seasonal part of $M(t)$. This yields the following parameters of $M(t)$:

$$b_1 \approx 0, \quad b_2 = -10.89, \quad b_3 = 14.55.$$

2.5.2 Estimation of the mean-reverting parameter θ

Now we use the de-trended and de-seasonalized temperature series to estimate the coefficient of (2.32). In other words, we regress the daily mean temperature against the one of the previous day. The value of the mean-reverting parameter, which is significant, is

$$\rho = 0.80$$

and corresponds to

$$\hat{\theta} = -\ln(\rho) = 0.22.$$

Before we estimate the volatilty σ, we take a look at the autocorrelation function (Figure 2.6) of the residuals of the AR(1) model of the de-trended and de-seasonalized mean temperature. For the sake of completeness, we note that the standard deviation of the residuals is $\bar{\sigma} = 2.08$. We observe high values of the autocorrelation for the first lags. For the higher lags it seems that the values are varying randomly around zero. The autocorrelation function of the sqared residuals (Figure 2.7) indicates the fact of time dependency. Moreover a clear seasonal variation exists.

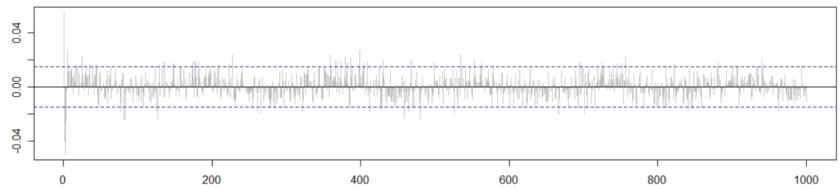

Figure 2.6: Autocorrelation function of the residuals of the mean temperature at the airport Graz

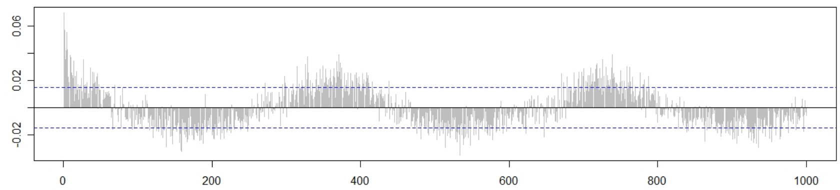

Figure 2.7: Autocorrelation function of the square residuals of the mean temperature at the airport Graz

2.5.3 Estimation of the volatility process σ

The estimation of the seasonal volatility of the residuals will be done in several steps. First, we define the form of $\sigma^2(t)$. We already mentioned above that the volatility will be modeled with a truncated Fourier series

$$\sigma^2(t) = c + \sum_{i=1}^{I} c_i \sin(2i\pi t/365) + \sum_{j=1}^{J} d_j \cos(2j\pi t/365). \tag{2.34}$$

At first we group the residuals of the AR(1)-model in 365 groups, that means we get 47 observations for a particular date (i.e. May 1). Taking the average of the square of each group we obtain the volatility. Next we choose (according to [13]) $I = J = 4$ in equation (2.34). Recall

$$\sigma^2(t) = \tilde{\sigma}^2(t)/\rho^2. \tag{2.35}$$

In Figure 2.8 we see the empirical volatility with the fitted function. The highest volatility occurs during the winter period, while fall and summer season have lower volatilities. This confirms our result of the simpler model form Section 2.3. In Tabel 2.4 we see the fitted parameters of $\sigma^2(t)$.

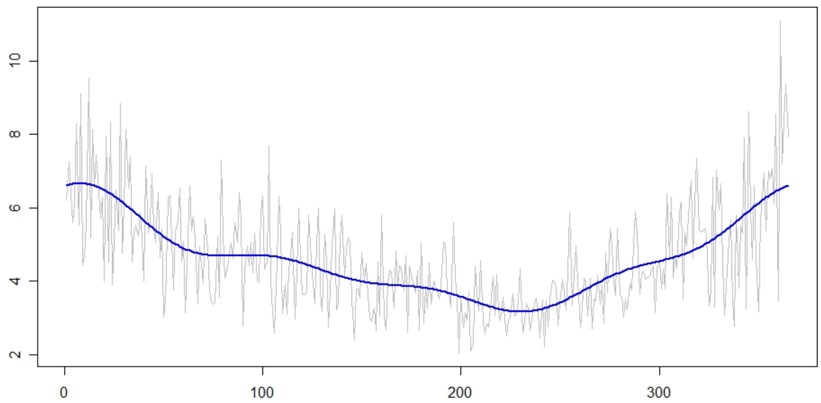

Figure 2.8: Empirical volatility with the fitted function $\tilde{\sigma}(t)$

c	c_1	d_1	c_2	d_2	c_3	c_3	c_4	d_4
6.74	0.74	1.95	-0.18	0.77	0.15	0.18	0.14	0.26

Table 2.4: Estimators of $\sigma^2(t)$

2.5.4 Non-normality

After removing the temporal phenomena in the volatility we obtain the autocorrelation functions, which are presented in Figure 2.9 and Figure 2.10. We can see that the seasonality in the autocorrelation function of the square residuals (Figure 2.10) has been removed. The autocorrelation in the first lags is still existing. This should be an indicator to use more refined models, but this would lead to a significant complication of later calculations of futures and options prices.

Although the histogram of the residuals (Figure 2.11) seems to be standard normal distributed, there still exists another problem. The Shapiro Wilk test rejects normal distribution

and the residuals are left skewed.
In Section 2.7 we suggest a Lévy-based Ornstein-Uhlenbeck process with marginals following the class of generalized hyperbolic distributions. Such models are able to capture small peaks in the center, which we see in the histogram. But Benth and Šaltytė-Benth [13] mention that such processes may be hard to use for pricing derivatives and they think that the assumption of $i.i.d$ standard normal residuals is remarkable.

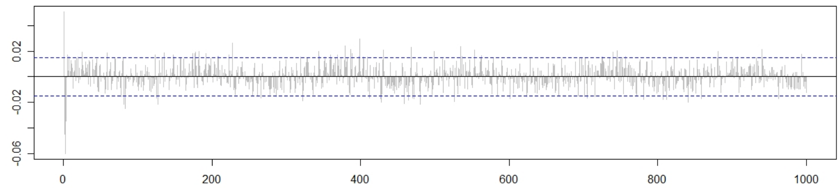

Figure 2.9: Autocorrelation function of the residuals of the mean temperature at the airport Graz, after dividing out the volatility function $\tilde{\sigma}$(t)

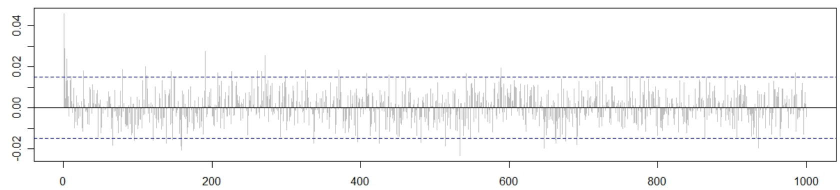

Figure 2.10: Autocorrelation function of the square residuals of the mean temperature at the airport Graz, after dividing out the volatility function $\tilde{\sigma}$(t)

2.6 Simulation

In this section we want to simulate the different models and compare the simulated values with the observed temperature in the period January 1, 2008 until Febrary 16, 2009, as already mentioned. For simulating the paths we have to discretise (2.4) and and (2.24), especially for Model B. Using the Euler Scheme of approximation (see Kloeden et al. [41]) we obtain the following equations

$$T(t+1) = T(t) + \theta(M(t) - T(t)) + \frac{dM(t)}{dt} + \sigma(\tau)Y_1 \tag{2.36}$$

$$\sigma(\tau) = \sigma(\tau - 1) + \theta_\sigma(M_\sigma - \sigma(\tau - 1)) + \sigma_\sigma Y_2 \tag{2.37}$$

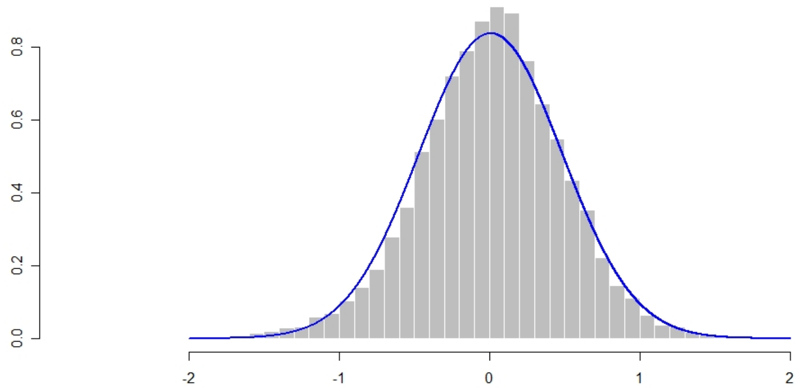

Figure 2.11: Histogram of the residuals of the temperature from Graz with the standard normal density

with Y_1, Y_2 $i.i.d$ standard normal distributed random variables. $\sigma(\tau)$ in (2.36) for Model A are given in Table 2.3 for the according month. In Model B the volatility will be simulated using (2.37) for the according month. For the latter model the volatility is given by the truncated Fourier Series (2.34). The other parameters and estimators are given in the previous sections.

The simulation will be done in several steps. A simulated path will be the average of 5 separately simulated paths (here the choice of 5 stems from empirical studies). This assumption makes sense, because if the number is too high the simulated temperature is too smooth and if the number is too low the variation of the paths is very high. To enable the comparison of the models, the generated normally distributed random variables will be the same for the according run of each model (i.e. the random variables of the first run of Model A, B and C are the same).

Carrying out various simulation runs shows that the solutions of the different models are nearly the same. Furthermore, the errors between the observed and fitted values of the different models are of similar size. Models of different complexity give us nearly the same result. It's impossible to say which model is the most suitable for simulating the observed temperature in the period from January 1, 2008 to February 16, 2009.

Figure 2.12 shows a simulated path of Model C and the observed values in the period from January 1, 2008 to February 16, 2009.

2.7 Improvement of the Models

It is clear that these models are not ideal, but more complex models may be hard to use for pricing derivatives. Nevertheless we want to mention two possible improvements of the models. On the on hand it could be helpful to consider larger models where the temperature

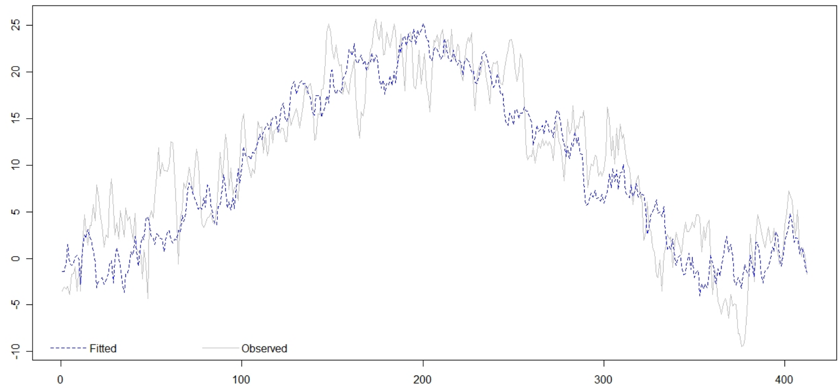

Figure 2.12: Simulated Path (Model C) using the Euler Scheme, together with the observed mean temperature in the period from January 1, 2008 to February 16, 2009

is only one of many variables. On the other hand we could upgrade Model C and use a Lévy-based Ornstein-Uhlenbeck process.

2.7.1 Lévy-based Ornstein-Uhlenbeck process

In this section we mainly follow Benth and Šaltytė-Benth [14]. A generalization of the Ornstein-Uhlenbeck process (2.4) for modeling the evolution of temperature is given by:

$$dT(t) = dM(t) + \theta(T(t) - M(t))dt + \sigma(t)dL(t). \tag{2.38}$$

Compared to (2.4), the Brownian motion is replaced by a Lévy process. A popular choice is to use a Lévy process with marginals following the class of generalized hyperbolic distributions, because this is a very flexible family of distributions. This choice makes it possible to model skewness and (semi-)heavy tails. Another advantage of this class of distributions is that the density, characteristic function and the moment generating function are known explicitly. The generalized hyperbolic distribution is a family of infinitely divisible distributions with density function

$$\begin{aligned} f_{gh}(x; \lambda, \mu, \alpha, \beta, \delta) = {} & c(\delta^2 + (x-\mu)^2)^{\frac{\lambda - \frac{1}{2}}{2}} e^{\beta(x-\mu)} \\ & \times K_{\lambda - \frac{1}{2}}\left(\alpha\sqrt{\delta^2 + (x-\mu)^2}\right) \end{aligned}$$

Here K_s denotes the modified Bessel function of the third kind with index s (see [2] Section 9.6) and the normalizing constant c is given as

$$c = \frac{(\alpha^2 - \beta^2)^{\frac{\lambda}{2}}}{\sqrt{2\pi}\alpha^{\lambda - \frac{1}{2}}\delta^\lambda K_\lambda\left(\delta\sqrt{\alpha^2 - \beta^2}\right)}.$$

The parameter μ controls the location of the distribution, α the fatness of the tails, β the skewness and δ is the scaling. Obviously the distribution is symmetric when $\beta = 0$. The parameter λ identifies the subfamily within the generalized hyperbolic class. In the financial context two special cases have been studied extensively. On the one hand the hyperbolic distribution with $\lambda = 1$, and on the other hand the normal inverse Gaussian distribution with $\lambda = 0.5$. The limiting cases of the generalized hyperbolic family are the normal, Student t and the Cauchy distribution. The explicit form of the moment generating function of generalized hyperbolic distributions is

$$\mathbb{E}\left[e^{uX}\right] = e^{\mu u}\left(\frac{\alpha^2-\beta^2}{\alpha^2-(\beta+\mu)^2}\right)^{\frac{\lambda}{2}} \frac{K_\lambda\left(\delta\sqrt{\alpha^2-(\beta+\mu)^2}\right)}{K_\lambda\left(\delta\sqrt{\alpha^2-\beta^2}\right)}, \tag{2.39}$$

whenever $|\beta+\mu| < \alpha$. Hence the family of distributions has finite moments of all orders.
A Lévy process $L(t)$ is called generalized hyperbolic Lévy process if the marginals $L(1)$ are distributed according to the generalized hyperbolic family. In this case the Lévy measure is given explicitly by

$$\ell_{GH}(dz) = |z|^{-1}e^{\beta z}\left(\frac{1}{\pi^2}\int_0^\infty \frac{\exp(-\sqrt{2y+\alpha^2}|z|)}{J_\lambda^2(\delta\sqrt{2y})+Y_\lambda^2(\delta\sqrt{2y})y}\frac{dy}{y} + \lambda e^{-\alpha|z|}\right)dz, \quad \lambda \geq 0 \tag{2.40}$$

and

$$\ell_{GH}(dz) = |z|^{-1}e^{\beta z}\frac{1}{\pi^2}\int_0^\infty \frac{\exp(-\sqrt{2y+\alpha^2}|z|)}{J_{-\lambda}^2(\delta\sqrt{2y})+Y_{-\lambda}^2(\delta\sqrt{2y})y}\frac{dy}{y}dz, \quad ,\lambda < 0. \tag{2.41}$$

Here J_λ and Y_λ are the Bessel functions of the first and second kind with index λ (see Abramowitz and Stegun [2] Section 9.1). The generalized hyperbolic Lévy processes $L(t)$ are pure-jump processes with paths of infinite variation.
Applying the Itô-Formula for semimartingales, we obtain the following solution of (2.38):

$$T(t) = M(t) + (T(0)-M(0))e^{\theta t} + \int_0^t \sigma(u)e^{\theta(t-u)}dL(u) \tag{2.42}$$

Now we can give an explicit formula for the cumulative temperature over a time interval $[\tau_1, \tau_2]$.

Theorem 2.3 *If the temperature $T(t)$ follows (2.38), the cumulative temperature over the time interval $[\tau_1, \tau_2]$ is explicitly given by*

$$\begin{aligned}\int_{\tau_1}^{\tau_2} T(t)dt = &\int_{\tau_1}^{\tau_2} M(t)dt + \theta^{-1}(T(0)-M(0))(e^{\theta\tau_2}-e^{\theta\tau_1}) \\ &+ \int_0^{\tau_2}\sigma(t)\theta^{-1}\left(e^{\theta(\tau_2-t)} - 1_{[0,\tau_1]}(t)e^{\theta(\tau_1-t)} - 1_{[\tau_1,\tau_2]}(t)\right)dL(t)\end{aligned}$$

Proof: See Benth and Šaltytė-Benth [14]. □

For more details on generalized hyperbolic distributions we refer to Barndorff-Nielsen and Shephard [8] and for applications of the Itô-Formula for semimartingales to Ikeda and Watanabe [37].

Chapter 3

Pricing of Temperature Derivatives

In this chapter we present a pricing method for temperature derivatives which is based on Model C of Chapter 2. After giving an overview of different pricing methods of weather derivatives, we analyse derivatives based on different temperature indices (see Section 3.1.2). We analyse temperature futures, which are also traded at the CME, and we discuss the pricing of options on these derivatives. We mainly follow in this chapter Benth and Šaltytė-Benth [13] and Benth and Šaltytė-Benth [15].

3.1 Introduction

3.1.1 Overview of other pricing methods

In the last decades different methods of pricing derivatives based on non-tradeable underlyings have been developed. We want to give an overview of some other different pricing methods for weather derivatives. A straightforward method is burn analysis or just "burn". The main idea of this method is how the contract would have performed in previous years. Although the derivation is very simple, but this method might be a good first step in pricing contracts. A more accurate possibility of pricing contracts is modeling different indices (cf. Jewson et al. [38]). Index modeling is a calculation of statistical parameters (mean, variance etc.) and fitting an adequate distribution for the index. This method can be seen as an actuarial approach. Platon and West [47] suggest a benchmark approach for pricing temperature derivatives, which is based on the existence of an optimal benchmark portfolio. Cao and Wei [22] introduce an equilibrium approach and apply an extended version of Lucas' equilibrium pricing model. On the other hand, Brockett et al. [19] suggests an indifference pricing approach and deals with portfolio effects and possible hedging strategies. In further chapters we will also give an indifference pricing approach which is based on an exponential utility function. A different view of pricing weather derivatives is given by Davis [28]. His idea is pricing by marginal values technique, he models the HDD-index directly by a geometric Brownian motion. There are still further pricing methods to be found in the literature.

3.1.2 Temperature Indices

At the CME there exists an organized market of futures and options written on different temperature indices. The most frequently traded futures are written on degree days over one month or one season in US cities. We consider four different indices in detail.

Heating degree days (HDD)

This index has its roots in the energy industry. It is used to measure the demand for heating, and thus a measure how cold it is. In the US, where temperature is measured in Fahrenheit, the baseline is taken to be 65. For European cities the baseline will be taken by 18 degree Celsius. In the following we will concentrate on the degree Celsius case.

Definition 3.1 (HDD) *The HDD index over an N_d day period is defined as*

$$\sum_{i=1}^{N_d} \max(18 - T_i, 0)$$

where 18 denotes the baseline and T_i the average temperature on day i.
In a continuous time setting, the HDD index over the time interval $[\tau_1, \tau_2]$ is defined as

$$\int_{\tau_1}^{\tau_2} \max(18 - T(\tau), 0) d\tau.$$

Cooling degree days (CDD)

Just as the HDD index, the CDD arises from the energy industry. It is used during the warm season to measure the demand for energy used for cooling. Derivatives written on this index are mainly offered for US cities. In European cities this index is rarely used, because the energy demand for cooling is very low. The use of air conditions is much popular in US cities than in European cities.

Definition 3.2 (CDD) *The CDD index over an N_d day period is defined as*

$$\sum_{i=1}^{N_d} \max(T_i - 18, 0)$$

where 18 denotes the baseline and T_i the average temperature on day i.
In a continuous time setting, the CDD index over the time interval $[\tau_1, \tau_2]$ is defined as

$$\int_{\tau_1}^{\tau_2} \max(T(\tau) - 18, 0) d\tau.$$

Cumulative average temperature (CAT)

This index is mainly used for European cities during the summer.

Definition 3.3 (CAT) *The CAT index over an N_d day period is defined as*

$$\sum_{i=1}^{N_d} T_i$$

where T_i denotes the average temperature on day i.
In a continuous time setting, the CAT index over the time interval $[\tau_1, \tau_2]$ is defined as

$$\int_{\tau_1}^{\tau_2} T(\tau)d\tau.$$

Pacific Rim index (PRIM)

Derivatives written on this index are only organized for Asian cities. This index measures the average daily temperature over a month or a season.

Definition 3.4 (PRIM) *The Pacific Rim index over an N_d day period is defined as*

$$\frac{1}{N_d}\sum_{i=1}^{N_d} T_i$$

where T_i denotes the average temperature on day i.
In a continuous time setting, the Pacific Rim index over the time interval $[\tau_1, \tau_2]$ is defined as

$$\frac{1}{\tau_2 - \tau_1}\int_{\tau_1}^{\tau_2} T(\tau)d\tau.$$

More information about the different indices and simple calculation examples can be found in Jewson et al. [38].

3.2 Analysis of derivatives on temperature

In this section we derive prices of different temperature derivatives. Before we start with the derivation we want to define some general assumptions.

3.2.1 General assumptions

First of all we assume that the market is arbitrage free ("there is no free lunch"). This assumption is fundamental for pricing financial assets. If the opportunity of an arbitrage exists, an investor can earn a profit without taking any risk. But such opportunities will be wiped out through competition in all efficient markets, so the lack of arbitrage seems to be a reasonable mathematical assumption. In particular it is not possible to make a profit by speculating with a tradeable asset in the market without taking any risk. An asset which is tradeable means that it can be bought or sold on the market. Temperature is obviously a non-tradeable asset. A main result of the pricing theory is the relation between no-arbitrage and the existence of a equivalent martingale probability measure.

Theorem 3.1 (Fundamental theorem of asset pricing) *The market model defined by* $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$ *and asset prices* $(S_t)_{t\in[0,T]}$ *is arbitrage-free if and only if there exists a probability measure* $\mathbb{Q} \sim \mathbb{P}$ *such that the discounted assets* $(\tilde{S}_t)_{t\in[0,T]}$ *are martingales with respect to* $\mathbb{Q}$.

Definition 3.5 *Two probability measures* $\mathbb{P}$ *and* $\mathbb{Q}$ *on* $(\Omega, \mathcal{F})$ *are equivalent if they define the same impossible events:*

$$\mathbb{P} \sim \mathbb{Q} \Longleftrightarrow [\forall A \in \mathcal{F}, \mathbb{P}(A) = 0 \Leftrightarrow \mathbb{Q}(A) = 0] .$$

For more details see Øksendal [45], Cont and Tankov [26] and Klebaner [42].
We denote by r the risk-free interest rate, which is also used for discounting the tradeable assets. The risk-free interest rate is in our case constant and is the interest rate which will be earned on a risk-free bond. The price of an derivative is established by the present expected value of the payoff from the derivative with respect to the pricing measure $\mathbb{Q}$. In many situations a claim or a derivative can be replicated by a dynamic investment strategy, e.g. hedged. In some market models (e.g. Black-Scholes) any contingent claim can be represented as the final value of a self-financing strategy. Such markets are called complete. The following theorem shows us a relation between complete markets and the equivalent martingale measure $\mathbb{Q}$.

Theorem 3.2 (Second fundamental theorem of asset pricing) *A market defined by the assets* $(S_t^0, S_t^1, \ldots, S_t^d)_{t\in[0,T]}$ *described as a stochastic processes on* $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$ *is complete if and only f there exists a unique martingale measure* $\mathbb{Q}$ *equivalent to* $\mathbb{P}$.

Many popular models are arbitrage-free, but incomplete. For more details see Cont and Tankov [26].
If we discount the risk-free bond, we obtain a constant process 1 which is trivially a martingale with respect to the martingale measure $\mathbb{Q}$. Note that in our framework we consider temperature derivatives like futures written on different temperature indices, which are obviously non-tradeable assets. In the considered market the only tradeable asset is the risk-free bond, which is a martingale after discounting with respect to all equivalent measures $\mathbb{Q}$. It means that all equivalent measures $\mathbb{Q}$ are risk neutral probabilities. We cannot perfectly hedge any temperature future, therefore the market is incomplete. Note that the futures are tradeable contracts, therefore their price dynamics should be arbitrage-free. To get this dynamics, we use any equivalent measure $\mathbb{Q}$ as a pricing measure. Since options are written on tradeable contracts, we use the same martingale measure $\mathbb{Q}$, which we have determined for

calculating the futures price.
We specify a class of risk neutral probabilities by using the Girsanov transform (see Øksendal [45]). Assume that $\zeta(t)$ is a real-valued measurable and bounded function. The stochastic process

$$Z^{\zeta}(t) = exp\left(\int_0^t \frac{\zeta(s)}{\sigma(s)} dB(s) - \frac{1}{2}\int_0^t \frac{\zeta^2(s)}{\sigma^2(s)}\right) \tag{3.1}$$

will become the density process of the probability measure

$$\mathbb{Q}^{\zeta}(A) = \mathbb{E}[1_A Z^{\zeta}(\tau_{max})], \tag{3.2}$$

where 1_A is the indicator function and τ_{max} is a fixed time horizon bigger than the trading times for all relevant futures. The probability is equivalent to $\mathbb{P}$, and the process

$$dW(t) = dB(t) - \frac{\zeta(t)}{\sigma(t)} dt \tag{3.3}$$

is a standard Brownian motion under $\mathbb{Q}^{\zeta}$. An important assumption is that the parameters of $\sigma^2(t)$ (in Chapter 2) are chosen that the function is bounded away from zero, thus creating no problems when using it as a divisor in the Girsanov transformation. The dynamics of $T(t)$ under $\mathbb{Q}^{\zeta}$ becomes

$$dT(t) = dM(t) + (\zeta(t) - \theta(T(t) - M(t)))dt + \sigma(t)dW(t). \tag{3.4}$$

Applying the Itô-Formula we get an explicit form of $T(t)$ which is given by

$$T(t) = M(t) + (T(0) - M(0))e^{-\theta t} + \int_0^t \zeta(u)e^{-\theta(t-u)}du + \int_0^t \sigma(u)e^{-\theta(t-u)}dW(u). \tag{3.5}$$

Here we call ζ the market price of risk. The expectation under $\mathbb{Q}^{\zeta}$ is denoted by $\mathbb{E}_{\zeta}[.]$. Obviously is $T(t)$ normally distributed under $\mathbb{Q}^{\zeta}$, with expectation

$$\mathbb{E}_{\zeta}[T(t)] = M(t) + (T(0) - M(0))e^{\theta t} + \int_0^t \zeta(t)e^{-\theta(t-u)}du \tag{3.6}$$

and variance

$$\text{Var}_{\zeta}[T(t)] = \int_0^t \sigma^2(u)e^{-2\theta(t-u)}du. \tag{3.7}$$

Before we start calculating prices of futures traded at the CME, we introduce a constant interest rate r and consider the price dynamics of futures over a specified period $[\tau_1, \tau_2]$, $\tau_1 < \tau_2$ (a month during the winter). The future price at time $t \leq \tau_1$ written on the HDD index is defined as the $\mathcal{F}_t$-adapted stochastic process $F_{HDD}(t, \tau_1, \tau_2)$ satisfying

$$0 = e^{-r(\tau_2 - t)}\mathbb{E}_{\zeta}\left[\int_{\tau_1}^{\tau_2} \max(18 - T(\tau), 0)d\tau - \mathrm{F}_{HDD}(t, \tau_1, \tau_2)|\mathcal{F}_t\right] \tag{3.8}$$

Using the fact that the future price is $\mathcal{F}_t-$adapted, we find

$$F_{HDD}(t, \tau_1, \tau_2) = \mathbb{E}_{\zeta}\left[\int_{\tau_1}^{\tau_2} \max(18 - T(\tau), 0)d\tau|\mathcal{F}_t\right]. \tag{3.9}$$

The future price written on an the CDD index is

$$F_{CDD}(t,\tau_1,\tau_2) = \mathbb{E}_\zeta\left[\int_{\tau_1}^{\tau_2} \max(T(\tau)-18,0)d\tau|\mathcal{F}_t\right]. \tag{3.10}$$

Now we also can derive the prices of CAT-futures and Pacific Rim futures, which are

$$F_{CAT}(t,\tau_1,\tau_2) = \mathbb{E}_\zeta\left[\int_{\tau_1}^{\tau_2} T(\tau)d\tau|\mathcal{F}_t\right]. \tag{3.11}$$

and

$$F_{PRIM}(t,\tau_1,\tau_2) = \mathbb{E}_\zeta\left[\frac{1}{\tau_2-\tau_1}\int_{\tau_1}^{\tau_2} T(\tau)d\tau|\mathcal{F}_t\right]. \tag{3.12}$$

Obviously there exists a trivial relation

$$F_{PRIM}(t,\tau_1,\tau_2) = \frac{1}{\tau_2-\tau_1}F_{CAT}(t,\tau_1,\tau_2). \tag{3.13}$$

Since $\max(x-18,0)-\max(18-x,0)=x-18$, we have

$$F_{HDD}(t,\tau_1,\tau_2) = 18*(\tau_2-\tau_1) - F_{CAT}(t,\tau_1,\tau_2) + F_{CDD}(t,\tau_1,\tau_2). \tag{3.14}$$

3.2.2 CAT-futures and options

Before we give an explicit CAT future price, which is defined by the risk neutral conditional expectation

$$F_{CAT}(t,\tau_1,\tau_2) = \mathbb{E}_\zeta\left[\int_{\tau_1}^{\tau_2} T(\tau)d\tau|\mathcal{F}_t\right], \tag{3.15}$$

we give an expression for the cumulative temperature under the risk neutral probability $\mathbb{Q}^\zeta$.

