Pricing and Hedging of Mortgage Options

Ivo Bodin

February-August 2017

Msc Thesis

Under the supervision of Sabrina Wandl and Dr. Peter Spreij and submitted to the Board of Examiners as a final project for the degree of

MSc in Stochastics and Financial Mathematics

at the Universiteit van Amsterdam.

Date of public defence: August 22, 2017

Members of the Exam Committee:
Prof. Dr. R. (Sindo) Nunez Queija
Prof. Dr. P.J.C. (Peter) Spreij
Dr. A. (Asma) Khedher
Contents

1 Introduction 4
  1.1 Background .................................................. 4
  1.2 Research question ........................................... 4
  1.3 Relevancy of the problem ..................................... 4
  1.4 Structure of the thesis ........................................ 5

2 Basic description of different types of mortgages 6
  2.1 Available amortization schemes ............................... 6
    2.1.1 Bullet loan ................................................ 6
    2.1.2 Level paying or Annuity loan .......................... 6
    2.1.3 Linear loan ............................................... 7
  2.2 Additional choices ............................................. 7
    2.2.1 Fixed versus floating interest rate .................... 7
    2.2.2 Hypotheekrente-aftrek .................................. 7
    2.2.3 Nationale Hypotheek Garantie ........................... 8

3 Basic description of different types of options on mortgages 9
  3.1 Loan-to-Value option ......................................... 9
  3.2 Meeeneemoptie ................................................ 9
  3.3 Rentemiddeling ............................................... 10
  3.4 Pipeline ..................................................... 10

4 Interest rate models 11
  4.1 Merton’s model ............................................... 11
    4.1.1 Advantages of Merton’s model ........................ 11
    4.1.2 Drawback of Merton’s model ............................ 11
  4.2 Vasicek model ............................................... 11
    4.2.1 Estimation within the Vasicek Model .................. 13
    4.2.2 Advantages of the Vasicek Model ....................... 13
    4.2.3 Drawbacks of the Vasicek Model ........................ 14
  4.3 Hull-White model ............................................ 14
    4.3.1 Advantages of the Hull-White model .................. 14
    4.3.2 Drawback of the Hull-White model ..................... 14
  4.4 Cox–Ingersoll–Ross (CIR) model ............................ 14
    4.4.1 Shifted CIR-model ........................................ 15
    4.4.2 Extended shifted CIR-model ............................ 15
    4.4.3 Estimation within the extended CIR-model ............ 15

5 Pricing techniques 16
  5.1 Mathematical framework ...................................... 16
  5.2 Hedging/Replicating portfolio in complete markets ......... 17
    5.2.1 Delta hedge .............................................. 17
    5.2.2 Replicating portfolio ................................... 18
    5.2.3 Black-Scholes approach ................................ 18
  5.3 Monte Carlo .................................................. 20
    5.3.1 Generating the underlying process ...................... 20
    5.3.2 Monte Carlo estimation of derivatives ................. 20
    5.3.3 Confidence intervals ..................................... 21
    5.3.4 Quasi-Monte Carlo method ............................... 21
    5.3.5 Variance reduction methods .............................. 24
  5.4 Finite difference methods .................................... 25
    5.4.1 Transforming the partial differential equation ....... 26
5.4.2 Boundary conditions ................................................. 26
5.4.3 Solving the partial differential equation ....................... 27
5.5 Lattice model ......................................................... 28
  5.5.1 Binomial tree ................................................... 28
  5.5.2 Trinomial tree ................................................... 28
  5.5.3 Similarities between binomial and trinomial trees .......... 29
  5.5.4 Hull-White trinomial tree ................................... 29
5.6 Linear regression .................................................... 32
  5.6.1 Generalized linear regression ................................. 32
  5.6.2 Estimation within the binomial logit model ................. 33
  5.6.3 Extension to the multinomial logit model .................... 35
6 Quadratic hedging in incomplete markets ......................... 36
  6.1 Lévy theory ......................................................... 36
  6.2 Market incompleteness ........................................... 38
  6.3 Local risk-minimization approach ................................ 39
  6.4 Mean-variance hedging ............................................ 40
7 Model choices and implementation .................................... 42
  7.1 Calibrating the Hull-White model ............................... 42
    7.1.1 Finding the mean-reversion and volatility parameters .. 43
    7.1.2 Finding confidence intervals ............................... 43
    7.1.3 Finding the long term mean function ...................... 44
  7.2 Calibrating the CIR-model ....................................... 46
    7.2.1 Finding the mean-reversion and volatility parameters .. 47
    7.2.2 Finding the long term mean function ...................... 47
  7.3 Estimations and calculations regarding the LtV process .... 47
    7.3.1 Modelling the outstanding notional ....................... 47
    7.3.2 Modelling the house price .................................. 49
    7.3.3 Calculating the pay-off of the LtV option ................. 51
  7.4 Hull-White Monte Carlo with variance reduction methods ... 52
  7.5 CIR Monte Carlo with variance reduction methods ............ 53
  7.6 Hull-White tree .................................................... 53
8 Results of numerical experiments .................................... 56
  8.1 LtV ................................................................. 56
    8.1.1 Log-normal versus hyperbolic distribution .............. 56
    8.1.2 Practical example ............................................ 56
  8.2 Meeneemoptie ....................................................... 56
    8.2.1 Meeneemoptie as an American option ...................... 61
    8.2.2 Conditioned on the time of moving ....................... 61
  8.3 Rentemiddeling ................................................... 64
  8.4 Pipeline .......................................................... 64
    8.4.1 Without lock-or-lower ..................................... 65
    8.4.2 With lock-or-lower ......................................... 65
    8.4.3 Comparison with current used value ...................... 66
9 Conclusion ............................................................. 68
  9.1 Further points of action ......................................... 68
  9.2 Model risk ........................................................ 68
    9.2.1 Choice of interest rate model .............................. 68
    9.2.2 Sensitivity to the input parameters ....................... 68
    9.2.3 Sensitivity to the calibration data ....................... 68
  9.3 Advice for the bank .............................................. 71
1 Introduction

Buying a house is for most people, both financially and personally, a big step in their life. Some are able to directly pay the house price and the number of people that do this is actually rising according to [1]. However most people need a mortgage in order to finance their new home. A mortgage contract is specified by many different things, which depending on the situation are either perks or drawbacks. Often cash flows and regulations within the mortgage depend on the future and therefore create risk. This thesis will focus on the interest rate risk created by different mortgage options, namely the loan-to-value option, pipeline option, 'rentemiddeling' and the 'meeneemoptie'. The last two are Dutch for respectively 'interest rate averaging' and 'take along' option. All will be further explained in section 3. Throughout this thesis, some concepts are named by their Dutch as their English counterpart is not available. A short Dutch-English dictionary can be found in table 1.

<table>
<thead>
<tr>
<th>Dutch</th>
<th>English</th>
</tr>
</thead>
<tbody>
<tr>
<td>Meeneemoptie</td>
<td>Take along option</td>
</tr>
<tr>
<td>Rentemiddeling</td>
<td>Interest rate averaging</td>
</tr>
<tr>
<td>Hypotheekrente aftrek</td>
<td>Mortgage rate deduction</td>
</tr>
<tr>
<td>Risico-opslag</td>
<td>Risk addition</td>
</tr>
<tr>
<td>Nationale Hypotheek Garantie</td>
<td>National Mortgage Guarantee</td>
</tr>
<tr>
<td>Oversluiten</td>
<td>Refinancing</td>
</tr>
</tbody>
</table>

Table 1: Dutch-English dictionary.

1.1 Background

Mortgage options are often featured on the national news. A lot of people try to profit from the historically low current interest rate. An example of this is the rise of mortgages with longer fixed interest rate periods, as explained in [2]. However a warning is also present about the meeneemoptie, as a lack of this option means that clients are often better off with a shorter fixed interest rate period. The low current interest rate also creates the possibility of rentemiddeling, as [3] states that around 20% of mortgages can profit from rentemiddeling. However, not one in five clients of ABN AMRO actually used rentemiddeling as according to [4] rentemiddeling was only used by 5,000 out of 800,000 clients. Another hot topic today is the risico-opslag, which is Dutch for risk-addition. Risico-opslag is very closely related to the loan-to-value option. Articles like [5] argue that the current risico-opslag is sometimes too high as it does not reduce in case the risk reduces. At ABN AMRO the risico-opslag drops when the risk decreases, but it is the responsibility of the client to show this.

1.2 Research question

The research question of the thesis is: What is the risk is of the mortgage options described above and how can the bank best manage this risk? In order to say something about this all mortgage options are quantified and different pricing models are developed. These models can then be used to simulate outcomes or sometimes even provide an entire hedging strategy.

1.3 Relevancy of the problem

As banks are risk averse, an obvious advantage of the analysis of mortgage options is the understanding and reduction of risk. However there is also a less obvious benefit. Most pricing models assume that both parties behave optimally. In real life this is not the case as clients are no experts. This means that currently the risk can be covered by an analysis of behavioural data.
via a regression model and this is what is done in the financial industry. However, it is possible that a third party helps clients with their financial problems. This is already done by for example www.ikbenfrits.nl. The pricing models give an upper limit to the profit of the clients and can therefore be used to calculate a worst case scenario. The generalized regression models are an exception to this, as they do not assume that clients behave optimally, but behave like they did in the past. This is what is now done in practice, but banks have to be aware that past data is still representative for the behaviour of the future. This risk is known as model risk.

1.4 Structure of the thesis

After this introduction, section 2 describes the general setup of a mortgage and the different mortgage options are explained in section 3. All pricing models need an interest model as input and the theory of these is covered in section 4. Section 5 provides a theoretical background to all relevant pricing models. The implementation of these pricing models is described in section 6 and the results of numerical experiment in section 7. Finally, the conclusion is given in section 8.
2 Basic description of different types of mortgages

A mortgage is in essence nothing else than a loan. During the time of the contract, interest rate and notional payments are made to make sure that the notional is paid back at maturity. The charged interest rate is known as the mortgage rate. The mortgage rate is calculated by adding a spread to the relevant interest rate and is represented by a coupon $c$. This spread represents the profit the bank makes on the mortgage. There exist three main types of mortgages, all characterized by their own amortization schedule. Let $i \in \{1, \ldots, M\}$ be an index that indicates the amount of payments made. An amortization schedule consists of the times $T_i$ and height of interest rate payments $I_i$ and notional payments $P_i$. The outstanding notional at each time point is given by $N_i$. In practice, payments are made every month and the interest rate payments equal to $\frac{c N_i}{12}$.

The different amortization schemes are explained below. The second part of this section focuses on additional choices clients have to make when entering a mortgage.

2.1 Available amortization schemes

As with every loan, at maturity the entire notional has to be paid back. There are amortization schemes to accomplish this, with different total interest payments at maturity. Clients can pick between three schemes. The choice is mainly based on when they want to pay, as later payments lead to higher interest payments. About 70% of the portfolio consists of bullet loans. Therefore most experiments and results consider bullet loans. Luckily, most models can easily be adjusted to work for any amortization scheme.

2.1.1 Bullet loan

The distinctive property of a bullet loan is that the only notional payment is at maturity. Consequently, the height of this payment is equal to the entire notional. This implies that all other payments consist of only an interest rate part. The amortization scheme is given by:

$$I_i + P_i = \begin{cases} \frac{c}{12}N_0 & \text{if } i \in \{1, \ldots, M - 1\} \\ \left(\frac{c}{12} + 1\right)N_0 & \text{if } i = M \end{cases}$$

Because every interest payment regards the entire notional, this amortization scheme is beneficial for the bank. Without further information, it seems that this type of mortgage is never best for the client. However, often clients use a more extensive construction. Every month they put some money into a savings account that guarantees an interest rate equal to the mortgage rate or they use the money to finance a savings insurance. Furthermore, most clients take advantage of the hypotheekrente-aftrek. These regulations make this type of loan an interesting option for both bank and client. Unfortunately for newer clients, the government is already restricting the possibilities of this regulation.

2.1.2 Level paying or Annuity loan

As the name suggests, a level paying loan characterizes itself by the fact that monthly payments are constant. As the interest rate payments depend on the outstanding notional, this implies that the interest payments decrease and the notional payments increase during the lifetime of the mortgage. The amortization scheme is given by:

$$I_i + P_i = \frac{N_0}{a_c}$$

where the value of $a_c$ is given by:

$$a_c = \frac{1 - \left(1 + \frac{c}{12}\right)^{-M}}{\frac{c}{12}}.$$
As the interest rate payments decrease over time, the advantage given by the hypotheekrente-aftrek does also decrease. Since every month a part of the notional is paid back, an advantage for both client and the bank is that there is little risk that the loan is not repaid fully at maturity. An advantage over linear loan is that the first payments are not the highest.

2.1.3 Linear loan

The defining property of a linear loan is that the notional payments are equal to the notional divided by the number of payments. A simple result of this is that the notional payments are constant. Consequently, the interest rate payments and therefore the total payments decrease linear in time. An advantage is that this results in less total interest payments. The amortization scheme is given by:

\[
I_t + P_t = \frac{c}{12} N_{t-1} + \frac{N_0}{M} = \frac{c}{12} \left(1 - \frac{i - 1}{M}\right) N_0 + \frac{N_0}{M}
\]

As with the level paying loan, the advantage given by the hypotheekrente-aftrek decreases in time. On the other hand, this amortization scheme makes sure that the notional is paid back at maturity.

2.2 Additional choices

A mortgage is not only defined by the amortization scheme. The coupon for example is input in the amortization scheme and depends on various choices. Also some regulations depend on the start time of the mortgage.

2.2.1 Fixed versus floating interest rate

The client has, for every mortgage type, the choice between fixed and floating interest rate. Fixed interest rate means that the coupon remains equal during a long period of time until the next interest rate reset date. This period is usually 10 years. If the client chooses to pay a floating interest rate, the coupon depends on the future interest rate. Paying floating interest rate is a risk for the client, since the client does not know what the payments will be in the future. The bank however does not have to take this risk into account since it receives the same floating interest rate. Therefore the difference in interest rate is constant and profit for the bank is guaranteed. In case the client pays a fixed mortgage rate, the situation is totally different. There is no uncertainty for the client since the future payments are deterministic. The uncertainty is now for the bank. It might be the case that the constant mortgage rate paid by the client is less than the current interest rate. This would result in a loss for the bank. To hedge this risk, the bank engages in swaps that trade floating rate coupons for fixed rate coupons. The received fixed rate coupons make sure that a profit is made on the fixed mortgage rate paid by the client.

2.2.2 Hypotheekrente-aftrek

The hypotheekrente-aftrek allows mortgage owners to deduct their interest payments from their taxable income. This is the main reason that most loans are bullet loans. It is initiated and supported by the Dutch government. The idea behind this principle is that tax is paid over income and that mortgage rate payments are regarded as a negative income. Since the hypotheekrente-aftrek results in less tax obtained by the government, not everyone is happy with this rule. Since 2013 rules are incorporated to reduce the advantage of the hypotheekrente-aftrek. However, since most mortgages in the portfolio of the bank started way before 2013, this regulation is still very relevant for client behaviour regarding mortgage types. The bank has to be aware, as a reduction of the advantage of the hypotheekrente-aftrek might lead to an increase in amortization schemes other than the bullet loan.
2.2.3 Nationale Hypotheek Garantie

The bank always has to keep in mind that there is a possibility that the client is not able to follow the amortization scheme. Reasons for this for example are that the client gets unemployed or divorced or in a worst case dies. In case this happens, the underlying property can be sold to obtain funds to repay the mortgage. However, this is not always enough to repay the entire notional. This risk differs per client and type of mortgage but can never be neglected. This risk is captured in an increase in the mortgage rate called risico-opslag. In some cases, the client can combine the mortgage with a Nationale Hypotheek Garantie (NHG). In that case the Stichting Waarborgfonds Eigen Woningen guarantees to pay back the remaining debt. The risk for the bank vanishes and therefore the client does not have to pay a risico-opslag. On the other hand the client pays the Stichting Waarborgfonds Eigen Woningen one time a fixed percentage (at the time of writing this is 1%) of the notional. The NHG only guarantees a maximum of 245,000 euro at the time of writing. The exact regulations are a bit more complicated, but not relevant for the scope of this thesis.
3 Basic description of different types of options on mortgages

This section thoroughly describes the properties of the mortgage options within the scope of this thesis. Both a practical background and a mathematical definition are provided.

3.1 Loan-to-Value option

Although almost always clients engage a mortgage in order to buy a house, the notional value is in general different from the price of the house. Often clients saved some money in order to pay a part of the house immediately. This is beneficial in two ways. Obviously, this payment decreases the notional value and therefore the total interest payments. The payment can be seen as a prepayment without any repercussions. Secondly, the risk for the bank decreases because of this payment. It is less likely that, for example in case the client wants to relocate, selling the house does not cover the remaining notional value. Banks therefore often block relocation events if the loan-to-value (LtV) is above 100%. However, in case of events like default or death, the bank will take a loss. The price of this risk is directly calculated through the client in the form of risico-opslag. This additional payment will decrease if the client pays part of the house using own funds. The risk for the bank that selling the underlying property does not guarantee repayment of the notional is captured in the loan-to-value ratio. This ratio is defined for every $t \in [0, M]$ as:

$$(LtV)_t = \frac{(Outstanding notional)_t}{(Value of the house)_t}$$

Recently, an upper limit to the LtV is introduced. It was common for clients to loan more than the value of the house in order to cover additional costs. In 2012 the maximum LtV was set to 106% and will be lowered with one percentage point every year until 2018 when the maximum LtV reaches 100%. Note that the LtV is not constant in time. Both notional payments and house price fluctuations will change the clients LtV ratio. It might be the case that the LtV drops below a certain percentage. In that case the risk for the bank is significantly lower and the client can choose profit from a lower risico-opslag. However, if the LtV increases, the client does not have to pay extra risico-opslag. The choice to match the risico-opslag to the current LtV is known as the loan-to-value option. As with every option, the loan-to-value option represents a non-negative value. At the time of writing, it is the responsibility of the client to show that the LtV is below a certain percentage and to match the risico-opslag accordingly. At interest rate reset dates, the bank matches the risico-opslag to the LtV at that time.

3.2 Meeneemoptie

The client’s current financial situation and personal preferences may be different than the financial situation and personal preferences at the start of the mortgage. This may help the client to decide to move to another house. The client will prepay the mortgage using the money obtained from selling the house and will engage into a new mortgage to buy the new house. In case selling the house is not enough to pay the outstanding notional, it might be difficult for the client to move. Since the new mortgage rate will depend on the new interest rate, it will most likely be different from the old mortgage rate. This is of course beneficial for the client if the mortgage rate decreased during that time period, but risks the possibility that the mortgage rate increased. The policy within ABN AMRO enables the client to use the old mortgage rate for the original notional in the new contract. If the notional of the new mortgage is higher than the notional of the original mortgage, the current mortgage rate is charged over the difference. The result is that over the original notional, the client profits if the current mortgage rate is higher than the original mortgage rate, but does not risk paying more than the original mortgage rate. This option is called the 'meeneemoptie', which translates in English to 'take along option'. The value of the meeneemoptie at a certain time point depends on the difference between the mortgage rate at
that time and the contractual mortgage rate as well as the remaining time left to profit from this
difference. Note that the value of the meeneemoptie does not depend on the new notional as it
covers only the original notional.

