Maastricht University

Master's Thesis

Portfolio Valuation of an Insurance Company

Author:
A.L.M.A. de Rijck, i6001514

Supervisor:
A.A.J. Pelsser

August 20, 2014

Contents

1 Introduction

Everyone is subject to risk. These can be all sorts of risk, such as an unexpected passing away. One of the most common methods to mitigate risks is to become insured against them. In case someone deceases, one might want to insure the bereaved against the costs incurred from the funeral. This security can be found in the form of life insurance contracts. Under a life insurance contract the benefit insured consists of a single payment, the sum insured, which is paid at the end of the contract (Gerber, 1997, p.23). In case of a life insurance, the end of the contract implies the moment of death.

A life insurance company has policyholders who pay a premium for a certain amount of time. The company and the policyholder have to agree upon the sum insured that the policyholder will receive at the end of the contract. The amount of premium a policyholder pays depends on the respective age and the amount he would like to have insured. This premium will be invested such that at the end of the contract the amount insured can be paid out to the policyholder's account.

One type of product a life insurance company can have is a profit-sharing contract. The profit-sharing contracts are based on the return on investments the company has made each year. If the rate of return exceeds a certain threshold, the company will give its policyholders a certain percentage increase on the total sum insured. An example would be if a policy holder has a profit-sharing contract which states that he will receive 6000 euros at the moment of death, and each year the company has made a profit of at least 3%, its insured amount will increase with 0.5%. Note that in this example, the threshold is equal to 3%.

This paper examines the contract of a policyholder who has a profit-sharing contract starting from January 2014, i.e. $t = 0$, with an insured amount of 1000 euros. The contract states that for each year the company makes a profit of 3% or more, its insured amount will increase. The increase in the insured amount depends on the value of $0.75 \max(R - 0.03)$, where R is the rate of return the company made that year. It is already known that the policyholder will pass away in exactly 50 years. Keeping this information in mind, it is possible to calculate the value of the insured amount at the end of the contract.

To calculate the problem at hand, this paper first examines the available data. Then it establishes a solid theoretical background. Subsequently, the methods which enable us to tackle this problem are elaborated upon. Afterwards, a practical application is introduced. Finally, the approximation methods are discussed and the outcomes are evaluated.

2 Data

To be able to compute the value of the contract, it is necessary to have a look at the data provided by the company. For that reason this section shortly outlines the data.

As mentioned, the increase in the insured amount depends on the company's rate of return. It is also stated that the insurance company invests the amount of premium they receive from policyholders. Therefore we focus on the investment portfolio. It is important to note that the company has solely invested in bonds. Hence, its rate of return will be based on these investments.

In total the investment portfolio consists of 11 coupon-bearing bonds and 23 zero-coupon bonds. These bonds all differ in time to maturity and volume, i.e. a larger amount of money is invested in some bonds than in others. Additionally, the coupon-bearing bonds do not have similar coupon rates or the same coupon date. To solve these issues, all cashflows will be evaluated together, keeping in mind the differences between the bonds.

As each bond has a different time to maturity, this results in a substantial amount of different data points. It has been decided that the bonds and their releases will be evaluated at the last day of each month. Hence, one year consists of twelve data points. However, the company will review once every year whether it has to increase the sum insured of its policyholders. For this purpose the company will take the average of the rate of return over 12 months.

The company makes a distinction between money that will be reinvested and money that will not be reinvested. Currently, the reinvestments are made against the forward rate, based on today's term structure of interest rates, i.e. $t = 0$. The term structure of interest rates is the relationship between interest rates and their maturities (Hull, 2014). For each time to maturity a corresponding interest rate is depicted. These interest rates exist as long as they can be replicated in the market. However, for longer time to maturities they cannot be replicated, because there are simply no bonds with these maturity dates. To find the missing values an extrapolation method has been used. The forward rates are calculated from this term structure model. This translates into yearly forward rates for each month.

3 Theoretical Framework

In this section a theoretical framework will be provided such that the reasoning behind and the derivations in the applied models and methods can be understood. These are the fundamental tools for pricing purposes.

3.1 Martingales and Numéraires

Two fundamental concepts in financial mathematics are the notions of martingales and numéraires. As these concepts play an important role in the methods used in this paper, their definitions will be given in this section. A martingale is defined in the following manner (Mikosch, 2008, p.80):

Definition 1 (Martingale) *Let $X = (X_t)_{t \geq 0}$ be a stochastic process, and $(\mathcal{F}_t)_{t \geq 0}$ a filtration, or information set, at time t. Then X is a martingale with respect to $(\mathcal{F}_t)_{t \geq 0}$ if*

- $\mathbb{E}|X_t| < \infty$ *for all* $t \geq 0$;
- X *is adapted to* $(\mathcal{F}_t)_{t \in \mathcal{T}}$;
- $\mathbb{E}(X_t|\mathcal{F}_s) = X_s$ *for all* $0 \leq s \leq t$.

Having a closer look at the third point of the definition, an interesting interpretation can be made. Given the information set at time s, the best possible prediction of X_t is X_s. Hence, if the information up until the present is known, the present value is the best predictor of the future value. In the literature martingales are often thought of as representations of fair games.

Another important concept is the numéraire. It can be defined as (Bingham and Kiesler, 2004, p.101):

Definition 2 (Numéraire) *A numéraire is a price process $B(t)$, which is strictly positive for all $t \in \{0, 1, ..., T\}$.*

A numéraire can be used to express asset price processes as relative prices, i.e. $Z'(t) = Z(t)/B(t)$, where $Z(t)$ is an asset price process and $B(t)$ a numéraire. Hence, a numéraire can be interpreted as a reference asset to which other asset prices are normalized to (Brigo and Mercurio, 2001, p.27).

The concepts of numéraires and martingales turn out to be useful in the following sections.

3.2 No Arbirtage

An important assumption often made in the financial world is the no arbitrage opportunity. Absence of arbritrage implies that it is not possible to invest nothing at $t = 0$, and receive a positive amount at $t = T$, i.e. making a profit without actually bearing any risk.

To define an arbitrage free market[1], some terminology has to be introduced first. A continuous-time economy is considered with a time horizon of $T > 0$. The probability space which measures the uncertainty is given by $(\Omega, \mathcal{F}, \mathbb{P})$, and the filtration defined by $\mathbb{F} = \{\mathcal{F}_t : 0 \leq t \leq T\}$ is being

[1] Note that the definitions which are used in this section are obtained from Brigo and Mercurio (2001) unless stated otherwise.

used. The filtration can be interpreted as the information set at time t. This economy consists of continuously traded (non-dividend paying) assets, whose prices follow a process of the form:

$$dS_t^0 = r_t S_t^0 dt,$$

where r_t is the instantaneous short rate. With this information, a trading strategy can be defined in the following manner:

Definition 3 (Trading strategy) *A trading strategy is a process $\phi = \{\phi_t : 0 \leq t \leq T\}$, whose components $\phi^0, \phi^1, ..., \phi^K$ are locally bounded and predictable. The value of the trading process is expressed as:*

$$V_t(\phi) = \phi_t S_t = \sum_{k=0}^{K} \phi_t^k S_t^k, \ 0 \leq t \leq T,$$

or in the differential form:

$$dV_t(\phi) = \sum_{k=0}^{K} \phi_t^k S_t^k$$

whereas the gains of the trading process are given by:

$$G_t(\phi) = \int_0^t \phi_u dS_u = \sum_{k=0}^{K} \int_0^t \phi_u^k dS_u^k, \ 0 \leq t \leq T$$

The predictability of the process implies that the value at time t is known immediately before time t (Brigo and Mercurio, 2001, p.25).

As it is known how the trading process and its gain can be expressed, it is possible to have a look at self-financing trading strategies, i.e. trading strategies without capital inflows or outflows after $t = 0$ (initial time).

Definition 4 (Self-financing) *A trading strategy ϕ is self-financing if $V(\phi) \geq 0$ and*

$$V_t(\phi) = V_0(\phi) + \sum_{k=0}^{K} \int_0^t \phi_u^k dS_u^k = V_0(\phi) + G_T(\phi), \ 0 \leq t \leq T.$$

This definition implies that changes in the underlying asset prices are the only ones who have an influence on the change in value.