Proposition 3.1 *If the temperature $T(t)$ follows (3.4), the cumulative temperature over the time interval $[\tau_1,\tau_2]$ is explicitly given by*

$$\begin{aligned}\int_{\tau_1}^{\tau_2} T(t)dt = &\int_{\tau_1}^{\tau_2} M(t)dt - \theta^{-1}(T(0)-M(0))\left(e^{\theta\tau_2}-e^{\theta\tau_1}\right)\\ &-\int_0^{\tau_2}\zeta(t)\theta^{-1}\left(e^{-\theta(\tau_2-t)} - 1_{[0,\tau_1]}(t)e^{-\theta(\tau_1-t)} - 1_{[\tau_1,\tau_2]}(t)\right)dt\\ &-\int_0^{\tau_2}\sigma(t)\theta^{-1}\left(e^{-\theta(\tau_2-t)} - 1_{[0,\tau_1]}(t)e^{-\theta(\tau_1-t)} - 1_{[\tau_1,\tau_2]}(t)\right)dW(t).\end{aligned}$$

Proof:

Let $\tilde{T}(t)=T(t)-M(t)$ be the deseasonalized temperature. From (3.4) we obtain

$$\tilde{T}(\tau_2) = \tilde{T}(\tau_1) - \theta\int_{\tau_1}^{\tau_2}\tilde{T}(t)dt + \int_{\tau_1}^{\tau_2}\zeta(t)dt + \int_{\tau_1}^{\tau_2}\sigma(t)dW(t). \tag{3.16}$$

Combining this expression with the explicit dynamics (3.5) of $T(t)$,

$$\tilde{T}(t) = T(t) - M(t) = (T(0) - M(0))\, e^{-\theta t} + \int_0^t \zeta(t) e^{-\theta(t-u)} du + \int_0^t \sigma(u) e^{-\theta(t-u)} dW(u) \tag{3.17}$$

and the result follows. □

This result is helpful for the calculation of the CAT-future price.

Theorem 3.3 *The CAT-futures price $F_{CAT}(t,\tau_1,\tau_2)$ at time $t \leq \tau_1$ where the index is measured over the interval $[\tau_1,\tau_2]$, is given by*

$$F_{CAT}(t,\tau_1,\tau_2) = \int_{\tau_1}^{\tau_2} M(\tau) d\tau + \theta^{-1}\left(e^{-\theta(\tau_1 - t)} - e^{-\theta(\tau_2 - t)}\right)(T(t) - M(t)) + \Theta(t,\tau_1,\tau_2), \tag{3.18}$$

where $\Theta(t,\tau_1,\tau_2)$ is given as function of the market price of risk and volatility as

$$\Theta(t,\tau_1,\tau_2) = \theta^{-1}\int_t^{\tau_2} \zeta(u)\left(1 - e^{-\theta(\tau_2 - u)}\right) du - \theta^{-1}\int_t^{\tau_1} \zeta(u)\left(1 - e^{-\theta(\tau_1 - u)}\right) du.$$

Proof:

The price of a CAT-future is given by

$$F_{CAT}(t,\tau_1,\tau_2) = \mathbb{E}_\zeta\left[\int_{\tau_1}^{\tau_2} T(\tau) d\tau | \mathcal{F}_t\right].$$

inserting the result form the proposition above leads to

$$\begin{aligned} F_{CAT}(t,\tau_1,\tau_2) = \mathbb{E}_\zeta \Bigg[& \int_{\tau_1}^{\tau_2} M(u) du - \theta^{-1}(T(t) - M(t))\left(e^{-\theta(\tau_2 - t)} - e^{-\theta(\tau_1 - t)}\right) \\ & - \int_t^{\tau_2} \zeta(u)\theta^{-1}\left(e^{-\theta(\tau_2 - u)} - 1_{[t,\tau_1]}(u) e^{-\theta(\tau_1 - u)} - 1_{[\tau_1,\tau_2]}(u)\right) du \\ & - \int_t^{\tau_2} \sigma(u)\theta^{-1}\left(e^{-\theta(\tau_2 - u)} - 1_{[t,\tau_1]}(u) e^{-\theta(\tau_1 - u)} - 1_{[\tau_1,\tau_2]}(u)\right) dW(u) | \mathcal{F}_t \Bigg]. \end{aligned}$$

We obtain

$$\begin{aligned} F_{CAT}(t,\tau_1,\tau_2) = & \int_{\tau_1}^{\tau_2} M(u) du - \theta^{-1}(T(t) - M(t))\left(e^{-\theta(\tau_2 - t)} - e^{-\theta(\tau_1 - t)}\right) \\ & - \int_t^{\tau_2} \zeta(u)\theta^{-1}\left(e^{-\theta(\tau_2 - u)} - 1_{[t,\tau_1]}(u) e^{-\theta(\tau_1 - u)} - 1_{[\tau_1,\tau_2]}(u)\right) du. \end{aligned}$$

Rewriting the last integral, the result follows. □

The CAT future price is given by the aggregated mean temperature over the measurement period plus a weighted dependency of $\tilde{T}(t)$, which means that the price depends on the temperature at time t. The two last terms only smooth the market price of risk over the considered period with a change, which happens at time τ_1.
A straightforward application of the Itô-Formula yields the following result.

Theorem 3.4 (Time dynamics of CAT-future) *The dynamics under the measure* $\mathbb{Q}^{\zeta}$ *of* $F_{CAT}(t, \tau_1, \tau_2)$ *is*

$$dF_{CAT}(t, \tau_1, \tau_2) = \Sigma(t, \tau_1, \tau_2) dW(t) \tag{3.19}$$

where

$$\Sigma(t, \tau_1, \tau_2) := \theta^{-1} \left(e^{-\theta(\tau_1 - t)} - e^{-\theta(\tau_2 - t)} \right) \sigma(t) \tag{3.20}$$

Proof:
Let

$$\tilde{T}(t) = T(t) - M(t).$$

We obtain by applying the Itô-Formula

$$\begin{aligned} dF_{CAT}(t, \tau_1, \tau_2) &= \frac{\partial F_{CAT}}{\partial t} dt + \frac{\partial F_{CAT}}{\partial \tilde{T}} d\tilde{T}(t) + \frac{1}{2} \frac{F_{CAT}}{\partial \tilde{T}^2} \left(d\tilde{T}(t) \right)^2 \\ &= \left(e^{-\theta(\tau_1 - t)} - e^{-\theta(\tau_2 - t)} \right) \left[\left(T(t) - M(t) - \theta^{-1} \zeta(t) \right) dt + \theta^{-1} d\tilde{T}(t) \right]. \end{aligned}$$

Note that

$$d\tilde{T}(t) = (\zeta(t) - \theta(T(t) - M(t))) dt + \sigma(t) dW(t)$$

and so

$$dF_{CAT}(t, \tau_1, \tau_2) = \left(\theta^{-1} \left(e^{-\theta(\tau_1 - t)} - e^{-\theta(\tau_2 - t)} \right) \sigma(t) \right) dW(t).$$

□

Here we can interpret $\Sigma(t, \tau_1, \tau_2)$ as the CAT-futures volatility function of the CAT-futures dynamics. In Figure 3.1 we plotted the volatility function for the case of Graz. The CAT volatility is a function of t for contracts with a monthly measurement period and we plotted the function until the beginning of each delivery month. The first curve is the volatility of a contract whose delivery is in February plotted for t starting January 1, and ending January 31. The second curve is $\Sigma(t, \tau_1, \tau_2)$ in the time range from January 1 until February 28 with delivery in March, and so on until delivery in January of the following year. We notice that the volatility is for each contract is close to zero when the time to measurement is large, but it increases sharply up to the start of the measurement period. This maturity effect of the volatility is called Samuelson effect and this effect here is in accordance with observations in the commodity markets. The CAT-volatility as a seasonally varying maximum volatility reflects the seasonality of the function $\sigma(t)$.

The fact that the CAT-future price is an additive Gaussian process makes the calculation of an explicit formula for a European call option quite easy. The price of a call with exercise date $\tau \leq t$ and strike price K is given in the following theorem.

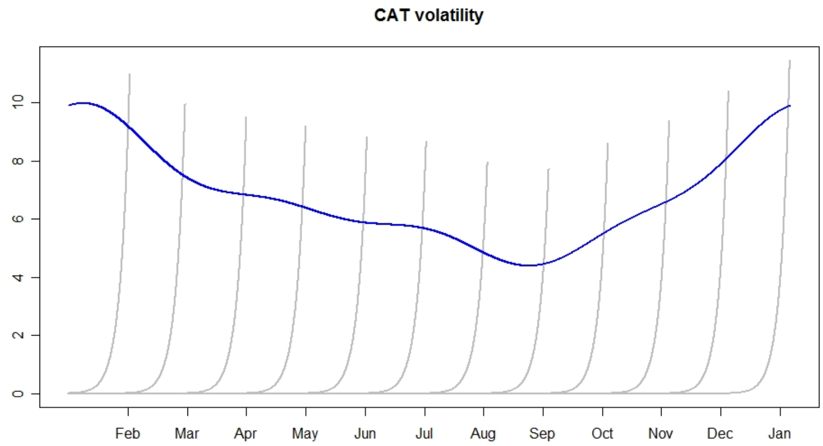

Figure 3.1: CAT-volatility with $\sigma^2(t)$

Theorem 3.5 *The price of a call option at the time t on the CAT-future contract with exercise date $t \leq \tau \leq \tau_1$ and the strike price K is*

$$C_{CAT}(t,\tau,\tau_1,\tau_2) = e^{-r(\tau-t)}(F_{CAT}(t,\tau_1,\tau_2) - K)\Phi(d(t,\tau)) + e^{-r(\tau-t)}\Sigma_{t,\tau}\frac{1}{\sqrt{2\pi}}e^{\frac{-d(t,\tau)^2}{2}} \tag{3.21}$$

where

$$d(t,\tau) = \frac{F_{CAT}(t,\tau_1,\tau_2) - K}{\Sigma_{t,\tau}}$$

$$\Sigma^2_{t,\tau} := \int_t^\tau \Sigma^2(u,\tau_1,\tau_2)du$$

Φ the cumulative probability function for the standard normal distribution and $\Sigma(t,\tau_1,\tau_2)$ is defined in (3.20).

Proof:

By risk neutral pricing the call price must satisfy

$$C_{CAT}(t,\tau,\tau_1,\tau_2) = e^{-r(\tau-t)}\mathbb{E}_\zeta\left[max(F_{CAT}(t,\tau_1,\tau_2) - K, 0)|\mathcal{F}_t\right].$$

We know that

$$\int_t^\tau dF_{CAT}(u,\tau_1,\tau_2)du = F_{CAT}(\tau,\tau_1,\tau_2) - F_{CAT}(t,\tau_1,\tau_2)$$

and by (3.19), we obtain

$$F_{CAT}(\tau,\tau_1,\tau_2) = F_{CAT}(t,\tau_1,\tau_2) + \int_t^\tau \Sigma(u,\tau_1,\tau_2)dW(u). \tag{3.22}$$

We require for $F_{CAT}(\tau,\tau_1,\tau_2) > K$

$$\int_t^\tau \Sigma(u,\tau_1,\tau_2)dW(u) > K - F_{CAT}(t,\tau_1,\tau_2). \tag{3.23}$$

To solve this problem, we have to determine the distribution of

$$\int_t^\tau \Sigma(u,\tau_1,\tau_2)dW(u) \tag{3.24}$$

at time t. We know that this integral is normally distributed with

$$\mathbb{E}_\zeta\left[\int_t^\tau \Sigma(u,\tau_1,\tau_2)dW(u)|\mathcal{F}_t\right] = 0$$

and by the Itô isometry

$$\mathbb{E}_\zeta\left[\left(\int_t^\tau \Sigma(u,\tau_1,\tau_2)dW(u)\right)^2|\mathcal{F}_t\right] = \int_t^\tau \Sigma^2(u,\tau_1,\tau_2)du.$$

Hence

$$\int_t^\tau \Sigma(u,\tau_1,\tau_2)dW(u) \sim N(0,\int_t^\tau \Sigma^2(u,\tau_1,\tau_2)du).$$

Let us return to (3.23), for $F_{CAT}(\tau,\tau_1,\tau_2) > K$, we require

$$X\sqrt{\int_t^\tau \Sigma^2(u,\tau_1,\tau_2)du} > K - F_{CAT}(t,\tau_1,\tau_2), \quad \text{where } X \sim N(0,1).$$

This expression is equivalent to

$$X > \frac{K - F_{CAT}(t,\tau_1,\tau_2)}{\sqrt{\int_t^\tau \Sigma^2(u,\tau_1,\tau_2)du}} := d_1(t,\tau).$$

We obtain

$$\begin{aligned}
C_{CAT}(t,\tau,\tau_1,\tau_2) &= e^{-r(\tau-t)}\int_{d_1(t,\tau)}^\infty \left(F_{CAT}(t,\tau_1,\tau_2) + x\sqrt{\int_t^\tau \Sigma^2(u,\tau_1,\tau_2)du} - K\right)\frac{e^{-\frac{x^2}{2}}}{\sqrt{2\pi}}dx \\
&= e^{-r(\tau-t)}(F_{CAT}(t,\tau_1,\tau_2) - K)\Phi(-d_1(t,\tau)) \\
&+ \sqrt{\int_t^\tau \Sigma^2(u,\tau_1,\tau_2)du}e^{-r(\tau-t)}\int_{d_1(t,\tau)}^\infty \frac{xe^{-\frac{x^2}{2}}}{\sqrt{2\pi}}dx.
\end{aligned}$$

Set

$$\Sigma_{t,\tau}^2 := \int_t^\tau \Sigma^2(u,\tau_1,\tau_2)du$$

and the result follows. □

Note that the market which consists of futures and options is complete. Hence we can hedge the option price perfectly and the option price is unique. The optimal hedging strategy for the call option will be described by the option's delta, which is the sensitivity of the option price with respect to the CAT-future price. In an analoguous way the hedging strategy in the Black-Scholes model can be described. The delta of the call option, is given by

$$\frac{\partial C_{CAT}(t,\tau,\tau_1,\tau_2)}{\partial F_{CAT}(t,\tau_1,\tau_2)} = \Phi(d(t,\tau)). \tag{3.25}$$

The notation is the same as above. To replicate the call perfectly, this delta-hedge ratio gives the number of CAT-futures that should be held in the hedging portfolio.

3.2.3 PRIM-futures

For the sake of completeness we give some explicit formulas for pacific rim futures which follows immediately from the cat future prices. Contracts which are written on this index are denominated in Japanese Yen and created for the Asian market. Recall that

$$F_{PRIM}(t,\tau_1,\tau_2) = \mathbb{E}_\zeta\left[\frac{1}{\tau_2-\tau_1}\int_{\tau_1}^{\tau_2} T(\tau)d\tau|\mathcal{F}_t\right] = \frac{1}{\tau_2-\tau_1}\mathbb{E}_\zeta\left[\int_{\tau_1}^{\tau_2} T(\tau)d\tau|\mathcal{F}_t\right]. \tag{3.26}$$

It follows trivially that

$$F_{PRIM}(t,\tau_1,\tau_2) = \frac{1}{\tau_2-\tau_1}F_{CAT}(t,\tau_1,\tau_2) \tag{3.27}$$

and

$$dF_{PRIM}(t,\tau_1,\tau_2) = \frac{1}{\tau_2-\tau_1}dF_{CAT}(t,\tau_1,\tau_2). \tag{3.28}$$

To obtain the price of an PRIM-call option we only have to modify the call price of the CAT-option. The price is given by

$$C_{PRIM}(t,\tau,\tau_1,\tau_2) = e^{-r(\tau-t)}\mathbb{E}_\zeta\left[\max(F_{PRIM}(t,\tau_1,\tau_2)-K,0)|\mathcal{F}_t\right]. \tag{3.29}$$

Dividing (3.22) by $(\tau_2-\tau_1)$ we get

$$F_{PRIM}(\tau,\tau_1,\tau_2) = F_{PRIM}(t,\tau_1,\tau_2) + \frac{1}{(\tau_2-\tau_1)}\int_t^\tau \Sigma(u,\tau_1,\tau_2)dW(u) \tag{3.30}$$

It follows that

$$F_{PRIM}(\tau,\tau_1,\tau_2) \sim N\left(F_{PRIM}(t,\tau_1,\tau_2), \frac{\Sigma_{t,\tau}^2}{(\tau_2-\tau_1)^2}\right). \tag{3.31}$$

For the call price of a PRIM future we obtain

$$\begin{aligned} C_{PRIM}(t,\tau,\tau_1,\tau_2) &= e^{-r(\tau-t)}(F_{PRIM}(t,\tau_1,\tau_2)-K)\Phi(d(t,\tau)(\tau_2-\tau_1)) \\ &+ e^{-r(\tau-t)}\frac{\Sigma_{t,\tau}}{(\tau_2-\tau_1)}\frac{1}{\sqrt{2\pi}}e^{\frac{-d(t,\tau)^2}{2}(\tau_2-\tau_1)^2}. \end{aligned}$$

3.2.4 HDD/CDD-futures

We now derive explicit prices of futures and options based on futures which are written on HDD or CDD index. These derivatives are the most traded weather derivatives at CME. For the European market only HDD-futures are relevant, while CDD-futures are especially traded for American cities. Recall the HDD-CDD parity (3.14) and the price of a HDD-future which is given by

$$F_{HDD}(t,\tau_1,\tau_2) = \mathbb{E}_\zeta \left[\int_{\tau_1}^{\tau_2} \max(18 - T(\tau), 0) d\tau | \mathcal{F}_t \right]. \tag{3.32}$$

The following theorem gives an explicit formula.

Theorem 3.6 (HDD-future) *The price of a HDD-future at time t, where the index is measured of the period $[\tau_1, \tau_2]$, $t \le \tau_1 < \tau_2$, is given by*

$$F_{HDD}(t,\tau_1,\tau_2) = \int_{\tau_1}^{\tau_2} \Sigma(t,\tau) \left[d(t,\tau,T(t)) \Phi(d(t,\tau,T(t))) + \phi(d(t,\tau,T(t))) \right] d\tau \tag{3.33}$$

where

$$d(t,\tau,x) = \frac{18 - M(\tau) + (M(t) - x)e^{-\theta(\tau-t)} - \int_t^\tau \zeta(u) e^{-\theta(\tau-t)} du}{\Sigma(t,\tau)}$$

$$\Sigma^2(t,\tau) = \int_t^\tau \sigma^2(u) e^{-2\theta(t-u)} du$$

and Φ denotes the cumulative probability function of a standard normal variable and ϕ its density.

Proof:

We can interchange expectation and integration by using the theorem of Fubini-Tonelli and we obtain

$$F_{HDD}(t,\tau_1,\tau_2) = \int_{\tau_1}^{\tau_2} \mathbb{E}_\zeta \left[\max(18 - T(\tau), 0) | \mathcal{F}_t \right] d\tau$$

Let us calculate

$$\mathbb{E}_\zeta \left[\max(18 - T(\tau), 0) | \mathcal{F}_t \right]$$

For the temperature we can write

$$T(\tau) = M(\tau) + (T(t) - M(t)) e^{-\theta(\tau-t)} + \int_t^\tau \zeta(u) e^{-\theta(t-u)} du + \int_t^\tau \sigma(u) e^{-\theta(t-u)} dW(u)$$

which is normally distributed with expectation

$$d_1(t,\tau,T(t)) := M(\tau) + (T(t) - M(t)) e^{-\theta(\tau-t)} + \int_t^\tau \zeta(u) e^{-\theta(t-u)} du$$

and variance

$$\Sigma^2(t,\tau) := \int_t^\tau \sigma^2(u) e^{-2\theta(t-u)} du.$$

Let us express the temperature dynamics under $\mathbb{Q}^{\zeta}$ in terms of the standard normal random variables $X \sim N(0,1)$ as

$$T(\tau) = d_1(t,\tau,T(t)) + \Sigma^2(t,\tau)^{\frac{1}{2}} \times X.$$

Let us consider

$$18 > T(\tau)$$

and it follows

$$18 - d_1(t,\tau,T(t)) > \Sigma(t,\tau) \times X$$
$$X < \frac{18 - d_1(t,\tau,T(t))}{\Sigma(t,\tau)} := d(t,\tau,T(t)).$$

We obtain

$$\begin{aligned}
\mathbb{E}_{\zeta}\left[\max\left(18 - T(\tau),0\right)\right] &= \int_{-\infty}^{d(t,\tau,x)} (18 - T(\tau))\frac{e^{-\frac{x^2}{2}}}{\sqrt{2\pi}}dx \\
&= \int_{-\infty}^{d(t,\tau,x)} \left(d(t,\tau,T(t))\Sigma(t,\tau) - \Sigma(t,\tau) \times x\right)\frac{e^{-\frac{x^2}{2}}}{\sqrt{2\pi}}dx \\
&= \Sigma(t,\tau)\left(d(t,\tau,x)\Phi(d(t,\tau,T(t))) - \int_{-\infty}^{d(t,\tau,x)} x\frac{e^{-\frac{x^2}{2}}}{\sqrt{2\pi}}dx\right) \\
&= \Sigma(t,\tau)\left(d(t,\tau,x)\Phi(d(t,\tau,T(t))) + \phi(d(t,\tau,T(t)))\right)
\end{aligned}$$

and the result follows immediately. □

Analogously we can derive the price of a CDD-future.

Theorem 3.7 (CDD-future) *The price of a CDD-future at time t, where the index is measured of the period $[\tau_1,\tau_2]$, $t \le \tau_1 < \tau_2$, is given by*

$$F_{CDD}(t,\tau_1,\tau_2) = \int_{\tau_1}^{\tau_2} \Sigma(t,\tau)\left[d(t,\tau,T(t))\Phi(d(t,\tau,T(t))) + \phi(d(t,\tau,T(t)))\right]d\tau \quad (3.34)$$

where

$$d(t,\tau,x) = \frac{M(\tau) - (M(t)+x)e^{-\theta(\tau-t)} + \int_t^{\tau}\zeta(u)e^{-\theta(\tau-t)}du - 18}{\Sigma(t,\tau)}$$
$$\Sigma^2(t,\tau) = \int_t^{\tau}\sigma^2(u)e^{-2\theta(t-u)}du$$

and Φ denotes the cumulative probability function of a standard normal variable and ϕ its density.

Proof: Analogous to the proof above. □

Recall the Itô-Formula. By applying it, we can calculate the dynamics of the HDD-future price, $dF_{HDD}(t, \tau_1, \tau_2)$.

Theorem 3.8 (Time dynamics of HDD-future) *The time-dynamics of a HDD-future where the index is measured over the period* $[\tau_1, \tau_2]$ *is given by*

$$dF_{HDD}(t, \tau_1, \tau_2) = -\sigma(t) \int_{\tau_1}^{\tau_2} e^{-\theta(\tau - t)} \Phi(d(t, \tau, T(t))) d\tau dW(t) \tag{3.35}$$

for $t \leq \tau_1 < \tau_2$, *with*

$$d(t, \tau, x) = \frac{18 - M(\tau) + (M(t) - x)e^{-\theta(\tau - t)} - \int_t^\tau \zeta(u) e^{-\theta(\tau - t)} du}{\Sigma(t, \tau)}.$$

By changing $d(t, \tau, x)$ we get the same result for CDD-futures.

Proof:

To proof this theorem we have to apply the Itô-Formula, already mentioned above.

$$dF_{HDD}(t, \tau_1, \tau_2) = \frac{\partial F_{HDD}}{\partial t} dt + \frac{\partial F_{HDD}}{\partial T} dT(t) + \frac{1}{2} \frac{F_{HDD}}{\partial T^2} (dT(t))^2$$

where $dT(t)$ is defined in (3.4). The fact that $F_{HDD}(t, \tau_1, \tau_2)$ is a $\mathbb{Q}^\zeta-$martingale, we only need to focus on the $dW-$term in the Itô-Formula, and so

$$dF_{HDD}(t, \tau_1, \tau_2) = \frac{\partial F_{HDD}}{\partial T} \sigma(t) dW(t) \tag{3.36}$$

$$\begin{aligned} \frac{\partial F_{HDD}}{\partial T} &= \int_{\tau_1}^{\tau_2} \Sigma(t, \tau) \left(\frac{e^{-\theta(\tau - t)}}{\Sigma(t, \tau)} \right) \Phi(d(t, \tau, T(t))) d\tau \\ &+ \int_{\tau_1}^{\tau_2} \Sigma(t, \tau) d(t, \tau, T(t)) \phi(d(t, \tau, T(t))) \left(\frac{e^{-\theta(\tau - t)}}{\Sigma(t, \tau)} \right) d\tau \\ &+ \int_{\tau_1}^{\tau_2} \Sigma(t, \tau) \phi(d(t, \tau, T(t))) \left(\frac{e^{-\theta(\tau - t)}}{\Sigma(t, \tau)} \right) (-d(t, \tau, T(t))) d\tau \\ &= \int_{\tau_1}^{\tau_2} e^{-\theta(\tau - t)} \Phi(d(t, \tau, T(t))) d\tau. \end{aligned}$$

After substitution in (3.36), the result follows. □

We see that the dynamics of F_{HDD} is complicated and it is not possible to derive explicit closed formulas of put or call options based on HDD-futures. If we want to calculate prices for put or call options written on HDD futures we have to approach the price numerically. A possible approach could be a Monte Carlo method, simulating the risk neutral temperature $T(\tau)$ at the strike time τ, and then integrate numerically to compute $F_{HDD}(\tau, \tau_1, \tau_2)$. To obtain HDD-future prices we have to specify the market price of risk ζ. By choosing an appropriate family ζ, we can fit these using today's observed HDD-future curve by appealing to the theoretical price curve yielded by (3.33):

$$F_{HDD}(0, \tau_1, \tau_2) = \int_{\tau_1}^{\tau_2} \Sigma(0, \tau) \left[d(0, \tau, T(0))\Phi(d(0, \tau, T(0))) + \phi(d(0, \tau, T(0)))\right] d\tau \tag{3.37}$$

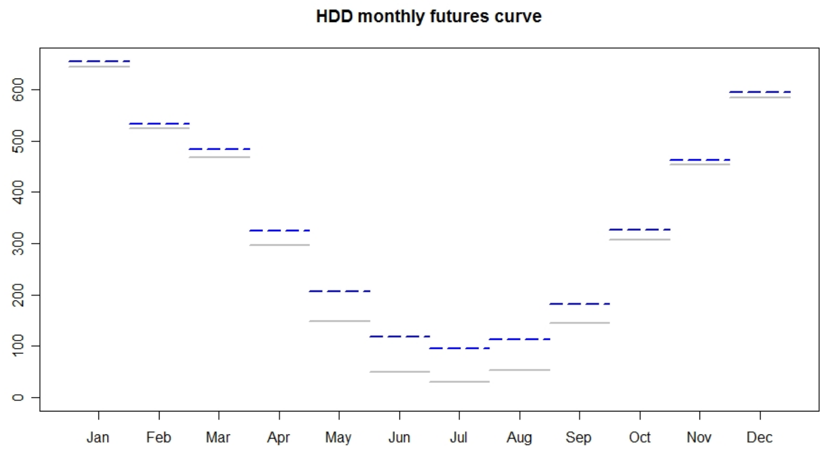

Figure 3.2: HDD-future curve (3.37) for each month based on the model of airport Graz. The constant lines represent the prices if the volatility is constant and the broken lines represent the prices if we choose the volatility function of model C in Chapter 2.

