ERASMUS UNIVERSITY ROTTERDAM

REPRINT PROHIBITED

Erasmus School of Economics

Master Thesis

Econometrics & Management Science

To Hedge, Or Not To Hedge

The Cost of Complying to Dutch Pension Fund Regulation

Author:
J.J.M. van Spronsen
Student no. 439615

Supervisor:
Prof. Dr. C.G. De Vries

Second Reader:
Dr. A. Pick

Rotterdam, June 17, 2019

Abstract

In this thesis I investigate the costs of complying with Dutch pension fund regulation. More specifically, this study investigates whether the benefits of interest rate hedging outweigh the costs by means of a simulation of a general pension fund. The simulation discriminates between periods of economic expansion and slump. I find that in periods of economic downturn a difference in transaction cost expenditures of 8 bps of total fund value exists between funds that maintain a high hedge ratio versus their less hedged counterparts. Furthermore, the gains in fund performance from hedging disappear at longer time horizons. In times of economic upswing the costs of hedging are a factor 5-6 lower compared to periods of economic distress. In such periods, differences in fund performance are tiny. While in times of economic bust regulation could be relaxed, a modest hedge is advised in a boom.

Keywords: nFTK, Pension Fund, interest Rate Swaps, Hedging, Heston model, Hull-White model, Kalman Filter, Particle Filter

"If people do not believe that mathematics is simple, it is only because they do not realize how complicated life is."

- JOHN VON NEUMANN, 1947.

Acknowledgements

This thesis took a lot of time to complete. First of all, I would like to thank Prof. Dr. Casper De Vries, Thijs Aaten, and Dr. Jaroslav Krystul for their patience. Although my acceptance into the MPhil programme at the Tinbergen Institute is a great opportunity for my academic career, the current project was still unfinished. Thijs Aaten helped me in the process of getting accepted for which I am forever grateful.

This project has taught me a lot about the field of financial mathematics, which was relatively new to me. I am grateful that Prof. Dr. De Vries saw me fit to start a project that touched upon more fields than econometrics alone. The discussions with Thijs Aaten not only helped me understand Dutch regulation and finance better, but revealed the world of pensions to me. Considered by most as a dusty field, the Dutch pension world is intriguing for those who are interested in mathematics, economics, and regulation. I would like to thank Dr. Jaroslav Krystul for his helpful mathematical comments on the derivations in this thesis.

Furthermore, I express my gratitude towards Prof. Dr. Peter Spreij who made some useful comments on the derivations of my models. Also, I would like to thank Dr. Lech Grzelak for his opinion on my estimation method.

And last but not least, I thank my family and friends for their support. Without their understanding this project would have been a lot heavier.

Josha van Spronsen
Amsterdam, August 9, 2018

Contents

List of Figures

List of Tables

List of Algorithms

1 Introduction

Since the Great Financial Crisis of 2007 EU regulation of institutional investors has increased drastically. The Solvency II, Basel III and IV are examples of EU insurance and banking regulation that came to life since 2009. To minimize the risk of insolvency of Dutch pension funds the Nederlandsche Bank (DNB) introduced the new Financial Assessment Framework (Dutch: nieuw Financieel Toetsingskader, short: nFTK) in 2015. One of the main goals of the nFTK is to limit interest rate risk. The interest rate risk of a pension fund stems from the duration mismatch between the pension benefits (the liabilities) and its interest rate sensitive assets. The nFTK strongly encourages a high hedge position regarding interest rate risk: a high hedge ratio drastically decreases the buffer the fund needs to maintain.

As swaps have an initial price of zero, they are regarded a low cost solution to interest rate risk hedging. There exist, however, several drawbacks to using swaps. First, the use of swaps adds interest rate volatility (rho) to the portfolio by construction. As such, in order to comply to the fund's investment strategy a fund should rebalance its portfolio more frequently, which implies higher transaction costs. Second, as a swap is an over-the-counter (OTC) product, the agreement can not be sold in the market. Only with mutual agreement the contract can be nullified. Another solution to revoke the swap is by entering into another swap agreement with an opposing character to the initial swap. In turn, this translates to higher transaction and management costs.

Although, there exist a rich literature on interest rate swaps, a thorough cost-benefit analysis of interest rate hedging with swaps for pension funds has, to the best of my knowledge, not been carried out. The impact of intensive interest rate hedging on fund performance has not been investigated either. This is an important question as Dutch pension funds are the largest institutional investors in the Netherlands. Moreover, most Dutch pension funds are defined benefit (DB) schemes, where liability risk is carried by the fund. Significant costs or other consequences due to hedging likely have substantial macro-economic effects. Not only a direct effect through investments is imaginable, but also income redistributive effects are plausible: by complying to strict regulation, pension funds often have to refrain from pension benefit indexation or increase the pension contributions. Such measures have a wealth redistributive character (from the young generation to the old, or vice versa).

The question becomes even more interesting if interest rates have a mean-reverting character. As the average duration of pension funds is around 18 years[1], there is ample time for the term-structure to revert to its mean. Then, hedging short term movements in the term-structure might seem redundant. The existence of a mean of the term-structure is disputable as interest rates are continuously subject to monetary policy. If true, regulation regarding interest rate risk might be too strict, especially if hedging turns out to be costly.

To this backdrop, this thesis aims to shed light on the usefulness of swaps for hedging purposes for pension funds. On the one hand, this research tries to unveil the direct (value impact) and indirect (extra rebalancing) costs of hedging practices for pension funds. On the other hand, this study investigates the beneficial aspects of hedging. Moreover, I study how the results to this cost-benefit analysis alter over time, and behave in the long run particularly. Therefore, the research question I try to answer is:
Do the gains from the stabilizing character of interest rate swaps outweigh the costs of interest rate hedging for pension funds? And, does the answer to this question change when focusing on different time horizons?

With the help of the following sub-questions I try to answer this question:

1. *What is the effect of interest rate hedging on the expenditures on transaction costs?*

[1] `https://www.dnb.nl/en/binaries/AV293139%20TK%20Bijlage_tcm47-334285.pdf?2016010115`

2. *What is the effect of interest rate hedging on the performance of the fund?*
3. *Does the funding ratio become more stable as the hedge ratio increases?*
4. *Are more hedged funds less frequent insolvent compared to their less hedged counterparts?*
5. *From a regulative viewpoint, what is the effect of hedging intensity on the level of indexation?*

I expect that transaction costs increase with the chosen hedge ratio of the fund. A larger rho of the portfolio in combination with other investment restrictions mean more frequent rebalancing. If a trend in interest rates movement is absent, then adding swaps should not alter the fund performance, but should make the returns of the pension fund more volatile due to increased rebalancing frequency. However, the volatility of the funding ratio should decrease in the hedging intensity, and hence the frequency of solvent funds should increases. This finding should be reflected in larger price level compensations for highly hedged funds.

I intent to answer these questions by means of a controlled simulation experiment. In this experiment funds are simulated over a length of 20 years, which (more or less) is the average duration of pension funds. To infer the impact of swaps on identical funds that differ in investment and hedging strategy requires many instances of a financial world, which simply do not exist. Moreover, by means of a simulation experiment I can isolate the rules of nFTK and its consequences. Although the assumptions made in creating the model are close to reality, the results should be understood as indicative.

This thesis touches two different strands of literature. First, the financial market model is a combination of the Heston model (1993) and Hull-White model (1990). This hybrid model is studied in Grzelak & Oosterlee (2010). However, whereas Grzelak & Oosterlee (2010) calibrate their model parameters to quoted prices in the financial market, I use historical data to estimate the parameters. Although Girsanov Theorem states that the diffusion term remains identical under the $\mathbb{Q}$- and $\mathbb{P}$-measure, the drift term is different. This is of importance as this study does not concern asset pricing (for which the $\mathbb{Q}$-measure is needed), but risk management (for which the $\mathbb{P}$-measure is key). With the help of a Cholesky decomposition, the model can be historically estimated in two steps. The first stage is a Kalman filter (1960), the second is an adaptation of the particle filter of Aihara et al. (2012).
Second, this thesis adds to the asset liability management (ALM) literature. ALM is a form of asset management that takes the liabilities into account in defining the investment strategy. When applied to pension funds, ALM considers risk on future liabilities and the available instruments, e.g. hedge ratio, premium contribution, indexation strategy, and investment strategy. This study makes use of such a model, focusing on interest rate hedging. This study fits into the literature of ALM modelling of pension funds, such as Drijver (2002), Dert (1995), Broeders et al. (2017). To the best of my knowledge this is the first stochastic ALM-model specification of a pension fund in the literature that focuses on interest rate hedging.

The rest of the thesis is organized as follows. Section 2 gives some background on interest rate risk for pension funds, and the nFTK. Section 3 presents the pension fund model, and the financial market model that lies at its core. Section 4 describes the data on which the financial market model is estimated. Section 5 presents the results of both the financial market model estimation, and the results of the pension fund model. Section 6 concludes the analysis and Section 7 presents some points of discussion.

2 Interest Rate Risk & the New Financial Assessment Framework

Interest rate risk is one of the main risk factors of pension funds[2]. Both sides of the balance sheet of a pension fund are very sensitive to changes in the interest rate curves. As well the value of the assets as the value of the liabilities of the fund are determined by the current yield curves. Due to the different size and maturities of the assets than those of the liabilities, the asset side responds in a different manner to a change in the yield curves than the liability side. This is what is known as the duration gap or duration mismatch, or more generally, the rho gap. The duration gap is the main reason why a pension fund is susceptible to interest rate risk.

In theory, a pension fund could invest in fixed income securities, such as bonds, in such a way that the change in the value of the assets would be congruent with the movement in the value of the liabilities. This is known as a natural hedge, since it covers the interest rate risk. A full natural hedge solely by bonds is infeasible due to the illiquidity of bonds with a maturity equal to the maturity of the liabilities far in the future[3][4]. A report by the DNB (2011)[5] publishes the size of the liabilities and prime grade bonds by maturity for the aggregate Dutch pension sector. Figure 1 shows the assets and liabilities of the aggregate pension sector per maturity. The figure shows the discrepancy (read: duration mismatch) between the liabilities and the bonds, both in value and maturity. Moreover, it depicts the illiquidity of bonds with a maturity higher than 10 years. Furthermore, as liabilities have to be discounted against the risk-free rate since the introduction of the Financial Assessment Framework (Nederlands: Financieel Toetsingskader, short: FTK) in 2007, bonds are not a perfect natural hedge, because they are valued on a different yield curve due to their inherent risk.

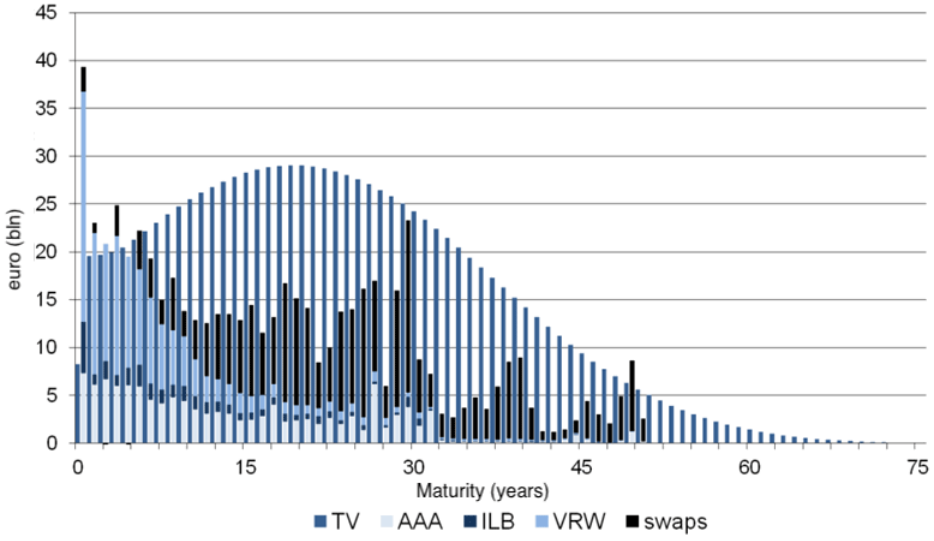


Figure 1: Duration Mismatch, *source: DNB*.

A pension fund can cover its remaining interest rate risk by investing in hedge products, such as interest rate swaps (short: swaps). Figure 1 depicts that Dutch funds add swaps to their portfolio to close the duration gap. Even though the height of the interest rate hedge can be chosen by the

[2] http://www.toezicht.dnb.nl/2/50-202312.jsp

[3] Note that pension benefits range up to roughly 70 years

[4] Even if bonds on the long end of the curve were liquid, pension funds would remain to invest partly in equities to achieve higher returns and to lower potentially the pension premiums.

[5] https://www.dnb.nl/nieuws/nieuwsoverzicht-en-archief/dnbulletin-2011/dnb253798.jsp

fund itself, the investment strategy of pension funds is restrained. These restrictions are denoted in the nFTK, of which the so-called square-root formula is the most important. The square-root formula negatively relates the amount of risk pension fund is exposed to with the buffer they need to maintain. Among the risks the nFTK regards are interest rate risk, equity and real estate risk, currency risk, commodity risk, and credit rate risk. The buffer is expressed in a lower bound on the equity of the fund, the so-called required own equity (Nederlands: Vereist eigen vermogen (VEV)), which should ensure solvency of a pension fund. The standard model of DNB that calculates the VEV as a funding ratio is defined as:

$$VEV = 1 + \frac{\sqrt{S_1^2 + S_2^2 + 2\rho_{12}S_1S_2 + S_3^2 + S_4^2 + S_5^2 + 2\rho_{15}S_1S_5 + 2\rho_{25}S_2S_5}}{100} \tag{2.1}$$

Where:

1. S_1 is interest rate risk, which is defined as the duration mismatch between the liabilities and portfolio (including interest rate hedging products). Therefore, S_1 is a function of the hedge ratio. For instance, if the duration of the liabilities is 40 and the fund targets a 50% hedge ratio, then $S_1 = 40 - 0.5 \times 40 = 20$.

2. S_2 is equity and real estate risk, which is calculated as the percentage change in the value of overall equity and real estate portfolio. For mature markets this shortfall percentage is 30% of the equity and real estate portfolio value.

3. S_3 is currency risk, which is calculated as the aggregate of depreciation of currencies is developed and emerging markets, controlling for correlations in the exchange rate movements.

4. S_4 is commodity risk, which is calculated as an overall decrease (35%) in the commodity portfolio.

5. S_5 is credit rate risk, which is calculated as the effect of a credit spread movement on the bond portfolio.

6. $\rho_{i,j}$ denotes the correlation between risk class i and j.

According to the nFTK, the VEV roughly ranges between 105% and 135%.

The goal of a pension fund is to compensate their nominal pension payments in order to account for inflation. The level of the price level correction depends on the VEV. Moreover, following article 15 of nFTK, a pension fund may not grant price level correction of the pension benefits, if it is not able to meet a funding ratio of at least 110%. If the funding ratio is larger than the VEV, the fund is allowed to grant full indexation. For the intermediate values (between a 110% funding ratio and the VEV) the fund may grant a partial indexation. The most straightforward method is a linear interpolation method between the maximum level of indexation and zero (Dutch: staffelmethode). The maximum level of indexation is calibrated at 2%, the ECB's target of yearly inflation. Hence, the level of indexation is determined by a piecewise linear function. Denote by π the level of indexation and by f the actual funding ratio. Then:

$$\pi = \begin{cases} 0, & f < 1.1 \\ \frac{2}{VEV - 1.1} f, & 1.1 \leq f < VEV \\ 2, & f \geq VEV \end{cases}$$

3 The ALM-Model: A General Pension Fund

This section describes the ALM-model, which consists of two layers. The first layer is the financial market model that provides the simulated equity and term-structure paths. On this financial market the assets and liabilities of the pension fund are priced. Moreover, this market provides the interest rate sensitivity of the fund's portfolio. The second layer is the pension fund model and resembles the asset management of a general pension fund.

The main characteristic of the pension fund model is its investment strategy. This strategy is expressed by two key ratio's: the asset mix and the hedge ratio. The asset mix represents the fractions of the total capital the fund allocates over the different asset classes. The asset mix of the pension fund model is captured by the distribution between equity and fixed income instruments. As a pension fund will use less swaps to cover interest rate sensitivity as it holds a larger bond portfolio, I investigate multiple asset mixes. Concerning equity, optimal portfolio construction as in stock picking or sector timing is outside the scope of this paper. A blue chip index, the Euronext100, represents the stock portfolio of the pension fund, as in Kaschutkze & Muerer (2016). This keeps the model tractable and captures the differences in performance due to different hedging strategies.

This chapter is organized as follows. First, I turn to the assumptions that ensure tractability of the model, which are supported data and/or reports of DNB. Second, I describe the first layer: the models that simulate the financial market. Then, based on the assumptions and the simulations, I provide the algorithm that produces the required results for the pension funds that differ in their investment strategy.

3.1 Assumptions

1. **Assets: equity, bonds, cash, and swaps.**
   The model is restricted to take positions in equity, bonds, cash and swaps. Bikker et al. (2007) states that for small and medium sized pension funds around 90% of the asset value is constituted by bonds and equity. For large pension funds this fraction is somewhat smaller, around 80%. Around 1% of the asset value is captured by cash. Cash is used as a reservation for running costs, such as transaction costs and swaps payments. Following data of DNB 1% of the total asset value is allocated to cash. The rest of the value comes from derivatives, real estate and commodities. The model refrains from real estate, which accounts for roughly 5% of the asset value. Moreover, the model does not consider commodities, which accounts for 1% of the asset value of Dutch pension funds, see Broeders et al. (2017). Concerning derivatives, data by DNB[6] shows that around 95% of the value of the derivative portfolio comes from swaps. As such, swaps are the only derivatives the model trades. Empirically, the considered asset classes constitute the largest part of pension fund portfolio's. As such, the modelled pension fund portfolio should capture the main effects and dynamics.

   Figure 2 below displays the asset mix of the Dutch pension sector in the aggregate. In particular, it shows the allocation of total assets to bonds and equity. The data indicates that most of pension fund capital is invested in either one of these asset classes. Moreover, of the sum of fixed interest securities and equity around 60% is allocated to fixed income and 40% to equity.

[6] https://statistiek.dnb.nl/downloads/index.aspx#/details/derivatenposities-naar-type/dataset/a02b1607-cc2c-4bdf-9887-2503608521bf/resource/1feee170-cf48-40cf-8b36-a09f82db9938

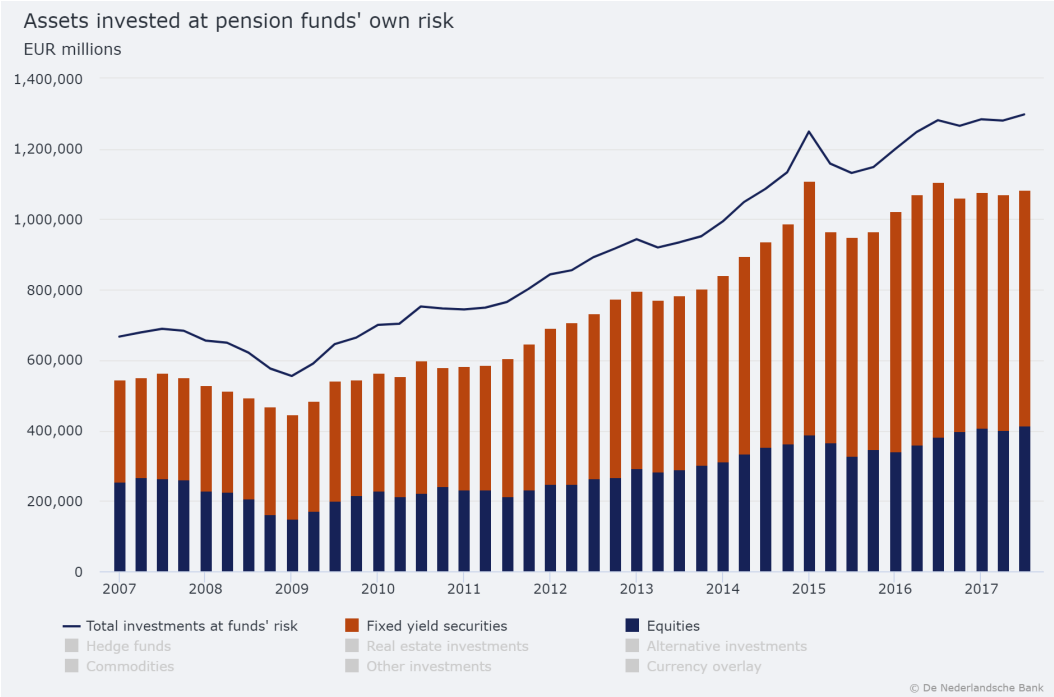


Figure 2: Asset mix of aggregate pension sector, Q1 2007- Q3 2017. *source: DNB.*

To ensure parsimony of the model the interest rate products are restricted to be traded on ten maturities. These pillars are the 1, 2, 3, 5, 7, 10, 15, 20, and 30 year maturities. These maturities are not equidistant, but tries to capture the nonlinearity of the term-structure. Moreover, the hypothetical fund can only trade in prime grade (AAA) bonds. Although not completely true, DNB data on the aggregate pension sector shows that 68% of the bonds in the portfolio have a credit rating A or higher[7]. Around 83% is of investment grade. Figure 3 shows these observations.

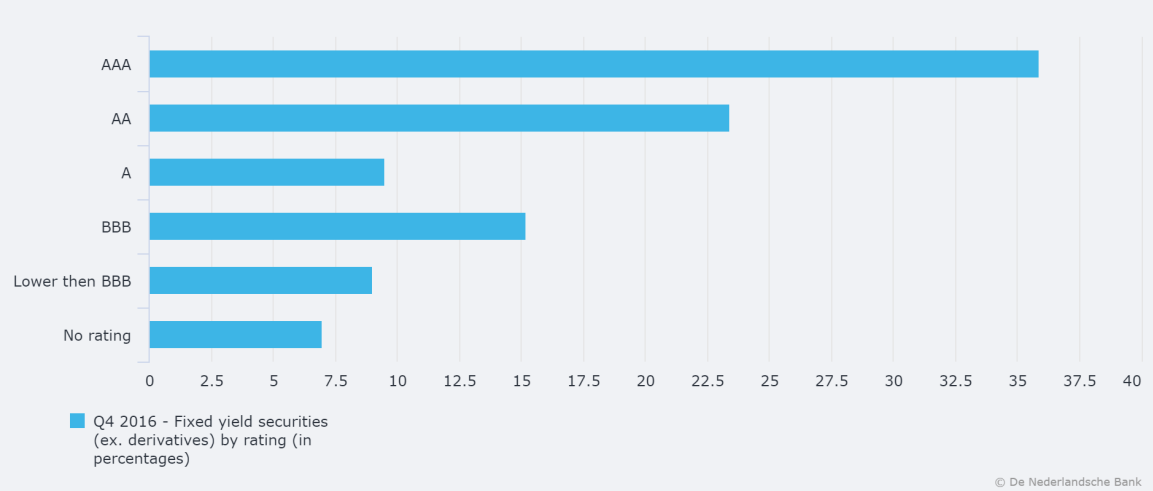


Figure 3: Credit rating of bonds in portfolio of aggregate sector, Q4 2016. *source: DNB.*

[7] https://statistiek.dnb.nl/en/downloads/index.aspx#/details/fixed-income-securities-by-credit-rating/dataset/32d24b43-6395-4102-b1a1-92df0647b24d/resource/302bc395-d89a-4225-9a94-c5f91aef65ed

2. **Liabilities: stationarity and indexation.**
The model assumes that pension contributions are as large as the payments. Data of DNB shows that the incoming pension premiums are of comparable in size to the outgoing pension payments, see Figure 4. This assumption enables the model to focus solely on investment decisions and it's consequences on the funding ratio rather than operational decisions. If this assumption would be violated, the fund could always increase the contribution rate or opt to not correct the pension benefits for inflation until the assumption is met.

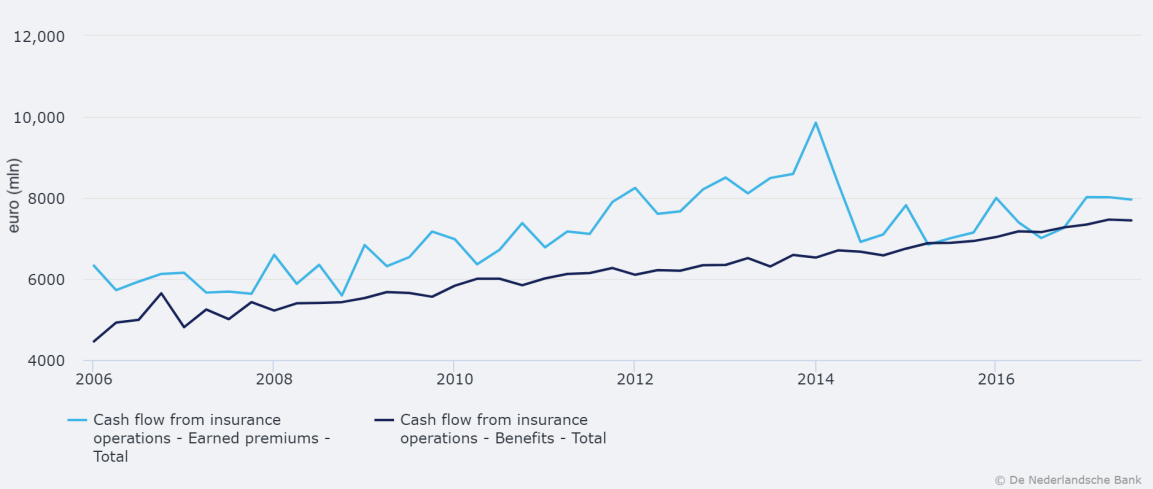


Figure 4: Pension premiums and payments of the aggregate sector. *source: DNB*

Further, I assume that the growth rate of the liabilities equals zero, apart from indexation. Thus, the ageing of the society is outside the scope of this thesis. Ageing has an effect on the shape of the liability curve, see Figure 1. As the life expectancy of society would increase, the fund would have to payout over a longer timespan. This would translate into a thicker tail of the liability curve to the right. Moreover, the probability that elderly will survive increases. As such, the height of the curve would increase at the short term liabilities. On the other hand, as western societies experience drops in fertility rates, the liability curve is expected to shift downwards. Furthermore, the average age of retirement also increases. As such, the time to accumulate pension assets increases and the liability curve shifts to the left as in expectation the time span of pension benefits decreases. Demographic effects can produce contrary effects on the liabilities and migration isn't even considered yet. Modelling these effects is a tough exercise and outside the scope of the thesis. Moreover, in the pension literature there is no consensus on the joint effect of ageing and migration, see e.g. Zaiceva & Zimmerman (2016).

This assumption would be violated if the composition of the population does change significantly over the course of the simulation (20 years) - if the distribution of the population is unstable over the time of roughly one generation. If this assumption does not hold, the results of this study would induce a lower bound. As the liabilities grow due to ageing, so does their interest rate sensitivity. As a result more hedging is needed, which likely translates to an increased use of swaps.

With this assumption in place, the liability curve of the pension fund is stationary with respect to ageing. As such, the effect of interest rate hedging works through a changing economy and not through a change in demographics. This is a desirable feature as it isolates the effect of the use of swaps for pension funds.

Considering indexation, the pension fund model potentially grants yearly price level correction of the pension benefits. The pension fund follows the staffelmethode, as described in previous section.

3. **Rebalancing.**
When the fund's portfolio does not satisfy the investment strategy - that is, the hedge- and equity-to-total-asset-value ratio - the portfolio should be rebalanced. In the asset pricing literature the exists a consensus on infeasibility of continuous rebalancing. More recent papers developed methods that involve so-called bound rebalancing. That is, rebalancing is triggered when the targeted strategy surpasses certain bounds, see e.g. Donohue & Yip (2003) or Woodside-Oriakhi et al. (2013). More interestingly, it has been shown that rebalancing to the point where the investment strategy is exactly satisfied might be sub-optimal as the market potentially moves towards the target after rebalancing already has taken place. Therefore, rebalancing towards a point between the bound and the target is preferred.

   The modelled pension fund has a two-step rebalancing method. The first step is triggered when the equity-to-total-asset-value ratio is off by 3%[8]. Then, the portfolio is rebalanced to 1.5% off target under under minimal transaction costs. Moreover, in this same step, the bonds and swaps are set to satisfy the hedge ratio. The first step is skipped if the fund still satisfies the equity-to-total-asset-value ratio when arriving in a new time point. If the portfolio does not satisfy the hedge ratio, the fund enters contracts to satisfy the hedge ratio. Note that this action does not distort the equity-to-total-asset-value ratio as new swap agreements have nil value when contracted. The result is a portfolio that satisfies the equity-to-total-asset-value ratio and the hedge ratio.

   Important to note is that the rebalancing objective is to establish a general portfolio that satisfies both restrictions under minimal transaction costs instead of making predictions which portfolio will generate maximum return. This model replicates a general pension fund and investigates the effects of interest rate hedging rather than the optimal portfolio. Secondly, swaps cannot be sold as they are OTC contracts. They can only be offset by new swaps agreements in the opposite direction. Therefore, every swap agreement is held until maturity.

4. **Market liquidity.**
The market is assumed to be perfectly liquid, i.e. the pension fund is able to trade all products its investment strategy dictates. It is known that fixed income securities, such as bonds and swaps, on the long end of the curve are less liquid.

   The effect of violation of this assumption on results is unclear. If the liquidity of fixed income products are uncorrelated with the chosen hedging strategy, results are likely to be robust against other specifications of market liquidity - any specification affects every fund in similar fashion. However, it is imaginable that the hedging strategy is endogenous to market liquidity. Pension funds cannot adapt their strategy on short notice responding to liquidity shocks, but probably incorporate the ongoing difference in liquidity of short- and long-term fixed income products. However, for simplicity it is assumed that the fund is freely able to make such trades.

5. **Initialization.**
Initially, the fund is assumed to hold sufficient own capital. This amount is calculated by the

[8]Remember that the target value is 40%

square-root formula of the DNB[9]. This assumption is made to enable honest comparison between pension funds that differ in strategy. For instance, a fund that has a low hedging ratio should hold more own equity as a buffer to compensate the risk. When endowing all different type of funds with the same amount of capital would yield a biased comparison in, for instance, the level on indexation.

The pension fund model investigates the effect of different hedging strategies. Therefore, the different funds considered in the analysis only differ in investment decisions involving hedging. These decisions affect the interest rate risk and equity risk. Other risk factors, such as currency and commodity risk, are homogeneous over the different funds to isolate the effect of the hedging strategy. For homogeneous risk factors the modal estimate is incorporated for all funds:

- Currency risk: currency denomination is not considered in the pension fund model as the model trades in European stock only. However, to create a acceptable resemblance of the VEV, a constant is taken for this type of risk. A modal estimate by DNB for this risk factor is 9.5%[10] of the value of assets denominated in foreign currencies. This estimate is based on a portfolio where 50,51%[11] is invested in foreign assets and 75% of total assets is denominated in developed currencies.
- Commodity risk: As the pension fund model cannot trade in commodities this risk factor is not considered, hence $S_4 = 0$.
- Credit risk: The pension fund model is assumed to invest in AAA sovereign bonds only. As such, S_5 is equal to zero, see DNB[12].
- Further, $\rho_{12} = 0.4$ (if interest rates increase and nil otherwise), $\rho_{15} = 0.4$, and $\rho_{25} = 0.5$ - as calibrated by DNB[13].

Initialization of the portfolio works as follows. When the required own funds are calculated as, for instance, 120% and the present value of liabilities equals €100,-, the fund is endowed with €120,- in cash. With this capital the fund takes positions as it would when rebalancing the portfolio, i.e. the fund buys a portfolio satisfying the equity-to-total-asset-value-ratio and the hedge ratio. The size of the fund is not important for the results. An increase in the number of pension contributors would translate into an increase in both liabilities and contributions.

6. **Constant Spread**.
To ensure tractability of the model, the financial market model simulates one curve, namely the risk-free curve used for discounting the liabilities. However, the model also needs to price the AAA government bonds. This might seem problematic at first sight. However, the model is evaluated on two different data sets. In one of the two samples, the pre-crisis period of Economic Expansion (EE), the two curves coincide. Thus one curve is enough to carry out the necessities, and no spread is needed.

In the other sample, the period of Quantative Easing (QE), the spread between the risk-free curve and the average of AAA government bonds remained relatively stable. Following Zhu (2012), a constant spread is added to the simulated risk-free curve to obtain a curve that can be used for pricing of the AAA government bonds.

[9] http://www.toezicht.dnb.nl/2/50-202138.jsp
[10] http://www.toezicht.dnb.nl/2/50-202274.jsp
[11] https://statistiek.dnb.nl/en/downloads/index.aspx#/details/pension-assets-by-currency-breakdown/dataset/f2031581-8b7d-48b3-9eea-e169f13b4373/resource/50ad2f1e-0eae-4e0b-a8ec-21086ea940f9
[12] http://www.toezicht.dnb.nl/2/50-202270.jsp
[13] http://www.toezicht.dnb.nl/2/50-202138.jsp

3.2 Simulation Financial Market

This section presents the models that are used to generate the equity and interest rate paths. First, I treat the Heston model, which is responsible for the stock prices. Second, I turn to the Hull-White model that generates the yield curves. Then, I present the hybrid model, which combines the Heston- and Hull-White model into one joint model.

3.2.1 Equity: Heston Model

To simulate equity paths I use a stochastic volatilty (SV) model. Since there exists a vast amount of SV models, I give an overview of the models that have been developed over the years and provide argumentation of the selection of the appropriate model.

Probably the most popular model where the dynamics of an asset price was modelled, is the Black-Scholes model (1973), which is closely related to the work of Bachelier (1900). The paper applies the link between a Brownian motion and the heat-diffusion equation, which was discovered by Einstein (1905), to the field of financial mathematics and was awarded with a Nobel prize. The popularity of this model stemmed from the analytical results for the value of options, the Black-Scholes formula. However, this model does not belong to the class of SV models, since the derivation rests on the assumption of constant volatility. Moreover, this assumption is proven not to hold in most cases. Mandelbrot already noticed in 1963 that there seemed to be a clustering effect of volatility: "Large changes tend to be followed by large changes, of either sign, and small changes tend to be followed by small changes." Black and Scholes noted this problem themselves in 1972 (p. 416): "... there is evidence of non-stationarity in the variance. More work must be done to predict variances using the information available." Next to the clustering, volatility is subordinated to the leverage effect, where positive changes in the stock price affect the volatility differently in a mild and negative manner, whereas negative changes in the stock price do the opposite, see e.g. Nelson (1991) and Engle & Ng (1993).

To incorporate the aforementioned effects and provide bias-free results, autoregressive conditional heteroskedasticity (ARCH) and SV models were developed in the late eighties and the beginning of the nineties. The ARCH, Engle (1982), and generalized ARCH models, Bollerslev (1986), is not the focus of this thesis, nor will be their further developments, such as Glosten, Jagannathan and Runkle (1993) who modelled the asymmetric leverage effect. The reason I refrain from ARCH models in the analysis is twofold. The first reason is that these models incorporate positivity constraints on the parameters to ensure that the variance process remains positive. In the process of estimation this constraint is often violated, see Nelson (1991). Although the exponential GARCH (EGARCH) model of Nelson (1991) relaxed these constraints, the second reason to abstract from ARCH models remains: in the class of ARCH models, the conditional variance, conditional on past returns and/or past variances, is explicitly modelled. As such, random oscillatory behaviour of the variance process is not likely to happen.

In SV models the distribution of the conditional variance is modelled indirectly through the structure of the model, the stochastic differential equations. This seems more natural, and more convenient in continuous time, since such a model does not model the predictive conditional distribution directly. At first, there was critique by econometricians that such models were not easy to estimate or test. But with the arrival of simulation based methods that were able to do so, these models gained a lot of popularity. Especially since these models are better able to explain the observations in the financial markets.

The first SV model alternative that was provided as an alternative to the ARCH models was by Taylor (1982). This model is now known as the log-normal SV model if the innovations are assumed

to be Gaussian. Johnson & Shanno (1987) and Hull-White (1987) were the first to cast SV models into continuous time. The latter model was the first to provide a coherent structure to incorporate the leverage effect and looked like the continuous counterpart of the GARCH model. Scott (1987) incorporated a long term mean by modelling the volatility as a mean reverting Ornstein-Uhlenbeck process. The inclusion of mean-reversion for the volatility is a desirable feature, as it is consistent with empirics. Wiggins (1987) went further and modelled the logarithm of the volatility to ensure positiveness of the process. The Stein and Stein (1991) model was the first SV model to provide a closed-form solution for as well the distribution of the asset price as the option price. One main drawback of the model is that it is able to produce negative values for the volatility. All the before-mentioned models lack the direct correlation between the stock dynamics and the volatility process. As such the models are unable to capture the skewness of the returns, see Heston (1993).

The model developed by Heston (1993) allows for non-zero correlation between the asset and volatility dynamics and is able to provide a closed-form solution of the option price. Moreover, since the dynamics of the volatility in this model are modelled as a Cox-Ingersoll-Ross (1985) process, the volatility cannot become negative. Bates (1996) and Scott (1997) both extended the Heston (1993) model to allow for jumps in the stock dynamics. The main difference between these two models is that the latter allows for stochastic interest rates. Schöbel and Zhu (1999) extended the Stein and Stein model (1991) to allow for non-zero correlation between the asset and volatility dynamics. There is much empirical evidence that the original Heston model outperforms the extensions, see e.g. Bakshi, Cao & Chen (1997), Kim and Kim (2004), and Moon et al. (2009). Moreover, the original Heston model is more parsimonious and is therefore easier to estimate. Later work, e.g. the papers by Barndorff-Nielsen & Shephard (2001), Chernov, Gallant, Ghysels and Tauchen (2003) and Nicolato and Venardos (2003), focuses more on long memory to better explain the behaviour of the dynamics of volatility in a high-frequency setting. As I will be working with long term horizons sampled on a less frequent basis, such models are not needed. Based on the evidence given above and what is needed for the research, the model chosen for the analysis is the Heston model.

Dynamics

Assuming the probability space $(\Omega, \mathcal{F}, \mathbb{P})$, the Heston model takes the form of:

$$\begin{aligned} dS(t) &= \mu S(t)dt + \sqrt{V(t)}S(t)dW_S(t) \\ dV(t) &= \kappa\left(\theta - V(t)\right)dt + \sigma\sqrt{V(t)}dW_V(t) \end{aligned} \tag{3.1}$$

where,

- μ is the rate of return on the asset. This can easier be seen by rewriting the dynamics of the value of the asset as $dS(t)/S(t) = \mu dt + \sqrt{V(t)}dW_S(t)$. The interpretation of this parameter depends on the time definition used in the model. For instance, when one looks at daily data, using $dt = 1/252$ implies that μ will represent the yearly expected rate of return.
- θ is the long term mean of the instantaneous variance of the asset $V(t)$.
- κ is the speed with which $V(t)$ reverts to it's long term mean θ.
- σ is the variance of $V(t)$.
- ρ is the correlation of the Wiener increments $dW_S(t)$ and $dW_V(t)$. Put differently, the quadratic variation $d\left[W_S, W_V\right](t) = \rho dt$. Typically, this parameter is negative due to the leverage effect; a decrease in the asset value typically goes hand in hand with an increase in the volatility.

Negative values of the variance process are prevented if $2\kappa\theta \geq \sigma^2$. This condition is known as the Feller condition, see Feller (1951). Intuitively, when the variance process reaches zero, the volatility term $\sigma\sqrt{V(t)}$ approaches zero as well. Subsequently, the process is pushed back towards θ at such a point. However, this only holds if the upward drift is large enough, i.e. if the Feller condition is satisfied.