If $V_0(\phi) = 0$, $\Pr(V_t(\phi) \geq 0) = 1$, and $\Pr(V_t(\phi) > 0) > 0$ then the self-financing strategy is an arbitrage opportunity. When examining arbitrage opportunities, the concept of equivalent martingal measures turns out to be of value. The definition given below is obtained from Bingham and Kiesler (2004, p.233):

Definition 5 (Equivalent martingale measure) *An equivalent martingale measure $\mathbb{Q}$ is a probability measure defined on the space $(\Omega, \mathcal{F})$ such that*

- *$\mathbb{P}$ and $\mathbb{Q}$ are equivalent*[2]*,*
- *the discounted price process $D(0, \cdot)S$ is a martingale under $\mathbb{Q}$, that is $\mathbb{E}^{\mathbb{Q}}(D(0,t)S_t^k|\mathcal{F}_u) = D(0,u)S_u^k$, for all $k = 0, 1, ..., K$ and all $0 \leq u \leq t \leq T$.*

Note that the discounted price process can be seen as a relative price process with the discount factor as the numéraire.

Brigo and Mercurio (2001) mention that the absence of arbitrage opportunities is implied by the existence of an equivalent martingale measure.

A contingent claim is a positive random variable on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ (Brigo and Mercurio, 2001, p.25).

[2]The fact that $\mathbb{P}$ and $\mathbb{Q}$ are equivalent measures implies that they have the same null sets, i.e. $\mathbb{P}(A) = 0$ if and only if $\mathbb{Q}(A) = 0$ for every $A \in \mathcal{F}$.

Definition 6 (Attainable contingent claim) *A contingent claim H is attainable if there exists at least one self-financing trading strategy ϕ such that*

$$V_T(\phi) = H.$$

In this case ϕ is called a replicating strategy for H, and $\pi_t = V_t(\phi)$ is the price at time t of the trading strategy associated with H.

The differential form of the value of the trading strategy implies that the following holds:

$$d\left(\frac{V_t(\phi)}{Z_t}\right) = \sum_{k=0}^{K} \phi_t^k d\left(\frac{S_t^k}{Z_t}\right).$$

Hence, the definition for an attainable claim holds under any numéraire.

Definition 7 (Complete market) *A financial market is said to be complete if and only if all contingent claims are attainable.*

The market is said to be free of arbitrage and complete if and only if there exists a unique martingale measure (Brigo and Mercurio, 2001, p.26). Next to that, it is interesting to note that in a market without arbitrage opportunities, any attainable claim has an uniquely given price. This market does not necessarily have to be complete.

3.3 Change of Measure

In the previous section the existence of an equivalent martingale measure has been demonstrated. Nevertheless, the existence of this measure does not have to mean that it is the most obvious or convenient measure for pricing the contingent claim. In the cases that it is not the most convenient measure, a change of measure can be applied.

Recall from section 3.1 the description of a numéraire. It was mentioned that the price of an asset expressed in terms of the numeraire, i.e. $Z(t)/B(t)$, is called the relative price of the asset. The relative price of an asset is a martingale under the equivalent martingale measure with respect to that numéraire, i.e. the stochastic differential equation does not have a drift term (Brigo and Mercurio, 2001, p.29). If you change the underlying numéraire it is no longer a martingale. Hence, it has a drift term. When applying a change of measure, the drift term of a stochastic differential equation can be eliminated. This will be explained later on in this section and turns out to be useful when looking at the convexity correction concept.

At first a framework for the change of measure is introduced. An important concept is the Radon-Nikodym derivative, and is defined as (Mikosch, 2008, p.193):

Definition 8 (Radon-Nikodym) *Assume $\mathbb{P}$ and $\mathbb{Q}$ are two probability measures. If and only if there exists a non-negative $\mathcal{F}$-measurable function f such that*

$$\mathbb{P}(A) = \int_A f d\mathbb{Q}(A), \ A \in \mathcal{F}$$

it is true that $\mathbb{P} \ll \mathbb{Q}$ holds.[3] *The function f is the density of $\mathbb{P}$ with respect to $\mathbb{Q}$, and is called the Radon-Nikodym derivative. The Radon-Nikodym derivative is denoted by $\frac{d\mathbb{P}}{d\mathbb{Q}}$.*

The Radon-Nikodym derivative connects the probabilities of two equivalent probability measures (Kwok, 2008, p.87). With the Radon-Nikodym derivative in mind, it is possible to state the theorem of the change of numéraire (Pelsser, 2000, p.11):

[3] This means that $\mathbb{P}$ is absolutely continuous with respect to $\mathbb{Q}$. Note that if $\mathbb{P} \ll \mathbb{Q}$ and $\mathbb{Q} \ll \mathbb{P}$, then $\mathbb{P}$ and $\mathbb{Q}$ are equivalent probability measures.

Definition 9 (Change of Numéraire) *Let $N(t)$ and $M(t)$ be two numéraires. Assume $\mathbb{Q}^N$ is the equivalent martingale measure w.r.t. $N(t)$, and $\mathbb{Q}^M$ the equivalent martingale measure w.r.t. $M(t)$. The Radon-Nikodym derivative that turns $\mathbb{Q}^M$ into $\mathbb{Q}^N$ is then given by*

$$\frac{d\mathbb{Q}^N}{d\mathbb{Q}^M} = \frac{N(T)M(t)}{M(T)N(t)}.$$

A theorem which also uses the the Radon-Nikodym derative is Girsanov's theorem. This theorem describes the change of measure for diffusion processes, and is defined as (Pelsser, 2000, p.11):

Definition 10 (Girsanov's Theorem) *For any stochastic process $\lambda(t)$ such that*

$$\Pr\left[\int_0^t \lambda(s)^2 ds < \infty\right] = 1,$$

consider the Radon-Nikodym derivative $\rho(t) = \frac{d\mathbb{Q}^*}{d\mathbb{Q}}$. The Radon-Nikodym derivative is given by

$$\rho(t) = \exp\left(\int_0^t \lambda(s)dW(s) - \frac{1}{2}\int_0^t \lambda(s)^2 ds\right),$$

where W is a Brownian motion under $\mathbb{Q}$. Then the following process under $\mathbb{Q}^*$ is also a Brownian motion:

$$W^*(t) = W(t) - \int_0^t \lambda(s)ds.$$

The process obtained with this theorem can be rewritten into $dW = dW^* + \lambda(t)dt$. Thus by applying Girsanov's theorem the drift term can be eliminated.

3.4 Ornstein-Uhlenbeck Process

The Ornstein-Uhlenbeck process can be used to model interest rate processes. Because this paper uses the Ornstein-Uhlenbeck process to model the interest rate process, this section shortly outlines the properties of an Ornstein-Uhlenbeck process. Its stochastic differential equation is given by:

$$dX(t) = a(\mu - X(t))dt + \sigma dW(t)$$

A gaussian stochastic process is said to be an Ornstein-Uhlenbeck process if it is stationary, Markovian and continuous in probability (Finch, 2004). It is normally distributed with mean and variance (Faris, 2001):

$$\mathbb{E}[X(t)] = X(0)e^{-at} + \mu(1 - e^{-at}),$$
$$\text{Var}[X(T)] = \frac{\sigma^2}{2a}(1 - e^{-2at})$$

with covariance function:

$$\text{Cov}(X(s), X(t)) = \frac{\sigma^2}{2a} e^{-a(s+t)} \left(e^{2a \min(s,t)} - 1\right)$$

4 Model and methods

This section outlines the underlying model and the different methods which are used to value the contract of the insurance company. At first, the model used by the company to calculate the rate of return is described. As this paper would like to replace the forward rate based on the term structure at $t = 0$ by one-year zero-rates derived from a dynamic term structure model, a dynamic term structure model is introduced afterwards. Having a closer look at the problem at hand, it turns out that a convexity correction has to be applied. For that reason, a section concerning convexity correction follows. Finally, the option pricing formula applicable to this problem is elaborated on.