By recalling (3.33) we see that the dependency on ζ is in the function d. In Figure 3.2 we plotted the HDD-future curve (3.33) at time 0 for each month of the following year based on model C which we fitted in Chapter 2. We assumed that the market price of risk is zero. In this figure we see two different line types. On the one hand we plotted the HDD-Price if $\sigma(t)$ is the function which we determined in Chapter 2 (broken lines) and on the other hand we represent the prices of HDD-futures if $\sigma(t)$ is constant. The constant is chosen to be the standard deviation of the residuals from the regression analysis, which we recall to be $\bar{\sigma} = 2.08$ (see Chapter 2). Obviously, a constant volatility leads to a permanent underestimation of the future curve during the whole year. We see that the error is smaller during the winter, but during the summer prices differs enormously. This figure emphasizes that it is important to have a good model for the temperature volatility.

Chapter 4

Hedging Spatial Temperature Risk

In the last chapter we gave explicit pricing formulas of weather derivatives based on different temperature indices. These contracts may help many companies to hedge their risk. But most firms have to deal with a temperature risk in their operations, which is not located in one of the cities on which temperature futures are traded (for instance, an energy producer is exposed to temperature at various sites simultaneously). In this chapter we analyse how to combine available future contracts traded at the market, so that the need of the company is reflected optimally. Here the optimality criterion will usually be to minimize the variance with a certain temperature index. Therefore we construct a synthetic future contract by using the contracts offered at the market, which is closest to the company's needs.
Our analysis mainly follows the ideas of Barth et al. [9] and Benth et al. [16]. The temperature dynamics follows Model C in Chapter 2 with an additional space variable.

4.1 Temperature Dynamics in Time and Space

In this section we introduce the temperature dynamics, which is a continuous spatial-temporal model. Before we define the model, we introduce the following notation. Let $C(\mathcal{A})$ denote the space of real-valued continuous functions on some Borel subset $\mathcal{A} \subset \mathbb{R}$. $L^2(\mathcal{A})$ is the space of square integrable functions (w.r.t. Lebesque measure) on $\mathcal{A}$. The space of continuous functions on $\mathbb{R}_+ \times \mathcal{A}$, which are continuously differentiable in the first variable, is denoted by $C^{1,0}(\mathbb{R}_+ \times \mathcal{A})$.
We define a compact domain in $\mathbb{R}^2$ with piecewise smooth boundary $\partial\mathcal{D}$ by $\mathcal{D}$. This domain is equipped with the Euclidean metric. We assume according to Chapter 2 that $M(t)$ is the mean temperature function, but with one difference. $M \notin C^1(\mathbb{R}_+)$, but $M \in C^{1,0}(\mathbb{R}_+ \times \mathcal{D})$ and describes the mean temperature in $\mathcal{D}$. We have seen in Chapter 2 that the mean reverting parameter θ is quite stable over time, so we choose the parameter to be time-independent. Before we can write a closed form of our time-spatio temperature model we have to add the random fluctuation by a Gaussian random field, which is a reasonable model for many natural phenomena.

Definition 4.1 (Random field) *Let a probability space, $(\Omega, \mathcal{F}, \mathbb{P})$, and a parameter set, T be given. A random field is then a finite or real valued function $X(t, \omega)$ which for every fixed $t \in T$ is a measurable function of $\omega \in \Omega$.*

In our framework we consider the 2-dimensional Euclidean space, $T = \mathbb{R}^2$. For a fixed $\omega \in \Omega$ the function $X(t, \omega)$ is a non-random function of t. The parameter t is called the coordinate or position, which leads to a simple intuitive interpretation:
A random field $X(t, \omega)$ on $\mathbb{R}^2$ is a function whose values are random variables for any $t \in \mathbb{R}^2$. Note that a one-dimensional random field is usually called a stochastic process.

Definition 4.2 (Gassian Random Fields) *A Gaussian random field is a random field where all finite-dimensional distributions, $F_{t_1,\ldots,t_k}$ are multivariate normal distributions for any choice of k and $\{t_1, \ldots, t_k\}$.*

For further readings about random fields we refer to Abrahamsen [1]. We will formulate the Gaussian random field within the framework of Hilbert space valued random processes. The explicit definition of the Gaussian measure and Gaussian processes on Hilbert spaces can be found in Chapter 2 and 3 of Da Prato and Zabcyk [27].
Let $q \in C(\mathcal{D} \times \mathcal{D})$ denote a symmetric, strictly positive definite function. By this function the covariance in the space of the random field will be modeled. Moreover it can be viewed as the integral kernel of an operator acting on $L^2(\mathcal{D})$. Recalling our assumptions on $\mathcal{D}$ and q, it follows that Q is a symmetric Hilbert-Schmidt operator on $L^2(\mathcal{D})$ with a strictly positive spectrum. Furthermore Q admits the Mercer expansion

$$Q = \sum_{i=1}^{\infty} \lambda_i e_1 \otimes e_i. \tag{4.1}$$

Here $\{\lambda_i\}_{i\in\mathbb{N}}$ is the sequence of eigenvalues of Q and $\{e_i\}_{i\in\mathbb{N}}$ is the associated set of eigenfunctions ($Qe_i = \lambda_i e_i$ for $i \in \mathbb{N}$). Moreover the functions $\{e_i\}_{i\in\mathbb{N}}$ should form an orthonormal basis of $L^2(\mathcal{D})$ and they should be continuous on $\mathcal{D}$. Next we remark that the expansion (4.1) converges in the operator norm, following from the Mercer theorem, which can be found in Riesz and Sz.-Nagy [48], see also Werner [50] for an introduction about Hilbert-Schmidt operator and integral kernels.
Let us consider a filtrated complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$. The filtration $\{\mathcal{F}_t\}_{t\in\mathbb{R}_+}$ is right-continuous and $\mathcal{F}_0$ contains all $\mathbb{P} - null$ sets. To retain generality we suppose that $\mathcal{F} = \mathcal{F}_\infty = \sigma(\mathcal{F}_t, t \in \mathbb{R}_+)$. Now let $\{B(t)\}_{t\in\mathbb{R}}$ be a $L^2(\mathcal{D})-$valued $Q-$Brownian motion with respect to the filtration $\{\mathcal{F}_t\}$ defined on $(\Omega, \mathcal{F}, \mathbb{P})$. By these assumptions, $\{B(t)\}_{t\in\mathbb{R}}$ has the following expansion in $L^2(\mathcal{D})$:

$$B(t) = \sum_{i=1}^{\infty} \sqrt{\lambda_i}\beta_i(t)e_i. \tag{4.2}$$

Here $\{\beta_i\}_{i\in\mathbb{N}}$ is a sequence of one-dimensional standard Brownian motions given by

$$\beta_i(t) = \frac{1}{\sqrt{\lambda_i}} < B(t), e_i >_{L^2(\mathcal{D})}, \quad i \in \mathbb{N}. \tag{4.3}$$

$< .,. >$ denotes the inner product in $L^2(\mathcal{D})$. $B(t,x)$ is a centered Gaussian random field with covariance given by

$$\text{Cov}(B(t,x), B(s,y)) = \min(s,t)q(x,y), \quad s,t \in \mathbb{R}_+, x,y \in \mathcal{D}. \tag{4.4}$$

Now we can give an explicit formula of our spatio-temporal temperature model. The temperature $T(t,x)$ at time $t \in \mathbb{R}$ and at the point $x \in \mathbb{R}$ is given by

$$dT(t,x) = \left[(\theta(x)(M(t,x) - T(t,x)) + \frac{dM(t,x)}{dt} \right] dt + \sigma(t,x)dB(t,x). \tag{4.5}$$

Here $\sigma(t,x)$ describes the time-space volatility of the temperature. The solution of this model is

$$T(t,x) = (T(0,x) - M(0,x))e^{-\theta(x)t} + M(t,x) + \int_0^t e^{-\theta(x)(t-u)}\sigma(u,x)dB(u,x). \tag{4.6}$$

4.2 Optimal Portfolio of Future Contracts

In the last chapter we presented the pricing method of the different future contracts. We know that the underlying index is only measured in a specific city. Now we expand these models and consider that $\mathcal{D}$ is a geographical area, being a country or a province. In this considered area $\mathcal{D}$, we suppose that the different futures are traded for n cities located in the area at the coordinates $x_1, x_2, \ldots, x_n$.

4.2.1 General Temperature Index

We consider now a company which is exposed to temperature risk in an area $\mathcal{A} \subset \mathcal{D}$ (with $\mathcal{A}$ Borel). This company wants to hedge its risk using futures based on indices that are offered on the market. Let $I(\tau_1, \tau_2, .)$ denote either a HDD, a CDD or a CAT index. We make the assumption that the company is exposed to a temperature risk of the form

$$\int_{\mathcal{A}} I(\tau_1, \tau_2, y)\mu(dy), \tag{4.7}$$

where μ denotes a measure on $(\mathcal{A}, \mathcal{B}(\mathcal{A}))$. If we want for example the average of $I(\tau_1, \tau_2, y)$ over $\mathcal{A}$, we can use the Lebesgue measure normalized by the total mass. The measure could also be a point mass in specific locations $y_1, \ldots, y_n$, describing an exposure to the $I(\tau_1, \tau_2, y)$ in these locations. We only make the assumption that $I(\tau_1, \tau_2, .)$ is integrable w.r.t. μ on $\mathcal{A}$. So we consider a minimal-variance hedging problem. The aim of the agent is to minimize the L^2-distance between the desired future, which reflects its needs best, defined in (4.7) and the futures, which are available on the market, given he enters the market at time $t \leq \tau_1$. To create a portfolio of temperature futures, the agent can combine any available futures in the locations $x_1, \ldots, x_n$. We assume that every location i has a specific measurement period $[\tau_1^i, \tau_2^i]$ with $\tau_1 \leq \tau_1^i$. The investor has to minimize the spatial and temporal risk. On the one hand there may not exist indices with the desired measurement period and on the other hand there exist no futures at the location of risk $y \in \mathcal{A}$. Therefore we have to find an optimal portfolio which minimizes the spatial and temporal risk and covers the desired non-traded future optimally. We add to our considerations that it is possible that the investor wants to have a CDD-future, but there are only CAT or HDD futures available.

Let us denote by $\mathrm{a}(t) := (a_1(t), \dots, a_n(t))$ the number of contracts invested in each of the temperature futures, which is available on the market at time $t \in \mathbb{R}_+$. We assume that $t \mapsto \mathrm{a}(t)$ is an $\{\mathcal{F}_t\}_{t\in\mathbb{R}_+}-$adapted stochastic process. This is a reasonable assumption, because the investment decision $\mathrm{a}(t)$ can only depend on the available market information up to this time. The residual risk measured in terms of the variance of the unhedgeable part of (4.7) is defined for a given strategy $\mathrm{a}(t)$ as

$$R(t, \mathrm{a}(t)) = \mathbb{E}\left[\left[\int_{\mathcal{A}} I(\tau_1, \tau_2, y) d\mu(y) - \sum_{i=1}^{n} a_i(t) I(\tau_1^i, \tau_2^i, x_i)\right]^2 |\mathcal{F}_t\right]. \tag{4.8}$$

The aim of the agent is to find an optimal hedging strategy $\hat{\mathrm{a}}(t)$ which minimizes $R(t, \mathrm{a}(t))$. In other words we want to solve the minimization problem

$$\hat{R}(t) = \min_{\mathrm{a}(t)} R(t, \mathrm{a(t)}). \tag{4.9}$$

The optimal solution of (4.9) is denoted by $\hat{\mathrm{a}}(t)$ and defines a synthetic future contract. This contract minimizes the distance, in the sense of the variance, between the desired future and the portfolio of existing futures on the market. The synthetic future contract is given by $\sum_{i=1}^{n} a_i(t) I(\tau_1^i, \tau_2^i, x_i)$, which are the positions in tradeable temperature futures taken at time t.

By rewriting $R(t, \mathrm{a}(t))$ we obtain

$$\begin{aligned} R(t, \mathrm{a}(t)) = &\ \mathbb{E}\left[\left[\int_{\mathcal{A}} I(\tau_1, \tau_2, y)\mu(dy)\right]^2 |\mathcal{F}_t\right] \\ &- 2\sum_{i=1}^{n} a_i(t)\mathbb{E}\left[\int_{\mathcal{A}} I(\tau_1, \tau_2, y)\mu(dy) I(\tau_1^i, \tau_2^i, x_i)|\mathcal{F}_t\right] \\ &+ 2 \sum_{i=1,j=1,i<j}^{n} a_i(t)a_j(t)\mathbb{E}\left[I(\tau_1^i, \tau_2^i, x_i) I(\tau_1^j, \tau_2^j, x_j)|\mathcal{F}_t\right] \\ &+ \sum_{i=1}^{n} a_i^2(t)\mathbb{E}\left[I(\tau_1^i, \tau_2^i, x_i)^2|\mathcal{F}_t\right]. \end{aligned}$$

Let us introduce the following notation

$$w(t, x, y) = \left[I(\tau_1^x, \tau_2^x, x) I(\tau_1^y, \tau_2^y, y)|\mathcal{F}_t\right], \tag{4.10}$$

for $(x, y) \in \mathcal{D}^2$. It is important that the measurement periods are labeled differently, because they may not coincide. Let us define $\mathrm{b}(t) := (b_1(t), \dots, b_n(t))^T$ with

$$b_i(t) = \frac{\int_{\mathcal{A}} w(t, x_i, y)\mu(dy)}{w(t, x_i, x_i)}, \tag{4.11}$$

and the matrix $A(t)$ given as

$$A(t) := \begin{pmatrix} 1 & A_{12}(t) & \dots & A_{1n}(t) \\ A_{21}(t) & 1 & \dots & A_{2n}(t) \\ \vdots & & \ddots & \vdots \\ A_{n1}(t) & A_{n2}(t) & \dots & 1 \end{pmatrix} \tag{4.12}$$

with

$$A_{ij}(t) = \frac{w(t, x_i, x_j)}{w(t, x_i, x_i)}. \tag{4.13}$$

Theorem 4.1 *The minimizer* $\hat{a}(t) := (\hat{a}_1(t), \dots, \hat{a}_n(t))^T$ *of* $R(t, a(t))$ *is given as the solution of the linear system:*

$$A(t)a(t) = b(t).$$

Proof:
The first order condition $\frac{\partial R}{\partial a_j} = 0$, $j = 1, \dots, n$ for the optimum yields

$$\sum_{i=1}^{n} a_i(t)\mathbb{E}\left[I(\tau_1^i, \tau_2^i, x_i)I(\tau_1^j, \tau_2^j, x_j)|\mathcal{F}_t\right] = \int_{\mathcal{A}} \mathbb{E}\left[I(\tau_1, \tau_2, y)I(\tau_1^i, \tau_2^i, x_i)|\mathcal{F}_t\right] d\mu(y).$$

The change of order of expectation and differentiation is valid in this situation due to the normality of the temperature dynamics. Hence, we find that $\hat{\mathrm{a}}(t)$ is the solution of $A(t)\mathrm{a}(t) = \mathrm{b}(t)$. □

It hence follows that the solution $\hat{\mathrm{a}}(t)$ is a function of the temperature at time t in the location $x_1, \dots, x_n$, where futures are traded. The optimal values of the weights $\hat{a}_1(t), \dots, \hat{a}_n(t)$ define the synthetic future price

$$F_{\mathcal{A}}^I(t, \tau_1, \tau_2) = \sum_{i=1}^{n} \hat{a}_i(t)F_i^I(t, \tau_1^i, \tau_2^i), \tag{4.14}$$

where $F_i^I(t, \tau_1^i, \tau_2^i)$ denotes the future price at time t for the traded contract written on the Index $I(\tau_1^i, \tau_2^i, x_i)$. In Chapter 3 we have already seen how to compute explicit future prices for the different underlying indices. In the following sections we calculate explicitly the optimal weights for the different underlying indices.

4.2.2 CAT future

In this section we assume that all traded contracts are based on the CAT index, i.e. $I = CAT$. By using the explicit expression of $T(t, x)$ (4.6), we obtain the following expression for

$CAT(\tau_1, \tau_2, x)$:

$$
\begin{aligned}
CAT(\tau_1^x, \tau_2^x, x) &= \int_{\tau_1^x}^{\tau_2^x} T(\tau, x) d\tau \\
&= \int_{\tau_1^x}^{\tau_2^x} M(\tau, x) d\tau + (T(t,x) - M(t,x))\tilde{\theta}(t, \tau_1^x, \tau_2^x, x) \\
&+ \int_t^{\tau_2^x} \sigma(u,x)\tilde{\theta}(u, \tau_1^x, \tau_2^x, x) dB(u,x)
\end{aligned}
$$

with

$$\tilde{\theta}(u, \tau_1^x, \tau_2^x, x) = \frac{1}{\theta}\left[\exp(-\theta(x)(\tau_1^x - \min(u, \tau_1^x))) - \exp(-\theta(x)(\tau_2^x - u))\right].$$

The conditioned expectation of $CAT(\tau_1^x, \tau_2^x, x)$ is a Gaussian random field with mean

$$m(t,x) := \mathbb{E}\left[CAT(\tau_1^x, \tau_2^x, x)|\mathcal{F}_t\right] = \int_{\tau_1^x}^{\tau_2^x} M(\tau, x) d\tau + (T(t,x) - M(t,x))\tilde{\theta}(t, \tau_1^x, \tau_2^x, x)$$

and covariance

$$
\begin{aligned}
c(t,x,y) :&= \text{Cov}\left[CAT(\tau_1^x, \tau_2^x, x), CAT(\tau_1^y, \tau_2^y, y)|\mathcal{F}_t\right] \\
&= q(x,y) \int_t^{\min(\tau_2^x, \tau_2^y)} \sigma(u,x)\sigma(u,y)\tilde{\theta}(u, \tau_1^x, \tau_2^x, x)\tilde{\theta}(u, \tau_1^y, \tau_2^y, y) du.
\end{aligned}
$$

Form Theorem 4.1 we obtain for the non-diagonal elements of the matrix $A(t)$ in this special case

$$A_{ij}(t) = \frac{c(t, x_i, x_j) + m(t, x_i)m(t, x_j)}{c(t, x_i, x_i) + m^2(t, x_i)}$$

for $i, j = 1, \ldots, n$ and $i \neq j$. Moreover the coordinates of the vector $\mathrm{b}(t)$ are given as

$$b_i(t) = \frac{\int_A c(t, y, x_i)\mu(dy) + \int_A m(t,y)\mu(dy)m(t, x_i)}{c(t, x_i, x_i) + m^2(t, x_i)},$$

for $i = 1, \ldots, n$. If we solve the the linear system $A(t)\mathrm{a}(t) = \mathrm{b}(t)$ with these specifications of $A(t)$ and $\mathrm{b}(t)$, we obtain the optimal weights for constructing a synthetic future contract based on CAT-contracts at different locations.

4.2.3 CDD/HDD futures

In the following the focus will be on CDD futures and we make the assumption that μ is a point mass at location $y \in \mathcal{D}$. This restriction is only for the avoidance of complex notation. The calculations for the optimal weights of HDD-futures can be done in an analogous way. To get $A(t)$ and $\mathrm{b}(t)$, we have to calculate the conditional expectation

$$
\begin{aligned}
w(t,x,y) :&= \mathbb{E}\left[\int_{\tau_1^x}^{\tau_2^x} \max(T(\tau, x) - 18, 0) d\tau \int_{\tau_1^y}^{\tau_2^y} \max(T(\tau, y) - 18, 0) d\tau |\mathcal{F}_t\right] \\
&= \int_{\tau_1^x}^{\tau_2^x} \int_{\tau_1^y}^{\tau_2^y} \mathbb{E}\left[\max(T(\tau, x) - 18, 0) \max(T(\tau', y) - 18, 0)|\mathcal{F}_t\right] d\tau d\tau'.
\end{aligned}
$$

at locations $(x, y) \in \mathcal{D}$, here for the combinations of $x_1, \ldots, x_n$ and y. Using the explicit form of $T(t, x)$ we obtain

$$w(t,x,y) = \int_{\tau_1}^{\tau_2}\int_{\tau_1}^{\tau_2} (1-\rho(\tau,\tau',x_i,x_j)^2)^{\frac{3}{2}} \times \sqrt{\text{Var}\left[T(\tau,x_i)|\mathcal{F}_t\right]\text{Var}\left[T(\tau',x_j)|\mathcal{F}_t\right]}\, g(\tau,\tau',x_i,x_j)d\tau d\tau',$$

where

$$g(\tau,\tau',x_i,x_j) = \int_{\hat{x}_j}^{\infty} (u-18)\text{exp}(-\frac{1}{2}(1-\rho(\tau,\tau',x_i,x_j)^2)u^2) \times \int_{\hat{x}_i}^{\infty} (v-18)\text{exp}(-\frac{1}{2}(v-\rho(\tau,\tau',x_i,x_j)u)^2)dvdu$$

with integration limits

$$\hat{x}_i = \frac{18-\mathbb{E}\left[T(\tau,x_i)|\mathcal{F}_t\right]}{\sqrt{(1-\rho(\tau,\tau',x_i,x_j)^2)\text{Var}\left[T(\tau,x_i)|\mathcal{F}_t\right]}},$$

$$\hat{x}_j = \frac{18-\mathbb{E}\left[T(\tau',x_j)|\mathcal{F}_t\right]}{\sqrt{(1-\rho(\tau,\tau',x_i,x_j)^2)\text{Var}\left[T(\tau',x_i)|\mathcal{F}_t\right]}}$$

and

$$\rho(\tau,\tau',x_i,x_j) = \frac{\text{Cov}\left[T(\tau,x_i),T(\tau',x_j)|\mathcal{F}_t\right]}{\sqrt{\text{Var}\left[T(\tau,x_i)|\mathcal{F}_t\right]\text{Var}\left[T(\tau',x_j)|\mathcal{F}_t\right]}}.$$

The expectation and the covariance are given by

$$\mathbb{E}\left[T(\tau,x)|\mathcal{F}_t\right] = M(t,x) + (T(t,x)-M(t,x)) * \text{exp}(-\theta(x)(\tau-t))$$

and

$$Cov\left[T(\tau,x),T(\tau',y)|\mathcal{F}_t\right] = q(x,y)\int_t^{\min(\tau,\tau')} \sigma(u,x)\sigma(u,y)\text{exp}(-\theta(x)(\tau-u)) \text{exp}(-\theta(y)(\tau'-u))du.$$

Now we have calculated the entries of $A(t)$ and $\text{b}(t)$.

4.2.4 CAT and CDD futures

In this section we consider a mixture of CAT and CDD futures. We suppose that at some locations only CAT futures are traded and at the others only CDD futures. It is also possible that at some locations both are traded. More specifically, we consider $x_1, \ldots, x_m$, which are locations where only CAT futures available for trade, and $x_{m+1}, \ldots, x_n$ are these locations with only CDD futures. If $x_i = x_j$ for $i \leq m$ and $j > m+1$, then there exists the possibility of trade CDD as well as CAT futures. For these assumptions we have to modify $A(t)$, which is given by

$$\begin{pmatrix} A^{CAT}_{m\times m}(t) & A^{CAT,CDD}_{m\times(n-m)}(t) \\ A^{CDD,CAT}_{(n-m)\times m}(t) & A^{CDD}_{(n-m)\times(n-m)}(t) \end{pmatrix}.$$

Here $A^{CAT}_{m\times m}$ denotes the matrix A derived in the Section 4.2.2 for m locations $x_1,\dots,x_m$. Analogously $A^{CDD}_{(n-m)\times(n-m)}$ denotes the matrix derived in Section 4.2.3 for CDD futures in this case for $n-m$ locations $x_{m+1},\dots,x_n$. Let us now consider the matrices $A^{CAT,CDD}$ and $A^{CDD,CAT}$.

$A^{CAT,CDD}_{m\times(n-m)}(t)$ denotes the matrix with elements $w(t,x_i,x_j)/w(t,x_i,x_i)$ for $i=1,\dots,m$ and $j=m+1,\dots,n$ with

$$w(t,x_i,x_j)=\mathbb{E}\left[CAT(\tau_1^i,\tau_2^i,x_i)CDD(\tau_1^j,\tau_2^j,x_j)|\mathcal{F}_t\right] \tag{4.15}$$

and

$$w(t,x_i,x_i)=\mathbb{E}\left[CAT^2(\tau_1^i,\tau_2,x_i)|\mathcal{F}_t\right].$$

The derivation of $w(t,x_i,x_i)$ we have already seen in Section 4.2.2. The elements of the matrix $A^{CDD,CAT}$ are $w(t,x_i,x_i)/w(t,x_i,x_j)$ for $i=m+1,\dots,n$ and $j=1,\dots,m$ with

$$w(t,x_i,x_j)=\mathbb{E}\left[CAT(\tau_1^j,\tau_2^j,x_j)CDD(\tau_1^i,\tau_2^i,x_i)|\mathcal{F}_t\right]$$

and

$$w(t,x_i,x_i)=\mathbb{E}\left[CDD^2(\tau_1^i,\tau_2^i,x_i)|\mathcal{F}_t\right].$$

We have already done the computation of $w(t,x_i,x_i)$ in Section 4.2.3. The derivation of $w(t,x_i,x_j)$ is similar to (4.15). Now we calculate $w(t,x_i,x_j)$ as defined in (4.15) with $1=1,\dots,m$ and $j=m+1,\dots,n$. By using (4.2.2) and the measurability of $T(t,x_i)$ we obtain

$$\begin{aligned} w(t,x_i,x_j) &= \left(\int_{\tau_1^i}^{\tau_2^i} M(\tau,x)d\tau + (T(t,x_i)-M(t,x_i))\tilde{\theta}(t,\tau_1^i,\tau_2^i,x_i)\right) \\ &\quad * \mathbb{E}\left[CDD(\tau_1^j,\tau_2^j,x_j)|\mathcal{F}_t\right] \\ &\quad + \mathbb{E}\left[\int_t^{\tau_2^i} \sigma(u,x_i)\tilde{\theta}(u,\tau_1^i,\tau_2^i,x_i)dB(u,x_i)CDD(\tau_1^j,\tau_2^j,x_j)|\mathcal{F}_t\right]. \end{aligned}$$

Calculating the first conditional expectation we obtain

$$\begin{aligned} \mathbb{E}\left[CDD(\tau_1^j,\tau_2^j,x_j)|\mathcal{F}_t\right] &= \int_{\tau_1^j}^{\tau_2^j}(m_j(t,\tau)-18)\Phi\left(\frac{m_j(t,\tau)-18}{\Sigma_j(t,\tau)}\right) \\ &\quad + \Sigma_j(t,\tau)\phi\left(\frac{m_j(t,\tau)-18}{\Sigma_j(t,\tau)}\right)d\tau, \end{aligned}$$

with

$$m_j(t,\tau)=M(\tau,x_j)+(T(t,x_j)-M(t,x_j))*\exp(-\theta(x_j)(\tau-t))$$

and

$$\Sigma_j^2(t,\tau) = \int_t^\tau \sigma^2(s,x_j) * \exp(-2\theta(x_j)(\tau-s))ds.$$

Φ denotes the standard normal distribution function and ϕ its density.
Before we calculate $\mathbb{E}\left[\int_t^{\tau_2^i} \sigma(u,x_i)\tilde{\theta}(u,\tau_1^i,\tau_2^i,x_i)dB(u,x_i)CDD(\tau_1^j,\tau_2^j,x_j)|\mathcal{F}_t\right]$, let us define the following expression:

$$\rho_{ij}(t,\tau) = \frac{\int_t^{\min(\tau_2^i,\tau)} \sigma(s,x_i)\sigma(s,x_j)\tilde{\theta}(s,\tau_1^i,\tau_2^i,x_i) * \exp(-\theta(x_j)(\tau-s))ds}{\sqrt{\int_t^{\tau_2^i} \sigma^2(s,x_i)\tilde{\theta}^2(s,\tau_1^i,\tau_2^i,x_i)ds}\sqrt{\int_t^\tau \sigma^2(s,x_j) * \exp(-2\theta(x_j)(\tau-s))ds}}.$$

This term denotes the correlation between the two normally distributed random variables

$$\int_t^{\tau_2^i} \sigma(s,x_i)\tilde{\theta}^2(s,\tau_1^i,\tau_2^i,x_i)dB(s,x_i)$$

and

$$\int_t^\tau \sigma(s,x_j) * \exp(\theta(x_j)(\tau-s))dW(s,x_j).$$

Furthermore, Y is a standard normally distributed random variable and we obtain for the conditional expectation

$$\begin{aligned}
&\mathbb{E}\left[\int_t^{\tau_2^i} \sigma(u,x_i)\tilde{\theta}(u,\tau_1^i,\tau_2^i,x_i)dB(u,x_i)CDD(\tau_1^j,\tau_2^j,x_j)|\mathcal{F}_t\right] \\
&= \int_{\tau_1^j}^{\tau_2^j} \rho_{ij}(t,\tau)\tilde{\Sigma}_i(t)\mathbb{E}\left[Y\max(m_j(t,\tau)-c+\Sigma_j(t,\tau)Y,0)|\mathcal{F}_t\right]d\tau \\
&= -2\tilde{\Sigma}_i(t)\int_{\tau_1^j}^{\tau_2^j} \rho_{ij}(t,\tau)\phi\left(\frac{m_j(t,\tau)-18}{\Sigma_j(t,\tau)}\right)\left(\frac{1}{2}+\frac{m_j(t,\tau)-18}{\Sigma_j(t,\tau)}\right)d\tau
\end{aligned}$$

with

$$\tilde{\Sigma}_i^2(t) = \int_t^{\tau_2^i} \sigma^2(s,x_i)\tilde{\theta}^2(s,\tau_1^i,\tau_2^i,x_i)ds.$$

Now we have an explicit expression for $w(t,x_i,x_j)$ and a description of the matrix A.
The calculation of the vector $\mathsf{b}(t)$ is much easier because it can be represented by

$$\mathsf{b}(t) = (\mathsf{b}_m^{CAT}(t)^T, \mathsf{b}_{n-m}^{CDD}(t)^T)^T.$$

Here $\mathsf{b}_m^{CAT}(t)^T$ denotes the b vector we have already computed for the CAT futures in Section 4.2.2 with m locations where CAT futures are traded. The vector $\mathsf{b}_{n-m}^{CDD}(t)$ can be derived as in Section 4.2.3.