In what follows it turns out useful to transform the dynamics of the asset to the log dynamics. By applying Itô's lemma, one obtains:

$$\begin{aligned} d\log S(t) &= \underbrace{\frac{\partial \log S(t)}{\partial t}}_{=0} dt + \underbrace{\frac{\partial \log S(t)}{\partial S(t)}}_{=1/S(t)} dS(t) + \frac{1}{2} \underbrace{\frac{\partial^2 \log S(t)}{\partial S(t)^2}}_{=-1/S(t)^2} \underbrace{[dS(t), dS(t)]}_{=V(t)S(t)^2 dt} \\ &= \left(\mu - \frac{1}{2}V(t)\right) dt + \sqrt{V(t)} dW_S(t) \end{aligned} \tag{3.2}$$

By applying a Choleksy decomposition on the correlation structure of the Wiener processes, one can rewrite the Heston model in terms of two independent Wiener processes; $\{W_1(t)\}_{t\geq 0}$ and $\{W_2(t)\}_{t\geq 0}$, $W_1(t) \perp\!\!\!\perp W_2(t), \forall t \geq 0$. With the help of equation (3.1) and (3.2) the Heston model can be rewritten as:

$$\begin{aligned} d\log S(t) &= \left(\mu - \frac{1}{2}V(t)\right) dt + \sqrt{V(t)} dW_1(t) \\ dV(t) &= \kappa(\theta - V(t)) dt + \sigma\sqrt{V(t)} \left(\rho dW_1(t) + \sqrt{1-\rho^2} dW_2(t)\right) \end{aligned} \tag{3.3}$$

In deriving the solution of the model by integrating the system, one stumbles upon the stochastic integral $\int_s^t V(u)du$. The difficulty of the the analytical solution of the Heston model lies in the impossibility to calculate this stochastic integral.

As the volatility process is a latent process, standard maximum likelihood methods are not applicable. Furthermore, due to the non-Gaussian transition distribution, the Kalman filter is not suitable either, since the updating equations are derived by invoking results for the normal distribution. A widely used method in historically estimating the Heston model is the particle filter. The particle filter rests on simulating the propagation of the model and inferring which paths are most likely. Hence, simulation of the process is a key element in order to estimate the model. As such, there exists a rich literature on the simulation of the Heston model, which I summarize below. In simulating the Heston model, there are two different approaches.

The first approach is to apply a discretization scheme. Discretization schemes partition the interval on which one wants to solve the SDE and approximate the solution to the SDE on this grid. The finer the grid is, the better the approximation. This approach is the most popular in the literature. The main rationale underlying this choice is that discretization delivers transition distributions for which quick sampling methods exist, which is key in particle filtering. However, this method is inevitably accompanied by an increase in bias.

The most simple discretization scheme is the Euler-Maruyama scheme. This scheme intuitively differences the SDE, which for the log asset dynamics of the Heston model in equation (3.2) looks like:

$$\log(S(t+\Delta t)) = \log(S(t)) + \left(\mu - \frac{1}{2}V(t)\right)\Delta t + \sqrt{V(t)}\Delta W(t) \tag{3.4}$$

As one can see, this scheme can produce negative values if the draw of $\Delta W(t) \overset{d}{=} \sqrt{\Delta t} Z$, where Z is a standard normal, is negative enough - an undesirable feature. The Milstein scheme outperforms the Euler-Maruyama scheme in accuracy as it considers expansions of both the drift and diffusion terms

coefficients via Itô's Lemma. However, the problem of possible negative values remains. Kahl & Jäckel (2006) developed the IJK scheme, an implicit Milstein scheme, to solve the possibility of negative values. The scheme is efficient, but seems to fail in practice due to the constraint on the parameters that has to be met to prevent negative values. Lord et al. (2010) proposed to use the Full Truncation scheme, which just replaces the value by zero if the new draw would be negative. Although this scheme is heuristic, it works relatively well and outperforms higher order Milstein schemes. However the discretization error remains quite large for a practical number of grid points.

The literature is consistent on the efficiency of the Quadratic-Exponential (QE) scheme developed by Andersen (2007). The QE scheme is based on drift interpolation and moment-matching techniques; the exact distribution is approximated by a similar distribution whose moments are (locally) matched with those of the exact distribution. Even for a practical number of grid points the scheme is very fast and highly accurate. The QE scheme suffers a serious drawback though: it alters the correlation structure between the asset and volatility which is the core of the Heston model. Andersen was aware of this problem and named it the leaking correlation problem.

The second method uses the exact distributions of the model. Per definition this eliminates the discretization bias. Broadie and Kaya (2006) derived the characteristic function of $\int_s^t V(u)du$. As such, the distribution function of $\int_s^t V(u)du$ can be obtained through the Fourier inverse of this characteristic function, which unfortunately has to be done numerically as this inversion cannot be computed analytically. A particle filter based on this method is very slow due to the complicated distribution of the integrated variance process which involves many evaluations of a modified Bessel function of the first kind. Moreover, since the Fourier inversion of the characteristic function of the integrated variance process involves a non-computable integral, this method is in fact already an approximation. Alongside the fact that the method is slow, it can also lead to numerical errors due to truncation.

Glasserman & Kim (2011) proposed another representation of the integrated variance process. They opt to rewrite the process in terms of infinite sums and mixtures of Gamma distributed random variables. Even though they improve on the speed of the simulation method of Broadie & Kaya (2006), their method still entails a computationally costly Fourier inversion. Van Haastrecht and Pelsser (2010) propose a hybrid solution that is highly accurate and fast. They combine the QE method of Andersen (they use the drift interpolation method of Andersen to approximate the stochastic integral $\int_s^t V(u)du$) with the exact transition distribution by using a Poisson representation of the non-central chi-squared distribution, which is slightly slower than the QE scheme due to the sampling of Poisson random numbers. Their scheme, called the NCI scheme, solves the leaking correlation problem and is in terms of overall accuracy and efficiency very comparable.

Due to the fact that the Hybrid model cannot be solved analytically and the aforementioned continuous simulation schemes cannot be applied either, I discuss the Euler-Maruyama discretization scheme. The continuous time solution, and their simulation and estimation schemes are discussed in the Appendix.

3.2.1.1 Euler-Maruyama discretization

For the Heston model, the Euler-Maruyama scheme takes the form:

$$\begin{aligned} \log S\left(t+\Delta t\right) &= \log S\left(t\right) + \left(\mu - \frac{1}{2}V\left(t+\Delta t\right)\right)\Delta t + \sqrt{V\left(t\right)}\Delta W_1(t) \\ V(t+\Delta t) &= V(t) + \kappa\left(\theta - V(t)\right)\Delta t + \sigma\sqrt{V(t)}\left(\rho\Delta W_1(t) + \sqrt{1-\rho^2}\Delta W_2(t)\right) \end{aligned} \tag{3.5}$$

Note that different time indices are used for the variance in the log asset dynamics, following Aihara et al. (2008). Otherwise, the instantaneous link between the log assets and variance dynamics would have

been lost. As such, using the log asset observation as a proxy for the variance at the same point in time would not have been possible. Subsequently, deriving a particle filter would have been problematic or even impossible.

One can readily see that the conditional transition distributions are Gaussian as the distribution of a Wiener increment is Gaussian, e.g. $\Delta W_1(t) \sim \mathcal{N}(0, \Delta t)$. This simplicity is the basis and the power of the Euler-Maruyama scheme.

In the estimation it is useful to rewrite the system by plugging $\Delta W_1(t)$ in the variance dynamics. The resulting equation incorporates of the observable log asset process, which yields valuable information on the unobservable variance process.

$$\begin{aligned} V(t+\Delta t) &= V(t) + \kappa(\theta - V(t))\Delta t + \sigma\rho\left(\log S(t+\Delta t) - \log S(t) - \left(\mu - \frac{1}{2}V(t+\Delta t)\right)\Delta t\right) \\ &\quad + \sigma\sqrt{V(t)}\sqrt{1-\rho^2}\Delta W_2(t) \\ &= \left(1 - \frac{\sigma\rho\Delta t}{2}\right)^{-1}\left[V(t) + \kappa(\theta - V(t) - \sigma\rho\mu)\Delta t + \sigma\rho(\log S(t+\Delta t) - \log S(t))\right] \\ &\quad + \left(1 - \frac{\sigma\rho\Delta t}{2}\right)^{-1}\left(\sigma\sqrt{V(t)}\sqrt{1-\rho^2}\right)\Delta W_2(t) \end{aligned} \tag{3.6}$$

3.2.1.2 Estimation

Per contra auxiliary particle filters where parameters and state are sampled jointly, such as the Liu-West (2001) particle filter, the particle filter used in this thesis generates two different sets of particles. The parameter particles are parallel to the state particles; for every parameter particle, the state particle is updated. As such, the likeliness of every parameter particle is considered per state particle. The particle filter presented is a decretized version of the particle filter developed by Aihara et al. (2012).

The optimal importance density (the original exact distribution), as described in Doucet et al. (2000), can be used as the exact distribution is known and easy to simulate from. From equation (3.6):

$$\begin{aligned} &g(V(t+\Delta t)|\log S(t+\Delta t), \log S(t), V(t)) = \mathcal{N}\left(m, s^2\right) \\ &m = \left(1 - \frac{\sigma\rho\Delta t}{2}\right)^{-1}\left[V(t) + \kappa(\theta - V(t) - \sigma\rho\mu)\Delta t + \sigma\rho(\log S(t+\Delta t) - \log S(t))\right] \\ &s^2 = \frac{\sigma^2 V(t)(1-\rho^2)}{\left(1 - \frac{\sigma\rho\Delta t}{2}\right)^2}\Delta t \end{aligned} \tag{3.7}$$

The observational equation directly follows from equation (3.5):

$$\begin{aligned} &p(\log S(t+\Delta t)|\log S(t), V(t+\Delta t), V(t)) = \mathcal{N}\left(m, s^2\right) \\ &m = \log(St) + \left(\mu - \frac{1}{2}V(t+\Delta t)\right)\Delta t \\ &s^2 = V(t)\Delta t \end{aligned} \tag{3.8}$$

The transition density $p(V(t+\Delta t|\log(S(t)), V(t))$ used by Aihara et al. (2008) is obtained by plugging in the observation equation $\log(S(t+\Delta t))$ into equation (3.6). However, this distribution should be equal to the original distributions stated in equation (3.5) as these two dynamics are mathematically equal. The fact that the transition distribution used in the particle filter of Aihara et al. (2008) does

not include $\log(S(t))$, just as the original transition distribution, supports this:

$$\begin{aligned} &p\left(V(t+\Delta t|\log S(t), V(t)\right) = \mathcal{N}\left(m, s^2\right) \\ &m = V(t) + \kappa\left(\theta - V(t)\right)\Delta t \\ &s^2 = \sigma\sqrt{V(t)}\Delta t \end{aligned} \tag{3.9}$$

To start the particle filter one starts by selecting an appropriate amount of state particles N and parameter particles M. The initial state particles, $\{V_j(0)\}_{j=1}^{N}$, are drawn from a Gaussian distribution, and the parameter particles, $\{\psi_j(0)\}_{j=1}^{M}$, are drawn from an uniform distribution. The mean μ_0 and variance σ_0^2 of the Gaussian distribution, and the lower lb_ψ and upper bounds ub_ψ of the uniform distribution need to be set. Set the initial weights - read likelihoods - of the state particles to $\omega_j^V(0) = \frac{1}{N}, \forall j = 1, 2, \ldots, N$, and of the parameter particles to $\omega_j^\psi(0) = \frac{1}{M}, \forall j = 1, 2, \ldots, M$. Negative initial state particle values need to be excluded. Two possibilities are to resample until no negative values are left, or to replace the negatives with small numbers. As I opt to replace negative values with small values, the simulation scheme is a Full Truncation scheme by Lord et al. (2010). To stimulate parameter diversity, resampling takes place if the filter reached time point τ. Moreover, a vector ϵ is multiplied with the variance of the parameter vector to mitigate the degeneracy problem. Formally, the particle filter is dictated by:

Algorithm 1 Parallel Particle Filter Discretized Heston Model

Initialize. Sample $\{(V_j(0))\}_{j=1}^N \sim \mathcal{N}\left(\mu_0, \sigma_0^2\right)$ and set $\omega_j^V(0) = \frac{1}{N}, \forall j = 1, 2, \ldots, N$.
Sample $\{(\psi_m(0))\}_{m=1}^M \sim U(lb_\psi, ub_\psi)$ and set $\omega_m^\psi(0) = \frac{1}{M}, \forall m = 1, 2, \ldots, M$.
Initialize ϵ

for $t = 1, 2, \ldots T$ **do**

for $m = 1, 2, \ldots, M$ **do**

1. $\forall j = 1, 2, \ldots N$, sample $V_j(t) \sim g\left(V(t)|V_j(t-1), \psi_m(t), \log S(t), \log S(t-1)\right)$

2. $\forall j = 1, 2, \ldots N$, calculate particle weights
$\tilde{\omega}_j^V(t) = \omega_j^V(t-1)\frac{p(\log S(t)|V_j(t), \psi_m(t), \log S(t-1), V_j(t-1)) \times p(V_j(t)|\psi_m(t), \log S(t-1), V_j(t-1))}{g\left(V_j(t)|\psi_m(t), \log S(t), \log S(t-1), V_j(t-1)\right)}$

3. Normalize $\omega_j^V(t) = \frac{\tilde{\omega}_j^V(t)}{\sum_j \tilde{\omega}_j^V(t)}$

4. Resample if $\left(\sum_j \omega_j^V(t)^2\right)^{-1} < \frac{2}{3}N$
Resample $\{(V_j(t))\}_{j=1}^N$ with probabilities $\left\{\omega_j^V(t)\right\}_{j=1}^N$. Set $\omega_j^V(t) = \frac{1}{N}, \forall j = 1, 2, \ldots N$.

5. Calculate $\tilde{\omega}_m^\psi(t) = \omega_m^\psi(t-1)\sum_j \omega_j^V(t) p(\log S(t)|V_j(t), \psi_m(t), \log S(t-1), V_j(t-1))$

6. Normalize $\omega_m^\psi(t) = \frac{\tilde{\omega}_m^\psi(t)}{\sum_m \tilde{\omega}_m^\psi(t)}$

7. Inference $\hat{V}(t) = \sum_m \omega_m^\psi(t) \sum_j V_j(t)\omega_j^V(t)$ ▷ Note: $\sum_j V_j(t)\omega_j^V(t) = f(\psi_m(t))$

8. Inference $\hat{\psi}(t) = \sum_m \omega_m^\psi(t)\psi_m(t)$

9. Resample if $\left(\left(\sum_j \omega_j^V(t)^2\right)^{-1} < \frac{2}{3}N\right) \wedge (t \geq \tau)$
Calculate $\hat{\Sigma} = \sum_m \omega_m^\psi(t)\left(\psi_m(t) - \hat{\psi}(t)\right)\left(\psi_m(t) - \hat{\psi}(t)\right)^\top$
If not first resample, set $lb_{\psi_p} = \max\left[lb_{\psi_p}, \hat{\psi}_p(t) - 3\hat{\Sigma}_{p,p}\right]$ and $ub_{\psi_p} = \min_m\left[ub_{\psi_p}, \hat{\psi}_p(t) + 3\hat{\Sigma}_{p,p}\right]$
$\forall p = 1, 2, \ldots, |\psi|$ to counter degeneracy
Create candidates $c(i) = lb_\psi + \frac{ub_\psi - lb_\psi}{M-1}(i-1), i = 1, 2, \ldots M$ ▷ Create $|\psi| \times M$ matrix
Calculate for each parameter $p(c_p(i)) = \frac{\phi\left(\frac{c_p(i) - \hat{\psi}_p(t)}{\epsilon_p \hat{\sigma}_p}\right)}{\Phi\left(\frac{ub_{\psi_p} - \hat{\psi}_p(t)}{\hat{\sigma}_p}\right) - \Phi\left(\frac{lb_{\psi_p} - \hat{\psi}_p(t)}{\hat{\sigma}_p}\right)}, \forall i = 1, 2, \ldots M,$
$\forall p = 1, 2, \ldots, |\psi|$
Resample each parameter independently $\{\psi_m^p(t)\}_{m=1}^M$ from $\{c_p(i)\}_{i=1}^M$ with probabilities $p(c_p(i))$
Set $\omega_m^\psi(t) = \frac{1}{M}, \forall m = 1, 2, \ldots M$

end for

end for

3.2.1.3 Simulation

Simulating the Heston model follows directly from equation (3.5). Given the time increment Δt, the end point T, the initial asset price S_0, the initial variance V_0, and the parameters ψ, the simulating algorithm is given by:

Algorithm 2 Simulation Discretized Heston Model

for $t = 1, 2, \ldots, T$ **do**

Draw $(Z_1(t), Z_2(t))^\top \sim \mathcal{N}\left((0,0)^\top, \begin{bmatrix} 1,0 \\ 0,1 \end{bmatrix}\right)$

$$V(t) = V(t-1) + \kappa\left(\theta - V(t-1)\right)\Delta t + \sigma\sqrt{V(t-1)}\sqrt{\Delta t}\left(\rho Z_1(t) + \sqrt{1-\rho^2}Z_2(t)\right)$$

$$\log S(t) = \log S\left(t\right) + \left(\mu - \tfrac{1}{2}V\left(t\right)\right)\Delta t + \sqrt{V\left(t\right)}\sqrt{\Delta t}Z_1(t)$$

end for

3.2.2 Term-Structure: Hull-White Model

This section starts with describing the most widely-used term-structure models, which are models that describe the progression of the whole curve of zero rates (the term-structure). As can be seen in Section 9.1.1.1 in the Appendix, Equation (9.7), the zero rate can be obtained from short rates. This relation enables the term-structure to be modelled solely by the short rate. If a process for the short rate is defined, these models are able to determine the initial term-structure and its evolution through time.

Model Selection: Equilibrium versus No-Arbitrage Models

Within the class of term-structure models, there exist two different types of models. Models of the first type are called equilibrium models. These models are based on economic assumptions about interest rates and typically can not perfectly replicate the initial term-structure. In contrast, models of the second type are meant to capture the initial term-structure and are called no-arbitrage models. The main distinction between the two models is that the initial term-structure is an output in equilibrium models and an input in no-arbitrage models. Furthermore, whereas the drift term in the dynamics of the short rate is constant in equilibrium models, this term is a deterministic function of time (a $\mathcal{F}(t)$-adapted process, see Section 9.1.2 in the Appendix) no-arbitrage models. This is due to the so-called expectations hypothesis, which postulates that long-term interest rates are fully explained by current and future expected short rates. In other words, the shape of the initial term-structure explains the direction of the future short rate in a no-arbitrage model. This proposition has been debated, see Froot (1989) and Sarno et al. (2007), but reaffirmed by Guidolin and Thornton (2008).

For the aforementioned reasons, I opt to exclude equilibrium models - such as the Vasicek model (1977), Rendleman & Bartter (1980) model, Brennan & Schwartz model (1980), Courtadon model (1982), the Cox-Ingersoll-Ross model (1985), and the Longstaff model (1989) - in my analysis and solely focus on no-arbitrage models. Moreover, the Cox-Ingersoll-Ross model and the Longstaff model are extensions of the Vasicek model in the sense that it uses the same drift term, but adapts the diffusion term (includes the square root of the short rate) such that the short rate cannot become negative. This property of these models used to make them very popular amongst practitioners. However, in the episode of QE2 the world has witnessed that interest rates can cross the Zero Lower Bound (ZLB) and can become negative. Therefore, the models that can not provide negative rates are not realistic.

The Ho-Lee (1986) model was the first no-arbitrage term-structure model. This model belongs to the class of the Affine Term-Structure (ATS) models, see Dai & Singleton (2000), which means that the price of a zero coupon bond can be expressed as an affine function (linear plus a constant) of the short rate. The Ho-Lee model does not provide a mean-reverting process. Not only is this property desirable for tractability, but since then it was believed that interest rates are pulled back to their long-run average level. The reason is based on a supply-and-demand argument; when rates are high, the economy tends to stagnate due to inter-temporal substitution (people are more willing to consume later) and due to low demand for funds from borrowers (since it is expensive to borrow against such rates). As a result, rates decline. When rates are low, the opposite applies and rates tend to rise. Another argument is that a central bank may hold onto a stable interest rate policy.

The Hull-White model (1990) is a no-arbitrage model that possesses mean-reversion and is a very popular model among practitioners due to its tractability. The ability to perfectly incorporate the initial term-structure is the extension that Hull and White provided in comparison with the Vasicek model. Also, the Hull-White model is an extension of the Ho-Lee model in the sense that it includes mean-reversion. The Hull-White model regained popularity due to the ability to produce negative interest rates. Occurrence of negative rates was perceived as something impossible due to the ZLB, but has been proven to be possible.

Due to the perception that interest rates could not become negative, several extensions have been made to the Hull-White model. The model of Black–Derman–Toy (1990) is based on the Hull-White model in the sense that it has the same dynamics, but applied to the logarithm of the short rate. The model assumes a direct relation between the mean-reversion parameter and the volatility parameter, which both are a function of time. The relation is such that the process only reverts to its mean if the volatility decreases over time. As most practitioners use a constant volatility, the mean-reversion parameter is zero. Hence, the model collapses to a log-normal version of the Ho-Lee model.

To fix this problem, Black & Karinski (1991) altered the Black-Derman-Toy model and dropped the direct relation between the mean-reversion and the volatility. The drawback of no analytical properties and the absence of negative rates remains. Moreover, Hogan and Weintraub (1993) pointed out that the Black-Derman-Toy model and the Black-Karinski model could provide exploding short rates. Therefore, Sandmann & Sondermann (1993) developed a model to circumvent this problem. However, this model could not replicate the initial term-structure. The model of Miltersen, Sandman & Sondermann (1997) provided a solution to this problem, but still abstracted from negative rates.

Table 1 below summarizes the properties of the no-arbitrage models discussed above. The model that is used in this thesis is the Hull-White model.

Table 1: Properties of No-Arbitrage Term-Structure Models

Model	**Initial TS matched**	**Mean reversion**	**Negative rates**	**Stable rates**
Ho-Lee (1986)	+	-	+	+
Hull-White (1990)	+	+	+	+
Black-Derman-Toy (1990)	+	+/-	-	-
Black-Karinski (1990)	+	+	-	-
Sandman & Sondermann (1993)	-	+	+	+
Sandman & Sondermann (1997)	+	+	-	+

Dynamics
As the Hull-White model has an explicit Gaussian solution, there is no role for discretization of the model for computational gains. The model describes the dynamics of the short rate r as follows:

$$dr(t) = (\Theta(t) - \kappa r(t))\, dt + \sigma dW(t) \tag{3.10}$$

Equation (3.10) can be seen as a mean reverting Ornstein–Uhlenbeck process with a time dependent level. In this equation κ determines the speed with which the short rate $r(t)$ reverts to the time dependent level $\Theta(t)/\kappa$. The parameter σ determines the volatility of the short rate and $W(t)$ is a Wiener process with respect to measure $\mathbb{Q}$. Moreover the function $\Theta(t)$ is chosen in such a way that the model can replicate the term-structure observed in the market. As can be seen from equation (3.10) the volatility in the model stems from one Wiener process. Therefore, the model is accounted amongst the one-factor models and is able to explain one factor of market risk. The propagation of the rate process is mainly determined by horizontal shifts and tilts of the term-structure.

In order to estimate the parameters and to obtain simulations for the term-structure of interest rates, I make use of the state space formulation of the model. To put the model in a state space formulation, one needs to solve the stochastic differential equation and obtain the closed form solution for the zero-coupon bonds prices (which have a one-to-one relation with the zero-rates) dictated by the model. First, I turn to the solution of the stochastic differential equation. Then, I provide the closed form solution of the price of a zero-coupon. In both derivations I follow the approach by Rom (2013).

The assumption of a single factor is not as restrictive as it might appear. A one-factor model implies that all rates move in the same direction over any short time interval, but not that they all move by the same amount. The shape of the zero curve can therefore change with the passage of time.

3.2.2.1 Affine Term-Structure & Solution of the Stochastic Differential Equation

As the Hull-White model is an ATS model the zero rates can be expresses as a function of the short rate plus a constant. More specifically, the price of a bond (with face value €1,-) can be formulated as:

$$P(t,T) = e^{\alpha(t,T)+\beta(t,T)r(t)} \tag{3.11}$$

Using (3.11), and the relation between the discount factor and the spot rate, see equation (9.13):

$$R(t,T) = -\frac{\log(P(t,T))}{T-t} = -\frac{\alpha(t,T)}{T-t} - \frac{\beta(t,T)}{T-t} r(t) \tag{3.12}$$

which is an affine function in the short rate. The expressions of $\beta(t,T)$ and $\alpha(t,T)$ are given by:

$$\begin{aligned} \beta(t,T) &= \frac{1}{\kappa}\left(e^{-\kappa(T-t)} - 1\right) \\ \alpha(t,T) &= -\beta(t,T) f(0,t) + \log\left(\frac{P(0,T)}{P(0,t)}\right) + \frac{\sigma^2}{4\kappa}\beta^2(t,T)\left(e^{-2\kappa t} - 1\right) \end{aligned} \tag{3.13}$$

Proof: See Section 9.2.1 in the Appendix.

The solution of the stochastic differential equation 3.10 is given by:

$$r(t)|r(s) \sim \mathcal{N}\left(e^{-\kappa(t-s)}\left(r(s) - \mu(s)\right) + \mu(t), \frac{\sigma^2}{2\kappa}\left(1 - e^{-2\kappa(t-s)}\right)\right) \tag{3.14}$$

with $\mu(t) = f(0,t) + \frac{\sigma^2}{2\kappa^2}\left(1 - e^{-\kappa t}\right)^2$.

Proof: See Section 9.2.2 in the Appendix.

3.2.2.2 State Space Formulation

The Hull-White model can be written in a linear Gaussian state space model. This is due to the fact that the model belongs to the class of affine term-structure models and the conditional distribution of the short rate, the state, is Gaussian. The observations in the state space formulation are the zero rates observed in the market. Assume that the observations are equidistant in time with a time difference of Δt. The system matrices, as in Eq. (9.28) in the Appendix, follow from equations (3.12) and (3.14). More specifically, when one uses p points of the term-structure per point in time, the Hull-White model can be represented in state space form as follows:

$$\begin{aligned}
y(t) &= (R_1(t), R_2(t), \ldots, R_p(t))^\top \\
\gamma(t) &= r(t),\ \gamma(\cdot) \text{ represents state variable } (\alpha(\cdot) \text{ taken}) \\
c(t) &= \left(-\frac{\alpha(t,T_1)}{T_1 - t}, -\frac{\alpha(t,T_2)}{T_2 - t}, \ldots, -\frac{\alpha(t,T_p)}{T_p - t} \right)^\top \\
Z(t) &= \left(-\frac{\beta(t,T_1)}{T_1 - t}, -\frac{\beta(t,T_2)}{T_2 - t}, \ldots, -\frac{\beta(t,T_p)}{T_p - t} \right)^\top \\
H(t) &= \begin{bmatrix} 0 & \ldots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \ldots & 0 \end{bmatrix} \\
d(t) &= \mu(t + \Delta t) - e^{-\kappa \Delta t} \mu(t) \\
T(t) &= e^{-\kappa \Delta t} \\
R(t) &= 1 \\
Q(t) &= \frac{\sigma^2}{2\kappa} \left(1 - e^{-2\kappa \Delta t}\right)
\end{aligned} \tag{3.15}$$

Since the model has constant parameters the process is stationary, see Hull & White (1996). Therefore, $\alpha_{t_1} \sim \mathcal{N}(a_{t_1}, P_{t_1})$, with a_{t_1}, P_{t_1} the unconditional mean and unconditional variance, respectively. In other words, the long-term mean and variance of the state are used as mean and variance for the initial state distribution. One obtains the long-term mean and variance by taking the limit of the mean and variance conditional on the initial state. Mathematically:

$$a_{t_1} = \lim_{t \to \infty} \mathbb{E}[r(t)|r(0)] = \lim_{t \to \infty} e^{-\kappa t} \left(r\left(0\right) - \mu(0) \right) + \mu(t) = \lim_{t \to \infty} f(0,t) + \frac{\sigma^2}{2\kappa^2} \left(1 - e^{-\kappa t}\right)^2 = \lim_{t \to \infty} f(0,t)$$

Substituting $f(0,t) = -\partial \log\left(P(0,t)\right) / \partial t$ in order to express the forward rate into model parameters is not useful, as the resulting expression involves the short rate. Since both the long-term forward and short rate are unknown, I opt to include a_{t_1} as a parameter in my estimation. An expression for long-term variance can be expressed as:

$$P_{t_1} = \lim_{t \to \infty} \mathbb{V}ar[r(t)|r(0)] = \lim_{t \to \infty} \frac{\sigma^2}{2\kappa} \left(1 - e^{-2\kappa t}\right) = \frac{\sigma^2}{2\kappa}$$

3.2.2.3 Estimation

In contrast with the parameter estimation approach of Hull & White (1990), who use derivatives, I use historical term-structures. When one calibrates the parameters to the price of derivatives observed in the market, one obtains parameter estimates of the short rate process under the risk-free measure $\mathbb{Q}$. This measure is useful for the pricing of derivatives and other market instruments, but not for scenario analysis. In scenario analysis the real world measure $\mathbb{P}$ is needed to identify the risks of the portfolio.

Hence, the model is estimated on historical term-structures.

Although the conditional distribution of the short rate given a previous observation of the short rate is known, maximum likelihood estimation of the parameters is not very useful. This is due to the fact that the short rate is not observed, and hence should be estimated first. In earlier work, such as Chan et al. (1992) or Nowman (1996), short-term interest rates (e.g. 1 month maturity) were used as a proxy for the short rate. The approaches of Brown & Dybvig (1986) and De Munnik & Schotman (1994) were based on a cross-sectional analysis (incorporating different maturity levels in the estimation). However, all these approaches rested on the assumption that the short-term interest rate is able to proxy for the short rate. In more recent work, application of the Kalman filter has gained popularity due to it's ability to extract the unobservable process of the short rate, see e.g. Babbs & Nowman (1999) and the references therein. The empirical work by Sapp (2009) provided evidence that the Kalman filter outperformed the earlier estimation methods in terms of estimation bias.

As the system matrices are defined,the Kalman filter can directly be applied to historical term-structures.

3.2.2.4 Simulation

Simulation of the Hull-White model is very simple because of the state space formulation of the model. The algorithm to sample a path of yield curves of length T with time increment Δt is given below.

Algorithm 3 Simulation Hull-White Model

for $t = 1, 2, \ldots, T$ **do**

Draw $Z(t) \sim \mathcal{N}(0, 1)$

$\gamma(t) = d(t) + T(t)\gamma(t-1) + Q(t)Z(t)$

$y(t) = c(t) + Z(t)\gamma(t)$

end for

3.2.3 Heston-Hull-White Model

The hybrid model combines the short rate dynamics of the Hull-White model and the stock dynamics of the Heston model. Due to more complicated correlation structure of the processes, the transition densities become very different. The Heston-Hull-White model is dictated by:

$$\begin{aligned} dr(t) &= \kappa_r(\Theta(t) - r(t))\,dt + \sigma_r dW_r(t) \\ d\log(S(t)) &= \left(\mu - \frac{1}{2}V(t)\right)dt + \sqrt{V(t)}dW_S(t) \\ dV(t) &= \kappa_V(\theta - V(t))\,dt + \sigma_V\sqrt{V(t)}dW_V(t) \end{aligned} \tag{3.16}$$

where the Wiener processes are correlated by covariance matrix $\Sigma = \begin{bmatrix} 1 & \rho_{12} & \rho_{13} \\ \rho_{12} & 1 & \rho_{23} \\ \rho_{13} & \rho_{23} & 1 \end{bmatrix}$. A Cholesky decomposition on $\Sigma = LL'$, yields $L = \begin{bmatrix} 1 & 0 & 0 \\ \rho_{12} & \sqrt{1-\rho_{12}^2} & 0 \\ \rho_{13} & \frac{\rho_{23}-\rho_{13}\rho_{12}}{\sqrt{1-\rho_{12}^2}} & \sqrt{1-\rho_{13}^2-\frac{(\rho_{23}-\rho_{13}\rho_{12})^2}{1-\rho_{12}^2}} \end{bmatrix} \equiv \begin{bmatrix} l_{11} & 0 & 0 \\ l_{21} & l_{22} & 0 \\ l_{31} & l_{32} & l_{33} \end{bmatrix}$.

Then, the model can be written as:

$$\begin{aligned} dr(t) &= \kappa_r(\Theta(t) - r(t))\,dt + \sigma_r dW_1(t) \\ d\log(S(t)) &= \left(\mu - \frac{1}{2}V(t)\right)dt + \sqrt{V(t)}\,(l_{21}dW_1(t) + l_{22}dW_2(t)) \\ dV(t) &= \kappa_V(\theta - V(t))\,dt + \sigma_V\sqrt{V(t)}\,(l_{31}dW_1(t) + l_{32}dW_2(t) + l_{33}dW_3(t)) \end{aligned} \tag{3.17}$$

where $W_1(t) \perp\!\!\!\perp W_2(t) \perp\!\!\!\perp W_3(t), \forall t$. By iterated substitution one obtains:

$$\begin{aligned} dr(t) &= \kappa_r(\Theta(t) - r(t))\,dt + \sigma_r dW_1(t) \\ d\log(S(t)) &= \left(\mu - \frac{1}{2}V(t) - \frac{l_{21}}{\sigma_r}\kappa_r(\Theta(t) - r(t))\sqrt{V(t)}\right)dt + \frac{l_{21}}{\sigma_r}\sqrt{V(t)}dr(t) + l_{22}\sqrt{V(t)}dW_2(t) \\ dV(t) &= \left(\kappa_V\theta - \frac{l_{32}\mu\sigma_V}{l_{22}} - \left(\kappa_V - \frac{l_{32}\sigma_V}{2l_{22}}\right)V(t) - \left(l_{31} - \frac{l_{21}l_{32}}{l_{22}}\right)\frac{\sigma_V}{\sigma_r}\kappa_r(\Theta(t) - r(t))\sqrt{V(t)}\right)dt \\ &\quad + \left(l_{31} - \frac{l_{21}l_{32}}{l_{22}}\right)\frac{\sigma_V\sqrt{V(t)}}{\sigma_r}dr(t) + \frac{l_{32}\sigma_V}{l_{22}}d\log(S(t)) + l_{33}\sigma_V\sqrt{V(t)}dW_3(t) \end{aligned} \tag{3.18}$$

Mathematically speaking, this decomposition does not alter the interpretation of the model as it is defined in continuous time. However, at any lower frequency the decomposition has some economic meaning. By applying the decomposition in this order, one assumes that the stock and variance processes have no contemporaneous effect on the short rate dynamics. The short rate dynamics affect both the stock and variance process directly. The stock process affects the variance process contemporaneously, but not vice versa. This decomposition order can be justified by the fact that interest rates do not respond directly to a sudden change in the equity markets, but equity markets respond directly to a change in interest rates (e.g. due to change in monetary policy). Since the system is decoupled into a regular Hull-White model and an adapted Heston model, the estimation of the models is similar to the earlier approaches. As the short rate dynamics are independent of the log asset and variance dynamics, the parameters can be obtained by applying the Kalman filter as described in the previous section. The dynamics in the Heston model include extra terms that give information on the processes, and hence alter the transition distributions. I present the discretized version of the model, and why a continuous solution is not feasible at this stage.