4.1 Rate of Return

In section 2 it was already mentioned that the company has monthly data points. To calculate the rate on which the profit sharing is based, the company looks at the following formula:

$$Rate_i = \left(1 + a_i + b_i[(1+c_i)^{\frac{1}{12}} - 1]\right)^{12} - 1$$

where a_i is the amount that will not be reinvested, b_i the amount that will be reinvested, and c the corresponding risk-free forward rate of that month. It is important to note that Rate_i are monthly values.

When applying a Taylor expansion, an approximation of this equation is given by:

$$\text{Rate}_i = \left(1 + a_i + b_i[(1+c_i)^{\frac{1}{12}} - 1]\right)^{12} - 1 \approx (1+a_i)^{12} + (1+a_i)^{11} b_i c_i - 1$$

It is easier to rewrite this equation into:

$$\text{Rate}_i \approx F_i + G_i c_i,$$

where $F_i = (1+a_i)^{12} - 1$ and $G_i = (1+a_i)^{11} b_i$.
The mean and variance of the variable $Rate_i$ are defined as:

$$\mathbb{E}[\text{Rate}_i] = F_i + G_i \mu_c,$$

$$\text{Var}[\text{Rate}_i] = G_i \sigma_c^2.$$

This is the rate per month in which yearly forward rates of each month are used. The company is actually interested in the rates per year, as the insurance company has based its profit-sharing contracts on yearly rates. To find the yearly rates, the average of the rates per month of each year is considered. Hence, for each year the values of the rates from January untill December are added up and divided by 12:

$$R = \sum_{j=1}^{12} \frac{\text{Rate}_j}{12} = \frac{1}{12} \sum_{j=1}^{12} (F_j + G_j c_j) = \sum_{j=1}^{12} (\frac{1}{12} F_j + \frac{1}{12} G_j c_j).$$

The expected value and variance of the yearly rate are given by:

$$\mathbb{E}(R) = \frac{1}{12} (\sum_{i=1}^{12} F_i + \sum_{i=1}^{12} G_i \mu_c),$$

$$\text{Var}(R) = \sum_{k=1}^{12} \sum_{l=1}^{12} G_k G_l Cov(c_k c_l) (\frac{1}{12})^{12}.$$

4.2 Hull-White model

The company has reinvested against the forward rate based on the term structure at $t = 0$. A disadvantage of this method is that the confidence interval becomes very wide over time. To overcome this disadvantage, another method can be used in which the reinvestments are made against the zero-rate. To be able to apply this method, a dynamic term structure model has to be introduced. With this dynamic term structure model it is possible to depict a term structure for each time in the future. Then for each term structure of interest rates, the one-year zero-rate will be used to calculate the rate of return. This results in a confidence interval which is narrower than in the case of using forward rates to calculate the rate of return.

At first it is important to decide what kind of term structure model is convenient to use. The simplest option is to define a one-factor model. In this case the interest rate movements are driven by one factor. As for larger times to maturity the movements are severely correlated, a one-factor model is a reasonable assumption to model the term structure of interest rates (Pelsser, 2013).

One of these one-factor models is the Hull-White model. In the Hull-White model the interest rates have a mean-reverting Gaussian distribution. The benefit of mean-reversion is that it takes into account economic theory concerning the level of interest rates. A disadvantage of this model is that the interest rates are normally distributed. Hence, it is theoretically possible that the interest rates become negative. An advantage of the Hull-White model is that it can be analytically fitted to the initial term-structure of interest rates, and the possibility to analytically value the prices of, and options on discount bonds (Pelsser 2000, p.45). This analytical tractability makes the Hull-White model very popular in spite of its drawback that interest rates can become negative. The analytically tractability is a main reason to use the Hull-White model in this research.

The dynamics of $r(t)$, as proposed by the Hull-White model used in this paper, are defined as:

$$dr(t) = [\vartheta(t) - ar(t)]dt + \sigma dW(t)$$

where a and σ are positive constants, respectively the mean-reverting parameter and the market volatility. The parameter $\vartheta(t)$ is chosen such that the model is fitted to the currently observed term-structure (Brigo and Mercurio, 2001 p.73).

By means of integration[4] we can derive the mean and variance of $r(t)$ conditioned on the information set at time s. It can be seen that

$$r(t)|\mathcal{F}_s \sim N(\mathbb{E}\{r(t)|\mathcal{F}_s\}, \text{Var}\{r(t)|\mathcal{F}_s\})$$

with

$$\mathbb{E}\{r(t)|\mathcal{F}_s\} = r(s)e^{-a(t-s)} + \alpha(t) - \alpha(s)e^{-a(t-s)}$$

$$\text{Var}\{r(t)|\mathcal{F}_s\} = \frac{\sigma^2}{2a}[1 - e^{-2a(t-s)}]$$

where $\alpha(t) = -\frac{\partial lnP^M(0,t)}{\partial t} + \frac{\sigma^2}{2a^2}(1 - e^{-at})^2$.
Under the Hull-White model the price at time t of the zero-coupon bond maturing at time T can be expressed as (Hull and White, 1994):

$$D(t,T) = e^{A(t,T) - B(t,T)r(t)},$$

with $B(t,T) = \frac{1}{a}[1 - e^{-a(T-t)}]$ and $A(t,T) = \ln \frac{P^M(0,T)}{P^M(0,t)} + B(t,T)f^M(0,t) - \frac{\sigma^2}{4a}(1 - e^{-2at})B(t,T)^2$

4.3 Convexity correction

The Hull-White model described in section 4.2 assumes that the pay-off date is equal to the observation date. However, if the pay-off date is not equal to the observation date, a convexity correction needs to be applied. This correction is necessary because the derivative is not longer priced under its natural martingale measure (Pelsser, 2001). In essence, this correction shows the adjustment the interest rate has to undergo to take into account the difference between the pay-off date and the observation date. In this paper, the pay-off date (T) is later than the observation date (S), i.e. $T \geq S$. For that reason a convexity correction is needed.

Hence, the distribution of $r(t)$ will be derived under the T-forward risk-adjusted measure $\mathbb{Q}^T$. With among others Ito's Lemma and Girsanov's theorem[5], the distribution of $r(t)$ can be obtained.

The mean and variance conditioned on the information set at time s, $\mathcal{F}_s$, are given by:

$$\mathbb{E}^T\{r(t)|\mathcal{F}_s\} = r(s)e^{-a(t-s)} + \alpha(t) - \alpha(s)e^{-a(t-s)} - M^T(s,t)$$

[4]See Appendix 8.1 for the derivations.
[5]The derivations can be found in Appendix 8.2.

$$\text{Var}^T\{r(t)|\mathcal{F}_s\} = \frac{\sigma^2}{2a}\left[1 - e^{-2a(t-s)}\right]$$

with the correction term: $M^T(s,t) = \frac{\sigma^2}{a^2}[(1 - e^{-a(t-s)}) - \frac{1}{2}(e^{-a(T-t)} - e^{-a(T+t-2s)})]$.
Please note that the variance remains unchanged.

4.4 Option Pricing

In most literature the option pricing formulae of Fisher Black and Merton Scholes are used to calculate derivatives. Brigo and Mercurio (2000) have formulated an option pricing formula in case of the Hull-White model. In their model the returns are based on a lognormal distribution. However, in the model of this paper the returns turn out to be normally distributed which makes the Bachelier model applicable. Bachelier has already developed this theorem in 1900 (Sullivan and Weithers, 1991). The option pricing formula of Bachelier is expressed as:

$$C = (R - K)\Phi\left(\frac{R-K}{\sigma\sqrt{T}}\right) + \sigma\sqrt{T}\phi\left(\frac{R-K}{\sigma\sqrt{T}}\right)$$

where σ is the market volatility and T the time to maturity of the option.

5 Practical Application

This section explains how the above model and methods can be implemented and used in practical situations. First, a small change will be made to the equations for pricing the zero-coupon bonds. These alterations are based on Pelsser's approach (2000) of the Hull-White model. Afterwards, the idea of using zero-rates instead of forward rates is implemented. Lastly, the option pricing is briefly discussed.