Chapter 5

Optimal Cross Hedging

After concentrating on modeling and pricing of temperature derivatives and considering spatial temperature models, we now present in the following Chapters 5 and 6 more general methods of pricing and hedging derivatives, which are based on non-tradeable underlyings. In Chapter 5 we will use standard stochastic control techniques to obtain an optimal strategy. Contrary to the approach in Chapter 5, we describe the price and the optimal strategy in Chapter 6 in terms of a forward-backward stochastic differential equation. The main idea of those models is closely related. The model in Chapter 6 is an extension of the model in Chapter 5 and permits that the asset is based on more than one underlying.
The main idea of the models is to circumvent the problem of non-tradability of the underlyings by considering tradeable assets (one or more) which are correlated to the non-tradeable underlying of the derivative. It is impossible to create a perfect hedge of the derivative by investing in correlated assets, because an unhedgeable risk remains.

5.1 Introduction

Before we start describing the model we want do motivate it. This general model can help companies who are involved in hedging weather risk or other risks which are based on non-tradeable underlyings. We want to focus on the idea of cross hedging, which is based on the negative correlation of risk exposures of different agents on the market. This idea of alternative risk transfer has been already considered for the well studied sea surface temperature anomaly of the South Pacific in Chaumont et al. [25].
Let us give a short introduction to the model which we consider in this chapter and those optimal cross hedging strategies we will estimate. The main process of the model is an non-tradeable external risk process X, whose uncertainty is modeled by a Brownian Motion W with volatility σ. Next we suppose that a derivative F, sold by an agent, depends on the external risk process. In other words, the income $F(X_T)$ is random and depends on the position X_T at the maturity T. The second process P is a security on a financial market, which is a possibility for hedging the derivative. Note that X and P are correlated. We will assume that the preferences of the investor are determined by an exponential utility function (see Section

6.1.1). In this section we mainly follow Ankirchner et al. [5] to which we also refer for more details.
Before describing the model explicitly, we want to mention that a better understanding of some mathematical tools will be helpful. Instead of giving an introduction we want to give some further readings. An introduction to Markov control processes, the Hamilton-Jacobi-Bellman equation, the verification theorem and the famous Feynman-Kac formula can be found in Øksendal [45]. Moreover, detailed information about the Markov processes and optimal control can be found in Fleming and Soner [30].

5.2 The model

Let $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$ a filtered probability space and W a Brownian motion. We consider derivatives, whose pay-off is based on non-tradeable external risks. We assume that the external risk can be modeled as a diffusion process

$$dX_t = b(t, X_t)dt + \sigma(t, X_t)dW_t \quad \text{and} \quad X_0 = x_0. \tag{5.1}$$

Next let F be a measurable and real-valued function and $T > 0$, then $F(X_T)$ denotes a derivative with maturity T. An essential assumption for this model is that the financial market consists of one-risky asset with zero interest rate and one risky asset whose price is correlated to the process X. We assume that the price dynamics of the risky asset satisfies

$$dP_t = P_t \left(g(t, X_t)dt + \beta_1 dW_t + \beta_2 dB_t\right) \quad \text{and} \quad P_0 > 0. \tag{5.2}$$

where $g : [0, T] \times \mathbb{R} \to \mathbb{R}$ is measurable, $\beta_1 \in \mathbb{R}$, $\beta \neq 0$ and B a Brownian motion, which is independent of W. Next we denote the volatility of P by $\beta = \sqrt{\beta_1^2 + \beta_2^2}$. We see that the Brownian motion W drives the risk process X and the price process P. The risk process influences the drift part of the price process. For the sake of completeness we recall the definition of an investment strategy, which is any $(\mathcal{F}_t)_{t \geq 0}$-predictable process π with $\pi_0 = 0$ and such that the stochastic integral process of π with respect to

$$\int_0^{\cdot} \left(g(t, X_t)dt + \beta_1 dW_t + \beta_2 dB_t\right)$$

exists on $[0, T]$. We can interpret π as the value of the portfolio fraction invested in the risky asset. This means that $\pi_t = P_t\theta_t$, where θ_t denotes the number of assets, shares in the portfolio at time t. Let us now define the wealth process which is given by

$$dV_t^\pi = \pi_t \left(g(t, X_t)dt + \beta_1 dW_t + \beta_2 dB_t\right) \quad \text{and} \quad V_0^\pi = v_0, \tag{5.3}$$

where v_0 denotes the initial wealth. We assume that the preferences of an investor are determined by an exponential utility function with risk-averse parameter η and we want to maximize the expected value of his wealth at time T. The optimal strategy π^{opt} is determined by

$$\mathbb{E}\left[U\left(V_T^{\pi^{opt}} + F(X_T)\right)\right] = \sup_\pi \mathbb{E}\left[U(V_T^\pi + F(X_T))\right], \tag{5.4}$$

if the investor has a derivative $F(X_T)$ in his portfolio. Notice that the supremum is taken over all admissible strategies. A strategy π is said to be a-admissible if for all $t \in [0,T]$ we have $V_t^\pi \geq -a$, almost surely for $a \geq 0$. We call π admissible if there exists an $a \geq 0$ such that π is a-admissible.

Definition 5.1 *A strategy π is called quasi-admissible with respect to U if there exists a sequence (π_n) of admissible strategies such that $U(V_t^{\pi_n})$ converges to $U(V_T^\pi)$ in L^1 as $n \to \infty$.*

We now want to find the indifference price p (see Section 6.1.1) of the derivative such that

$$\sup_\pi \mathbb{E}\left[U\left(V_t^\pi F(X_T) - p\right)\right] = \sup_\pi \mathbb{E}\left[U\left(V_T^\pi\right)\right] \tag{5.5}$$

is fulfilled. We denote by $\pi^\#$ the strategy which optimizes the right-hand side of the equation and by π^* the strategy optimizing the left-hand side. In the following section we solve the control problem with the two processes X and V^π.

5.3 The control problem

We modify X and V^π and define for all $(t,x,v) \in [0,T] \times \mathbb{R}^2$:

$$V_r^{\pi,t,x,v} = v + \int_t^r \pi_s \left(g(s, X_s^{t,x})ds + \beta_1 dW_s + \beta_2 dB_s\right) \quad \forall r \geq t; \tag{5.6}$$

$$X_r^{t,x} = x + \int_t^r b(s, X_s^{t,x})ds + \int_t^r \sigma(s, X_s^{t,x})dW_s \quad \forall r \geq t. \tag{5.7}$$

$$\tag{5.8}$$

To get unique strong solutions of those SDE we have to assume that there exists $\mu \in \mathbb{R}$ and $K \geq 0$ so that

1. $\forall t \in [0,T], \forall (x,x') \in \mathbb{R}^2$:

$$(x - x')(b(t,x) - b(t,x')) \leq \mu |x - x'|^2,$$
$$|g(t,x) - g(t,x')| + |\sigma(t,x) - \sigma(t,x')| \leq K|x - x'|$$

2. $\forall t \in [0,T], \forall x \in \mathbb{R}, |g(t,x)| + |b(t,x)| + |\sigma(t,x)| \leq K(1 + |x|)$.

3. $\forall t \in [0,T]$, the function $x \mapsto b(t,x)$ is continuous.

Let us now define a function $K^G : [0,T] \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ by

$$K^G(t,x,v) = \sup_\pi \mathbb{E}\left[U\left(V_T^{\pi,t,x,v} + G(X_T^{t,x})\right)\right],$$

with G a given function. In (5.4) G is equal to 0 or $F(.) - p$. To solve this control problem we use the Hamilton-Jacobi-Bellman equation which leads to the following equation. For all $(t, x, v) \in [0, T) \times \mathbb{R} \times \mathbb{R}$

$$\frac{\partial h}{\partial t} + \frac{\sigma^2(t,x)}{2}\frac{\partial^2 h}{\partial x^2} + b(t,x)\frac{\partial h}{\partial x} + \sup_{\pi}\left\{\frac{\pi^2\beta^2}{2}\frac{\partial^2 h}{\partial v^2} + \pi\beta_1\sigma(t,x)\frac{\partial^2 h}{\partial v \partial x} + \pi g(t,x)\frac{\partial h}{\partial v}\right\} = 0. \tag{5.9}$$

h satisfies

$$h(T, x, v) = U(v + G(x)).$$

Note that the supremum from the equation above is finite if $\frac{\partial^2 h}{\partial v^2} < 0$. We are searching for an explicit solution h for (5.9) we will show that $h = K^G$, according the variation theorem. Recall that we assume that the preferences of the investor are determined by an exponential utility function. That is why we can write the utility of the whole portfolio as a product of the utility of the derivative and the utility of the investment. This multiplicative property motivates us to search h such that

$$h(t, x, v) = U(v)\text{exp}(-\eta\Gamma(t, x)),$$

where $\Gamma(T, x) = G(x)$.
Then the condition $\frac{\partial^2 h}{\partial v^2} < 0$ holds and

$$\frac{\partial \Gamma}{\partial t} + \frac{\sigma^2(t,x)}{2}\frac{\partial^2 \Gamma}{\partial x^2} + b(t,x)\frac{\partial \Gamma}{\partial x} - \frac{\sigma^2(t,x)\eta}{2}\left(\frac{\partial \Gamma}{\partial x}\right)^2 - \inf_{\pi}\left\{\frac{\pi^2\beta^2\eta}{2} + \pi\beta_1\sigma(t,x)\eta\frac{\partial \Gamma}{\partial x} - \pi g(t,x)\right\} = 0$$

By differentiating and setting equal to zero, we obtain the optimal strategy which is given by

$$\pi^* = \pi^*(t, x) = \frac{1}{\beta^2}\left(\frac{g(t,x)}{\eta} - \beta_1\sigma(t,x)\frac{\partial \Gamma}{\partial x}(t,x)\right), \tag{5.10}$$

and we obtain

$$\frac{\partial \Gamma}{\partial t} + \frac{\sigma^2(t,x)}{2}\frac{\partial^2 \Gamma}{\partial x^2} - \frac{\beta_2^2}{\beta^2}\frac{\sigma^2(t,x)\eta}{2}\left(\frac{\partial \Gamma}{\partial x}\right)^2 + \left(b(t,x) - \frac{\alpha\beta_1\sigma(t,x)}{\beta^2}\right)\frac{\partial \Gamma}{\partial x} + \frac{g(t,x)^2}{2\eta\beta^2} = 0. \tag{5.11}$$

Let us define

$$\varphi(t, x) = \text{exp}\left(-\frac{\eta\beta_2^2}{\beta^2}\Gamma(t, x)\right)$$

which is equivalent to

$$\Gamma(t,x) = \frac{\ln(\varphi(t,x))}{k}, \quad \text{with} \quad k = \frac{\beta_2^2}{\beta^2}\eta.$$

Then

$$\varphi(T,x) = \exp(-kG(x)),$$

and we obtain

$$\frac{\partial \varphi}{\partial t} + \frac{\sigma^2(t,x)}{2}\frac{\partial^2 \varphi}{\partial x^2} + \left(b(t,x) - \frac{\beta_1 g(t,x)\sigma(t,x)}{\beta^2}\right)\frac{\partial \varphi}{\partial x} - \frac{kg(t,x)^2}{2\beta^2\eta}\varphi = 0. \tag{5.12}$$

To solve this partial differential equation we use the Feynman-Kac formula. Then a solution of (5.12) is given by

$$\varphi^G(t,x) = \mathbb{E}\left[\exp\left(kG(Y_T^{t,x})\right)\exp\left(-\frac{k}{2\beta^2\eta}\int_t^T g(r,Y_r^{t,x})^2 dr\right)\right].$$

Here $Y^{t,x}$ is defined as follow

$$Y_r^{t,x} = x + \int_t^r \hat{b}(s,Y_s^{t,x})ds + \int_t^r \sigma(s,Y_s^{t,x})d\hat{W}_s, \tag{5.13}$$

with

$$\hat{b}(t,x) = b(t,x) - \frac{\beta_1 g(t,x)\sigma(t,x)}{\beta^2},$$

and $\hat{W}$ is an arbitrary one-dimensional Wiener process. Immediately we obtain the solution of (5.9) which is given by

$$\begin{aligned} h(t,x,v) &= U(v)\varphi^G(t,x)^{\frac{\eta}{k}} \qquad (5.14)\\ &= U(v)\left(\mathbb{E}\left[\exp\left(-kG(Y_T^{t,x})\right)\exp\left(-\frac{k}{2\beta^2\eta}\int_t^T g(r,Y_r^{t,x})^2\right)dr\right]\right)^{\frac{\eta}{k}} \end{aligned}$$

If for any t and x the stochastic integral $(\frac{\beta_1}{\beta^2}g(.,X^{t,x})W)$ satisfies the Novikov condition, we can define a new probability measure $\mathbb{Q}$ with density

$$\frac{d\mathbb{Q}}{d\mathbb{P}} = \exp\left(-\int_t^{T-t}\frac{\beta_1 g(s,X_s^{t,x})}{\beta^2}dW_s - \frac{1}{2}\int_t^{T-t}\frac{\beta_1^2 g^2(s,X_s^{t,x})}{\beta^4}ds\right). \tag{5.15}$$

The distribution of $Y^{t,x}$ under the measure $\mathbb{P}$ coincides with the distribution of $X^{t,x}$ under the measure $\mathbb{Q}$ and we can write φ^G in terms of $X^{x,t}$ and $\mathbb{Q}$ to obtain

$$h(t,x,v) = U(v)\left(\mathbb{E}^{\mathbb{Q}}\left[\exp\left(-kG(X_T^{t,x})\right)\exp\left(-\frac{k}{2\beta^2\eta}\int_t^T g(r,X_r^{t,x})^2 dr\right)\right]\right)^{\frac{\eta}{k}}.$$

Let us now derive the right hand side of (5.4). We set $G = 0$, and the value function $K^0(t,x,v)$ satisfies

$$K^0(t,x,v) = \sup_\pi \mathbb{E}\left[U(V_T^{\pi,t,x,v})\right] \tag{5.16}$$

and using the Hamilton-Jacobi-Bellman equation we obtain the solution

$$h^0(t,x,v) = U(v)\mathbb{E}\left[\exp\left(-\frac{k}{2\beta^2\eta}\int_t^T g(r,Y_r^{t,x})^2 dr\right)\right]^{\frac{\eta}{k}} = U(v)\varphi^0(t,x)^{\frac{\eta}{k}}. \tag{5.17}$$

Let us assume that the optimal strategy $\pi^\#$ is quasi-admissible and that the following property is satisfied.

Definition 5.2 (Regularity Property 1) *There exists an open set $\mathcal{V} \subseteq \mathbb{R}$ such that h^0 belongs to the class $C^{1,2}([0,T) \times \mathcal{V} \times \mathbb{R})$, and for all $x \in \mathcal{V}$ the process $X^{t,x}$ stays in $\mathcal{V}$.*

This guarantees us that h^0 is a classical solution of the HJB equation (5.9) on $[0,T] \times \mathcal{V} \times \mathbb{R}$. Let us apply the verification theorem and we obtain $h^0(t,x,v) = K^0(t,x,v)$ on $[0,T] \times \mathcal{V} \times \mathbb{R}$. It follows that the optimal strategy is given by

$$\pi^\#(t,x) = \frac{1}{\beta}\left(\frac{g(t,x)}{\eta} + \frac{\beta_1\sigma(t,x)}{k\varphi(t,x)}\frac{\partial\varphi^0}{\partial x}(t,x)\right). \tag{5.18}$$

If g is constant and equal to α we obtain

$$K^0(t,x,v) = K^0(t,v) = \exp\left(-\frac{\alpha^2}{2\beta^2}(T-t)\right)U(v)$$

and

$$\pi^\#(t,x) = \pi^\# = \frac{\alpha}{\eta\beta^2}.$$

Let us now consider the right hand side of (5.4). Here we have

$$K^F(t,x,v) = \sup_\pi \mathbb{E}\left[U(V_T^{\pi,t,x,v} + F(X_T^{t,x}) - p)\right]$$

and a solution of (5.9) is

$$h^F(t,x,v) = U(v)\varphi^F(t,x)^{\frac{\eta}{k}} \tag{5.19}$$

$$= U(v)\mathbb{E}\left[\exp(k*p)\exp\left(-kF(Y_T^{t,x})\right)\exp\left(-\frac{k}{2\beta^2\eta}\int_t^T g(r,Y_r^{t,x})^2 dr\right)\right]^{\frac{\eta}{k}}. \tag{5.20}$$

We assume that the optimal strategy is quasi-admissible and h^F is regular in the following sense.

Definition 5.3 (Regularity Property 2) *There exists an open set $\mathcal{U} \subseteq \mathcal{V}$ such that h^F belongs to the class $C^{1,2}([0,T) \times \mathcal{U} \times \mathbb{R})$ and for all $x \in \mathcal{U}$, the process $X^{t,x}$ stays in U.*

By using the verification theorem we get

$$h^F(t,x,v) = K^F(t,x,v)$$

if $v \in \mathbb{R}$ and $x \in \mathcal{U}$.

In Section 5.5 we give conditions for the optimal strategies under which the are quasi-admissible.

5.3.1 Estimation of the indifference price

Here we consider the indifference price of the asset at any time t which is given by

$$\sup_{\pi} \mathbb{E}\left[U(V_T^{\pi,t,x,v} + F(X_T^{t,x}) - p(t,x))\right] = \sup_{\pi} \mathbb{E}\left[U(V_T^{\pi,t,x,v})\right]. \tag{5.21}$$

By setting $h^0(t,x,v) = h^F(t,x,v)$ we obtain for the price

$$p(t,x) = -\frac{1}{k}\ln\left[\frac{\mathbb{E}\left[\exp\left(-kF(Y_T^{t,x})\right)\exp\left(-\frac{\beta_2^2}{2\beta^4}\int_t^T g(r,Y_r^{t,x})^2 dr\right)\right]}{\mathbb{E}\left[\exp\left(-\frac{\beta_2^2}{2\beta^4}\int_t^T g(r,Y_r^{t,x})^2 dr\right)\right]}\right]$$

where Y denotes the solution of (5.13). Let us introduce a new probability measure $\hat{\mathbb{P}}$, whose density is given by

$$\frac{d\hat{\mathbb{P}}}{d\mathbb{P}} = \frac{\exp\left(-\frac{\beta_2^2}{2\beta^4}\int_t^T g(r,Y_r^{t,x})^2 dr\right)}{\mathbb{E}\left[\exp\left(-\frac{\beta_2^2}{2\beta^4}\int_t^T g(r,Y_r^{t,x})^2 dr\right)\right]}, \tag{5.22}$$

then we get

$$p(t,x) = -\frac{1}{k}\ln\left[E^{\hat{\mathbb{P}}}\left[\exp\left(-kF(Y_T^{t,x})\right)\right]\right].$$

We now can represent the indifference price, if the Novikov condition for $(\frac{\beta_1}{\beta_2}g(.,X^{t,x}) \cdot W)$ is satisfied. The indifference price is given by

$$p(t,x) = -\frac{1}{k}\ln\left[E^{\hat{\mathbb{Q}}}\left[\exp\left(-kF(X_T^{t,x})\right)\right]\right] \tag{5.23}$$

with

$$\frac{d\hat{\mathbb{Q}}}{d\mathbb{Q}} = \frac{\exp\left(-\frac{\beta_2^2}{2\beta^4}\int_t^T g(r,X_r^{t,x})^2 dr\right)}{\mathbb{E}\left[\exp\left(-\frac{\beta_2^2}{2\beta^4}\int_t^T g(r,X_r^{t,x})^2 dr\right)\right]}. \tag{5.24}$$

We see that the indifference price does not depend on the initial wealth v and the price is independent of the Brownian Motion B and is only driven by W.

5.3.2 Cross Hedging strategy

Here we want to analyse the relation between the optimal strategies π^* and $\pi^{\#}$. We will see that the difference depends on the sensitivity of the indifference price due to changes in X_t. Recall that the price is given by

$$p(t,x) = -\frac{1}{k}\frac{\varphi^F(t,x)}{\varphi^0(t,x)}.$$

The price is differentiable with respect to x (the initial capital), if the two properties from above are satisfied. From equation (5.10) we obtain a optimal strategy π^* which is given by

$$\pi^*(t,x) = \frac{g(t,x)}{\eta\beta^2} - \frac{\beta_1}{\beta^2}\sigma(t,x)\frac{\partial\Gamma}{\partial x}(t,x) \tag{5.25}$$

$$= \frac{g(t,x)}{\eta\beta^2} + \frac{\beta_1\sigma(t,x)}{k\beta^2}\frac{1}{\varphi^0(t,x)}\frac{\partial\varphi^0}{\partial x}(t,x) - \frac{\beta_1\sigma(t,x)}{\beta^2}\frac{\partial p}{\partial x}(t,x) \tag{5.26}$$

(5.27)

The strategy $\pi^{\#}$ is given by (5.18). After some elementary calculations we obtain a simple formula

$$\pi^*(t,x) = \pi^{\#}(t,x) - \frac{\beta_1\sigma(t,x)}{\beta^2}\frac{\partial p}{\partial x}(t,x). \tag{5.28}$$

Remark that the optimal strategy π^* depends on the risk sensitivity $\frac{\partial p(t,x)}{\partial x}$ of the indifference price of the derivative at time t and risk level x. Let us introduce the following notion.

Definition 5.4 *The function defined by*

$$\delta(t,x) = -\beta_1\sigma(t,x)\frac{\partial p(t,x)}{\partial x}, \quad 0 \le t \le T, x \in \mathcal{U} \tag{5.29}$$

is called diversification pressure at time t and initial capital x.

The choice of the name will be explained by the following considerations. We suppose that the risk sensitivity is positive. The derivative $F(X_T)$ diversifies the portfolio risk if the correlation $\beta_1\sigma(t,x)$ between the external risk and the price is negative. This means that the investment in the asset is higher i.e. $\pi^* > \pi^{\#}$. On the other hand, if the the correlation is negative, the derivative $F(X_T)$ amplifies the financial risk.
The following formula summarizes the relationship between the strategies in terms of the diversification pressure.

Theorem 5.1 *The optimal cross-hedging strategy π^* differs from the classical optimal investment strategy $\pi^{\#}$ only by the diversification pressure which means more precisely*

$$\pi^*(t,x) = \pi^{\#}(t,x) + \frac{1}{\beta^2}\delta(t,x). \tag{5.30}$$

5.4 Example

In this section we want to consider a particular case, which is underpinned by an example of external risk source. This example will give us a connection between the mathematical tools and different practical problems. The example is based on an idea of Davis [28]. He proposes another method for pricing HDD derivatives which differs from the methods we have presented in the former chapters. He showed that the accumulated HDD, which are equivalent to the definition of the HDD in Chapter 2, are almost lognormally distributed. Because of that we assume that X is a geometric Brownian motion. The corresponding open set $\mathcal{U}$ is $\mathbb{R}^* = \mathbb{R}\backslash\{0\}$. We will prove the two properties form the section above for this special case.

5.4.1 HDD (The geometric case)

We assume that the price process P and the risk process X are geometric Brownian motions with drift. We set $b(t,x) = \mu x$, $\sigma(t,x) = \nu x$ and $g = \alpha$, with $\mu, \alpha \in \mathbb{R}$ and $\nu \in \mathbb{R}\backslash\{0\}$. Because of these assumptions, we do not need any further assumption for the derivative function F. F only has to satisfy the condition

$$F(x) \geq -K(1 + \ln(x)) \quad \forall x \geq 1 \tag{5.31}$$

with K constant, because the expectation in (5.19) must be finite. A further assumption is that the initial value of X satisfies $X_0 = x_0 > 0$.
We obtain for φ^G the following form

$$\varphi^G(t,x) = \mathbb{E}\left[\exp\left(-kG(xZ_{T-t})\right)\exp\left(-\frac{\beta_2^2}{2\beta^4}\alpha^2(T-t)\right)\right]$$

where Z_{r-t} denotes a normally distributed random variable with mean

$$a = \left(\mu - \frac{\nu\alpha\beta_1}{\beta^2} - \frac{\nu^2}{2}\right)(r-t)$$

and variance

$$b^2 = \nu^2(r-t).$$

Next we continue by proving the regularity properties.
The first property is satisfied because

$$h^0(t,x,v) = U(v)\exp\left(-\frac{\beta_2^2}{2\beta^4}\alpha^2(T-t)\right).$$

The second property will follow from the next theorem where we show the regularity of the function h^F. For notational simplification we define

$$I(t,x) = \mathbb{E}\left[\exp\left(-kF(xZ_{T-t})\right)\right],$$

which implies that we can write

$$h^F = \exp\left(-\frac{\beta_2^2}{2\beta^4}\alpha^2(T-t)\right) I(t,x).$$

Theorem 5.2 *Assume that the growth condition (5.31) is satisfied. The function h^F is a classical solution, i.e. h belongs to the class $C^{1,2}([0,T) \times \mathbb{R}_+^* \times \mathbb{R}; \mathbb{R})$. Moreover if F is continuous, h^F is in $C^0([0,T] \times \mathbb{R}_+^* \times \mathbb{R}; \mathbb{R})$.*

Proof:
We have to prove that I belongs to $C^{1,2}([0,T) \times \mathbb{R}_+^* \times \mathbb{R}; \mathbb{R})$. Observe that for all $x > 0$ and $t < T$

$$\begin{aligned} I(t,x) &= \int_{\mathbb{R}} \exp(-kF(xe^z))\exp\left(-\frac{(z-a)^2}{2b^2}\right)\frac{1}{b\sqrt(2\pi)}dz \\ &= \int_{\mathbb{R}} \exp(-kF(e^v))\exp\left(-\frac{(v-\ln(x)-a)^2}{2b^2}\right)\frac{1}{b\sqrt(2\pi)}dv. \end{aligned}$$

If follows from the dominated convergence theorem, that for all $x > 0$ and $t < T$

$$\begin{aligned} \frac{\partial I}{\partial x}(t,x) &= \int_{\mathbb{R}} \exp(-kF(e^v))\exp\left(-\frac{(v-\ln(x)-a)^2}{2b^2}\right)\frac{(v-\ln(x)-a)}{b^2x}\frac{1}{b\sqrt(2\pi)}dv \\ &= \frac{1}{x}\int_{\mathbb{R}} \exp(-kF(xe^z))\exp\left(-\frac{(z-a)^2}{2b^2}\right)\frac{(z-a)}{b^2}\frac{1}{b\sqrt(2\pi)}dz \\ &= \frac{1}{x}\mathbb{E}\left[\exp\left(-kF(xe^Z)\right)\frac{(Z-a)}{b^2}\right]. \end{aligned}$$

By an analogous way we obtain

$$\frac{\partial^2 I}{\partial x^2}(t,x) = \frac{1}{x^2}\mathbb{E}\left[\exp\left(-kF(xe^Z)\right)\left(\frac{(Z-a)}{b^2} - \frac{1}{2}\right)^2\right] - \frac{4+b^2}{4b^2x^2}I(t,x)$$

Applying the dominated convergence theorem we show that

$$\begin{aligned} \frac{\partial I}{\partial t}(t,x) &= \frac{1}{2(T-t)}I(t,x) + \frac{a'}{b^2}\mathbb{E}\left[\exp\left(-kF(xe^Z)\right)(Z-a)\right] \\ &\quad - \frac{1}{2\nu^2(T-t)^2}\mathbb{E}\left[\exp\left(-kF(xe^Z)\right)(Z-a)^2\right] \end{aligned}$$

with

$$a' = \left(\frac{\nu\alpha\beta_1}{\beta^2} + \frac{\nu^2}{2} - \mu\right).$$

Note that if F is continuous then $I(T,x) = \exp\left(-kF(x)\right)$ is also continuous and we obtain that h^F is continuous. □

Next we want to given an explicit formula for the diversification pressure for this special case. Recall that the optimal strategy is given by (5.10) and satisfies

$$\pi^*(t,x) = \frac{\alpha}{\beta^2}\left(1 + \frac{\beta\nu x}{\alpha k}\frac{1}{I(t,x)}\frac{\partial I}{\partial x}(t,x)\right).$$

Consequently is the diversification pressure given by

$$\delta(t,x) = \frac{\beta_1\nu}{k}\frac{x}{I(t,x)}\frac{\partial I}{\partial x}(t,x).$$

For the geometric case the diversification pressure has the following properties.