3.3 Rentemiddeling

Even in a fixed rate mortgage, in general the interest rate is not constant during the entire period.
Often interest reset times are specified, usually every 10 years. At these time point the mortgage
rate is set to the mortgage rate at that time and fixed until the next interest reset time. If the
current mortgage rate is lower than the contractual mortgage rate, this results in lower monthly
costs for the client. The client may not want to wait until the next interest reset time and profit
from the low interest rates immediately. This can be done using rentemiddeling. The fixed interest
rate will change to a value between this rate and the current interest rate. The bank calculates
the new fixed coupon with:

\[ c_{\text{new}} = c_{\text{current}} + \text{penalty} + \text{rentemiddelingsopslag} \]  

The current rentemiddelingsopslag is 0.2\% and the penalty is calculated to compensate for the
missed interest income. The \( \delta t_{\text{old}} \) and \( \delta t_{\text{new}} \) represent respectively the time left until the next
interest reset time following the old contract and the time left until next interest reset time
following the new contract. The \( \Delta c \) represents the difference between the contractual mortgage
rate and the current mortgage rate for the remaining period. The penalty \( \Pi \) is calculated as
follows:

\[ \Pi = \frac{\Delta c \Delta t_{\text{old}}}{\Delta t_{\text{new}}} \]

This makes sure that the bank does not take a loss as long as the contract is not terminated.
Rentemiddeling results in a short term reduction of the monthly payments. Therefore clients can
profit from this regulation if they are planning to move soon, as they leave the contract before fully
paying the penalty. To counter this, bank sometimes decides to add a relocation penalty when
the client decides to use rentemiddeling. ABN AMRO however, does not do this and therefore
the rentemiddelingsoptie comes at a price. A client considering rentemiddeling should always also
consider the prepayment event oversluiten. Oversluiten means to engage in a new mortgage and
use that mortgage to fully prepay the original mortgage. Both oversluiten and rentemiddeling
result in lower interest rate payments but have different consequences. ABN AMRO does not
advise clients to use rentemiddeling as it is almost never in best interest of the client.

3.4 Pipeline

Clients have some time to think about a mortgage offer. The time between the offer and the
time of accepting is called the offer period. Typical offer periods are three or nine months. If the
mortgage rate drops within the offer period, the client has the option to use this lower mortgage
rate as the mortgage rate in the contract. If the mortgage rate increases during the offer period,
the client can just use the initial offered mortgage rate. The right, but not the obligation, to use a
different mortgage rate is known as the pipeline option. There are two different kinds of pipeline
options. The first one allows the client to engage the contract at a time of their choosing, while
the second one takes the minimum mortgage rate within the offer period. This makes the first one
an American option and the second one a European option. The European option is known as the
lock-or-lower variant and is by construction more valuable. The American option does, in contrast
to the lock-or-lower variant, take some knowledge to use optimally, as it requires the client to pick
the best time to engage in the mortgage, depending on the interest rate path.
4 Interest rate models

The pay-off of all mortgage options within the scope of this thesis depends on the interest rate. Therefore an interest rate model is required. As negative interest rates are present in the current market, it is important that the interest rate model can handle these. This section describes and discusses the most common options. However, before starting, bond prices are introduced. A bond $P(0, T)$ is a contract that pays the holder one euro/dollar at maturity $T$. Bond prices for different maturities are quoted directly or indirectly in the market, mostly in the form of a yield or forward curve.

4.1 Merton’s model

This model assumes that the interest rate follows the equation:

$$r_t = r_0 + at + \sigma W_t$$

for a starting point $r_0$, drift $a$, volatility $\sigma$ and a Brownian motion $W_t$. From this equation it is easy to compute the mean:

$$\mu_t = \mathbb{E}[r_t] = \mathbb{E}[r_0 + at + \sigma W_t] = r_0 + at$$

and the variance:

$$\sigma^2_t = \text{Var}(r_t) = \text{Var}(r_0 + at + \sigma W_t) = \sigma^2 \text{Var}(W_t) = \sigma^2 t$$

Historical interest data can be used to calculate these moments and in the process determine $a$ and $\sigma$.

4.1.1 Advantages of Merton’s model

Due to the simple model equation, estimation and implementation is very easy. From the differential equation, it is also clear that Merton’s model allows negative interest rate. While this used to be a drawback of the model, currently negative interest rates are observed in the market.

4.1.2 Drawback of Merton’s model

There is not enough freedom in the model to match a given interest rate structure. There is more information available in the market than the first and second moment of the interest rate. For example bond prices and the yield curve provide information about the future interest rate that cannot be captured within the model. One could try to solve this by allowing $a$ to be time dependent, but this is not done in practice.

4.2 Vasicek model

This model assumes that the short rate follows the stochastic differential equation:

$$dr_t = a(b - r_t) dt + \sigma dW_t$$

Where again $W_t$ is a Brownian motion. The value of $b$ describes the long term average level and the value of $a > 0$ describes how fast the process tend towards this level. This is also known as the mean reversion parameter. The value of $\sigma$ will represent the volatility of the interest rate. This differential equation can be solved by looking at:

$$d(e^{at} r_t) = e^{at} dr_t + r_t de^{at}$$

$$= e^{at} dr_t + ae^{at} r_t dt$$

$$= e^{at} (a(b - r_t) dt + \sigma dW_t) + ae^{at} r_t dt$$

$$= abe^{at} dt + \sigma e^{at} dW_t$$
This final expression makes it possible to integrate everything together with the initial condition \( r_{t=0} = r_0 \) to obtain:

\[
e^{at}r_t = r_0 + ab \int_0^t e^{as} \, ds + \sigma \int_0^t e^{as} \, dW_s
\]

Solving for \( r_t \) and calculating the first integral yields:

\[
r_t = e^{-at}r_0 + b(1 - e^{-at}) + \sigma \int_0^t e^{-a(t-s)} \, dW_s
\]

This direct expression shows that \( r_t \) is a Gaussian process and also allows for calculation of the parameters. The mean is given by:

\[
\mu_t = \mathbb{E}[r_t] = \mathbb{E} \left[ e^{-at}r_0 + b(1 - e^{-at}) + \sigma \int_0^t e^{-a(t-s)} \, dW_s \right] = e^{-at}r_0 + b(1 - e^{-at}) \quad (2)
\]

and the variance:

\[
\sigma_t^2 = \text{Var}(r_t) = \mathbb{E} \left[ \left( \sigma \int_0^t e^{-a(t-s)} \, dW_s \right)^2 \right] = \sigma^2 \int_0^t e^{-2a(t-s)} \, ds = \frac{\sigma^2}{2a} (1 - e^{-2at}) \quad (3)
\]

Note that the long term limits are given by:

\[
\lim_{t \to \infty} \mu_t = b
\]

\[
\lim_{t \to \infty} \sigma_t^2 = \frac{\sigma^2}{2a}
\]

These equations support the claim about the values of \( a \) and \( b \) made above. In order to use this model for numerical experiments, the stochastic differential equation needs to be discretized. This is done by introducing \( \delta > 0 \) and time points \( t_i = \delta i \). The Euler scheme of this model is then given by:

\[
r_{t_i} = r_{t_{i-1}} + a(b - r_{t_{i-1}})\delta + \sigma \sqrt{\delta} Z_i,
\]

where \( Z_i \) are standard normally distributed. Note that an Euler scheme is only correct in the first order and therefore one has to make sure that \( \delta \) is small enough. Fortunately, it is possible to make a discretization that is correct in every order. Observe that \( r_t \) is, as mentioned before, a Gaussian process and using the expression for the moments the following holds.

\[
r_t = \mu_t + \sigma_t W_t = r_{t_0} e^{-at} + b(1 - e^{-at}) + \sigma \sqrt{\frac{1 - e^{-2at}}{2a}} Z_i
\]

By taking \( t_i \) as a starting point instead of \( t_0 \) a different discretization scheme is obtained.

\[
r_{t_i} = r_{t_{i-1}} e^{-a\delta} + b(1 - e^{-a\delta}) + \sigma \sqrt{\frac{1 - e^{-2a\delta}}{2a}} Z_i
\]

Note that, as expected, a calculation of the first order Taylor approximation of above expression yields the Euler scheme.

\[
(5) = r_{t_{i-1}} e^{-a\delta} + b(1 - e^{-a\delta}) + \sigma \sqrt{\frac{1 - e^{-2a\delta}}{2a}} Z_i
\]

\[
\approx r_{t_{i-1}} (1 - a\delta) + b(1 - (1 - a\delta)) + \sigma \sqrt{\frac{1 - (1 - 2a\delta)}{2a}} Z_i
\]

\[
= r_{t_{i-1}} - a\delta r_{t_{i-1}} + ba\delta + \sigma \sqrt{\delta} Z_i
\]

\[
= r_{t_{i-1}} + a(b - r_{t_{i-1}})\delta + \sigma \sqrt{\delta} Z_i = (4)
\]
4.2.1 Estimation within the Vasicek Model

The respective roles of the mean-reversion parameter $a$, long-term average parameter $b$ and volatility parameter $\sigma$ have already been explained. However, it is not obvious how to choose the parameters in order to best fit the current situation. Here two approaches are presented, linear regression and maximum likelihood estimation. Both start with the analytical formula (5).

$$r_t = r_{t-1}e^{-a\delta} + b(1 - e^{-a\delta}) + \sigma\sqrt{\frac{1 - e^{-2a\delta}}{2a}}Z_t$$

Linear regression Observe that the relationship between $r_{t+1}$ and $r_t$ is linear and of the form

$$r_{t+1} = a'r_t + b' + \epsilon_t$$

where $a', b' \in \mathbb{R}$ and $\epsilon_t$ are independent, identically distributed normal random variables. Comparing yields:

$$e^{-a\delta} = a' \iff a = \frac{-\log a'}{\delta}$$

$$b(1 - e^{-a\delta}) = b' \iff b = \frac{b'}{1 - a'}$$

$$\sigma\sqrt{\frac{1 - e^{-2a\delta}}{2a}} = \sigma_e \iff \sigma = \sigma_e\sqrt{-2\log a'}$$

Least squares regression is the most common approach to calculate $a'$, $b'$ and $\sigma_e$. This procedure is available in most statistical packages. For completeness, details can be found in [6].

Maximum likelihood estimation The conditional probability density of $r_{t+1}$ given observation $r_t$ is given by

$$f(r_{t+1}|r_t; a, b, \hat{\sigma}) = \frac{1}{\sqrt{2\pi\hat{\sigma}^2}}\exp\left[-\frac{(r_t - r_{t-1}e^{-a\delta} - b(1 - e^{-a\delta}))^2}{2\hat{\sigma}^2}\right]$$

where

$$\hat{\sigma}^2 = \sigma^2\frac{1 - e^{-2a\delta}}{2a}$$

The logarithm of this expression is taken and the partial derivatives are set to zero. Details of this procedure can be found in [6], but the results are presented here.

$$b = \frac{S_yS_{xx} - S_xS_{xy}}{n(S_{xx} - S_{xy}) - (S_x^2 - S_xS_y)}$$

$$a = -\frac{1}{\delta}\log \frac{S_{xy} - bS_x - bS_y + nb^2}{S_{xx} - 2bS_x + nb^2}$$

$$\hat{\sigma}^2 = 2a\frac{S_{yy} - 2e^{-a\delta}S_{xy} + e^{-2a\delta}S_{xx} - 2b(1 - e^{-a\delta})(S_y - e^{-a\delta}S_x) + nb^2(1 - e^{-a\delta})^2}{n(1 - e^{-2a\delta})}$$

where $S_x = \sum_{i=1}^n r_i$, $S_y = \sum_{i=1}^n r_{i-1}$, $S_{xy} = \sum_{i=1}^n r_ir_{i-1}$, $S_{xx} = \sum_{i=1}^n r_i^2$ and $S_{yy} = \sum_{i=1}^n r_{i-1}^2$.

4.2.2 Advantages of the Vasicek Model

Historically, it has been observed that high interest rates hamper economic activity, resulting in a decrease in interest rates. By a similar reason, low interest rates help the interest rates to increase. The Vasicek model captures this mean reversion. Even though the model is not as simple as Merton’s model, analytical results for bond prices, estimation and the conditional distribution of $r_t$ exist, making implementation is very straightforward. Just like in Merton’s model, interest rate can become negative in the Vasicek model.
4.2.3 Drawbacks of the Vasicek Model

Since none of the parameters depend on time, there is not enough freedom in the model to match a given interest rate structure, like a bond prices or a forward or yield curve.

4.3 Hull-White model

This model assumes that the short rate follows the stochastic differential equation:

\[ dr_t = a(b - r_t) \, dt + \sigma \, dW_t \quad (6) \]

The only difference with the Vasicek model is that \( b_t \) is not a constant anymore, but a function of \( t \). This allows the Hull-White model to be fitted to an initial yield or forward curve. This flexibility comes at a price as the model cannot be handled analytically. Before fitting the model to a yield curve, the mean-reversion parameter \( a \) and volatility parameter \( \sigma \) have to be chosen. This can be done by applying linear regression or maximum likelihood estimation to a realized sample path, similar as in the Vasicek model. These parameters are now input in order to determine \( b_t \) such that the model produces a given forward rate curve. The solution is given by Hull himself in [7]:

\[ b_t = \frac{1}{a} \frac{\partial f(0,t)}{\partial t} + f(0,t) + \frac{\sigma^2}{2a^2} (1 - e^{-at}) \quad (7) \]

If the forward rate is not present, it can be approximated by its definition for small \( \Delta t \):

\[ f(0,t_i) = -\frac{\partial \log P(0,t)}{\partial t} \bigg|_{t=t_i} \approx -\frac{\log P(0,t_i + \Delta t) - \log P(0,t_i)}{\Delta t} \]

The values of the bonds for different maturities can be calculated for example from LIBOR or EURIBOR data for short maturities and from the yield curve for long maturities. An interpolation method is then used to approximate the rest of the curve. Common interpolation methods are linear interpolation, polynomial interpolation and spline interpolation. The derivative of the forward rate can be approximated by a finite difference method

\[ \frac{\partial f(0,t)}{\partial t} \bigg|_{t=t_i} \approx \frac{f(0,t_i + \Delta t) - f(0,t_i - \Delta t)}{2\Delta t} \]

In [8] a warning is present about the interpolation method, as linear interpolation will produce oscillations near the nodes of the interpolation. A similar issue known as Runge’s phenomenon can appear when using polynomial interpolation. Because of these unwanted properties, cubic splines are used in this thesis. However, also when using cubic splines, [8] claims that some oscillations may be present if \( t \) is small. This can be circumvented by not using bond prices for small \( t \).

4.3.1 Advantages of the Hull-White model

The main difference with the Vasicek model is the time-dependency of \( b_t \). This allows the model to be calibrated to for example bond prices or a forward or yield curve. Negative interest rates are still possible.

4.3.2 Drawback of the Hull-White model

Analytical formulas for bond prices are no longer available. All experiments and results are strictly numerical.

4.4 Cox–Ingersoll–Ross (CIR) model

This model was introduced in 1985 by Cox, Ingersoll and Ross in [9]. They assumed that the short rate follows the stochastic differential equation:

\[ dr_t = a(b - r_t) \, dt + \sigma \sqrt{r_t} \, dW_t \quad (8) \]
The first and second moment conditioned on \( r_0 \) are given in [9] as:

\[
E[r_t] = r_0 e^{-at} + b(1 - e^{-at})
\]
\[
\text{Var}(r_t) = r_0 \frac{\sigma^2}{a} (e^{-at} - e^{-2at}) + \frac{b\sigma^2}{2a} (1 - e^{-at})^2
\]

and Euler scheme

\[
r_{t+i} = r_t + a(b - r_t)\delta + \sigma \sqrt{r_t} Z_i
\]

Note that equation (8) only makes sense if \( r \geq 0 \).

### 4.4.1 Shifted CIR-model

As mentioned before, negative interest rates are present is the current market and the standard CIR-model cannot capture this. This problem is solved by introducing a shift parameter \( \nu \). The stochastic differential equation then becomes:

\[
dr_t = a((b - \nu) - r_t) \, dt + \sigma \sqrt{r_t + \nu} \, dW_t
\]

This allows the CIR-model to reach negative interest rates up to \(-\nu\).

### 4.4.2 Extended shifted CIR-model

It is still not possible to match a given interest rate structure. This is done by introducing the extended CIR-model. Just like the Hull-White model, it allows \( b \) to be time dependent. For convenience, the shift parameter is also put into this long term mean function \( b_t \) to obtain the extended CIR-model stochastic differential equation:

\[
dr_t = a(b_t - r_t) \, dt + \sigma \sqrt{r_t + \nu} \, dW_t.
\]

### 4.4.3 Estimation within the extended CIR-model

Similarly to the estimation methods regarding the Hull-White model, the mean-reversion parameter and the volatility parameter are estimated by applying maximum likelihood estimation to a realized sample path. A MATLAB documentation can be found in [10]. The long term mean function is found recursively starting from the long term mean function of the Hull-White model. Each iteration a fraction of the difference between the predicted bond prices and the input bond prices is added to the long term mean function until the difference between the two is smaller than a given tolerance level. More details can be found in section 7.2.2.
5 Pricing techniques

A great part of financial mathematics is to find techniques to calculate the fair price of a financial product. But what exactly is the fair price? In practice, the price of a product is the lowest price someone wants to sell the product for or the highest someone wants to buy the product for. The difference between these two prices is called the bid–ask spread and corresponds to the liquidity of the market. Most mathematical models assume a frictionless market, that is a bid-ask spread of zero. When the price of the product is available in the market, results of pricing techniques can be used to spot 'wrong' priced products. Trading a wrong priced asset does not guarantee any short-term money and is very risky, but, if the pricing techniques are right, will make money in the long run. The main point of pricing techniques in this thesis is to find fair prices for products that cannot be observed in the market. This is the case when the main supplier of the product will not buy the same product for a similar price. For example, a supermarket will sell you a liter of milk for about one euro, but refuses to buy the same liter of milk of you for a similar amount of money. Financial products following the same principle are for example trades with clients like mortgages or loans or over-the-counter trading with another party. Finding a fair price for these trades is done by various pricing techniques that all require a mathematical framework in order to be well-defined.

5.1 Mathematical framework

Let \((\Omega, \mathcal{F}, F, P)\) be a filtered probability space, where \(\mathcal{F} = \mathcal{F}_T\) and the filtration \(F = (\mathcal{F}_t)_{t \in [0,T]}\) satisfies the usual conditions, thus is right-continuous and \(\mathcal{F}_0\) contains all \(P\)-null-sets. Let \(T \in [0, \infty]\) denote the time horizon which, in general, can be infinite. First the assumption is made that there exists a riskless asset with a constant risk-free interest rate \(r\). This assumption is in practice not true but can be relaxed. Note that this interest rate is not equal to the interest rate quoted by banks or newspapers. After a year one euro increases to \(e^r\) euro instead of \((1 + r_{\text{quoted}})\) euro. This means that there exists a connection \(r = \log(1 + r_{\text{quoted}})\). Even though the difference is small for small \(r\), it cannot be neglected when dealing with big sums of money.