5.1 Alternative Pricing Equations

Pelsser (2000) has derived the analytical formulae of the Hull-White model in a different manner than Hull and White (1994). His alternative derivation makes it easier to implement the model in practice. The underlying dynamics of $r(t)$ remain unchanged and are given by:

$$dr(t) = [\vartheta(t) - ar(t)]dt + \sigma dW(t).$$

He starts with a transformation of the variables:

$$y(t) = r(t) - \alpha(t),$$

where $\alpha(t) = e^{-at}(r_0 + \int_0^t e^{au}\vartheta^*(u)du)$. The parameter $\alpha(t)$ is chosen such that $y(0) = 0$. Next to that, the unfamiliar parameter $\vartheta^*(t)$ appears in the formula for $\alpha(t)$. This parameter is defined as:

$$\vartheta^*(t) = \vartheta(t) - \lambda(t)\sigma$$

where $\lambda(t)$ is the market price of risk.

According to Pelsser (2000) the transformation of variables has the advantage that y is now not dependent anymore on $\alpha(t)$, and $y(t)$ follows an Ornstein-Uhlenbeck process. Under the T-forward-risk-adjusted measure, $y(t)$ has a normal distribution with:

$$\mathbb{E}\{y(t)|\mathcal{F}_s\} = y(s)e^{-a(t-s)} - \frac{\sigma^2}{2a^2}(1 - e^{t-s})^2,$$

and

$$\text{Var}\{y(t)|\mathcal{F}_s\} = \frac{\sigma^2}{2a}[1 - e^{-2a(t-s)}].$$

The formulae below can be used to value the prices of discount bonds[6].

$$B(t,T) = \frac{1-e^{-a(T-t)}}{a}$$

$$A(t,T) = \frac{\sigma^2}{2a^3}\big(a(T-t) - 2(1-e^{-a(T-t)}) + \frac{1}{2}(1-e^{-2a(T-t)})\big) - \int_t^T \alpha(s)\mathrm{d}s$$

where $\int_t^T \alpha(s)\mathrm{d}s = -\log\frac{D(0,T)}{D(0,t)} + \frac{\sigma^2}{2a^3}(a(T-t) - 2(e^{-at}-e^{-aT}) + \frac{1}{2}(e^{-2at}-e^{-2aT}))$.
The value of the zero-coupon bond can be calculated with:

$$D(t,T) = e^{A(t,T)-B(t,T)y(t)}$$

Also in this case the convexity correction can and needs to be applied. The convexity correction term can be derived in a similar manner as in Appendix 8.2. This means that the interest rate r can be modelled as a Gaussian distribution with mean and variance

$$\mathbb{E}\{y(t)|\mathcal{F}_s\} = y(s)e^{-a(t-s)} - \frac{\sigma^2}{a^2}[(1-e^{-a(t-s)}) - \frac{1}{2}(e^{-a(T-t)} - e^{-a(T+t-2s)})],$$

$$\mathrm{Var}\{y(t)|\mathcal{F}_s\} = \frac{\sigma^2}{2a}(1-e^{-2a(t-s)}).$$

With these formulae on hand, the stochastic y-process can be modelled and the discount bonds can be priced.

5.2 Zero-Rate

In chapter 4 the Hull White model, a dynamic term structure model, has been examined. As mentioned, in the past the company has reinvested against the risk-free forward rate based on the term structure of today, i.e. $t = 0$. Since we have a dynamic term structure now, the company can reinvest against the zero-rates derived from the term structures of interest rates based on the Hull-White model. Therefore, the zero-rate is the variable of interest in this paragraph.

The price of a zero-coupon bond with a face value equal to one can be calculated as $e^{-z_0\tau}$. Next to that, it is possible to calculate this value with the formulae obtained in section 5.1. Hence, to calculate the zero-rate the following equation has to be solved:

$$e^{-z_0\tau} = e^{A(t,t+\tau)-B(t,t+\tau)y(t)}$$

$$z_0(t) = \frac{-A(t,t+\tau)}{\tau} + \frac{B(t,t+\tau)}{\tau}y(t)$$

where τ represents time-steps of 1 month, i.e. $\tau = \frac{1}{12}$.
Then it can be seen that z_0 is normally distributed with mean and variance:

$$\mathbb{E}(z_0) = \frac{-A(t,t+\tau)}{\tau} + \frac{B(t,t+\tau)}{\tau}\mu_y$$

$$\mathrm{Var}(z_0) = \left(\frac{B(t,t+\tau)}{\tau}\right)^2\sigma_y^2$$

[6] See Pelsser (2000) chapter 3 for the alternative derivations.

The distribution of the zero-rates can be used to calculate the rates of return again. The same logic as in paragraph 4.1 is applied to find the distribution of Rate_i. The mean and variance of the normally distributed variable Rate_i with the zero-rate instead of the forward rate are defined as:

$$\mathbb{E}[\text{Rate}_i] = F_i - \frac{A_i(t,t+\tau)G_i}{\tau} + \frac{B_i(t,t+\tau)G_i}{\tau}\mu_y,$$

$$\text{Var}[\text{Rate}_i] = \left(\frac{B_i(t,t+\tau)G_i}{\tau}\right)^2 \sigma_y^2.$$

Note that the subscripts of A and B correspond to the particular month. E.g. $A_i(t,t+\tau)$ corresponds to the same month as G_i.
These are again the monthly rates in which yearly zero-rates are used. The company is interested in the yearly rates, because the profit-sharing contracts are based on these rates. To find the yearly rates, the average of the rates per month of each year is considered.

The mean and variance of the yearly rate are given by:

$$\mathbb{E}(R) = \frac{1}{12}\sum_{i=1}^{12}\big(F_i - \frac{A_i(t,t+\tau)G_i}{\tau} + \frac{B_i(t,t+\tau)G_i}{\tau}\mu_y\big),$$

$$\text{Var}(R) = \left(\frac{1}{12}\right)^{12}\sum_{j=1}^{12}\sum_{k=1}^{12}\frac{B_j(t,t+\tau)G_j}{\tau}\frac{B_k(t,t+\tau)G_k}{\tau}\text{Cov}\big(y(j),y(k)\big).$$

The variance of the R-process depends on the covariance of the y-process, which is an Ornstein-Uhlenbeck process. The covariance of an Ornstein-Uhlenbeck process is given by:

$$\text{Cov}\big(y(j),y(k)\big) = \frac{\sigma^2}{2a}e^{-a(j+k)}(e^{2a\min(j,k)} - 1).$$

5.3 Average Rate Option Pricing

In the previous section it is explained how the rate of return is established. The rate of return is an average rate over 12 months. The company would like to know the value of $\mathbb{E}(R-K)^+$. It is possible to do monthly option pricings, however this is a very time-consuming approach and in the end the company is only interested in the average result (Levy, 1992). Hence, we need to take the theory of average rate options into account, also known as Asian options. Usually, the distribution of the underlying process is difficult to find. However, at this point it is already known that the average rate is normally distrisbuted with the mean and variance expressed as in paragraph 5.2. As the distribution is known, it is possible to apply the Bachelier option pricing formula of section 4.4. Recall that the option pricing formula is given by:

$$C = (R-K)\Phi\left(\frac{R-K}{\sigma\sqrt{T}}\right) + \sigma\sqrt{T}\phi\left(\frac{R-K}{\sigma\sqrt{T}}\right),$$

where R is the average rate per year, K the threshold, σ the market volatility, and T the time to maturity.

6 Results

Up until now the methods show how the portfolio of the life insurance company can be valued. This section shows the results of the practical application, and it explains the outcomes. Next to that, it compares the outcome of the analytical approach with the outcome of a numerical approach.

Recall that the object of this research is a policyholder who has a profit-sharing contract starting from January 2014. This particular policyholder has an insured amount of 1000 euros on January 2014, and will pass away after exactly 50 years. Hence, we have 600 data points. Our goal is to calculate the accrued sum insured at the end of December 2063.