Theorem 5.3 *If F is bounded, then δ is finite on $[0,T) \times \mathbb{R}_+^*$. If F is differentiable and satisfies*

$$\exists M \geq 0, \forall x \geq 0, \quad |xF(x)| \leq M, \tag{5.32}$$

then

$$\forall t \in [0,T], \forall x > 0, \quad |\delta(t,x)| \leq \beta_1\nu M.$$

Proof:

Let F be bounded. Recall from Theorem 5.2 that for all $x > 0$ and $t < T$

$$\frac{\partial I}{\partial x}(t,x) = \frac{1}{b^2 x}\mathbb{E}\left[\exp\left(-kF(xe^Z)\right)(Z-a)\right].$$

By the boundedness of F we obtain

$$\delta(t,x) = \frac{\beta_1\nu}{k}\frac{1}{b^2}\frac{\mathbb{E}\left[\exp\left(-kF(xe^Z)\right)(Z-a)\right]}{\mathbb{E}\left[\exp\left(-kF(xe^Z)\right)\right]} \leq C\mathbb{E}\left[\left|\frac{Z-a}{b^2}\right|\right] \leq \tilde{C}\frac{1}{b}$$

and the first result follows.

We assume now that F is differentiable and (5.32) holds. Note that $I(t,.) = \exp(-kF(.))$, we have

$$\delta(T,x) = \beta_1\nu x F'(x)$$

and this is bounded by $\beta_1\nu M$, caused by (5.32). For $t < T$ we have

$$\frac{\partial I}{\partial x}(t,x) = (-k)\mathbb{E}\left[e^Z F'(xe^Z)\exp(-kF(xe^Z))\right].$$

Hence

$$\delta(t,x) = (-\beta_1\nu)\frac{\mathbb{E}\left[xe^Z F'(xe^Z)\exp\left(-kF(xe^Z)\right)\right]}{\mathbb{E}\left[\exp\left(-kF(xe^Z)\right)\right]}$$

which implies

$$|\delta(t,x)| \leq \beta_1\nu M.$$

□

Theorem 5.4 *If F is Lipschitz-continuous and non-decreasing, then there is a constant K such that*

$$|\delta(t,x)| \leq K(1+|x|)$$

and δ is non-incrasing.

Proof:
Recall that

$$\frac{\partial I}{\partial x}(t,x) = (-k)\mathbb{E}\left[\exp\left(-kF(xe^Z)\right)e^Z\frac{\partial F}{\partial x}(xe^z)\right].$$

Suppose that $\frac{\partial F}{\partial x}$ is bounded by M. Then

$$\begin{aligned}\left|\frac{\partial I}{\partial x}(t,x)\right| &\leq kM\mathbb{E}\left[\exp\left(-kF(xe^Z)\right)e^Z\right]\\ &= kM\text{exp}\left(a+\frac{b^2}{2}\right)\mathbb{E}\left[\exp\left(-kF\left(xe^{b^2}e^{bN+a}\right)\right)\right],\end{aligned}$$

where N denotes a normally distributed random variable. Thus

$$\left|\frac{1}{I(t,x)}\frac{\partial I}{\partial x}(t,x)\right| \leq kM\text{exp}\left(a+\frac{b^2}{2}\right)\frac{\mathbb{E}\left[\exp\left(-kF\left(xe^{b^2}e^{bN+a}\right)\right)\right]}{\mathbb{E}\left[\exp\left(-kF(xe^{bN+a})\right)\right]}.$$

If F is increasing, then for all $(t,x) \in [0,T) \times \mathbb{R}_+^*$ we have

$$\left|\frac{1}{I(t,x)}\frac{\partial I}{\partial x}(t,x)\right| \leq kM\text{exp}\left(a+\frac{b^2}{2}\right) \tag{5.33}$$

and

$$\frac{1}{I(T,x)}\frac{\partial I}{\partial x}(T,x) = (-k)\frac{\partial F}{\partial x}(x).$$

We see that inequality (5.33) holds for $t = T$. □

5.5 When is the cross hedging strategy admissible?

In this section we analyse some conditions under which the optimal strategies π^* and $\pi^\#$ are quasi-admissible. Here we want to give some results (for the proofs, see Ankirchner et al. [5]). First we give conditions under which the derivatives of φ^0 and φ^F are bounded, before we give sufficient conditions for the optimal strategies $\pi^\#$ and π^*.

Proposition 5.1 *Suppose that there exists a constant C such that σ and $\hat{b}$ are globally Lipschitz continuous and for all $(t,x) \in [0,T] \times \mathbb{R}$ we have $\frac{|\hat{b}(t,x)|+|\sigma(t,x)|}{1+|x|} \leq C$.*
If g is bounded and Lipschitz continuous in x, then the partial derivative $\frac{\partial \varphi^0}{\partial x}$ is bounded and if in addition F is Lipschitz continuous and bounded from below, then $\frac{\partial \varphi^F}{\partial x}$ is bounded.

The following proposition follows immediately.

Proposition 5.2 *Let the assumptions of Proposition 5.1 be satisfied and suppose that σ is bounded. The the optimal strategies $\pi^{\#}$ and π^* are bounded.*

Theorem 5.5 *Let the assumptions of Proposition 5.1 be satisfied and additional we assume that σ is bounded. Then $U(V_T^{\#})$ and $U(V_T)$ are integrable and the strategies $\pi^{\#}$ and π^* are quasi-admissible.*

Let us now consider some special case of our model.

5.5.1 $\beta_1 = 0$

If $\beta_1 = 0$, then the price process and the process of the underlying non-tradeable risk are only correlated via the drift part. In this case we obtain

$$\pi_t^* = \pi_t^{\#} = \frac{g(t, X_t)}{\eta \beta_2}$$

and obviously the strategies do not depend on the structure of the derivative $F(X_T)$. If g is bounded or continuous in t and x, then the stochastic integral of $g(., X)$ relative to B is defined. We will also assume that this integral is defined.

Theorem 5.6 *Let $\beta_1 = 0$. Then $U(V_T^*)$ is integrable and π^* is quasi-admissible.*

5.5.2 The geometric case

Now we consider the case were the price and risk process are geometric Brownian motions. We only give the key results without any proof.

Theorem 5.7 *Under the assumption*

$$\pi^{\#} = \frac{\alpha}{\eta \beta^2},$$

the optimal strategy $\pi^{\#}$ maximizing the right-hand side of (6.6) is quasi-admissible.

Now we want to give a sufficient criterion for the optimal strategy π^* to be quasi-admissible.

Theorem 5.8 *Suppose that there exist constants $A, C > 0$ such that the diversification pressure satisfies*

$$|\delta(u, x)| \leq C + A|ln(x)| \quad \forall (u, x) \in [0, T] \times \mathbb{R}_+^*.$$

If A is sufficiently small, then the optimal control π^* is quasi-admissible.

Chapter 6

Pricing and hedging of derivatives based on non-traded underlyings

In the following chapter we present an amplification of the model in Chapter 5 and we describe the indifference price and hedges in terms of solution processes of forward-backward stochastic differential equations. Here we mainly follow Ankirchner et al. [4]. We will calculate utility based indifference prices and explicit hedging strategies. Before we start with the pricing model we give a short introduction of some mathematical terms, which will be helpful for a better understanding of the following sections.

6.1 Introduction

6.1.1 Indifference Pricing

For pricing claims in an incomplete market, indifference pricing based on a utility criterion is a popular method. Provided that the agent has a utility function U, the indifference price corresponds to the maximal amount π, that an agent is willing to pay for a claim X. This means that π is the amount the agent pays such the expected utility remains unchanged when doing the transaction:

$$\mathbb{E}\left[U(X-\pi)\right] = U(0).$$

Here π does not denote the transaction price, but it is an upper bound for the price of the claim.

Definition 6.1 *A risk-averse utility function is a concave and monotone increasing function that assigns to every value x a utility $U(x)$.*

Exponential Utility

We introduce this special utility function because it is widely used in the financial literature and we also use it in the following sections. We consider the exponential utility function

$$U(x) = -\gamma \text{exp}(-\frac{1}{\gamma}x) \tag{6.1}$$

with γ the risk-tolerance parameter. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space with time horizon T. The wealth W of the agent at this future date T is uncertain. W can be seen as a particular position on a given portfolio. To reduce the risk of the agent in his position, he can decide to buy a contingent claim with payoff X at time T or not. To simplify the model we do not consider an interest rate and assume that both variables X and W are bounded.
The agent can decide if he buys the claim or not. The maximum price he is willing to pay for the claim X is the indifference price $\pi(X)$ which is given by

$$\mathbb{E}\left[U(W + X - \pi(X))\right] = \mathbb{E}\left[U(W)\right].$$

Rewriting this equations leads to

$$\mathbb{E}\left[\text{exp}(-\frac{1}{\gamma}(W + X - \pi(X)))\right] = \mathbb{E}\left[\text{exp}(-\frac{1}{\gamma}W)\right].$$

The indifference price for the claim X given the wealth W is given by

$$\pi(X|W) = e_\gamma(W) - e_\gamma(W + X)$$

where e_γ is given by

$$e_\gamma(\Psi) := \gamma \text{ln}\mathbb{E}\left[\text{exp}\left(-\frac{1}{\gamma}\Psi\right)\right] \tag{6.2}$$

for any bounded variable Ψ.
In the following sections we will use a utility exponential utility function of the form

$$U(x) = -\text{exp}(-\eta x) \tag{6.3}$$

with η the risk averse parameter. In general, the risk aversion parameter is the inverse of the risk-tolerance parameter.

6.1.2 Backward Stochastic Differential Equations

In this section an introduction to backward stochastic differential equations BSDE is given. A better understanding of BSDEs may be helpful for the later calculations. After the definition of BSDEs and references to some key results, we focus on mathematical tools which are helpful and essential for some proofs in the further sections.

Definition 6.2 *Backwards stochastic differential equations BSDE are equations of the following type:*

$$Y_t = \zeta + \int_t^T f(s, Y_s, Z_s)ds - \int_t^T Z_s dW_s \quad 0 \le t \le T, \tag{6.4}$$

where $(W_t)_{0 \le t \le T}$ *is a standard* $d-$*dimensional Brownian motion on a filtered probability space* $(\Omega, \mathcal{F}, \mathbb{P})$ *with* $(\mathcal{F}_t)_{0 \le t \le T}$ *the standard Brownian filtration. The random function* $f : [0,T] \times \mathbb{R}^n \times \mathbb{R}^{n \times d} \to \mathbb{R}$ *is called a coefficient or generator,* T *the terminal time, which may be a stopping time, and the* $\mathcal{F}_T-$*adapted random variable* $\zeta = Y_T$ *a terminal condition. The parameters of (6.4) are* (f, T, ζ).

A solution of (6.4) is a pair of measurable processes $(Y, Z) \in \mathcal{S}^2(\mathbb{R}) \times \mathcal{H}^2(\mathbb{R}^d)$. Here $\mathcal{H}^2(\mathbb{R}^d)$ denotes the set of all $\mathbb{R}^d-$valued predictable processes ξ such that

$$\mathbb{E}\left[\int_0^T |\xi_t|^2 dt\right] < \infty$$

and $\mathcal{S}^2(\mathbb{R})$ denotes the set of all $\mathbb{R}-$valued predictable processes δ satisfying

$$\mathbb{E}\left[\sup_{s \in [0,T]} |\delta_s|^2\right] < \infty.$$

The set of one-dimensional progressively measurable processes which are almost surely bounded, for every t, is denoted by $\mathcal{H}^\infty(\mathbb{R})$.
Pardoux and Peng [46] introduced nonlinear BSDEs first. They assumed that f is Lipschitz continuous in the variables Y and Z and ζ is square integrable. They proved the existence and uniqueness of the solution, which is a pair (Y, Z) of square-integrable adapted processes. We will concentrate on BSDEs where the generator is continuous and has a quadratic growth in Z.

Definition 6.3 (Growth condition) *Let* $\alpha_0, \beta_0, b \in \mathbb{R}$ *and* c *be a continuous increasing function. We say that the coefficient* f *satisfies the growth condition (6.5) with* α_0, β_0, b, c *if for all* $(\omega, t, z) \in \mathbb{R}^+ \times \mathbb{R} \times \mathbb{R}^d$,

$$f(\omega, t, z) = a_0(\omega, t, z) + f_0(\omega, t, z),$$

with

$$\beta_0 \le a_0(\omega, t, z) \le \alpha_0 \quad a.s.$$
$$|f_0(\omega, t, z)| \le b + c(|t|)|z|^2 \quad a.s. \tag{6.5}$$

The following theorem is the main result about the existence of a solution of a BSDE with quadratic growth.

Theorem 6.1 *Let* (f, τ, ζ) *be a set of parameters of BSDE (6.4) and suppose that the coefficient* f *satisfies the growth condition with* $\alpha_0, \beta_0, b \in \mathbb{R}$ *and* $c : \mathbb{R}^+ \to \mathbb{R}^+$ *continuous increasing,* $\zeta \in L^\infty(\Omega)$, *and:*

1. *The terminal time τ is either bounded, ($\tau \leq T$ a.s) or*
2. *The terminal time is such that $\tau < \infty$ a.s and $\alpha_0 < 0$.*

Then the BSDE (6.4) has at least one solution $(Y, Z) \in \mathcal{H}^\infty(\mathbb{R}) \times \mathcal{H}^2(\mathbb{R}^d)$ such that the process Y has continuous paths.
Moreover there exists a minimal solution (Y_*, Z_*) [resp. a maximal solution (Y^*, Z^*)] such that for any set of parameters (g, τ, ξ), if

$$f \leq g \text{ and } \zeta \leq \xi \quad (\text{resp. } f \geq g \text{ and } \zeta \geq \xi)$$

and for any solution (Y_g, Z_g) of the BSDE with parameters (g, τ, ξ),

$$Y_* \leq Y_g \quad (\text{resp. } Y_* \geq Y_g).$$

Proof: See Kobylanski [43] □

For further information about the monotone stability which is helpful for the proof of the theorem and the main argument of the existence see Kobylanski [43]. Kobylanski also shows the uniqueness (Theorem 2.6) of the solution. We now give some mathematical properties of quadratic FBSDEs. A better understanding of BMO-martingales is helpful for the understanding of the following explanations. For this reason we refer for an introduction about BMO-martingales to Kazamaki [40]. We show the following results without any proof, which can be found in Ankirchner et al. [4].

Moment estimates for BSDE

We now give moment estimates for BSDE with coefficients that satisfy Lipschitz conditions with random bounds for the slopes. We assume that for our coefficient $f : \Omega \times [0, T]\mathbb{R}^d \to \mathbb{R}$, there exists an $\mathbb{R}^+-$valued predictable process H such that for all $(\omega, t, z) \in \Omega \times [0, T] \times \mathbb{R}^d$ we have

$$|f(\omega, t, z) - f(\omega, t, z')| \leq H_t|z - z'|. \tag{6.6}$$

Next we assume that H is such that the stochastic integral $\int_0^\cdot HdB$ with respect to a Brownian motion B is a so-called BMO martingale. $\int_0^\cdot HdB$ is a BMO martingale if and only if there exists a constant $C \in \mathbb{R}^+$ independent of ω such that for all stopping times τ with values in $[0, T]$ we have

$$\mathbb{E}\left[\int_\tau^T H_s^2 ds|\mathcal{F}_t\right] \leq C \quad a.s. \tag{6.7}$$

We use the definition and refer the smallest $C \in R^+$ such that the inequality (6.7) is satisfied as BMO-norm of H. Consider the BSDE

$$Y_t = \zeta \int_t^T Z_s dW_s + \int_t^T f(s, Z_s)ds, \quad 0 \leq t \leq T, \tag{6.8}$$

where W is a $d-$dimensional Brownian motion and ζ is a bounded $\mathcal{F}_t-$ measurable variable. The generator f satisfies (6.6) relative to a predictable H with finite BMO-norm. The moment estimator will be needed for establishing the smoothness of the solution of the quadratic BSDE with respect to the parameters, the terminal condition depends on.

Proposition 6.1 *Suppose that for all $\beta \geq 1$ we have $\int_0^T |f(s,0)|ds \in L^\beta(P)$. Let $p > 1$. Then there exist constants $q > 1$ and $C > 0$, depending only on p, T, and the BMO-norm of H, such that we have*

$$\mathbb{E}\left[\sup_{t\in[0,T]} |Y_t|^{2p}\right] + \mathbb{E}\left[\left(\int_0^T |Z_s|^2 ds\right)^p\right] \leq C\left(\mathbb{E}\left[|\zeta|^{2pq} + \left(\int_0^T |f(s,0)|ds\right)^{2pq}\right]\right)^{\frac{1}{q}}.$$

Differentiability of quadratic FBSDE

Consider the FBSDE

$$X_s^x = x + \int_0^t b(s, X_s^x)ds + \int_0^t \rho(s, X_s^x)dW_s \tag{6.9}$$

$$Y_s^x = F(X_T^x) - \int_t^T Z_s^x dW_s + \int_t^T f(s, X_s^x, Z_s^x)ds, \tag{6.10}$$

where $b : [0,T] \times \mathbb{R}^m \to \mathbb{R}^m$ and $\rho : [0,T] \times \mathbb{R}^m \to \mathbb{R}^{m\times d}$ and W is a $d-$dimensional Brownian motion. Note that ρ is a $n \times d$ matrix. The generator of the backward part is assumed to be a $\mathbb{P}(\mathcal{F}_t)\otimes\mathcal{B}(\mathbb{R}^m)\otimes\mathcal{B}(\mathbb{R}^d)-$measurable process $f : \Omega\times[0,T]\times\mathbb{R}^m\times\mathbb{R}^d \to \mathbb{R}$ such that there exists a constant $M \in \mathbb{R}^+$ such for all $(t,x,z) \in [0,T] \times \mathbb{R}^m \times \mathbb{R}^d$ we have

$$|f(t,x,z)| \leq M(1 + |z|^2) \quad a.s. \tag{6.11}$$

The $\sigma-$field of predictable sets with respect to the filtration $\mathcal{F}_t$ is denoted by $\mathbb{P}(\mathcal{F}_t)$. Furthermore we assume that

$$f \text{is differentiable in } x \text{ and } z \text{ and}$$

$$|\nabla_z f(t,x,z)| \leq M(1 + |z|) \quad \text{for all } (t,x,z) \in [0,T] \times \mathbb{R}^m \times \mathbb{R}^d \text{a.s.} \tag{6.12}$$

We want to give conditions for the process Y^x in the solution of the FBSDE (6.9) to be differentiable in x. We need a further assumption which assures us that the coefficients of the forward equation belong to the function space $\mathrm{B}^{m\times d}$ and $\mathrm{B}^{m\times 1}$, see Definition 6.8. To simplify the notation we introduce for the pair (b,ρ) of coefficient functions the second order differential operator

$$\mathcal{L} = \sum_{i=1}^m b_i(.)\frac{\partial}{\partial x_i} + \frac{1}{2}\sum_{i,j=1}^m [\rho\rho^T]_{ij}(.)\frac{\partial^2}{\partial x_i \partial y_j}.$$

Now we give some restrictions to the coefficients of the forward equation (6.9) which ensures us that X^x is differentiable in x.

(D1) $\rho \in \mathrm{B}^{m\times d}$, $b \in \mathrm{B}^{m\times 1}$ and

(D2) $F : \mathbb{R}^m \to \mathbb{R}$ is a twice differentiable function such that $\nabla F \cdot \rho \in \mathrm{B}^{1\times d}$ and $\mathcal{L}\mathcal{F} \in \mathrm{B}^{\infty\times\infty}$.

Next we give a standard result which will be helpful for the later derivations. e_i denotes the unit vector in $\mathbb{R}^m$.

Proposition 6.2 *Suppose that (D1) and (D2) are satisfied. For all $x \in \mathbb{R}^m$, $h \neq 0$ and $i \in \{1, \dots, m\}$, let*

$$\xi^{x,h,i} = \frac{1}{h}\left(F(X_T^{x+he_i}) - F(X_T^x)\right).$$

Then for every $p > 1$ there exists a $C > 0$, dependent only on p and the bounds of b, ρ, F and its derivatives, such that for all $x, x' \in \mathbb{R}^m$ and $h, h' \neq 0$,

$$\mathbb{E}\left[|\xi^{x,h,i} - \xi^{x',h',i}|^{2p}\right] \leq C(|x-x'|^2 + |h-h'|^2)^p. \tag{6.13}$$

The facts that F is bounded and growth condition (6.11) holds, ensures a unique solution $(Y^x, Z^x) \in \mathcal{H}^\infty(\mathbb{R}) \otimes \mathcal{H}^2(\mathbb{R}^d)$ of the BSDE in (6.9) for all $x \in \mathbb{R}^d$. We will choose a family $(Y^x)_{x\in\mathbb{R}^m}$ which is continuous in x.

Proposition 6.3 *Let (D1), (D2), (6.11) and (6.12) be satisfied, and assume that F is bounded and that there exists a constant $K \in \mathbb{R}^+$ such that for all $t \in [0, T]$, $x, x' \in \mathbb{R}^m$ and $z \in \mathbb{R}^d$*

$$|f(t, x, z) - f(t, x', z)| \leq K(1 + |z|)|x - x'|. \tag{6.14}$$

Then for all $p > 1$ there exists a constant $C \in \mathbb{R}^+$ such that for all $x, x' \in \mathbb{R}^m$,

$$\mathbb{E}\left[\sup_{t\in[0,T]} |Y_t^x - Y_t^{x'}|^{2p}\right] \leq C|x - x'|^{2p}, \tag{6.15}$$

$$\mathbb{E}\left[\left(\int_0^T |Z_t^x - Z_t^{x'}|^2 dt\right)^p\right] \leq C|x - x'|^{2p}. \tag{6.16}$$

In particular, Kolmogorov's continuity criterion implies that there exists a measurable process $\tilde{Y} : \Omega \times [0, T] \times \mathbb{R}^m$ such that $(t, x) \mapsto \tilde{Y}_t^x$ is continuous for almost all ω, and for all $(t, x) \in [0, T] \times \mathbb{R}^m$ we have $\tilde{Y}_t^x = Y_t^x$ a.s.

Pathwise continuous differentiability will be guaranteed by the following theorem.

Theorem 6.2 *Let (D1), (D2), (6.11) and (6.12) be satisfied, and suppose that F is bounded and f satisfies (6.14). Besides suppose that $\nabla_z f$ is globally Lipschitz continuous in (x, z) and that $\nabla_x f$ satisfies for all $t \in [0, T]$, $x, x' \in \mathbb{R}^m$ and $z, z' \in \mathbb{R}^d$*

$$|\nabla_x f(t, x, z) - \nabla_x f(t, x', z')| \leq K(1 + |z| + |z'|)(|x - x'| + |z - z'|). \tag{6.17}$$

Then there exists a function $\Omega \times [0,T] \times \mathbb{R}^m \to \mathbb{R}^{m+1+d}$, $(\omega, t, x) \mapsto (X_t^x, Y_t^x, Z_t^x)(\omega)$, *such that for almost all* ω, X_t^x *and* Y_t^x *are continuous in* t *and continuously differentiable in* x, *and for all* x, (X_t^x, Y_t^x, Z_t^x) *is a solution of FBSDE (6.9). Moreover, there exists a process* $\nabla_x Z^x \in \mathcal{H}^2$ *such that the pair* $(\nabla_x Y^x, \nabla_x Z^x)$ *solves the BSDE*

$$\nabla_x Y_t^x = \nabla_x F(X_T^x)\nabla_x X_T^x - \int_t^T \nabla_x Z_s^x dW_s \tag{6.18}$$

$$+ \int_t^T \left(\nabla_x f(s, X_s^x, Z_s^x)\nabla_x X_s^x + \nabla_z f(s, X_s^x, T_s^x)\nabla_x Z_s^x\right) ds. \tag{6.19}$$

Markov property of FBSDE

A time-inhomogeneous Markov process will solve the forward part of the FBSDE (6.9). Now we give some consequences of this fact. Fix an initial time $t \in [0,T)$ and x denotes the initial state of the forward process as time t. The forward process conditioned on taking the value x at time t satisfies the SDE

$$X_s^{t,x} = x + \int_t^s b(r, X_r^{t,x})dr + \int_t^s \rho(r, X_r^{t,x})dW_r, \tag{6.20}$$

with $x \in \mathbb{R}^m$ and $s \in [t,T]$. Let us assume that the coefficients satisfy a growth and a Lipschitz condition. Let $C \in R^+$ be a constant such that for all $x, x' \in \mathbb{R}$ and $t \in [0,T]$

$$\begin{aligned} |b(t,x) - b(t,x')| + |\rho(t,x) - \rho(t,x')| &\leq C(|x - x'|), \\ |b(t,x)| + |\rho(t,x)| &\leq C(1 + |x|). \end{aligned} \tag{6.21}$$

This condition guarantees that there exists a unique solution of (6.20). It also implies that $X_r^{t,x}$ is Malliavin differentiable and that its Malliavin gradient has a representation, which involves for (t,r) fixed the global flow on the space of nonsingular linear operators $\Phi^{t,x}$ on $\mathbb{R}^m$ defined by

$$\Phi_s^{t,x} = I_m + \int_t^s \nabla_x b(u, X_u^{t,x})\Phi_u^{t,x} du + \int_t^s \nabla_x \rho(u, X_u^{t,x})\Phi_u^{t,x} dW_u. \tag{6.22}$$

Here $\nabla_x b$ and $\nabla_x \rho$ denotes the gradient under the condition (6.21) and I_m the $m \times m$ unit matrix. The Malliavin gradient is given by

$$D_\vartheta X_s^{t,x} = \Phi_s^{t,x}(\Phi_\vartheta^{t,x})^{-1}\rho(\vartheta, X_\vartheta^{t,r}), \quad t \leq \vartheta \leq s. \tag{6.23}$$

$X^{t,x}$ denotes the Markov process starting at time t in x. Let us consider the following BSDE

$$Y_s^{t,x} = F(X_T^{t,x}) - \int_s^T Z_r^{t,x} dW_r + \int_s^T f(r, X_r^{t,x}, Z_r^{t,x})dr. \tag{6.24}$$

Here and in the further explanations we assume that the generator is a deterministic Borel measurable function $f : [0,T] \times \mathbb{R}^m \times \mathbb{R}^d \to \mathbb{R}$. We also assume that f is differentiable in

(x, z) and that there exists a constant $M \in \mathbb{R}^+$ such that for all $(t, x, z) \in [0, T] \times \mathbb{R}^m \times \mathbb{R}^d$ we have

$$|f(t, x, z)| \leq M(1 + |z|^2)a.s. \quad \text{and} \quad |\nabla_z f(t, x, z)| \leq M(1 + |z|)a.s. \tag{6.25}$$

for all $(t, z) \in [0, T] \times \mathbb{R}^m$. We know that, if F is bounded that there exists a unique solution $(Y^{t,x}, Z^{t,x}) \in \mathcal{H}^\infty(\mathbb{R}) \otimes \mathcal{H}^2(\mathbb{R}^d)$ of the BSDE (6.24). The following theorem shows that the solution of the BSDE is already determined by the forward process $X^{t,x}$. $\mathcal{D}^m$ denotes for all $m \in \mathbb{N}$ the $\sigma-$algebra on $\mathbb{R}^m$ generated by the family of functions

$$x \mapsto \mathbb{E}\left[\int_t^T \varphi(s, X_s^{t,x})ds\right]$$

with $x \in \mathbb{R}^m$, $t \in [0, T]$ and $\varphi : [0, T] \times \mathbb{R}^m \to \mathbb{R}$ is bounded and continuous.