3.2.3.1 Discrete

Discretization of the system (3.17) yields:

$$\begin{aligned} r(t+\Delta t) &= r(t) + \kappa_r(\Theta(t) - r(t))\,\Delta t + \sigma_r\Delta W_1(t) \\ \log(S(t+\Delta t)) &= \log(S(t)) + \left(\mu - \frac{1}{2}V(t+\Delta t)\right)\Delta t + \sqrt{V(t)}\,(l_{21}\Delta W_1(t) + l_{22}\Delta W_2(t)) \\ V(t+\Delta t) &= V(t) + \kappa_V(\theta - V(t))\,\Delta t + \sigma_V\sqrt{V(t)\Delta t}\,(l_{31}\Delta W_1(t) + l_{32}\Delta W_2(t) + l_{33}\Delta W_3(t)) \end{aligned} \tag{3.19}$$

Isolating $\Delta W_1(t)$ and $\Delta W_2(t)$ gives:

$$\begin{aligned} \Delta W_1(t) &= \frac{r(t+\Delta t) - r(t) - \kappa_r(\Theta(t) - r(t))\,\Delta t}{\sigma_r\sqrt{\Delta t}} \\ \Delta W_2(t) &= \frac{\log(S(t+\Delta t)) - \log(S(t)) - \left(\mu - \frac{1}{2}V(t+\Delta t)\right)\Delta t}{l_{22}\sqrt{V(t)\Delta t}} - \frac{l_{21}\Delta W_1(t)}{l_{22}} \\ &= \frac{\log(S(t+\Delta t)) - \log(S(t)) - \left(\mu - \frac{1}{2}V(t+\Delta t)\right)\Delta t}{l_{22}\sqrt{V(t)\Delta t}} - \frac{l_{21}}{l_{22}}\left(\frac{r(t+\Delta t) - r(t) - \kappa_r(\Theta(t) - r(t))\,\Delta t}{\sigma_r\sqrt{\Delta t}}\right) \end{aligned} \tag{3.20}$$

Plugging $\Delta W_1(t)$ and $\Delta W_2(t)$ into (3.19) yields:

$$\begin{aligned}
r\left(t+\Delta t\right) &= r\left(t\right)+\kappa_r\left(\Theta\left(t\right)-r\left(t\right)\right)\Delta t+\sigma_r\Delta W_1\left(t\right)\\
\log\left(S\left(t+\Delta t\right)\right) &= \log\left(S\left(t\right)\right)+\left(\mu-\frac{1}{2}V(t)-\frac{l_{21}\sqrt{V(t)}}{\sigma_r}\kappa_r\left(\Theta\left(t\right)-r\left(t\right)\right)\right)\Delta t+\frac{l_{21}\sqrt{V(t)}}{\sigma_r}\left(r\left(t+\Delta t\right)-r\left(t\right)\right)\\
&\quad +l_{22}\sqrt{V(t)}\Delta W_2\left(t\right)\\
V(t+\Delta t) &= V\left(t\right)+\left(\theta-\frac{l_{32}\mu\sigma_V}{l_{22}}-\left(\kappa_V-\frac{l_{32}\sigma_V}{2l_{22}}\right)V(t)-\left(l_{31}-\frac{l_{21}l_{32}}{l_{22}}\right)\frac{\sigma_V\sqrt{V(t)}}{\sigma_r}\kappa_r\left(\Theta\left(t\right)-r\left(t\right)\right)\right)\Delta t\\
&\quad +\left(l_{31}-\frac{l_{21}l_{32}}{l_{22}}\right)\frac{\sigma_V\sqrt{V(t)}}{\sigma_r}\left(r\left(t+\Delta t\right)-r\left(t\right)\right)+\frac{l_{32}\sigma_V}{l_{22}}\left(\log\left(S\left(t+\Delta t\right)\right)-\log\left(S\left(t\right)\right)\right)\\
&\quad +l_{33}\sigma_V\sqrt{V(t)}\Delta W_3\left(t\right)
\end{aligned} \tag{3.21}$$

From this system, the densities that are needed for the particle filter are easily obtained. In principle, one could estimate the parameters of the short rate alongside the parameters of the stochastic volatility model. As the Cholesky decomposition enables one to use the Kalman filter, the estimated state and parameters can be used as the basis of the second stage of the estimation procedure:

$$\begin{aligned}
\log\left(S\left(t+\Delta t\right)\right) &= \log\left(S\left(t\right)\right)+\left(\mu-\frac{1}{2}V(t)-\frac{l_{21}\sqrt{V(t)}}{\hat{\sigma}_r}\hat{\kappa}_r\left(\hat{\Theta}\left(t\right)-\hat{r}\left(t\right)\right)\right)\Delta t+\frac{l_{21}\sqrt{V(t)}}{\hat{\sigma}_r}\left(\hat{r}\left(t+\Delta t\right)-\hat{r}\left(t\right)\right)\\
&\quad +l_{22}\sqrt{V(t)}\Delta W_2\left(t\right)\\
V(t+\Delta t) &= V\left(t\right)+\left(\theta-\frac{l_{32}\mu\sigma_V}{l_{22}}-\left(\kappa_V-\frac{l_{32}\sigma_V}{2l_{22}}\right)V(t)-\left(l_{31}-\frac{l_{21}l_{32}}{l_{22}}\right)\frac{\sigma_V\sqrt{V(t)}}{\hat{\sigma}_r}\hat{\kappa}_r\left(\hat{\Theta}\left(t\right)-\hat{r}\left(t\right)\right)\right)\Delta t\\
&\quad +\left(l_{31}-\frac{l_{21}l_{32}}{l_{22}}\right)\frac{\sigma_V\sqrt{V(t)}}{\hat{\sigma}_r}\left(\hat{r}\left(t+\Delta t\right)-\hat{r}\left(t\right)\right)+\frac{l_{32}\sigma_V}{l_{22}}\left(\log\left(S\left(t+\Delta t\right)\right)-\log\left(S\left(t\right)\right)\right)\\
&\quad +l_{33}\sigma_V\sqrt{V(t)}\Delta W_3\left(t\right)
\end{aligned} \tag{3.22}$$

From this system of equations the densities used in the particle filter are apparent:

$$p\left(\log\left(S\left(t+\Delta t\right)\right)\mid\log\left(S\left(t\right)\right),V\left(t+\Delta t\right),\hat{r}\left(t+\Delta t\right),\hat{r}\left(t\right)\right)\sim\mathcal{N}\left(\mu_S,l_{22}^2V(t)\Delta t\right)\text{ ,where}$$

$$\mu_S=\log\left(S\left(t\right)\right)+\left(\mu-\frac{1}{2}V(t)-\frac{l_{21}\sqrt{V(t)}}{\hat{\sigma}_r}\hat{\kappa}_r\left(\hat{\Theta}\left(t\right)-\hat{r}\left(t\right)\right)\right)\Delta t+\frac{l_{21}\sqrt{V(t)}}{\hat{\sigma}_r}\left(\hat{r}\left(t+\Delta t\right)-\hat{r}\left(t\right)\right)$$

$$p\left(V\left(t+\Delta t\right)\mid\log\left(S\left(t+\Delta t\right)\right),\log\left(S\left(t\right)\right),V(t),\hat{r}\left(t+\Delta t\right),\hat{r}\left(t\right)\right)\sim\mathcal{N}\left(\mu_V,\left(l_{33}\sigma_V\right)^2V(t)\Delta t\right)\text{ ,where}$$

$$\begin{aligned}
\mu_V &= V\left(t\right)+\left(\theta-\frac{l_{32}\mu\sigma_V}{l_{22}}-\left(\kappa_V-\frac{l_{32}\sigma_V}{2l_{22}}\right)V(t)-\left(l_{31}-\frac{l_{21}l_{32}}{l_{22}}\right)\frac{\sigma_V\sqrt{V(t)}}{\hat{\sigma}_r}\hat{\kappa}_r\left(\hat{\Theta}\left(t\right)-\hat{r}\left(t\right)\right)\right)\Delta t\\
&\quad +\left(l_{31}-\frac{l_{21}l_{32}}{l_{22}}\right)\frac{\sigma_V\sqrt{V(t)}}{\hat{\sigma}_r}\left(\hat{r}\left(t+\Delta t\right)-\hat{r}\left(t\right)\right)+\frac{l_{32}\sigma_V}{l_{22}}\left(\log\left(S\left(t+\Delta t\right)\right)-\log\left(S\left(t\right)\right)\right)
\end{aligned} \tag{3.23}$$

The only density that still is required is $p\left(V\left(t+\Delta t\right) \mid \log\left(S\left(t\right)\right), V(t), \hat{r}\left(t+\Delta t\right), \hat{r}\left(t\right)\right)$, which is obtained by replacing the first equation in the system of (3.22) into the second equation:

$$\begin{aligned} V(t+\Delta t) = V\left(t\right) + \left(\theta - \kappa_V V(t) - \frac{l_{31}\sigma_V\sqrt{V(t)}}{\hat{\sigma}_r}\hat{\kappa}_r\left(\hat{\Theta}\left(t\right) - \hat{r}\left(t\right)\right) + \frac{l_{32}\sigma_V}{l_{22}}\log\left(S\left(t\right)\right)\right)\Delta t \\ + \frac{l_{31}\sigma_V\sqrt{V(t)}}{\hat{\sigma}_r}\left(\hat{r}\left(t+\Delta t\right) - \hat{r}\left(t\right)\right) + \sigma_V\sqrt{V(t)}\left(l_{32}\Delta W_2\left(t\right) + l_{33}\Delta W_3\left(t\right)\right) \end{aligned} \tag{3.24}$$

Since $W_2(t)$ and $W_3(t)$ are independent $\forall t$:

$$\begin{aligned} p\left(V\left(t+\Delta t\right) \mid \log\left(S\left(t\right)\right), V(t), \hat{r}\left(t+\Delta t\right), \hat{r}\left(t\right)\right) \sim \mathcal{N}\left(c, \sigma_V^2 V(t)\left(l_{32}^2 + l_{33}^2\right)\Delta t\right) \text{ ,where} \\ c = V\left(t\right) + \left(\theta - \kappa_V V(t) - \frac{l_{31}\sigma_V\sqrt{V(t)}}{\hat{\sigma}_r}\hat{\kappa}_r\left(\hat{\Theta}\left(t\right) - \hat{r}\left(t\right)\right) + \frac{l_{32}\sigma_V}{l_{22}}\log\left(S\left(t\right)\right)\right)\Delta t \\ + \frac{l_{31}\sigma_V\sqrt{V(t)}}{\hat{\sigma}_r}\left(\hat{r}\left(t+\Delta t\right) - \hat{r}\left(t\right)\right) \end{aligned} \tag{3.25}$$

Note that $l_{32}^2 + l_{33}^2$ simplifies to $1 - \rho_{13}^2$. In other words, the correlation between the Wiener processes of the stock and variance dynamics do not contribute to the variance of the conditional transition distribution of the variance process. Hence, the adapted Heston dynamics parameters can be estimated by the earlier stated particle filter adapted for the correct conditional transition distributions.

3.2.3.2 The Continuous Time Problem

The hybrid model decoupled in this particular Cholesky ordering creates a difficult distribution for the volatility dynamics. The process looks like a squared Ornstein-Uhlenbeck process. If so, the this transition distribution would become a chi-squared distribution. This would make sense as its marginal transition distribution is chi-squared as well. However, the resulting dynamics are not a squared Ornstein-Uhlenbeck process exactly. The drift function is different. In the squared Ornstein-Uhlenbeck process the drift has a direct link with the diffusion parameter. Applying these dynamics to the model would be very restrictive. Moreover, it would create indeterminacy in the parameters. Of course this is undesirable. A more formal and elaborate discussion of the restrictiveness of this solution can be found in the Section 9.4 in the Appendix.

Another solution is to solve the conditional transition distribution by the Kolmogorov forward equations (under physicians known as the Fokker-Planck equation). The solution not is straightforward, and I doubt whether it exists.

Of course, one could try to do a useful transformation to get rid of one of the stochastic drift terms. However, as the distance in powers is a half, a useful transformation is hard to find. It might be non-existent even.

A last solution would be by reversing the Choleksy ordering. This seems to be a nice solution as we know the marginal transition distribution of the volatility process. However, this turns out problematic for the reason that not only a $\int_s^t V(u)du$ term pops up in the conditional transition distribution of the short rate, but also a $\int_s^t \sqrt{V(u)}du$ term. The $\int_s^t V(u)du$ term can be sampled by the Broadie & Kaya (2006) approach, the $\int_s^t \sqrt{V(u)}du$ term not. Then one could apply a drift interpolation method. However, this method would have to be applied twice. The continuous time solution method drifts towards a discretization scheme and loses a lot of its unbiasedness.

I discussed the problem of the continuous hybrid model and proposed some directions where the

solution may lie. However, I dedicate this problem to future research and proceed with the discrete hybrid model for now. Although there are undoubtedly effects on the results of the pension fund analysis, these effects are likely to be small. The reason is that discretization bias increases in the step size Δt of the simulation/estimation grid. As our time step is relatively small (especially in the estimation of the model, $\Delta t = 1/252$), the discretization bias is not likely to be that large.

3.3 Pension Fund Model

In this section I present the main model that generates the results. With the simulated financial market from the hybrid model as input, the pension fund model iterates through time constantly complying to its investment strategy. At every point in time the fund interprets the financial market, evaluates its portfolio, calculates its risk, and undertakes transactions accordingly. The transactions are made such that the investment strategy, the equity-to-total-asset-value ratio and hedge ratio, are satisfied under minimal costs. Once a year the fund possibly grants price level compensation. The formula's lying at the basis of the pricing functions can be found in Section Preliminaries in the Appendix.

Suppose the fund is evaluated up to a horizon of T years with time increments Δt year. Then, the time grid is defined by $\mathcal{G} = (\Delta t, 2\Delta t, \ldots, T)$. Denote the set of key pillars of the term-structure by $\mathcal{P} = \{1y, 2y, 3y, 5y, 7y, 10y, 15y, 20y, 30y\}$, such that $|\mathcal{P}| = 9$. Then, we can denote the financial market as a set $\left\{\mathcal{E}, \mathcal{R}^{rf}, \mathcal{R}^{AAA}\right\}$ that contains the time evolution of the equity prices $\mathcal{E}$, the term-structure $\mathcal{R}^{rf}$ of risk-free rates, and of the yield curve of AAA government bonds $\mathcal{R}^{AAA}$. A typical element of the sequence $\mathcal{E}$, of size T, is the scalar $P_{E,t}$ being the price of equity at time t. An element of $\mathcal{R}^{rf}$ is the vector R_t^{rf}, of size $|\mathcal{P}|$, representing the risk-free term-structure characterized by the key pillars. The notation for the AAA yield curve is analogous.

Turning to the instruments in the market, all traded bonds have a face value of FV_{bond}. Without loss of generality, bonds are specified as zero coupon bonds. As coupon paying bonds are a combination of a series of zero coupon bonds, the fund is perfectly able to replicate any coupon bearing bond it desires. Given the term-structure $R_t^{AAA} \in \mathcal{R}^{AAA}$, the price of a bond maturing m years from now is characterized by the function ZCBondPrice:

function ZCBondPrice(R_t^{AAA}, m, FV_{bond})
  DF = DiscFactor(R_t^{AAA});
  $P = FV_{bond} \times DF_m$;
  return P;
end function

Note that the index m does not have to coincide with p. As time progresses in the model, it is perfectly possible that a bond has a time to maturity of two and a half years, which is is not included in the set $\mathcal{P}$.

Swap prices are specified as a fixed-for-floating agreement. If the fund desires a floating-for-fixed contract, the contract is shorted such that wanted position is obtained. To price the swap agreement the vector of fixed rate payments needs to be defined. These payments are a function of the maturity m and the fixed rate of the swap r_{swap}. Swap payments are assumed to be exchanged semi-annually. Then, the fixed cash flow payments of a swap of maturity m is defined by a sequence of length $m/\Delta t$, where at each $\frac{1/\Delta t}{2}$th position the element equals r_{swap} times the notional FV_{swap} of the swap, and zero elsewhere. The last element of the the sequence $CF(m, r_{swap})$ is the sum of the last fixed leg payment $r_{swap} \times FV_{swap}$ and the notional FV_{swap}. Normally, swap agreements are specified as one agreement of a certain notional. Without loss of generality, I specify swap agreements with a fixed notional. The fund can enter several (or partial) identical agreements to obtain the desired hedge position. A more thorough discussion on swap pricing can be found in the Section Preliminaries in the Appendix. Given the term-structure $R_t^{rf} \in \mathcal{R}^{rf}$, the price of a swap of maturity m with contractual predetermined fixed rate r_{swap} and notional FV_{swap} is given by the function SwapPrice:

function SwapPrice($R_t^{rf}, CF(m, r_{swap}), FV_{swap}$) ▷ Fixed-for-floating
DF = DiscFactor(R_t^{rf});
return $DF_{m/\Delta t} \times \underbrace{CF_{m/\Delta t}}_{=FV_{swap}} - DF_{(1:m/\Delta t)}^{\top} CF_{(1:m/\Delta t)}$;
end function

With the help of these pricing functions the interest rate sensitivity of the fixed income securities can be calculated. Denote the interest rate exposure of a bond of maturity m by $\Delta_{B,m}$. Notation for swaps is analogous: $\Delta_{S,m}$. The interest rate sensitivity is specified as the difference in the present value as a consequence of a basis point increase in pillar $p \in \mathcal{P}$. This measure is called the Price Value of a Basis Point (PVBP). The PVBP of a fixed income security $\mathcal{S}$ of maturity m, given the term-structure $R_t^{\mathcal{S}} \in \mathcal{R}^{\mathcal{S}}$, is calculated with the function PVBP:

function PVBP($R_t^{\mathcal{S}}, \mathcal{P}, \mathcal{S}, m$)
$\Delta_{\mathcal{S}} = (0, 0, \ldots, 0)$; where $|\Delta_{\mathcal{S}}| = |\mathcal{P}|$ ▷ initialize delta's as sequence of zero's

for $p \in \mathcal{P}$ **do**

if $\mathcal{S}$ =Bond **then**
$R^* = R^{AAA}$;
$R_p^* = R_p^{AAA} + 0.0001$;
$\Delta_{B,p}$ =ZCBondPrice(R^{AAA}, m, FV_{bond})−ZCBondPrice(R^*, m, FV_{bond});

else if $\mathcal{S}$ =Swap **then**
$R^* = R^{rf}$;
$R_p^* = R_p^{rf} + 0.0001$;
$\Delta_{S,p}$ =SwapPrice($R^{rf}, CF(m, r_{swap}), FV_{swap}$)−SwapPrice($R^*, CF(m, r_{swap}), FV_{swap}$);
end if

end for

return $\Delta_{\mathcal{S}}$;
end function

The PVBP of the pension benefits, or liabilities, $\mathcal{L}$ is calculated in similar fashion. The liabilities are captured in a sequence $\mathcal{L}$, with $|\mathcal{L}| = K$. Every element of the sequence represent the outgoing pension benefits at that particular time. The length of the sequence K typically runs up to about 70 years. The PVPB of the pension benefits are calculated with the following function:

function $\text{PVBP}^{\mathcal{L}}(R_t^{rf}, \mathcal{P}, \mathcal{L})$
  $\Delta_{\mathcal{L}} = (0, 0, \ldots, 0)$; where $|\Delta_{\mathcal{L}}| = |\mathcal{P}|$

  for $p \in \mathcal{P}$ **do**
    $R^* = R_t^{rf}$;
    $R_p^* = R_{t,p}^{rf} + 0.0001$;
    $DF = \text{DiscFactor}(R_t^{rf})$;
    $DF^* = \text{DiscFactor}(R^*)$;
    $\Delta_{B,p} = \sum_k DF_k \mathcal{L}_k - \sum_k DF_k^* \mathcal{L}_k$;
  end for

  return $\Delta_{\mathcal{L}}$;
end function

A rebalance is triggered when the fund does not satisfy its equity-to-total-asset-value ratio, or when the cash reserves are depleted. The optimization problem minimizes the transaction costs to obtain the portfolio that is aligned with the investment strategy. More specifically, the selected portfolio should comply with the fund's hedge ratio h, equity-to-total-asset-value ratio a, and cash reserves. Denote by $\Pi = \{\Pi_E, \Pi_B, \Pi_S, C\}$ the portfolio of the fund, consisting out of the equity, bond, and swap portfolio, and cash, respectively. A typical element of Π contains all information about that specific asset class. For instance, Π_E contains the number of stocks in the portfolio $N_{E,t}$ at time $t \in \mathcal{G}$ at price P_E. The bond portfolio Π_B contains for every bond $b \in \Pi_B$ the amount held $N_{B,b,t}$ at time $t \in \mathcal{G}$, its maturity m_b, and its the interest rate sensitivity $\Delta_{B,b,t}$. The notation for the swap portfolio is analogous. Denote by $\Pi^{mkt} = \left\{\Pi_B^{mkt}, \Pi_S^{mkt}\right\}$, the collection of bond and swap securities available in the market. Note that for equities the product in the portfolio of the fund is identical as the security in the market. However, as for bonds and swap the maturity changes over time the products become heterogeneous. Transaction costs for security $\mathcal{S}$ of maturity $m_{\mathcal{S}}$ are denoted by $\tau_{\mathcal{S},m_{\mathcal{S}}}$. For product $\mathcal{S} \in \Pi^{mkt}$ available in the market, it obviously holds that $N_{\mathcal{S},t-1} = 0$. Then the fund, that implements equity-to-total-asset-value ratio a and hedge ratio h, rebalances its portfolio according to the following optimization problem:

function PORTFOLIO(Π, Π^{mkt}, a, h)

$$\min_{N_{E,t}, N_{B,t}, N_{S,t}} |N_{E,t} - N_{E,t-1}|P_E\tau_E + \sum_{b\in\Pi_B\cup\Pi_B^{mkt}} |N_{B,b,t} - N_{B,b,t-1}|P_{B,b}\tau_{B,m_b} + \sum_{s\in\Pi_S\cup\Pi_S^{mkt}} |N_{S,s,t} - N_{S,s,t-1}|P_{S,s}\tau_{S,m_s}$$

s.t.

1. $$\left|\frac{N_{E,t}P_E}{N_{E,t}P_E + \sum_{b\in\Pi_B\cup\Pi_B^{mkt}} N_{B,b,t}P_{B,b} + \sum_{s\in\Pi_S\cup\Pi_S^{mkt}} N_{S,s,t}P_{S,s}} - a\right| \leq 0.015$$

2. $$\sum_{b\in\Pi_B\cup\Pi_B^{mkt}} N_{B,b,t}\Delta_{B,b} + \sum_{s\in\Pi_S\cup\Pi_S^{mkt}} N_{S,s,t}\Delta_{S,s} \geq h\Delta_{\mathcal{L}}\ ,\ \forall p \in \mathcal{P}$$

3. $$\frac{C_t}{N_{E,t}P_E + \sum_{b\in\Pi_B\cup\Pi_B^{mkt}} N_{B,b,t}P_{B,b} + \sum_{s\in\Pi_S\cup\Pi_S^{mkt}} N_{S,s,t}P_{S,s} + C_t} = 0.01$$

4. $$N_{E,t}P_E + \sum_{b\in\Pi_B\cup\Pi_B^{mkt}} N_{B,b,t}P_{B,b} + \sum_{s\in\Pi_S\cup\Pi_S^{mkt}} N_{S,s,t}P_{S,s} + C_t - T = N_{E,t-1}P_E + \sum_{b\in\Pi_B} N_{B,b,t-1}P_{B,b} + \sum_{s\in\Pi_S} N_{S,s,t-1}P_{S,s} + C_{t-1}$$

   where $T = |N_{E,t} - N_{E,t-1}|P_E\tau_E + \sum_{b\in\Pi_B\cup\Pi_B^{mkt}} |N_{B,b,t} - N_{B,b,t-1}|P_{B,b}\tau_{B,m_b} + \sum_{s\in\Pi_S\cup\Pi_S^{mkt}} |N_{S,s,t} - N_{S,s,t-1}|P_{S,s}\tau_{S,m_s}$

return $\{N_{E,t}, N_{B,t}, N_{S,t}, C_t\}$
end function

The first restriction sets the equity-to-total-asset-value ratio within bounds of the investment strategy. The second equation establishes the appropriate hedge level per pillar of the term-structure. The third equation takes care of the cash reserve. The fourth equation is the budget constraint.

When the fund satisfies its equity-to-total-asset-value ratio and holds enough cash, but the hedge ratio is not met, it would be suboptimal to completely rebalance the portfolio. As swaps have zero net present value initially, the fund can alter its interest risk exposure by entering new swap agreements while preserving the equity-to-total-asset-value ratio. Only the cash reserves are reduced to pay for the transaction/contract costs of the agreements. The fund enters new contracts in the market Π^{mkt} until it satisfies the hedge ratio again:

function ADDSWAPS(Π, Π^{mkt}, h)

$$\min_{N_{S,t}} \sum_{s\in\Pi_S^{mkt}} N_{S,s,t}P_{S,s}\tau_{S,m_s}$$

s.t.

1. $$\sum_{b\in\Pi_B} N_{B,b,t}\Delta_{B,b} + \sum_{s\in\Pi_S\cup\Pi_S^{mkt}} N_{S,s,t}\Delta_{S,s} \geq h\Delta_{\mathcal{L}}\ ,\ \forall p \in \mathcal{P}$$

2. $$C_t - T = C_{t-1}$$

   where $T = \sum_{s\in\Pi_S^{mkt}} N_{S,s,t}P_{S,s}\tau_{S,m_s}$

return $\{N_{S,t}, C_t\}$
end function

The modelled fund in is initialized as healthy: it is endowed with the amount of capital such that the VEV is met. Letting funds start at a equal level of capital would bias the comparison. For instance, the level of indexation would be much smaller for risky firms only because it would be much harder to obtain the returns that are needed to grant full indexation. See the Assumptions section for the calibrated parameters for the buffers required per risk factor. The VEV of a fund implementing hedge ratio h and equity-to-total-asset-value ratio h is calculated with the function SquareRootFormula:

function SQUAREROOTFORMULA($a, h, \mathcal{L}, DF$)
  $\rho_{S1,S2} = 0.4$; ▷ calibrated by DNB[a]
  $\rho_{S1,S5} = 0.4$;
  $\rho_{S2,S5} = 0.5$;
  $f = 0$; ▷ initialize as zero
  $VEV = 105\%$; ▷ initialize as MVEV
  while $f \neq VEV$ **do** ▷ funding ratio should equal VEV
    $f = VEV$;
    $L = \sum_k DF_k \mathcal{L}_k$;
    $A = f \times L$; ▷ value assets
    $A_F = 0.5051 \times A$;
    $E = a \times A$; ▷ Equity
    $FI = (1 - a)\, A$; ▷ Fixed Income
    $S1_L = \frac{\sum_k kDF_k \mathcal{L}_k}{L}$; ▷ duration liabilities
    $S1_B = h \times S1_L$; ▷ interest rate risk of investments
    $S1_T = S1_B - S1_L$; ▷ total interest rate exposure
    $S2 = 0.3 \times E$; ▷ 30% potential loss in stocks
    $S3 = 0.095 \times A_F$; ▷ denomination risk of foreign assets
    $S4 = 0$; ▷ no commodities in model
    $S5 = 0$; ▷ no credit risk: AAA sovereign bonds only[b]
    $VEV = 1 + \frac{\sqrt{S1_T^2 + S2^2 + 2\rho_{S1,S2} S1 S2 + S3^2}}{100}$; ▷ square-root formula (nFTK)
  end while
  return f; ▷ required funding ratio
end function

[a] http://www.toezicht.dnb.nl/2/50-202138.jsp
[b] http://www.toezicht.dnb.nl/2/50-202270.jsp

The main model is an algorithm that implements the aforementioned functions as follows. Initially the fund is endowed with capital such that it satisfies its VEV in cash. The cash is used to form its initial portfolio. When the fund arrives at a new time point, the portfolio is updated. In particular, prices and values are updated, swap payments are exchanged, and the maturity of fixed income securities are reduced by Δt. Moreover, the fund recalculates its delta risks. All these steps are summarized by an update of Π.

The available securities in the market are updated too. More specifically, the fund receives new quotes it can buy the products for and information on the delta exposure of these products. This information is renewed every timestep and captured by Π^{mkt}.

At the arrival of a new time point the fund rebalances its portfolio if its current (realized) equity-to-total-asset-value ratio deviates more than 3%-point from its target equity-to-total-asset-value ratio (according to its investment strategy). Furthermore, when the fund's cash reserves are less than 0.5% of the total fund value, the fund is forced to sell of some assets to restore the cash reserves at 1%. When the fund satisfies both the equity-to-total-asset-value ratio and cash reserves within bounds, but

the hedge ratio is not satisfied, the fund enters new swap contracts to satisfy the hedge ratio.

When a full year passed, the fund evaluates its position and its risks. Based on this scenario, the fund tries to grant price level correction on the pension benefits. The maximum level of indexation is calibrated at 2% as this is the ECB's target of yearly inflation.

To compare the impact of hedging levels on a pension fund, I analyze a set of funds indexed by their hedge ratio $h \in \mathcal{H}$ and equity-to-total-asset-value ratio $a \in \mathcal{A}$. Thus a total of $|\mathcal{H}| \times |\mathcal{A}|$ different types of funds are investigated. The interaction between the two parameters is interesting. A fund that invests relatively more into equities but wants to maintain a high hedge ratio uses more swaps than a fund that invests more in bonds, because of the natural hedge character of bonds. Naturally, this procedure is repeated N times, where N is the number of simulations.

Amongst the results that are recorded are the value of both the assets and liabilities. By measuring these statistics, we can evaluate the performance of the portfolio as well as the funding ratio of the fund. Note that although the liabilities are assumed to be stationary, they are subject to potential growth due to price level compensation. This level of possible indexation is stored. Furthermore, the model records the transaction costs paid.

Algorithm 4 Pension Fund

for $n \in N$ **do**
  for $h \in \mathcal{H}$ **do**
    for $a \in \mathcal{A}$ **do**
      Load $\left\{\mathcal{E}, \mathcal{R}^{rf}, \mathcal{R}^{AAA}\right\}_n$;
      VEV =SquareRootFormula$(a, h, \mathcal{L}, DF_0^{rf})$;
      $L = \sum_k DF_{0,k} \times L_k$; ▷ present value liabilities
      $C = VEV \times L$; ▷ initially portfolio is only cash

      Update Π, Π^{mkt}; ▷ Update prices and delta's
      $\{N_{E,t}, N_{B,t}, N_{S,t}, C_t\}$ = Portfolio(Π, Π^{mkt}, a, h);
      Update Π; ▷ Update quantities held

      for $t \in T$ **do**
        Update Π, Π^{mkt}; ▷ Update prices and delta's

        Calculate $\hat{a} = \dfrac{N_{E,t}P_E}{N_{E,t}P_E + \sum_{b\in\Pi_B} N_{B,b,t}P_{B,b} + \sum_{s\in\Pi_S} N_{S,s,t}P_{S,s}}$;

        Calculate $\hat{c} = \dfrac{C_t}{N_{E,t}P_E + \sum_{b\in\Pi_B} N_{B,b,t}P_{B,b} + \sum_{s\in\Pi_S} N_{S,s,t}P_{S,s} + C_t}$;

        Calculate $\hat{h} = \dfrac{\Delta_{\mathcal{L},p}}{\sum_{b\in\Pi_B} \Delta_{B,b,t,p}N_{B,b,t,p} + \sum_{s\in\Pi_S} \Delta_{S,b,t,p}N_{S,s,t,p}}$;

        if $|\hat{a} - a| > 0.03 \vee \hat{c} < 0.05$ **then**

          $\{N_{E,t}, N_{B,t}, N_{S,t}, C_t\}$ = Portfolio(Π, Π^{mkt}, a, h);
        else if $\hat{h} < h$ **then**
          $\{N_{S,t}, C_t\}$ =AddSwaps(Π, Π^{mkt}, h);
        end if
        Update Π; ▷ Update quantities held

        if $t \times \Delta t$ mod $= 0$ **then**
          VEV =SquareRootFormula$(a, h, \mathcal{L}, DF_t)$;
          Calculate $f = \dfrac{\sum_k DF_{t,k} \times L_k}{N_{E,t}P_E + \sum_{b\in\Pi_B} N_{B,b,t}P_{B,b} + \sum_{s\in\Pi_S} N_{S,s,t}P_{S,s} + C_t}$;

          if $f > VEV$ **then**
            $\mathcal{L} = \mathcal{L} \times 1.02$; ▷ Full Indexation
          else if $1.10 \leq f \leq VEV$ **then**
            $\mathcal{L} = \mathcal{L} \times \left(1 + \frac{0.02}{VEV - 1.10} \times f\right)$; ▷ Partial Indexation
          end if
        end if
      end for
    end for
  end for

4 Data

First, I discuss the data on which the Heston-Hull-White model is estimated: historical zero rates and equity data. The analysis is performed on two different data sets. More specifically, one regime covers the recent QE period, while the other ranges over a period of economic expansion (EE). Lastly, the calibration of the transaction costs and pension benefits is discussed.

4.1 Interest Rates

The model uses interest rates for two purposes: discounting and pricing. Following the nFTK (2015), liabilities of the pension fund are discounted using risk-free rates. As before the great financial crisis of 2007 sovereign interest rates were considered risk-free, the discounting curve and the curve used for prime grade government bond pricing coincide. The European debt crisis and the subsequent QE period have proven that this no longer is the case. In that period the model uses swap curves for discounting and a different for bond pricing.

4.1.1 Risk-free Rate

4.1.1.1 QE

Following the nFTK (2015), the risk-free rates are proxied by swap rates after the great financial crisis. I use EONIA zero curves obtained from Bloomberg[14]. The zero curves range from the 11th of February 2014 until the 9th of September 2016. Every zero curve contains information on the yields of the 1 to 30 year maturity. This translates to a panel with 661 time observations and 30 cross-section observations, totaling 19830 observations. Figure 5 below depicts the evolution of the EONIA term-structure over time. The figure clearly depicts the downward trend of the interest rates due to QE.

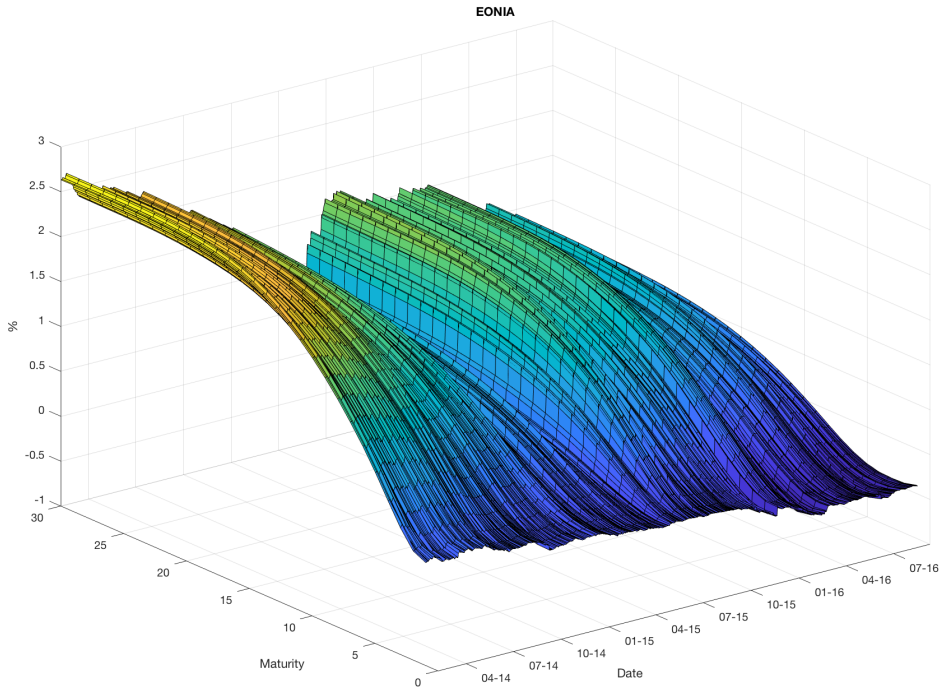


Figure 5: Evolution of EONIA term-structure

[14]These curves are bootstrapped using EONIA swaps that are quoted on the Bloomberg terminal.

Table 2 displays the summary statistics of the term-structure of the interest rates. Over the whole sample the short end of the curve denotes negative values. The yield curve is upward sloping. From the 10 year maturity almost all observations are positive as the 1%-percentile indicates. The statistics show a clear trend in the volatility; the longer the horizon, the higher the volatility.

Table 2: Descriptive Statistics of the Term-Structure During QE

Maturity	Mean	Std. Dev.	1%-Perc.	Median	99%-Perc.
$1Y$	-0.001	0.002	-0.005	-0.001	0.001
$2Y$	-0.001	0.002	-0.005	-0.001	0.002
$3Y$	-0.001	0.002	-0.005	-0.001	0.004
$5Y$	0.001	0.003	-0.004	0.001	0.009
$7Y$	0.004	0.004	-0.003	0.004	0.013
$10Y$	0.008	0.005	0.000	0.007	0.019
$15Y$	0.012	0.005	0.004	0.012	0.024
$20Y$	0.014	0.006	0.005	0.014	0.026
$25Y$	0.015	0.006	0.006	0.014	0.026
$30Y$	0.015	0.006	0.006	0.014	0.026

Over the different years the mean declined consistently for all maturities, see Table 3. Furthermore, the standard deviation of all maturities, except for the five and seven years, decreased consistently as well. This finding is in line with the policy of the ECB to maintain lower and stable interest rates.

Table 3: Descriptives Statistics of the Term-Structure per Year During QE

	Mean			Std. Dev.		
Maturity	2014	2015	2016	2014	2015	2016
$1Y$	0.001	-0.002	-0.004	0.001	0.001	0.004
$2Y$	0.001	-0.001	-0.004	0.001	0.001	0.004
$3Y$	0.002	-0.001	-0.004	0.001	0.001	0.001
$5Y$	0.004	0.001	-0.003	0.002	0.001	0.001
$7Y$	0.008	0.003	-0.001	0.003	0.002	0.002
$10Y$	0.012	0.006	0.003	0.004	0.002	0.002
$15Y$	0.017	0.010	0.007	0.004	0.003	0.002
$20Y$	0.020	0.012	0.009	0.004	0.003	0.003
$25Y$	0.021	0.013	0.009	0.003	0.003	0.003
$30Y$	0.021	0.013	0.009	0.003	0.003	0.003

4.1.1.2 EE

The course of EE is reflected by the run-up to the crisis. More specifically, the period starts at September 6th 2004 and ends at December 29th 2006. Where post-crisis EONIA or LIBOR swap curves are considered to be risk-free, sovereign yields were considered to be risk-free before the crisis. Figure 6 depict the course of the average of European prime graded government bonds with a 3-month

time to maturity[15] (in blue) against the 3-month Libor (based on Euro) rate[16](in orange). The yield on prime graded government bonds was lower than the 3 month Libor rate over the whole course, which supports the use of European prime graded government bond yields as risk-free rates.

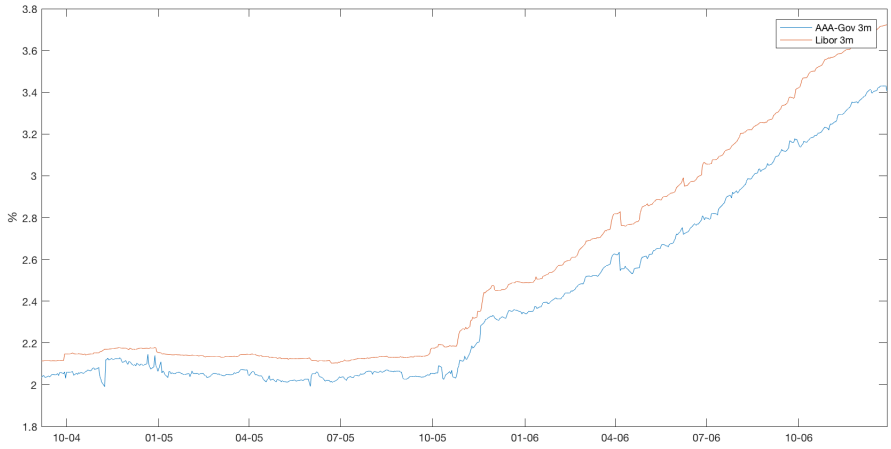


Figure 6: AAA Sovereign (blue) versus Libor Yields (red)

The evolution of the term-structure of the European prime graded government bonds[17] over the sample period is given in 7. Every zero curve contains information on the yields of the 1 to 30 year maturity. This translates to a panel with 597 time observations and 30 cross-section observations, totaling 17910 observations. Compared to the zero curves during the QE period the term-structure remains at a higher level.

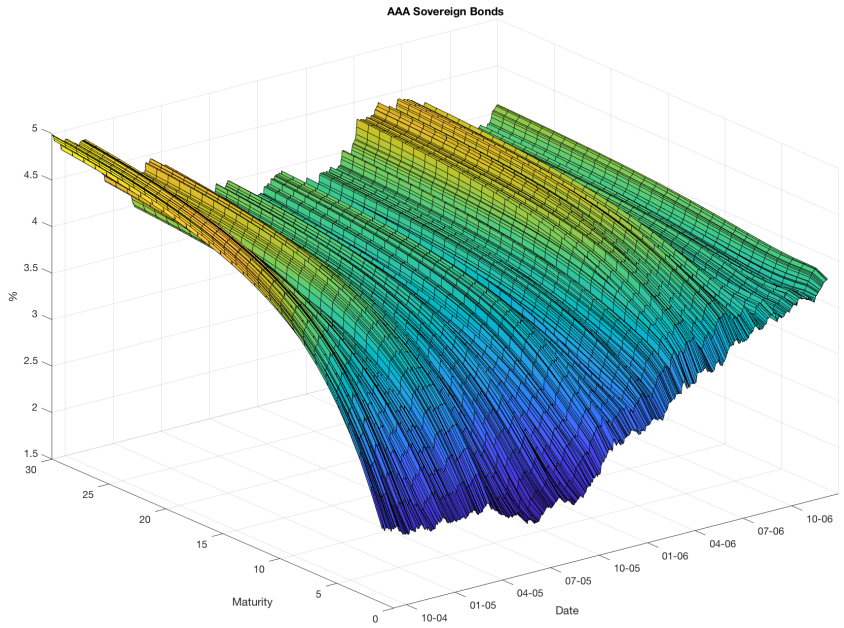


Figure 7: Evolution of European AAA Government Yield Curves

[15]Data is obtained at the ECB: `http://sdw.ecb.europa.eu/browseExplanation.do?node=9689726`
[16]Data is obtained at the FRED St. Louis: `https://fred.stlouisfed.org/search?st=libor+euro`
[17]Data is obtained at the ECB:`http://sdw.ecb.europa.eu/browseExplanation.do?node=9689726`

Table 4 displays the summary statistics of the term-structure of the interest rates. The yield curve is upward sloping with less volatility at the end of the curve. By comparing the mean and median one sees that the distribution of the yields are quite symmetric for all maturities.

Table 4: Descriptive Statistics of the Term Structure in EE

Maturity	Mean	Sigma	1%-Perc.	Median	99%-Perc.
$1Y$	2.644	0.557	1.952	2.409	3.700
$2Y$	2.809	0.536	2.015	2.646	3.758
$3Y$	2.940	0.478	2.173	2.833	3.757
$5Y$	3.181	0.374	2.534	3.118	3.799
$7Y$	3.389	0.313	2.800	3.383	3.914
$10Y$	3.626	0.278	3.095	3.649	4.130
$15Y$	3.865	0.281	3.402	3.830	4.500
$20Y$	3.996	0.297	3.515	3.927	4.708
$25Y$	4.076	0.311	3.572	3.991	4.838
$30Y$	4.129	0.322	3.609	4.032	4.925

4.1.2 Government Bond Spread

Following Zhu (2012), a spread is added to the simulated risk-free zero curves to create prime grade bond pricing curves. As such, only one term-structure needs to be modelled. In the EE period, risk-free rates and prime grade government bond yields coincide such that no spread needs to be added. In the QE regime, I add the average spread between the prime grade government bond yield curves and the EONIA term-structure[18]. Figure 8 depicts the spread between AAA European government bonds and EONIA over the sample period.

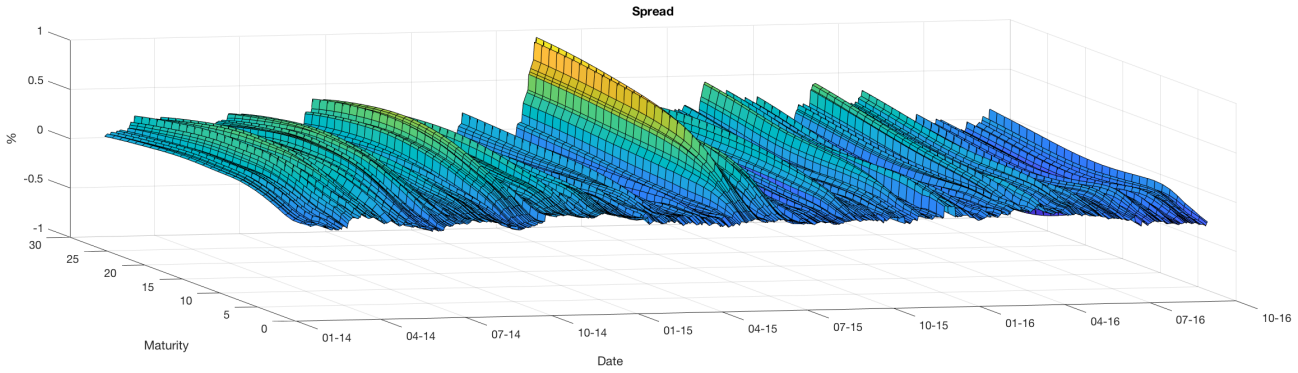


Figure 8: Evolution of AAA Government-EONIA Spread

The spread hovers around zero as expected, reflecting that prime grade government bonds are close to being risk-free. Although the spread is not constant, it remains relatively stable over the years.