6.1 Analytical Approximation

In the previous sections it was shown how we can calculate the options of this model taking into account the different methods. The value of interest in this paper is:

$$\mathbb{E}\big[\prod_{i=1}^{n}\big(1+0.75\max[R_i(y)-K,0]\big)\big]$$

As it is a path-dependent process, it is known that $\mathbb{E}(XY)\neq\mathbb{E}(X)\mathbb{E}(Y)$. Hence, it is not possible to just multiply the options to get the final result. If doing so, one does not take into account the cross-products which turn out to be of important value. However, it can be a tedious or even impossible job to account for all cross-products. Therefore, this paper uses an approximation to calculate the final result.

6.1.1 Lower Bound

As there are a lot of cross-products that need to be considered, it is hard to find an analytical solution. In 1995 Rogers and Shi already came across a similar problem and found a method to obtain a lower bound for average rate options. This method is based on conditioning the payoff on a zero-mean Gaussian variable Z. It is important to note that in this case the process remains Gaussian. Rogers and Shi (1995) used Jensen's inequality to show that the following equation holds:

$$\mathbb{E}(Y^{+})=\mathbb{E}(\mathbb{E}(Y^{+}|Z))\geq\mathbb{E}(\mathbb{E}(Y|Z)^{+}).$$

This holds because $(Y^{+}|Z)$ is a convex function. If it was a concave function, an upper bound would have been found. The function used in this paper is also convex. therefore we use the described method to calculate the lower bound. It concretely transformed into the following problem:

$$\mathbb{E}[f(y_1,...,y_n)]=\mathbb{E}\big[\mathbb{E}[f(y_1,...,y_n)|Z]\big]\geq\mathbb{E}\big[f[\mathbb{E}(y_1|Z)...\mathbb{E}(y_n|Z)]\big]$$

where $Z=\beta^{'}y$. For each β a lower bound is found. However, if you maximize β, the highest lower bound is found which should be very close to the true value.

It is known that $\begin{pmatrix} y_1\\ \vdots \\ y_n\end{pmatrix}\sim N(\mu_y;\Sigma)$ and $\begin{pmatrix} y\\ Z\end{pmatrix}\sim N\Big(\begin{pmatrix}\mu\\ \beta^{'}\mu\end{pmatrix};\begin{pmatrix}\Sigma & \Sigma\beta\\ \beta^{'}\Sigma & \beta^{'}\Sigma\beta\end{pmatrix}\Big)$

where $\beta^{'}\Sigma\beta=1$.

According to the properties of multivariate normal conditional distributions, we get:

$$y|Z=z\sim N\big(\mu_y+(Z-\beta^{'}\mu_y)\Sigma\beta;(\Sigma-\Sigma\beta\beta^{'}\Sigma)\big)$$

As we know the conditional distribution of y given Z, we can calculate the lower bound. To find the appropriate lower bound, $\varepsilon=(z-\beta\mu_y)$ is a standard normal random variable.

$$\mathbb{E}[f(y)]\geq\mathbb{E}[f(\mu_y+(z-\beta^{'}\mu)\Sigma\beta)]=\mathbb{E}[f(\mu_y+\varepsilon\Sigma\beta)]=\int_{-\infty}^{\infty}f(\mu_y+\varepsilon\Sigma\beta)\phi(\varepsilon)d\varepsilon$$

Note that $f(y)=\prod_{i=1}^{n}\big(1+0.75\max[R_i(y)-K,0]\big)$.

Hence, it can be rewritten into:

$$\mathbb{E}\big[\prod_{i=1}^{n}\big(1+0.75\max[R_i(y)-K,0]\big)\big]=\mathbb{E}\big[\mathbb{E}[\prod_{i=1}^{n}\big(1+0.75\max[R_i(y)-K,0]\big)|Z]\big]$$

$$\geq\mathbb{E}\big[\prod_{i=1}^{n}\big(1+\mathbb{E}[0.75\max(R_i(y)-K,0)|Z]\big)\big]=\mathbb{E}\big[\prod_{i=1}^{n}\big(1+[0.75\max(R_i(\mu_y+\varepsilon\Sigma\beta)-K,0)]\big)\big]$$

$$= \int_{-\infty}^{\infty} \prod_{i=1}^{n} \big(1 + g_i(\mu_y + \varepsilon\Sigma\beta)\big)\phi(\varepsilon)d\varepsilon.$$

We can now calculate $g_i(\mu_y + \varepsilon\Sigma\beta)$ by means of the Bachelier formula, and find:

$$\int_{-\infty}^{\infty} \prod_{i=1}^{n} \big(1 + \ g_i(\mu_y + \varepsilon\Sigma\beta)\big)\phi(\varepsilon)d\varepsilon = 0.26961.$$

The value of the lower bound is equal to: 0.26961. Hence, if the policyholder has a sum insured of 1000 euros, he will have at the end of the 50 years a sum insured of at least $1269,61$ euros.

6.1.2 Upper Bound

Eventhough a lower bound has been found which is according to Rogers and Shi (1995) spot on, this paper also derives an upper bound. To find an upper bound, comonotonic random variables are used. This implies that the variables are very strongly positively dependendent (Dhaene, Denuit, Goovaerts, Kaars and Vyncke, 2002).
Recall that

$$y \sim N(\mu, \Sigma).$$

However, at this point the distribution of the comonotonic variable, y_c, is of interest. Dhaene et al. (2002) mention that y is only comonotonic if all couples of (y_i, y_j) are comonotonic. This implies a correlation of 1, i.e. $\rho(y_i, y_j) = 1$. In this case, the variance can be found by looking at the diagonal of the variance of y, i.e.

$$\sigma = \sqrt{\text{diag}(\Sigma)}.$$

The variance of the new monotonic variable can be obtained by $\Sigma_c = \sigma\sigma'$.
The distribution is now given by:

$$y_c \sim N(\mu, \Sigma_c).$$

If we look at the equation to which we would like to have the value for, we get:

$$\mathbb{E}[f(y)] \leq \mathbb{E}[f(y_c)] = \mathbb{E}\big[\mathbb{E}[f(y_c)|Z]\big].$$

Hence, also for the upper bound the conditional distribution is needed.
It can be seen that

$$y_c|Z \sim N\big(\mu_y + (Z - \beta^{'}\mu_y)\Sigma\beta; 0\big).$$

In this case the conditional covariance matrix is equal to zero, which implies that the variance and standard deviation are also equal to zero. As there is no standard deviation, it is not necessary to apply an option pricing formula. Hence, to find the expected value we merely have to plug in the mean to get the value of:

$$\mathbb{E}\big[\prod_{i=1}^{n} \big(1 + 0.75 \max[R_i(y_c) - K, 0]\big)\big] = \mathbb{E}\big[\prod_{i=1}^{n} \big(1 + 0.75 \max[R_i(\mu_{y_c|Z}) - K, 0]\big)\big].$$

Plugging in the conditional mean of y_c gives a value of 0.27630. Hence, the value of the upper bound is equal to: 0.27630. This implies that if the policyholder has a sum insured of 1000 euros, he will have at the end of the 50 years a sum insured of at most 1276.30 euros.

6.2 Numerical Approximation

Next to the analytical approximation, a numerical approximation has been considered: the Monte Carlo simulation. A Monte Carlo simulation is a discrete approximation method which simulates the underlying process a certain amount of times. For an infinite number of simulations, the outcome of the Monte Carlo should converge to the true solution of the problem at hand. As it is impossible to do an infinite number of simulations, often a very large number of simulations is done, which results in an acceptable level of convergence. The Monte Carlo simulation is a very labor-intensive approach to compute the solution. Eventhough today's computers have a lot of computing power, it is not the most efficient method to find the answer, and with large data sets it can become very time-consuming. However, it is an excellent tool to perform checks on the analytical method. For that reason, the numerical approximation obtained in this section is used to check the values of the analytical approximation.

The Monte Carlo simulation is conducted in this research. The purpose of this simulation is to generate random variables which should represent the y-proces according to the distribution we have found previously. The Cholesky decomposition allows us to do a Monte Carlo simulation in which the values of the underlying process are correlated. As the y-process is a path-dependent process, the Cholesky decomposition has to be applied to do a proper simulation.