Theorem 6.3 *Let $F : \mathbb{R}^m \to \mathbb{R}$ be a bounded Borel function, suppose that f satisfies (6.25) and the coefficients of the forward diffusion (6.21). Suppose that there exists functions $f_n : [0, T] \times \mathbb{R}^m \times \mathbb{R}^d \to \mathbb{R}$, globally Lipschitz continuous in (x, z), such that for almost all ω and for all compact sets $K \subset \mathbb{R}^m \times \mathbb{R}^d$ the sequence f_n converges to f uniformly on $[0, T] \times K$. Then there exist two $\mathcal{B}[0, T] \otimes \mathcal{D}^m-$ and $\mathcal{B}[0, T] \otimes \mathcal{D}^m-$measurable deterministic functions u and v on $[0, T] \times \mathbb{R}^m$ such that*

$$Y_s^{t,x} = u(s, X_s^{t,x}) \quad \text{and} \quad Z_s^{t,x} = v(s, X_s^{t,x})\rho(s, X_s^{t,x}), \tag{6.26}$$

for $P \otimes \lambda-$ almost all $(\omega, s) \in \Omega \times [t, T]$.

Using Theorem 6.3 and Theorem 6.2 we can represent the control process $Z^{t,r}$ in terms of the derivative of $Y^{t,r}$ with respect to x.

Theorem 6.4 *Suppose that the assumptions of Theorem 6.3 are satisfied. Besides assume that $\nabla_z f$ is globally Lipschitz continuous, that (6.14) and (6.17) are satisfied, and further that the forward coefficients satisfy the stronger conditions (D1) and (D2). Then $u(t, x)$ is differentiable in x for almost all $t \in [0, T]$. Moreover*

$$Z_s^{t,x} = \nabla_x u(t, X_s^{t,x})\rho(s, X_s^{t,x}). \tag{6.27}$$

Differentiability of quadratic BSDE with parameterized terminal condition

Let us consider the BSDE

$$Y_t^x = \xi(x) - \int_t^T Z_s^x dW_s + \int_t^T f(s, Z_s^x)ds, \quad t \in [0, T], x \in \mathbb{R}^m \tag{6.28}$$

and we assume that

(E1) $\mathbb{R}^m \ni x \mapsto \xi(x) \in \mathbb{R}$ is a bounded random field which as a function of X is differentiable with bounded partial derivatives. $\nabla\xi(x)$ is also Lipschitz in x; also $f(t, 0)$ is $\mathcal{F}_t-$adapted and satisfies $f(t, 0) \in L^p$ for $p \geq 1$.

(E2) there exists $M \in \mathbb{R}_+$ such that $|f(t,z)| \leq M(1+|z|^2)$a.s.; f is differentiable in z such that $|\nabla_z f(t,z)| \leq M(1+|z|)$ for all $(t,z) \in [0,T] \times \mathbb{R}^d$ a.s.

(E3) for all $x \in \mathbb{R}^m$, $h \neq 0$ and $i \in \{1,\dots,m\}$, let $\zeta^{x,h,i} = \frac{1}{h}(\xi(x+he_i) - \xi(x))$. Then for every $p > 1$ there exists a $C > 0$, dependent only on p, such that for all x, $x' \in \mathbb{R}^m$ and h, $h' \neq 0$,

$$\mathbb{E}\left[|\zeta^{x,h,i} - \zeta^{x',h',i}|^{2p}\right] \leq C(|x-x'|^2 + |h-h'|^2)^p. \tag{6.29}$$

The assumptions (E1)-(E3) allow us to apply the methods of subsection "Differentiability of quadratic FBSDE" and we obtain the following theorem.

Theorem 6.5 *Let (E1)-(E3) be satisfied. Then there exists a function* $\Omega \times [0,T] \times \mathbb{R}^m \to \mathbb{R}^{1+d}$, $(\omega,t,x) \mapsto (Y_t^x, Z_t^x)(\omega)$, *such that for almost all* ω, *the process* Y_t^x *is continuous in* t *and continuously differentiable in* x, *and for all* x, (Y_t^x, Z_t^x) *is a solution of BSDE (6.28). Moreover, there exists a process* $\nabla_x Z^x \in \mathcal{H}^2(\mathbb{R}^{m\times d})$ *such that the pair* $(\nabla_x Y^x, \nabla_x Z^x)$ *solves the BSDE*

$$\nabla_x Y_t^x = \nabla_x \xi(x) - \int_t^T \nabla_x Z_s^x dW_s + \int_t^T (\nabla_z f(s, Z_s^x) \nabla_x Z_s^x)\, ds.$$

6.2 The Model

Let us consider a d-dimensional Brownian motion W, $d \in \mathbb{N}$ on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. The filtration generated by W is denoted by $(\mathcal{F}_t)_{0\leq t\leq T}$. We assume that a derivative is based on a m-dimensional non-tradeable index with maturity $T > 0$. Its dynamics is given by

$$dR_t = b(t,R_t)dt + \rho(t,R_t)dW_t, \tag{6.30}$$

with $b : [0,T] \times \mathbb{R}^m \to \mathbb{R}^m$ and $\rho : [0,T] \times \mathbb{R}^m \to \mathbb{R}^{m\times d}$ measurable deterministic functions. Next we assume that there exists a constant $C \in \mathbb{R}_+$ such that for all $t \in [0,T]$ and $x, x' \in \mathbb{R}^m$

$$\begin{aligned} |b(t,x) - b(t,x')| + |\rho(t,x) - \rho(t,x')| &\leq C\,|x-x'|\,, \\ |b(t,x)| + |\rho(t,x)| &\leq C(1+|x|). \end{aligned} \tag{6.31}$$

The derivative which we consider is given by $F(R_T)$, where $F : \mathbb{R}^m \to \mathbb{R}$ is a bounded and measurable function. The expected payoff at time t of $F(R_T)$, conditioned on $R_t = r$, is given by $F(R_T^{t,r})$, where $R^{t,r}$ is the solution of the SDE

$$R_s^{t,r} = r + \int_t^s b(u, R_u^{t,r})du + \int_t^s \rho(u, R_u^{t,r})dW_u \tag{6.32}$$

for $s \in [t,T]$. To create a hedge we also need assets which are correlated to the non-tradeable index (6.30). The considered finance market consists of k risky assets and one non-risky asset. The non-risky asset will be used as numeraire and we suppose that the prices of the risky assets in the units of the numeraire evolve according to the SDE

$$dS_t^i = S_t^i(\alpha_i(t, R_t))dt + \beta_i(t, R_t)dW_t) \tag{6.33}$$

for $i = 1, \dots, k$. Here $\alpha_i(t, r)$ denotes the ith component of a measurable and vector-valued map $\alpha : [0, T] \times \mathbb{R}^m \to \mathbb{R}^k$ and $\beta_i(t, r)$ is the ith row of a measurable and matrix-valued map $\beta : [0, T] \times \mathbb{R}^m \to \mathbb{R}^{k \times d}$. The Brownian Motion driving the risk-assets is the same $\mathbb{R}^d-$dimensional Brownian motion as the one driving the process of the index (6.30). The correlation between the non-tradeable index and the tradeable is determined by the matrices ρ and β.

We assume $d \geq k$, to guarantee that the market is arbitrage-free. For technical reasons we suppose that

(M1) α is bounded,

(M2) there exist constants $0 < \epsilon < K$ such that $\epsilon I_k \leq (\beta(t, r)\beta^T(t, r)) \leq KI_k$ for all $(t, r) \in [0, T] \times \mathbb{R}^m$,

where $\beta^T(t, r)$ is the transpose of $\beta(t, r)$ and I_k the $k-$dimensional unit matrix.

Before we derive explicit hedging strategies and we want to give some examples where our model can be applied. In the first example we present a derivative which is based on one underlying only.

Example 6.1 *In the previous chapter we have already seen a possibility to derive degree futures and options. Davis proposes in [28] another method for pricing HDD derivatives. He shows that the accumulated HDD are almost lognormally distributed. For this reason he models them by a geometric Brownian Motion. If we apply this fact to our framework in (6.30) we would have to choose* $b(t, R_t) = \alpha_1 R_t$ *and* $\rho(t, R_t) = \alpha_2 R_t$ *with* $\alpha_1 \in \mathbb{R}$ *and* $\alpha_2 \in \mathbb{R}\backslash\{0\}$*. Tradeable assets which are correlated with temperature indices are for example electricity futures and stocks of electricity producers.*

In the next example we look at a derivative which is based on more than one underlying.

Example 6.2 *Here we consider a special type of spread option. Spread options involve in general more underlying structures (prices, indices, interest rates and other quantities), and measure the distance between them. Carmona and Durrleman gives in [24] an overview of pricing and hedging spread options. Now we analyse a 2-dimensional crack spread.*

This special type consists of the simultaneous purchase of crude against the sale or purchase of refined petroleum products. We look at a kerosene crack spread, which pits crude oil price (co) *against kerosene* (ke)*. A firm which produces kerosene from crude oil wants to cover a part of its risk arising from a sudden increase of the crude oil price by buying a kerosene crack spread. The problem here is that no sufficiently liquid market for kerosene exists. This means it is impossible to buy a kerosene future or other products whose underlying is*

kerosene. Hence contracts of this type can only be arranged over the counter.
A possibility to hedge kerosene crack spreads can be hedged by using the high correlation between the price of the heating oil (he) *and the kerosene price. The modeling of the prices will be done in the following way:*

$$\begin{aligned} dR_t^{ke} &= R_t^{ke}(b_1 dt + \gamma_2 dW_t^1 + \gamma_3 dW_t^2 + \gamma_4 dW_t^3) \\ dR_t^{co} &= R_t^{co}(b_2 dt + \gamma_1 dW_t^1) \\ dS_t^{ho} &= S_t^{ho}(b_3 dt + \beta_1 dW_t^1 + \beta_2 dW_t^2) \end{aligned}$$

with $b_1, b_2, b_3 \in \mathbb{R}$, $\gamma_1, \gamma_2, \gamma_3, \gamma_4, \beta_1, \beta_2 \in \mathbb{R}\backslash\{0\}$ *and the correlation between heating oil and kerosene is given by*

$$\sigma = \frac{\gamma_2\beta_1 + \gamma_3\beta_2}{\sqrt{(\gamma_2^2 + \gamma_3^2 + \gamma_4^2)(\beta_1^2 + \beta_2^2)}}.$$

A European call on the spread with strike K *is of the form* $F(R_T^{ke}, S_T^{co}) = (R_T^{ke} - S_T^{co} - K)^+$.

Before we estimate the hedging strategies we define the wealth process. First of all let U be an exponential utility function with risk tolerance parameter $\eta > 0$. We define an investment strategy as any predictable process $\lambda = (\lambda^i)_{1 \le i \le k}$ with values in $\mathbb{R}^k$ such that the integral process $\int_0^t \lambda_r^i \frac{dS_r^i}{S_r^i}$ is defined for all $i \in \{1, \dots, k\}$. In our framework λ^i denotes the value of the portfolio fraction invested in the $i-$th asset. Investing according to a strategy λ leads to a total gain due to trading during the time interval $[t, s]$ which amounts to

$$G_s^{\lambda,t} = \sum_{i=1}^k \int_t^s \lambda_u^i \frac{dS_u^i}{S_u^i}.$$

The gain conditioned on $R_t = r$ will be denoted by $G_s^{\lambda,t,r}$. Using the definition of $\frac{dS_u^i}{S_u^i}$ we obtain

$$G_s^{\lambda,t,r} = \sum_{i=1}^k \int_t^s \lambda_u^i [\alpha_i(u, R_u^{t,r})du + \beta_i(u, R_u^{t,r})dW_u]. \tag{6.34}$$

We see that the wealth process does not depend on the value of the correlated price process.

Definition 6.4 (Stopping Time) *A nonnegative random variable* τ*, which is allowed to take value* ∞*, is called a stopping time if for each t, the event*

$$\{\tau \le t\} \in \mathcal{F}_t.$$

The definition of admissible strategies guarantees that the market is arbitrage-free.

Definition 6.5 (Admissible strategies) *Let* $\mathcal{A}^{t,r}$ *be a set of all strategies* λ *such that*

$$\mathbb{E}\left[\int_t^T |\lambda_s \beta(s, R_s^{t,r})|^2 ds\right] < \infty$$

and the family

$$\left\{exp(-\eta G_\tau^{\lambda,t,r}) : \tau \textit{ is a stopping time with values in}[t,T]\right\}$$

is uniformly integrable. If $\lambda \in \mathcal{A}^{t,r}$, then λ is admissible.

The maximal expected utility at time T, conditioned on the wealth v at time t and the index satisfy $R_t = r$ is defined by

$$V^0(t,v,r) = \sup\left\{\mathbb{E}\left[U(v + G_T^{\lambda,t,r})\right] : \lambda \in \mathcal{A}^{t,r}\right\}. \tag{6.35}$$

In the next section we show that there exists an optimal strategy π which solves this maximization problem. If an investor holds a derivative $F(R_T)$ in his portfolio until maturity T, the problem differs from the expected maximal utility above and has the following form:

$$V^F(t,v,r) = \sup\left\{\mathbb{E}\left[U(v + G_T^{\lambda,t,r} + F(R_T^{t,r}))\right] : \lambda \in \mathcal{A}^{t,r}\right\}. \tag{6.36}$$

The optimal strategy of this problem will be denoted by $\hat{\pi}$. If an investor holds a derivative $F(R_T)$, the optimal strategy must be changed form π to $\hat{\pi}$. The difference

$$\Delta = \hat{\pi} - \pi$$

which evolves from the presence of the derivative in the portfolio should be hedged. Therefore we call Δ derivative hedge. In the following sections we derive an explicit expression for Δ. We have already seen in the introduction of this chapter that for all $(t,r) \in [0,T] \times \mathbb{R}$ there exists a real number $p(t,r)$ such that for all $v \in \mathbb{R}$

$$V^F(t, v - p(t,r), r) = V^0(t,v,r).$$

Then $p(t,r)$ is called the indifference price at time t and level r.

6.3 Solving the control problem

After describing the model we want to show that the supremum is taken by a particular strategy $\hat{\pi}$ which is admissible. Before we start solving this problem, let us introduce the following notations and assumptions. Let

$$\vartheta(t,r) = \beta^T(t,r)(\beta(t,r)\beta^T(t,r))^{-1}\alpha(t,r)$$

and

$$C(t,r) = \left\{x\beta(t,r) : x \in \mathbb{R}^k\right\}.$$

Our assumptions imply that $\upsilon(t,r)$ is bounded. The set $C(t,r)$ is a closed and convex set. The distance of a vector $z \in \mathbb{R}^d$ to $C(t,r)$ is defined as

$$\text{dist}(z, C(t,r)) = \min\left\{|z-u| : u \in C(t,r)\right\}.$$

At last we define $\prod_{C(t,r)}(z)$ which is the projection of a vector $z \in \mathbb{R}^d$ onto the linear subspace $C(t,r)$.
Recall the maximization problem (6.36)

$$V^F(t,v,r) = \sup\left\{\mathbb{E}\left[U(v + G_T^{\lambda,t,r} + F(R_T^{t,r}))\right] : \lambda \in \mathcal{A}^{t,r}\right\}. \tag{6.37}$$

Before estimating the optimal strategy we introduce the following notation:

$$l_s := \lambda_s \beta(s,r)$$

and note that the set of admissible strategies $\mathcal{A}^{t,r}$ is equivalent to a set $\tilde{\mathcal{A}}^{t,r}$ of $\mathbb{R}^d$-valued predictable processes l with $l \in \tilde{\mathcal{A}}$, which we will also call strategies.

Definition 6.6 (New admissible strategies) *Let $\tilde{\mathcal{A}}^{t,r}$ be a set of all strategies l such that*

$$\mathbb{E}\left[\int_t^T |l_s|^2 ds\right] < \infty$$

and the family

$$\left\{exp(-\eta G_\tau^{l,t,r}) : \tau \textit{ is a stopping time with values in} [t,T]\right\}$$

is uniformly integrable. If $l_s \in \tilde{\mathcal{A}}^{t,r}$ then l_s is admissible.

Note that $l_s \in C(t,r)$ and $G_s^{l,t,r}$ is given by

$$G_s^{l,t,r} = \sum_{i=1}^{k} \int_t^s l_u^i [dW_u + \vartheta(u, R_u^{t,r})du].$$

Now the maximization problem is evidently equivalent to

$$V^F(t,v,r) = \sup\left\{\mathbb{E}\left[U(v + G_T^{l,t,r} + F(R_T^{t,r}))\right] : l \in \tilde{\mathcal{A}}^{t,r}\right\}. \tag{6.38}$$

We construct a family of stochastic processes J^l, to find the value function and the optimal strategy $\hat{\pi}$. The family of processes has the following properties:

- $J_T^l = U(v + G_T^{l,t,r} + F(R^{t,r}))$ for all $l \in \tilde{\mathcal{A}}^{t,r}$
- $J_0^l = J_0$ is constant for all $l \in \tilde{\mathcal{A}}^{t,r}$
- J^l is a supermartingale for all $l \in \tilde{\mathcal{A}}^{t,r}$ and there exists a $\hat{l} \in \tilde{\mathcal{A}}^{t,r}$ such that $J^{\hat{l}}$ is a martingale.

The process J^l and the initial value J_0 depend on the initial wealth v. Now we can compare the expected utilities of the strategies $l \in \tilde{\mathcal{A}}^{t,r}$ and $\hat{l} \in \tilde{\mathcal{A}}^{t,r}$ by

$$\mathbb{E}\left[U(v + G_T^{l,t,r} + F(R_T^{t,r}))\right] \le J_0(t,v,r) = \mathbb{E}\left[U(v + G_T^{\hat{l},t,r} + F(R_T^{t,r}))\right] = V^F(t,v,r) \tag{6.39}$$

with $\hat{l}$ is the desired optimal strategy. To construct the family, we set

$$J_s^{l,t,r} := U(v + G_s^{l,t,r} - \hat{Y}_s^{t,r}) \quad s \in [t,T],$$

where the pair of processes $(\hat{Y}_s^{t,r}, \hat{Z}_s^{t,r})$ is the solution of the BSDE

$$\hat{Y}_s^{t,r} = F(R_t^{t,r}) - \int_s^T \hat{Z}_u^{t,r} dW_u - \int_s^T f(u, R_u^{t,r}, \hat{Z}_u^{t,r}) du \quad s \in [t,T]. \tag{6.40}$$

To guarantee that the solution of the BSDE is unique, f must be differentiable in z and satisfies the growth condition. For more details see Section 6.1.2. We have to choose a function f for which $J_s^{\lambda,t,r}$ is a supermartingale for $l \in \tilde{\mathcal{A}}^{t,r}$ and $J_s^{\hat{l},t,r}$ for $\hat{l} \in \tilde{\mathcal{A}}^{t,r}$ is a martingale. We get

$$V^F(t,v,r) = J_0^{l,t,r} = U(v - \hat{Y}_0^{t,r}) \quad \text{for all } l \in \tilde{\mathcal{A}}^{t,r} \tag{6.41}$$

For the calculation of f, we write J as product of a martingale M^l and a not strictly decreasing process K^l that is constant for a $\hat{l} \in \tilde{\mathcal{A}}^{t,r}$. For $t \in [0,T]$, we define

$$M_t^l = U(v - \hat{Y}_0^{t,r})) * \exp\left(-\int_0^t \eta(l_s - \hat{Z}_s^{t,r}) dW_s - \frac{1}{2}\int_0^t \eta^2 (l_s - \hat{Z}_s^{t,r})^2 ds\right).$$

By comparing J^l with $M^l K^l$ we obtain

$$K_t^l = -\exp\left(\int_0^t \psi(s, l_s, \hat{Z}_s^{t,r})\right)$$

with

$$\psi(t,l,z) = -\eta l \vartheta(t,r) + \eta f(t,r,z) + \frac{1}{2}\eta^2 |l - z|^2. \tag{6.42}$$

To obtain a decreasing process K^l, f has to satisfy

$$\psi(s, l_s, \hat{Z}_s^{t,r}) \geq 0 \quad \text{for all } l \in \tilde{\mathcal{A}}^{t,r}$$

and

$$\psi(s, \hat{l}_s, \hat{Z}_s^{t,r}) = 0$$

for $\hat{l} \in \tilde{\mathcal{A}}^{t,r}$. For $t \in [0,T]$, we have

$$\begin{aligned}
\frac{1}{\eta}\psi(t, l_t, \hat{Z}_t^{t,r}) &= \frac{\eta}{2}|l_t|^2 - \eta l_t \left(\hat{Z}_t^{t,r} + \frac{1}{\eta}\vartheta(t,r)\right) + \frac{\eta}{2}|\hat{Z}_t^{t,r}|^2 + f(t,r,\hat{Z}_t^{t,r}) \\
&= \frac{\eta}{2}\left|l_t - \left(\hat{Z}_t^{t,r} + \frac{1}{\eta}\vartheta(t,r)\right)\right|^2 - \frac{\eta}{2}\left|\hat{Z}_t^{t,r} + \frac{1}{\eta}\vartheta(t,r)\right|^2 + \frac{\eta}{2}(\hat{Z}_t^{t,r})^2 + f(t,r,\hat{Z}_t^{t,r}) \\
&= \frac{\eta}{2}\left|l_t - \left(\hat{Z}_t^{t,r} + \frac{1}{\eta}\vartheta(t,r)\right)\right|^2 - \hat{Z}_t^{t,r}\vartheta(t,r) - \frac{1}{2\eta}|\vartheta(t,r)|^2 + f(t,r,\hat{Z}_t^{t,r}).
\end{aligned}$$

Now we set

$$f : [0,T] \times \mathbb{R}^m \times \mathbb{R}^d \to \mathbb{R}, (t,r,z) \mapsto z\vartheta(t,r) + \frac{1}{2\eta}|\vartheta(t,r)|^2 - \frac{\eta}{2}\text{dist}^2(z + \frac{1}{\eta}\vartheta(t,r), C(t,r)).$$

For this choice we get $\psi(t,\lambda,z) \geq 0$ and for

$$\hat{l} = \hat{\pi}_s\beta(s,r) \in \prod_{C(s,r)} \left(\hat{Z}_s^{t,r} + \frac{1}{\eta}\vartheta(s,r)\right), \quad s \in [t,T] \tag{6.43}$$

we obtain $\psi(.,\hat{l},\hat{Z}^{t,r}) = 0$. To find the maximal expected wealth, or in other words the value function of our stochastic control problem we have to find the orthogonal projection of the $d-$dimensional vector z to the linear space $C(t,r)$ of strategies. The fact that $C(t,r)$ is convex leads to the uniqueness of the minimizer. The following theorem summarizes the essential information about the optimal strategy.

Theorem 6.6 *The value function of the optimization problem (6.36) is given by*

$$V^F(t,v,r) = -exp\left(-\eta(v - \hat{Y}_t^{t,r})\right),$$

where $\hat{Y}_s^{t,r}$ is defined by the unique solution $(\hat{Y}^{t,r}, \hat{Z}^{t,r})$

$$\hat{Y}_s^{t,r} = F(R_t^{t,r}) - \int_s^T \hat{Z}_u^{t,r} dW_u - \int_s^T f(u, R_u^{t,r}, \hat{Z}_u^{t,r})du \quad s \in [t,T].$$

with

$$f(t,r,z) = z\vartheta(t,r) + \frac{1}{2\eta}|\vartheta(t,r)|^2 - \frac{\eta}{2}dist^2(z + \frac{1}{\eta}\vartheta(t,r), C(t,r))$$

There exists an optimal trading strategy $\hat{\pi} \in \mathcal{A}^{t,r}$, if $R_t = r$ with

$$\hat{\pi}_s\beta(s,R_s^{t,r}) = \prod_{C(s,R_s^{t,r})} \left[\hat{Z}_s^{t,r} + \frac{1}{\eta}\vartheta(s,R_s^{t,r})\right], \quad s \in [t,T].$$

Proof: See Hu et al. [36] □

In an analogous calculation we obtain the optimal strategy for the portfolio without the derivative. Let $(Y^{t,r}, Z^{t,r})$ be the solution of

$$Y_s^{t,r} = -\int_s^T Z_u^{t,r} dW_u - \int_s^T f(u, R_u^{t,r}, Z_u^{t,r})du, \quad s \in [t,T]. \tag{6.44}$$

We obtain for the maximal expected utility

$$V^0(t,v,r) = -exp\left(-\eta(v - Y_t^{t,r})\right)$$

and the optimal strategy π on $[t, T]$ satisfies

$$\hat{\pi}_s \beta(s, R_s^{t,r}) = \prod_{C(s,R_s^{t,r})} \left[Z_s^{t,r} + \frac{1}{\eta} \vartheta(s, R_s^{t,r}) \right] \quad s \in [t, T]. \tag{6.45}$$

Using the fact that $\prod_{C(s,R_s^{t,r})}$ is a linear operator, the derivative hedge is given by

$$\Delta_s \beta(s, R_s^{t,r}) = \prod_{C(s,R_s^{t,r})} \left[\hat{Z}_s^{t,r} - Z_s^{t,r} \right] \tag{6.46}$$

which we derive in the following sections.

6.4 Markov property of indifference prices

Now we want to show the Markov property of indifference prices.

Definition 6.7 (Markov property) *X is a Markov process if for any t and $s > 0$, the conditional distribution of $X(t+s)$ given $\mathcal{F}_t$ is the same as the conditional distribution of $X(t+s)$ given $X(t)$, that is,*

$$P(X(t+s) \leq y | \mathcal{F}_t) = P(X(t+s) \leq y | X(t)), \quad a.s.$$

Using the fact that the solutions of the BSDEs (6.40) and (6.44) are deterministic functions of time and the underlying, the property will follow. We introduce the following σ-algebras to give a complete and correct proof. Fixing $t \in [0, T]$, we denote by $\mathcal{D}^m$ the σ-algebra generated by the functions

$$r \rightarrow \mathbb{E} \left[\int_t^T \phi(s, R_s^{t,r}) ds \right]$$

with $t \in [0, T]$ and ϕ is a bounded continuous $\mathbb{R}$-valued function. Moreover we assume that the map $(t, r) \rightarrow \vartheta(t, r)$ is Lipschitz continuous in r. Using the facts

- α is bounded,
- there exist constants $0 < \epsilon < K$ such that $\epsilon I_k \leq (\beta(t, r)\beta^T(t, r)) \leq K I_k$ for all $(t, r) \in [0, T] \times \mathbb{R}^m$,

this is guaranteed if β and α are Lipschitz continuous.