The second assumption is that the market is arbitrage free and is the backbone of many pricing techniques. This means that there exists no strategy with non-negative pay-off almost surely and a positive chance of a positive pay-off. In other words, there is no possibility to make money without taking risk. This seems like a fair assumption since, if there would exist arbitrage opportunities, everyone would do it and the possibility would vanish. However, for some financial institutes arbitrage opportunities are the only source of income. Keeping the no-arbitrage principle in mind, a first try to calculate a fair price might be to calculate the discounted expected pay-off. However, in practice, the real world measure is unknown and, maybe counterintuitively, this does not guarantee an arbitrage free market. Both problems are solved by introducing a risk-neutral measure. A probability measure \(P^*\) is a risk-neutral measure if under \(P^*\) for every asset price \(S_i\):

\[
S_i^0 = e^{-rT}E_{P^*}[S_i^T]
\]

So, under the risk neutral measure, the discounted expectation of an asset is equal to the initial price. The existence of the risk-neutral measure in arbitrage-free markets is guaranteed by the first Fundamental Theorem of Asset Pricing presented below.

**Theorem 5.1** (Fundamental Theorem of Asset Pricing). A market is free of arbitrage if and only if there exists at least one risk neutral measure such that the Radon-Nikodym derivative with respect to the real world measure is bounded.

Using a risk-neutral measure, an important principle can be derived. This principle tells how the no-arbitrage principle can be used to determine a fair price.

**Theorem 5.2** (Principal of one price). If two strategies result in the same pay-off at maturity in every situation, then the price of the two strategies at \(t = 0\) must be equal. This price is regarded the fair price.
This result is intuitively very clear. If the price was not equal, then buying the lowest and selling the highest at \( t = 0 \) would yield a risk-less positive amount of money, while there are no cash flows at maturity. Therefore an arbitrage opportunity exists.

Now add a financial product \( X \) to the arbitrage-free market. This product is defined by its pay-off and can in general be any nonnegative random variable. The price of this product should be in such a way that the extended market is also free of arbitrage. This price is known as an arbitrage free price. The next theorem guarantees the existence of an arbitrage price and provides a formula. Proof can be found in the appendix.

**Theorem 5.3** (General pricing formula). Let \( X \) be a financial product. If the original market is arbitrage free, then for every risk-neutral measure \( \mathbb{P}^* \) a fair price \( \pi^X \) is given by:

\[
\pi^X = e^{-rT} \mathbb{E}_{\mathbb{P}^*}[X]
\]

Where the expectation is taken with respect to this risk-neutral measure.

Note that the theorem does not guarantee a unique arbitrage free price. In general there may be multiple arbitrage free prices. However, it turns out that all arbitrage free prices form an interval. The uniqueness of the arbitrage free price corresponds directly to another important property of the financial product. This property is whether \( X \) is attainable or not. A product is attainable if there exists a strategy in the original market that replicates the pay-off of \( X \). This strategy is known as a hedge. This correspondence is made precise in the next theorem.

**Theorem 5.4.** If \( X \) admits a unique arbitrage-free price then \( X \) is attainable and if every product is attainable, the extended market is complete.

The first part of the theorem is equivalent to saying that there exists exactly one risk-neutral measure and the second part of the theorem is equivalent to saying that a replicating portfolio or hedge exists. A portfolio is a vector with predictable amounts of each available financial product at every time point. The price of the hedge is then equal to the arbitrage free price of \( X \) by the Fundamental Theorem of Asset Pricing. Finally, if the interest rate is not deterministic, but admits a stochastic process, the \( e^{-rT} \) outside the expectation should be replaced with an \( e^{-\int_0^T r_s \, ds} \) inside the expectation. Of course if \( r_s = r \) almost everywhere, the two coincide. The mathematical framework already includes a general pricing formula. Before this formula can be used in practice, the risk-neutral measure has to be known. If the risk-neutral measure is given, the problem reduces to an integral that can be computed either analytically or numerically.

### 5.2 Hedging/Replicating portfolio in complete markets

Hedging an arbitrary option is not always possible in view of the theory above, but even when a theoretical hedge exists, it is not obvious how to find this hedge and if it works in practice. The value of product \( X \) at a given time point may depend on many different things. Most common are, the underlying stock prices, the current interest rate and the possibility of default of the counterparty or a volatility change. A perfect hedge should protect the holder of the option against all sort of outcomes, but this is generally not done in practice. Common forms of practically usable hedge can be found below.

#### 5.2.1 Delta hedge

This hedge protects the holder against changes in the underlying asset. It is only applicable for financial products with a price for which the first derivative with respect to the underlying exists. This derivative can either be found from the analytical pricing formula or by numerical procedures. The basic idea is to hold a fraction of the underlying asset for each product. Therefore a change of the underlying will influence both the value of the financial product as well as the hold underlying assets. If the correct fraction is used, these changes balance. The mathematical backbone is given by the Taylor series:

\[
X(S + \epsilon) = X(S) + \epsilon \frac{\partial X}{\partial S} + \frac{\epsilon^2}{2} \frac{\partial^2 X}{\partial S^2} + O(\epsilon^3)
\]
If the change in the underlying is small enough, the second order term in the Taylor series is negligible. Thus a portfolio consisting of the financial product together with \( \frac{\partial X}{\partial S} \) of the underlying hedges against a small change in the underlying. This portfolio is called a delta neutral portfolio. Theoretically, in order to maintain a delta neutral portfolio, the derivative has to be known all the time and the amount of the underlying asset in the portfolio has to change accordingly. In practice, this is of course impossible and daily or weekly rebalancing is done. This comes at the price of a bit of risk the higher order terms in the Taylor expansion are also relevant in discrete time, so one has to be careful that \( \epsilon \) is small enough for their risk appetite.

5.2.2 Replicating portfolio

This strategy is a more elaborate version of delta hedging. Delta hedging eliminates market risk by creating a portfolio of the derivative together with a corresponding fraction of the underlying asset. In a way the delta neutral portfolio replicates a constant pay-off and therefore eliminates market risk. It does not for example eliminate currency risk. The replication idea can be extended. Not only underlying assets can be used to replicate pay-offs, but also other derivatives. Look for example at the well-known put-call parity:

\[
C_t - P_t = S_t - K \cdot \frac{P(0,T)}{P(0,t)}
\]

This can be used to replicate the pay-off of a call/put with strike \( K \), given by respectively \( C_t \) and \( P_t \), by the other one, some bonds and the underlying asset. A call option gives the holder the right to buy the underlying at maturity \( T \) at strike price \( K \), while a put option allow the holder to sell the underlying at maturity at price \( K \). For example, the risk of a put can be hedged by buying a call, selling an asset and holding \( K \) amounts of corresponding bonds. If the put-call parity does not hold, there exists an arbitrage opportunity by the Principle of One Price. The strategy would be to either buy or sell a call option and immediately hedge the corresponding risk with the replicating portfolio. Other commonly used products to create a replicating portfolio are swaps to hedge against interest rate changes or forward exchange contracts to hedge against a change in ratio between currencies. This last one might be necessary if contracts in the portfolio are in different currencies. In complete markets there always exists a replicating portfolio, although no description is given how to find one. Even in incomplete markets a replicating portfolio can exists and one should always try to find products and strategies that exactly or closely replicate the pay-off of a certain financial product.

5.2.3 Black-Scholes approach

The general pricing formula presented in the mathematical framework is not directly usable in all situations. In [11] Black and Scholes proposed a model that allowed for analytical prices of the common call and put options and presented an indirect way to calculate any European type option through a partial differential equation. While this all seems nice, the model requires some assumptions that are known not to be true. However, when comparing the results to real world data, the outcomes of the model are fairly close. The model is by no means perfect as also some inconsistencies can be found, like the volatility smile and the inability to deal with path-dependent or American style options. Black and Scholes made the following assumptions:

- There exists a riskless asset with a constant return.
- The log return of the stock price is a drifted geometric Brownian motion under the risk-neutral measure.
- The stock does not pay dividend.
- The market is free of arbitrage.
- The market is frictionless.
The first and second assumption can be relaxed to allow for deterministic functions in time and the third assumption can be relaxed to allow continuous yield or discrete proportional dividends. Both relaxations however yield more complicated results, but do not add anything in understanding the process. The last two assumptions are known not to hold but are required in the derivation. Luckily, the assumptions ‘almost’ hold in practice and therefore the results are still usable. The second assumption in stochastic differential notation is equivalent to:

\[ dS = \mu S \, dt + \sigma S \, dW. \]

Here \( W \) is a Brownian motion. After switching to a risk-neutral measure, this stochastic differential equation can be solved by applying Itô formula on \( \log S_t \) and a no-arbitrage argument. The solution is given by:

\[ S_t = S_0 e^{(r - \frac{\sigma^2}{2}) t + \sigma W_t}. \]

This solution in combination with the general pricing formula is already enough to find the analytical Black-Scholes prices of the vanilla call and put option with strike price \( K \) and maturity \( T \) at \( t = 0 \). The results are presented here. Let

\[ d_1 = \frac{1}{\sigma \sqrt{T}} \left[ \log \left( \frac{S_0}{K} \right) + \left( r + \frac{\sigma^2}{2} \right) T \right], \]

then the price of a European call option is given by:

\[ C(S_0, K, T) = S_0 \Phi(d_1) - Ke^{-rT} \Phi(d_1 - \sigma \sqrt{T}), \]

where \( \Phi(x) \) is the standard normal cumulative distribution function. The value of a put option can now easily be calculated using the put-call parity. A direct formula is not available anymore, but a partial differential equation with the value of the derivative as a solution can be derived. The start is building a delta-neutral portfolio. This portfolio \( \Pi \) consists of \(-1\) of the derivative and \( \frac{\partial V}{\partial S} \) shares. By Itô’s formula and the stochastic differential equation for the stock price the differential of the portfolio \( \Pi \) can be derived:

\[ d\Pi = -dV + \frac{\partial V}{\partial S} \, dS = \left( - \frac{\partial V}{\partial t} - \frac{1}{2} \frac{\partial^2 V}{\partial S^2} S^2 \sigma^2 \right) dt \]

The differential of the portfolio can be calculated in a similar way and the result above can be directly substituted together with the differential for the stock price to find:

\[ r \Pi \, dt = d\Pi \]

\[ r \left( - V + \frac{\partial V}{\partial S} S \right) \, dt = \left( - \frac{\partial V}{\partial t} - \frac{1}{2} \frac{\partial^2 V}{\partial S^2} S^2 \sigma^2 \right) dt \]

\[ 0 \, dt = \left( \frac{\partial V}{\partial t} + \frac{1}{2} \frac{\partial^2 V}{\partial S^2} S^2 \sigma^2 - rV + r \frac{\partial V}{\partial S} S \right) dt \]

\[ 0 = \frac{\partial V}{\partial t} + \frac{1}{2} \sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + rS \frac{\partial V}{\partial S} - rV \]

This is the Black-Scholes partial differential equation. To solve it the boundary condition at \( t = T \) is set to the pay-off function of the derivative. Then the solution \( V \) can be found numerically using finite difference methods. In some cases however, like the vanilla call and put option, an analytical solution is available.
5.3 Monte Carlo

The general pricing formula states that a fair price \( V_0 = V(0, S_0) \) of an option is equal to the discounted expected payoff \( X \) under a risk-neutral measure. Problem is that this integral in practice often is very hard to compute. Monte Carlo methods offer an approach to approximate the integral by simulating the underlying process. The law of large numbers states that for \( N \to \infty \):

\[
V_0 = \mathbb{E}[e^{-\int_0^T r_s \, ds} X] = \lim_{N \to \infty} \frac{1}{N} \sum_{\omega \in \Omega} e^{-\int_0^T r_s(\omega) \, ds} X(\omega).
\]

Of course simulating an infinite amount of sample paths is not feasible. Luckily, by the central limit theorem the following holds.

\[
\sqrt{N} \left( V_0 - \frac{1}{N} \sum_{\omega \in \Omega} e^{-\int_0^T r_s(\omega) \, ds} X(\omega) \right) \overset{d}{\to} \mathcal{N}(0, \sigma^2).
\]

provided that \( \sigma^2 < \infty \), that is the variance of the payoff is finite. In theory, given the required computational power and an underlying stochastic process, Monte Carlo methods can be used to calculate the fair price for every tolerance level.

5.3.1 Generating the underlying process

In general, the underlying process can be any stochastic process. In practice however, most underlying processes are specified via a stochastic differential equation. If an exact solution is available, this one can be used to directly generate sample paths. Most of the time however, an exact solution is not available so a different approach is required. This starts by discretization of the continuous equation. To make sure that the approximation is close, the time steps should be small. The error can be controlled by comparing the results for \( m \) and \( 2m \) time steps, the results of \( 2m \) and \( 4m \) time steps et cetera until the difference is acceptably small. The stochastic differential equation can now be translated into a discretization scheme. In practice the Euler scheme is most commonly used. The result can then be used to determine the distribution of the next time step given the current value. This process can be repeated from \( t = 0 \) till maturity. An example of a discretization scheme is encountered when generating interest rate paths following the Vasicek model:

\[
dr_t = a(b - r_t) \, dt + \sigma \, dW_t \Rightarrow r_{t_i} = r_{t_{i-1}} + a(b - r_{t_{i-1}}) \delta + \sigma \sqrt{\delta} Z_i.
\]

This example also gives insight in how to create the scheme from the differential equation. \( Z_i \) is a standard normal distributed random variable and therefore generating sample paths can easily be done with for example MATLAB.

5.3.2 Monte Carlo estimation of derivatives

Financial institutes are not only interested in the price of an option, but also in the sensitivities. These sensitivities are known as the Greeks and are necessary to correctly hedge a derivative. If a Monte Carlo estimation of the fair price is available, then the derivative with respect to a given parameter \( \theta \) can be approximated by:

\[
\frac{\partial V_0}{\partial \theta} \approx \frac{V_0(\theta + \epsilon) - V_0(\theta)}{\epsilon}
\]

This method is known as the bump and revalue method. The value of \( \epsilon \) is positive and it is important to pick a suitable \( \epsilon \). For large \( \epsilon \) the approximation of the derivative by the finite difference is not defendable as second order terms start to interfere and a small \( \epsilon \) results in a large variance due to the division by \( \epsilon \). To combat this variance, some variance reduction methods are described in section 5.3.5. In theory, for a given \( \epsilon \), a given underlying process and the required computation power, the confidence interval can be made arbitrarily small. However, one should be careful as every \( \epsilon \) comes with a bias caused by the finite difference approximation.
5.3.3 Confidence intervals

As stated before, the central limit theorem is the backbone of Monte Carlo estimation. Recall that the scaled difference of the realized average and the mean converges in distribution to a normal distribution:

$$\sqrt{\frac{N}{\sigma}} \left( V_0 - \frac{1}{N} \sum_{\omega \in \Omega} e^{-\int_0^T r_s(\omega) ds} X(\omega) \right) \xrightarrow{d} \mathcal{N}(0, 1)$$

This statement can very easily be used to calculate confidence intervals. Let $\bar{X}$ be the Monte Carlo estimate, then:

$$P\left( \bar{X} - \frac{\xi_{\frac{\alpha}{2}} \sigma}{\sqrt{N}} \leq V_0 \leq \bar{X} + \frac{\xi_{\frac{\alpha}{2}} \sigma}{\sqrt{N}} \right) = 1 - \alpha,$$

where $\xi_{\frac{\alpha}{2}}$ are quantiles of the standard normal distribution function. The expression within the probability is therefore a confidence interval with a significance level of $1 - \alpha$. From this formula, it is evident that decreasing the length of the confidence interval should be done by either increasing $N$ or by reducing $\sigma$. The first option is simply translates in increasing computational power and/or computation time. For the second option some other methods exists. One common method is quasi-Monte Carlo, described in section 5.3.4. Three other methods are described in section 5.3.5. In practice often a combination of different methods is used.

5.3.4 Quasi-Monte Carlo method

Monte Carlo estimation requires random number generators that create pseudo-random numbers. Nowadays, these numbers are not distinguishable anymore from pure random numbers. The randomness guarantees convergence by the law of large numbers, but the rate of convergence is, by the central limit theorem, quite slow. A possible solution lies in quasi-random numbers. These numbers are deterministic and try to best represent the underlying space by equally distributing themselves over this space. Sequences of quasi-random numbers are called low-discrepancy sequences. Examples are given by van der Corput, Halton or Sobol sequences. These sequences can be used as input for Monte Carlo estimations.

**Low-discrepancy sequences**

The discrepancy of a set $X = \{x_1, \ldots, x_N\}$ is defined as:

$$D_N^*(X) = \sup_{B \in J} \left| \frac{A(B; X)}{N} - \lambda_d(B) \right|,$$

where $A(B; X)$ is the amount of points in $X$ that are in $B$, $\lambda_d(B)$ is the $d$-dimensional Lebesgue measure and $J$ is the set of rectangular boxes in the $d$-dimensional unit cube with one point fixed in the origin:

$$\prod_{i=1}^d [0, a_i) = \{x \in \mathbb{R}^d : 0 \leq x_i < a_i\}.$$

A sequence $X$ is known as a low-discrepancy sequence if and only if the following holds for a certain constant $C$:

$$D_N^*(X) \leq C \left( \frac{\log N}{N} \right)^d.$$

These low-discrepancy sequences are used to calculated integrals. The error can be characterized by the Koksma-Hlawka inequality. A complete discussion and proof can be found in section 5.4 of [12]. Let $V(f)$ be the total variation of $f$, as in equations (5.7) and (5.8) of [12], then:

$$\left| \frac{1}{N} \sum_{i=1}^N f(x_i) - \int_{[0,1]^d} f(u) du \right| \leq V(f) D_N^*(X)$$
This upper bound is useless in practice, since for normal values like \( N = 100 \) and \( d = 360 \), the discrepancy can be of order \( 2^{786} \). In [13] is claimed that under appropriate conditions, the order of convergence is \( O(N^{\epsilon - 1}) \) for all \( \epsilon > 0 \). For comparison, for the normal Monte Carlo method the bound is found by the central limit theorem and Chebyshev’s inequality. However, this bound only holds with a minimum probability of \( 1 - \delta \):

\[
\left| \frac{1}{N} \sum_{i=1}^{N} f(x_i) - \int_{[0,1]^d} f(u) \, du \right| \leq \sqrt{\frac{\operatorname{Var}(f(U))}{\delta N}}
\]

where \( U \) is a uniform random variable on \([0,1]^d\). Therefore \( \operatorname{Var}(f(U)) \) is easy to estimate whereas \( V(f) \) and \( D_N(X) \) are more often than not harder to estimate than the integral itself. The big difference in the bounds of the quasi-Monte Carlo and ordinary Monte Carlo is the order of \( N \). Namely, again under the conditions of [13], \( O(N^{\epsilon - 1}) \) versus \( O(N^{-1/2}) \). Unfortunately for many practical situations, it is very hard to show that these conditions hold.