To generate normal random variables which are correlated and are drawn from the same distribution as y, two new variables have to be introduced, X and C. X is a standard normal variable, and C is a matrix. Then, the values can be drawn from:

$$C^{'}X \sim N(0, C^{'}C).$$

As the y-process has a covariance matrix Σ, we should find a C such that $C^{'}C = \Sigma$. This is the point where the Cholesky decomposition comes in place. The Cholesky decomposition has been used to determine the value for C.

Due to the limitations of the program, the y-process could be simulated only a maximum number of 75.000 times. The simulated values of y are used to calculate the zero-rates in a manner as explained in the previous sections. These zero-rates are then again used to calculate the payoffs, and to find the final result. Each simulation of the Monte Carlo approach gave a different value. However, the useful information is the expected value and the standard deviation. The expected value found is: 0.26921 with a standard deviation of 0.0014. Hence, the initial amount of 1000 euros is expected to become 1269.21 euros by the end of the contract.

Below the results of the three different approximation methods are tabulated:

Lower bound	Upper bound	Monte Carlo simulation
0.26961	0.27630	0.26921 (0.0014)

If we compare these values, it can be seen that the lower bound differs only by 0.29 standard deviations, whereas the upper bound differs by 5.06 standard deviations. As was already expected, the lower bound approximation of Rogers and Shi (1995) is close to the true value of the problem.

7 Conclusion

This paper examined the profit-sharing contract of one policyholder. The contract started in January 2014 and ends in exactly 50 years, with an initial sum insured of 1000 euros. The goal is to calculate the accrued sum insured at the end of December 2063. The profit-sharing depends on the rate of return the company made in that specific year. If it is above a certain threshold the insured amount will be increased. The increase per year is defined by the value of: $0.75 \max(R - 0.03)$. To calculate the accrued sum insured the increases per year are evaluated and connected.

To value this contract, it is necessary to model the rate of return. The return of return is based on the company's investment portfolio which consists of zero-coupon bonds and coupon-bearing bonds. The rate of return consists of two parts: one part is reinvested, whereas the other part is not. In the past the company used the forward rate based on today's term structure to reinvest. This approach gave quite a wide confidence interval. This paper has decided to use one-year zero rates to reinvest. To obtain these zero-rates, a dynamic term structure model has been introduced, the Hull-White model. As the pay-off date was later than the observation date, a convexity correction needed to be applied. This approach of using one-year zero-rates instead of forward rates gave a narrower confidence interval.

As it turned out that the company's rate of return followed a Gaussian distribution, the Bachelier option pricing formula has been applied to calculate the value of the increase per year. The value of the contract is given by: $\mathbb{E}\big[\prod_{i=1}^{n}\big(1+0.75\max[R_i(y)-K,0]\big)\big]$. As the R-process is path-dependent, it was not possible to just multiply the outcomes of the Bachelier pricing formula. Therefore, an approximation method has been used.

An analytical approximation as well as a numerical approximation has been calculated. The analytical approximation methods gave us a lower bound and an upper bound, whereas the Monte Carlo simulation gave us an expected value with a standard deviation. When comparing the outcome of the analytical approximations to the outcome of the Monte Carlo simulation, it turned out that the lower bound was very close to the numerical solution which is in line with the findings of Rogers and Shi (1995).

In the particular example used in this research, the policyholder had a profit-sharing contract and it was known that he would pass away in exactly 50 years. Obviously, this is not a representative situation of the real world. In the real world, the moment of death of a policyholder is unknown and cannot be determined. However, probabilities can be assigned to the moment a person passes away. Taking these probabilities into account, it is possible to value a profit-sharing contract. The analytical approximation to obtain a lower bound can be used to calculate the value for each possible year of death, i.e. the value of the contract in case he would pass away in exactly 48 years, 49 years, etc. These values should be multiplied by the probability that a person lives until that age, and passes away in that specific year. When using this approach it is possible to value the profit-sharing contract of a policyholder whose moment of death is unknown.

One of the drawbacks of the used model, the Hull-White model, is the possibility of negative interest rates. Although the probability is already lower for the Hull-White model than it is for some other models, it is still larger than zero. If it is possible to use a one-factor model for the interest rates which is analytically tractable and does have a probability of zero for negative interest rates, it would be an improvement.

8 Appendix

8.1 Deriving the Distribution of the Interest Rate

The dynamics of $r(t)$ in the Hull-White model are given by:

$$dr(t) = [\vartheta(t) - ar(t)]dt + \sigma dW(t).$$

To find the mean and variance of $r(t)$, it is necessary to integrate above equation. According to Mikosch (2008) the differential equation $dX_t = c_1(t)X_t dt + \sigma_1(t)X_t dB_t$ can equivalently be written as $X_t = X_0 + \int_0^t c_1(s)X_s ds + \int_0^t \sigma_1(s)X_s dB_s$. This logic also applies to this differential equation:

$$dr(t) = [\vartheta(t) - ar(t)]dt + \sigma dW(t)$$

$$dr(t) + ar(t)dt = \vartheta(t)dt + \sigma dW(t)$$

$$d\big(r(t)e^{\int_s^t a du}\big) = \vartheta(t)e^{\int_v^t a du}dt + \sigma e^{\int_v^t a du}dW(t)$$

$$r(t) = r(s)e^{\int_s^t a du} + \int_s^t \vartheta(v)e^{\int_v^t a du}dv + \int_s^t \sigma e^{\int_v^t a du}dW(v)$$

$$= r(s)e^{-a(t-s)} + \int_s^t \vartheta(v)e^{-a(t-v)}dv + \sigma \int_s^t e^{-a(t-v)}dW(v)$$

It is important to note that : $\vartheta(v) = \frac{\partial f^M(0,v)}{\partial v} + af^M(0,v) + \frac{\sigma^2}{2a}(1 - e^{-2av})$.
Then:

$$\int_s^t \vartheta(v)e^{-a(t-v)}dv = \int_s^t \frac{\partial f^M(0,v)}{\partial v}e^{-a(t-v)} + af^M(0,v)e^{-a(t-v)} + \frac{\sigma^2}{2a}(1 - e^{-2av})e^{-a(t-v)}dv$$

$$= \int_s^t \frac{\partial f^M(0,v)}{\partial v}e^{-a(t-v)} + af^M(0,v)e^{-a(t-v)}dv + \int_s^t \frac{\sigma^2}{2a}(1 - e^{-2av})e^{-a(t-v)}dv$$

$$= f^M(0,t)e^{-a(t-t)} - f^M(0,s)e^{-a(t-s)} + \frac{\sigma^2}{2a}\int_s^t (1 - e^{-2av})e^{-a(t-v)}dv$$

Let's have a look at $\frac{\sigma^2}{2a}\int_s^t (1 - e^{-2av})e^{-a(t-v)}dv$ solely:

$$\frac{\sigma^2}{2a}\int_s^t (1 - e^{-2av})e^{-a(t-v)}dv$$

$$= \frac{\sigma^2}{2a}\Big(\int_s^t e^{-a(t-v)}dv - \int_s^t e^{-2av}e^{-a(t-v)}dv\Big)$$

$$= \frac{\sigma^2}{2a}\left(\frac{1}{a}\big[1 - e^{-a(t-s)}\big] + \frac{1}{a}\big[e^{-2at} - e^{-a(t+s)}\big]\right)$$

$$= \frac{\sigma^2}{2a^2}(1 + e^{-2at}) - \frac{\sigma^2}{2a^2}e^{-a(t-s)}\big(1 + e^{-2as}\big)$$

$$= \frac{\sigma^2}{2a^2}(1 + e^{-2at}) - \frac{\sigma^2}{2a^2}\big(2e^{-at} - 2e^{-at}\big) - \frac{\sigma^2}{2a^2}e^{-a(t-s)}\big(1 + e^{-2as}\big)$$

$$= \frac{\sigma^2}{2a^2}(1 - e^{-at})^2 - \frac{\sigma^2}{2a^2}e^{-a(t-s)}\big(1 - e^{-as}\big)^2.$$

If we plug this in again, we get:

$$r(t) = r(s)e^{-a(t-s)} + \alpha(t) - \alpha(s)e^{-a(t-s)} + \sigma \int_s^t e^{-a(t-v)}dW(v),$$

where $\alpha(t) = f^M(0,t) + \frac{\sigma^2}{2a^2}(1-e^{-at})^2$ and the market instantaneous forward rate is given by: $f^M(0,t) = -\frac{\partial \ln P^M(0,t)}{\partial t}$.