Proposition 6.4 *There exist $\mathcal{B}[0, T] \otimes \mathcal{D}^m-$measurable deterministic functions u and $\hat{u}$: $[0, T] \times \mathbb{R}^m \rightarrow \mathbb{R}$ such that*

$$Y_s^{t,r} = u(s, R_s^{t,r})$$

and

$$\hat{Y}_s^{t,r} = \hat{u}(s, R_s^{t,r})$$

for $P \otimes \lambda - a.s.(w, s) \in \Omega \times [t, T]$.

Proof: See Ankirchner et al. [4]. □

This proposition implies that the indifference price p can be given as function of (t, r).

Theorem 6.7 *There exists a* $\mathcal{B}[0,T] \otimes \mathcal{D}^m-$*measurable deterministic function* $p : [0,T] \times \mathbb{R}^m \to \mathbb{R}$ *such that for all* $x \in \mathbb{R}, (t,r) \in [0,T] \times \mathbb{R}^m$

$$V^F(t, v - p(t,r), r) = V^0(t, v, r). \tag{6.47}$$

Proof:
Let $v \in \mathbb{R}, (t,r) \in [0,T] \times \mathbb{R}^m$ be given. We know that

$$V^F(v,t,r) = -\exp\left(-\eta(v - \hat{Y}_t^{t,r})\right)$$

and

$$V^0(v,t,r) = -\exp\left(-\eta(v - Y_t^{t,r})\right).$$

Now put

$$p(t,r) = u(t,r) - \hat{u}(t,r),$$

where u and $\hat{u}$ are given from the proposition above and the result follows. □

We assume that the function p is measurable in both t and r. It inherits this property from the functions u and $\hat{u}$.
Now we turn to the description of the optimal strategy and the derivative hedge. We want to derive that from the solutions of the BSDE, which we considered in the last section. We describe the optimal strategy in terms of time and the index process R. In other words it should only depend on the time and the index process.

Theorem 6.8 *There exist* $\mathcal{B}[0,T]\otimes\mathcal{D}^m-$*measurable deterministic functions* v *and* $\hat{v}$*, defined on* $[0,T] \times \mathbb{R}^m$ *and taking values in* $\mathbb{R}^d$ *such that for* $(t,r) \in [0,T] \times \mathbb{R}^m$*, the optimal strategies, conditioned on* $R_t = r$*, are given by*

$$\pi_s = v(s, R_t^{0,r})$$

and

$$\hat{\pi}_s = \hat{v}(s, R_s^{t,r})$$

for all $s \in [t,T]$.

Proof:
Fix $(t,r) \in [0,T] \times \mathbb{R}^m$. Theorem 6.3 implies that there exist $\mathcal{B}[0,T] \otimes \mathcal{D}^m-$measurable deterministic functions v and $\hat{v}$ mapping $[0,T] \times \mathbb{R}^m$ to $\mathbb{R}^m$ such that for all $s \in [0,T]$

$$Z_s^{t,r} = v(s, R_s^{t,r})\rho(s, R_s^{t,r})$$

and

$$\hat{Z}_s^{t,r} = \hat{v}(s, R_s^{t,r})\rho(s, R_s^{t,r}).$$

Let

$$\gamma(t,r) = \prod_{C(t,r)} \left[v(t,r)\rho(t,r) + \frac{1}{\eta} v(t,r) \right]$$

and

$$\hat{\gamma(t,r)} = \prod_{C(t,r)} \left[\hat{v}(t,r)\rho(t,r) + \frac{1}{\eta} v(t,r) \right].$$

Then, by (6.43) and (6.45) the optimal strategies conditioned on $R_t = r$ satisfy

$$\pi_s \beta(s, R_s^{t,r}) = \gamma(s, R_s^{t,r})$$

and

$$\hat{\pi}_s \beta(s, R_s^{t,r}) = \hat{\gamma}(s, R_s^{t,r}),$$

for all $s \in [t, T]$. The fact that the rank of $\beta(t,r)$ is k implies that

$$v(t,r) = \gamma(t,r)\beta^T(t,r)(\beta(t,r)\beta^T(t,r))^{-1}$$

and

$$\hat{v}(t,r) = \hat{\gamma}(t,r)\beta^T(t,r)(\beta(t,r)\beta^T(t,r))^{-1}$$

are well defined. The uniqueness of π and $\hat{\pi}$ yields the result.

□

This theorem implies that the optimal strategies are so-called Markov controls.

This section will be closed by noting that Theorem 6.7 implies a dynamic principle for the indifference price. Let $\mathcal{A} = \mathcal{A}^{0,r}$ for some $r \in \mathbb{R}^m$. For any stopping time $\tau \leq T$ and $\mathcal{F}_\tau-$measurable random variable G_τ, let

$$V^F(\tau, G_\tau) = \text{esssup} \left\{ \mathbb{E} \left[U(G_\tau + G_T^{\lambda,\tau} + F(R_T^{0,r})) | \mathcal{F}_\tau \right] : \lambda \in \mathcal{A} \right\}$$

and similarly we define $V^0(\tau, G_\tau)$.

Proposition 6.5 *We have*

$$V^F(\tau, G_\tau - p(\tau, R_\tau^{0,r})) = V^0(\tau, G_\tau).$$

Proof: See Ankirchner et al. [4] □

6.5 Explicit hedging strategies

We want to obtain an explicit form of the derivative hedge in terms of the price gradient. To get this result we have to impose stronger conditions on the coefficients of the index process R and the function F and we have to show that the price function is differentiable in r. Let us introduce the following class of functions.

Definition 6.8 *Let $n, p \geq 1$. We denote by $\boldsymbol{B}^{n\times p}$ the set of all functions $h : [0,T] \times \mathbb{R}^m \to \mathbb{R}^{n\times p}$, $(t,x) \mapsto h(t,x)$, differentiable in x, for which there exists a constant $C > 0$ such that*

$$\sup_{(t,x)\in[0,T]\times\mathbb{R}^m} \sum_{i=1}^m \left| \frac{\partial h(t,x)}{\partial x_i} \right| \leq C$$

for all $t \in [0,T]$ we have $\sup_{x\in\mathbb{R}^m} \frac{|h(t,x)|}{1+|x|} \leq C$, and $x \mapsto \frac{\partial h(t,x)}{\partial x}$ is Lipschitz continuous with Lipschitz constant C.
In addition to (6.31) we assume that the coefficients of the index diffusion satisfy

$$\rho \in \boldsymbol{B}^{m\times d}, b \in \boldsymbol{B}^{m\times 1} \tag{6.48}$$

and F is bounded and twice differentiable function such that

$$\nabla F \cdot \in \boldsymbol{B}^{1\times d} and \sum_{i=1}^m b_i(t,r)\frac{\partial}{\partial r_i}F(r) + \frac{1}{2}\sum_{i,j=1}^m [\rho\rho^T]_{i,j}(t,r)\frac{\partial^2}{\partial r_i \partial r_j}F(r) \in \boldsymbol{B}^{1\times 1}. \tag{6.49}$$

The following theorem guarantees Lipschitz continuity and differentiability of the functions u and $\hat{u}$ which we obtained from Proposition 6.4.

Theorem 6.9 *Suppose that (6.31), (6.48), (6.49) are satisfied. Besides, suppose that the volatility matrix β and the drift density α are bounded, Lipschitz continuous in r, differentiable in r and that for all $1 \leq i \leq k, 1 \leq j \leq d$ the derivatives $\nabla_r \beta_{ij}$ and $\nabla_r \alpha_i$ are also Lipschitz continuous in r. Then the functions u and $\hat{u}$ are Lipschitz continuous in r and continuously differentiable in r.*

Proof: See Ankirchner et al. [4] □

The next result follows immediately.

Proposition 6.6 *Suppose that the assumptions of Theorem 6.9 are satisfied. Then the indifference price function p is continuously differentiable in r.*

Now we can derive an explicit formula for the derivative hedge in terms of the price gradient. Let us denote the conditional derivative hedge by $\Delta(t,r) = \hat{v}(t,r) - v(t,r)$, $(t,r) \in [0,T] \times \mathbb{R}^m$.

Theorem 6.10 *Under the assumptions of Theorem 6.9, and with the notation of the previous sections, the derivative hedge satisfies*

$$\Delta(t,r) = -\nabla_r p(t,r)\rho(t,r)\beta^T(t,r)(\beta(t,r)\beta^T(t,r))^{-1}, \quad (t,r) \in [0,T] \times \mathbb{R}^m. \tag{6.50}$$

Obviously Theorem 6.10 implies that the derivative hedge at time t only depends on R_t.
Proof:
Note that $C(t,r)$ is a linear subspace of $\mathbb{R}^d$ for all $(t,r) \in [0,T] \times \mathbb{R}^m$. Therefore, the projection operator $\prod_{C(t,r)}$ is linear and hence

$$\begin{aligned}\Delta(t,r) &= (\hat{\gamma}(t,r) - \gamma(t,r))\beta^T(t,r)(\beta(t,r)\beta^T(t,r))^{-1} \\ &= \left(\prod_{C(t,r)}\left[\hat{Z}_t^{t,r} + \frac{1}{\eta}\upsilon(t,r)\right] - \prod_{C(t,r)}\left[Z_t^{t,r} + \frac{1}{\eta}\upsilon(t,r)\right]\right)\beta^T(t,r)(\beta(t,r)\beta^T(t,r))^{-1} \\ &= \left(\prod_{C(t,r)}\left[\hat{Z}_t^{t,r} - Z_t^{t,r}\right]\right)\beta^T(t,r)(\beta(t,r)\beta^T(t,r))^{-1}.\end{aligned}$$

It follows from Theorem 6.4 that

$$\hat{Z}_t^{t,r} - Z_t^{t,r} = (\nabla_r \hat{u}(t,r) - \nabla_r u(t,r))\,\rho(t,r) = -\nabla_r p(t,r)\rho(t,r), \tag{6.51}$$

and hence the result follows. □

Proposition 6.7 *Let $k = 1$. Then the derivative hedge is given by*

$$\begin{aligned}\Delta(t,r) &= -\frac{\langle \beta(t,r), \nabla_r p(t,r)\rho(t,r)\rangle}{|\beta(t,r)|^2} \\ &= -\frac{\sum_{i=1}^d \beta_i(t,r)\sum_{j=1}^m \frac{\partial}{\partial r_j}p(t,r)\rho_{ji}(t,r)}{\sum_{i=1}^d \beta_i^2(t,r)}, \quad (t,r) \in [0,T] \times \mathbb{R}^m.\end{aligned}$$

Proof:
Fix $(t,r) \in [0,T] \times \mathbb{R}^m$.
Note that $C(t,r) = \{x\beta(t,r) : x \in \mathbb{R}\}$ is a one-dimensional subspace of $\mathbb{R}^d$.
For all $z = (z_i)_{i \le i \le d} \in \mathbb{R}^d$ let

$$g(z) = \frac{\langle \beta(t,r), z\rangle}{|\beta(t,r)|^2} = \frac{\sum_{i=1}^d \beta_i(t,r)z_i}{\sum_{i=1}^d \beta_i^2(t,r)}.$$

Then $g(z)\beta(t,r)$ is the orthogonal projection of z onto $C(t,r)$. Thus Theorem 6.10 yields that $\Delta(t,r) = -g(\nabla_r p(t,r)\rho(t,r))$. □

We close this section with some remarks and the explicit derivative hedge for both examples.

1. We assume that the derivative $F(R_T)$ is traded on an exchange. By pretending the price observed is approximately equal to an indifference price, the hedging formula (6.50) provides a very simple tool for hedging the derivative. The risk aversion coefficient η does not appear in the formula (6.50).

2. If $k = d$ and the matrices $\beta(t,r)$ are all invertible, then our financial market is complete and the derivative $F(R_T)$ can be fully replicated. The derivative hedge satisfies

$$\Delta(t,r) = -\nabla_r p(t,r)\rho(t,r)\beta^{-1}(t,r).$$

If $S = R$ (which means that S is chosen to be the index), it follows that Δ coincides the classical delta hedge.

Let us now give the derivative hedge for both examples which we considered at the begining of this chapter.

Example 6.3 *In the first example we assumed that the HDD index is modeled as a geometric Brownian motion, and assume that there exists one tradeable correlated risky asset. In detail we set:* $d = 2$, $k = m = 1$, $\rho = (\alpha_2, 0)$, $\beta = (\beta_1, \beta_2)$ *with* $\alpha_2, \beta_1, \beta_2' \in \mathbb{R}\backslash\{0\}$. *Furthermore we obtain*

$$\Delta(t,r) = -\alpha_a \frac{\partial p(t,r)}{\partial r} \frac{\beta_1}{\beta_1^2 + \beta_2^2}$$

Example 6.4 *If we apply our results to the second example we have to take* $m = 2$, $k = 2$, *and* $d = 3$. *It follows that:*

$$\rho = \begin{pmatrix} \gamma_1 & 0 & 0 \\ \gamma_2 & \gamma_3 & \gamma_4 \end{pmatrix}, \quad \beta = \begin{pmatrix} \gamma_1 & 0 & 0 \\ \beta_1 & \beta_2 & 0 \end{pmatrix}, \quad \beta^T(\beta\beta^T)^{-1} = \begin{pmatrix} \beta_2 & 0 \\ -\beta_1 & \gamma_1 \\ 0 & 0 \end{pmatrix}.$$

We compute

$$\prod_{C(t,r)} [\nabla_r p(t,r)\rho(t,r)] = \left(\gamma_1 \frac{\partial}{\partial r_1} p(t,r) + \gamma_2 \frac{\partial}{\partial r_2} p(t,r), \gamma_3 \frac{\partial}{\partial r_2} p(t,r), 0\right).$$

We obtain the following delta hedge for our example:

$$\Delta(t,r) = \left(-\frac{\partial}{\partial r_1} p(t,r) + \left(\frac{\beta_1\gamma_3}{\gamma_1\beta_2} - \frac{\gamma_2}{\gamma_1}\right) \frac{\partial}{\partial r_2} p(t,r), -\frac{\gamma_3}{\beta_2} \frac{\partial}{\partial r_2} p(t,r)\right)$$

for $(t,r) \in [0,T] \times \mathbb{R}^2$, *where* r_1 *represents the crude oil and* r_2 *the kerosene variable. If* $\gamma_4 = 0$, *we have a perfect hedge and if* $\gamma_3 = 0$, *then the price of heating oil does not play a role in the hedge.*

6.6 Pricing by marginal utility

We assume that there does not exist a market for our derivative $F(R_T)$ and it is sold over-the-counter. In the following we want to find a reasonable price, which a seller could ask for the derivative. Of course one possibility is the indifference price, but the disadvantage of the indifference price is that it is not linear. Note that the indifference price of $2 \times F(R_T)$ is not the same as twice the indifference price of $F(R_T)$. To obtain a linear version we may take the limit of the indifference price as the quantity converges to 0. Thus we derive the indifference price for a vanishing amount of derivatives, and it is therefore called marginal utility price (MUP). If an investor has to pay the MUP for each derivative he is indifferent between buying and not buying the infinitesimal amount of the derivative. We further assume that (R1)-(R3) are satisfied. Next we update the notation and we define by $p(t,r,q)$ the indifference price of q units of $F(R_T^{t,r})$ for $q \in \mathbb{R}$ and $(t,r) \in [0,T] \times \mathbb{R}^m$. Therefore is $p(t,r,q)$ the unique real satisfying

$$\sup_{\lambda} \left\{ \mathbb{E}\left[U(v + G_T^{\lambda,t,r} + qF(R_T^{t,r}) - p(t,r,q)) \right] \right\} = \sup_{\lambda} \left\{ \mathbb{E}\left[U(v + G_T^{\lambda,t,r}) \right] \right\}.$$

The price of one unit is equal to $\frac{p(t,r,q)}{q}$, $(q \neq 0)$ and the MUP is defined by

$$\text{MUP}(t,r) = \frac{\partial}{\partial q} p(t,r,q)|_{q=0}.$$

Recall that $p(t,r,q) = Y_t^{t,r} - \hat{Y}_t^{t,r,q}$, where $(\hat{Y}^{t,r,q}, \hat{Z}^{t,r,q})$ is the solution of the BSDE

$$\hat{Y}_s^{t,r,q} = qF(R_T^{t,r}) - \int_s^T \hat{Z}_u^{t,r,q} dW_u - \int_s^T f(u, R_u^{t,r}, \hat{Z}_u^{t,r,q}) du \quad s \in [t,T].$$

Set $\xi(q) = qF(R_T^{t,r})$, then $\xi(q)$ is a globally bounded differentiable Lipschitz with bounded derivatives. ξ inherits the boundedness from F and we are only interested in the differentiability of the process with relation to q in a neighborhood of zero. Due the boundedness of F and the quadratic growth of f the conditions of Theorem 6.5 are satisfied. Therefore, the process $\hat{Y}^{t,r,q}$ is continuous in t and continuously differentiable in q.
Differentiating the BSDE with respect to q we obtain

$$\frac{\partial}{\partial q}\hat{Y}_s^{t,r,q} = F(R_T^{t,r}) - \int_s^T \frac{\partial}{\partial q}\hat{Z}_u^{t,r,q} dW_u - \int_s^T \nabla_z f(u, R_u^{t,r}, \hat{Z}_u^{t,r,q}) \frac{\partial}{\partial q}\hat{Z}_u^{t,r,q} du \quad s \in [t,T]$$

Setting $q = 0$ and renaming the process we obtain

$$U_s^{t,r} = F(R_T^{t,r}) - \int_s^T V_s dW_s - \int_s^T \nabla_z f(s, R_s^{t,r}, Z_s^{t,r}) V_s ds.$$

We obtain the following explicit formula for the MUP of our derivative.

Theorem 6.11 *The explicit formula for the Marginal Utility PRice of the derivative $F(R_T)$ is given by*

$$MUP(t,r) = U_t^{t,r},$$

where $U_t^{t,r}$ *is the first component of the solution pair of the BSDE*

$$U_s^{t,r} = F(R_T^{t,r}) - \int_s^T V_s dW_s - \int_s^T \nabla_z f(s, R_s^{t,r}, Z_s^{t,r}) V_s ds.$$

Chapter 7

Hybrid CAT-Bonds

In the previous chapters we dealt with the modeling and pricing of temperature derivatives and weather derivatives. The discussed methods are suited for high-frequency events with low severity and low losses (compared to natural catastrophes like floods and earthquakes). For events with low frequency, but high impact we present in the last chapter a popular method for managing such events.
In this chapter we mainly follow Barrieu and Louberge [11], but for further readings about managing catastrophic risk we refer to Banks [6], Banks [7] and Grossi and Kunreuther [34]. Those readings give a general introduction to the analysis and modeling of catastrophic risk.

7.1 Introduction

In the past decades the frequency of natural catastrophes such as hurricanes, floods and earthquakes arises. Some of them led to huge human losses, others caused only economic losses. To deal with this source of risk, different catastrophe-linked securities have been developed. Not only national insurance programs have been installed i.e. Katastrophenfond Österreich, but also other catastrophe-linked securities have emerged. Among the more successful products on the market are CAT-bonds. This special possibility of alternative risk transfer is designed in such a way that its return is contingent upon the occurrence of a natural catastrophe. These bonds are standard coupon or zero-coupon bonds with a higher yield than the yield obtained on normal bonds, if no catastrophe occurs during a time interval. If the catastrophe occurs the bondholder loses a part or the whole return and the coupons are not (or only partially) paid.
Important to note is that CAT-bonds are not substitutes for reinsurance. They are used to complete the market, to fill gaps in market niches. In most cases CAT-bonds are used for low-frequency, high-severity layers, where reinsurance is unattractive, because of its high premium rates.
We will consider in the following sections two models of CAT-bonds. On the one hand a very simple CAT-bond and on the other hand a hybrid CAT-bond which is linked to the stock market. The hybrid bond could be a possibility to make the catastrophe-linked securities

more attractive. Moreover this type of contract protects the investor against downside risk.

7.2 Framework

Let $(\Omega, \mathcal{F}, \mathbb{P})$ a filtrated probability space with $\mathbb{P}$ the probability measure. We denote the agent, typically an insurer or reinsurer, who is exposed to an occurrence of a natural catastrophe, by A. We suppose that the amount of losses, if the catastrophe occurs, is known and denote it by X. In this framework the randomness does not evolve from the size of losses, but from the occurrence of the catastrophe itself. The time horizon is denoted by $T > 0$ and the random date of the catastrophe by τ.
Agent A wants to hedge his exposure to the risk of the occurrence of a natural catastrophe. To reduce his losses he will put a CAT-bond on the market. In the following section we consider two different products and analyse the impact on the transaction volume. In each case different agents are involved. First agent A, who is the issuer of the bond and exposed to the natural catastrophe. We denote the investor, possibly a hedge fund, by C. The third agent in our framework is an intermediary between A and C, typically a Special Purpose Vehicle (SPV) sponsored by an investment bank, who is acting on behalf of agent A to issue whose cat bond, is denoted by B. He only performs an advisory function and makes the transactions possible. He acts as a pure intermediary and retains no risk. We already mentioned that we consider two different CAT-bonds, let us now consider them more precisely.

Simple CAT-bond (Case A)

In this case agent A issues a CAT-bond, agent B buys it and transfers it to agent C. This bond is characterized by two parameters: The price, which is equivalent to the volume of the capital flowing into the CAT-bond market and the nominal amount N. The nominal amount is completely transfered by agent B. In contrary to the nominal amount agent B pays a price π_A to agent A and then agent B receives the price π_C from agent C. Here we consider a very simple and theoretical structure of a CAT-bond. We assume that no coupons are paid and the nominal amount to the buyer is paid at maturity T if no catastrophe occurs in the considered time interval. If a catastrophe occurs before T nothing is paid and the return of the cat bond is zero.

Hybrid CAT-bond (Case B)

The characteristics of the first transaction between agent A and agent B remains the same. Agent B buys the CAT-bond from agent A. Next agent B will issue a hybrid product which agent C buys. As in the simple case agent C pays π_C^h to agent B and receives the prospect of the nominal amount N^h, if no catastrophe occurs and the prospect of a fixed amount H if there is a catastrophe and a market crash before the maturity T. In the case of no catastrophe and a market boom, agent C receives only $N^h - H$, because he has to pay H to agent B. The hybrid product includes the sale of a digital put, paying H if there is no catastrophe and a

market boom, and the purchase of a digital put, paying H if there occurs a catastrophe and a market crash. We assume that the initial prices of the call option and the put option are the same. Agent B transfers the payoff of the call option, if exercised, to the financial market, because he only acts as intermediary and does not retain any risk in his book.
This special hybrid structure depends on the existence of such hybrid options on the market. In our framework we assume that this condition is fulfilled. The existence of those options can be relaxed to the consideration of a put and a call written only on a market index, without any reference to a natural catastrophe. In Section 7.4 we will declare the meaning of market crash and market boom.

Risk assessment

We assume that agent A and agent C assess their risk using a convex risk measure.

Definition 7.1 *The functional*

$$\begin{aligned} \rho :& \mathcal{X} \to \mathbb{R} \\ & \Psi \to \rho(\Psi) \end{aligned}$$

is a convex risk measure (in the sense of Föllmer and Schied) if, for any X and Y in $\mathcal{X}$, it satisfies the following properties:

1. *Convexity:* $\forall \lambda \in [0,1] \quad \rho(\lambda X + (1-\lambda)Y) \leq \lambda\rho(X) + (1-\lambda)\rho(Y)$
2. *Monotonicity:* $X \leq Y \Rightarrow \rho(X) \geq \rho(Y)$;
3. *Translation invariance:* $\forall m \in \mathbb{R} \quad \rho(X+m) = \rho(X) - m$

where $\mathcal{X}$ denotes a linear space of bounded functions.

We consider an exponential utility framework, therefore we use a very common risk measure. The entropic measure is defined as the functional (6.2). With γ_i we denote the risk tolerance coefficient of agent i $(i = A, C)$ and the risk measure associated with the terminal investment payoff Ψ is given by

$$\rho_i(\Psi) = \gamma_i \ln\mathbb{E}\left[\exp\left(-\frac{1}{\gamma_i}\Psi\right)\right].$$

We do not need to introduce the agent's initial wealth and can fix it equal to zero because of the translation invariance property of the entropic risk measure. Further readings about convex risk measures are Föllmer and Schied [31] and Föllmer and Schied [32].
The role of agent B is completely different, he acts only as intermediary. He is not really exposed to any risk and can only generate servicing fees. We assume that agent B is risk neutral and has no particular measure of risk.

For our further considerations we modify the risk measures. Agent A and C will determine his optimal investment by solving the following problem:

$$\min_{\xi_T \in \mathcal{V}_T} \rho_i(\Psi - \xi_T)$$

where Ψ is agent i's exposure $(i = A, C)$ and $\mathcal{V}_T$ is a convex set of bounded terminal gains at time T. Here the net gain corresponds to the spread between the terminal wealth and the capitalized initial wealth. The new convex risk measure is characterized by the value function of this optimization problem. It corresponds to the risk measure which agent i will have after having optimally chosen his investment on the market. The measure is denoted by ρ_i^m and called market modified risk measure. In the entropic framework we get

$$\rho_i^m(\Psi) = \gamma_i \ln \mathbb{E}_{\hat{\mathbb{Q}}}\left[\exp\left(-\frac{1}{\gamma_i}\Psi\right)\right]$$

with $\hat{\mathbb{Q}}$ the minimal entropy probability measure. For further readings we refer to Barrieu and El Karoui [10] and Barrieu and Louberge [11].
In the following we consider directly the market modified risk measure. This means that we will not work with the historical probability measure $\mathbb{P}$, but instead with the minimal entropy probability measure $\hat{\mathbb{Q}}$, which is common to all agents as they have the same access to the financial market.

7.3 Analysis of Case A

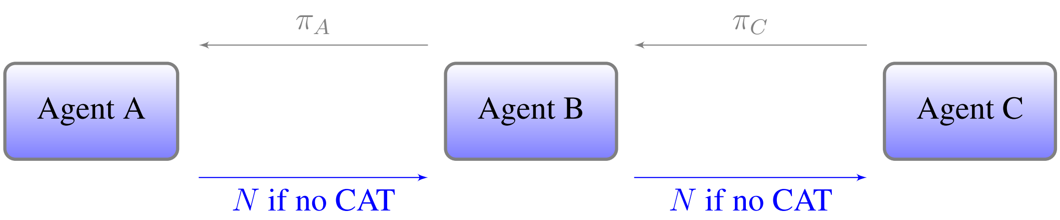

Figure 7.1: Simple CAT-Bond

Here we want to study and analyse the simple transaction. The structure is given in Figure 7.1.
In this transaction different prices are involved. On the one hand π_A, which is the price for the simple CAT-bond issued by agent A and purchased by agent B and on the other hand π_C, which denotes the price paid by agent C in his transaction with agent B. We also assume that agent B transfers the risk of the CAT-bond completely. We know that he acts only as intermediary. We intend to determine the optimal structure of the CAT-bond. In other words, we derive the prices and the nominal amount. We will do it in several steps, after describing the different payoffs, we determine the role of the agents in the optimal transaction.

7.3.1 The model

In our framework we distinguish two possible states, which depends on the occurrence of a natural catastrophe, before T or not. The probability of an event is given by the reference probability measure, in our study we use the minimal entropy probability measure $\hat{\mathbb{Q}}$ and $\hat{p} := \hat{\mathbb{Q}}(\tau \leq T)$. We see in the following Table 7.1 we see the payoffs of the different agents for the possible cases. We see that agent B is indifferent between the occurrence and the non-occurrence of the catastrophe.