**Van der Corput sequence** First described in 1935, this is one of the simplest low-discrepancy sequences. Like most sequences the output is a number in the unit interval. There exists a van der Corput sequence for every positive integer \( p \), but it turns out that in practice it works best if \( p \) is prime. To calculate the \( n^{th} \) element of a van der Corput sequence for a given number \( p \), first write down the number in base-\( p \):

\[
n = \sum_{k=0}^{L-1} d_k(n)p^k
\]

Here \( L = \lfloor \log_p(n) + 1 \rfloor \). These corresponding digits \( d_k(n) \) are the input to calculate the \( n^{th} \) number in the van der Corput sequence. The formula is given by:

\[
g_p(n) = \sum_{k=0}^{L-1} d_k(n)p^{-k-1}
\]

In other words, what happens is that the number in base-\( p \) gets flipped and scaled to the unit interval. Then the number is translated back into base-10. For example, the first eight numbers of the van der Corput sequence with prime numbers 2, 3, 5 and 7 are given in figure 2.

The proof that this procedure indeed generates a low-discrepancy sequences for every \( p \) follows from Theorem 3.6 in [14] after setting the dimension equal to 1. In [15] Henri Faure even found...
<table>
<thead>
<tr>
<th>n</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>$g_2(n)$</td>
<td>0.5000</td>
<td>0.2500</td>
<td>0.7500</td>
<td>0.1250</td>
<td>0.6250</td>
<td>0.3750</td>
<td>0.8750</td>
<td>0.0625</td>
</tr>
<tr>
<td>$g_3(n)$</td>
<td>0.3333</td>
<td>0.6667</td>
<td>0.1111</td>
<td>0.4444</td>
<td>0.7778</td>
<td>0.2222</td>
<td>0.5556</td>
<td>0.8889</td>
</tr>
<tr>
<td>$g_5(n)$</td>
<td>0.2000</td>
<td>0.4000</td>
<td>0.6000</td>
<td>0.8000</td>
<td>0.0400</td>
<td>0.2400</td>
<td>0.4400</td>
<td>0.6400</td>
</tr>
<tr>
<td>$g_7(n)$</td>
<td>0.1429</td>
<td>0.2857</td>
<td>0.4286</td>
<td>0.5714</td>
<td>0.7143</td>
<td>0.8571</td>
<td>0.0204</td>
<td>0.1633</td>
</tr>
</tbody>
</table>

Figure 2: First eight outcomes for the first four prime numbers.

(a) First 10000 points with $p_1 = 2$ and $p_2 = 4$

(b) First 1000 points with $p_1 = 101$ and $p_2 = 103$

Figure 3: Examples of two situations that need to be avoided when using van der Corput sequences.

asymptotic results for the constant $C_{\text {Corput}}(p)$.

$$C = \frac{p^2}{4(p+1) \log p} \quad \text{if } p \text{ is even}$$

$$C = \frac{p-1}{4 \log p} \quad \text{if } p \text{ is odd.}$$

**Halton sequence** The van der Corput sequence is one dimensional. Halton generalized this in [16] to an arbitrary dimension by introducing the Halton sequence. The base are van der Corput series for different primes. Let $g_{p_i}$ be the van der Corput sequence for prime number $p_i$. Then:

$$H_p(n) = (g_{p_1}(n), \ldots, g_{p_d}(n))$$

is a $d$-dimensional low-discrepancy sequence, again by Theorem 3.6 of [14]. In theory, all numbers can be chosen as long as they are relative primes, but in practice, because of computational power, the first $d$ prime numbers are used. To see the low-discrepancy in practice, the first 500 elements of a 2-dimensional Halton-sequence are calculated, plotted into a lattice and compared with 500 uniformly random 2D-points in figures 1a and 1b. It immediately stands out that the points in the left picture are more evenly placed, while the right picture contains clusters and more open space. It is necessary that the numbers are relative primes because of correlation. However, even if the numbers are relative prime, a burn-in period might be necessary. Realizations of both situations are plotted in 3a and 3b. It is immediately clear that for larger primes a burn-in period is needed to fill the entire space. However, after this burn-in period, everything is fine. This is not the case when using numbers that are not relative primes. This will result in a pattern repetition and some fractal like structures appear. The result is clearly not uniform.
**Transforming uniform random variables** Every low discrepancy series outputs a number between zero and one. To transform this into an arbitrary random variable, the following calculation for $X = F^{-1}(U)$ is considered, provided that $F^{-1}$ is well defined:

$$
P(X \leq x) = P(F^{-1}(U) \leq x) = P(U \leq F(x)) = F(x)
$$

In other words, the random variable $X = F^{-1}(U)$ has distribution function $F$. This can be used to transform a uniform $(0, 1)$ distribution into an arbitrary distribution as long as the inverse of the distribution can be found either analytically or numerically. If this is not the case then generalized inverse can be used. Generalized inverse defined as:

$$F^{-1}(u) = \inf \{x | F(x) \geq u\}.$$

This principle is used to generate the quasi-random normal distributed variables needed in most quasi-random Monte Carlo estimators.

### 5.3.5 Variance reduction methods

Every path requires a number of samples from a random variable. Whatever the quantity of interest is, the problem lies most of the time in finding acceptable small confidence intervals. In case one has unlimited time, the central limit theorem allows the confidence interval to become arbitrarily small. While this theoretical property is very convenient, the rate of convergences is only the square root of the number of paths. In particular when estimating derivatives, like the Greeks, the process is very time consuming. Luckily, other methods to decrease the variance exists. More details can be found in chapter 4 of [12]. It is worth noting that these methods only try to affect the implicit constant in $O(N^{-\frac{1}{2}})$, so these are not as ambitious as the quasi-Monte Carlo as this tries to change the order of convergence.

**Antithetic variates** The generated path is a direct function of the generated random variables. Using the same random variables is it possible to generate a new path by using the same random variables but multiplied by $-1$. Note that the distribution function needs to be centered and symmetric in order to guarantee that the generated paths are from the same distribution. Since generating the random numbers takes in general more time than the rest of the simulation, these ‘free’ paths are very welcome. However, the variance reduction goes a bit further than that. Suppose antithetic variates are used to estimate $\theta = \mathbb{E}(X)$. Define $\hat{\theta}_1$ and $\hat{\theta}_2$ as respectively the original Monte Carlo estimate and the ‘negative’ Monte Carlo method. When calculating the variance of the average, denoted as $\hat{\theta}$, between these two, observe that:

$$Var(\hat{\theta}) = Var\left(\frac{\hat{\theta}_1 + \hat{\theta}_2}{2}\right) = Var(\hat{\theta}_1) + Var(\hat{\theta}_2) + 2Cov(\hat{\theta}_1, \hat{\theta}_2) \leq Var(\hat{\theta}_1)$$

as $Var(\hat{\theta}_1) = Var(\hat{\theta}_2)$ and by the Cauchy-Schwartz inequality:

$$Cov(\hat{\theta}_1, \hat{\theta}_2) \leq \sqrt{Var(\hat{\theta}_1)Var(\hat{\theta}_2)} = \sqrt{Var(\hat{\theta}_1)^2} = Var(\hat{\theta}_1)$$

From the formula it is clear that this methods works best if the covariance is negative and does nothing if $\hat{\theta}_1 = \hat{\theta}_2$. Intuitively, the first possibility is expected, since the input has negative correlation. In any case, using antithetic variates never increases the variance. Due to the easy implementation and the lack of added computational time, this method is commonly used to reduce the variance of a Monte Carlo estimator.

**Control variates** The control variates technique uses information of a simpler option $B$ to improve the evaluation of a more difficult option $A$. Define $\hat{C}_A$ and $\hat{C}_B$ be the Monte Carlo estimates of options $A$ and $B$. Let $C_A$ and $C_B$ be the accurate value of option $A$ and $B$, respectively.
Because option B is simpler, the variance of the value is small compared to the variance of the price of option A. The control variate estimate of derivative A is denoted as $\tilde{C}_A$ and defined as:

$$\tilde{C}_A = \hat{C}_A - \beta(\hat{C}_B - C_B)$$

Because of the law of large numbers, the Monte Carlo estimators for both option A and option B are without bias. It follows that the control variate estimator is also without bias:

$$E(\tilde{C}_A) = E(\hat{C}_A - \beta(\hat{C}_B - C_B)) = E(\hat{C}_A) + \beta E(\hat{C}_B - C_B) = E(\hat{C}_A) = E(C_A)$$

Above equation holds for any value of $\beta$. For now, the only constraint for $\beta$ is that it must have the same sign as the correlation coefficient $\rho$ of options A and B. To obtain more information about the value of $\beta$, the variance of the control variate estimator is calculated and expressed in terms of the standard deviations $\sigma_A$ and $\sigma_B$:

$$\sigma^2 = \sigma_A^2 + \beta^2 \sigma_B^2 - 2\rho \beta \sigma_A \sigma_B$$

From this it can be easily verified that the variance is reduced if and only if:

$$\rho > \frac{\beta \sigma_B}{2 \sigma_A}$$

This equation puts a constraint on the value of $\beta$, but it is always possible to pick a value such that above equation holds and the variance gets reduced. It is even possible to calculate an optimal $\beta^*$. This is done by calculating the derivative of the variance with respect to $\beta$ and calculate the root. The result of this trivial calculation is:

$$\beta^* = \frac{\sigma_A}{\sigma_B} \rho$$

Of course in practice these standard deviations and correlation are unknown, so estimators are used to calculate the optimal $\beta$. The resulting variance is:

$$\sigma^2 = (1 - \rho^2)\sigma_A^2$$

As expected, the control variates methods works best if option A and option B undergo correlation, that is $\rho^2$ close to 1.

**Same initial seed** This method is only concerning Monte Carlo estimations of derivatives. The basic idea is to use the same generated sample paths for both $V_0(\theta + \epsilon)$ and $V_0(\theta)$. This will reduce the variance in the difference between the two if the fair price is sufficiently smooth in the given parameter. One should therefore take care when implementing this idea into for example digital options. As the total variance is just only different with factor $\frac{1}{\epsilon^2}$ from the variance of the difference, this procedure will decrease the total variance of the Monte Carlo estimation of the Greeks.

### 5.4 Finite difference methods

The Black-Scholes model provides an indirect way of pricing options through a partial differential equation. This is very common in financial mathematics as the underlying process is often defined by a stochastic differential equation. Therefore it will be useful to take a closer look at numerical methods to solve partial differential equations. There exist many methods with their own advantages and drawbacks, but the dominant approach to numerically solve partial differential equations at the moment are finite difference methods. The main idea of these methods is to replace the derivative by the finite difference approximation. This is done is a similar way as the delta hedge approach. For example:

$$X(S + \epsilon) = X(S) + \epsilon \frac{\partial X}{\partial S} + O(\epsilon^2) \iff \frac{\partial X}{\partial S} = \frac{X(S + \epsilon) - X(S)}{\epsilon} + O(\epsilon)$$
Thus the approximation is valid as $\epsilon \to 0$. The second derivative is written as the finite difference of the derivatives at $X(S)$ and $X(S-\epsilon)$ and is therefore approximated by:

$$\frac{\partial^2 X}{\partial S^2} \approx \frac{X(S+\epsilon) - 2X(S) + X(S-\epsilon)}{\epsilon^2}$$

Because of these approximations, two questions need to be answered. Is the solution stable and if so, what is the order of convergence?

### 5.4.1 Transforming the partial differential equation

The goal is to calculate the value of the option at $t=0$, while the value at maturity is known. Therefore it is convenient to invert the time using a transformation like $\tau(t) = T - t$. This guarantees that the price is known at $\tau = 0$ and that the partial differential equation can be solved up to $\tau(0)$. The solution at maturity can then be interpreted as the value at $t = 0$. While this transformation is a standard procedure, in general any continuous transformation can be done in order to simplify the partial differential equation before solving it, as long as there exists a bijection between the old set of variables and the new one. For example the Black-Scholes partial differential equation can be transformed into the heat equation

$$\frac{\partial u}{\partial \tau} = \frac{\partial^2 u}{\partial z^2}$$

The complete procedure can be found in [17]. While these kinds of transformations might be convenient, they are rarely necessary in order to solve the partial differential equation. The found solution has to be transformed back into the original form by applying the inverses of all transformations to the found solution. These inverses exist because all transformations are bijections. In order to find the solution for all time points and all initial values of the underlying, linear interpolation is used. As long as the time and space steps are not too large, this yields accurate results.

### 5.4.2 Boundary conditions

The start to solving a partial differential equation using finite difference methods is to setup a grid. This grid has a time-axis and an axis for every underlying asset. The description given here is limited to one underlying asset, but the idea can easily be generalized to higher dimensions. Keep in mind though, that solving for high dimensions can yield stability and computational power problems. The boundary of the time axis is, after the time inversion, given by $[0, \tau(0)]$. The value of the underlying $S$ is in general not bounded, but the assumption is made that in finite time this value does not exceed a certain $S_{\text{max}}$. As the value of the financial product is known at maturity, this gives one of the boundary conditions, namely:

$$V(0, S) = X(S)$$

where $X(S)$ is the payoff function of the financial product. Another boundary condition can be found by observing that if the value of the underlying reaches zero, it will always be zero from then on. Therefore:

$$V(\tau, 0) = X(0)$$

The final boundary condition cannot be logically derived, but is an approximation in order to meet the assumption that $S$ does not exceed $S_{\text{max}}$. Therefore the final boundary condition is given by:

$$V(\tau, S_{\text{max}}) = X(S_{\text{max}})$$

This boundary condition is an approximation and to make sure that this approximation does not influence the solution too much, the value of $S_{\text{max}}$ needs to be so large that in practice the value of the underlying does not come too close very often. When additional transformations are used to rewrite the partial differential equation, the boundary conditions need to be transformed equally. These three boundary conditions cover three of the four sides of the grid. The inside and fourth side are now calculated using a discretization scheme.
5.4.3 Solving the partial differential equation

In order to calculate the values that are not given by the boundary condition, three methods are described. All use the values of the given time point before in order to calculate the values of the next time point. Because the values are known at \( \tau = 0 \), iterating this process can be used to calculate the entire grid. The new values can be calculated from the old ones using an explicit, implicit or Crank-Nicolson method. The first two follow from section 8.4 from [18] while the last method was proposed in 1947 by Crank and Nicolson in [19]. The horizontal grid distance is given by \( \Delta \tau \) and the vertical grid distance is given by \( \Delta S \).

Explicit method  In order to calculate \( V(\tau_i, S_j) \) the explicit method substitutes the values of \( V(\tau_{i-1}, S_{j-1}), V(\tau_{i-1}, S_j) \) and \( V(\tau_{i-1}, S_{j+1}) \) into the finite difference version of the partial differential equation.

\[
\begin{align*}
\frac{\partial V}{\partial \tau} &= V(\tau_i, S_j) - V(\tau_{i-1}, S_j) \\
\frac{\partial V}{\partial S} &= \frac{V(\tau_{i-1}, S_{j+1}) - V(\tau_{i-1}, S_j)}{\Delta S} \\
\frac{\partial^2 V}{\partial S^2} &= \frac{V(\tau_{i-1}, S_{j+1}) - 2V(\tau_{i-1}, S_j) + V(\tau_{i-1}, S_{j-1})}{(\Delta S)^2}
\end{align*}
\]

As \( V(\tau_i, S_j) \) is only used in the finite difference approximation of \( \frac{\partial V}{\partial \tau} \), the result is a direct formula to calculate this value as a function of the values of the time point before. Therefore this method is known as a forward in time, central in space (FTCS) method. Because of this, the explicit method is very fast, but the downside is that stability issues might arise. The order of convergence is linear over the time step and quadratic over the space step.

Implicit method  This method substitutes the values of \( V(\tau_{i-1}, S_{j-1}), V(\tau_i, S_j) \) and \( V(\tau_i, S_{j+1}) \) into the finite difference version of the partial differential equation in order to calculate \( V(\tau_i, S_j) \). This yields a system of linear equations that has to be solved by numerical methods to obtain the values of \( V(\tau_i, S_j) \).

\[
\begin{align*}
\frac{\partial V}{\partial \tau} &= V(\tau_i, S_j) - V(\tau_{i-1}, S_j) \\
\frac{\partial V}{\partial S} &= \frac{V(\tau_{i-1}, S_{j+1}) - V(\tau_i, S_j)}{\Delta S} \\
\frac{\partial^2 V}{\partial S^2} &= \frac{V(\tau_{i-1}, S_{j+1}) - 2V(\tau_{i-1}, S_j) + V(\tau_i, S_{j-1})}{(\Delta S)^2}
\end{align*}
\]

Therefore this method is implicit and the method is known to be backwards in time, central in space (BTCS). While this method requires more computations, the advantage is that it is always stable. The order of convergence is equal to that of the explicit methods.

Crank-Nicolson  This method uses a combination of the explicit and the implicit method in order to calculate \( V(\tau_i, S_j) \). Crank and Nicolson proposed to use the following set of equations:

\[
\begin{align*}
\frac{\partial V}{\partial \tau} &= V(\tau_i, S_j) - V(\tau_{i-1}, S_j) \\
\frac{\partial V}{\partial S} &= \frac{1}{2} \left[ \frac{V(\tau_{i-1}, S_{j+1}) - V(\tau_{i-1}, S_j)}{\Delta S} + \frac{V(\tau_i, S_{j+1}) - V(\tau_i, S_j)}{\Delta S} \right] \\
\frac{\partial^2 V}{\partial S^2} &= \frac{1}{2} \left[ \frac{V(\tau_{i-1}, S_{j+1}) - 2V(\tau_{i-1}, S_j) + V(\tau_{i-1}, S_{j-1})}{(\Delta S)^2} + \frac{V(\tau_i, S_{j+1}) - 2V(\tau_i, S_j) + V(\tau_i, S_{j-1})}{(\Delta S)^2} \right].
\end{align*}
\]
This method is also known as central in time, central in space (CTCS) method. This yields the same advantages as the implicit method, but the order of convergence is quadratic in the time-step size, instead of linear. The only downside is the additional computational time.

5.5 Lattice model

The biggest drawback of Monte Carlo is reducing the variance of the outcome to an acceptable level. The lattice model does not have this problem since the outcome is deterministic given the input. Instead of the randomness in the paths of Monte Carlo, the lattice model only allows for certain paths. The amount of these paths is in such a way that all can be evaluated within a reasonable amount of time. The outcome therefore is deterministic. Drawbacks are that it is not easy to deal with complicated products like Asian or basket options on different but correlating products. Pricing using the lattice model uses the following steps. First, if necessary, the value of the short rate at each node is calculated following a given interest rate model. Then the value of the underlying is calculated at each node. Finally, the value of the option is calculated back from maturity to the current time point using a risk-neutral measure.