The variance can be calculated by the well-known formula:

$$\mathbb{E}\big[\big(r(t)-\mathbb{E}[r(t)]\big)^2\big] = \mathbb{E}\big[\big(\sigma \int_s^t e^{-a(t-v)}dW(v)\big)^2\big]$$

$$= \sigma^2 \mathbb{E}\big[\int_s^t e^{-a(t-v)}dv\big] = \frac{\sigma^2}{2a}[1-e^{-2a(t-s)}]$$

Hence, $r(t)|\mathcal{F}_s \sim N\big(\mathbb{E}\{r(t)|\mathcal{F}_s\}, \mathrm{Var}\{r(t)|\mathcal{F}_s\}\big)$ and the mean and variance are expressed as

$$\mathbb{E}\{r(t)|\mathcal{F}_s\} = r(s)e^{-a(t-s)} + \alpha(t) - \alpha(s)e^{-a(t-s)}$$

$$\mathrm{Var}\{r(t)|\mathcal{F}_s\} = \frac{\sigma^2}{2a}[1-e^{-2a(t-s)}],$$

where $\mathcal{F}_s$ denotes the information set at time s.

8.2 Deriving Distribution including Convexity Correction

To derive the distribution of the r-process including the convexity correction, it is necessary to have a look at $dR(t) = d\big(\frac{D(t)}{N(t)}\big)$ where D(t) is the asset pricing process and N(t) the numéraire. The asset pricing process is given by $D(t) = e^{A(t,T)-B(t,T)r}$, and as a numéraire the money market account is used, i.e. $N(t) = e^{\int_0^t r(s)ds}$. In this case, we actually need $\frac{1}{N(t)} = e^{-\int_0^t r(s)ds}$.

To solve $dR(t)$, Ito's product rule turns out to be useful:

$$dR(t) = d\big(\frac{D(t)}{N(t)}\big) = d\big(D(t)\frac{1}{N(t)}\big) = \frac{1}{N(t)}dD(t) + D(t)d\frac{1}{N(t)} + dD(t)d\big(\frac{1}{N(t)}\big).$$

At this moment, it is important to derive the expressions for $dD(t)$ and $d\frac{1}{N(t)}$.
Ito's Lemma is given by:

$$df(t,r) = \left(\frac{\partial f}{\partial t} + \mu\frac{\partial f}{\partial r} + \frac{\sigma^2}{2}\frac{\partial^2 f}{\partial r^2}\right)dt + \sigma\frac{\partial f}{\partial r}dB(t)$$

By means of Ito's Lemma it can be seen that $d\big(\frac{1}{N(t)}\big) = -r\frac{1}{N(t)}dt$.
Also $dD(t)$ can be derived by using Ito's Lemma. However, this outcome is less obvious and more labor-intensive.

For ease of notation, $D(t)$ is renamed into $f(t,r)$ during the period that Ito's Lemma is applied. Then

$$f(t,r) = e^{A(t,T)-B(t,T)r} = e^{A(t,T)}e^{-B(t,T)r},$$

where $A(t,T) = ln\frac{P^M(0,T)}{P^M(0,t)} + \big[B(t,T)f^M(0,t) - \frac{\sigma^2}{4a}(1-e^{-2at})B(t,T)^2\big]$ and $B(t,T) = \frac{1}{a}\big[1-e^{T-t}\big]$
At first, the partial derivatives have to be derived:

$$\frac{\partial f(t,r)}{\partial r} = e^{A(t,T)}(-B(t,T))e^{-B(t,T)r} = B(t,T)f(t,r)$$

$$\frac{\partial^2 f(t,r)}{\partial r^2} = e^{A(t,T)}B(t,T)^2e^{-B(t,T)r} = B(t,T)^2 f(t,r)$$

$$\frac{\partial f(t,r)}{\partial t} = \frac{\partial A(t,T)}{\partial t}e^{A(t,T)}e^{-B(t,T)r} + r\frac{\partial(-B(t,T))}{\partial t}e^{A(t,T)}e^{-B(t,T)r}$$

$$= \frac{\partial A(t,T)}{\partial t} f(t,r) + r\frac{\partial(-B(t,T))}{\partial t} f(t,r)$$

Plugging these values into Ito's Lemma, the following is obtained:

$$df(t,r) = \Big[\frac{\partial A(t,T)}{\partial t} f(t,r) + r\frac{\partial(-B(t,T))}{\partial t} f(t,r) + (\vartheta(t) - ar)B(t,T)f(t,r) + \frac{\sigma^2}{2}B(t,T)^2 f(t,r)\Big]dt$$

$$+\sigma(-B(t,T))f(t,r)dW(t)$$

$$= \Big[\Big(\frac{\partial A(t,T)}{\partial t} + r\frac{\partial(-B(t,T))}{\partial t} + (\vartheta(t) - ar)B(t,T) + \frac{\sigma^2}{2}B(t,T)^2\Big) f(t,r)\Big]dt + \sigma(-B(t,T))f(t,r)dW(t)$$

At this point, we have to take a look at the two differential equations: $\frac{\partial A(t,T)}{\partial t}$ and $\frac{\partial B(t,T)}{\partial t}$.
The solution to $\frac{\partial B(t,T)}{\partial t}$ is rather easy to obtain:

$$\frac{\partial B(t,T)}{\partial t} = -e^{-a(T-t)}$$

The solution to $\frac{\partial A(t,T)}{\partial t}$ is more time-consuming.
Recall:

$$A(t,T) = ln\frac{P^M(0,T)}{P^M(0,t)} + \big[B(t,T)f^M(0,t) - \frac{\sigma^2}{4a}(1 - e^{-2at})B(t,T)^2\big]$$

$$= ln(P^M(0,T)) - ln(P^M(0,t)) + \big[B(t,T)f^M(0,t) - \frac{\sigma^2}{4a}B(t,T)^2 + \frac{\sigma^2}{4a}e^{-2at})B(t,T)^2\big]$$

Then,

$$\frac{\partial A(t,T)}{\partial t} = 0 - \frac{\partial ln P^M(0,t)}{\partial t} + B(t,T)\frac{\partial f^M(0,t)}{\partial t} + \frac{\partial B(t,T)}{\partial t} f^M(0,t) - \frac{\sigma^2}{4a}\frac{\partial B(t,T)^2}{\partial t} + \frac{\sigma^2}{4a}e^{-2at}(-2ae^{-2at})B(t,T)^2$$

$$= f^M(0,t) + B(t,T)\frac{\partial f^M(0,t)}{\partial t} - e^{-a(T-t)} f^M(0,t) + \frac{\sigma^2}{2a}B(t,T)e^{-a(T-t)} - \frac{\sigma^2}{2a}e^{-2at}B(t,T)e^{-a(T-t)} - \frac{\sigma^2}{2}e^{-2at}B(t,T)^2$$

where $\frac{\partial B(t,T)^2}{\partial t} = 2B(t,T)\frac{\partial B(t,T)}{\partial t} = -2e^{-a(T-t)}B(t,T)$

Let's have a look again at:

$$df(t,r) = \Big[\Big(\frac{\partial A(t,T)}{\partial t} + r\frac{\partial(-B(t,T))}{\partial t} + (\vartheta(t) - ar)B(t,T) + \frac{\sigma^2}{2}B(t,T)^2\Big) f(t,r)\Big]dt + \sigma(-B(t,T))f(t,r)dW(t).$$