[18]Again, the prime grade European government bonds data are from the ECB:`https://sdw.ecb.europa.eu/browseSelection.do?type=series&q=spot+rate+triple+A&node=SEARCHRESULTS&ec=&oc=&rc=&cv=&pb=&dc=&df=`

Moreover, there seems to be some structure in the spread curves. Since October 2014 - the start of the Covered Bond Purchase Programme 3 (CBPP3) - the spread is declining due to the various asset purchase programs of the ECB. The spread series depict its peak at 14 May 2015. This short increase in the spread is, amongst other reasons, due to unexpected increased inflation (due to the increase in oil prices) and the increased supply of government bonds on the secondary market (as the ECB was willing to buy anything on the market)[19]. Because of these peculiarities I will use the 2014 sample to set the constant spread, see Figure 9. The resulting spreads are in line with the spreads reported by the ECB[20]. The negative short end can be explained by the results of QE.

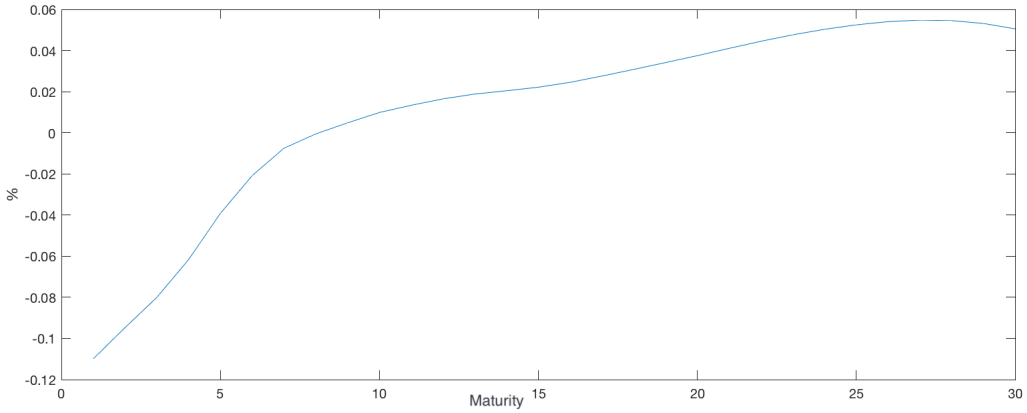


Figure 9: Average of EONIA AAA Government Spread

4.2 Equity

Equity in the model is represented by the Euronext100, which is regarded as the blue-chip index of Europe. This index includes national exchanges such as the AEX-index, BEL 20, CAC 40 and PSI 20. Specifically, it includes stocks of firms with the highest market capitalization on Euronext. Moreover, as a liquidity constraint, every stock on the Euronext 100 must reach a trade volume of at least 20% of its issued shares per year.

4.2.1 QE

Figure 10 shows the course of the Euronext100 during the QE period, which ranges from the 11th of February 2014 until the 9th of September 2016. The data is obtained from Yahoo Finance. The sample contains 661 observations. The most negative return of -6.96% occurred at the 24th of June 2016 when the English population voted to leave the EU. The second most negative return was -5.31% which happened on the 24th of August 2015 as a consequence of China's worst trading days in history: "China's Black Monday".

[19] http://www.investmenteurope.net/opinion/reasons-for-the-recent-bond-market-sell-off/

[20] https://www.ecb.europa.eu/pub/pdf/other/art1_mb201407_pp63-77en.pdf?0166cbc8f40410fb99cbcc51c1b07bf2

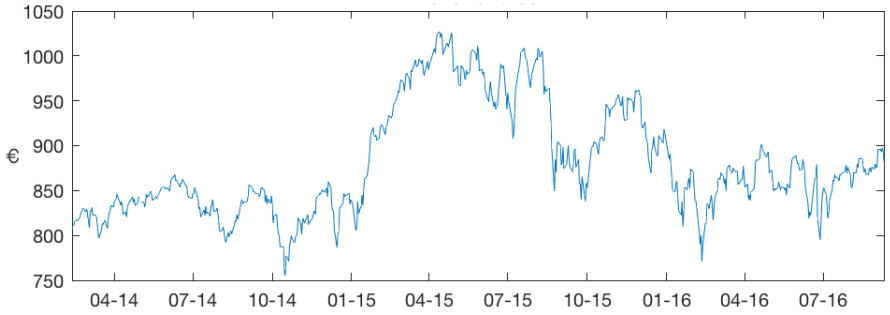


Figure 10: Evolution Euronext100 During QE

The level of the Euronext100 in 2015 seems a lot higher than the level in 2014 and 2016. The decline of crude oil prices initiating in 2014 probably had a positive effect on the Euronext100. Also, the announcement and initiation of the ECB's PSPP program (both in 2015) that tried to stimulate economic growth with low interest rates had a positive effect on equity prices. Moreover, due to the low interest rates fixed income markets were less attractive with the result of investors deviating to the equity markets pushing up stock prices.

Table 5 displays the descriptive statistics of the Euronext100 and its return series, taken as the log difference $r_t = \log(P_t) - \log(P_{t-1})$, during QE. On average the Euronext realizes a basis point return per day. The return series is skewed to the left as we can see from the mean and median. This is in correspondence with heavy tails theory. From Table 6 two regimes stand out. In 2015 - with the start of the PSPP - the Euronext100 switched to a different level and higher volatility. In 2016 the Euronext100 reverted to it's old mean, denoting negative returns on average. The volatility remained relatively high.

Table 5: Descriptive Statistics of the Euronext100 During QE

	Mean	Std. Dev.	1%-Perc.	Median	99%-Perc.
Euronext100	880.766	61.069	786.836	862.820	1016.718
Returns (%)	0.014	1.228	-3.407	0.076	3.176

Table 6: Descriptive Statistics of the Euronext100 per Year During QE

	Mean			Std. Dev.		
	2014	2015	2016	2014	2015	2016
Euronext100	829.383	941.742	859.230	20.299	49.981	23.856
Returns (%)	0.018	0.033	-0.0120	0.919	1.341	1.397

4.2.2 EE

Figure 11 shows the course of the Euronext100 during the EE, which ranges from the September 6th 2004 until the December 29th 2016. The data is obtained from Yahoo Finance. The sample contains 597 observations. Compared to the QE sample, this period is characterized by a strong upward trend.

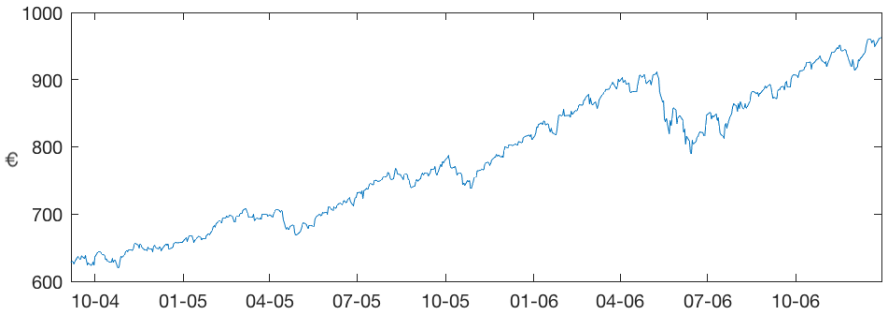


Figure 11: Evolution Euronext100 in EE

Table 7 displays the descriptive statistics of the Euronext100 and its return series in EE. On average the Euronext realizes a tenth of a basis point return per day. The return series are relatively symmetric as we can see from the mean and percentiles - especially when comparing to the returns during QE. The average level of the Euronext100 is lower than the level during QE. A possible explanation is the portfolio substitution effect of investors during QE. Investors resort to equities as fixed income securities are less attractive, pushing up stock prices. The higher standard deviation of the series stems from the strong upward trend in the series. When comparing to the QE sample, there is much more variation in the level of the series. However, the volatility of the returns is much lower. Although the returns are lower on average, the series slopes upwards steady and smooth. The 1%- and 99%-percentiles of the returns show that there are no large swings in the returns, whereas during QE the returns series denoted heavy negative and positive returns.

Table 7: Descriptive Statistics of the Euronext100 in EE

	Mean	Std. Dev.	1%-Perc.	Median	99%-Perc.
Euronext100	780.939	97.525	625.301	770.740	958.217
Returns (%)	0.001	0.007	-0.022	0.001	0.019

4.3 Transaction Costs

Transaction costs are calibrated to estimates from the asset pricing literature. I consider three different regimes of transaction costs to assess the sensitivity of the results to transaction cost: low, medium, and high. The medium regime is the input in the main analysis. The other regimes form robustness tests. For large institutional investors, such as pension funds, transaction costs might include costs associated with slippage. Slippage costs are the costs that an investor incurs by spreading it's investment over time in order to minimize market impact. By analysing different regimes of transaction costs, I incorporate different transaction cost factors in a general manner.

In the literature there is a vast amount of estimates for equity transaction costs for institutional investors, see e.g. Mei et al. (2016) or Garleanu & Pedersen (2009). The transaction costs in the literature are modelled as a fixed or variable costs, or a combination thereof. I stick to variable costs for simplicity. Transaction costs for equity range from 20 to 30 bps of the size of the trade, see Table

8.

Bond transactions costs are typically increasing in maturity. However, for government bonds this effect is less pronounced, see Chakravarty and Sarkar (1999). Moreover, for investment grade bonds this difference is relatively small for large trades, see Elkamhi et al. (2017) and references therein. For liquid products the difference diminishes even further. In line with A study of the Rabobank that estimates average trading costs of government bonds[21], I opt to include a transaction costs that differ over four maturity buckets. Again, I use three different transaction cost curves for the various regimes that define the general costliness. The medium regime transaction curve is calibrated to the estimates obtained by the Rabobank. The curves of the other regimes are obtained by adding (high regime) or subtracting (low regime) a basis point to the whole curve. Bond trading cost are given in Table 8.

The ISDA Margin Survey (2011) reports that 78.6% of fixed income derivatives are subject to a collateral agreement[22]. Allowed types of collateral are typically cash and government bonds. Pension funds, typically with a large bond portfolio, specify their CSA in terms of government bonds[23]. In my analysis I assume that the pension fund trades in swaps collateralized by government bonds. Swaps specified with a CSA are subject to small transaction cost as indicated by the estimates by the Rabobank report. Table 8 summarizes the transaction costs associated with swap trading.

Table 8: Transaction costs in basis points

		Low	Medium	High
Equity		20	25	30
Fixed Income	0-2Y	0	0.5	1
	3-7Y	0.5	1	1.5
	8-12Y	1	1.5	2
	>12Y	1.5	2	2.5
Swaps	0-5Y	0	0.5	1
	6-10Y	0	0.5	1
	>10Y	0.5	1	1.5

4.4 Pension Benefits

Data on the liabilities of the pension fund is, unfortunately, classified data. The data is obtained at APG Asset Management NV. The shape of the pension benefits curve is similar to the one presented in Figure 1 the Introduction. The actual size of the pension benefits do not matter for the results. Only the relative size of the payouts compared to each other are relevant. This follows as the asset side of the fund is initialized as a factor, being the VEV, times the present value of the liabilities. Moreover, the results are presented as relative measures, making actual size unimportant. To conclude, the research is replicable as long as pension benefits reflect the shape of Figure 1.

[21] https://www.rabobank.com/en/images/ex-ante-costs-and-charges-disclosure.pdf

[22] https://www.isda.org/a/neiDE/isda-margin-survey-2011-final.pdf

[23] Potter & Lansink, 2011, Meer waardering voor onderpand - Hoe onderpand de waarde van derivaten bepaalt, *VBA Journaal*, 105, p. 12-15.

5 Results

This chapter is organized as follows. First, I discuss the estimation results of the Heston-Hull-White model. I first demonstrate the performance of the algorithm with the help of a simulation experiment. Then, I apply the estimation algorithms to the two different data sets discussed in the previous section. Second, I examine the results of the pension fund algorithm applied to the financial market simulations generated by the Heston-Hull-White model.

5.1 Financial Market Estimation

5.1.1 Simulation Experiment

Following Aihara et al. (2012), the simulation parameters are set to:

1. $(\kappa_r, \sigma_r) = (0.022, 0.009)$ for the Hull-White part of the model.
2. $(\mu, \kappa_V, \theta_V, \sigma_V) = (0.05, 1.8, 0.6, 2.1)$ for the Heston part of the model.
3. $\begin{bmatrix} 1 & \rho_{r,S} & \rho_{r,V} \\ \rho_{r,S} & 1 & \rho_{S,V} \\ \rho_{r,V} & \rho_{S,V} & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0.2 & -0.1 \\ 0.2 & 1 & -0.4 \\ -0.1 & -0.4 & 1 \end{bmatrix}$ for the correlation of the Wiener processes.
4. $(r_0, S_0, V_0) = (0.001, 1, 0.3)$ as initialisation.
5. $\Delta t = \frac{1}{250}$ and $T = 1$. The first parameter particles resampling is set after the first 20 evaluations.
6. $(iN_{vol}, iN_{par}) = (200, 200)$ as number of volatility and parameter vector particles for the Aihara-Adapted algorithm. Therefore, a total number of 4000 particles are evaluated.
7. The initial term-structure taken to initialize the Hull White model is the EONIA spot curve of February 11th of 2014. The simulated pillars are the 3, 5, 7, 10, 15, 20, and 25 year maturities.

Table 9 depicts the parameter estimates. Figure 12 shows the simulated short rate in green and the estimated short rate in blue. Figure 13 displays the simulated and estimated zero curves in green and blue, respectively. The two plots in both figures are indistinguishable as the Kalman Filter provides an excellent estimate. It is not surprising that the Kalman filter performs very well; the short rate dynamics are relatively simple as they are Gaussian and not heavily non-linear. The parameters and short rate are estimated without any bias. The downward trend in the short rate is the effect of the initial term-structure.

Table 9: Simulation Experiment: Parameter Estimates

Parameter	True	Kalman	Aihara-Adapted
κ_r	0.022	0.022 (0.001)	
σ_r	0.009	0.009 (0.029)	
μ	0.05		0.059 (0.025)
κ_V	1.8		1.689 (0.532)
θ_V	0.6		0.594 (0.164)
σ_V	2.1		2.642 (0.145)
$\rho_{r,S}$	0.2		0.197 (0.088)
$\rho_{r,V}$	-0.1		-0.102 (0.044)
$\rho_{S,V}$	-0.4		-0.407 (0.131)
MSE		1e-09	0.0280
T		250	250
N		7	1

standard errors reported in parentheses.

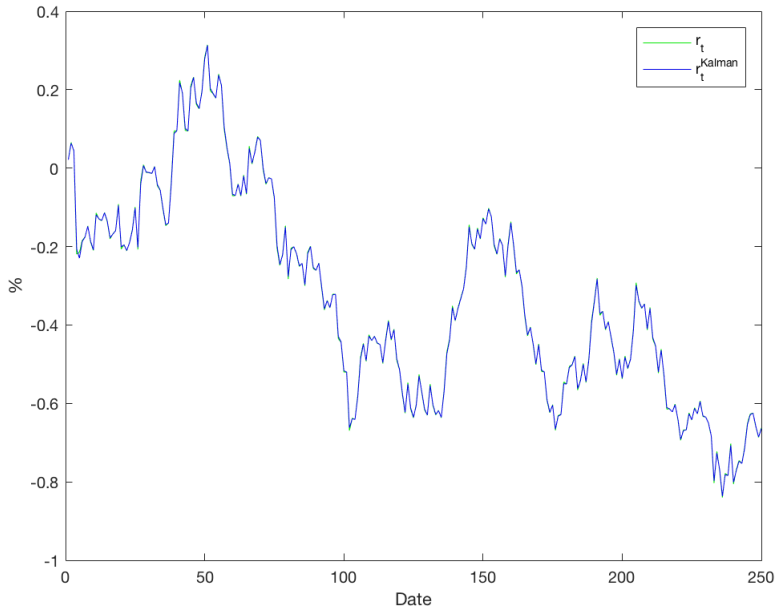


Figure 12: Simulation Experiment: Kalman Filtered Short Rate.

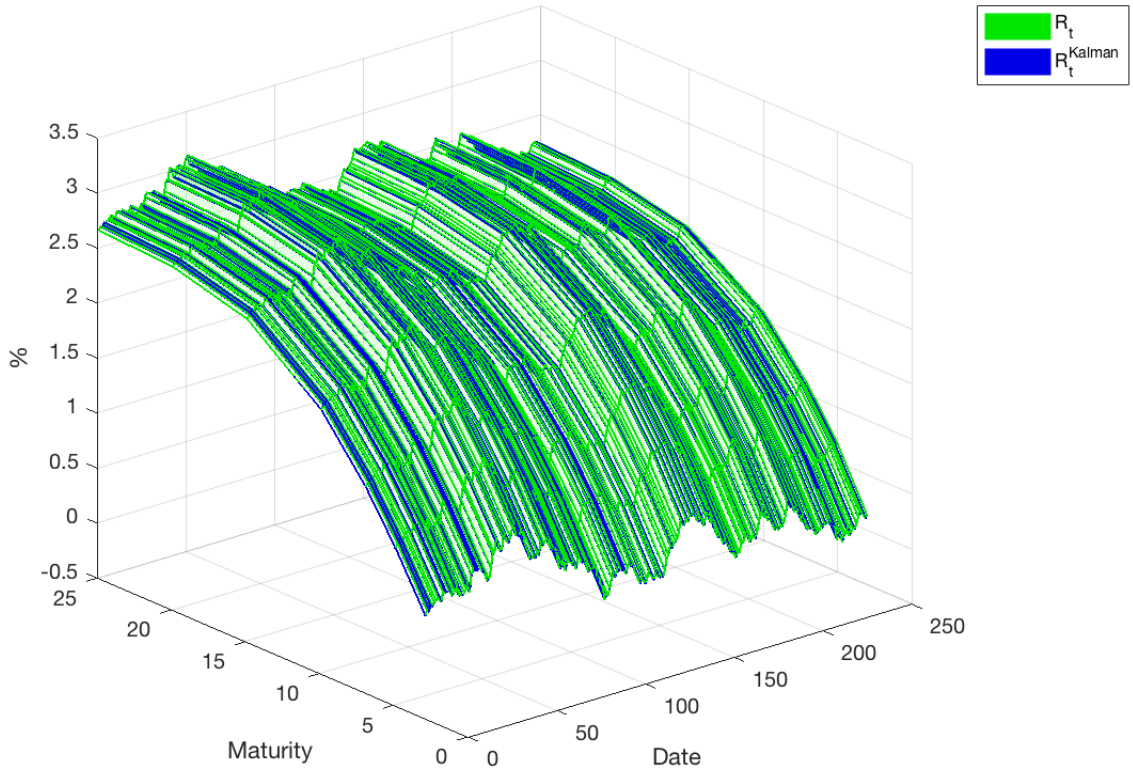


Figure 13: Simulation Experiment: Kalman Fit of Zero Curves

The results of Table 9 and Figure 14 point out that the Aihara-Adapted particle filter performs very well in volatility and parameter estimation. The only problematic parameter for the Aihara-Adapted particle filter is the volatility of volatility parameter σ_V: the bias of the estimate such that the true parameter lies outside the 95% confidence interval. In a lot of particle filters in the literature this parameter seems to cause problems. The Aihara-Adapted particle filter provides very low standard errors for the volatility estimates. This is a mechanical result of preselecting state variables based on the observation weights. Of course one could increase the number of particles, but increasing the number of volatility particles by a factor 10, increases the total particles by 10 times the number of parameter vector particles. This is a trade-off between accuracy in terms of standard errors and computing time. The particle filters are tested on a range of parameter values, time grids, and number of particles. The results are robust against these different specifications.

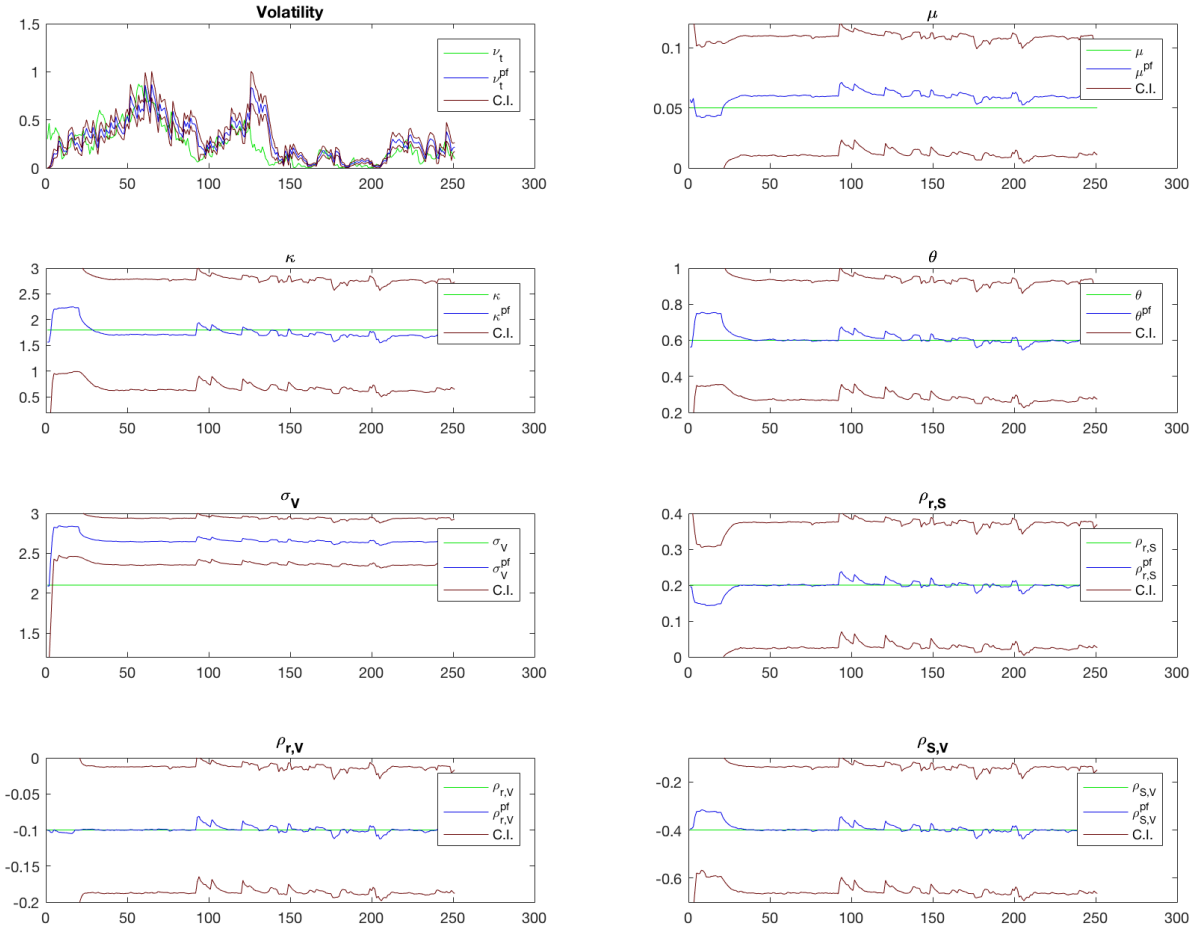


Figure 14: Simulation Experiment: Aihara-Adapted Particle Filter

5.1.2 Model Estimation

In this section the algorithms are applied to the data discussed in the Section Data. First, I demonstrate the model dynamics estimates during the QE episode. Second, model estimates in times of EE are discussed.

5.1.2.1 QE

The dynamics estimates during the QE period from February 2th 2014 to September 9th 2016 are shown in Table 10. The Kalman filter estimates are based on the 3, 5, 7, 10, 15, 20, and 25 year maturities. The Aihara-Adapted particle filter is based on 200 volatility particles by 200 parameter vector particles.

Table 10: Parameter Estimates During QE

Parameter	Kalman	Aihara-Adapted
κ_r	0.000 (0.000)	
σ_r	0.003 (0.018)	
μ		0.045 (0.019)
κ_V		1.345 (0.577)
θ_V		0.131 (0.056)
σ_V		0.753 (0.086)
$\rho_{r,S}$		0.298 (0.134)
$\rho_{r,V}$		-0.152 (0.067)
$\rho_{S,V}$		-0.352 (0.112)
T	661	661
N	7	1

standard errors reported in parentheses.

The Kalman filtered short rate and zero curves fit are displayed in Figure 15 and 16, respectively. As expected during QE, the short rate appears to be mostly negative. The fit of the zero curves is quite good overall. As the Hull-White model consists of one factor the short end of the curve is estimated a bit too steep. The Kalman filter estimates the mean-reversion at 0.0002. Although a low mean-reversion parameters is not uncommon in the literature, a possible explanation might be that the initial term-structure that determines the long-term mean in the Hull-White model is not representative in periods of unconventional monetary policy. To circumvent explosive behaviour in the simulation (as the factor $e^{-\kappa_r}$ turns up in the variance term of the short rate) I set this parameter to 0.001.

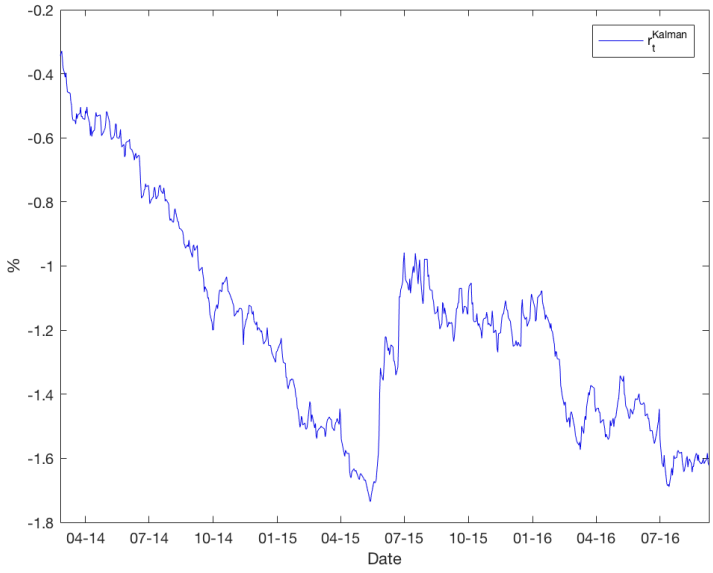


Figure 15: Kalman Filtered Short Rate During QE

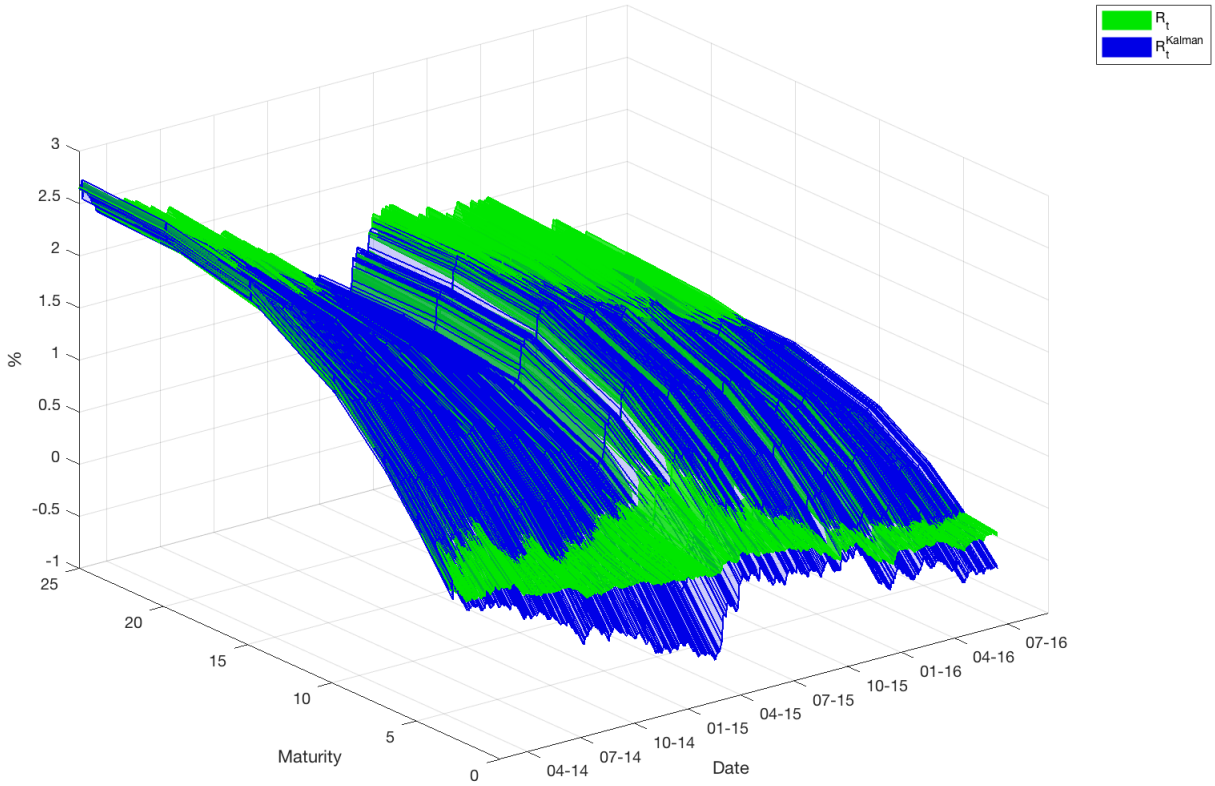


Figure 16: Kalman Fit of Zero Curves During QE

The parameter estimates obtained by the Aihara-Adapted particle filter are similar to those in the literature, see e.g. Grzelak & Oosterlee (2011) or Kienitz & Kammeyer (2009). Although their parameters are obtained through calibration, the Girsanov theorem states that when changing from the risk-neutral measure $\mathbb{Q}$ to the real measure $\mathbb{P}$ only the diffusion term of the stochastic differential equation changes, but the drift term remains. And indeed, comparing results show that the parameters estimates are very alike for the drift term, and diverse for the diffusion parameter.

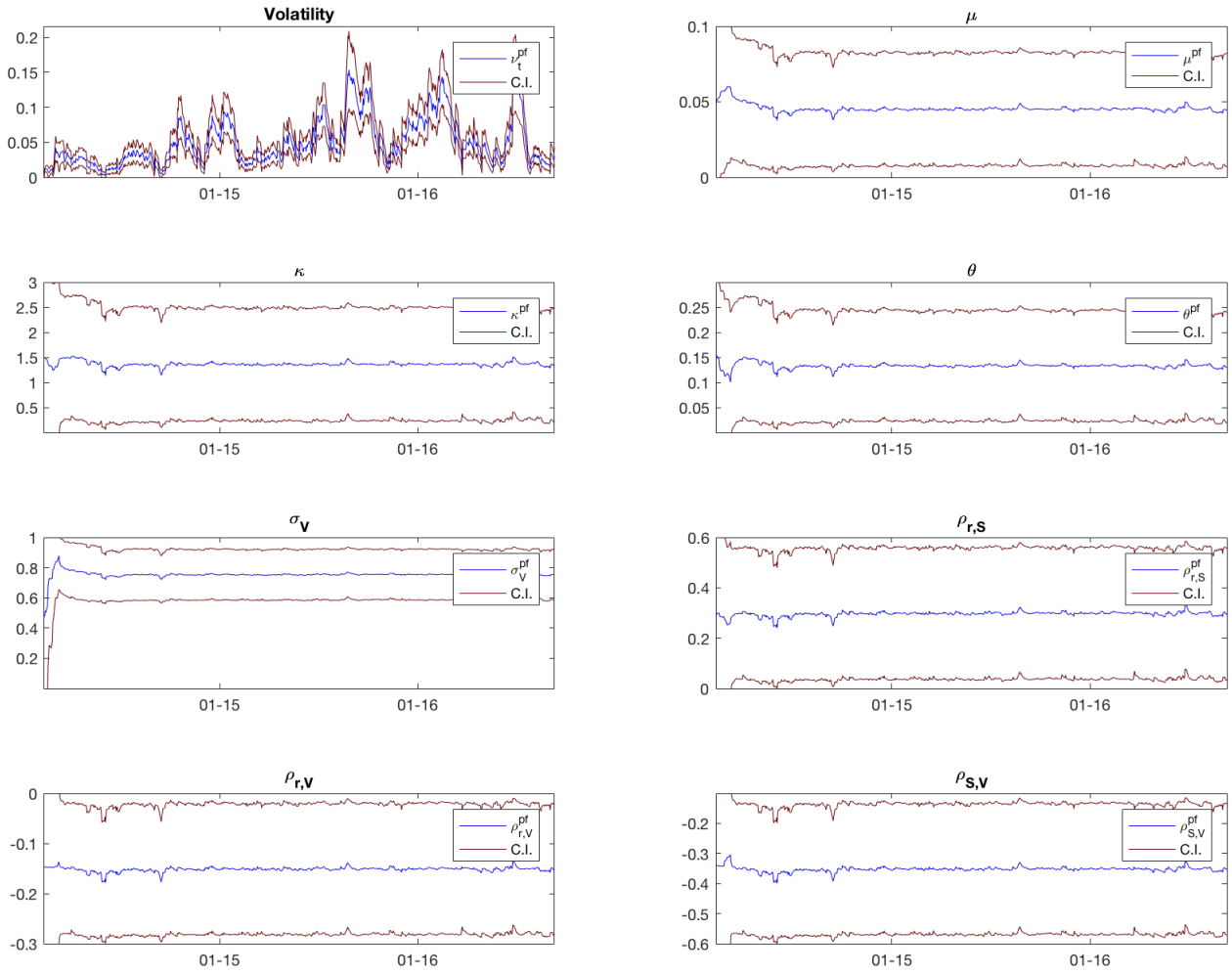


Figure 17: Aihara-Adapted Particle Filter Estimates During QE

5.1.2.2 EE

The estimates of the dynamics during the period from September 6th 2004 to December 29th 2006 are shown in Table 11.

Table 11: Parameter Estimates During EE

Parameter	Kalman	Aihara-Adapted
κ_r	0.000 (0.000)	
σ_r	0.002 (0.003)	
μ		0.130 (0.032)
κ_V		1.960 (0.470)
θ_V		0.065 (0.016)
σ_V		0.748 (0.114)
$\rho_{r,S}$		0.347 (0.119)
$\rho_{r,V}$		-0.125 (0.059)
$\rho_{S,V}$		-0.312 (0.100)
T	597	597
N	7	1

standard errors reported in parentheses.

The Kalman filtered short rate and fit of the zero curves are given in Figure 18 and 19, respectively. The estimated short rate is mostly increasing over the interval, as expected during EE. The fit of the zero curves is quite good overall. Again, the short end is estimated a bit to steep as the Hull-White consist of one factor. The Kalman filter has estimated the mean-reversion practically at zero again. To circumvent explosive behaviour in the simulation (as the factor $e^{-\kappa_r}$ turns up in the variance term of the short rate) I set this parameter to 0.001.

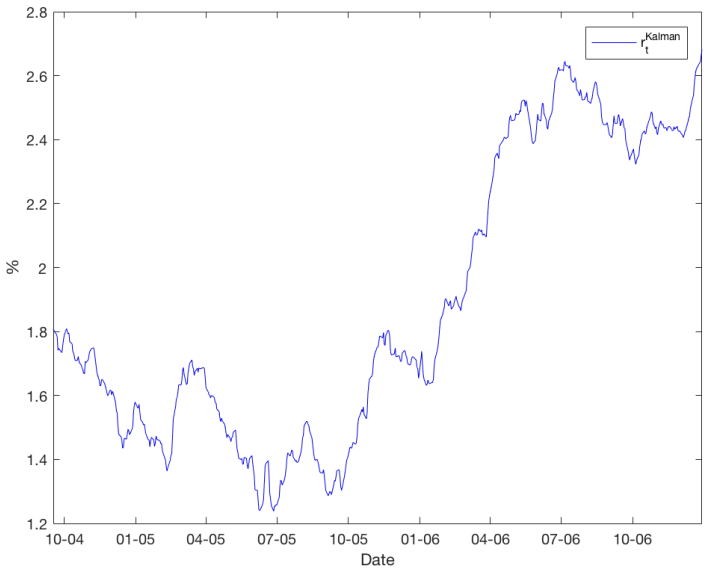


Figure 18: Kalman Filtered Short Rate in EE

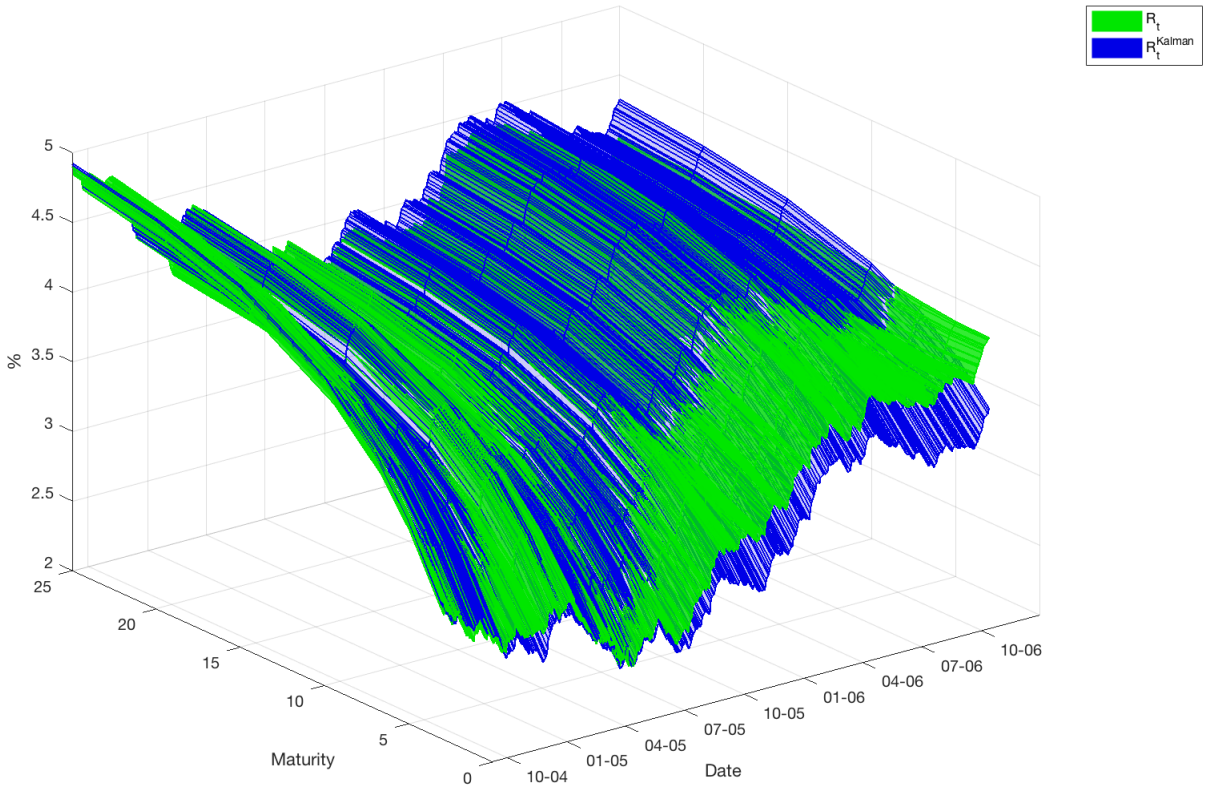


Figure 19: Kalman Fit of Zero Curves in EE

The Heston parameter estimates are given in Table 11, and can be seen from Figure 20.