Occurrence	Probability	Agent A	Agent B	Agent C
$T < \tau$	$1 - \hat{p}$	$\pi_A - N$	$\pi_C - \pi_A$	$-\pi_C - N$
$T \geq \tau$	$\hat{p}$	$\pi_A - X$	$\pi_C - \pi_A$	$-\pi_C$

Table 7.1: Payoffs

Let us consider the structure and the related cash-flows of the simple model. The transactions are done in the following way. First agent A determines the volume of the transaction with agent B. Agent A sets the reservation price π_A of the first transaction. For the second transaction agent C determines the price π_C. Finally agent B , who is the advisor of the whole transaction, determines the nominal amount N, such that he maximizes the fees he can generate, under the participation constraint imposed by both agents A and C. These fees, called service fees, cover the expenses which are related to the issue of transaction and the set-up. This problem can be written as the following optimization program:

$$\begin{aligned} &\max_{N,\pi_A,\pi_C} (\pi_C - \pi_A) \\ &s.t. : \\ &\quad \rho_A^m(\pi_A - N1_{\tau>T} - X1_{\tau\leq T}) \leq \rho_A^m(-X1_{\tau\leq T}) \\ &\quad \rho_C^m(-\pi_C + N1_{\tau>T}) \leq \rho_C^m(0) \end{aligned} \tag{7.1}$$

By using the translation invariance property of the risk measure and inserting each constraint in the maximization function we obtain:

$$\begin{aligned} \pi_A &= \rho_A^m(-N1_{\tau>T} - X1_{\tau\leq T}) - \rho_A^m(-X1_{\tau\leq T}) \\ \pi_C &= \rho_C^m(0) - \rho_C^m(N1_{\tau>T}) \end{aligned}$$

Let us determine the price of each part of the transaction which represents the volume of the capital flowing into the CAT-bond market. We will show that both pricing rules can be estimated explicitly in the entropic framework. First we consider the lower bound of the volume of the transaction between agent A and agent B. Recall

$$\pi_A = \rho_A^m(-N1_{\tau>T} - X1_{\tau\leq T}) - \rho_A^m(-X1_{\tau\leq T})$$

and by using the entropic risk measure we get the following result.

Theorem 7.1 *The indifference volume for the CAT-bond issue by agent A is given by:*

$$\pi_A = \gamma_A ln \left(\frac{(1-\hat{p})exp\left(\frac{1}{\gamma_A}N\right) + \hat{p}exp\left(\frac{1}{\gamma_A}X\right)}{(1+\hat{p}) + \hat{p}exp\left(\frac{1}{\gamma_A}X\right)} \right). \tag{7.2}$$

Note that $\pi_A > 0$ if and only if $N > 0$.

On the other hand we can estimate an upper bound for the transaction between agent B and C. From above we know that

$$\pi_C = \rho_C^m(0) - \rho_C^m(N1_{\tau>T})$$

and we obtain, by using the entropic risk measure, the following result.

Theorem 7.2 *The indifference volume for the CAT-bond issue by agent B is given by:*

$$\pi_C = -\gamma_C ln \left((1-\hat{p})exp\left(-\frac{1}{\gamma_C}N\right) + \hat{p} \right). \tag{7.3}$$

note that π_C if and only if $N > 0$ and $\frac{\partial \pi_C}{\partial N} > 0$ for any N.

Now it remains to determine the nominal amount of the transaction. Agent B has to solve the optimization problem (7.1) under the pricing constraint imposed by agent A and agent C which are given by (7.2) and (7.3).

Theorem 7.3 *The optimal nominal amount for the simple CAT-bond is (up to a constant) given as:*

$$N = \frac{\gamma_C}{\gamma_A + \gamma_C} X \tag{7.4}$$

We see that agent A retains a part of the risk X, which increases with his risk tolerance, for a given risk tolerance of agent C. Agent B makes the transaction feasible and he receives the difference between π_C and π_A. This difference would not exist if agent A deals directly with agent C.

Proof:

Using the optimal pricing rule together with the translation invariance property of the risk measure we obtain

$$\min_N \left\{ \rho_A^m(-X1_{\tau\leq T} - N1_{\tau<T}) - \rho_A^m(-X1_{\tau\leq T}) + \rho_C^m(N1_{\tau>T}) - \rho_C^m(0) \right\}.$$

Using the entropic risk measure we get

$$\min_N \left\{ \rho_A^m(-X1_{\tau\leq T} - N1_{\tau<T}) - \rho_A^m(-X1_{\tau\leq T}) + \rho_C^m(N1_{\tau>T}) \right\}$$

because

$$\rho_C^m(0) = 0.$$

Inserting (7.2) and (7.3) we obtain

$$\min_N \left\{ \gamma_A \ln \left(\frac{(1-\hat{p})\exp\left(\frac{1}{\gamma_A}N\right) + \hat{p}\exp\left(\frac{1}{\gamma_A}X\right)}{(1+\hat{p}) + \hat{p}\exp\left(\frac{1}{\gamma_A}X\right)} \right) + \gamma_C \ln \left((1-\hat{p})\exp\left(-\frac{1}{\gamma_C}N\right) + \hat{p} \right) \right\}$$

which is equivalent to

$$\min_N \left\{ \gamma_A \ln \left((1-\hat{p})\exp\left(\frac{1}{\gamma_A}N\right) + \hat{p}\exp\left(\frac{1}{\gamma_A}X\right) \right) + \gamma_C \ln \left((1-\hat{p})\exp\left(-\frac{1}{\gamma_C}N\right) + \hat{p} \right) \right\}$$

We write the first order optimality condition:

$$\gamma_A \frac{\frac{1}{\gamma_A}(1-\hat{p})\exp\left(\frac{1}{\gamma_A}N\right)}{(1-\hat{p})\exp\left(\frac{1}{\gamma_A}N\right) + \hat{p}\exp\left(\frac{1}{\gamma_A}X\right)} + \gamma_C \frac{-\frac{1}{\gamma_C}(1-\hat{p})\exp\left(-\frac{1}{\gamma_C}N\right)}{(1-\hat{p})\exp\left(-\frac{1}{\gamma_C}N\right) + \hat{p}} = 0$$

By applying the reciprocal and using the logarithm the result follows immediately. Without proving we note that the second derivative w.r.t. N is positive. □

To complete this section about the simple transaction we show a result which ensures the feasibility of the transaction. We will see that agent B earns a positive fee for organizing the CAT-bond.

Theorem 7.4 *The indifference buyer's price of agent C is larger than the indifference seller's price of agent A.*

Proof:
We consider the difference $\pi_A - \pi_C$ and use the fact that the optimum $N = \alpha X$, with $\alpha = \frac{\gamma_C}{\gamma_A + \gamma_C}$. Therefore,

$$\begin{aligned}
\pi_A - \pi_C = {} & \gamma_A \ln \left(\frac{(1-\hat{p})\exp\left(\frac{1}{\gamma_A}\alpha X\right) + \hat{p}\exp\left(\frac{1}{\gamma_A X}\right)}{(1-\hat{p}) + \hat{p}\exp\left(\frac{1}{\gamma_A X}\right)} \right) \\
& + \gamma_C \ln \left((1-\hat{p})\exp\left(-\frac{1}{\gamma_C}\alpha X\right) + \hat{p} \right) \\
= {} & \gamma_A \ln \left((1-\hat{p})\exp\left(\frac{1}{\gamma_A}(\alpha - 1)X\right) + \hat{p} \right) - \gamma_A \ln \left((1-\hat{p})\exp\left(-\frac{1}{\gamma_A}X\right) + \hat{p} \right) \\
& + \gamma_C \ln \left((1-\hat{p})\exp\left(-\frac{1}{\gamma_C}\alpha X\right) + \hat{p} \right)
\end{aligned}$$

Set $\alpha = \frac{\gamma_C}{\gamma_A+\gamma_C}$, and we obtain

$$\begin{aligned}\pi_A - \pi_C &= \gamma_A \ln\left((1-\hat{p})\exp\left(-\frac{1}{\gamma_A+\gamma_C}X\right)+\hat{p}\right) - \gamma_A \ln\left((1-\hat{p})\exp\left(-\frac{1}{\gamma_A}X\right)+\hat{p}\right)\\ &+ \gamma_C \ln\left((1-\hat{p})\exp\left(-\frac{1}{\gamma_A+\gamma_C}X\right)+\hat{p}\right)\\ &= (\gamma_A+\gamma_C)\ln\left((1-\hat{p})\exp\left(-\frac{1}{\gamma_A+\gamma_C}X\right)+\hat{p}\right)\\ &- \gamma_A \ln\left((1-\hat{p})\exp\left(-\frac{1}{\gamma_A}X\right)+\hat{p}\right)\end{aligned}$$

Note that the function

$$\gamma \longrightarrow \gamma \ln\left((1-\hat{p})\exp\left(-\frac{1}{\gamma}X\right)+\hat{p}\right)$$

is strictly decreasing for $\gamma > 0$. Using the fact that $\gamma_A + \gamma_C > \gamma_A$ and the result

$$\pi_A - \pi_C < 0$$

follows. □

7.4 Analysis of Case B

Let us now analyse the hybrid transaction, whose structure is given in Figure 7.2. Analogous to the simple CAT-bond π_A^h denotes the price for the CAT-bond issued by agent A and purchased by agent B and π_C^h denotes the price of the hybrid CAT-bond issued by agent B. The nominal amount N^h of the CAT-bond differs in general from the nominal amount N of simple CAT-bond transaction. In this case the role of agent B as intermediary differs from the role in the simple case. Here agent B has to handle with risk management considerations. The additional payment H deters agent A from the direct issue of the bond. The hybrid product which is received by agent C includes the sale of a call, paying H if there is no natural catastrophe and a market boom and the purchase of a put, which is paying H if there is a catastrophe and a market crash. The initial price of the options coincide, and hence it only remains to exchange the contingent payoff at maturity. This assumption imposes some constraint on the market events. If the transaction is completed, no participator of the transaction is simultaneously exposed to both risks. For agent A, the risk of a catastrophe remains, if $N^h < X$ and agent C has to take over the remaining risk of the catastrophe against a random suitable compensation. Here, analogous to the simple case, agent B does not retain any risk and interacts only as intermediary and transfers only the cash flows. Let us describe the payoffs and the transaction more precisely.

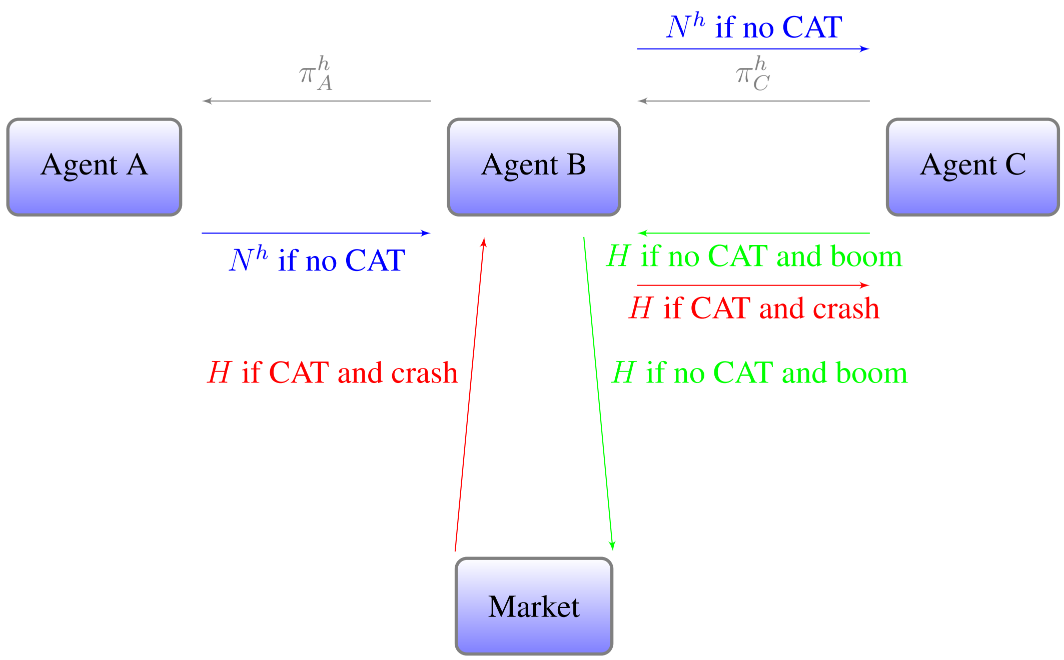

Figure 7.2: Hybrid CAT-Bond

7.4.1 The model

Note that the hybrid CAT-bond depends on the joint occurrence of the catastrophe and a crash, and the market boom and non-occurrence of the catastrophe. To reflect this effect in the model we have to introduce the joint distribution. We note as in the simple model the random time of occurrence of a natural catastrophe by τ, τ_c will represent the occurrence of a market crash and τ_b that of a market boom.
For simplification we assume that the occurrence of a natural catastrophe and the market-based events are independent. This means that if the catastrophe and then a market crash occurs before the time horizon T, the amount H will be paid. A further assumption is that the events "market boom" and "market crash" are disjoint, which means that they cannot occur both before T. To justify this assumption the maturity T of the transaction is close.
Under the minimal entropy probability measure $\hat{\mathbb{Q}}$ the joint distribution of the events can be written by:

Occurrence	Probability	States of nature	Probability
$T < \tau$	$1 - \hat{p}$	$\tau_b < T < \tau$	$(1-\hat{p})q_b$
		$T < \tau_b < \tau$ and $T < \tau < \tau_b$	$(1-\hat{p})(1-q_b)$
$T \geq \tau$	$\hat{p}$	$\tau_c < \tau < T$ and $\tau < \tau_c < T$	$\hat{p}q_c$
	$\hat{p}$	$\tau < T < \tau_c$	$\hat{p}(1-q_c)$

with $q_b = \hat{\mathbb{Q}}(\tau_b < T)$ and $q_c = \hat{\mathbb{Q}}(\tau_c < T)$. We now can summarize the payoffs of the different agents in Table 7.2.

Occurrence	Probability	Agent A	Agent B	Agent C
$\tau_b < T < \tau$	$(1-\hat{p})q_b$	$\pi_A^h - N^h$	$\pi_C^h - \pi_A^h$	$-\pi_C^h + N^h - H$
$T < \tau_b < \tau$ and $T < \tau < \tau_b$	$(1-\hat{p})(1-q_b)$	$\pi_A^h - N^h$	$\pi_C^h - \pi_A^h$	$-\pi_C^h + N^h$
$\tau_c < \tau < T$ and $\tau < \tau_c < T$	$\hat{p}q_c$	$\pi_A^h - X$	$\pi_C^h - \pi_A^h$	$-\pi_C^h + H$
$\tau < T\tau_c$	$\hat{p}(1-q_c)$	$\pi_A^h - X$	$\pi_C^h - \pi_A^h$	$-\pi_C^h$

Table 7.2: Payoffs

We previously mentioned that both options should have the same initial price, computed under the pricing measure $\hat{\mathbb{Q}}$. Note that the market is not complete, which is equivalent to the non-uniqueness of the equivalent martingale measure. We assume that both options are traded and prices by the risk neutral pricing rule with respect to the minimal entropy probability measure. We should have that

$$\mathbb{E}_{\hat{\mathbb{Q}}}\left[H1_{\tau,\tau_c\leq T}\right] = \mathbb{E}_{\hat{\mathbb{Q}}}\left[H1_{\tau_b\leq T\leq\tau}\right]$$

which is equivalent to

$$\hat{\mathbb{Q}}(\tau,\tau_c \leq T) = \hat{\mathbb{Q}}(\tau_b \leq T \leq \tau).$$

Using the independence of the market and the occurrence of a natural catastrophe, we can write the condition as follows:

$$\hat{p}q_c = (1-\hat{p})q_b. \tag{7.5}$$

This condition influences the definition of the events market-crash and market-boom. By defining a crash as a fall of a stock market index below a certain level L_c implies the characterization of another level L_b for the index such that the event market-boom is defined by exceeding this level. A market-boom corresponds to a particular market-crash and vice versa by condition (7.5).

We now analyse the transactions and the related cash flows of the hybrid CAT-bond structure. The transaction will proceed in the following way. Agent A determines the volume of the transaction π_A^h with agent B first. As in the simple case, agent C determines the price π_C^h. The role of agent B is in the hybrid transaction more responsible, because he advises the whole transaction. He fixes the optimal nominal amount N^h of both transactions and the payoff H of the put and call option. As in the simple case, agent B wants to maximize the fees from the transaction. Therefore we can write the optimization program for agent B as follows:

$$\begin{aligned}
&\max_{N^h,H,\pi_A,\pi_c} (\pi_C^h - \pi_A^h) \qquad (7.6)\\
&s.t:\\
&\rho_A^m\left(-N^h 1_{\tau>T} + \pi_A^h - X1_{\tau\leq T}\right) \leq \rho_A^m\left(-X1_{\tau\leq T}\right)\\
&\rho_C^m\left(N^h 1_{\tau>T} - \pi_C^h + H1_{\tau,\tau_c\leq T} - H1_{\tau_b\leq T\leq\tau}\right) \leq \rho_C^m(0)
\end{aligned}$$

We already know that the price of each part of the transaction represents the volume of capital flow in the CAT-bond market. We give the pricing rules of both transactions in the entropic framework in the following two theorems below. The lower bound of the transaction between agent A and agent B is given by

$$\pi_A^h = \rho_A^m \left(-N^h 1_{\tau>T} - X1_{\tau\leq T}\right) - \rho_A^m \left(-X1_{\tau\leq T}\right).$$

Theorem 7.5 *The indifference volume for the CAT-bond issue by agent A is given by:*

$$\pi_A^h = \gamma_A ln\left(\frac{(1-\hat{p})exp\left(\frac{1}{\gamma_A}N^h\right) + \hat{p}exp\left(\frac{1}{\gamma_A}X\right)}{(1-\hat{p}) + \hat{p}exp\left(\frac{1}{\gamma_A}X\right)}\right) \tag{7.7}$$

Note that $\pi_A^h > 0$ *if* $N^h > 0$.

The upper bound of the transaction between agent B and agent C is given by

$$\pi_C^h = \rho_C^m(0) - \rho_C^m\left(N^h 1_{\tau>T} + H1_{\tau,\tau_c\leq T} - H1_{\tau_b\leq T\leq\tau}\right).$$

Using the entropic risk measure we obtain:

Theorem 7.6 *The indifference volume for the hybrid CAT-bond issue by agent B is given by:*

$$\pi_C^h = -\gamma_C ln\left((1-\hat{p})\left(q_b exp\left(\frac{1}{\gamma_C}(h-N^h)\right) + (1-q_b)exp\left(-\frac{1}{\gamma_C}N^h\right)\right)\right. \tag{7.8}$$
$$\left.+\hat{p}q_c exp\left(-\frac{1}{\gamma_C}H\right) + \hat{p}(1-q_c)\right).$$

It remains to determine the nominal amount N^h and the hybrid amount H. Agent B has to maximize problem (7.6) by considering (7.7) and (7.8). We express the problem as follows:

$$\max_{H,N^h}(\pi_C^h - \pi_A^h)$$
$$s.t.:$$
$$\pi_C^h = -\gamma_C \ln\left((1-\hat{p})\left(q_b \exp\left(\frac{1}{\gamma_C}(H-N^h)\right) + (1-q_b)\exp\left(-\frac{1}{\gamma_C}N^h\right)\right)\right.$$
$$\left.+\hat{p}q_c \exp\left(-\frac{1}{\gamma_C}H\right) + \hat{p}(1-q_c)\right)$$
$$\pi_A^h = \gamma_A \ln\left(\frac{(1-\hat{p})\left(\frac{1}{\gamma_A}N^h\right) + \hat{p}\exp\left(\frac{1}{\gamma_A}X\right)}{(1-\hat{p}) + \hat{p}\exp\left(\frac{1}{\gamma_A}X\right)}\right)$$

First we optimize this problem in N^h for any given H and we have

$$\max_{N^h}\left\{\pi_C^h(N^h,H) - \pi_A^h(N^h,H)\right\}.$$

Theorem 7.7 *The optimal nominal amount of the transaction is given by:*

$$N^h = \frac{\gamma_C}{\gamma_A + \gamma_C} X + \frac{\gamma_A}{\gamma_A + \gamma_C} \mathcal{L}_C(q_b, q_c, H)$$

where

$$\mathcal{L}_C(q_b, q_c, H) = \rho_C^m(-H 1_{\tau_b \le T}) - \rho_C^m(H 1_{\tau_c \le T})$$

Note that the nominal amount of the hybrid CAT-bond depends on two terms. The first one is the same as in the simple CAT-bond. The second term links the financial market conditions and the contingent payment H in a non-linear way with the nominal. If H is positive, the second term is also positive and the nominal amount N^h is larger than the nominal amount N of the simple transaction.

Proof:

This proof can be done by the same arguments used in the Proof of Theorem 7.3, using the first order optimality condition and equating to zero and the result follows. □

We solve the maximization problem with respect to H and obtain the following result.

Theorem 7.8 *The optimal hybrid amount to be paid to the investor when both a crash and a catastrophe occur before maturity is given by:*

$$H = \frac{1}{2} N^h.$$

Proof:

Using the first order optimality condition and equating to zero, the result follows. □

Bibliography

[1] Abrahamsen P.: A Review of Gaussian Random Fields and Correlation Functions, *Technical Report 917, Norwegian Computing Center, second Editon,* 1997.

[2] Abramowitz M. and Stegun I.A.: Handbook of Matematical Functions, *New York: Dover Publications, http://www.math.ucla.edu/ cbm/aands/,* 1972.

[3] Alaton P., Djehiche B. and Stillberger D.: On Modelling and Pricing Weather Derivatives, *Appl. Math. Finance, 9, 1, 1-20,* 2002.

[4] Ankirchner S., Imkeller P. and Dos Reis G.: Pricing and Hedging of Derivatives based on non-tradeable Underlyings, *Preprint, HU Berlin,* 2007.

[5] Ankirchner S., Imkeller P. and Popier A.: Optimal cross hedging of insurance derivatives, *Stochastic Analysis and Applications, 26, 4, 679-709,* 2008.

[6] Banks E.: Alternative risk transfer, *John Wiley & Sons, Ltd,* 2004.

[7] Banks E.: Catastrophic Risk, *John Wiley & Sons, Ltd,* 2005.

[8] Barndorff-Nielsen O.E. and Shephard N.: Financial Volatility: Stochastic Volatility and Lévy based Models, *Cambridge University Press,* 2004.

[9] Barth A., Benth F.E. and Potthoff J.: Hedging of Spatial Temperature Risk with Market-Traded Futures *E-print No 2, http://folk.uio.no/fredb/,* 2008.

[10] Barrieu P. and El Karoui N.: Pricing, hedging and optimally designing derivatives via minimization of risk measures, *forthcoming in Volume on Indifference Pricing (ed. R. Carmona), Princeton University Press,* 2005.

[11] Barrieu P. and Louberge H.: Hybrid Cat-bonds, *Swiss Finance Institute, Research Paper Series, 07, 27,* 2007.

[12] Basawa I. and Prasaka Rao B.L.S.: Statistical Inference for Stochastic Processes, *Academic Press*, 1980.

[13] Benth F.E. and Šaltytė-Benth J.: The Volatility of Temperature and Pricing of Weather Derivatives, *Quantitative Finance, 7, 5, 553-561,* 2007.

[14] Benth F.E. and Šaltytė-Benth J.: Stochastic Modelling of Temperature Variations with a View Towards Weather Derivatives, *Appl. Math. Finance, 12, 1, 53-85,* 2005.

[15] Benth F.E., Šaltytė-Benth J. and Koekebakker S.: Putting a Price on Temperature, *Scandinavian Journal of Statistics, 34, 4, 746-767,* 2007.

[16] Benth F.E., Šaltytė-Benth J. and Jalinskas P.: A Spatial-Temporal Model for Temperature with Seasonal Variance *Journal of Applied Statistics, 34, 7, 823-841,* 2007.

[17] Bhowan A.: Temperature Derivatives, *Honours Project, University of Wiwatersrand*, 2003.

[18] Bibby B.M. and Sørensen M.: Martingal Estimation Functions for Discretely Observed Diffusion Processes, *Bernoulli, 1, 1-2, 17-39,* 1995.

[19] Brockett P.L., Wang M., Yang C. and Zou H.: Portfolio Effects and Valuation of Weather Derivatives, *The Financial Review, 41, 55-76,* 2006.

[20] Brockwell P.K. and Davis R.A.: Time Series: Theory and Methods, *Springer,* 1990.

[21] Campbell S.D. and Diebold F.X.: Weather forecasting for weather derivatives, *Journal of the American Statistical Association, 100, 6-16,* 2005.

[22] Cao M. and Wei J.: Weather Derivatives Valuation and Market Price of Weather Risk, *Journal of Futures Market, 24, 11, 1065-1089,* 2004.

[23] Cao M. and Wei J.: Equilibrium Valuation of Weather Derivatives, *Department of Economics, Queen's University, Working Paper,* 2001.

[24] Carmona R. and Durrleman V.: Pricing and Hedging Spread Options, *SIAM Review, 45, 4, 627-685,* 2003.

[25] Chaumont S., Imkeller P. and Müller M.: Equilibrium trading of climate and weather risk and numerical simulations in a markovian framework, *SERRA, 20, 3, 184-205,* 2006.

[26] Cont R. and Tankov P.: Financial Modelling With Jump Processes, *Chapman & Hall/CRC London,* 2004.

[27] Da Prato G. and Zabcyk J.: Stochastic Equations in Infinite Dimensions, *Cambridge University Press,* 1992.

[28] Davis M.: Pricing Weather Derivatives by Marginal Value, *Quantitative Finance, 1, 3, 309-317,* 2001.

[29] Dornier F. and Querel M.: Caution to the wind, *Energy Risk Management, Weather Risk special report,* 2000.

[30] Fleming W.H. and Soner H.M.: Controlled markov processes and viscosity solutions, *Applications of Mathematics, Springer,* 1993.

[31] Föllmer H. and Schied A.: Convex Measures of risk and trading constraints, *Finance and Stochastics, 6, 4, 429-447,* 2002.

[32] Föllmer H. and Schied A.: Stochastic Finance: An introduction in discrete time *De Gruyter studies in Mathematics,* 2002.

[33] Frittelli M.: The Minimal Entropy Martingal Measure and the Valuation in Incomplete Markets *Mathematical Finance, 10, 1, 39-52,* 2000.

[34] Grossi P. and Kunreuther H.: Catastrophe Modeling: a new approach to risk managing, *Springer,* 2005.

[35] Henderson V.: Valuation of claims on nontraded assets using utility maximization, *Mathematical Finance, 12, 4, 351-373,* 2002.

[36] Hu Y., Imkeller P. and Müller M.: Utility maximization in incomplete markets, *Annals of Applied Probability, 15, 4, 1691-1712,* 2005.

[37] Ikeda N. and Watanabe S: Stochastic Differential Equations and Diffusion Processes, *North-Holland Mathematical Library,* 1981.

[38] Jewson S., Brix A. with Ziehmann C.: Weather Derivatives Valuation, *Cambridge University Press,* 2005.

[39] Kabas T.: Das Klima in Südösterreich 1961-2004: Die alpine Region Hohe Tauern und die Region Südoststeiermark im Vergleich, *Diplomarbeit, Karl-Franzens-Universität Graz,* 2005.

[40] Kazamaki N.: Continuous exponential martingales and BMO, *Lecture Notes in Mathematics Vol.: 1579, Springer,* 1994.

[41] Kloeden P.E., Platen E. and Schurz H.: Numerical Solution of SDE through Computer Experiments, *Springer,* 1994.

[42] Klebaner F.C.: Introduction to Stochastic Calculus with Applications, *Imperial College Press,* 1999.

[43] Kobylanski M.: Backward stochastic differential equations and partial differential equations with quadratic growth *The Annals of Probability, 28, 2, 558-602,* 2000.

[44] Mraoua M. and Bari D.: Temperature stochastic Modeling and Weather Derivatives pricing: empirical Study with Moroccan Data, *Afrika Statistika, 2, 1, 22-43,* 2007.

[45] Øksendal B.: Stochastic differential equations, an introduction with apllications, *6th Edition, Springer,* 2007.

[46] Pardoux E. and Peng S.: Adapted solution of a backward stochastic differential equation, *Systems Control Letters, 14, 55-61,* 1990.

[47] Platen E. and West J.: A Fair Pricing Approach to Weather Derivatives, *Asia-Pacific Financial Markets, 11, 1,* 2004.

[48] Riesz F. and Sz.-Nagy B.: Functional Analysis *Dover, New York. Originally published by Frederick Ungar Publ. Co, New York, 1955,* 1995.

[49] Turvey C.G.: The Pricing of Degree-Day Weather Options, Agricultural Finance Review, Spring 2005, 59-85, 2005.

[50] Werner D.: Funktionalanalysis, *Springer,* 1997.

[51] Xu W., Odening M. and Musshoff O.: Indifference Pricing of Weather Derivatives, *American Journal of Agricultural Economics, 90, 4, 979-993,* 2008.

[52] Zapranis A. and Alexandridis A.: Weather Derivatives Pricing: Modeling the Seasonal Residual Variance of Ornstein-Uhlenbeck Temperature Process with Neural Networks, *Conference Paper, EANN, Thessaloniki,* 2007.