To be able to keep an overview, it is better to start with the first part of the equation, and plug in the values for $\frac{\partial A(t,T)}{\partial t}$ and $\frac{\partial B(t,T)}{\partial t}$. It becomes:

$$\Big(\frac{\partial A(t,T)}{\partial t} + r\frac{\partial(-B(t,T))}{\partial t} + (\vartheta(t) - ar)B(t,T) + \frac{\sigma^2}{2}B(t,T)^2\Big)$$

$$= f^M(0,t) + B(t,T)\frac{\partial f^M(0,t)}{\partial t} - e^{-a(T-t)} f^M(0,t) + \frac{\sigma^2}{2a}B(t,T)e^{-a(T-t)} - \frac{\sigma^2}{2a}e^{-2at}B(t,T)e^{-a(T-t)} - \frac{\sigma^2}{2}e^{-2at}B(t,T)^2$$

$$+re^{-a(T-t)} + arB(t,T) - \big[\frac{\partial f^M(0,t)}{\partial t} + af^M(0,t) + \frac{\sigma^2}{2a}(1 - e^{-2at})\big]B(t,T) + \frac{\sigma^2}{2}B(t,T)^2$$

$$= \frac{\sigma^2}{2a}B(t,T)e^{-a(T-t)} - \frac{\sigma^2}{2a}e^{-2at}B(t,T)e^{-a(T-t)} - \frac{\sigma^2}{2}e^{-2at}B(t,T)^2 + r - \frac{\sigma^2}{2a}(1 - e^{-2at})]B(t,T) + \frac{\sigma^2}{2}B(t,T)^2.$$

At this point it is possible to plug in $B(t,T) = \frac{1}{a}\big[1 - e^{T-t}\big]$ in the equation we just obtained. It turns out that almost all terms cancel against each other, except for one: r. $df(t,r)$ can now be expressed as:

$$df(t,r) = rf(t,r)dt - \sigma B(t,T)f(t,r)dW(t).$$

Recall that for ease of notation, $D(t)$ has been called $f(t,r)$. Hence, an expression for $dD(t)$ is derived :

$$dD(t) = rD(t)dt - \sigma B(t,T)D(t)dW(t).$$

Note that the value of $d\big(\frac{1}{N(t)}\big)$ was already derived at the beginning of this section. Now the point is reached that these values can be plugged into Ito's product rule:

$$dR(t) = d(D(t)\frac{1}{N(t)}) = \frac{1}{N(t)}dD(t) + D(t)d\frac{1}{N(t)} + dD(t)d\big(\frac{1}{N(t)}\big)$$

$$= \frac{1}{N(t)}\big(rD(t)dt - \sigma B(t,T)D(t)dW(t)\big) + D(t)(-r\frac{1}{N(t)}dt) + \big(rD(t)dt - \sigma B(t,T)D(t)dW(t)\big)\big(-r\frac{1}{N(t)}dt\big)$$

$$= r\frac{D(t)}{N(t)}dt - \frac{D(t)}{N(t)}\sigma B(t,T)dW(t) - r\frac{D(t)}{N(t)}dt$$

Hence, it can be seen that:

$$dR(t) = R(t)(-\sigma B(t,T))dW(t)$$

By means of Girsanov's theorem, the following can be derived:

$$dW^T(t) = dW(t) - \frac{1}{D(t)}\big(rD(t)dt - \sigma B(t,T)D(t)dW(t)\big)dW(t)$$

$$= dW(t) - rdtdW(t) + \sigma B(t,T)dW(t)^2$$

$$= dW(t) + \sigma B(t,T)dt$$

Rewriting this equation and plugging in gives:

$$dr(t) = \big(\vartheta(t) - ar(t)\big)dt + \sigma\big(dW^T - \sigma B(t,T)dt\big)$$

$$= \big(\vartheta(t) - \sigma^2 B(t,T) - ar(t)\big)dt + \sigma dW^T$$

At this point it is possible to start deriving the distribution of the process. If you look closely at the formula, you can see the similarities with the formula of the original Hull-White formula. However, in this case there is one additional term, $-\sigma^2 B(t,T)dt$. For that reason, the focus will only be on this term at the moment, and for the rest of the derivations you can have a look at Appendix 8.1. Following a similar reasoning as in Appendix 8.1, we get:

$$\int_s^t e^{-a(t-u)}\sigma^2 B(u,T)du$$

$$= \frac{\sigma^2}{a}\int_s^t e^{-a(t-u)}\big(1 - e^{-a(T-u)}\big)du$$

$$= \frac{\sigma^2}{a}\int_s^t e^{-a(t-u)}du - \int_s^t e^{-a(T+t-2u)}du$$

$$= \frac{\sigma^2}{a^2}\big[\big(1 - e^{-a(t-s)}\big) - \frac{1}{2}\big(e^{-a(T-t)} - e^{a(T+t-2s)}\big)\big] = M^T.$$

If the findings of Appendix 8.1 and Appendix 8.2 are combined, the following mean and variance correspond to the distribution of the interest rate including the convexity correction:

$$\mathbb{E}^T\{r(t)|\mathcal{F}_s\} = r(s)e^{-a(t-s)} + \alpha(t) - \alpha(s)e^{-a(t-s)} - M^T(s,t)$$

$$\mathrm{Var}^T\{r(t)|\mathcal{F}_s\} = \frac{\sigma^2}{2a}\big[1 - e^{-2a(t-s)}\big].$$

As the additional term, $-\sigma^2 B(t,T)$, does not have an influence on the variance, the variance remains unchanged.

References

Bingham, N.H. and R. Kiesel (2004). Risk-Neutral Valuation - Pricing and Hedging of Financial Derivatives, 2nd edition. *Springer Finance*

Black, F.(1975). The pricing of commodity contracts. *Journal of Financial Economics 3, pp. 167-179.*

Brigo, D. and Mercurio, F. (2001). Interest Rate Models - Theory and Practice. *Springer Finance.*

Dhaene, J., M. Denuit, M.J. Goovaerts, R. Kaas and D. Vyncke (2002). The Concept of Comonotonicity in Actuarial Science and Finance: Theory. *Insurance: Mathematics and Economics 31, pp. 3-33.*

Faris, W.G. (2001). Lectures on Stochastic Processes. *Retrieved from: http://math.arizona.edu/ faris/stoch.pdf.*

Finch, S. (2004). Ornstein-Uhlenbeck Process. *Retrieved from: http://www.people.fas.harvard.edu/ sfinch/csolve/ou.pdf.*

Gerber, H.U. (1997). Life Insurance Mathematics, 3rd edition. *Springer-Verlag.*

Hull, J.C. (2014). Fundamentals of Futures and Options Markets, 8th edition. *Pearson.*

Hull, J. and A. White (1994). Numerical Procedures for Implementing Term Structure Models I: Single-Factor Models. *Journal of Derivatives 2(1), pp. 7-16.*

Kwok, Y.-K. (2008). Mathematical Models of Financial Derivatives, 2nd edition. *Springer Finance.*

Levy, E. (1992). Pricing European Average Rate Currency Options. *Journal of International Money and Finance 11, pp. 474-491.*

Mikosch, T. (2008). Elementary Stochastic Calculus with Finance in View. *World Scientific, Singapore.*

Pelsser, A.A.J. (2000). Efficient Methods for Valuing Interest Rate Derivatives. *Springer-Verlag.*

Pelsser, A.A.J. (2001). Mathematical Foundation of Convexity Correction. *Quantitative Finance 3(1), 2003.*

Pelsser, A.A.J. (2013). Convexity Correction Calculations. *Kleynen Consultants.*

Pizzi, M.A. and G.L. Montgomery (1998). The Pricing of Path-Dependent European Options via Monte Carlo Simulation. *Journal of The Franklin Institute 335A, pp. 57-63.*

Rogers, L. and Z. Shi (1995). The Value of an Asian Option. *Journal of Applied Probability 32, pp. 1077-1088.*

Schaefer, M.P. (2002). Pricing and Hedging European Options on Futures Spreads Using the Bachelier Spread Option Model. *Conference, April 22-23, 2002, St. Louis, Missouri 19055, NCR-134 Conference on Applied Commodity Price Analysis, Forecasting, and Market Risk Management.*

Sullivan, E.J. and T.M. Weithers (1991). Louis Bachelier: The Father of Modern Option Pricing Theory. *Journal of Economic Ecudation 22(2), pp. 165-171.*