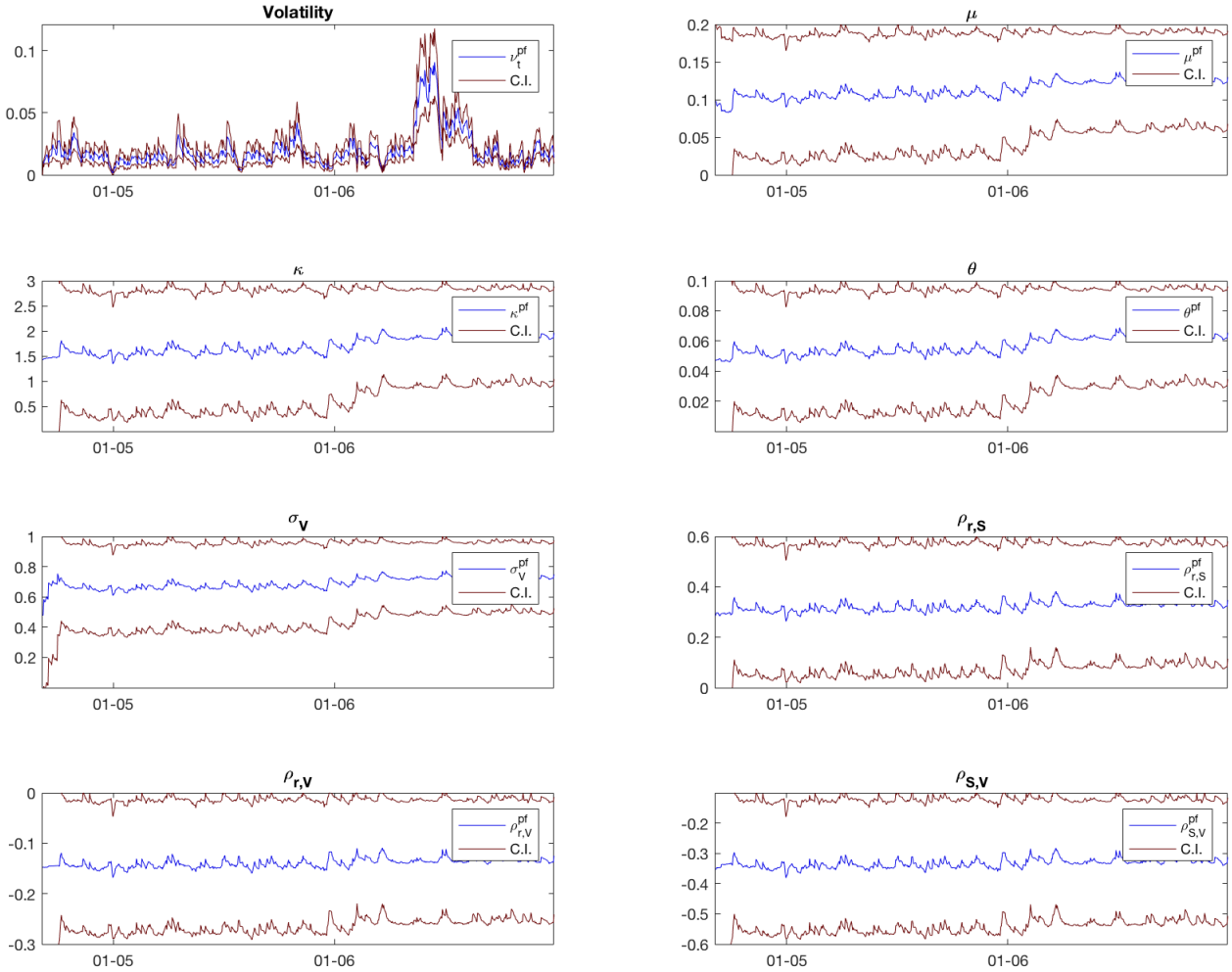


Figure 20: Aihara-Adapted Particle Filter Estimates in EE

5.1.2.3 Differences Between Regimes

Comparing the estimation results in both regimes, a few observations stand out. First, the short rate evolves systematically different during the QE episode: it is subject to a higher variance and remains for a great deal in negative territory. These findings are congruent with the QE objective of keeping interest rates and uncertainty in the interest rate market low. Although the real interest rate is bounded below at the so-called Zero-Lower Bound (ZLB), the nominal interest rate is not. But interestingly, there seems to exist a plateau that the nominal interest rate does not exceed: from Figure 15 it seems that the nominal short rate is bounded below around -1.75%. This observation is shared by the IMF[24].

Regarding the Heston part, the differences lie in the rate of return of the process and the long-term volatility. The other parameters are quite similar in both regimes. The long-term volatility is estimated lower during EE than QE. The summary statistics in Section Data pointed out that the volatility of the returns was higher during the QE episode than in the period of EE. Moreover, implied volatility indexes (VSTOXX for Europe and VIX for USA) show that volatility levels were lower in

[24] Viñals, Gray, and Eckhold:
`https://blogs.imf.org/2016/04/10/the-broader-view-the-positive-effects-of-negative-nominal-interest-rates`

times of EE than during the recent QE episode, see Figure 21. The mean level during QE were 22.36 and 15.76 for the VSTOXX and VIX, respectively. While during EE the mean of the VSTOXX and VIX were 15.39 and 12.94, respectively. The difference in volatility is larger in European data used as input in the algorithms than the difference in USA volatility (a 31.2% decrease for Europe versus 17.9% for the USA). A possible explanation lies in the fact that European corporations use more debt financing than equity compared to USA corporations, making them more susceptible to changes in interest rates. These observations support the estimates of the model.

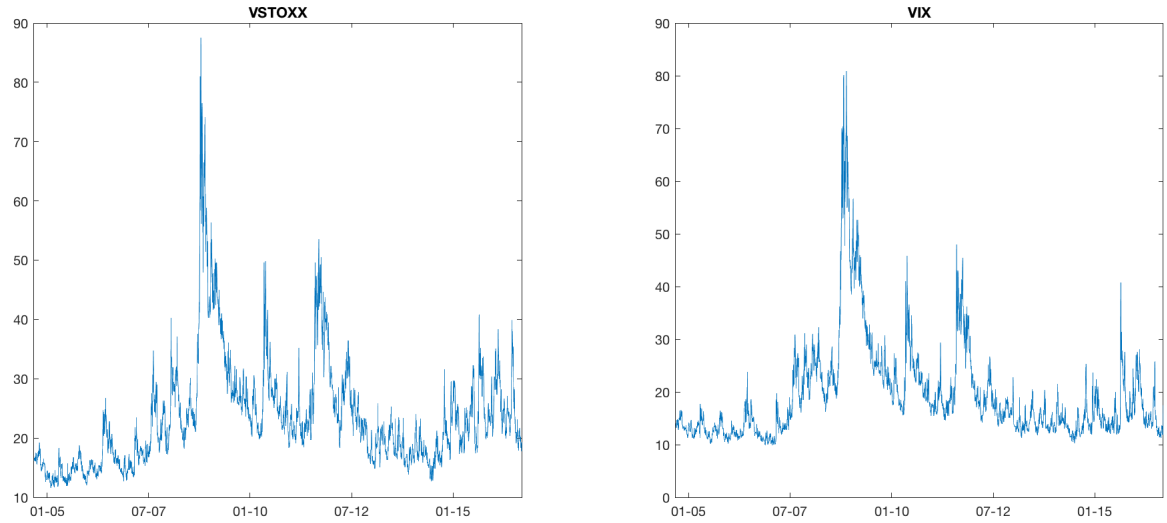


Figure 21: VSTOXX and VIX Over Sample Periods

Equity returns are typically very high in times of EE. Corporate expansion generally cause stock prices to rise through an increase in expectation of future dividends. The model estimates an average annual return of 12.9%. As described in Section Data the returns during QE were quite large, but went in both directions. In contrast, the returns in times of EE were lower on average. The long run steadiness of these returns creates a high annual return estimate.

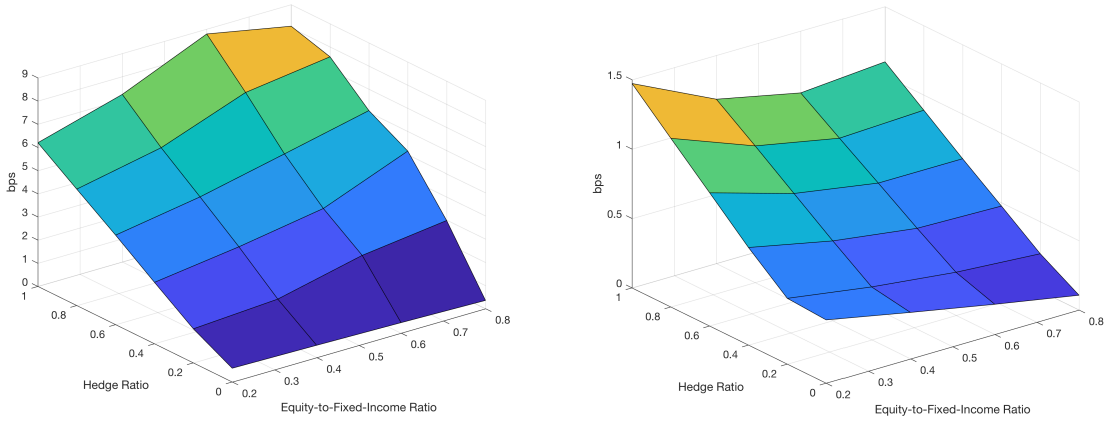
Empirical evidence denotes a positive correlation between stock returns and QE, see e.g. Bernanke & Kuttner (2005). Due to low interest rates safe fixed income securities are less attractive. Investors prefer to hold relatively risky securities in order to secure some return pushing stock market prices upward. Moreover, falling interest rates mean lower corporate borrowing costs. Companies have greater incentive to invest and take on more leverage, see Bernanke et al. (1999). This translates to an increase in investments and thus corporate expansion. However, this mechanism only applies if the central bank is successful in convincing the public. If investors do not believe the asset purchase program to be successful, QE will not have an effect on output, see Krugman (1998). Caballero et al. (2015) advocate that the QE mechanism works through exchange rates. Increasing the monetary base translates to an appreciated currency. This appreciated currency basically exports the recession abroad. As such, QE is called a beggar-thy-neighbor policy by some, see Caballero et al. (2015). Although a different mechanism, a beggar-thy-neighbor mechanism would also increase the stock price. Compared to pre QE returns the mean return is higher, but compared to the level of EE it is much lower. The explanation could be that higher uncertainty about the future (interest rate policy) induces lower stock prices through similar mechanisms as described above. This explanation would be in line with the lower volatility estimate of the model during EE.

5.2 Pension Fund Algorithm

In this section I present the results of the pension fund model. First, I discuss the costs of hedging. These costs can be decomposed into two parts: transaction costs and impact on fund performance. Second, I address the benefits of hedging. In particular, I examine the stability of the funding ratio and insolvency probabilities. The cost and benefits are aggregated into a measure to assess the optimal hedging ratio from the perspective of the nFTK: the level of indexation. In presenting the results I discriminate between the QE and EE regimes.

5.2.1 Transaction Costs

The estimated transaction costs are presented as the amount spent on transaction costs relative to the value of the fund. As every fund starts with a funding ratio equal to its VEV, and the VEV is a function of the hedge- and equity-to-total-asset-value ratio, riskier funds start relatively with a lot of capital. Therefore, by correcting for the value of the fund, transaction costs become comparable. The estimated relative expenditures on transaction costs cumulative over the simulation period are shown in Figure 22. Panel (a) shows the estimates for for the QE, and Panel (b) for the EE period. The data used to create the figures can be found in Table 12 and 13 in the Appendix.



(a) During QE

(b) In EE

Figure 22: Cumulative Relative Transaction Costs

The first salient result is the large difference in transaction cost expenditures between periods of QE and EE. As stated in the previous section, market volatility is higher in periods of QE. To comply with the chosen investment strategy the fund must rebalance more frequently. For fully hedged funds, expenditures become 4 to 5 times higher in periods of QE than in episodes of EE. A fully hedged pension fund spends between 6 to 9 basis points of total fund value on transaction costs, while in EE the costs would lie at a maximum of 1.5 basis point. To put these numbers into perspective, as the value of the ABP pension fund's assets are around €400 billion[25] a basis point corresponds to €40 million.

Second, expenditures are increasing in the hedge ratio. During the QE episode the difference in transaction costs expenditures can become 7-8 basis points of total fund value, which corresponds to €280-320 million for the ABP fund. Whereas a typical fund with an equity-to-total-asset-value ratio of 40% pays 0.5 bps for a zero hedge, the same fund pays 7.1 bps of total fund value for a full hedge.

[25] https://www.abp.nl/over-abp/financiele-situatie/actuele-financiele-situatie/

In EE the difference is less pronounced: the maximum difference in expenditures is around 1 basis point (0.3 bps for a zero hedge and 1.1 for a full hedge, based on a 40% equity-to-total-asset-value ratio fund). There are two mechanisms that explain this result. First, if swap payments eat up cash reserves, the fund needs to sell off some assets to replenish their 1% cash reserve, thus incurring some transaction costs. Second, adding swaps to the portfolio translates to more interest risk exposure (rho) of the portfolio. This could be beneficial for hedging purposes, but increases the portfolio volatility with respect to interest rate movements. Adhering to the equity-to-total-asset-value ratio means more frequent rebalancing. The increased market volatility during QE explains why this difference is larger in such episodes than in moments of EE.

Moreover, during QE there seems to be a positive relation between transaction costs and the amount invested in equities. This is a straightforward result as the transaction costs are highest for equities. In EE the relation is not linear. The transaction costs are U-shaped in the equity-to-total-asset-value ratio. The fact that intermediate values of the equity-to-total-asset-value ratio deliver lower transaction costs on average might be the result of an interplay between diversification and transaction costs. If interest rates are more volatile relative to equities, an investment strategy that invests less in fixed income securities might deliver lower transaction costs. On the other hand, an investment strategy that invests heavily in equities yields high transaction costs due to the aforementioned mechanism.

5.2.2 Fund Performance

The most straightforward measure to evaluate fund performance is the evolution of the fund value. First, this measure is presented. Second, to correct for the overall risk of the investment strategy the Sharpe ratio of the funds are discussed.

5.2.2.1 Fund Value

The estimates are presented as fund value relative to the initial fund value. As such, the presented statistic is a basic measure of return. The measure eliminates differences due to higher initial capital resulting from buffers required by the nFTK. Figure 23 and 24 depict the evolution of fund value during QE and EE, respectively. The data underlying the figures can be found in Table 14 and 15 in the Appendix. There is a clear relation between fund value and the amount invested in equities. This is not surprising as equities yield higher returns on average. However, this high return comes at the cost of a higher standard deviation. Obviously, as returns are higher in economic upturn, the pattern is more pronounced.

Interestingly, a high hedge ratio delivers a lower return on average in the short run (5 years) and a higher mean return in the long run (20 years) for the QE period. This pattern is present in the period of the economic upswing, but much less pronounced, see Table 15 in the Appendix. Ceteris paribus, a high hedge ratio means more intensive use of swap agreements. Then, swap payments and/or value have a negative (positive) effect on the portfolio in the short (long) run on average. There are two ways in which swaps have direct impact on the fund value. Indirect impact exists through increased rebalancing due to more interest rate exposure and compliance to the investment strategy. The first direct impact of swap agreements is the change in value. The second impact is the biannual payment. The impact on the portfolio must be the effect of an opposing movement relative to the leg taken on the swap agreement. For instance, pension funds are in general overexposed to interest rate movements in the short-end and underexposed to movements in the long end of the yield curve (see the duration mismatch between the liabilities and assets in both ends of the curve depicted in Figure 1). To diminish the duration mismatch to the level of the chosen hedge ratio, funds enter fixed-for-floating swap agreements in the short run and floating-for-fixed in the long run (note that I use the jargon of Glasserman (2003) where a floating-for-fixed swap is an agreement where the holder pays a floating rate and receives the fixed rate). Then, a decline in the short end of the curve decreases the value of

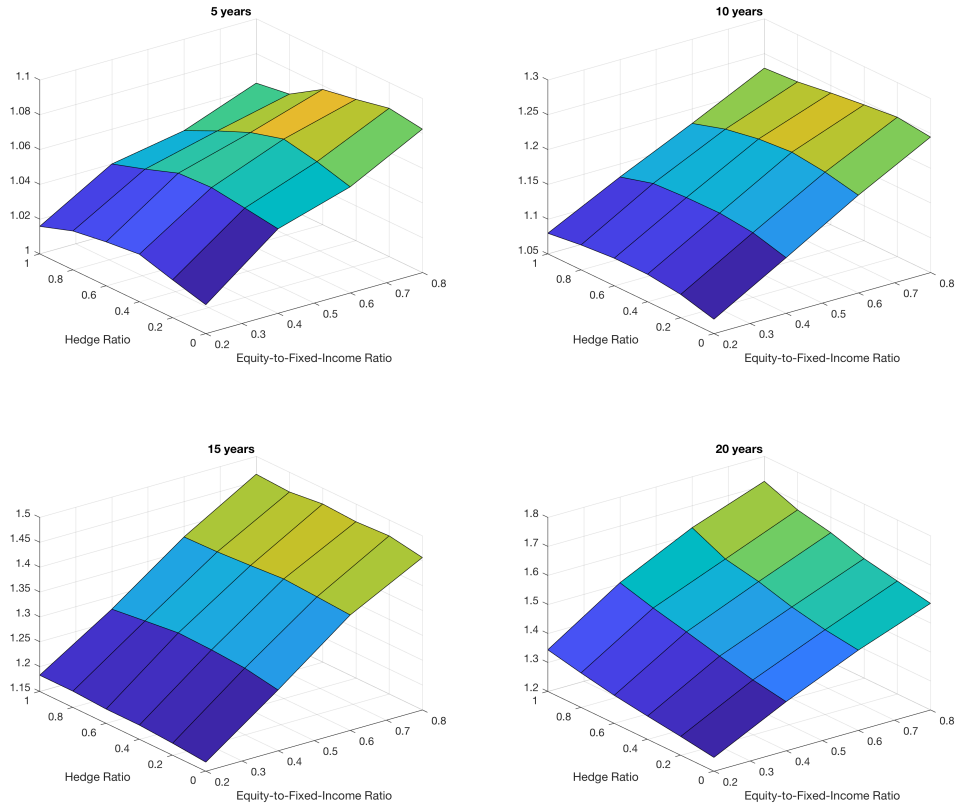


Figure 23: Fund Value Relative to Initial Value During QE

the swap portfolio, while a decline at the long end increases the value.

However, note that, in contrast with Figure 1, in my model the fund is not necessarily overexposed to interest rate risk at the short end of the curve. The market is perfectly liquid and the fund invests according to the investment strategy. As such, to satisfy the equity-to-total-asset-value ratio, the fund does not have to over-invest in bonds with a lower maturity due to market liquidity issues. The consequence is that the model does not enter different swap agreements for the short and long end of the curve systematically. The only exception is of course when it is cheaper, in terms of transaction costs, to satisfy the constraints.

As the model produces both positive and negative interest rate movements in the near future, it advices a modest hedge ratio in the short run. The difference is almost absent for funds with relatively more bonds. This is intuitive as a large bond portfolio act as a natural hedge and hence translates to a smaller usage of swaps. For equity intensive funds that invest more heavily in swaps to close the duration gap, the difference is around 2 percent points. However, as the term-structure reverts to its mean in the long run due to the shape of the initial term-structure, high hedge ratios become profitable in the long run. As the swap contracts cannot be opted out, all fixed-for-floating agreements already in the portfolio turn out profitable in the future. This result is stronger for funds that maintain high hedge ratios as they make more use of these type of swaps. The differences can become as large as 14 percent point.

The effect of hedging on fund performance becomes smaller when the volatility of the term-structure is smaller due to the ad hoc character of the hedging strategy: as pension funds continuously try to

maintain their hedge ratio, large swings in the term-structure translate to large impacts on the value of the fund. However, as the period of EE does not experiences a shift from low rates to positive rates (and thus experiences lower volatility), these results are absent. For this episode the fund value is almost the same for a certain equity-to-total-asset-value ratio (thus similar over the hedge ratio dimension)[26]. Only in the short run a hedge ratio of 20 percent seems to pay off. However, this difference is very small and diminishes over time.

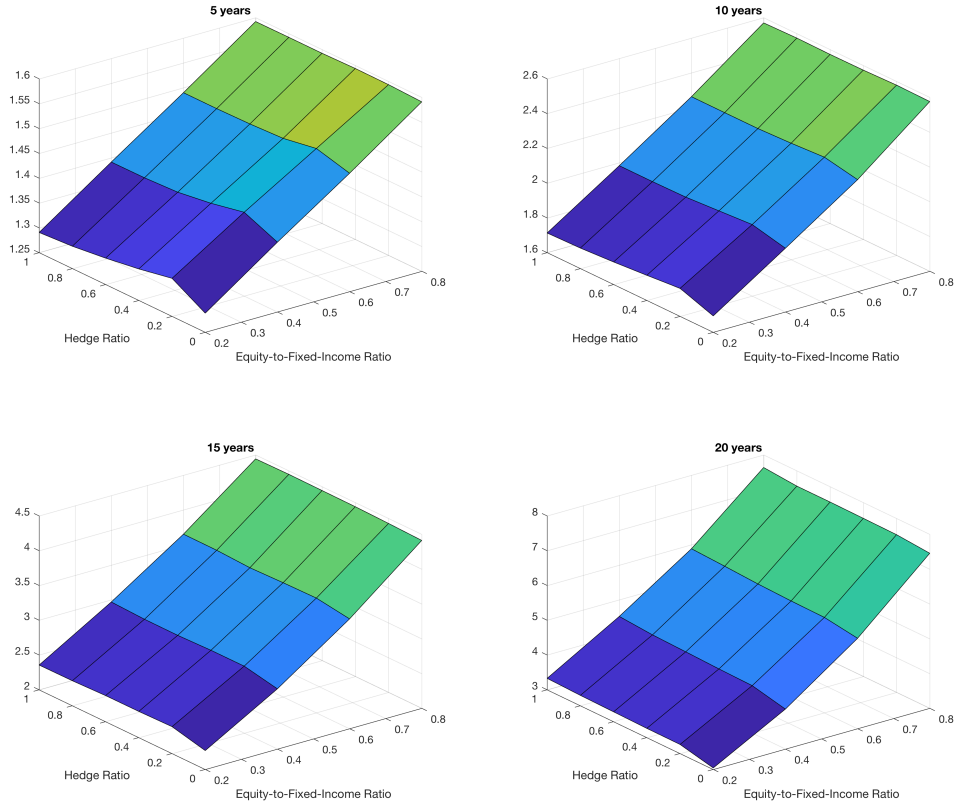


Figure 24: Fund Value Relative to Initial Value in EE

The results of the QE period show that the use of swaps can be beneficial in the short run, but also have lasting effects in the long run due to the stickiness of the contract. In this period high hedge ratios had a positive effect on fund value due to the positions in the fixed-for-floating swaps and the mean-reversion of the term-structure. However, the validity of these long run results is questionable as QE periods probably will not last for such an amount of time. Moreover, if the term-structure would not revert to its mean, but instead experience a change from a regime of high rates to low rates the long run effects on fund value of using swaps would have been negative. Therefore, it would be interesting to incorporate unanticipated shocks to the term-structure. Not only would the long run effects become more credible, also the hedging character of the swaps would stand out more clearly.

Interestingly, for both regimes, there are no clear differences in fund value over the hedge ratio dimension around 10 to 15 years (for the 10 year horizon in QE a small inverted U-shape still is

[26]Note that this does not mean that it is beneficial to invest in a bond only portfolio with duration matching. There is a substantial equity premium which can be seen from the upward slope in the equity-to-total-asset-value dimension.

present). Thus, on average, adding swaps to the portfolio has no influence on the fund performance in the medium run, irrespective of the movement of interest rates.

5.2.2.2 Sharpe Ratio

In order to correct the returns of the funds for their underlying uncertainty, Sharpe ratios are calculated. Figure 25 and 26 depict the returns of the fund divided by the corresponding volatility for the QE and EE period, respectively. The returns are calculated as the log difference of fund value over time. The risk-free rate used in the Sharpe ratio is the rate which is used to discount the liabilities (simulated interest rates). Data that lies at the basis of the figures can be found in Table 16 and 17 in the Appendix. NOTA BENE: the axes are mirrored compared to the other figures. This is done to make the pattern better visible.

In the short run, for both regimes, the highest Sharpe ratios can be found at funds with a modest hedge ratio. In the previous paragraph it was shown that a high hedge ratio in times of QE for equity intensive funds did worse relative to a similar fund with a lower hedge ratio in the short run. And indeed, Figure 25 shows that the Sharpe ratios are highest for the funds that maintain a hedge ratio around 40 percent. The fact that the results show the same pattern as the fund value measures means that the portfolio volatility is similar over the hedge ratio dimension. However, fixed income intensive funds have lower volatility in their portfolio. As a consequence, the differences in hedge ratio for such funds are more pronounced.

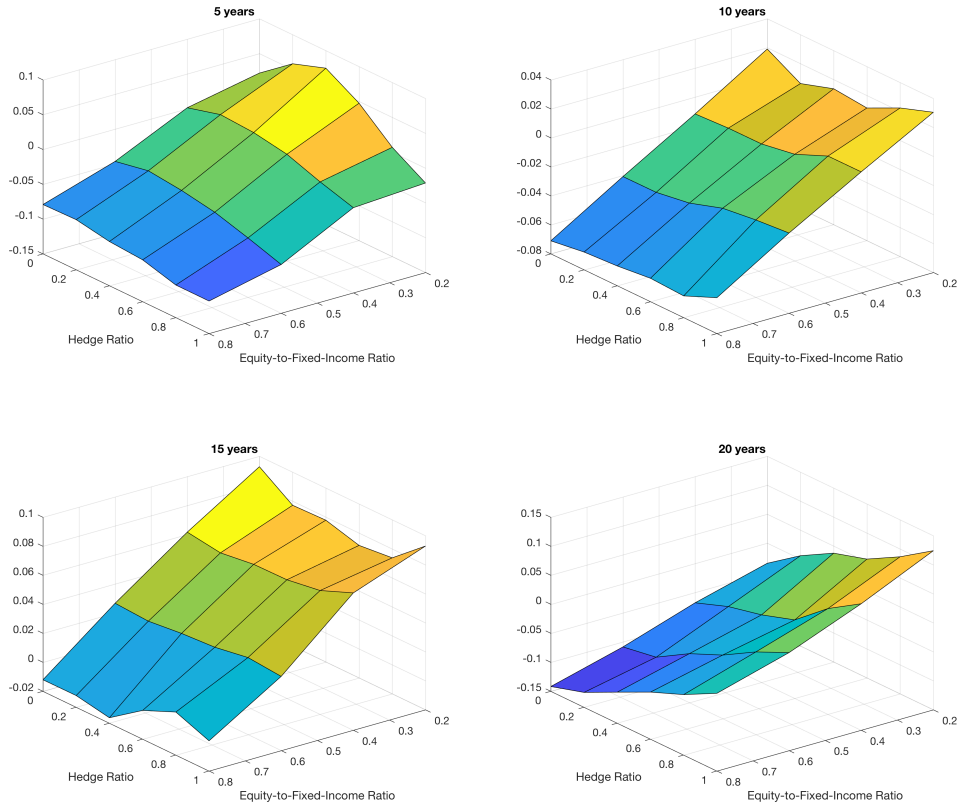


Figure 25: Sharpe Ratio During QE

In line with the previous paragraph, the hedge ratio does not seem to influence the Sharpe ratio for the 10 and 15 year horizons for the QE period that much. The amount of extra volatility the swap agreements bring into the portfolio exactly offset their possible return. For the EE period, higher hedge ratios denote higher Sharpe ratios at these horizon due to the fact that interest rates do not experience a shift from decline to growth, but rather tend to rise consistently. Again, the differences in Sharpe ratio over the hedge ratio dimension are largest for the fixed-income intensive funds due to their lower portfolio volatility.

In the long run, Sharpe ratios are highest for funds that maintain a full hedge. The mechanism is the same as for the high fund value. Again the differences are more pronounced when comparing funds with a low equity-to-total-asset-value ratio.

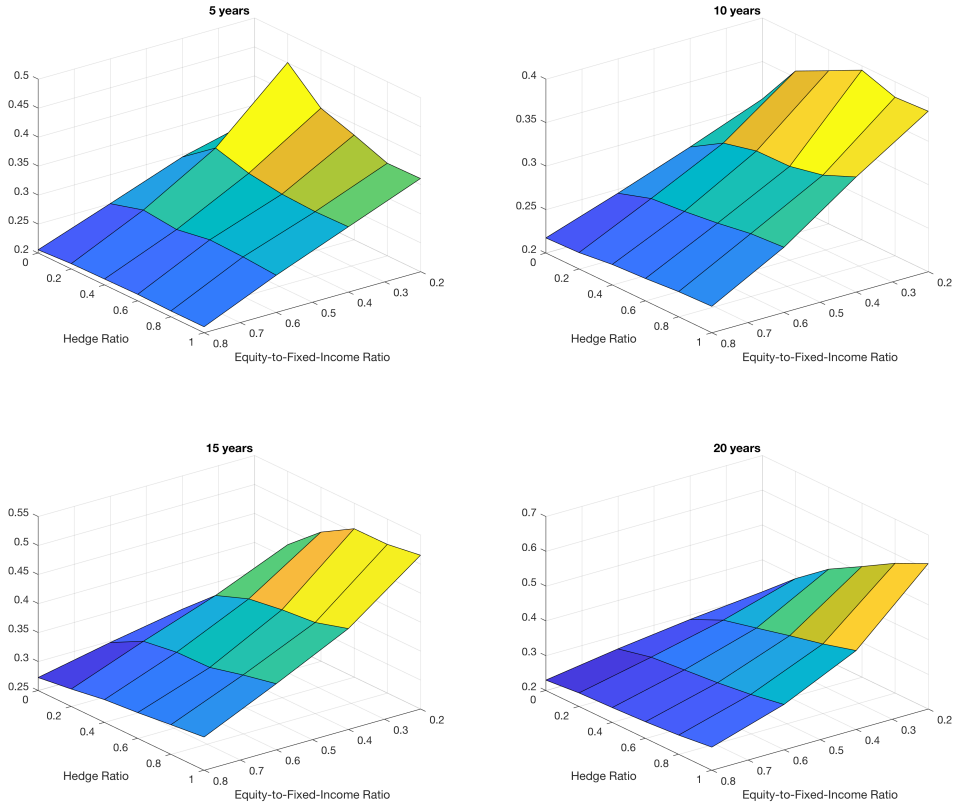


Figure 26: Sharpe Ratio in EE

For both regimes it holds that an intermediate hedge level is recommended. No hedging at all is undesirable as the portfolio does not benefit from a decrease in interest rates. However, a full hedge is not profitable either as this strategy might yield large negative returns when the yield curve experiences a positive shift. Long run effects point out that a full hedge is preferred. However, the validity of these results are again debatable as business cycles do not (all) last 20 years.

5.2.3 Fund Stability

In this section I present the impact of swaps on the stability of the fund. First, I discuss the stability of the funding ratio. Second, I turn to the insolvency probabilities.

5.2.3.1 Funding Ratio

The presented funds all start with the same liabilities. Moreover, not taking into account the indexation granted on the pension benefits, the liabilities are equal for every fund. Therefore, the funding ratios reflect the very same pattern as the pattern in fund value. The funding ratios for the QE and EE period are shown in Figure 36 and 37 in the Appendix, respectively.

The stability of the funding ratio is more interesting. Figure 27 and 28 show the volatility in the funding ratio for the QE and EE period, respectively. The data used to create the figures can be found in Table 18 and 19 in the Appendix.

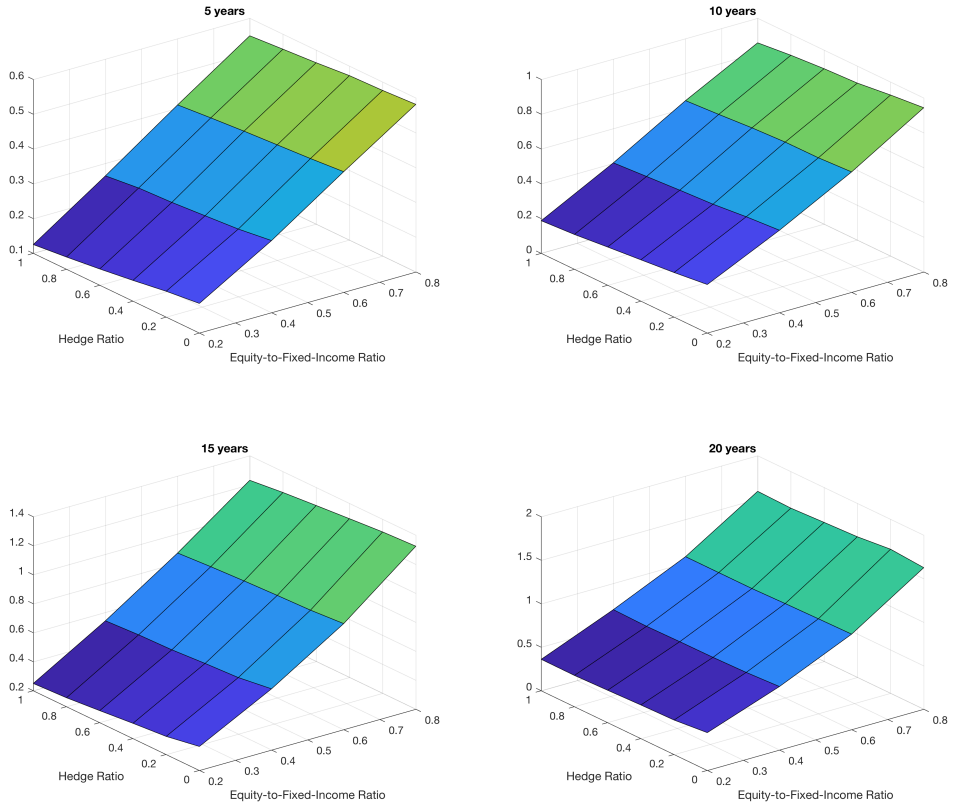


Figure 27: Volatility of Funding Ratio During QE

Unsurprisingly, the variability in the funding ratio is higher for funds that invest heavily into equities. Comparing funds with an equity-to-total-asset-value ratio of 80 percent to those with a ratio of 20 percent, the difference in variability is roughly a factor five. This pattern holds for both regimes and all horizons.

Interestingly, even though the volatility in funding ratio declines with the amount of hedging, the differences are very small, negligible almost. Moreover, this consistency in volatility over the hedge

ratio dimension holds at every horizon. Thus, speaking from a probability theory point of view, the use of swaps shift the distribution to another epicentre, keeping the variance more or less constant. Strictly speaking, the use of swaps does not create more certain fund results. The use of swaps just generates a windfall gain that counters the movement in the present value of the liabilities.

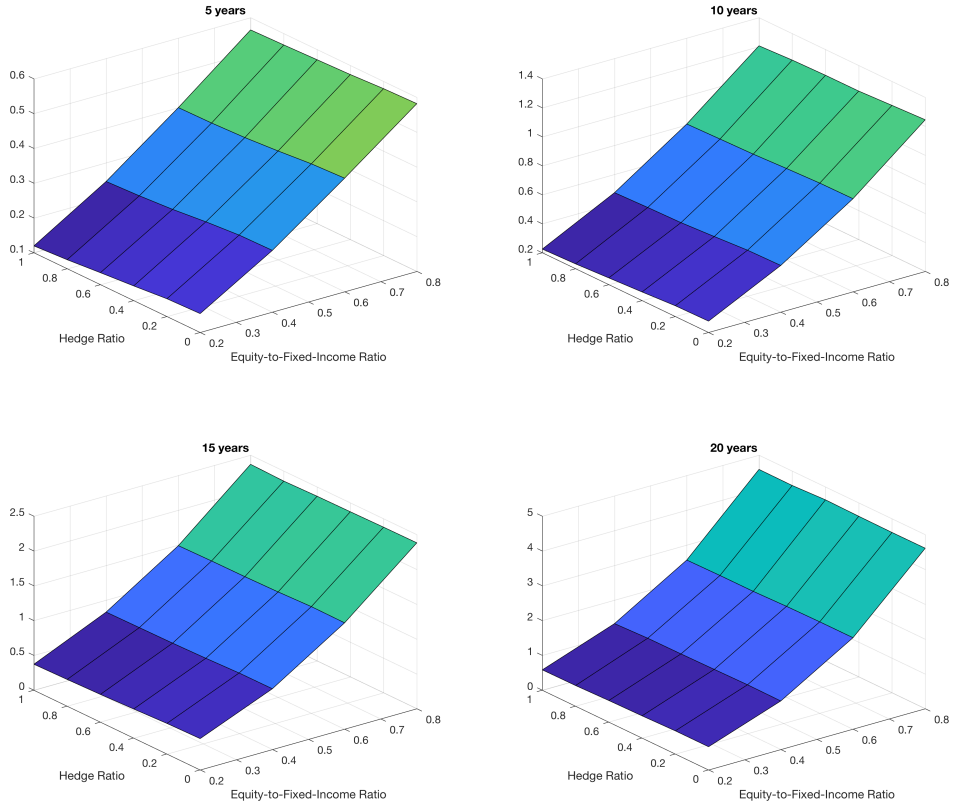


Figure 28: Volatility of Funding Ratio in EE

5.2.3.2 Insolvency Probability

The two main solvency statistics DNB uses are the MVEV and the VEV. Therefore, I present the probabilities that a fund is short to meet its MVEV and VEV.

Figure 29 and 29 depict the probabilities of a fund having a shortfall relative to their MVEV in times of QE and EE, respectively. The data used to construct the figures can be found in Table 20 and 22 in the Appendix. Note the differences in scale for both figures. As the MVEV for all modelled funds is the same (104.5%) the results are driven by the performance of the fund. Whereas in times of QE funds have a significant probability (in the 20-40% range) of having a shortfall relative to the MVEV, these probabilities do not exceed the 10% for the most risky funds in times of EE. Lower shortage probabilities are found at funds that have more certain (read less volatile) returns on average. And indeed, the figures are the mirror image of the fund value figures stated above. The mechanisms that drive these results are the exact same as stated in the Fund Performance section.