5.5.1 Binomial tree

As suggested by the name, the binomial tree model only allows two jumps. To decrease the required computational power, the up factor is the inverse of the down factor. This property makes the tree recombining and reduces the amount of nodes at time point \( n \) from \( 2^n \) to \( n + 1 \). The factors are often given by:

\[
\begin{align*}
    u &= e^{\sigma \sqrt{\Delta t}} = \frac{1}{d} \\
    p &= \frac{e^{r \Delta t} - d}{u - d} = 1 - p_d
\end{align*}
\]

The factor may seem random at first sight but ensures that the variance of the log of the price is \( \sigma^2 \Delta t \), similar to the Black-Scholes model. The risk neutral measure can be calculated and is given by:

\[
    p_u = \frac{e^{r \Delta t} - d}{u - d} = 1 - p_d
\]

provided that \( r^2 \Delta t < \sigma^2 \). If this does not holds, the market allows for arbitrage opportunities. Using this risk neutral measure and a discount factor, it is possible to translate the known option values at maturity back to the option value at \( t = 0 \). It is proven that binomial option pricing models in general do not have closed-form solutions but in case of vanilla put and call options the corresponding value converges to the analytical by Black-Scholes results for \( \Delta t \to 0 \).

5.5.2 Trinomial tree

The binomial tree model can be extended into the trinomial tree model. The key change is that the trinomial model allows for three possible paths at each node. The factor of the additional option is equal to one to guarantee that the trinomial tree is recombining. An example of factors is given by:

\[
\begin{align*}
    u &= e^{\sigma \sqrt{2 \Delta t}} = \frac{1}{d} \\
    m &= 1
\end{align*}
\]
This factor again makes sure that the variance of the log of the price is equal to \( \sigma^2 t \). An example of a risk-neutral measure is given, as long \( r^2 \Delta t < \sigma^2 \), by:

\[
p_u = \left( \frac{e^{\frac{r \Delta t}{2}} - e^{-\sigma \sqrt{\Delta t}}}{e^{\sigma \sqrt{\Delta t}} - e^{-\sigma \sqrt{\Delta t}}} \right)^2
\]

\[
p_d = \left( \frac{e^{\sigma \sqrt{\Delta t}} - e^{\frac{r \Delta t}{2}}}{e^{\sigma \sqrt{\Delta t}} - e^{-\sigma \sqrt{\Delta t}}} \right)^2
\]

\[
p_m = 1 - p_u - p_d,
\]

but in general there exist infinitely risk-neutral measures. This extra degree of freedom is an advantage of the trinomial model over the binomial model. The rest of the procedure is analogous to the procedure used in the binomial tree model. Under the right parametrization the trinomial tree is equivalent to the explicit finite difference method for option pricing. Note that a binomial tree that is only evaluated every other time period has an equivalent structure as the trinomial tree.

5.5.3 Similarities between binomial and trinomial trees

The trinomial tree possesses an extra degree of freedom. It is possible to use this to create a trinomial tree equivalent to the binomial tree. Start with the following observation: the amount of nodes at time \( n \) in the trinomial tree is equal to \( 2^n - 1 \) and in the binomial tree equal to \( n \).

It is easily checked that the amount of nodes in the binomial tree at odd time points is equal to the amount of nodes in the trinomial tree at all time points. Using the example of the previous section, the jump sizes, times and probabilities of the modified binomial tree can be compared. By definition the respective jump times compare like \( 2 \Delta t_{\text{bin}} = \Delta t_{\text{tri}} \). The jump sizes can be calculated and compared like:

\[
(u_{\text{bin}})^2 = e^{\sigma \sqrt{2 \Delta t_{\text{bin}}}} = e^{\sigma \sqrt{4 \Delta t_{\text{bin}}}} = e^{\sigma \sqrt{2 \Delta t_{\text{tri}}}} = u_{\text{tri}}
\]

As the size of a jump in the other direction is equal to the inverse of both numbers, \( (d_{\text{bin}})^2 = d_{\text{tri}}^2 \) also holds. Of course, jumping up and down in the modified binomial tree is equivalent to not jumping at all in the trinomial tree. Finally the jump probabilities are compared, note that the jump time in the trinomial tree is twice the jump time of the binomial tree:

\[
p_{uu}^{\text{bin}} = \left( \frac{e^{r \Delta t_{\text{bin}}}}{u - d} \right)^2 = \left( \frac{e^{r \Delta t_{\text{bin}}}}{e^{\sigma \sqrt{2 \Delta t_{\text{bin}}}} - e^{-\sigma \sqrt{2 \Delta t_{\text{bin}}}}} \right)^2 = \left( \frac{e^{\frac{r \Delta t_{\text{tri}}}{2}}}{e^{\sigma \sqrt{2 \Delta t_{\text{tri}}}} - e^{-\sigma \sqrt{2 \Delta t_{\text{tri}}}}} \right)^2 = p_{uu}^{\text{tri}}
\]

Calculating the other probabilities is similar and will also yield the required results. As a final observation, note that the discount factors behave like:

\[
(e^{-r \Delta t_{\text{bin}}})^2 = e^{-2r \Delta t_{\text{bin}}} = e^{-r \Delta t_{\text{tri}}}
\]

These calculations show that pricing using the trinomial tree under a certain measure and modified binomial tree is equivalent. The binomial tree can therefore be seen as an element of the collection of possible trinomial trees. In [20] Rubinstein shows that under a certain parametrization, binomial and trinomial trees are equivalent to finite difference methods.

5.5.4 Hull-White trinomial tree

As mentioned before, there exist infinitely many risk-neutral measures that can be used to create a trinomial tree. This freedom can be used to match the tree to a given interest rate model. This thesis will use the Hull-White model. Section 6 provides more details about this choice and the
implementation procedure. Following [21], an equivalent representation for the interest structure is used:

$$dr_t = (b_t - ar_t) \, dt + \sigma \, dW_t$$  \hspace{1cm} (11)

As an attempt to capture the mean-reversing property of the Hull-White model, alternative branching methods are used when the interest rate is far from the mean. Instead of the usual possible paths, the possible outcomes are: staying equal, going one step towards the mean and going two steps towards the mean. Figure 4a schematically shows this procedure. Constructing the tree is split up into two stages.

**The first stage.** First equation (11) is simplified by setting $r_0 = 0$ and $b_t = 0$. This defines new process $x_t$ that follows the stochastic differential equation:

$$dx = -ax \, dt + \sigma \, dW_t$$  \hspace{1cm} (12)

Note that this is a special case of the Vasicek model, namely with $b = 0$. Figure 4b shows an example of the shape of this tree at stage 1. The nodes are characterized by a horizontal index $i$ and a vertical index $j$. The horizontal distance between two nodes $\Delta t$ is seen as an input parameter. The vertical distance $\Delta x$ between nodes is calculated from input parameters as:

$$\Delta x = \sqrt{3} \sqrt{V} := \sqrt{3} \sqrt{\frac{\sigma^2}{2a} (1 - e^{-2a\Delta t})},$$

where $V$ is given by equation (3). Hull and White claim in [21] that theoretical work in numerical procedures suggest that this is a good choice from the standpoint of error minimization. Now define $M$ as expected change of $x_t$ in a $\Delta t$ number of time units. Calculate using equation (2) that:

$$Mx_t = \mathbb{E}(x_{t+\Delta t} - x_t | \mathcal{F}_t)$$
$$= \mathbb{E}(x_{t+\Delta t} | \mathcal{F}_t) - x_t$$
$$= e^{-a\Delta t}x_t - x_t$$
$$\Rightarrow M = (e^{-a\Delta t} - 1)$$

When $r$ is in node $(i,j)$ the expected change during a time step $\Delta t$ is $j\Delta r M$ and the variance of this change is $V$. According to [21], the probabilities that match this expected change and variance are:

$$p_u(A) = \frac{1}{6} + \frac{j^2M^2 + jM}{2}$$
$$p_m(A) = \frac{2}{3} - j^2M^2$$
$$p_d(A) = \frac{1}{6} + \frac{j^2M^2 - jM}{2}$$

for branching type A, see figure 4a. The probabilities of type B are given by:

$$p_u(B) = \frac{7}{6} + \frac{j^2M^2 + 3jM}{2}$$
$$p_m(B) = -\frac{1}{3} - j^2M^2 - 2jM$$
$$p_d(B) = \frac{1}{6} + \frac{j^2M^2 + jM}{2}$$
and the probabilities for type C are given by:

\[ p_u(C) = \frac{1}{6} + \frac{j^2M^2 - jM}{2} \]
\[ p_m(C) = -\frac{1}{3} - j^2M^2 + 2jM \]
\[ p_d(C) = \frac{7}{6} + \frac{j^2M^2 - 3jM}{2} \]

It is not yet clear when the branching type switches from type A to type B or C. If \( a > 0 \), it is necessary to switch to guarantee positive probabilities. Let \( j^* \) be the smallest integer larger than \(-\frac{0.184}{M}\). Following [22] all probabilities are positive when the switch to type C is at \( j_{max} := j^* \) and the switch to type B at \( j_{min} := -j^* \). This ends stage 1.

![Diagram of branching types](image)

**Figure 4:** Alternative branching in the Hull-White trinomial tree construction.

**The second stage** In order to implement the long term mean function \( b_i \), an \( \alpha_i \) is added to each node with index \( i \). To be able to calculate these \( \alpha_i \), first a value \( Q_{i,j} \) is added to each node. This \( Q_{i,j} \) can be interpreted as the value of a contract that pays one euro if node \((i,j)\) is reached and zero otherwise. Obviously \( Q_{0,0} = 1 \). Given the values of \( Q_{i,j} \) and the bond price \( P\left(0, (i+1)\Delta t\right)\), the value of \( \alpha_i \) can be calculated by looking at the equation:

\[ P\left(0, (i+1)\Delta t\right) = \sum_j Q_{i,j} e^{-\left(\alpha_i + j\Delta r\right)\Delta t}. \]

Thus the bond price at time \( i + 1 \) is equal to the sum of the discounted contracts \( Q_{i,j} \). This holds by the definition of \( Q_{i,j} \). This equation can be solved for \( \alpha_i \) by splitting the exponential, taking it out of the summation, taking the logarithm and finally shuffling the terms. The resulting formula is:

\[ \alpha_i = \log \sum_j Q_{i,j} e^{-j\Delta r\Delta t} - \log P\left(0, (i+1)\Delta t\right) \]

The resulting \( \alpha_i \) is then used to calculate the set \( Q_{i+1,j} \) using:

\[ Q_{i+1,j} = \sum_k Q_{i,k} p_{k,j} e^{-\left(\alpha_i + k\Delta r\right)\Delta t}. \]

That is the summation of the value of the states at time \( i \) multiplied by the probability \( p_{k,j} \) of getting from node \((i,k)\) to node \((i+1,j)\) multiplied by the relevant discount factor. After the new set \( Q_{i+1,j} \) is calculated, the value of \( \alpha_{i+1} \) can be calculated and the process is repeated until the last stage of the tree to obtain all \( \alpha_i \). Adding these \( \alpha_i \) to the stage one tree yields the Hull-White trinomial tree.
5.6 Linear regression

The goal of linear regression is to estimate the probability of an event with a set of explanatory variables. These explanatory variables are also known as risk drivers. Linear regression assumes that the relationship between the explanatory variables and the results is linear. This assumption is less strict than it seems at first sight, because the linearity is only present in the parameters. The explanatory variables can be arbitrarily transformed and one has to be careful not to overfit the model. Once the parameters are set, the relationship is modeled using an error variable \( \epsilon_i \). The model takes the form:

\[
Y = X\beta + \epsilon
\]

Here \( Y \) is a vector with realizations and \( X \) is a matrix with the explanatory variables for each individual as row vectors. Given this data, a vector \( \beta \) is proposed. This vector can be interpreted as a vector with partial derivatives of the explanatory variables. Eventually the vector \( \epsilon \) is such that the expression above holds. The proposed vector \( \beta \) is in such a way that it minimizes \( \epsilon \) for a certain norm. Usually this is done by a least squares method as this allows for the following closed-form expression for \( \beta \) using the generalized inverse of \( X \):

\[
\hat{\beta} = (X^T X)^{-1} X^T Y
\]

5.6.1 Generalized linear regression

The value of \( X\beta \) is called the linear predictor. It is easily verified that the range of the linear predictor is \( \mathbb{R}^d \). In some cases, this is not a desired property. Some situations restrict the outcomes to be positive or, when dealing with probabilities, within \([0, 1]^d\). This is fixed by adding a convenient link function \( g \) that makes \( g(Y) \) is linear. Models with a non-trivial link function are known as generalized linear regression models. In these models, the outcome values can be predicted by:

\[
\mathbb{E}Y = g^{-1}(X\beta)
\]

Most commonly used link functions to model probabilities are the logit and the quantile function of the normal distribution. The logit function is defined in equation (13).

A simple generalized linear regression model is the binomial model as there are only two outcomes for every individual. The general assumption within the binomial logit model is that the logit of the probabilities \( \pi_i \) for every individual follows a linear regression. The logit is defined as:

\[
\text{logit}(x) = \log\left(\frac{x}{1-x}\right), \quad x \in (0, 1)
\]

Note that the logit maps the unit interval onto the real numbers and that it is strictly increasing in its argument. Using the logit as a link function yields the following set of formulas to model the probabilities \( \pi_i \) given explanatory variables \( x_i' \):

\[
\text{logit}(\pi_i) = \log\left(\frac{\pi_i}{1-\pi_i}\right) = x_i'\beta
\]

The vector \( \beta \) in this case can be estimated for example by maximum likelihood estimation. Details of this procedure can be found in section 5.6.2. For now the probabilities are only described in an indirect way. Luckily it is not hard to find direct analytical formulas for \( \pi_i \). After taking the exponential on both sides and some rewriting:

\[
\frac{\pi_i}{1-\pi_i} = e^{x_i'\beta} \iff \pi_i = (1-\pi_i)e^{x_i'\beta}
\]

When combining this relation with the obvious constraint \( \pi_i + (1-\pi_i) = 1 \), it is easily verified that:

\[
\pi_i = \frac{e^{x_i'\beta}}{1+e^{x_i'\beta}} = F(x_i'\beta)
\]

\[
1-\pi_i = \frac{1}{1+e^{x_i'\beta}} = 1 - F(x_i'\beta)
\]
The function $F$ is a known distribution function called the cumulative logistic distribution function. This function could of course also have been calculated directly as the inverse of the link function:

$$F(x) = \frac{e^x}{1 + e^x} = \logit^{-1}(x)$$

The logit function and the cumulative logistic distribution function are plotted in figures 5a and 5b. Distribution functions have additional properties that are desired by the binomial logit model. An obvious but important consequence is that it maps $\mathbb{R}$ onto the unit interval, so all probabilities are well-defined. Distribution functions are also always monotone. This implies that an increase in an explanatory variable $x_j$ cannot decrease the probability if $\beta_j > 0$ and cannot increase if $\beta_j < 0$. This is of course in line with how a regression model is supposed to behave.

### 5.6.2 Estimation within the binomial logit model

Estimation within the binomial logit model is done using maximum likelihood estimation. This requires the likelihood function. To obtain this function, the following observation is made. In the binomial logit model every event $i$ is Bernoulli distributed with parameter $\pi_i$. Let $k_i$ be a random variable that is equal to 1 if event $k$ occurs and zero otherwise. Then the following expression for the density of $k_i$ can be found:

$$P(k_i) = \pi_i^{k_i} (1 - \pi_i)^{1-k_i}$$

Where $k_i \in \{0, 1\}$. Using the known expressions for $\pi_i$, this can be written as:

$$P(k_i) = \pi_i^{k_i} (1 - \pi_i)^{1-k_i} = F(x'_i \beta)^{k_i} (1 - F(x'_i \beta))^{1-k_i} = \left(\frac{e^{x'_i \beta}}{1 + e^{x'_i \beta}}\right)^{k_i} \left(1 - \frac{e^{x'_i \beta}}{1 + e^{x'_i \beta}}\right)^{1-k_i} = \left(\frac{e^{x'_i \beta}}{1 + e^{x'_i \beta}}\right)^{k_i} \left(\frac{1}{1 + e^{x'_i \beta}}\right)^{1-k_i} = \frac{e^{k_i (x'_i \beta)}}{1 + e^{x'_i \beta}}$$

This expression can be used to deduce the likelihood function $L$, simply by taking the product of this expression for all possible values of $i$. This requires independency of all event and while this
is in general not true, the assumption is made that all events are independent. If for example, in the prepayment model, a high curtailment event with corresponding penalty has taken place at a certain time, this client is less likely to relocate, since this is penalty free. The client has therefore paid a penalty that he would not have paid if he relocated without a preceding high curtailment event. The influences of these dependencies however are considered small and are therefore neglected.

\[
L(\beta) = \prod_{i=1}^{n} \mathbb{P}(k_i) = \prod_{i=1}^{n} \frac{e^{k_i(x_i'\beta)}}{1 + e^{x_i'\beta}}
\]

Because the goal is to estimate \( \beta \), the likelihood function is explicitly written as function of \( \beta \). Observe that the log-likelihood is given by:

\[
\log L(\beta) = \sum_{i=1}^{n} k_i(x_i'\beta) - \log(1 + e^{x_i'\beta})
\]

As usual, this function is maximized by setting the first order derivative with respect to \( \beta \) to zero:

\[
\frac{\partial \log L(\beta)}{\partial \beta} = \frac{\partial}{\partial \beta} \left( \sum_{i=1}^{n} k_i(x_i'\beta) - \log(1 + e^{x_i'\beta}) \right)
\]

\[
= \sum_{i=1}^{n} \left( k_i x_i' - \frac{e^{x_i'\beta} x_i'}{1 + e^{x_i'\beta}} \right)
\]

\[
= \sum_{i=1}^{n} \left( k_i - \frac{e^{x_i'\beta}}{1 + e^{x_i'\beta}} \right) x_i'
\]

\[
= 0
\]

Note that both the number of equations and the number of unknowns are equal to the length of vector \( \beta \). These equations can be numerically solved for \( \beta \) using numerical approximations methods, like Fisher scoring or Newton-Raphson. These are most often already included in statistical packages. A solution exists because for the range of the terms the following holds, provided not all \( k_i \) are equal:

\[
1 \leq \sum_{i=1}^{n} k_i \leq n - 1
\]

in combination with

\[
0 < \sum_{i=1}^{n} \frac{e^{x_i'\beta}}{1 + e^{x_i'\beta}} = \sum_{i=1}^{n} F(x_i'\beta) < n.
\]

As seen before \( F(x) \) can take all values between zero and one. Therefore there exists an \( \beta \) such that both terms are equal. If all \( k_i \) are equal, building a regression model is pointless. To check that the found solution is unique, the Hessian matrix of the likelihood function is calculated.