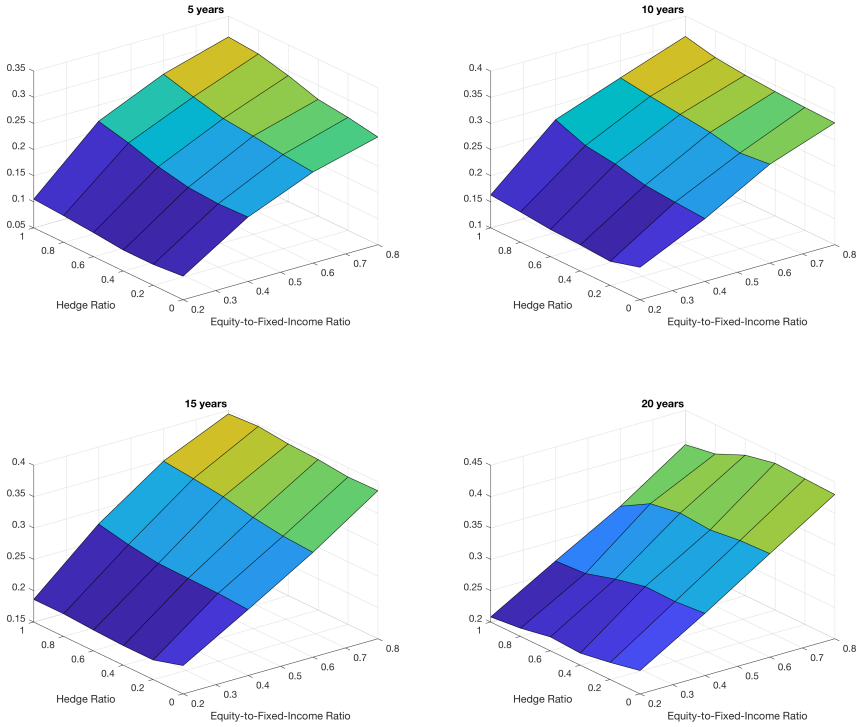


Figure 29: Probability Insolvency Relative to MVEV during QE

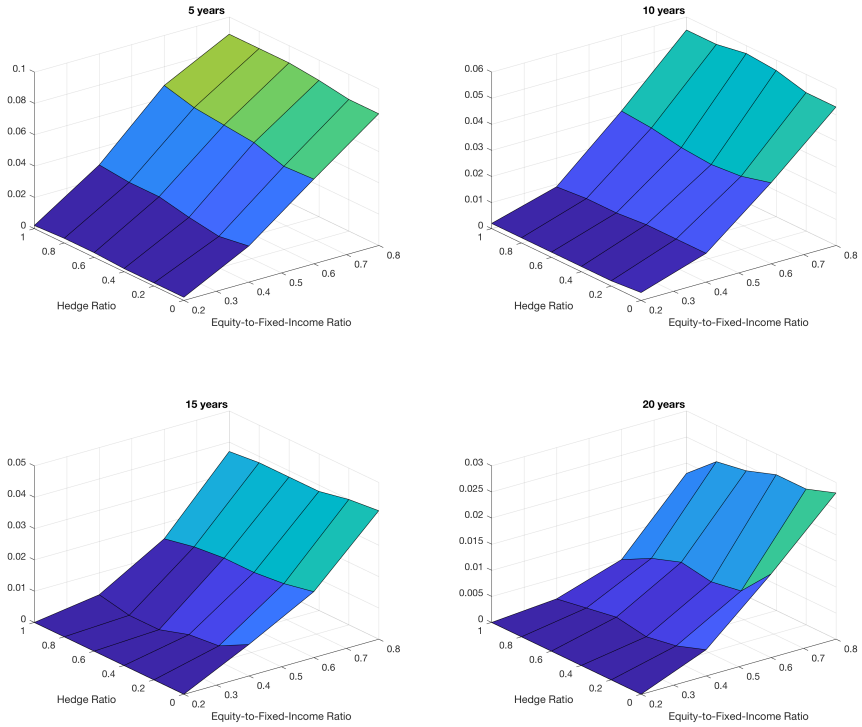


Figure 30: Probability Insolvency Relative to MVEV in EE

More interesting are the probabilities of having less funds than the VEV prescribes. Note that funds with a low equity-to-total-asset-value ratio (high fixed-income-to-equity ratio) have a lower VEV in the model relative to reality. This is the consequence of the assumption that the bonds a fund can invest in are prime grade government bonds for which extra buffers are not required by DNB. Figure 31 and 32 show the insolvency probabilities relative to the VEV for the QE and EE period, respectively. Note that the scale for both figures is very different. Whereas in times of economic growth probabilities of having a VEV shortage are very small, there is a substantial probability for such events in times of economic downturn. For both regimes it holds that more equity-intensive funds have a higher probability of having a shortage relative to their VEV. This result is intuitive as the VEV is relative high (for equity intensive funds) and equities are relative risky. Furthermore, funds with higher hedge ratios have a smaller probability of being short to meet their VEV. This is straightforward as funds with higher hedge ratios are required to hold smaller buffers.

However, during the unstable QE period, for the shorter horizons (5 and 10 years), equity-intensive funds that are highly hedged are insolvent more frequently, even though they are required to hold smaller buffers compared to their less hedged counterparts. Equity-intensive funds have a smaller natural hedge as they invest less into bonds. As such, their use of swaps is more intensive to maintain their desired hedge level compared to fixed-income intensive funds. This result shows that swaps are volatile and can have adverse effects to the portfolio value (due to transaction costs and returns on the swap portfolio). This holds especially when interest rates increase. This result is in line with the results presented in the Fund Performance section.

During the EE period the hedge ratio does not seem to alter the insolvency probabilities on the shorter horizons. As the volatility of the market is relative low in this period the fund adapts its hedge position less frequent relative to the more volatile QE period. As such, the effects are less pronounced in the EE period than the QE period. Moreover, the modelled funds mainly invest in floating-for-fixed agreements. Such contracts are of use when interest rates decline, but will have a negative effect on the portfolio value when interest rates rise. Whereas in the previous paragraph it was shown that a high hedge ratio can have negative effects on stability when the market is less stable, now the negative effect on fund value of the swap portfolio is offset by the positive returns on other asset classes such that insolvency probabilities remain similar.

Comparing the probabilities of having a shortage relative to MVEV and those relative to VEV, the VEV probabilities are the amplified versions of the MVEV probabilities. This is not surprising, as the VEV is a measure to bring stability into pension funds by setting buffers. However, a higher buffer increases the probability of passing it.

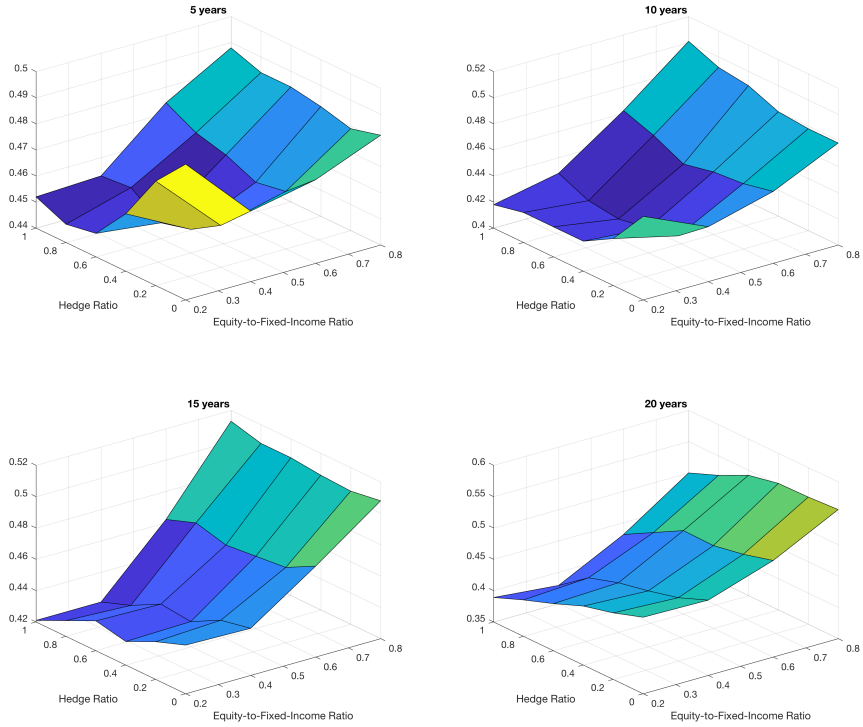


Figure 31: Probability Insolvency Relative to VEV during QE

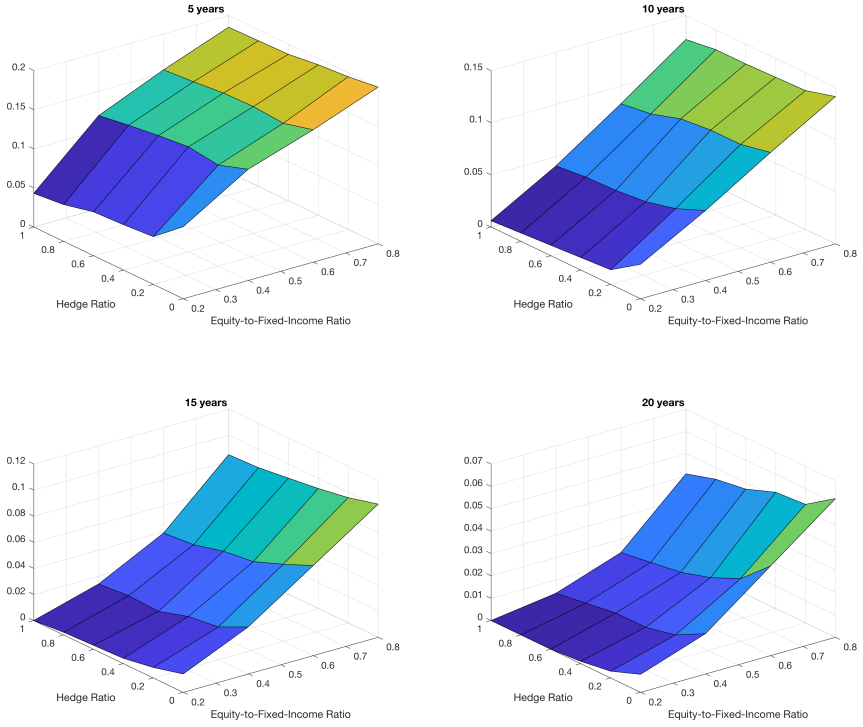


Figure 32: Probability Insolvency Relative to VEV in EE

5.2.4 Indexation

The price level correction that can be granted is determined by a piecewise linear function, see Section 3.1 on the assumptions regarding the liabilities. This function has as arguments the funding ratio and the VEV. As such, the granted level of indexation is a function of performance and stability of the fund. However, it is important to note that the stability is from a regulative point of view (as the VEV is determined/approved by DNB). Thus, the funds that are able to grant a maximal price level correction (of 2%, see section 3.1) are regarded most reliable by DNB.

Figure 33 and 34 depict the granted levels of indexation in times of QE and EE, respectively. The data used to construct the figures can be found in Table 24 and 25 in the Appendix. Note that for the EE period the indexation levels are almost indistinguishable over the hedge ratio dimension. The general pattern in the short run is that the higher the hedge level and the more the fund invests in fixed income securities, the higher the level of indexation. This is in line with the risk assessment of pension funds by DNB: a higher hedge ratio and more fixed income relative to equities is less risky, see the square root formula in Section 3.1. However, for the QE period, in the short run the level of indexation takes a dip for higher hedge ratios. This is the consequence of the negative short run returns of high hedge practices in episodes of low and declining interest rates.

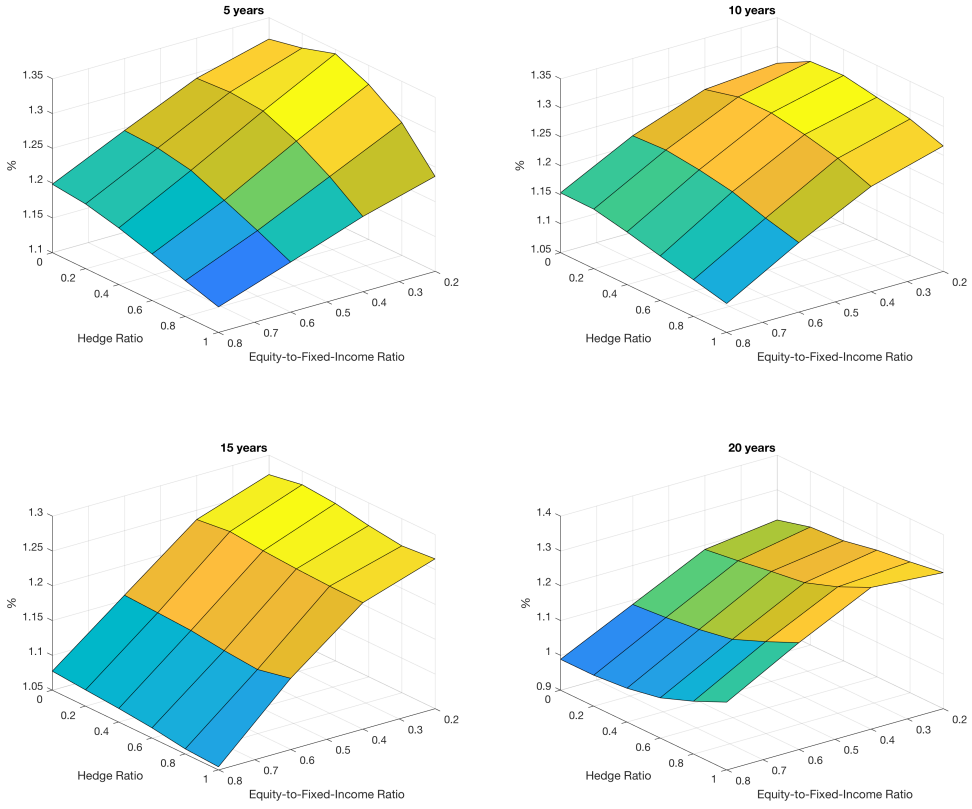


Figure 33: Indexation during QE

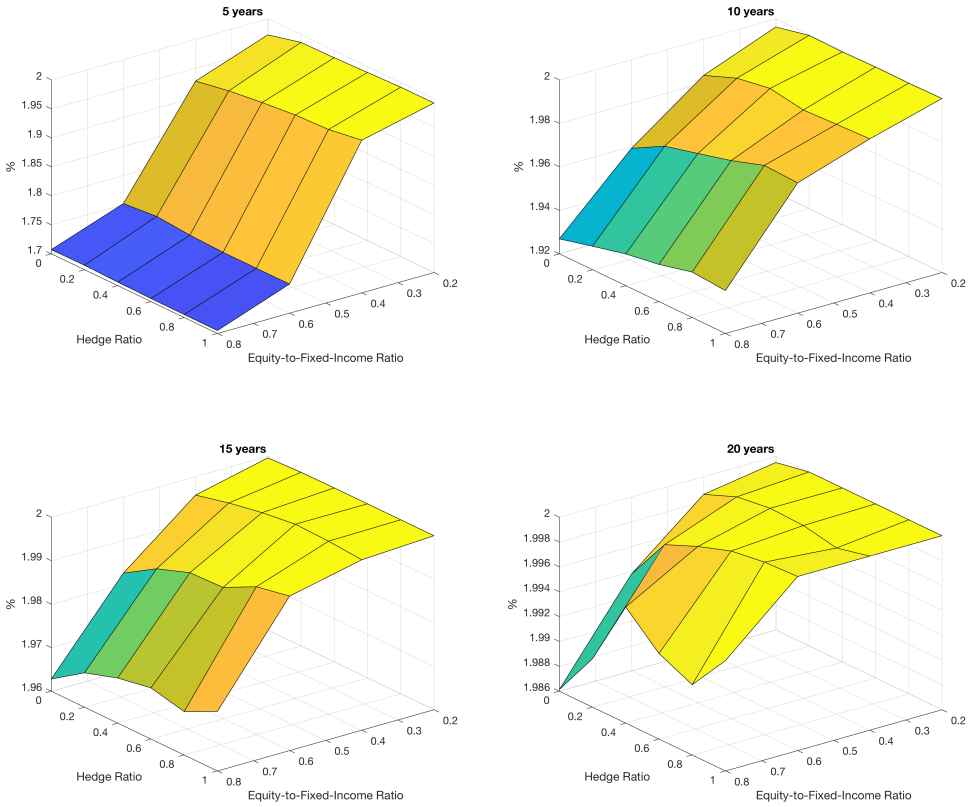


Figure 34: Indexation in EE

5.2.5 Robustness

The first robustness test I performed is to check how the results change under different constant spread specifications. That is, instead of the average spread added to the zero curve, I added a spread $\mathcal{S}$. The robustness tests performed included a zero spread and a spread that was double as large.
The results were robust in the sense that the patterns presented remained the same. The results are not presented for conciseness and can be requested at the author.

The second robustness check is an alternative transaction cost specification. Except for transaction cost part of the results, the results are similar. The transaction costs still increased in the level of hedging. However, the overall size of transaction costs expenditures inclined (declined) with an increase (decrease) of transaction costs. Again, exact results can be obtained at the author.

6 Conclusion

Interest rate hedging is paramount in DNB's nFTK. Indisputably, swaps are useful for hedging purposes, but the costs of using swaps for pension funds had not yet been investigated. As the duration of the pension liabilities typically is 20 years the usefulness of swaps for pension funds might disappear if the term-structure is mean reverting. In answering this main question I investigated different important aspects of a typical pension fund.

On average, pension funds with a hedge ratio of at least 80% spend between 5-9 basis points (bps) of total fund value on total transaction costs in periods of QE, while funds that do not hedge merely spend a half basis point. In times of economic growth this difference diminishes. In such episodes, highly hedged funds spend between 1-1.5 bps on transaction costs versus 0.1-0.5 bps for non hedged funds. There are two mechanisms that explain this result. First, if swap payments eat up all cash reserves, the fund needs to sell off some assets to replenish their 1% cash reserve and thus incur some transaction costs. Second, adding swaps to the portfolio translates to more interest risk exposure (rho) of the portfolio. Adhering to the fund's investment strategy, e.g. the equity-to-total-asset-value ratio, means more frequent rebalancing. The increased market volatility during QE explains why this difference is larger in such episodes than in moments of EE. The results indicate a strong positive effect between transaction costs expenditures and the hedge ratio.

During periods of QE, funds with higher hedge ratios perform less on average. In the long run differences in fund performance over the hedge ratio dimension is almost non-existent. In the EE period the fund value is almost the same over the hedge ratio dimension. Interestingly, for both regimes, there are no clear differences in fund value over the hedge ratio dimension around 10 to 15 years. On average, adding swaps to the portfolio has no influence on the fund performance in the medium run, irrespective of the movement of interest rates. Although there seems to be a negative relationship between the volatility of the funding ratio and the hedge ratio, this relation is negligible. This result holds for all horizons. Hence, the Sharpe ratios point towards the same conclusion as those based on the fund performance results.

As a higher hedge ratio is regarded safer in the nFTK, DNB requires heavily hedged funds to maintain a smaller buffer. Hence, funds with higher hedge ratios have a smaller probability of being short to meet their MVEV. Whereas in times of economic growth probabilities of having a VEV shortage are very small, there is a substantial probability for such events in times of economic downturn. However, during the unstable QE period, equity-intensive funds that are highly hedged are more frequent insolvent on the shorter horizons (5 and 10 years), even though they are required to hold smaller buffers compared to their less hedged counterparts. Their use of swaps is relatively more intensive to maintain their desired hedge level compared to fixed-income intensive funds. This result shows that swaps are volatile and can have adverse effects to the portfolio value. During the EE period the hedge ratio does not seem to alter the insolvency probabilities drastically. As the volatility of the market is relative low in this period the fund adapts its hedge position less frequent relative to the more volatile QE period. As such, the effects are less pronounced in the EE period than the QE period.

In times of economic upswing the indexation levels are almost indistinguishable over the hedge ratio dimension. This is the result of the strong positive returns of the fund that enables it to grant full indexation. The general pattern is that the higher the hedge level and the more the fund invests in fixed income securities, the higher the level of indexation. This is in line with the risk assessment of pension funds by DNB: a higher hedge ratio and more fixed income relative to equities is less risky. However, for the QE period, in the short run the level of indexation takes a dip for higher hedge ratios. This is the consequence of the negative short run returns of high hedge practices in episodes of low and declining interest rates.

The results constitute a Keynesian answer to the main research question. Even though the results did not show a particular argument pro hedging in times of economic expansion, a modest hedge ratio still is advisable: a reasonable hedge position does not alter transaction costs dramatically, nor does it affect the fund performance. However, if a negative shock to interest rates occurs, the fund is partly defended. There appears to exist an interior region on the hedge ratio dimension that turns out to be a wise strategy. This non-linear (but inverted U-Shaped) relationship is not reflected in nFTK and can be a point of improvement in Dutch pension fund regulation. On the other hand, maintaining a high hedge position is very costly in the QE period. Not only do transaction costs expenditures increase drastically in the hedge ratio, but a high hedge position also limits the profitability of the fund when the economy experiences a transition to a regime with higher interest rates. Furthermore, as a lower bound exists in the negative rate territory, hedging against a decline in the term-structure might seen redundant: the costs of the hedge are too high compared to the potential windfall gain. Moreover, adding swaps to the portfolio does not stabilize the funding ratio of the pension fund. At horizons of 10 to 15 years, identical funds that differ in the hedge ratio only are indistinguishable in financial performance and stability. Hence, in economic slump regulation might be eased.

7 Discussion and Further Research

A first interesting avenue for further research is to test the robustness of the results under different specifications of the financial market model. Allowing for more factors in the term-structure models potentially better capture the correlation structure between the maturities. This would be particularly interesting for the periods when correlations across the yield curve are not that large - periods of economic upswing. A model that includes jumps could reveal the true usefulness of swaps for pension funds. Although jumps do not occur frequently, it is interesting how the model processes the occurrence of a jump. A multi-curve framework would relax the constant spread assumption and provide a valuable generalization. Of course, this comes at the price of extra parameters and computing time.

DNB requires funds that do not meet their VEV to create a conditional restructuring plan. Such conditional plans can consist of increased pension premiums, restriction on price level correction on the pension benefits, changing investment behaviour, etc.. The restructuring is conditional on funds being short to their VEV, and becomes unconditional when the fund does not meet the MVEV for the fifth year in a row. Restructuring plans are designed to bring stability within funds, but have consequences on society. If a fund proposes to stick to its investment strategy and change the pension premiums and/or benefits and DNB agrees to the proposal, then the restructuring plan has a wealth distributive (from the young generation to the old, or vice versa) character. Thus, the outcome has a welfare effect on society. However, if the fund alters its investment strategy and keeps its operational characteristics the same, funds that are regarded as unsafe disappear in the long run. In itself, this is a desirable evolution.

But from another point of view, it is very likely that funds converge to an investment strategy that is regarded safe by DNB. From the analysis above it seems that funds with a low to intermediate equity-to-total-asset-value ratio (0.2-0.4) and intermediate to high hedge ratio (≥ 0.5) will remain in the economy. And indeed, such type of funds are the norm in the Dutch pension sector[27][28]. An important point regarding this convergence of characteristics is the build up of systemic risk. One could even advocate that this constitutes a volatility paradox as in Brunnermeier & Sannikov (2014). Pension funds are the largest institutional investors in the Netherlands. With regulation, DNB tries to stabilize the fund. However, if the result is that funds become comparable in investments and interest rate exposure, they go sour simultaneously. Thus, idiosyncratic risk is exchanged for systemic risk.

Regulation that switches in strictness between the regimes might counter this convergence. Strict regulation in economic boom limits idiosyncratic risk. However, as results point out, all funds where able to grant almost full indexation. As such, convergence to a certain type of fund is prevented and systemic risk is untouched. However, easing regulation in times of economic slump such that funds restructure - but not all to a identical fund - can counter the aforementioned volatility paradox. This is a point for further research.

[27] https://statistiek.dnb.nl/downloads/index.aspx#/details/derivatenposities-naar-type/dataset/a02b1607-cc2c-4bdf-9887-2503608521bf/resource/1feee170-cf48-40cf-8b36-a09f82db9938

[28] https://www.dnb.nl/nieuws/nieuwsoverzicht-en-archief/dnbulletin-2013/dnb295970.jsp

8 References

Aihara, S., Bagchi, A., Saha, S. (2008). Estimating volatility and model parameters of stochastic volatility models with jumps using particle filter. In *Proceedings of the 17th World Congress The International Federation of Automatic Control*, 6490-6495.

Aihara, S., Bagchi, A., Saha, S. (2012). Identification of Bates stochastic volatility model by using non-central chi-square random generation method. In *Proceedings of the 2012 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, 3905-3908. USA: IEEE Computer Society. DOI: 10.1109/ICASSP.2012.6288771

Anderson, L. (2007). Efficient Simulation of the Heston Stochastic Volatility Model. *Journal of Computational Finance*, 11, 1–42.

Babbs, S. H., & Nowman, K. B. (1999). Kalman Filtering of Generalized Vasicek Term Structure Models. *Journal of Financial and Quantitative Analysis*, 34(1), 115–130.

Bakshi, G. S., Cao, C., Chen, Z. W. (1997). Empirical Performance of Alternative Option Pricing Models. *Journal of Finance*, 52(5), 2003–2049.

Barndorff-Nielsen, O. E. (2001). Superposition of Ornstein-Uhlenbeck Type Processes. *Theory of Probability and its Applications*, 46, 175–194.

Barndorff-Nielsen, O. E. & Shephard, N. (2001). Non-Gaussian Ornstein–Uhlenbeck-Based Models and Some of Their Uses in Financial Economics. *Journal of the Royal Statistical Society, Series B*, 63, 167–241.

Bates, D. (1996). Jumps and Stochastic Volatility: Exchange Rate Processes Implicit in Deutsche Market Options. *Reviews of Financial Studies*, 9, 69-107.

Bernanke, B., Gertler, M., Gilchrist, S. (1999). The financial accelerator in a quantitative business cycle framework, *Handbook of Macroeconomics*, Elsevier, Volume 1, Part C, 1999, 1341-1393.

Bernanke, B. and Kuttner, K. (2005). What Explains the Stock Market's Reaction to Federal Reserve Policy?. *The Journal of Finance*, 60, 1221-1257.

Bikker, J., Broeders, D., and De Dreu, J. (2007). Stock market performance and pension fund investment policy: rebalancing, free float, or market timing? *DNB Working Paper*, 154.

Black, F. & Scholes, M. (1972). The Valuation of Option Contracts and a Test Of Market Efficiency. *Journal of Finance*, 27, 399–418.

Black, F. & Scholes, M. (1973). The Pricing of Options and Corporate Liabilities, *Journal of Political Economy*, 81, 637–659.

Brennan, M. & Schwartz, E. S. (1980). Analyzing Convertible Bonds. *Journal of Financial and Quantitative Analysis*, 15(4), 907-929.

Broadie, M. & Kaya, Ö. (2006), Exact Simulation of Stochastic Volatility and Other Affine Jump Processes. *Operations Research*, 54(2), 217-231.

Broeders, D., Jansen, K. & Werker, B. (2017). Pension fund's illiquid assets allocation under liquidity

and capital constraints. *DNB Working Paper*, 555.

Brown, S. J. & Dybvig, P. H. (1986). The Empirical Implications of The Cox, Ingersoll, Ross Theory Of The Term Structure of Interest Rates. *Journal of Finance*, 41, 617—30.

Brunnermeier, M. & Sannikov, Y. (2014). A Macroeconomic Model with a Financial Sector. *American Economic Review*, 104 (2): 379-421.

Caballero, R., Farhi, E., Gourinchas, P. (2015). Global Imbalances and Currency Wars at the ZLB. *NBER Working Papers 21670*.

Chakravarty, S. & Sarkar, A. (1999). Liquidity in U.S. Fixed Income Markets: A Comparison of the Bid-Ask Spread in Corporate, Government and Municipal Bond Markets. *FRB of New York Staff Report*, 73, Available at SSRN: *http://dx.doi.org/10.2139/ssrn.163139*.

Chan, K. C., Karolyi, G. A., Longstaff, F. A., Sanders, A. B. (1992). An Empirical Comparison of Alternative Models of the Short-Term Interest Rate. *Journal of Finance*, 47(3), 1209-27.

Chernov, M., Gallant, A. R., Ghysels, E., Tauchen, G. (2003). Alternative Models of Stock Price Dynamics. *Journal of Econometrics*, 116, 225–257.

Courtadon, G. (1982). The Pricing of Options on Default-Free Bonds. *Journal of Financial and Quantitative Analysis*, 17(1), 75-100.

Cox, J.C., Ingersoll, J.E. and Ross, S.A. (1985). A Theory of the Term Structure of Interest Rates. *Econometrica*, 53, 385–407.

Dai, Q. and Singleton, K. J. (2000). Specification Analysis of Affine Term Structure Models. *The Journal of Finance*, 55: 1943-1978.

Dert, C.L. (1995) . Asset Liability Management for Pension Funds: A Multistage Chance Constrained Programming Approach. Erasmus University Rotterdam.

Donohue, C. & Yip, K. (2003). Optimal Portfolio Rebalancing with Transaction Costs. *The Journal of Portfolio Management*, 29(4), 49-63.

Doucet, A., Godsill, S. & Andrieu, C. (2000). On Sequential Monte Carlo Sampling Methods for Bayesian Filtering. *Statistics and Computing*, 10(3), 197–208.

Drijver, S. J., Klein Haneveld, W. K., & van der Vlerk, M. H. (2002). ALM model for pension funds: numerical results for a prototype model. *SOM Research Reports*, No. A44. Groningen: University of Groningen, SOM research school.

Einstein, A. (1905). Über die von der Molekularkinetischen Theorie der Wärme Geforderte Bewegung von in Ruhenden Flüssigkeiten Suspendierten Teilchen. *Annalen der Physik*, 322, 549-560.

Elkamhi, R., Groba, J., Nolte, I. (2017), All Things both Great and Small: Transaction Cost Persistence in Corporate Bonds. *Working Paper*. Available at: `http://www.efmaefm.org/0EFMAMEETINGS/EFMA%20ANNUAL%20MEETINGS/2017-Athens/papers/EFMA2017_0258_fullpaper.pdf`

Engle, R. & Ng, V. K. (1993). Measuring and Testing the Impact of News on Volatility. *Journal of Finance*, 48(5), 1749-1778.

Feller, W. (1951). Two Singular Diffusion Problems. *Annals of Mathematics*, 54(1), 173–182.

Froot, K. (1989). New Hope for the Expectations Hypothesis of the Term Structure of Interest Rates. *The Journal of Finance*, 44(2), 283–305.

Garleanu, N. & Pedersen, L. (2009). Dynamic Trading With Predictable Returns and Transaction Costs. *NBER Working Paper Series 15205*, Available at: `http://www.nber.org.eur.idm.oclc.org/papers/w15205.pdf`

Glasserman, P. (2003). *Monte Carlo Methods in Financial Engineering*. New-York, U.S.A.: Springer-Verlag.

Glasserman, P. & Kim, K. (2011). Gamma Expansion of the Heston Stochastic Volatility Model. *Finance and Stochastics*, 15(2), 267–296.

Grzelak, L.& Oosterlee, C. (2010). On the Heston Model with Stochastic Interest Rates. *SIAM Journal on Financial Mathematics*, 2, 255–286.

Guidolin, M., Thornton, D. (2008). Predictions of Short-Term Rates and the Expectations Hypothesis of the Term Structure of Interest Rates. *European Central Bank Working Paper Series*, 977.

Haastrecht, A. & Pelsser, A. (2010). Efficient, Almost Exact Simulation of the Heston Stochastic Volatility Model. *International Journal of Theoretical and Applied Finance*, 13(1), 1-44.

Heston, S. L. (1993). A closed-Form Solution for Options With Stochastic Volatility With Applications to Bonds and Currency Options. *The Review of Financial Studies*, 6, 327–343.

Ho, T.S.Y. & Lee, S.B. (1986). Term Structure Movements and Pricing Interest Rate Contingent Claims. *Journal of Finance*, 41, 1011–1029

Hogan, M. & Weintraub, K. (1993). The Log-Normal Interest Rate Model and Eurodollar Futures, *Working Paper Citibank, New York*.

Hull, J.C. and White, A. (1990). Pricing Interest-Rate-Derivative Securities, *Review of Financial Studies*, 3(4), p. 573-592.

Hull, J.C. (1996). Using Hull-White Interest Rate Trees. *Journal of Derivatives*, Available at SSRN: `https://ssrn.com/abstract=7242`.

Johnson, J. & Shanno, D. (1987). Option Pricing When the Variance is Changing. *Journal of Financia1 and Quantitative Analysis*, 22 , 143–151.

Kahl, C. & Jäckel, P. (2006). Fast Strong Approximation Monte-Carlo Schemes for Stochastic Volatility Models. *Journal of Quantitative Finance*, 6, 513–536.

Kalman, R. E. (1960). A New Approach to Linear Filtering and Prediction Problems. *Journal of Basic Engineering*, 82, 35-45.

Kaschützke, B. and R. Maurer (2016). Investing and Portfolio Allocation for Retirement, in Piggott, J. and A. Woodland (eds.), *Handbook of Economics of Population Ageing*, Elsevier, Volumn 1B, Chapter 9.

Kienitz, J. & Kammeyer, H. (2009). An Implementation of the Hybrid-Heston-Hull-White Model. *Deutsche Postbank AG Internal Paper*. Available at SSRN: `https://ssrn.com/abstract=1399389`

Kim, I. J. & Kim, S. (2004). Empirical Comparison of Alternative Stochastic Volatility Option Pricing Models: Evidence from Korean KOSPI 200 Index Options Market. *Pacific-Basin Finance Journal*, 12, 117–142.

Krugman, P (1998). It's Baaack: Japan's Slump and the Return of the Liquidity Trap, *Brookings Papers on Economic Activity, Economic Studies Program, The Brookings Institution*, 29(2), 137-206.

Liu, J. & West, M. (2001). Combined Parameter and State Estimation in Simulation-Based Filtering. In Doucet, A., Freitas, N. de, Gordon, N. *Sequential Monte Carlo Methods in Practice*. New-York, U.s.A.: Springer-Verlag.

Longstaff, F. A. (1989). Temporal Aggregation and the Continuous-Time Capital Asset Pricing Model. *The Journal of Finance*, 44, 871-887.

Lord, R., Koekkoek, R., van Dijk, D. (2010). A Comparison of Biased Simulation Schemes for Stochastic Volatility Models, *Quantitative Finance*, 10(2), 177-194.

Mandelbrot, B. B. (1963) The Variation of Certain Speculative Prices, *Journal of Business*, 36, 392–417.

Mei, X., DeMiguel, V., Nogales, F. (2016). Multi-Period Portfolio Optimization With Multiple Risky Assets and General Transaction Costs. *Journal of Banking & Finance*, 69, 108-120.

Miltersen, K.R., Sandmann K., and Sonderman, D. (1997). Closed Form Solutions for Term Structure Derivatives with Log–Normal Interest Rates. *Journal of Finance*, 52, 409–430.

Moon, K., Seon, J., Wee, I., Yoon, C. (2009). Comparison of Stochastic Volatility Models: Empirical Study on KOSPI 200 Index Options. *Bulletin of the Korean Mathematical Society*, 46, 209-227.

Munnik, J. de & Schotman, P. (1994). Cross Sectional Versus Time Series Estimation of Term Structure Models: Empirical Results for the Dutch Bond Market. *Journal of Banking and Finance*, 18, 997—1025.

Nelson, C.R. & Siegel, A.F. (1987). Parsimonious Modeling of Yield Curves, *Journal of Business*, 60(4), 473–489.

Nelson, D. B. (1990). ARCH Models as Diffusion Approximations. *Journal of Econometrics*, 45, 7–38.

Nelson, D. B. (1991). Conditional Heteroskedasticity in Asset Returns: A New Approach. *Econometrica*, 59(2), 347-70.

Nicolato, E. & Venardos, E. (2003). Option Pricing in Stochastic Volatility Models of the Ornstein-Uhlenbeck Type. *Mathematical Finance*, 13, 445–466.

Nowman, K. B. (1997). Gaussian Estimation of Single-Factor Continuous Time Models of the Term Structure of Interest Rates. *Journal of Finance*, 52, 1695—1706.

Pitt, M., & Shephard, N. (1999). Filtering via Simulation: Auxiliary Particle Filters. *Journal of the American Statistical Association*, 94(446), 590-599.

Rendleman, R. & Bartter, B. (1980). The Pricing of Options on Debt Securities. *Journal of Financial and Quantitative Analysis*, 15, 11–24.

Rom, N. (2013). Calculations in the Hull-White Model. Available at SSRN: `https://ssrn.com/abstract=2306170`

Sandmann, K. & Sondermann, D. (1993). A Term Structure Model and the Pricing of Interest Rate Derivatives. *The Review of Futures Markets*, 12, 391-423.

Sapp, T. (2009). Estimating Continuous-Time Stochastic Volatility Models of the Short-Term Interest Rate: A Comparison of the Generalized Method of Moments and the Kalman Filter. *Review of Quantitative Finance and Accounting*, 33(4), 303-326.

Sarno, L., Thornton, D., Valente, G. (2007). The Empirical Failure of the Expectations Hypothesis of the Term Structure of Bond Yields. *Journal of Financial and Quantitative Analysis*, 42(1), 81–100.

Scott, L. (1987). Option Pricing When the Variance Changes Randomly: Theory, Estimation, and an Application. *The Journal of Financial and Quantitative Analysis*, 22, 419–438.

Scott, L. (1997). Pricing Stock Options in a Jump-Diffusion Model With Stochastic Volatility and Interest Rates: Applications of Fourier Inversion Methods. *Mathematical Finance*, 7, 413–424.

Stein, E. & Stein, J. (1991). Stock Price Distributions With Stochastic Volatility: An Analytic Approach. *Review of Financial Studies*, 4, 727–752.

Taylor, S. J. (2012). Financial Returns Modelled by the Product of Two Stochastic Processes, a Study of Daily Sugar Prices. In F. X. Diebold (Ed.), *Financial Risk Measurement and Management*, 441-464. Cheltenham: Edward Elgar.

Vasicek, Oldrich (1977). An Equilibrium Characterisation of the Term Structure. *Journal of Financial Economics*, 5(2), 177–188.

Wiggins, J. B. (1987). Option Values Under Stochastic Volatility: Theory and Empirical Estimates. *Journal of Financial Economics*, 19, 351–372.

Woodside-Oriakhi, W., Lucas, C., Beasley, J.E. (2013). Portfolio Rebalancing With an Investment Horizon and Transaction Costs, *Omega*, 41(2), 406-420.

Zaiceva, A. & Zimmermann, K. (2016). Chapter 3: Migration and the Demographic Shift, John Piggott & Alan Woodland *Handbook of the Economics of Population Aging*, North-Holland, Volume 1, Pages 119-177.

Zhang, J. E. & Shu, J. (2003). Pricing S&P500 Index Options With Heston's Model, *Proceedings, 2003 IEEE International Conference on Computational Intelligence for Financial Engineering*

Zhu, J. (2011). A Simple and Accurate Simulation Approach to the Heston Model. *The Journal of Derivatives*, 18, 26-36.

Zhu, J. (2012). Multiple-Curve Valuation with One-Factor Hull-White Model. *SSRN Electronic*

Journal. Available at SSRN: http://dx.doi.org/10.2139/ssrn.2049507.

9 Appendix

9.1 Preliminaries

This section serves as a brief summary of the economic and mathematical theory the model rests upon and can be skipped if the reader is familiar with the theory.

9.1.1 Economic Background

9.1.1.1 Interest rates, Compounding and Discounting

A fundamental economic concept in the framework of pension funds is the so-called time value of money. This concept means that the present value of money will be less or equal than the future value of this money due to the interest that can be earned on current money (at least when interest rates are non-negative). Define the value of a bank savings account at time t by $M(t)$. If one puts m at time $t = 0$ in the account the initial value of the bank account is known, say $M(0) = m_0$. Let us now consider the most straightforward type of interest, namely the *simple interest rate*. If the annualized interest rate for depositing money for a period of $(0, t)$ year equals $R(0, t)$, then after a period t years the simple interest rate gained on the money deposited is defined by $R(0, t)\, t$. If the bank account would grow according to simple interest then after t years the money in the bank account would equal:

$$M(t) = (1 + R(0, t)\, t)\, m_0 \tag{9.1}$$

The mechanism of *compounding* accounts for the fact that the interest gained on a deposit can be reinvested and can gain interest on itself. If the annual interest rate for depositing your money for a period of $(0, t)$ equals $R(0, t)$, the value of bank account after time t years, compounded annually, corresponds to:

$$M(t) = (1 + R(0, t))^t\, m_0 \tag{9.2}$$

Then, $R(0, t)$ is called the *zero or spot interest rate* (short: *zero or spot rate*). Interest on bank accounts often are paid more frequently than once a year. If interest is paid n times a year and the annual interest rate for the holding period of $(0, t)$ is $R(0, t)$, the bank account after t years of investment equals:

$$M_n(t) = \left(1 + \frac{R(0, t)}{n}\right)^{nt} m_0 \tag{9.3}$$

The subscript n in M_n denotes the compounding frequency per year. Taking the limit of Equation (9.3) in the compounding frequency n to infinity, yields the value of the bank account after t years with a *continuously compounded spot interest rate*:

$$\lim_{n\to\infty} M_n(t) = \lim_{n\to\infty} \left(1 + \frac{R(0, t)}{n}\right)^{nt} m_0 = e^{R(0,t)t} m_0 \tag{9.4}$$

Let us now consider the evolution of the bank account in continuous time. In this setup the bank account accrues with interest paid continuously, the so-called *short interest rate* (short: *short rate*) $r(t)$. After a infinitesimal step dt the interest paid on the bank account equals:

$$dM(t) = M(t)\, r(t)\, dt \tag{9.5}$$

The initial value put in the account is the initial condition for the ordinary differential equation. The solution to this differential equation is:

$$M(t) = e^{\int_0^t r(s) ds} m_0 \tag{9.6}$$

Combining Equation (9.4) and (9.6) yields the relation between the continuously compounded spot interest rate and the integral of the short rates over the same interval:

$$R(0,t) = \frac{1}{t}\int_0^t r(s)\,ds \tag{9.7}$$

In the calculation of interest rate swap prices *forward interest rates* (short: *forward rate*) play an important feature. A forward rate $f(t,T_1,T_2)$ is the interest rate that is paid over the period (T_1,T_2) and is determined at time t, for $t \leq T_1 \leq T_2$. In order to make the forward rate clear consider the following example. At time t you receive $\$1,-$. You would like to put this money in a savings account for two years. Denote by T_1 and T_2 one and two years from today respectively. Suppose there are two different options to invest the money. Option one is putting the $\$1,-$ in an account for the full two years paying a $R(t,T_2)$ continuous compounded interest. The second option is to save the money for a year in an account gaining continuous compounded interest $R(t,T_1)$ after which the gross return is invested again against the best rate available at different banks. The interest rate you will receive for the second year is the forward interest rate that is determined now over the period from one year from now until two years from today, $f(t,T_1,T_2)$. Assume that this rate is continuously compounded as well. This rate is unknown, but by the absence of *arbitrage*, the two different options should earn the same. With first option the end result equals $\$1 \times e^{R(t,T_2)(T_2-t)}$. By going with the second option the terminal value of the investment is $\$1 \times e^{R(t,T_1)(T_1-t)}e^{f(t,T_1,T_2)(T_2-T_1)}$. Then the forward rate can be expressed in terms of the other two zero rates:

$$\begin{aligned} e^{f(t,T_1,T_2)(T_2-T_1)} &= \frac{e^{R(t,T_2)(T_2-t)}}{e^{R(t,T_1)(T_1-t)}} \\ f(t,T_1,T_2) &= \frac{R(t,T_2)(T_2-t) - R(t,T_1)(T_1-t)}{(T_2-T_1)} \end{aligned} \tag{9.8}$$

With Equation (9.7) the forward rate can be expressed in terms of short rate integrals. Letting $T_2 \to T_1$, the limit of the forward rate is the so-called *instantaneous forward rate*. The instantaneous forward rate is the interest rate that is paid over a infinitesimal holding period somewhere in the future, say T_1 and is determined at time t:

$$f(t,T_1) = \lim_{T_2 \to T_1^+} f(t,T_1,T_2) \tag{9.9}$$

Note that $\lim_{T \to t^+} f(t,T) = r(t)$.

The concept of interest payments is closely related with the mechanism of discounting. Discounting is the method of calculating the present value of a future cash flow. Because of the time value of money (the possibility of earning interest), money now is less in quantity than it is in the future, if interest rates are positive. Assuming that at time T the bank account is worth m_T, then by discounting one calculates the value of the account any time before T, $m_t\ \forall t \leq T$. The values m_t can be found by inverting the growth formulas of the bank account stated before. The list below gives the discount factors for t years from now for every equation:

- simple interest (equation 9.1):

$$DF(0,t) = (1 + R(0,t)\,t)^{-1} \tag{9.10}$$

- annually compounded (equation 9.2):

$$DF(0,t) = (1 + R(0,t))^{-t} \tag{9.11}$$

- n times compounded (equation 9.3):

$$DF(0,t) = \left(1 + \frac{R(0,t)}{n}\right)^{-nt} \tag{9.12}$$

- continuously compounded (equation 9.4):

$$DF(0,t) = e^{-R(0,t)t} \tag{9.13}$$

- in terms of short rate (equation 9.6):

$$DF(0,t) = e^{-\int_0^t r(s)ds} \tag{9.14}$$

9.1.1.2 Bonds

A bond is a type of security sold by corporations and governments to raise funding today in exchange for future payments. A bond is a form of debt. The point up until the payments are made is called the maturity or expiration date of the bond. The remaining time to maturity is called the term of the bond. There are two types of payment, namely coupon payments and the face value payment. The face value (or principal) is the amount that is paid by the seller at the maturity date to the buyer of the bond. Coupons are payments that are made periodically (e.g. quarterly, semi-annually, etc.) and are constant over the lifespan of the bond. Each coupon payment is set by the coupon rate and is calculated as:

$$\text{Coupon payment} = \frac{\text{Coupon Rate } \times \text{ Face Value}}{\text{Number of Coupons per year}}$$

Denote the price of a bond at issue date with maturity T by $P(0,T)$. If this bond has face value N and coupon payments c per period, all the cash flows of buying such a bond can be depicted as in Figure 35.

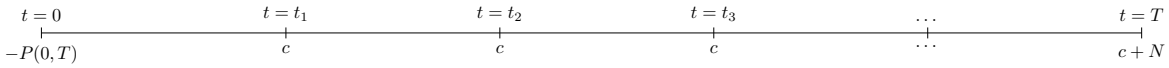


Figure 35: Cash flows of a coupon-paying bond

Consider a bond with maturity T and face value N that pays a coupon of c. Fix a period δ for the interval in between payments. Denote by $M = T/\delta$ the number of payments. Then, the set of payment dates is defined by $T = \{T_n = n\delta : n \in \{1, 2, \dots M\}\}$. Define by $T_0 = 0$ the issue date of the bond. The price of a bond with coupon payments of c and maturity T at time t, $P(t,T)$, equals the discounted future cash flows. At time t the first upcoming coupon payment occurs at T_n, where $n = \{n \in \{1, 2, \dots M\} : T_{n-1} < t \leq T_n\}$. Then, the bond price, maturing at T, at time t equals:

$$P(t,T) = \sum_{n=\{n\in\{1,2,\dots M\}:T_{n-1}<t\leq T_n\}}^{M} DF(t,T_n)\, c + N \times DF(t,T) \tag{9.15}$$

A bond's yield to maturity or yield is the interest rate such that the discounted cash flows equal the observed market price. In case of continuous compounding the yield of a coupon-paying bond equals the rate y such that Equation (9.15) equals the observed market price $P^M(t,T)$:

$$P^M(t,T) = \sum_{n=\{n\in\{1,2,\dots M\}:T_{n-1}<t\leq T_n\}}^{M} e^{-y(T_n-t)} c + N \times e^{-y(T-t)} \tag{9.16}$$

Solving Equation (9.16) for y delivers the required result.

A bond with a coupon rate of zero, is called a zero-coupon bond. Denote the price of a zero-coupon bond at time t that matures at time T with a face value of N by $Z_N(t,T)$. The bond consist of one cash flow at the maturity date. Therefore, the price of a zero-coupon bond equals the discounted value of the face value paid at time T. Note that the cash flows of a zero-coupon bond are in fact similar to

those of a savings bank account with corresponding initial and terminal date. With this in mind the price of the zero-coupon bond should equal:

$$Z_N(t,T) = DF(t,T)\,N \tag{9.17}$$

In the case of simple interest the price of the bond is:

$$Z_N(t,T) = (1 + R(t,T)(T-t))^{-1} N \tag{9.18}$$

Where $R(t,T)$ is the zero or spot rate for a holding period of (t,T). The yield of a zero-coupon coincides with the zero rate.

A coupon-paying bond can be decomposed in a batch of zero coupon bonds. Every coupon payment is a zero-coupon bond with face value equal to the coupon payment and with maturity the date that the coupon payment has to be made. The face value payment of the coupon-paying bond equals a zero-coupon bond with same face value and maturity. Therefore, the price of the coupon-paying bond can be expressed as:

$$P(t,T) = \sum_{n=\{n\in\{1,2,\dots M\}:T_{n-1}<t\leq T_n\}}^{M} Z_c(t,T_n) + Z_N(t,T) \tag{9.19}$$

The last type of bond considered in this section is a so-called floating rate bond or floating rate note (FRN). A FRN is a bond with a varying coupon rate. This coupon rate is determined by a certain benchmark interest rate. Typical benchmarks include the US Treasury rates or the London Interbank Offered Rates (LIBOR). At every coupon payment date the coupon rate for the next payment is set. Therefore, in the context of a FRN a coupon date is also called a reset date. Denote the price of a FRN with maturity T at time t by $FRN(t,T)$. Fix a period δ for the interval in between payment dates. Denote by $M = T/\delta$ the number of payments. Then, the set of payment dates is defined by $T = \{T_n = n\delta : n \in \{1,2,\dots M\}\}$. Define by $T_0 = 0$ the issue date of the bond. First consider the value of the FRN at the issue date. The first coupon payment at T_1 is known, since the rate has been fixed. The remaining coupon rates are yet unknown at the issue date. The unbiased estimates for the remaining payments are the corresponding forward rates. Therefore, the coupon rate at time T_2 equals $f(0,T_1,T_2)$. Since the coupons do not compound interest, we should define the forwards based on the simple interest rates. Then, by following the same approach in Equation (9.8):

$$\begin{aligned}
(1 + R(0,T_1)T_1)(1 + f(0,T_1,T_2)(T_2 - T_1)) &= (1 + R(0,T_2)T_2) \iff \\
f(0,T_1,T_2)(T_2 - T_1) &= \frac{1 + R(0,T_2)T_2}{1 + R(0,T_1)T_1} - 1 \iff \\
f(0,T_1,T_2) &= \frac{R(0,T_2)T_2 - R(0,T_1)T_1}{(1 + R(0,T_1)T_1)(T_2 - T_1)}
\end{aligned} \tag{9.20}$$

The value of the FRN equals the discounted values of the coupons and the face value. The discount factors used to calculate the present value should therefore also be based on the simple interest rate.

All together:

$$
\begin{aligned}
FRN(0,T) &= \sum_{i=1}^{M} N \times f(0,T_{i-1},T_i)(T_i - T_{i-1}) DF(0,T_i) + N \times DF(0,T) \\
&= \frac{N \times R(0,T_1) T_1}{1 + R(0,T_1) T_1} + N \left[\sum_{i=2}^{M} \left(\frac{1 + R(0,T_i) T_i}{1 + R(0,T_{i-1}) T_{i-1}} - 1 \right) \frac{1}{1 + R(0,T_i) T_i} \right] + \frac{N}{1 + R(0,T) T} \\
&= \frac{N \times R(0,T_1) T_1}{1 + R(0,T_1) T_1} + N \left[\sum_{i=2}^{M} \frac{1}{1 + R(0,T_{i-1}) T_{i-1}} - \frac{1}{1 + R(0,T_i) T_i} \right] + \frac{N}{1 + R(0,T) T} \\
&= \frac{N \times R(0,T_1) T_1}{1 + R(0,T_1) T_1} + N \left(\frac{1}{1 + R(0,T_1) T_1} - \frac{1}{1 + R(0,T_N) T_N} \right) + \frac{N}{1 + R(0,T) T} \\
&= \frac{N(1 + R(0,T_1) T_1)}{1 + R(0,T_1) T_1} = N
\end{aligned}
\tag{9.21}
$$

In the second line we have used the fact that the first forward rate equals the spot rate, $f(0,0,T_1) T_1 = R(0,T_1) T_1$. From the third to the fourth line we have observed that the summation is a telescoping series. The rest of the result is based on simple algebra.

From the derivation in Equation (9.21) it is clear that the value of the floating rate bond at the issue date does not depend on the level of interest rates, but purely on the notional. This is due to the fact that the coupon rate resets at every payment date. An increase in interest rates will result in a higher coupon payment, but also in a lower discount factor. Hence, the effect of an increase in the interest rates has no effect on the price.

Let us now turn to the price of the FRN at a time $t \in (0,T)$. At time t the first upcoming coupon payment occurs at T_n, where $n = \{n \in \{1,2,\dots M\} : T_{n-1} < t \leq T_n\}$ is the index of the first upcoming payment. At time t the coupon rate for the payment at T_n is already determined at T_{n-1}. This coupon payment is equal to $N \times R(T_{n-1},T_n)(T_{n-1} - T_n)$. For the remaining payments at times $T_{n+1}, T_{n+2}, \dots, T$ the coupon rate is determined from the forward rates. The value of the FRN at time t yields:

$$
\begin{aligned}
FRN(t,T) &= N \times R(T_{n-1},T_n)(T_n - T_{n-1}) DF(t,T_n) + N \left[\sum_{i=n+1}^{M} f(t,T_{i-1},T_i)(T_i - T_{i-1}) DF(t,T_i) \right] \\
&\quad + N \times DF(t,T) \\
&= \frac{N \times R(T_{n-1},T_n)(T_n - T_{n-1})}{1 + R(t,T_n)(T_n - t)} + N \left[\sum_{i=n+1}^{M} \frac{1}{1 + R(t,T_{i-1})(T_{i-1} - t)} - \frac{1}{1 + R(t,T_i)(T_i - t)} \right] \\
&\quad + \frac{N}{1 + R(t,T)(T - t)} \\
&= \frac{N \times R(T_{n-1},T_n)(T_n - T_{n-1})}{1 + R(t,T_n)(T_n - t)} + N \left(\frac{1}{1 + R(t,T_n)(T_n - t)} - \frac{1}{1 + R(t,T)(T - t)} \right) \\
&\quad + \frac{N}{1 + R(t,T)(T - t)} \\
&= \frac{N(1 + R(T_{n-1},T_n)(T_n - T_{n-1}))}{1 + R(t,T_n)(T_n - t)}
\end{aligned}
\tag{9.22}
$$

In the derivation above the same line of reasoning is used as in the derivation of Equation (9.21). Note that the result in 9.22 makes intuitively sense. Because of the reset characteristic of the FRN at a

payment date, the FRN at any time t equals the discounted upcoming payment plus the discounted face value at the same time of the upcoming payment. Furthermore, the result shows that a FRN carries little interest rate risk as the value at t only depends on the discount rate for the upcoming payment. After that time any changes in interest rates will not have any consequence on the price as the coupon rate and discount factor cancel each other out.

9.1.1.3 Interest Rate Swaps

An interest rate swap (short: swap) is a contract between two parties to exchange interest rates for a certain time period. More specifically, these parties exchange interest rate cash flows on a specified amount, called the notional principal. The principal is called notional because this amount is never exchanged, but only determines the size of the payments. One party receives a fixed interest rate, the swap rate, over the agreed upon period in return for a floating rate. This party or side of the swap is called the fixed leg. The other party receives the floating rate and pays the fixed rate. This side is called the floating leg of the swap. In most cases the fixed coupon payments and floating coupon payments are made annually and biannually respectively. At each floating coupon payment the floating coupon rate for the next payment is determined. These dates are called the reset dates.

To value a swap, note that a swap is a combination of two bonds. Consider the value of the swap from the receiver position. The fixed swap rate payments received can be represented by a long position in a coupon-paying bond where the coupon rate equals the swap rate, the face value is the same amount as the notional, and with the same maturity. The floating rate payments can be replicated by a short position in a floating rate bond where again the face value is the same amount as the notional and with the same maturity. Note that the face value exchange in the bond representation cancel each other out. As such, a swap can be perfectly replicated by the two bonds.

In the following the superscripts R and L will represent the fixed and floating side, respectively. Fix two periods δ^R and δ^L for the intervals in between fixed and floating payments, respectively. Denote by T the maturity of the swap and by $M^R = T/\delta^R$ and $M^L = T/\delta^L$ the number of fixed and floating payments respectively. Then, for a swap with maturity T, we define two sets of payment dates $T^R = \left\{T_n^R = n\delta^R : n \in n^R\right\}$, where $n^R = \left\{1, 2, \dots M^R\right\}$ and $T^L = \left\{T_n^L = n\delta^L : n \in n^L\right\}$, where $n^L = \left\{1, 2, \dots M^L\right\}$. Define by $T_0^R = T_0^L = 0$ the issue date of the swap. Note that $T_{M^R}^R = M^R\delta^R = T = M^L\delta^L = T_{M^L}^L$ being the maturity of the swap. Also notice that when fixed rate payments are made annually and floating rate payments biannually that $\delta^R = 2\delta^L$ and $M^R = \frac{1}{2}M^L$. Moreover, note that simple interest rates are used as the payments are fixed and cannot be reinvested. Denote by r the annualized swap rate and by $L_{T_{n-1}^L}\left(T_n^L\right)$ the annualized interest rate determined at T_{n-1}^L for the period $\left[T_{n-1}^L, T_n^L\right], \forall n \in n^L$. Denote by N the notional value of the swap. The value of the fixed rate bond at time $t \in (0, T)$ equals (see Equation (9.15)):

$$B^R(t,T) = Nr\delta^R \sum_{n=\left\{n\in n^R : T_{n-1}^R < t \leq T_n^R\right\}}^{M^R} DF\left(t, T_n^R\right) + N \times DF(t,T) \tag{9.23}$$

The argumentation for the valuation of the short position in the floating rate bond is similar to argumentation of the FRN. Consider an upfront incoming payment of N at $t = 0$. This payment can be used to save against a annualized rate of $L_0\left(T_1^L\right)$ over the period $\left[0, T_1^L\right]$, yielding a total gain of $N\delta^L L_0\left(T_1^L\right)$. This gain exactly equals the coupon that has to be paid (since it is a short position) to the buyer of the bond at T_1^L. What is left is the initial endowment of N. For every period $\left[T_n^L, T_{n+1}^L\right], 1 \leq n < M^L$ this procedure can be repeated. At $T_{M^L}^L$ the last coupon and the face value payment N is paid. Therefore, the value of this floating rate bond at $t = 0$ equals the upfront incoming payment, so $B^L(0,T) = -N$. Since at every coupon payment the initial endowment is preserved, the value will be $-N$ at those points in time, $\forall t \in T_n^L : B^L(0,t) = -N$. Now examine the value of a short position in a floating bond in between reset dates. Consider a $t \in (0,T) \setminus \left\{T^L\right\}$. Then, the index of the

upcoming coupon payment $n = \left\{n \in n^L : T_{n-1}^L < t < T_n^L\right\}$. At time t the upcoming coupon payment at T_n^L is already determined at T_{n-1}^L, being $-N\delta^L L_{T_{n-1}^L}\left(T_n^L\right)$. At the point in time of the coupon payment, T_n^L, the previous procedure of reinvesting can again be repeated. Therefore, the value of the short floating bond equals $-N$ at time T_n^L. The value of the short floating bond at t equals the sum of the discounted value at time T_n^L and the discounted payment at time T_n^L. All together:

$$B^L(t,T) = DF\left(t,T_n^L\right) N\left(-1 - \delta^L L_{T_{n-1}^L}\left(T_n^L\right)\right) \tag{9.24}$$

The value of the swap from the receiver's perspective equals the sum of Equation (9.23) and (9.24):

$$\begin{aligned} V(t,T) &= Nr\delta^R \sum_{n=\left\{n\in n^R : T_{n-1}^R < t \le T_n^R\right\}}^{M^R} DF\left(t,T_n^R\right) + N \times DF(t,T) + DF\left(t,T_n^L\right) N\left(-1 - \delta^L L_{T_{n-1}^L}\left(T_n^L\right)\right) \\ &= N\left(r\delta^R \sum_{n=\left\{n\in n^R : T_{n-1}^R < t \le T_n^R\right\}}^{M^R} DF\left(t,T_n^R\right) + DF(t,T) - DF\left(t,T_n^L\right) - \delta^L L_{T_{n-1}^L}\left(T_n^L\right)\right) \end{aligned} \tag{9.25}$$

To determine the fixed swap rate r, consider the value of the swap at $t = 0$. Since $t = 0$ is a reset date the value of the swap at $t = 0$ is the sum of Equation (9.23) and $-N$, :

$$V(0,T) = N\left(r\delta^R \sum_{n=1}^{M^R} DF\left(0,T_n^R\right) + DF(0,T) - 1\right) \tag{9.26}$$

The swap rate is the value for r that makes the value of the swap at time $t = 0$ equal to zero. For that fixed rate both the receiving and paying party are willing to enter the contract. Equating Equation (9.26) to zero and solving for r yields the swap rate:

$$r = \frac{1 - DF(0,T)}{\delta^R \sum_{n=1}^{M^R} DF\left(0,T_n^R\right)} \tag{9.27}$$

As discussed earlier, floating rate bonds are very insensitive to interest rate changes relative to other types of bonds due to the reset characteristic of the value. Fixed coupon-paying bonds on the contrary are very sensitive to a change in the level of interest rates. Although a swap is the difference between these two bonds, the sensitivity to changes in interest rates remains strong. Therefore, swaps are a useful instrument to hedge interest rate risk. Another attractive feature of swaps is that the value of the swap equals zero at issuance. Therefore, the two parties initiating the swap do not initially have to withhold any capital in order to hedge any possible interest rate risk. The parties only have to reserve cash to make the required payments.

As a last note on swaps: although swaps are very attractive to hedge interest rate risks due to the features discussed above, the volatility of swaps with respect to interest rate changes create an extra volatile component in the portfolio. When one has a strict investment strategy, adding swaps could end up being rather costly. The added swaps to the portfolio could force the portfolio manager to rebalance the portfolio more often, yielding transaction costs, to meet the requirements of the investment strategy.

9.1.2 Mathematical Background

The analysis in this thesis uses numerous mathematical definitions and results. This section is written to provide a basis to the reader with little knowledge of stochastic calculus, and econometric techniques.

9.1.2.1 Stochastic Calculus The stochastic calculus relies heavily on measure theory. As such, this will be the starting point. For a more elaborate discussion on these topics, I refer to Shreve (2004).

Definition: σ-algebra
Let Ω be a non-empty set, and let $\mathcal{F}$ be a collection of subsets of Ω. We say that $\mathcal{F}$ is a σ-algebra when:

(i) $\emptyset \in \mathcal{F}$

(ii) $A \in \mathcal{F} \implies A^c \in \mathcal{F}$

(iii) $A_1, A_2, \cdots \in \mathcal{F} \implies \bigcup_{n=1}^{\infty} A_n \in \mathcal{F}$

Note that because of (i) and (ii) the whole sample space is also included in the σ-algebra. Furthermore the σ-algebra is closed under intersection by applying DeMorgan's laws to (iii).

Definition: Borel set and Borel-σ-algebra A Borel set is a set in a topological space that is built with open sets through a countable union, countable intersection, relative complement or any combination of these. Equivalently, a Borel set can be constructed with closed sets, since each of the two can be expressed in terms of countable operations of the other. For instance $(a, b) = \bigcup_{i=1}^{\infty} \left[a + \frac{1}{i}, b - \frac{1}{i}\right]$.
The collection of all Borel sets on a topological space X form a σ-algebra. The Borel-σ-algebra on X, often denoted $\mathcal{B}(X)$, is the smallest collection of these sets such that the properties of a σ-algebra are satisfied. Put differently, it is the intersection of all σ-algebra's containing all subsets of X.

Definition: Stochastic variable
Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. A stochastic variable X, or random variable X, is a function that assigns to every outcome in the sample space Ω, a number on the real line $\mathbb{R}$. More precisely, $X : \Omega \to \mathbb{R}$ such that for all Borel subsets of $\mathbb{R}$, $\forall B \in \mathcal{B}(\mathbb{R})$, the set $\{\omega \in \Omega | X(\omega) \in B\} \in \mathcal{F}$. Most common is to write this set short-handedly as $\{X \in B\}$.

Definition: Probability space
A probability space is a mathematical space that translates an experiment to mathematics. More precisely, a probability space is a tuple with three elements $(\Omega, \mathcal{F}, \mathbb{P})$:

1. Ω, the sample space, which contains all possible outcomes.
2. $\mathcal{F}$, the set of events containing zero or more outcomes. This is a σ-algebra of subsets of Ω.
3. $\mathbb{P}$, the probability measure, which a assigns a number between zero and one for every event in $\mathcal{F}$.

The collection $\mathcal{F}$ and probability measure $\mathbb{P}$ is not defined very precise above. In the next two definitions both concepts are stated. The definition of the probability measure are also known as the Kolmogorov axioms.

Definition: Probability measure
Let Ω be a non-empty set, and let $\mathcal{F}$ be a σ-algebra of subsets of Ω. A probability measure $\mathbb{P}$ is a function that, to every set $A \in \mathcal{F}$, assigns a number in $[0, 1]$, called the probability of A and written $\mathbb{P}(A)$. We require:

(i) $\mathbb{P}(\Omega) = 1$

(ii) countable additivity, whenever $A_1, A_2, \cdots$ is a sequence of disjoint sets in $\mathcal{F}$, then $\mathbb{P}\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \mathbb{P}(A_n)$

At first glance the definition of a probability space might seen a bit abstract. Consider the experiment of a coin toss. The coin can land on two sides, namely heads (H) or tails (T). Therefore, the sample space $\Omega = \{H, T\}$. The σ-algebra $\mathcal{F}$ on Ω is the collection $\{\emptyset, \Omega, H, T\}$. This set captures all relevant events in the experiment. Note that the event $\emptyset$ correspond to the event that the coin will not show heads nor tails and the event Ω to the event that the coin will show heads or tails. The probability space is completed by defining the probability measure $\mathbb{P}$: $\mathbb{P}(\emptyset) = 0$, $\mathbb{P}(H) = \mathbb{P}(T) = \frac{1}{2}$, $\mathbb{P}(\Omega) = 1$.

Definition: Filtration
Let Ω be non-empty and $T > 0$. Assume $\forall t \in [0, T] : \exists \mathcal{F}(t)$, a σ-algebra. Further assume that $\forall s \leq t : \mathcal{F}(s) \subset F(t)$. Then, the collection of σ-algebras $\{\mathcal{F}(t), t \geq 0\}$ is called a filtration. In other words, the filtration is the information set that is available at time t. One can infer at time t whether or not the true event ω lies in each set in the filtration.

Definition: Stochastic process
Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. Assume that the indexing set $I \subset \mathbb{R}$ is a set of infinite cardinality. This indexing set typically represents time. Then, a stochastic process is a function $X : \Omega \times I \to \mathbb{R}$. Formally the process is written as $\{X_i, i \in I\}$. We speak of a discrete process if the index is restricted to the natural numbers, $I = \{\mathbb{N} \cup 0\}$. Then the stochastic process is the countable infinite collection $\{X_n, n \in \{\mathbb{N} \cup 0\}\}$. A continuous process is a stochastic process where $I = \mathbb{R}_+$, $\{X(t), t \in \mathbb{R}_+\}$.

Definition: Measurability
Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. Let X be a stochastic variable defined on the probability space. X is $\mathcal{F}$-measurable if $\sigma(X) \in \mathcal{F}$.

Definition: Adapted process
Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and $\mathcal{F}(t)$ a filtration. The stochastic process $\{X_i, i \in I\}$ is a $\mathcal{F}(t)$-adapted process if $\forall i \in I : X_i$ is $\mathcal{F}(t)$-measurable.

Definition: Brownian Motion / Wiener process
The Brownian motion or Wiener process is a continuous stochastic process $\{W(t), t \in \mathbb{R}_+\}$ with the following properties:

1. $W(0) = 0$ *almost surely*.
2. Independent increments: for $0 \leq s < t < \infty : W(t) - W(s) \perp\!\!\!\perp W(s)$.
3. Gaussian increments: for $0 \leq s < t < \infty : W(t) - W(s) \sim \mathcal{N}(0, t - s)$
4. the sample paths $t \mapsto W(t)$ are continuous *almost surely*.

Definition: Itô process
Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. An Itô process $\{X(t), t \in \mathbb{R}_+\}$ with drift function $\mu(X_s, s)$ an diffusion function $\sigma(X_s, s)$ is an adapted stochastic process that can be expressed as the sum of an integral with respect to time and an integral with respect to Brownian motion:

$$X_t = X_0 + \int_0^t \mu(X_s, s)ds + \int_0^t \sigma(X_s, s)dW_s$$

Lemma: Itô lemma
Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. Let $\{X(t), t \in \mathbb{R}_+\}$ an Itô process. Let $f(X_t, t)$ be a measurable function. Then the dynamics of $f(\cdot)$ are given by:

$$df(X_t, t) = \frac{\partial(X_t, t)}{\partial t}dt + \frac{\partial(X_t, t)}{\partial X_t}dX_t + \frac{1}{2}\frac{\partial^2(X_t, t)}{\partial X_t^2}d[X, X]_t$$

9.1.2.2 State Space Analysis Several economic time series are unobserved and have to be estimated. A well-known example of such a series is the volatility of a certain time series (e.g. a stock index). As one has only one observation per time point, the volatility at that point can not be extracted by applying the standard techniques. State space analysis serves the purpose of identifying the relation of such a state (the volatility) to its observation (stock price). Furthermore, state space analysis enables one to extract the state, estimate the parameters, and to forecast. In this thesis state space analysis is particularly useful for estimation. More specifically, as the short rate is an unobserved entity, one needs to extract this state variable in order to estimate the parameters of the model. In the case of the Heston model, the non-Gaussian nature of the model prevents the use of standard likelihood methods. State space methods and their further developments serve this cause very well.

Almost any model can be formulated into a state space model. A state space model is a general form that can be used for both linear and non-linear models. Define the state by α_t and the series of observations by y_t, $t \in \{1, 2, \cdots, n\}$. An important note is that both α_t and y_t can be scalars or vectors. The most general form of the state space formulation looks like:

$$\begin{aligned} y_t &= c_t + Z_t(\alpha_t) + I_t(\epsilon_t) \\ \alpha_{t+1} &= d_t + T_t(\alpha_t) + R_t(\eta_t) \end{aligned} \tag{9.28}$$

where c_t, d_t capture the levels of the series, the function $Z_t(\cdot)$ defines the relation between the observation and the state, the function $T_t(\cdot)$ identifies the progression of the state process and the functions $I_t(\cdot)$ and $R_t(\cdot)$ capture the irregular parts of the series. Furthermore, $I_t(\cdot) \sim (0, H_t)$ and $R_t(\cdot) \sim (0, Q_t)$. Note that all functions above also include the standard linear matrix multiplication, e.g. $Z_t\alpha_t \in Z_t(\alpha_t)$. Moreover, all system functions or matrices have to be known ex-ante and will typically depend on parameters. In other words, these entities are fixed in time. These system entities do not depend on parameters when the series do not have a level (such that $c_t = 0$ and/or $d_t = 0$) or follow a random walk (such that $Z(\cdot) = 1$ and/or $T(\cdot) = 1$ with the appropriate dimension). Now I turn to the methods to extract the unobserved state $\alpha_t, \forall t$, from the series y_t. First, I treat the method used for linear models, which is known as the Kalman filter. Then, I discuss the non-linear particle filter.

9.1.2.3 Kalman Filter

The Kalman filter is a method to extract the state from the observation series when the model is linear and the distribution of the irregular component are normal, and is developed by Rudolf Kalman (1960). The state space formulation of the most general linear Gaussian model takes the form of:

$$\begin{aligned} y_t &= c_t + Z_t\alpha_t + \epsilon_t \\ \alpha_{t+1} &= d_t + T_t\alpha_t + R_t\eta_t \end{aligned}$$

where $\epsilon_t \sim \mathcal{N}(0, H_t)$ and $\eta_t \sim \mathcal{N}(0, Q_t)$ and $t \in \{1, 2, \ldots, n\}$. The Kalman filter is an iterative scheme and rests on two lemmata. These lemmata propose an update of the distribution when new information comes into play. In other words, when new information is of significance, the expectation of the state gets altered conditional on this new information.

Lemma 9.1. *Let the stochastic vector (X, Y) be jointly normally distributed with $\mathbb{E}[X] \equiv \mu_X$, $\mathbb{E}[Y] \equiv \mu_Y$, $\mathbb{V}ar[X] \equiv \Sigma_X$, $\mathbb{V}ar[Y] \equiv \Sigma_Y$ and $\mathbb{C}ov[X, Y] \equiv \Sigma_{XY}$. Then, the conditional distribution of X given Y is normally distributed with mean and variance:*

$$\begin{aligned} \mathbb{E}[X|Y = y] &= \mu_X + \Sigma_{XY}\Sigma_Y^{-1}(y - \mu_Y) \text{ and} \\ \mathbb{V}ar[X|Y = y] &= \Sigma_X - \Sigma_{XY}\Sigma_Y^{-1}\Sigma_{XY}' \end{aligned}$$

Proof. Transform (X, Y) into (E, Y) by $E = X - \Sigma_{XY}\Sigma_Y^{-1}(Y - \mu_Y)$. Since this transformation is linear the vector (E, Y) is normally distributed.

Moreover, $\mathbb{E}[E] = \mathbb{E}[X] - \mathbb{E}\left[\Sigma_{XY}\Sigma_Y^{-1}(Y - \mu_Y)\right] = \mu_X$ and
$\mathbb{V}ar[E] = \mathbb{E}\left[(E - \mathbb{E}[E])(E - \mathbb{E}[E])'\right] = \mathbb{E}\left[(E - \mu_X)(E - \mu_X)'\right] =$
$\mathbb{E}\left[\left(X - \Sigma_{XY}\Sigma_Y^{-1}(Y - \mu_Y) - \mu_X\right)\left(X - \Sigma_{XY}\Sigma_Y^{-1}(Y - \mu_Y) - \mu_X\right)'\right] =$
$\mathbb{E}[XX' - X\mu_X' - \mu_X X' + \mu_X\mu_X'] - \mathbb{E}\left[\Sigma_{XY}\Sigma_Y^{-1}(Y - \mu_Y)(Y - \mu_Y)'\Sigma_Y^{-1}\Sigma_{XY}\right] = \Sigma_X - \Sigma_{XY}\Sigma_Y^{-1}\Sigma_{XY}'$.

Furthermore, $\mathbb{C}ov[E, Y] = \mathbb{E}\left[(E - \mu_E)(Y - \mu_Y)'\right] = \mathbb{E}\left[\left(X - \Sigma_{XY}\Sigma_Y^{-1}(Y - \mu_Y) - \mu_X\right)(Y - \mu_Y)'\right] =$
$\mathbb{E}\left[(X - \mu_X)(Y - \mu_Y)'\right] - \Sigma_{XY}\Sigma_Y^{-1}\mathbb{E}\left[(Y - \mu_Y)(Y - \mu_Y)'\right] = 0$.

Since E and Y are normal and uncorrelated, they are independent. Hence, the conditional distribution of E is the same as the unconditional distribution. But also $\mathbb{E}[E|Y = y] = \mathbb{E}[X|Y = y] - \Sigma_{XY}\Sigma_Y^{-1}(y - \mu_Y)$. Then, $\mu_X = \mathbb{E}[X|Y = y] - \Sigma_{XY}\Sigma_Y^{-1}(y - \mu_Y)$, which yields the result for the conditional expectation. The same argument applies to the conditional variance. □

Lemma 9.2. *Let the stochastic vector (X, Y, Z) be jointly normally distributed and suppose $\mathbb{E}[Z] = 0$ and $\Sigma_{YZ} = 0$. Then, the conditional distribution of X given (Y, Z) is normally distributed with mean and variance:*

$$\mathbb{E}[X|Y = y, Z = z] = \mathbb{E}[X|Y] + \Sigma_{XZ}\Sigma_{ZZ}^{-1}z \text{ and}$$
$$\mathbb{V}ar[X|Y = y, Z = z] = \mathbb{V}ar[X|Y] - \Sigma_{XZ}\Sigma_{ZZ}^{-1}\Sigma_{XZ}'$$

Proof. By applying Lemma 1 to X and $Y^* \equiv (Y', Z')'$ the result immediately follows. □

With the two lemmata one can derive the Kalman filter. First, assume that the parameters are known such that the system entities are known. Later, parameter estimation based on the Kalman filter will be discussed. Let $Y_t = (y_1, y_2, \ldots, y_t)'$ be the set of observations up until time t or more general $Y_t = (y_1', y_2', \ldots, y_t')'$. Define the one-step ahead estimate of the state by $a_t \equiv \mathbb{E}[\alpha_t|Y_{t-1}]$ and it's variance by $P_t \equiv \mathbb{V}ar[\alpha_t|Y_{t-1}]$. Furthermore, define the filtered state by $a_{t|t} \equiv \mathbb{E}[\alpha_t|Y_t]$ and it's variance by $P_{t|t} \equiv \mathbb{V}ar[\alpha_t|Y_t]$. Then:

$$v_t \equiv y_t - \mathbb{E}[y_t|Y_{t-1}] = y_t - \mathbb{E}[Z_t\alpha_t + \epsilon_t|Y_{t-1}] \underbrace{=}_{\epsilon_t \perp\!\!\!\perp Y_{t-1}} y_t - Z_t a_t = Z_t(\alpha_t - a_t) + \epsilon_t$$

$$\mathbb{E}[v_t|Y_{t-1}] = Z_t\mathbb{E}[\alpha_t|Y_{t-1}] - Z_t a_t = 0 \underbrace{\rightarrow}_{\text{Law iterated expectation}} \mathbb{E}[v_t] = 0$$

$$F_t \equiv \mathbb{V}ar[v_t|Y_{t-1}] = \mathbb{V}ar[Z_t(\alpha_t - a_t) + \epsilon_t|Y_{t-1}] \underbrace{=}_{\alpha_t \perp\!\!\!\perp \epsilon_t} \mathbb{V}ar[Z_t(\alpha_t - a_t)|Y_{t-1}] + \mathbb{V}ar[\epsilon_t|Y_{t-1}]$$
$$= Z_t P_t Z_t' + H_t$$

First derive the covariance between α_t and v_t given Y_{t-1}, which is needed when invoking Lemma 2 to incorporate new information.

$$\mathbb{C}ov[\alpha_t, v_t|Y_{t-1}] = \mathbb{E}[\alpha_t(Z_t(\alpha_t - a_t) + \epsilon_t)'|Y_{t-1}] = \mathbb{E}[\alpha_t\alpha_t'Z_t'|Y_{t-1}] = P_t Z_t'$$

Note that $\mathbb{E}[\alpha_t|Y_t] = \mathbb{E}[\alpha_t|Y_{t-1}, v_t]$, since this only entails a linear transformation of the data. Using Lemma 2 to incorporate the new information yields:

$$a_{t|t} = \mathbb{E}[\alpha_t|Y_t] = \mathbb{E}[\alpha_t|Y_{t-1}, v_t] \underbrace{=}_{\text{Lemma 2}} \mathbb{E}[\alpha_t|Y_{t-1}] + \mathbb{C}ov[\alpha_t, v_t|Y_{t-1}]\mathbb{V}ar[v_t|Y_{t-1}]^{-1}v_t = a_t + P_t Z_t' F_t^{-1} v_t$$

$$P_{t|t} = \mathbb{V}ar[\alpha_t|Y_t] = \mathbb{V}ar[\alpha_t|Y_{t-1}, v_t] \underbrace{=}_{\text{Lemma 2}} \mathbb{V}ar[\alpha_t|Y_{t-1}] - \mathbb{C}ov[\alpha_t, v_t|Y_{t-1}]\mathbb{V}ar[v_t|Y_{t-1}]^{-1}\mathbb{C}ov[\alpha_t, v_t|Y_{t-1}]'$$
$$= P_t - P_t Z_t' F_t^{-1} Z_t \underbrace{P_t}_{P_t' = P_t}$$

The prediction of the state at time $t+1$ becomes:

$$a_{t+1} = \mathbb{E}[\alpha_{t+1}|Y_t] = \mathbb{E}[T_t\alpha_t + R_t\eta_t|Y_t] = T_t\mathbb{E}[\alpha_t|Y_t] + R_t\mathbb{E}[\eta_t|Y_t] = T_t a_{t|t} = T_t a_t + K_t v_t$$

Where the Kalman gain matrix $K_t \equiv T_t P_t Z_t' F_t^{-1}$

$$P_{t+1} = \mathbb{V}ar[\alpha_{t+1}|Y_t] = \mathbb{V}ar[T_t\alpha_t + R_t\eta_t|Y_t] = T_t\mathbb{V}ar[\alpha_t|Y_t]T_t' + R_tQ_tR_t' = T_tP_{t|t}T_t' + R_tQ_tR_t'$$
$$= T_tP_t\left(T_t - K_tZ_t\right)' + R_tQ_tR_t'$$

The only thing that rests is the initialization of the algorithm. To start up the iteration scheme, the best guess is $\alpha_1 \sim \mathcal{N}(a_1, P_1)$ for some a_1 and P_1. There are three different possibilities for a_1 and P_1. Firstly, when one deals with a stationary model (such as a stationary ARMA model), the initial values of a_1 and P_1 can be initialized as the unconditional mean and variance of the process. Secondly, when it is known that the process is non-stationary, one has to incorporate an uninformative prior, that is, $a_1 = 0$ and $P_1 \to \infty$. Note that in this case the value of a_1 does not really matter as the variance approaches infinity. The third option is to estimate the initialization simultaneously with the other parameters. In this scenario one treats the initial state estimate and variance as components of the parameter vector.