\[
\frac{\partial}{\partial \beta} \sum_{i=1}^{n} \left( k_i - \frac{e^{x_i'\beta}}{1 + e^{x_i'\beta}} \right) x_i' = -\sum_{i=1}^{n} \frac{e^{x_i'\beta}}{(1 + e^{x_i'\beta})^2} x_i x_i' = -\sum_{i=1}^{n} F(x_i'\beta)(1 - F(x_i'\beta)) x_i x_i'
\]

Negative definiteness can be checked using this expression. Let \( z \in \mathbb{R}^J \setminus \{0\} \), then:

\[
z^T \left( -\sum_{i=1}^{n} F(x_i'\beta)(1 - F(x_i'\beta)) x_i x_i' \right) z = -\sum_{i=1}^{n} F(x_i'\beta)(1 - F(x_i'\beta))(x_i^T z)^2 (x_i^T z)
\]

\[
= -\sum_{i=1}^{n} F(x_i'\beta)(1 - F(x_i'\beta)) \|x_i^T z\|^2 < 0
\]

because every non-zero norm is positive and \( 0 < F(x_i'\beta) < 1 \). Thus the Hessian is negative definite and therefore a unique maximizing vector \( \beta \) exists.
5.6.3 Extension to the multinomial logit model

In some practical cases, the observed data is not binomial, but divided into multiple categories. Instead of just zero or one, the data is given by a vector with zeros everywhere except for one value of one that corresponds to the observed event. It is possible to extend the results from the binomial logit model if an additional assumption is made. This assumption is known as the independence of irrelevant alternatives and is explained in the appendix. The idea is to pick one event as pivot and to model the relative odds of every other outcome with respect to the pivot using the binomial logit model. Finally, the sum of all probabilities is set to one to obtain the final model. This model is known as the multinomial logit model. Let there be a total of \( J \) possible outcomes and select the last outcome as a pivot. Let \( \pi_{ij} \) be the probability that individual \( i \) has outcome \( j \). Then the generalized logit formula is given by:

\[
\log \left( \frac{\pi_{ij}}{\pi_{iJ}} \right) = x_{ij}' \beta_j
\]

Note that, as expected, in the case of \( J = 2 \), the generalized logit formula is equal to the binomial logit formula. Following an analogous calculation as in the binomial logit model, this can be rewritten as:

\[
\pi_{ij} = \pi_{iJ} e^{x_{ij}' \beta_j}
\]

By substituting this equation for all values of \( j \) into the constraint that all probabilities sum up to one the normalization factor:

\[
1 = \sum_{j=1}^{J-1} \pi_{ij} e^{x_{ij}' \beta_j} + \pi_{iJ} \Rightarrow \pi_{iJ} = \frac{1}{\sum_{j=1}^{J-1} e^{x_{ij}' \beta_j} + 1}
\]

This result is substituted to find exact formulas for all probabilities.

\[
\pi_{ij} = \frac{e^{x_{ij}' \beta_j}}{\sum_{j=1}^{J-1} e^{x_{ij}' \beta_j} + 1}
\]

Comparing these expressions with the expression for \( \pi_{iJ} \), it can be checked that \( \beta_J = 0 \). Estimation of \( \beta_j \) from the data is also done using maximum likelihood estimation, although the procedure is somewhat more difficult. However, when all \( \beta_j \) are determined, it is easy to calculate the predicted probabilities given a matrix with vectors of explanatory variables.
6 Quadratic hedging in incomplete markets

The dynamic delta-hedging strategy of section 5.2.1 works excellent in theory in complete markets. However, because of the frequent rebalancing, the transaction costs can be high. This problem only occurs in practice as the models assumes a frictionless market. In an incomplete market even theoretically there does not exists a perfect hedge. A perfect hedge is both self-financing and replicating the pay-off. Following the beginning of [23], two partial hedging strategies are analyzed. The first approach, called local risk-minimization, replicates the pay-off, but is not self-financing. The risk process that is introduced by this fact is minimized over all strategies. The second approach, known as mean-variance hedging, is self-financing but is not able to perfectly replicate the pay-off. Instead, the quadratic hedging error at maturity is minimized. The choice between these strategies introduces model risk. While these strategies are valid for all incomplete markets, this thesis focusses on markets where the underlying follows a geometric Lévy process \( L_t \) with Lévy triplet \( (b, c, \nu) \) under a measure \( P \) which, in general, is not a risk-neutral measure. The resulting process is assumed to be a semi martingale. The considered financial products are European options with payoff \( F(S_T) \). Define the function \( f \) by

\[
f(x) := F(e^x).
\]

Introduce a suitable damping factor \( R \neq 0 \) to write down the dampened payoff function in order to make the corresponding integrals converge

\[
g(x) := e^{-Rx}f(x).
\]

Finally denote the Fourier transform of \( g \) as \( \hat{g} \) and introduce the notation \( \hat{f} \) to write

\[
\hat{g}(u) := \int_{\mathbb{R}} e^{iux}g(x) \, dx = \hat{f}(u + iR).
\]

6.1 Lévy theory

Every underlying process encountered until now has been continuous. Sometimes however, the underlying process should allow for jumps. This for example because jumps have been observed in market data or that the underlying process contains jumps by definition. This second reason is relevant for the LTV-option. A prepayment event is a jump in the outstanding notional and since this is part of the underlying process, a continuous underlying process does not represent the situation very well. A class of stochastic processes that allows for jumps are Lévy processes. Some general ideas and properties of Lévy processes will be shown using the following example. Let the process \( L_t = (L_t)_{t \in [0,T]} \) be a Lévy jump-diffusion process, i.e. the sum of a drift, a Brownian motion and a compensated compound Poisson process. Therefore the paths are described by

\[
L_t = bt + \sigma W_t + \left( \sum_{k=1}^{N_t} J_k - t\lambda \right)
\]
where \( b \in \mathbb{R}, \sigma \in \mathbb{R}_{\geq 0}, W \) is a standard Brownian motion, \( N \) is a Poisson process with parameter \( \lambda \) and \( J = (J_k)_{k \geq 1} \) is a i.i.d. sequence of random variables with known distribution function \( F \) and mean \( \kappa \). This process is known as a compensated Lévy jump-diffusion process. It will be convenient to know the characteristic function of \( L_t \). This can be calculated straight from the definition. More details can be found in section 3 and appendix B of [24].

\[
\mathbb{E}[e^{iuL_t}] = \mathbb{E}\left[ \exp\left( iu(\frac{t}{n} \sum_{k=1}^{N_t} J_k - t\lambda \kappa) \right) \right]
\]

\[
e^{iubt} \mathbb{E}\left[ \exp\left( iu(\sigma W_t + \lambda t\kappa) \right) \right]
\]

\[
e^{iubt} \mathbb{E}\left[ \exp\left( iu\sigma W_t \right) \right] \mathbb{E}\left[ \exp\left( iu \sum_{k=1}^{N_t} J_k - iut\lambda \kappa \right) \right]
\]

\[
e^{iubt} \exp\left[ -\frac{1}{2} \sigma^2 u^2 t \right] \mathbb{E}\left[ \exp\left( iu J - iJ u \right) \right]
\]

\[
e^{iubt} \exp\left[ -\frac{1}{2} \sigma^2 u^2 t \right] \mathbb{E}\left[ \int_{\mathbb{R}} (e^{iux} - 1 - iux) \lambda F(dx) \right]
\]

\[
\exp\left[ t (iub - \frac{u^2 \sigma^2}{2} + \int_{\mathbb{R}} (e^{iux} - 1 - iux) \lambda F(dx)) \right]
\]

Observe that the time factorizes out and that the drift, diffusion and jump components are nicely separated. This is not only true in this example, but it turns out this property holds for general Lévy processes. Also note that the jump part factorizes into the number of jumps and the distribution of the jump size. This property does not hold in general. Lévy processes are closely related to infinitely divisible random variables. First the definition of infinitely divisible is presented

**Definition 6.0.2.** The law \( P_X \) of a random variable \( X \) is infinitely divisible if for all \( n \in \mathbb{N} \) there exists i.i.d. random variables \( X_1, \ldots, X_n \) such that:

\[ X^{\dagger} \triangleq X_1 + \ldots + X_n \]

or equivalently in terms of the characteristic functions \( \phi_X(u) \) if for all \( n \in \mathbb{N} \) there exists i.i.d. random variables \( X_1, \ldots, X_n \) such that:

\[ \phi_X(u) = \left( \phi_{X_1}(u) \right)^n \]

The compensated Lévy jump-diffusion process is infinitely divisible as:

\[ \mathbb{E}[e^{iuL_t}] = \left( \exp\left[ \frac{t}{n} \left( iub - \frac{u^2 \sigma^2}{2} + \int_{\mathbb{R}} (e^{iux} - 1 - iux) \lambda dF(dx) \right) \right] \right)^n = \left( \mathbb{E}[e^{iuL_{\frac{t}{n}}}] \right)^n \]

The next theorem generalizes the formula above. The result is known as the Lévy-Khintchine formula and introduces a triplet \( (b, c, \nu) \), consisting of a drift term, a diffusion coefficient and measure. This triplet be referred to as the Lévy triplet and the measure as the Lévy measure. For a proof, see lemma 7.8 of [25].

**Theorem 6.1.** The law \( P_X \) of a random variable \( X \) is infinitely divisible if and only if there exists a triplet \( (b, c, \nu) \), with \( b \in \mathbb{R}, c \in \mathbb{R}_{\geq 0} \) and a measure \( \nu \) satisfying \( \nu(\{0\}) = 0 \) and \( \int_{\mathbb{R}} (1 \wedge x^2) \nu(dx) < \infty \), such that

\[ \phi_X(u) = \mathbb{E}[e^{iuX}] = \exp\left[ iub - \frac{u^2 c}{2} + \int_{\mathbb{R}} (e^{iux} - 1 - iux \mathbb{1}_{(|x| < 1)}) \nu(dx) \right] \]

As the characteristic function of \( X \) completely determines its distribution by numerical inversion, a Lévy process is totally described by its Lévy triplet. Comparing the expressions, the Lévy
triplet of the Lévy jump-diffusion process is given by \( (t(b - \lambda \int_{|x| \geq 1} x F(dx)), \sigma^2 t, (\lambda \cdot F)t) \). Note that the time factors out, as expected from the alternative definition of infinitely divisible. The connection between the Lévy triplet and the corresponding Lévy process is made clear in the following theorem. This theorem is also known as the Lévy-Itô decomposition. A proof can be found in chapter 2 of [26].

**Theorem 6.2 (Lévy-Itô decomposition).** Let \((b, c, \nu)\) be a triplet with \(b \in \mathbb{R}, c \in \mathbb{R}_{\geq 0}\) and a measure \(\nu\) satisfying \(\nu(\{0\}) = 0\) and \(\int_{\mathbb{R}} (1 \wedge x^2) \nu(dx) < \infty\). Then there exists a probability space \((\Omega, \mathcal{F}, \mathbb{P})\) on which four independent Lévy processes exist

- \(L^{(1)}\), a constant drift \(b\)
- \(L^{(2)}\), a Brownian motion with variance parameter \(c\)
- \(L^{(3)}\), a compound Poisson process
- \(L^{(4)}\), a martingale with an a.s. countable number of jumps less than 1 on each finite time interval

Taking \(L = L^{(1)} + L^{(2)} + L^{(3)} + L^{(4)}\) there exists a probability space on which a Lévy process \(L = (L_t)_{t \in [0,T]}\) is defined with characteristic function

\[
\phi_t(u) = \exp \left[ t \left( iub - \frac{u^2 c}{2} + \int_{\mathbb{R}} \left( e^{ix} - 1 - ix 1_{|x| < 1} \right) \nu(dx) \right) \right], \quad t \in [0,T].
\]

The expression in round brackets is known as the Lévy exponent and denoted as \(\psi(u)\). Note that the cut off value 1 inside the indicator function is arbitrarily. There has to be a boundary between large and small jumps, but the value of this boundary could be any \(\epsilon > 0\). To force the characteristic function in the form of the Lévy-Khintchine formula, the correction in the integral in order to create the indicator function is captured in the drift. Look for example at the Lévy triplet of the compensated Lévy jump-diffusion process.

### 6.2 Market incompleteness

The hedging strategies discussed before required the derivative to the underlying process. For Lévy processes however, this derivative is not well-defined at jump spots. This makes the delta-hedge approach impossible. In general, there exists no perfect hedging method when the underlying process is Lévy. This is shown by the following argument. Assume that the price process of a financial asset is modeled as an exponential Lévy process under the real measure \(\mathbb{P}\) and a risk-neutral measure \(\mathbb{Q}\). Denote the corresponding Lévy triplets with \((b, c, \nu)\) and \((b', c', \nu')\). By theorem 1.20 of [27] the Lévy triplets are related as:

\[
\begin{align*}
    c' &= c \\
    b' &= b + c \beta + \int_{|x| < 1} x(Y(x) - 1) \nu(dx) \\
    d\nu' &= Y \, d\nu
\end{align*}
\]

for some \(\beta \in \mathbb{R}\) and a function \(Y\) from the support of \(\nu\) to \(\mathbb{R}_{\geq 0}\). By definition of the risk-neutral measure, the discounted stock price process \((e^{-rt} S_t)_{t \in [0,T]}\) is a martingale and thus also a local martingale. By equation (3.2) from [28] it holds that

\[
    b' + \frac{c'}{2} + \int_{\mathbb{R}} (e^x - 1 - x 1_{|x| < 1}) \, d\nu' = r.
\]
Combining both relations for \( b' \) the following equation has to hold
\[
0 = b + c\beta + \int_{\mathbb{R}} x 1_{|x|<1}(Y(x) - 1)\nu(dx) - r + \frac{c'}{2} + \int_{\mathbb{R}} (e^x - 1 - x 1_{|x|<1})\nu'(dx)
\]
\[
\Leftrightarrow 0 = b - r + c\left(\beta + \frac{1}{2}\right) + \int_{\mathbb{R}} ((e^x - 1)Y(x) - x 1_{|x|<1})\nu(dx).
\]

This is one equation with unknown \( \beta \in \mathbb{R} \) and an unknown function \( Y \). Every solution \((\beta, Y)\) to above equation corresponds to a different risk-neutral measure and therefore the market is in general incomplete. Note that if no jumps are present, the integral equals zero and the risk-neutral measure only depends on the value of \( \beta \). The only solution is given by
\[
\beta = \frac{2r - 2b - c}{2c}
\]
and the market becomes complete again, as there is only one risk-neutral measure. This is of course in line with for example the delta hedging method and the Black-Scholes approach.

### 6.3 Local risk-minimization approach

As the discounted stock price process is a special semi-martingale, it possesses a unique Doob-Meyer decomposition
\[
\hat{S}_t = e^{-rt}S_t = S_0 + M + A
\]
with \( S_0 \) the stock price at \( t = 0 \), \( M \) a local martingale and \( A \) a predictable process with finite variation. Both \( M \) and \( A \) start at zero. It turns out that in order to find a local risk-minimization approach it is very convenient to write down the Föllmer-Schweizer decomposition of the discounted stock price process.

**Definition 6.2.1.** Let \( \hat{S}_t \) be a special semimartingale with Doob-Meyer decomposition \( \hat{S}_t = S_0 + M + A \). An \( \mathcal{F}_T \)-measurable random variable \( H_T \) admits a Föllmer-Schweizer decomposition if there exist a constant \( H_0 \), a predictable process \( \chi \) and a martingale \( N \) starting at zero such that \( N \) is orthogonal to \( M \) and
\[
H_T = H_0 + \int_0^T \chi_t \, d\hat{S}_t + N_T
\]
The process \( \chi_t \) represent the number of assets in the portfolio at time \( t \) and is also referred to as the hedging number. In [29] is shown that this decomposition exists for exponential Lévy models.

It also provides formulas that allow the process \( \chi \) to be expressed as
\[
\chi(t, S_{t-}) = \frac{c\Delta(t, S_{t-}) + \frac{1}{S_{t-}}\int_{\mathbb{R}}(e^z - 1)|P(t, S_{t-} - e^z) - P(t, S_{t-})|\nu(dz)}{c + \int_{\mathbb{R}}(e^z - 1)^2\nu(dz)}
\]
In case no jumps are present, both integrals equal zero so the hedging number is equal to delta. The formulas for \( P(t, S_{t-}) \) and delta below are from [30]. A proof and the required assumptions can also be found there.

\[
P(t, S_t) = e^{-r(T-t)}E[F(S_T)|\mathcal{F}_t] = \frac{e^{-r(T-t)}}{2\pi} \int_{\mathbb{R}} \hat{f}(u + iR)\phi_{T-t}(-u - iR)S_t^{R-1-iu} \, du
\]
\[
\Delta(t, S_t) = \frac{\partial P}{\partial S_t}(t, S_t) = \frac{e^{-r(T-t)}}{2\pi} \int_{\mathbb{R}} (R - iu)\hat{f}(u + iR)\phi_{T-t}(-u - iR)S_t^{R-1-iu} \, du
\]
where the expectation is taken under the minimal martingale measure. The minimal martingale measure is the measure equivalent to the real world measure \( \mathbb{P} \) that minimizes the Kullback-Leibler divergence. The Kullback-Leibler divergence is also known as the relative entropy. For distributions \( \mathbb{P} \) and \( \mathbb{Q} \) of a continuous random variable, the relative entropy is given by :
\[
D_{KL}(\mathbb{P}||\mathbb{Q}) = \int_{-\infty}^{\infty} p(x) \log \left( \frac{p(x)}{q(x)} \right) dx,
\]
where \( p \) and \( q \) denote the density functions of \( \mathbb{P} \) and \( \mathbb{Q} \). Because the measures are equivalent, it is not possible that \( p(x) \neq 0 \) while \( q(x) = 0 \) or \( p(x) = 0 \) while \( q(x) \neq 0 \). In case \( p(x) = q(x) = 0 \) the integrand is assumed to be zero as \( \lim_{x \to 0} x \log \left( \frac{1}{x} \right) = 0 \). The relative entropy says something about how close the measures are. For example, the relative entropy of \( \mathbb{P} \) and \( \mathbb{P} \) is zero. More details and examples of the minimal martingale measure can be found in [31]. Even though the local risk-minimization strategy is, at least theoretically, known, formulas that allow for easier implementation exist. These formulas and their proofs are found in Lemma 3.3 and Proposition 3.1 of [30]. The results are presented in terms of a complex measure \( \Pi \) that satisfies

\[
F(S) = \int_{\mathbb{C}} S^z \Pi(\,dz) \]

If the measure exists, it is given by

\[
\Pi(\,dz) = \frac{1}{2\pi i} \mathbf{1}_{R+i\mathbb{R}}(z) \hat{f}(iz) \,dz
\]

Examples of this measure for common European options can be found in [32] and [33] and are given by:

\[
\Pi_{\text{call}}(\,dz) = \frac{1}{2\pi i} \mathbf{1}_{R+i\mathbb{R}} \frac{K^{1-z}}{z(z - 1)} \,dz \tag{14}
\]

\[
\Pi_{\text{digital call}}(\,dz) = \frac{1}{2\pi i} \mathbf{1}_{R+i\mathbb{R}} \frac{K^{-z}}{z} \,dz \tag{15}
\]

\[
\Pi_{\text{digital put}}(\,dz) = -\frac{1}{2\pi i} \mathbf{1}_{R+i\mathbb{R}} \frac{K^{-z}}{z} \,dz \tag{16}
\]

Note that this measure is only non-zero when the real part of the argument equals the damping factor \( R \). The components of the Föllmer-Schweizer decomposition are given by

\[
H_t = e^{-rt} \int_{\mathbb{C}} e^{\eta(z)(T-t)} S^z_t \Pi(\,dz) \tag{17}
\]

\[
\chi_t = \int_{\mathbb{C}} \mu(z)e^{\eta(z)(T-t)} S^z_t \Pi(\,dz) \tag{18}
\]

\[
N_t = H_t - H_0 - \int_0^t \chi_u \,d\tilde{S}_u \tag{19}
\]

where the functions \( \mu(z) \) and \( \eta(z) \) are given by

\[
\mu(z) = \frac{\psi(-iz - i) - \psi(-iz) - \psi(-i)}{\psi(-2i) - 2\psi(-i)}
\]

\[
\eta(z) = \psi(-iz) - \mu(z)(\psi(-i) - r) - r
\]

This direct formula for the hedging number can be used for numerical experiments, as is done in section 8.1.

### 6.4 Mean-variance hedging

This strategy is in general hard to compute, but in case the price process is modelled by exponential Lévy models the results follow from the local risk-minimization strategy. In [30] a formula is presented to calculate the mean-variance hedging number from the local risk-minimization hedging number. The variance-optimal initial capital \( V_0 \) is still equal to \( H_0 \) and variance-optimal hedging strategy \( \phi \) is given by

\[
\phi_t = \chi_t + \frac{\lambda}{\tilde{S}_t^{-}} \left( H_{t^-} - V_0 - G_{t^-}(\phi) \right)
\]
where the constant $\lambda$ and the cumulative gain process $G(\phi)$ are given by

$$
\lambda = \frac{\psi(-i) - r}{\psi(-2i) - 2\psi(-i)} \quad \text{and} \quad G_t(\phi) = \int_0^t \phi_s \, d\hat{S}_s
$$

The denominator in the last expression is non-zero as this would imply that the stock price process is deterministic. Note that because $V_0 = H_0$, the hedging number $\phi_0$ is equal to $\chi_0$. 
Model choices and implementation

For most experiments, an interest rate model is required. There are many different models available and most common ones are described with their own advantages and drawbacks in section 4. A first characterization can be made by noting whether or not the model allows for negative interest rates. At the time of writing, negative interest rates are no rarity to the market and therefore it will be convenient that the used model will allow for negative interest rates as well. The most common interest rate models are described in section 7. The choice between Merton’s model, Vasicek model, Hull-White model and the CIR-model is made by looking at the flexibility of the model. As Merton’s model is not mean-reversing and the Vasicek model cannot be calibrated to a yield curve, the final choices will be the Hull-White mode and the CIR-modell. Currently at ABN AMRO, a regime switch model is used. This is a model where the stochastic differential equation depends on the current interest rate. For example a regime switch model can behave like a Hull-White for $r_t < 2$ and like a CIR model for $r_t \geq 2$.

7.1 Calibrating the Hull-White model

In order to use the Hull-White model as input in the option pricing models, it should represent the current market the best way possible. This is done by using the freedom in the parameters. The model is calibrated to a short-term interest rate structure and the yield curve of 7 April 2017 provided by the European Central Bank, see figure 9. The short-term interest rates are provided by the Organization for Economic Cooperation and Development and shown in figure 6.
7.1.1 Finding the mean-reversion and volatility parameters

The total available data consists of 189 monthly observations. Going back further is not possible as before 2001 the data is relative to the gulden instead of the euro. Implementing the linear regression and maximum likelihood estimation methods of section 4.2.1 yields the following results:

\[
\begin{align*}
    a_{LR} &= 0.09902 \\
    \sigma_{LR} &= 0.005338 \\
    a_{MLE} &= 0.09902 \\
    \sigma_{MLE} &= 0.005305
\end{align*}
\]

These values are in line with the observations in [8], as they claim that common values in financial markets for the parameters are \(a \sim 0.01 - 0.1\) and \(\sigma \sim 0.001 - 0.01\). As more data leads to more reliable estimations, it would be convenient if additionally also 187 gulden data points from January 1986 until June 2001 could be used as input. However this gulden data alone yields the following results:

\[
\begin{align*}
    a_{LR} &= 0.06474 \\
    \sigma_{LR} &= 0.0082938
\end{align*}
\]

These values differ about 50\% with the euro parameters and are therefore not used as input to calibrate the model as the volatility structure does not seem to be identical.

7.1.2 Finding confidence intervals

As all the available data is used to estimate the parameters, it is not possible to calculate confidence intervals the usual way. Therefore first a number of new outcomes is simulated using the Vasicek model and the estimated parameters. The length of these paths will also be equal to 189. From a sample of one million paths, the corresponding 99.99\%-confidence intervals are calculated to check for bias. The 99.99\%-confidence level is arbitrary and other confidence levels could also have been considered. The results are:

\[
\begin{align*}
    a &\in (0.1134514, 0.1136929) \\
    \sigma &\in (0.0053612, 0.0053634)
\end{align*}
\]

Note that both seem to be biased as the initial estimates are not within the 99.99\%-confidence intervals, but the bias of \(\sigma\) is very small and can be neglected in comparison to the sample standard deviation of \(\sigma\), which is equal to 0.000289. The bias of \(a\) is relatively large but still within one sample standard deviation of 0.041045 away from the estimated value. Both biases are caused by the low sample length, as can be seen by increasing the length of the paths to 10,000. The confidence intervals then become:

\[
\begin{align*}
    a &\in (0.099423, 0.099459) \\
    \sigma &\in (0.00533457, 0.00533488)
\end{align*}
\]

Both are clearly closer to the input values although a small bias is still present. An explanation for the bias is that the low sample size creates a standard error in \(a\) that clashes with the fact that \(a\) cannot become negative. It is not possible to differ more than 0.1 in the negative direction, but it is possible in the positive direction as can be seen from figure 7. To compensate for this phenomenon a new \(a\) found using trial and error such that the outcome of the low sample size estimate is close to 0.09902. This way the calibrated Hull-White model matches the available data as good as possible. This new found value is 0.085 and is still within a standard deviation. The standard deviation in the mean reversion parameter is very large and therefore a more sophisticated method is required before using the results in practice. All further numerical experiments in this thesis
assume that $a = 0.085$ and $\sigma = 0.00533$, but sensitivity analysis will be provided. Finally, note that the difference between the linear regression method and maximum likelihood estimation is well within the standard deviations of both parameters. Therefore there is no need to use both and from now on only the linear regression method is used. This because experiments show that $\sigma_{LR} \geq \sigma_{MLE}$, as can be seen in figure 8 and therefore the linear regression method provides the worst case scenario.

7.1.3 Finding the long term mean function

Using the found volatility and mean reversion parameter, the long term mean function can be calculated. The method is already described in section 4.3. The input curve is based on AAA rated bonds and is provided by the European Central Bank and shown in figure 9.

The yield can be transformed into bond prices by the well-known formula:

$$P(0, T) = e^{-\text{yield}(T) \cdot T}$$

The resulting bond price curve is shown in figure 10. As recommended by [8] cubic splines were used as interpolation method. Finishing the calibration results in the long term mean function $b_t$.
Figure 9: Yield curve of the seventh of April provided by the ECB. The horizontal axis represents the time in years.
presented in figure 11. The warning of [8] about linear interpolation is very relevant as is shown in figure 12.

7.2 Calibrating the CIR-model

Similarly to the Hull-White model, the CIR-model should represent the market as close as possible. This starts by picking a reasonable shift parameter \( \nu \). This parameter can be interpreted as a lower bound of the interest rate. Based on current observations in the market, this thesis uses \( \nu = 0.04 \) for numerical experiments. The same data is used for calibration, namely the past short term interest rates provided by the Organization for Economic Cooperation and Development and the yield curve of April 7 provided by the European Central Bank. These can respectively be found

Figure 10: Bond Prices. The horizontal axis represents the time in years.

Figure 11: Long term mean function. The horizontal axis represents the time in years.

Figure 12: Long term mean when using linear interpolation
7.2.1 Finding the mean-reversion and volatility parameters

The mean-reversion and volatility parameters are found by implementing the maximum likelihood estimation methods described in [10]. This yields the following results.

\[ \alpha_{CIR} = 2.02 \]
\[ \sigma_{CIR} = 0.086 \]

These values are used as input parameters to find the time-dependent long term mean function.

7.2.2 Finding the long term mean function

Ideally, one would have a similar formula to equation (7) to find the long term mean function. Unfortunately, such a formula was not available. Instead a recursive method is used to find the long term mean function. The starting point is the long term mean function of the Hull-White model. Using this input, bond prices are calculated and compared with the real bond prices. A fraction of the difference is added to the initial long term mean function to obtain the second step long term mean function. This process is repeated until the difference in the calculated and observed bond prices is within a given tolerance level. For the experiments in this thesis, the fraction was set to 0.2 and the tolerance level was set to $10^{-6}$.

7.3 Estimations and calculations regarding the LtV process

Recall the definition of the loan-to-value process:

\[ (LtV)_t = \frac{(Outstanding notional)_t}{(House price)_t} \]

In order to model the LtV, a model for the two processes is required. Note that the implicit assumption is made that the processes are independent of each other. A prepayment event translates to a jump in the LtV. Therefore it is only logical to try to use Lévy theory and quadratic hedging to price and hedge the LtV option. As stated in [23], this requires a Lévy process \( L_t \) such that:

\[ (LtV)_t = (LtV)_0 e^{L_t} \]

This thesis discusses two approaches to approximate the required Lévy process.

7.3.1 Modelling the outstanding notional

The outstanding notional can change in two ways. By contractual repayments or non-contractual repayments. Non-contractual repayments will from now on be denoted as prepayments. The prepayment part will be estimated from data. As prepayments sometimes come with a penalty the model distinguishes the contracts that can prepay for free, since prepayment behaviour will differ in both cases. From a prepayment dataset the prepayment behaviour is estimated. Figure 13 shows the results for penalty-free contracts. Figure 14 shows the results for contracts with penalties. As from now on the focus is only on bullet loans, prepayments are fully responsible for changes in the outstanding notional. The only unknown in the jump size distribution is the jump frequency. This is estimated by the total number of partial prepayments in the dataset divided by the total number of years the relevant contracts have been in a linked dataset containing all loanparts that could be prepaid. The results are:

\[ \lambda_{\text{penalty}} = \frac{29466}{7815158} = 0.00377 \]
\[ \lambda_{\text{penalty free}} = \frac{93817}{7975887} = 0.0118 \]
As expected, clients with contracts that can prepay for free have a higher prepayment rate than the clients that have to pay a penalty. Surprisingly, the jump size distribution is almost equal. Note that only partial prepayments are used as a full prepayment ends the mortgage and therefore sets the value of every option to zero.

![Figure 13: Prepayments are penalty-free](image)

![Figure 14: Prepayments come with penalty.](image)
7.3.2 Modelling the house price

To model the house price, a dataset containing house prices at different dates is used to find an empirical distribution of the yearly log-returns. As the LtV is inversely proportional to the house price, the result of the empirical distribution is multiplied by $-1$ to find the relevant distribution. This empirical distribution is shown in figure 15.

![Normalized histogram of minus the yearly log-returns, obtained from a dataset containing house prices.](image)

**Figure 15**

**Log-normal returns** The first approach is to assume that the log-returns are distributed $N(\mu, \sigma^2)$. The LtV process in this case is given by:

$$ (LtV)_t = \frac{\text{(Outstanding notional)}_t}{\text{(House price)}_t} = \frac{N_0 e^{Pt}}{V_0 e^{L(0, \sigma^2)}} = (LtV)_0 e^{-bt + \sigma L(0,1) + Pt} = (LtV)_0 e^{L_t} \quad (20) $$

The best approximation in the $L^2$ sense of $\mu$ and $\sigma$ is calculated by brute force and given by:

$$ \mu = 0.02485 $$

$$ \sigma = 0.0503 $$

The brute force method simply loops over the variables and finds the best approximation within the loop. Then the search area is decreased with a factor close to one and the process is repeated until one reaches a given tolerance level. The factor used in experiment is 0.9 and each loop contains 10,000 possible values per parameter. In view of equation (20) the drift of the Lévy process $L_t$ is equal to $-\mu$ and the diffusion coefficient is equal to $\sigma$. From figure 16a it is evident that there is no perfect fit with the fitted normal distribution. The QQ-plot, shown in figure 16b and a Kolmogorov-Smirnov test confirm this observation.
Mortgage Options

Hyperbolic Distribution  The tails of a normal distribution are not fat enough to correctly capture the data. In finance, often hyperbolic distributions are used to model data with fatter tails. There are four input parameters that determine the distribution, namely location parameter \( \mu \), shape parameter \( \alpha \), asymmetry parameter \( \beta \) and scale parameter \( \delta \). The probability density function of a hyperbolic distribution is given by:

\[
P(X \leq x) = \frac{\sqrt{\alpha^2 - \beta^2}}{2\alpha \delta K_1(\delta \sqrt{\alpha^2 - \beta^2})} e^{-\alpha \sqrt{\delta^2 + (x-\mu)^2} + \beta (x-\mu)}
\]

The function \( K_\alpha(x) \) is known as the modified Bessel function of the second kind. An expression is given in [34]:

\[
k_\alpha(x) = \frac{\pi}{2\sin(\alpha \pi)} \sum_{m=0}^{\infty} \left[ \frac{1}{m! \Gamma(m - \alpha + 1)} \left( \frac{x}{2} \right)^{2m-\alpha} + \frac{1}{m! \Gamma(m + \alpha + 1)} \left( \frac{x}{2} \right)^{2m+\alpha} \right]
\]

where the limit is taken if \( \alpha \in \mathbb{Z} \). The modified Bessel function of the second kind is also one of the independent solutions to the second-order differential equation:

\[
u^2 \frac{\partial^2 y}{\partial u^2} + u \frac{\partial y}{\partial u} - (u^2 + \alpha^2) y = 0 \tag{21}
\]

Note that for now, this function is only used in the normalization factor and explicit calculation can therefore be avoided. The task is to find estimates for the parameters of the hyperbolic distribution based on the data. Different methods are suggested by [35]. However, as the procedure only needs to be carried out one time, a brute force method that minimizes the difference with the data in \( L^2 \) sense is sufficient, although one has to be careful not to find only a local minimum. This brute force method was used to find that the best approximation is very close to:

\[
(\mu, \alpha, \beta, \delta) = (-0.01232, 21.860, -5.9827, 0.005938)
\]

The corresponding fit and QQ-plot are shown in figure 17. Writing down the resulting Lévy process is somewhat difficult due to the indirect definition of the modified Bessel function of the second kind. Luckily this is not needed to calculate the resulting value and hedging number as only the Lévy exponent is required. An expression for the Lévy exponent of hyperbolic distributions can, after setting \( \lambda = 1 \) and taking the logarithm, be found in [35] and is given by:

\[
\psi_{hyp}(z) = i \mu \nu + \log \left[ \frac{\sqrt{\alpha^2 - \beta^2}}{\sqrt{\alpha^2 - (\beta + iz)^2}} \frac{K_1(\delta \sqrt{\alpha^2 - (\beta + iz)^2})}{K_1(\delta \sqrt{\alpha^2 - \beta^2})} \right] \tag{22}
\]
Note that now explicit values of the Bessel function of the second kind need to be calculated. This can be done by using the finite difference methods discussed in section 5.4 on equation (21) or by using the following integral formula found in [36]:

\[ K_1(x) = \int_0^\infty e^{-x \cos(it)} \cos(it) \, dt \]

Fortunately, Matlab already has an implemented function that does this by itself. Expression (22) is used as input of the quadratic hedging formulas (17) and (18).

Figure 17: Fit of the hyperbolic distribution found using brute force on the top and the corresponding QQ-plot below.

7.3.3 Calculating the pay-off of the LtV option

At this point, the only unknown input of the quadratic hedging formulas (17) and (18) is the measure \( \Pi(dz) \). Figure 18 shows the dependency of the mortgage rate on the LtV ratio for a linear loan with different maturities. Also for different types of loans, with almost no exceptions, the risico-opslag equals 0.2% if \( \text{LtV} \in [0.65, 0.85] \) and 0.7% if the LtV ratio is above 0.85. From the perspective of the client, this can be seen as the pay-off of the sum of two digital put options like:

\[ X_{LtV} = 0.002 \cdot 1_{\{(LtV)_T<0.65\}} 1_{\{(LtV)_T>0.65\}} + 0.005 \cdot 1_{\{(LtV)_T<0.85\}} 1_{\{(LtV)_T>0.85\}} \]

Of course, from the perspective of the bank, the pay-off is minus \( X_{LtV} \). Using equation (16), the complex measure \( \Pi(dz) \) is given by:

\[ \Pi(dz) = \frac{1}{2\pi i} R + \Re \left( 0.002 \frac{0.65}{z} 1_{\{(LtV)_T>0.65\}} + 0.005 \frac{0.85}{z} 1_{\{(LtV)_T>0.85\}} \right) dz \] (23)
with \( R < 0 \). This measure completes the input of the quadratic hedging formulas (17) and (18). From these formulas it is also immediately clear that both the hedging number and the value of the \( \text{LtV} \) option equal zero if the initial \( \text{LtV} \) ratio is below 0.65.

### 7.4 Hull-White Monte Carlo with variance reduction methods

The results of section 7.1 are used as input of a Monte Carlo simulation as described in section 5.3. The starting point of each interest rate path is set equal to the starting point of the forward curve at \( r_0 = -0.0078 \). This is motivated by the following calculation using Fubini’s theorem and Leibniz integral rule:

\[
 f(0, 0) = -\frac{\partial}{\partial T} \log \left( P(0, T) \right) \bigg|_{T=0} = -\frac{\partial}{\partial T} \log \left( \mathbb{E} \left[ e^{-\int_0^T r_s \, ds} \right] \right) \bigg|_{T=0} = \mathbb{E} \left[ \frac{\partial}{\partial T} \left( \int_0^T r_s \, ds \right) \right] \bigg|_{T=0} = r_0
\]

The Euler scheme of the Hull-White stochastic differential equation (6) is given by:

\[
 r_i = r_{i-1} + a(\theta_i - r_{i-1}) \, dt + \sigma \sqrt{dt} Z
\]

The spline interpolation used to determine the function \( \theta \) should at least contain the time points used in the Monte Carlo simulation. As an example, 10,000 sample path are generated with the Euler scheme with \( a = 0.085 \), \( \sigma = 0.005335 \), \( T = 10 \) and \( dt = 0.01 \) and plotted in figure 19. Due to the mean reversion, according to this model the interest rate stays between -0.02 and 0.06. In order to calculate confidence intervals, also the variance is required. To compare different variance reduction methods, the variance is determined by calculating bond prices using Monte Carlo multiple times and taking the variance of these outcomes. The variance reduction methods used are antithetic variables and a Halton series. To avoid situations as in figures 3a and 3b the \( p_i \) are the first 360 prime numbers and a burn in period in the order of \( 10^8 \) is present. To be able to calculate confidence intervals an extra \( 10^6 \) is added to the burn in period every time as otherwise the results will always be the same due to the deterministic nature of the Halton series. The
step size is such that no overlapping input data is present. The results of the variance reduction methods are found in figure 22. The input values are $T = 10$ and $dt = 0.1$. The input bond price for $T = 10$ is equal to 0.973. Check that quasi-Monte Carlo is only helpful if the number of sample paths is large enough, as the normal Monte Carlo has lower variances for $N = 10$ and $N = 100$. As described by the central limit theorem the convergence of the Monte Carlo method seems to be of order $\frac{1}{\sqrt{N}}$, while the quasi-Monte Carlo seems to be of (nearly) order $\frac{1}{N}$ as described in [13]. Antithetic variables do not seem to change the order of convergence, as it decreases the variance with a factor around 16 for every number of sample paths. This is also in line with [13].

7.5 CIR Monte Carlo with variance reduction methods

The results of section 7.2 are used as input of a Monte Carlo simulation as described in section 5.3. The starting point of each interest rate path is set equal to the starting point of the forward curve at $r_0 = -0.0078$. The Euler scheme of the CIR stochastic differential equation (10) is given by:

$$r_t = r_{t-1} + a(b_t - r_{t-1}) dt + \sigma \sqrt{r_{t-1}} + \nu dW_t$$

The spline interpolation used to determine the function $b_t$ should at least contain the time points used in the Monte Carlo simulation. As an example, 10,000 sample path are generated with the Euler scheme with $a = 2.02$, $\sigma = 0.086$, $T = 10$ and $dt = 0.01$ and plotted in figure 20. Due to the mean reversion, according to this model the interest rate stays between -0.02 and 0.06. The variance reduction methods are again antithetic variables and Halton series described in sections 5.3.4 and 5.3.5.

7.6 Hull-White tree

Implementation is done by the method described in section 5.5.4. Bond prices are calculated using the Hull-White tree and compared to the input bond prices. Results are shown in figure 21. The input parameter $dt$ determines the size of the tree and therefore influences the calculation time. An unreasonably small $dt$ leads to impractical computational time as it increases $j^*$, the size of...
Figure 20: 10,000 sample paths generated with a CIR-model with \( a = 2.02, \sigma = 0.086, T = 10 \) and \( dt = 0.01 \).

Figure 21: Result of bond prices using a Hull-White Monte Carlo simulation without variance reduction methods above and with variance reduction methods below with \( N = 100 \).
the tree. However, if \( dt \) is too big, a discretization error is visible in the results, as is shown in figure 24.

<table>
<thead>
<tr>
<th>( N )</th>
<th>Monte Carlo</th>
<th>Antithetic Variables</th>
<th>quasi-Monte Carlo</th>
<th>AV+QMC</th>
</tr>
</thead>
<tbody>
<tr>
<td>( 10^1 )</td>
<td>( 2.1 \cdot 10^{-2} )</td>
<td>( 1.3 \cdot 10^{-3} )</td>
<td>( 4.5 \cdot 10^{-2} )</td>
<td>( 1.7 \cdot 10^{-3} )</td>
</tr>
<tr>
<td>( 10^2 )</td>
<td>( 6.7 \cdot 10^{-3} )</td>
<td>( 4.7 \cdot 10^{-4} )</td>
<td>( 9.6 \cdot 10^{-3} )</td>
<td>( 5.4 \cdot 10^{-4} )</td>
</tr>
<tr>
<td>( 10^3 )</td>
<td>( 2.2 \cdot 10^{-3} )</td>
<td>( 1.2 \cdot 10^{-4} )</td>
<td>( 9.7 \cdot 10^{-4} )</td>
<td>( 1.1 \cdot 10^{-4} )</td>
</tr>
<tr>
<td>( 10^4 )</td>
<td>( 6.9 \cdot 10^{-4} )</td>
<td>( 4.2 \cdot 10^{-5} )</td>
<td>( 9.6 \cdot 10^{-5} )</td>
<td>( 1.4 \cdot 10^{-5} )</td>
</tr>
</tbody>
</table>

Figure 22: Table with estimated variances of different methods with different number of paths.

Figure 23: On the top a plot is presented with the bond prices predicted by the Hull-White and the input bond prices. Below is a plot with the absolute difference between these bond prices.
8 Results of numerical experiments

This section provides numerical results for all four mortgage options. Often 3D-plots are given in order to check if the model behaves like expected. Data from the bank is used as much as possible.

8.1 LtV

Combining the house price process with the model of the outstanding notional makes it possible to perform numerical experiments. The first experiment regards a digital put option to show the characteristics of this basic option. The second experiment regards the currently used loan-to-value option. The first choice is whether to use the normal or the hyperbolic distribution. The first one is simpler but the second one has a better fit, but this can also be caused by the additional parameters. In order to make the choice, both possibilities are compared.

8.1.1 Log-normal versus hyperbolic distribution

The fatter tails of the hyperbolic distribution make it more likely to pass the strike price. Therefore the slope around the strike price is less steep as can be seen in figure 25. This difference enlarges when calculating the hedging number. Some 3D graphs are given in figures 26 and 27 for both house price models.

8.1.2 Practical example

The loan-to-value option relevant to ABN AMRO is not an ordinary digital option. The measure of this relevant example is given by formula (23). This measure is used as input to the hedging formulas (17) and (18). The four resulting 3D-plots can be found in figures 28 and 29.

8.2 Meeneemoptie

There are two ways to look at the meeneemoptie. The first one is to see it like an American option as it is possible to exercise the option at any time. This way the meeneemoptie can be analyzed with a Hull-White trinomial tree. However, this strategy implies that a client has to be able to move at any time and this assumption is not very sound in practice. This first way to look at the meeneemoptie should therefore be considered as an upper bound. The second way to look at the meeneemoptie is that the time of moving has no causal connection to the pay-off of the meeneemoptie. This assumption makes the option an European option with the time of moving as maturity time.
Figure 25: Plots of the value and hedging number of the option with the log-normal distribution and the hyperbolic distribution.
Figure 26: Value using Log-normal distribution above and the hyperbolic distribution below.

Figure 27: Hedging number using Log-normal distribution above and the hyperbolic distribution below.
Figure 28: 3D-plots of the value and hedging number of the loan-to-value option with respect to the initial loan-to-value ratio and the maturity using the hyperbolic distribution.
Figure 29: 3D-plots of the value and hedging number of the loan-to-value option with respect to the initial loan-to-value ratio and the maturity using the normal distribution.
8.2.1 Meeneemoptie as an American option

The pay-off of the meeneemoptie is required to implement the trinomial tree. At an interest reset date, the pay-off is equal to zero as there is no time left to profit from the lower interest rate. The interest reset date is an input parameter of the model and is denoted by $T$. In practice, the interest reset dates are every 10 years. At each node $(i,j)$, the pay-off $X_{ij}$ is given by:

$$X_{ij} = \max \left\{ (r_0 - r_{ij}) \frac{T - T_i}{T}, 0 \right\}$$

The pay-off is thus equal to the decrease in interest rate weighted by the fraction of time left until the next interest reset date. At each node $(i,j)$ this value is compared to the discounted expected pay-off at $T_{i+1}$. The situation when $X_{ij}$ is bigger corresponds to exercising the option. A 3D-plot of the results of a numerical experiment is shown in figure 30. As expected it shows that the value is strictly increasing in both the volatility and the maturity time. Surprisingly a change in the mean-reversion rate does not influence the value very much, as is shown in figure 31.

![3D-plot](image)

Figure 30: 3D-plot of the value of the American meeneemoptie with respect to the volatility and maturity time.

8.2.2 Conditioned on the time of moving

The trinomial tree model determined the optimal time of moving. The lack of causality between the pay-off and the time of moving makes it possible to see the time of moving as an input parameter. The interest reset dates are assumed to be every ten years, as is most common in practice. The pay-off $X_t$ of the meeneemoptie at moving time $t$ is given by:

$$X = \max \left\{ (r_0 - r_t) \frac{10 - t}{10}, 0 \right\}$$

The pay-off is thus equal to the decrease in interest rate times the fraction of time left until the next interest reset date. Due to the European nature of the option, a quasi-Monte Carlo method can be used to calculate confidence intervals for the value conditioned on the time of moving. The results of a numerical experiment are shown in 3D-figure 33. In order to say something about the unconditional value, the distribution of the time of moving is required. This distribution is found in figure 34 and table 2. An expected value for each notional class can be calculated by taking the
The values using the two interest models for different notional classes differ from 8% to 16%. The value of the meeneemoptie using the CIR-model is greater when the time of moving is small and agrees with the value using the Hull-White model if the time of moving is close to the interest reset date. This can be explained by looking at figures 19 and 20. Clearly the behaviour for small $t$ is different, but the distribution for large $t$ is almost equal.

Figure 31: 3D-plot of the value of the American meeneemoptie with respect to the mean reversion rate and maturity time.

Figure 32: Value of the meeneemoptie with respect to the time of moving for the Hull-White model in green and the CIR-model in blue.
Figure 33: 3D-plot of the value of the conditioned meeneemoptie with respect to the time of moving and volatility.

Table 2: Probabilities of moving after a T number of years for different notional classes according to the Rabobank.
8.3 Rentemiddeling

The regulations of rentemiddeling guarantee that the clients total payments do not decrease when using rentemiddeling as long as the client is in the contract until the new next interest reset date. However, there can be different reasons why cashflows do not follow the contractual amortization scheme. The most common one is when the client moves and fully pays back the mortgage with the money from selling his old house. When timed right, a client can profit from the lower initial payments. This risk, introduced through rentemiddeling, can be priced with the Hull-White trinomial tree model. To make plots, the assumption is that the client moves at a certain time point. The time is the time between moving and the next original interest reset date. The interest reset dates are assumed to be every 10 years. The pay-off is calculated with formula 1 weighted with the time left to profit. The results of a numerical experiment of this scenario are given in figure 35. As expected, the value is increasing in time and decreasing in the rentemiddelingsopslag. Note that this situation is quite rare. If the client does not move, rentemiddeling does by construction not come with any risk for the bank as the initial profit is paid back at later stages. If the client moves before the next interest reset date, rentemiddeling is worth less than in figure 35 which provides the worst case scenario for the bank.

8.4 Pipeline

As already mentioned in section 3.4 there are two variants of the pipeline option, with or without the lock-or-lower regulation. An immediate consequence of the regulation is that the lock-or-lower regulation increases the value of the pipeline option. This is what makes it interesting for the client. From an analyst’s point of view, the regulation also totally changes the nature of the contract. With the lock-or-lower regulation, the pipeline option is a path-depended European option and the value can therefore be calculated by a quasi-Monte Carlo simulation. Without the lock-or-lower regulation, the pipeline option is an American option and the value can be calculated with a Hull-White trinomial tree. It turns out that both behave similarly, but the variant with
lock-or-lower is worth almost twice as much, as can be seen from figures 36 and 37.

8.4.1 Without lock-or-lower

This American variant of the option is by construction the least valuable. With the Hull-White trinomial tree model, 3D-figure 36 is created. As expected the value is increasing in both the maturity and the volatility.

8.4.2 With lock-or-lower

By definition, this European variant of the pipeline option is the most valuable in every situation. Using a quasi-Monte Carlo method, 3D-figure 37 is created. Due to the variance reduction methods, the figure is almost entirely smooth. As expected the value is increasing in both the maturity...
and the volatility.

Figure 37: 3D Plot of the value of the pipeline option with lock-or-lower with respect to the maturity in months and the volatility.

8.4.3 Comparison with current used value

For most mortgage options it is very difficult to test the outcome of the model with the real world. Luckily, there was data available about the structure of the mortgage rate. The mortgage rate consists of three parts. The current interest rate, the spread and the funds transfer price (FTP). The FTP covers the costs and risks of the mortgage. One of the components of the FTP is the pipeline premium. Other components cover for example liquidity risk and prepayment risk. It is very important to know that in the process of calculating the pipeline premium also a behavioural aspect is taken into account. This explains why the data shows that the pipeline premium is different for different amortization schemes and also different for different maturities. This is not the case in the model, as this only takes data during the offer period into account and assumes that the client behaves optimally. Unfortunately only data where the offer period is three months was available, so the dependency on the offer period could not be compared.

With lock-or-lower. The pipeline premium of different contracts with the lock-or-lower variant is shown in figure 38. The outcome of the Monte Carlo simulation with the same input is 0.19%. This value is in line with the values in figure 38, from which the conclusion can be made that clients behave close to optimal. This is no surprise due to the European nature of this variant.

Without lock-or-lower. The pipeline premium of different contracts without the lock-or-lower variant is shown in figure 39. The outcome of the trinomial tree model with the same input is 0.096%. This value is more than double the pipeline premiums in figure 39. This difference is no surprise as the American aspect of the option makes it very difficult to use this variant optimally. The difference is therefore explained by the behavioural aspect of the calculation as the behaviour of the client is in general not optimal.
Figure 38: Pipeline premium with lock-or lower.

Figure 39: Pipeline premium with lock-or-lower.
9 Conclusion

With the use of well-known theories and principles from mathematical finance pricing models for all mortgage options are developed. All models behave as expected, but there was not enough data available to check if the values make sense. With the pipeline model this was possible and the results matched when required. This section describes further points of action, the things one should be careful of when using the models and finally advice for the bank based on the developed pricing models.

9.1 Further points of action

At the time of writing, the bank has to ways to price a mortgage option. The first one is using a logistic regression model. This is for example done with the lock-or-lower pipeline option. It would be very interesting to compare outcomes of a logistic regression model for the other mortgage options with the results of the pricing models. This way, one could say something about the behaviour of clients and the difference it would make if clients suddenly start to behave optimally. Unfortunately, the bank did not have the required data to build such a logistic regression model. Section 5.6 describes how the model would have been build if the data was available.

9.2 Model risk

One should keep in mind that a model is just a model and every model is susceptible to model risk. Model risk is the risk introduced by the fact that the model does not perfectly respresent the current situation. Below three examples of model risk and their impact are described.

9.2.1 Choice of interest rate model

A motivation for the choice of interest rate models is already given in section 7. The choice beteen the Hull-White and CIR-model is arbitrary. Figures 19 and 20 show the behaviour of each model and both seem in the long run to produce sensible paths. However, as can be seen in equations (24) and (25) and figures 32 and 40, the value of the option can change with more than 10%. This is mainly caused by the difference in short term behaviour between the two models. The interest rate paths generated by the CIR-model seem to spread out faster than the paths generated by the Hull-White model. This results in a high value as can be seen from figure 32. Before implementing anything, the bank should decide what short term behaviour respresents the current market the best.

9.2.2 Sensitivity to the input parameters

The sensitivity with respect to the volatility was already shown in figures 30, 33, 36 and 37. All show that even a small change in volatility can have a substantial impact on the value. One should therefore be very careful when estimating the volatility. Examples of the sensitivity with respect to the mean reversion parameter are shown in figures 31 and 41. Clearly the value is very stable against a change in the mean reversion parameter.

9.2.3 Sensitivity to the calibration data

Unless stated otherwise, all experiments are done using a interest rate model calibrated to the yield curve of April 7, 2017. It is important to know how the value reacts when a different yield curve is used. The European mceneemoption is used as an example. Figure 42 shows the difference in value for two yield curves with a time difference of four months. Combining the values of figure 42 with the probabilities of table 2, the following values are found.

\[
V_{0}^{April} = 0.00214 \cdot 1_{\{0 < N_0 < 150k\}} + 0.00174 \cdot 1_{\{150k < N_0 < 250k\}} + 0.00155 \cdot 1_{\{250k < N_0 < \infty\}}
\]

\[
V_{0}^{August} = 0.00235 \cdot 1_{\{0 < N_0 < 150k\}} + 0.00189 \cdot 1_{\{150k < N_0 < 250k\}} + 0.00168 \cdot 1_{\{250k < N_0 < \infty\}}
\]
Figure 40: Relative difference of the value of the European meeneemoptie when using the Hull-White model and when using the CIR-model.

Figure 41: 3D-plot of the value of the European meeneemoptie with respect to the mean reversion rate and time of moving.
Figure 42: Graph of the value of the European meeneem option with respect to the time of moving. The values using the yield curve of 7 April 2017 are in blue, the values using the yield curve of 7 August 2017 are in green.
Observe that the different yield curve made an impact of 8%. One is therefore advised to calibrate the model to the yield curve of that day.

9.3 Advice for the bank

It is important to note that all pricing models assume that the counterparty behaves optimally. For most mortgage options, this is not a fair assumption. For example, a client will not move purely to use the meeneemooption optimally. Therefore the bank mainly uses logistic regression models to calculate the value and risk of mortgage options. A key assumption within these models is that past behaviour is representative for future behaviour. If clients do not change their behaviour, the outcome of the models should simply be interpreted as a worst case scenario and not as a fair price. However, there might be reasons that the behaviour of the client changes over time. One reason could be that clients are better informed or helped by a third party. Another reason can be a regulation change that forces the bank to make the best choices for the client. It is important to note that all pricing models assume that the counterparty behaves optimally. The bank should be aware of the situation when this will happen, like help by a third party or a new regulation that forces the bank to make the best choices for the client. When this happens, the results from the pricing models should be used as a base fair price as clients in this situation behave (close to) optimal. The bank can then decide to add whatever spread they like to this fair price. In any case, the bank should always keep the outcomes of the model in mind as a worst case scenario.

References


[20] Mark Rubinstein. On the relation between binomial and trinomial option pricing models, 
2000.

I: Singel-Factor Models.


[23] Catherine Daveloose, Asma Khedher, and Michèle Vanmaele. Robustness of quadratic hedging 
strategies in finance via Fourier transforms. Computational and Applied Mathematics, 296:56– 
88, 2016.


[26] Andreas E. Kyprianou. Introductory Lectures on Fluctuations of Lévy Processes with Appli-

[27] Jens Wannenwetsch. Lévy processes in finance: The change of measure and non-linear de-
pendence.

[28] Ernst Eberlein, Antonis Papapantoleon, and Albert N. Shiryaev. On the duality princeipe in 

[29] Tahir Choulli, Nele Vandaele, and Michèle Vanmaele. The Föllmer–Schweizer decomposition: 

with stationary independent increments. The Annals of Applied Probability, 16(2):853–885, 
2006.

[31] Jan Dhaene, Ben Stassen, Pierre Devolder, and Michel Vellekoop. The Minimal Entropy 
Martingale Measure in a Market of Traded Financial and Actuarial Risks. Tinbergen Institute 


[33] Ernst Eberlein, Kathrin Glau, and Antonis Papapantoleon. Analysis of fourier transform 
valuation formulas and applications.

[34] M. Abramowitz and I. A. Stegun. Handbook of mathematical functions. National Bureau of 


1922.