The parameters of the linear Gaussian state space model can be estimated with maximum likelihood by making use of the prediction error decomposition. Denote the vector of parameters by ψ. Then, the joint density of the observations can be decomposed as follows:

$$p(y_1, y_2, \ldots, y_n|\psi) = p(y_1|\psi)\prod_{i=2}^{n} p(y_i|Y_{i-1}\psi) \underbrace{=}_{Y_0=\emptyset} \prod_{i=1}^{n} p(y_i|Y_{i-1}\psi)$$

$$\log\left(p(y_1, y_2, \ldots, y_n|\psi)\right) = \sum_{i=1}^{n} \log\left(p(y_i|Y_{i-1}\psi)\right)$$

In the linear Gaussian model $\mathbb{E}[y_t|Y_{t-1}] = Z_t a_t$ and $\mathbb{V}ar[y_t|Y_{t-1}] = \mathbb{E}[(y_t - Z_t a_t)(y_t - Z_t a_t)'|Y_{t-1}] = \mathbb{E}[v_t v_t'|Y_{t-1}] = F_t$. Then the log-likelihood function boils down to:

$$\log\left(p(y_1, y_2, \ldots, y_n|\psi)\right) = \sum_{i=1}^{n} -\frac{|y_t|}{2}\log(2\pi) - \frac{1}{2}\log\left[\det(F_t)\right] - \frac{1}{2}v_t'F_t^{-1}v_t$$
$$= -\frac{n|y_t|}{2}\log(2\pi) - \sum_{i=1}^{n}\frac{1}{2}\log\left[\det(F_t)\right] - \frac{1}{2}v_t'F_t^{-1}v_t$$

where $|\cdot|$ denotes the cardinality of the vector. The parameters of the model are obtained by optimizing the log-likelihood with standard optimization algorithms such as the Broyden-Fletcher-Goldfarb-Shanno (BFGS) algorithm.

9.1.2.4 Importance Sampling and Particle Filter

This section discusses how to filter the signal when the model is non-linear and/or non-Gaussian. Instead of the linear Gaussian case, where $y_t|\alpha_t \sim \mathcal{N}(c_t + Z_t\alpha_t, Q_t)$ and $\alpha_{t+1}|\alpha_t \sim \mathcal{N}(d_t + T_t\alpha_t, R_tH_tR_t')$, the model has now a very general form:

$$\begin{aligned} y_t &\sim p(y_t|\alpha_t) \\ \alpha_t &\sim p(\alpha_t|\alpha_{t-1}) \end{aligned} \tag{9.29}$$

In the same spirit as the Kalman filter, $\mathbb{E}[\alpha_t|Y_t]$ is still of main interest. Or more general, any function of the state $\mathbb{E}[x_t(\alpha_t)|Y_t]$. As in the most general case no analytical results exist for this expression, such as in the Kalman filter, these expectations are obtained by means of simulation. This method is justified as long as the number of samples at every time point t is large enough The justification is established by the law of large numbers, which states that for $N >> 0$ independent draws, the mean of the sample of realizations $X^{(1)}, X^{(2)}, \ldots, X^{(N)}$, where e.g. $X \sim \mathcal{N}(0,1)$, will be equal to the expectation of the stochastic variable X, $\mathbb{E}[X]$ almost surely. Mathematically, $\forall\epsilon > 0, \lim_{N\to\infty}\mathbb{P}\left(\left\|\frac{\sum_{i=1}^{N}X^{(i)}}{N} - \mathbb{E}[X]\right\| > \epsilon\right) = 0$.

Importance sampling was originally developed for cases when sampling from the original distribution was slow or even impossible. The basic idea is to find a distribution, the importance density and denoted by $g(\cdot)$, that is as close as possible to the original distribution $g(\cdot)$ from which draws are easily obtained, i.e. $g(\cdot) \simeq p(\cdot)$. Importance sampling models the full posterior, i.e. $p(\alpha_{1:t}|Y_t)$, where $\alpha_{1:t} \equiv (\alpha_1, \alpha_2, \ldots, \alpha_t)$, and properties of this posterior distribution. Rewriting $\mathbb{E}[x_t(\alpha_{1:t})|Y_t]$ gives the main idea behind importance sampling:

$$\begin{aligned} \mathbb{E}_p[x_t(\alpha_{1:t})|Y_t] &= \int x_t(\alpha_{1:t})\, p(\alpha_{1:t}|Y_t)\, d\alpha_{1:t} = \int x_t(\alpha_{1:t})\, \frac{p(\alpha_{1:t}|Y_t)}{g(\alpha_{1:t}|Y_t)} g(\alpha_{1:t}|Y_t)\, d\alpha_{1:t} \\ &= \mathbb{E}_g\left[x_t(\alpha_{1:t})\, \frac{p(\alpha_{1:t}|Y_t)}{g(\alpha_{1:t}|Y_t)}|Y_t\right] = \frac{g(Y_t)}{p(Y_t)}\mathbb{E}_g\left[x_t(\alpha_{1:t})\, \frac{p(\alpha_{1:t}, Y_t)}{g(\alpha_{1:t}, Y_t)}|Y_t\right] \end{aligned} \tag{9.30}$$

When evaluating $x_t(\alpha_{1:t}) = 1, \forall\alpha_{1:t}$:

$$1 = \frac{g(Y_t)}{p(Y_t)}\mathbb{E}_g\left[\frac{p(\alpha_{1:t}, Y_t)}{g(\alpha_{1:t}, Y_t)}|Y_t\right] \tag{9.31}$$

Dividing equation (9.30) by (9.31):

$$\mathbb{E}_p[x_t(\alpha_{1:t})|Y_t] = \frac{\mathbb{E}_g[x_t(\alpha_{1:t})\, w(\alpha_{1:t}, Y_t)|Y_t]}{\mathbb{E}_g[w(\alpha_{1:t}, Y_t)|Y_t]} \tag{9.32}$$

where $w(\alpha_{1:t}, Y_t) \equiv \frac{p(\alpha_{1:t}, Y_t)}{g(\alpha_{1:t}, Y_t)}$ are the so-called importance weights. Thus, at the core of importance sampling lies a change of measure, where the importance weight is basically the Radon-Nikodym derivative. Equation (9.32) gives the basis for importance sampling and one can obtain a result for the desired expectation with the help of the law of large numbers:

$$\frac{\mathbb{E}_g[x_t(\alpha_{1:t})\, w(\alpha_{1:t}, Y_t)|Y_t]}{\mathbb{E}_g[w(\alpha_{1:t}, Y_t)|Y_t]} = \text{plim}_{N\to\infty}\frac{\sum_{i=1}^{N} x_t\left(\alpha_{1:t}^{(i)}\right) w\left(\alpha_{1:t}^{(i)}, Y_t\right)}{\sum_{i=1}^{N} w\left(\alpha_{1:t}^{(i)}, Y_t\right)} \tag{9.33}$$

where $\alpha_{1:t}^{(i)}$ is a draw from the posterior density $g(\alpha_{1:t}|Y_t)$. Filtering by importance sampling is very inefficient since draws have to be obtained from the full posterior. So one would obtain n estimates for the first state, $n-1$ for the second state, and so on. For long time series this would translate to lots of unnecessary work.

Whereas importance sampling estimates functions of the state from the full posterior $p(\alpha_1, \alpha_2, \ldots, \alpha_t | y_1, y_2, \ldots, Y_t)$ for every t, a particle filter retains the sample from $t-1$ and samples conditional on this sample following the known transition density. Particle filtering boils down to sampling $\alpha_t^{(i)} | \alpha_{1:t-1}^{(i)}, \forall i$. The samples in the set $\left\{\alpha_{1:t}^{(1)}, \alpha_{1:t}^{(2)}, \ldots \alpha_{1:t}^{(N)}\right\}$ are called particles and each particle contains a realization of the hidden process until t. These draws should be consistent with the distribution from the importance density:

$$g\left(\alpha_{1:t} | Y_t\right) = \frac{g\left(\alpha_{1:t}, Y_t\right)}{g\left(Y_t\right)} = \frac{g\left(\alpha_t | \alpha_{1:t-1}, Y_t\right) g\left(\alpha_{1:t-1}, Y_t\right)}{g\left(Y_t\right)} = g\left(\alpha_t | \alpha_{1:t-1}, Y_t\right) g\left(\alpha_{1:t-1} | Y_t\right) \tag{9.34}$$

To obtain a recursion, the following assumption is needed. Suppose $\alpha_{1:t-1}$ depends only on information contained in Y_{t-1}. Furthermore, the latent process is already established and new information y_t is established by a different process which does not depend directly on $\alpha_{1:t-1}$. Therefore, one can assume $g\left(\alpha_{1:t-1} | Y_t\right) \equiv g\left(\alpha_{1:t-1} | Y_{t-1}\right)$. Then:

$$g\left(\alpha_{1:t} | Y_t\right) = g\left(\alpha_t | \alpha_{1:t-1}, Y_t\right) g\left(\alpha_{1:t-1} | Y_{t-1}\right) \tag{9.35}$$

This assumption and the recursion above lie at the heart of particle filtering. For more details, I refer the interested reader to, e.g., Durbin & Koopman (2012) or Doucet, De Freitas and Gordon (2001, paragraph 1.3).

The recursion of the importance weights is based on the second to last equality in equation (9.30):

$$\mathbb{E}_g\left[x_t\left(\alpha_{1:t}\right) \frac{p\left(\alpha_{1:t} | Y_t\right)}{g\left(\alpha_{1:t} | Y_t\right)} | Y_t\right] = \frac{1}{p\left(Y_t\right)} \mathbb{E}_g\left[x_t\left(\alpha_{1:t}\right) \frac{p\left(\alpha_{1:t}, Y_t\right)}{g\left(\alpha_{1:t} | Y_t\right)} | Y_t\right] = \frac{1}{p\left(Y_t\right)} \mathbb{E}_g\left[x_t\left(\alpha_{1:t}\right) \tilde{w}_t | Y_t\right] \tag{9.36}$$

where $\tilde{w}_t = \frac{p(\alpha_{1:t}, Y_t)}{g(\alpha_{1:t} | Y_t)}$. Now:

$$\begin{aligned} \tilde{w}_t &= \frac{p\left(\alpha_{1:t}, Y_t\right)}{g\left(\alpha_{1:t} | Y_t\right)} = \frac{p\left(\alpha_t, y_t | \alpha_{1:t-1}, Y_{t-1}\right) p\left(\alpha_{1:t-1}, Y_{t-1}\right)}{g\left(\alpha_t | \alpha_{1:t-1}, Y_t\right) g\left(\alpha_{1:t-1} | Y_{t-1}\right)} \overset{\text{eq. (9.29)}}{=} \frac{p\left(y_t | \alpha_t\right) p\left(\alpha_t | \alpha_{t-1}\right) p\left(\alpha_{1:t-1}, Y_{t-1}\right)}{g\left(\alpha_t | \alpha_{1:t-1}, Y_t\right) g\left(\alpha_{1:t-1} | Y_{t-1}\right)} \\ &= \frac{p\left(y_t | \alpha_t\right) p\left(\alpha_t | \alpha_{t-1}\right)}{g\left(\alpha_t | \alpha_{1:t-1}, Y_t\right)} \tilde{w}_{t-1} \end{aligned} \tag{9.37}$$

Estimates are obtained by setting $\tilde{w}_t^{(i)} = 1$ for $i = 1, 2, \ldots, N$ and $t = 1, 2, \ldots, \tau$, for a certain burn-in τ. At $t = \tau + 1$ the recursions can be used to obtain the weights. By using the same approach as in importance sampling, by setting $x_t\left(\alpha_{1:t}\right) = 1$ and then dividing the two equations, starting from equation (9.36):

$$\mathbb{E}_g\left[x_t\left(\alpha_{1:t}\right) \frac{p\left(\alpha_{1:t} | Y_t\right)}{g\left(\alpha_{1:t} | Y_t\right)} | Y_t\right] = \frac{1}{p\left(Y_t\right)} \mathbb{E}_g\left[x_t\left(\alpha_{1:t}\right) \tilde{w}_t | Y_t\right] = \frac{\mathbb{E}_g\left[x_t\left(\alpha_{1:t}\right) \tilde{w}_t | Y_t\right]}{\mathbb{E}_g\left[\tilde{w}_t | Y_t\right]} \tag{9.38}$$

Which is almost identical to the case of importance sampling with the only difference being the weights.

Particle filters sometimes suffer from degeneracy, which means that particle weights are becoming small. It is computationally inefficient to retain particles with little weight as their impact on the Monte Carlo estimate is small. To circumvent degeneracy, one can replace particles with little weight by the particles with a large weight when the effective sample size (ESS) is below a certain threshold. Before one obtains new draws for the state process, the particles should be re-sampled from the full set of particles, with replacement, according to the normalized weights $w_t^{(i)} = \frac{\tilde{w}_t^{(i)}}{\sum_{i=1}^N \tilde{w}_t^{(i)}}$. The ESS is defined as $ESS = \frac{1}{\sum_{i=1}^N \tilde{w}_t^{(i)2}}$ and takes a value between zero and one. The threshold is normally set to 0.75 or 0.5. So when effectively only 75% or 50% percent of the particles are informative, re-sampling

takes place.

Parameter estimates follow from maximizing the simulated log-likelihood numerically. When denoting $\alpha = (\alpha_1', \alpha_1', \ldots, \alpha_n')'$, the likelihood function is:

$$L(\psi) = \int p(\alpha, Y_n)\, d\alpha = \int \frac{p(\alpha, Y_n)}{g(\alpha|Y_n)} g(\alpha|Y_n)\, d\alpha = \int \tilde{w}_n g(\alpha|Y_n)\, d\alpha = \mathbb{E}_g[\tilde{w}_n|Y_n] \tag{9.39}$$

One obtains the MLE by optimizing this function with standard optimization techniques. For a detailed description see Durbin & Koopman (2012). The covariance matrix of the parameter estimates follow by the standard formula:

$$\hat{\Omega} = \left[-\frac{\partial \log L(\psi)}{\partial \psi \partial \psi'} \right]^{-1} \Big|_{\psi=\hat{\psi}} \tag{9.40}$$

9.2 Proofs

9.2.1 Hull-White: Spot Rate Affine Formulation

In order to find expressions for $\alpha(t,T)$ and $\beta(t,T), \forall(t,T)$, the first step is to find the dynamics of the bond price process. By applying Itô's Lemma one gets:

$$dP(t,T) = \frac{\partial P(t,T)}{\partial t}dt + \frac{\partial P(t,T)}{\partial r(t)}dr(t) + \frac{1}{2}\frac{\partial^2 P(t,T)}{\partial r(t)^2}(dr(t))^2 \tag{9.41}$$

Plugging in equation 3.10 (the short rate dynamics), the quadratic variation, and the derivatives obtained from equation 3.11 yields:

$$\frac{dP(t,T)}{P(t,T)} = \left(\frac{\partial \alpha(t,T)}{\partial t} + \frac{\partial \beta(t,T)}{\partial t}r(t) + \beta(t,T)\Theta(t) - \beta(t,T)\kappa r(t) + \frac{1}{2}\beta^2(t,T)\sigma^2\right)dt + \beta(t,T)\sigma dW(t) \tag{9.42}$$

From the martingale representation theorem it follows that the zero-coupon bond process is not a martingale (the drift coefficient can not be zero $\forall t, r(t)$). In order to create a martingale, the dynamics need to be transformed into deflated bond dynamics. By using the money market account as numeraíre or equivalently by applying Itô's product rule to find the discounted bond dynamics, since the discount factor equation (9.6) is the reciprocal of the money market account, one obtains:

$$\begin{aligned} d\left(DF(t)P(t,T)\right) = & \underbrace{dDF(t)}_{-r(t)DF(t)dt} P(t,T) + DF(t)dP(t,T) + \underbrace{dDF(t,T)dP(t,T)}_{=0} \\ & = -r(t)DF(t)P(t,T)dt + DF(t)P(t,T) \\ & \left(\left(\frac{\partial \alpha(t,T)}{\partial t} + \frac{\partial \beta(t,T)}{\partial t}r(t) + \beta(t,T)\Theta(t) - \beta(t,T)\kappa r(t) + \frac{\sigma^2}{2}\beta^2(t,T)\right)dt + \beta(t,T)\sigma dW(t)\right) \end{aligned} \tag{9.43}$$

Rewriting yields:

$$\begin{aligned} \frac{d\left(DF(t)P(t,T)\right)}{DF(t)P(t,T)} = & \left(-r(t) + \frac{\partial \alpha(t,T)}{\partial t} + \frac{\partial \beta(t,T)}{\partial t}r(t) + \beta(t,T)\Theta(t) - \beta(t,T)\kappa r(t) + \frac{\sigma^2}{2}\beta^2(t,T)\right)dt \\ & + \beta(t,T)\sigma dW(t) \end{aligned} \tag{9.44}$$

The only extra term in comparison with the normal bond dynamics is $-r(t)dt$. This correction is intuitive as it corrects for the accumulation in the money market account, which has to be equal to the value accrual of the bond under no-arbitrage. By the martingale representation theorem, the deflated dynamics is a martingale if the drift term is zero, $\forall(t,r(t))$. The solutions to $\alpha(t,T), \beta(t,T), \Theta(t)$ follow from solving the equations imposed by the martingale representation theorem and the boundary conditions:

$$\begin{aligned} \frac{\partial \alpha(t,T)}{\partial t} + \beta(t,T)\Theta(t) + \frac{\sigma^2}{2}\beta^2(t,T) &= 0 \\ -r(t) + \frac{\partial \beta(t,T)}{\partial t}r(t) - \beta(t,T)\kappa r(t) = 0 \iff -1 + \frac{\partial \beta(t,T)}{\partial t} - \beta(t,T)\kappa &= 0 \\ \alpha(T,T) &= 0 \\ \beta(T,T) &= 0 \end{aligned} \tag{9.45}$$

From the second and fourth line from equation (9.45) the solution for $\beta(t,T)$ yields:

$$\beta(t,T) = \frac{1}{\kappa}\left(e^{-\kappa(T-t)} - 1\right) \tag{9.46}$$

Integrating the first line of equation (9.45):

$$\alpha(t,T) = \int_t^T \beta(u,T)\Theta(u)du + \frac{\sigma^2}{2}\int_t^T \beta^2(u,T)du - \underbrace{\alpha(T,T)}_{=0,\ \text{eq.}(9.45)} \tag{9.47}$$

Plugging $\Theta(\cdot)$ to match the initial curve:

$$\begin{aligned}
\frac{\partial\alpha(0,T)}{\partial T} &\overset{\text{Leibniz}}{=} \underbrace{\beta(T,T)}_{=0}\Theta(T) + \frac{\sigma^2}{2}\underbrace{\beta^2(T,T)}_{=0} + \int_0^T \underbrace{\frac{\partial\beta(u,T)}{\partial T}}_{=-e^{-\kappa(T-u)}}\Theta(u)du + \sigma^2\int_0^T \beta(u,T)\frac{\partial\beta(u,T)}{\partial T}du \\
&\overset{\text{eq.}(9.45)\ \&\ (9.46)}{=} -\int_0^T e^{-\kappa(T-u)}\Theta(u)du + \sigma^2\int_0^T e^{-\kappa(T-u)}\frac{1}{\kappa}(e^{-\kappa(T-u)}-1)du \\
&= -\int_0^T e^{-\kappa(T-u)}\Theta(u)du - \frac{\sigma^2}{\kappa^2}\left[\frac{1}{2}\left(1-e^{-2\kappa T}\right) - \left(1-e^{-\kappa T}\right)\right]
\end{aligned} \tag{9.48}$$

Combining the fact that $f(0,T) = -\frac{\partial\log(P(0,T))}{\partial T} = -\underbrace{\frac{\partial\alpha(0,T)}{\partial T}}_{\text{eq. (9.48)}} - \underbrace{\frac{\partial\beta(0,T)}{\partial T}}_{\text{eq. (9.46)}} r(0) = -\frac{\partial\alpha(0,T)}{\partial T} + e^{-\kappa T}r(0)$

and equation (9.48) yields:

$$f(0,T) = \int_0^T e^{-\kappa(T-u)}\Theta(u)du + \frac{\sigma^2}{\kappa^2}\left[\frac{1}{2}\left(1-e^{-2\kappa T}\right) - \left(1-e^{-\kappa T}\right)\right] + e^{-\kappa T}r(0) \tag{9.49}$$

Using Leibniz' rule to get an expression in $\Theta(T)$:

$$\frac{\partial f(0,T)}{\partial T} = \Theta(T) - \kappa\int_0^T e^{-\kappa(T-u)}\Theta(u)du + \frac{\sigma^2}{\kappa}\left[e^{-2\kappa T} - e^{-\kappa T}\right] - \kappa e^{-\kappa T}r(0) \tag{9.50}$$

Using equation (9.49) to replace integral:

$$\begin{aligned}
\Theta(T) &= \frac{\partial f(0,T)}{\partial T} + \kappa\left[f(0,T) - \frac{\sigma^2}{\kappa^2}\left[\frac{1}{2}\left(1-e^{-2\kappa T}\right) - \left(1-e^{-\kappa T}\right)\right] - e^{-\kappa T}r(0)\right] - \frac{\sigma^2}{\kappa}\left[e^{-2\kappa T} - e^{-\kappa T}\right] + \kappa e^{-\kappa T}r(0) \\
&= \frac{\partial f(0,T)}{\partial T} + \kappa f(0,T) + \frac{\sigma^2}{2\kappa}\left(1-e^{-2\kappa T}\right)
\end{aligned} \tag{9.51}$$

The expression for $\alpha(t,T)$ follow from plugging equation (9.51) into equation (9.48):

$$\begin{aligned}
\alpha(t,T) &= \frac{1}{\kappa}\int_t^T \left(e^{-\kappa(T-u)}-1\right)\left(\frac{\partial f(0,u)}{\partial u}+\kappa f(0,u)+\frac{\sigma^2}{2\kappa}\left(1-e^{-2\kappa u}\right)\right)du + \frac{\sigma^2}{2\kappa^2}\int_t^T\left(e^{-\kappa(T-u)}-1\right)^2 du \\
&= \underbrace{\frac{1}{\kappa}\int_t^T e^{-\kappa(T-u)}\frac{\partial f(0,u)}{\partial u}du}_{IBP:\frac{1}{\kappa}\left[e^{-\kappa(T-u)}f(0,u)\right]_{u=t}^T-\int_t^T e^{-\kappa(T-u)}f(0,u)du} + \int_t^T e^{-\kappa(T-u)}f(0,u)du + \frac{\sigma^2}{2\kappa^2}\int_t^T e^{-\kappa(T-u)}\left(1-e^{-2\kappa u}\right)du \\
&\quad - \frac{1}{\kappa}\int_t^T \frac{\partial f(0,u)}{\partial u}du + \int_t^T f(0,u)du + \frac{\sigma^2}{2\kappa^2}\int_t^T\left(1-e^{-2\kappa u}\right)du + \frac{\sigma^2}{2\kappa^2}\int_t^T\left(e^{-\kappa(T-u)}-1\right)^2 du \\
&= \underbrace{\frac{1}{\kappa}f(0,t)-\frac{1}{\kappa}e^{-\kappa(T-t)}f(0,t)}_{=-\beta(t,T)f(0,t)} - \int_t^T f(0,u)du \\
&\quad + \frac{\sigma^2}{2\kappa^2}\left[\int_t^T e^{-\kappa(T-u)}\left(1-e^{-2\kappa u}\right)du - \int_t^T\left(1-e^{-2\kappa u}\right)du + \int_t^T\left(e^{-\kappa(T-u)}-1\right)^2 du\right] \\
&= -\beta(t,T)f(0,t) - \int_t^T f(0,u)du + \frac{\sigma^2}{4\kappa}\beta^2(t,T)\left(e^{-2\kappa t}-1\right) \\
&= -\beta(t,T)f(0,t) + \log\left(\frac{P(0,T)}{P(0,t)}\right) + \frac{\sigma^2}{4\kappa}\beta^2(t,T)\left(e^{-2\kappa t}-1\right)
\end{aligned} \tag{9.52}$$

Where in the last step the equation $P(0,T) = e^{-\int_0^T f(0,u)du}$ is used.

9.2.2 Hull-White: Solution of the Stochastic Differential Equation

The first step to find the solution of the stochastic differential equation 3.10 is to find the dynamics of the process $e^{\kappa t}r(t)$. By applying Itô's Lemma, this process is dictated by:

$$\begin{aligned}
de^{\kappa t}r(t) &= e^{\kappa t}dr(t) + \kappa e^{\kappa t}r(t)\,dt + \frac{1}{2}\kappa^2 e^{\kappa t}r(t)\underbrace{(dt)^2}_{\to 0} + \frac{1}{2}\underbrace{\frac{\partial^2 e^{\kappa t}r(t)}{\partial r(t)^2}}_{=0}(dr(t))^2 + \kappa e^{\kappa t}\underbrace{dr(t)\,dt}_{\to 0} \\
&= e^{\kappa t}dr(t) + \kappa e^{\kappa t}r(t)\,dt \\
&\overset{\text{eq. (3.10)}}{=} e^{\kappa t}\Theta(t)dt + e^{\kappa t}\sigma dW(t)
\end{aligned} \tag{9.53}$$

Then taking the integral from s to t yields:

$$e^{\kappa t}r(t) - e^{\kappa s}r(s) = \int_s^t e^{\kappa u}\Theta(u)du + \sigma\int_s^t e^{\kappa u}dW(u) \tag{9.54}$$

Rewriting gives:

$$
\begin{aligned}
r(t) &= e^{-\kappa(t-s)} r(s) + \underbrace{\int_s^t e^{-\kappa(t-u)} \Theta(u) du}_{=-\int_s^t \frac{\partial \beta(u,t)}{\partial t} \Theta(u) du} + \sigma \int_s^t e^{-\kappa(t-u)} dW(u) \\
&= e^{-\kappa(t-s)} r(s) - \underbrace{\int_s^t \frac{\partial \beta(u,t)}{\partial t} \Theta(u) du}_{\text{Leibniz: } \frac{\partial}{\partial t}\int_s^t \beta(u,t)\Theta(u)du - \underbrace{\beta(t,t)}_{=0}\Theta(t)} + \sigma \int_s^t e^{-\kappa(t-u)} dW(u) \\
&= e^{-\kappa(t-s)} r(s) - \frac{\partial}{\partial t} \int_s^t \beta(u,t) \Theta(u) du + \sigma \int_s^t e^{-\kappa(t-u)} dW(u) \\
&\overset{\text{eq. (9.51)}}{=} e^{-\kappa(t-s)} r(s) - \frac{\partial}{\partial t} \int_s^t \beta(u,t) \left(\frac{\partial f(0,u)}{\partial u} + \kappa f(0,u) + \frac{\sigma^2}{2\kappa} \left(1 - e^{-2\kappa u}\right) \right) du + \sigma \int_s^t e^{-\kappa(t-u)} dW(u)
\end{aligned}
\tag{9.55}
$$

Note that by integration by parts (IBP):

$$
\begin{aligned}
\int_s^t \beta(u,t) \frac{\partial f(0,u)}{\partial u} du &= \underbrace{\beta(t,t)}_{=0} f(0,t) - \beta(s,t) f(0,s) - \int_s^t \frac{\partial \beta(u,t)}{\partial u} f(0,u) du \\
&= -\beta(s,t) f(0,s) - \int_s^t \left(\kappa \beta(u,t) + 1\right) f(0,u) du
\end{aligned}
\tag{9.56}
$$

Hence, $\int_s^t \beta(u,t) \left(\frac{\partial f(0,u)}{\partial u} + \kappa f(0,u) \right) du = -\beta(s,t) f(0,s) - \int_s^t f(0,u) du$ such that:

$$
\begin{aligned}
r(t) &= e^{-\kappa(t-s)} r(s) - \frac{\partial}{\partial t} \Bigg(-f(0,s)\beta(s,t) - \int_s^t f(0,u) du + \\
&\quad \frac{\sigma^2}{2\kappa^3} \left(1 - e^{-\kappa(t-s)} + \frac{1}{2} e^{-2\kappa t} - e^{-\kappa(t+s)} + \frac{1}{2} e^{-2\kappa s} \right) + \frac{\sigma^2}{2\kappa^2} (t-s) \Bigg) + \sigma \int_s^t e^{-\kappa(t-u)} dW(u) \\
&= e^{-\kappa(t-s)} r(s) + f(0,t) + \frac{\sigma^2}{2\kappa^2} \left(1 - e^{-\kappa t}\right)^2 - e^{-\kappa(t-s)} \left(f(0,s) + \frac{\sigma^2}{2\kappa^2} \left(1 - e^{-\kappa s}\right)^2 \right) \\
&\quad + \sigma \int_s^t e^{-\kappa(t-u)} dW(u) \\
&= e^{-\kappa(t-s)} \left(r(s) - \left(f(0,s) + \frac{\sigma^2}{2\kappa^2} \left(1 - e^{-\kappa s}\right)^2 \right) \right) + f(0,t) + \frac{\sigma^2}{2\kappa^2} \left(1 - e^{-\kappa t}\right)^2 \\
&\quad + \sigma \int_s^t e^{-\kappa(t-u)} dW(u)
\end{aligned}
\tag{9.57}
$$

The quadratic variation of the Itô integral equals $\left[\sigma \int_s^t e^{-\kappa(t-u)} dW(u), \sigma \int_s^t e^{-\kappa(t-u)} dW(u)\right] = \sigma^2 \int_s^t e^{-2\kappa(t-u)} du = \frac{\sigma^2}{2\kappa} \left(1 - e^{-2\kappa(t-s)}\right)$. Now one can conclude that:

$$
r(t) | r(s) \sim \mathcal{N} \left(e^{-\kappa(t-s)} \left(r(s) - \mu(s) \right) + \mu(t), \frac{\sigma^2}{2\kappa} \left(1 - e^{-2\kappa(t-s)} \right) \right) \tag{9.58}
$$

with $\mu(t) = f(0,t) + \frac{\sigma^2}{2\kappa^2} \left(1 - e^{-\kappa t}\right)^2$.

9.3 Heston Model: Continuous Time

Although exact simulation schemes are developed, it appears that the literature does not make use of the exact results in deriving a particle filter. An exception is the paper by Aihara et al. (2012). The authors use the results of Broadie and Kaya (2006) to estimate the Bates model. However, I propose to use the exact distribution for the variance process and the QE scheme for the asset process as in the NCI scheme. An objection against using the results of Broadie and Kaya (2006) is that for small time increment Δt, the simulation of the integrated variance process is very unstable, see Van Haastrecht and Pelsser (2008). Moreover, the NCI scheme performs very well for small time increments Δt. From these results one could argue that the exact scheme is desirable for simulating series on a grid with higher granularity, but a hybrid method such as the NCI scheme should perform very comparable on a fine grid and is computationally much cheaper.

The derivation for the continuous counterpart starts again from equation (3.3). In this derivation I follow the approach of Aihara et al. (2013) combined with the method of drift interpolation method of Andersen (2007). Plugging $dW_1(t) = \dfrac{d\log S(t) - \left(\mu - \frac{1}{2}V(t)\right)dt}{\sqrt{V(t)}}$ into the variance process yields:

$$\begin{aligned} d\log S(t) &= \left(\mu - \frac{1}{2}V(t)\right)dt + \sqrt{V(t)}dW_1(t) \\ dV(t) &= \kappa\left(\theta - V(t)\right)dt + \sigma\rho\left(d\log S(t) - \left(\mu - \frac{1}{2}V(t)\right)dt\right) + \sigma\sqrt{1-\rho^2}\sqrt{V(t)}dW_2(t) \end{aligned} \tag{9.59}$$

Collecting terms gives:

$$\begin{aligned} d\log S(t) &= \left(\mu - \frac{1}{2}V(t)\right)dt + \sqrt{V(t)}dW_1(t) \\ dV(t) &= \kappa\left(\theta - \left(1 - \frac{\sigma\rho}{2\kappa}\right)V(t)\right)dt + \sigma\rho\left(d\log S(t) - \mu dt\right) + \sigma\sqrt{1-\rho^2}\sqrt{V(t)}dW_2(t) \end{aligned} \tag{9.60}$$

Defining $\tilde{\kappa} = \kappa - \dfrac{\sigma\rho}{2}$ and $\tilde{\theta} = \dfrac{\theta}{1 - \dfrac{\sigma\rho}{2\kappa}} = \dfrac{\kappa\theta}{\tilde{\kappa}}$ the model can be written as:

$$\begin{aligned} d\log S(t) &= \left(\mu - \frac{1}{2}V(t)\right)dt + \sqrt{V(t)}dW_1(t) \\ dV(t) &= \tilde{\kappa}\left(\tilde{\theta} - V(t)\right)dt + \sigma\rho\left(d\log S(t) - \mu dt\right) + \sigma\sqrt{1-\rho^2}\sqrt{V(t)}dW_2(t) \end{aligned} \tag{9.61}$$

This representation forms the basis for the estimation of the Heston model. This is because the conditional variance dynamics are written as a CIR process with extra drift. With the results of Cox, Ingersoll and Ross (1985) the conditional transition distributions of the variance process can be derived as a non-central chi squared distribution. The conditional transition density of the log asset dynamics is a Gaussian distribution which can easily be seen from equation (9.61).

9.3.1 Estimation

Integration of the variance dynamics from equation (9.61) yields:

$$\begin{aligned} V(t+\Delta t) &= \tilde{V}(t) + \int_t^{t+\Delta t}\tilde{\kappa}\left(\tilde{\theta} - V(u)\right)du + \sigma\sqrt{1-\rho^2}\int_t^{t+\Delta t}\sqrt{V(u)}dW_2(u) \\ &\text{where:} \\ \tilde{V}(t) &= V(t) + \sigma\rho\left(d\log S(t) - \mu\Delta t\right) \end{aligned} \tag{9.62}$$

The results of Cox, Ingersoll, and Ross (1985) state, conditional on $\tilde{V}(t) > 0$, that:

$$\begin{aligned}
(V(t+\Delta t)|\log S(t+\Delta t), \log S(t), V(t)) &\sim c\chi_d^{'2}(\lambda) \\
&\text{where:} \\
c &= \frac{\sigma^2\left(1-\rho^2\right)\left(1-e^{-\tilde{\kappa}\Delta t}\right)}{4\tilde{\kappa}} \\
d &= \frac{4\tilde{\theta}\tilde{\kappa}}{\sigma^2\left(1-\rho^2\right)} \\
\lambda &= \frac{4\tilde{\kappa}e^{-\tilde{\kappa}\Delta t}}{\sigma^2\left(1-\rho^2\right)\left(1-e^{-\tilde{\kappa}\Delta t}\right)}\tilde{V}(t) \\
\tilde{V}(t) &= V(t) + \sigma\rho\left(\log S\left(t+\Delta t\right) - \log S(t) - \mu\Delta t\right)
\end{aligned} \tag{9.63}$$

This density is known and can easily be simulated from. As such, the optimal importance density can be used in the particle filter. The only problem is that $\tilde{V}(t)$ can become negative during simulations. This is inconsistent with the definition of variance and problematic for the distribution as it is not defined for such values. For computational speed, a negative value for $\tilde{V}(t)$ is replaced with a arbitrarily small number in my approach.

In similar vein the transition distribution for the variance process conditional on the previous variance realization is a non-central chi-square distribution. This distribution is exactly the same transition distribution as in Cox, Ingersoll, and Ross (1985). The intuition behind this distribution is that with no information on the asset value, the best you can base your prediction on is the original dynamics. Moreover, just as in equation (3.9), the dynamics with $dW_1(t)$ plugged in is mathematically equivalent to the original dynamics in equation (3.1). As this is the original CIR process, the transition density directly follows:

$$\begin{aligned}
V(t+\Delta t)|V(t) &\sim c\chi_d^{'2}(\lambda) \\
&\text{where:} \\
c &= \frac{\sigma^2\left(1-e^{-\kappa\Delta t}\right)}{4\kappa} \\
d &= \frac{4\kappa\theta}{\sigma^2} \\
\lambda &= \frac{4\kappa e^{-\kappa\Delta t}}{\sigma^2\left(1-e^{-\kappa\Delta t}\right)}V(t)
\end{aligned} \tag{9.64}$$

For the derivation of the conditional observation density I use the drift interpolation method for the integrated volatility and the integrated square root of volatility, see Andersen (2007). This method entails:

$$\begin{aligned}
\int_t^{t+\Delta t} V(u)du &= \Delta t\left[\gamma V(t) + (1-\gamma)V(t+\Delta t)\right] \\
&\text{where:} \\
&\gamma \in [0,1]
\end{aligned} \tag{9.65}$$

Integration of the log asset dynamics yields:

$$\log S(t+\Delta t) = \log S(t) + \mu\Delta t - \frac{1}{2}\int_t^{t+\Delta t} V(u)du + \int_t^{t+\Delta t}\sqrt{V(t)}dWu \tag{9.66}$$

As the quadratic variation of the process equals $\int_t^{t+\Delta t} V(u)du$, the conditional on $V(t+\Delta t), V(t), S(t)$, the observation distribution equals:

$$\log S(t+\Delta t)|V(t+\Delta t), S(t), V(t) \sim \mathcal{N}\left(\log S(t) + \mu\Delta t - \frac{1}{2}\int_t^{t+\Delta t} V(u)du, \int_t^{t+\Delta t} V(u)du\right) \overset{eq.(9.65)}{\Rightarrow}$$
$$\mathcal{N}\left(\log S(t) + \mu\Delta t - \frac{1}{2}\Delta t\left[\gamma V\left(t\right) + (1-\gamma)\, V\left(t+\Delta t\right)\right], \Delta t\left[\gamma V(t) + (1-\gamma)V(t+\Delta t)\right]\right) \tag{9.67}$$

Now all distributions are defined to establish the particle filter.The continuous particle filter is just as in Algorithm 1 if one replaces the discrete process distributions by the continuous counterpart. The resulting particle filter an adaption of the algorithm of Aihara et al. (2012).

9.3.2 Simulation

The derivation of the simulation scheme of the continuous Heston model starts from a different Cholesky decomposition. This is due to the simultaneity of the process in the conditional transition distributions. As one can see the observation transition distribution depends on $V(t+\Delta t)$ and the state transition distribution simultaneously depends on $S(t+\Delta t)$. However, this problem disappears when decomposing the original model as:

$$\begin{aligned} d\log S\left(t\right) &= \left(\mu - \frac{1}{2}V(t)\right)dt + \sqrt{V(t)}\left(\rho dW_1(t) + \sqrt{1-\rho^2}dW_2(t)\right) \\ dV(t) &= \kappa\left(\theta - V(t)\right)dt + \sigma\sqrt{V(t)}dW_2(t) \end{aligned} \tag{9.68}$$

Plugging $dW_2(t)$ into the first equation - making the observation equation conditional on the current volatility:

$$\begin{aligned} d\log S\left(t\right) &= \left(\mu - \frac{\kappa\theta\sqrt{1-\rho^2}}{\sigma} + \left(\frac{\kappa\sqrt{1-\rho^2}}{\sigma} - \frac{1}{2}\right)V(t)\right)dt + \frac{\sqrt{1-\rho^2}}{\sigma}dV(t) + \rho\sqrt{V(t)}dW_1(t) \\ dV(t) &= \kappa\left(\theta - V(t)\right)dt + \sigma\sqrt{V(t)}dW_2(t) \end{aligned} \tag{9.69}$$

The second equation is just a CIR process. The transition density is thus given by equation (9.64). The observation transition distribution follow from integrating the equation:

$$\begin{aligned} \log S(t+\Delta t) = &\log S(t) + \left(\mu - \frac{\kappa\theta\sqrt{1-\rho^2}}{\sigma}\right)\Delta t + \left(\frac{\kappa\sqrt{1-\rho^2}}{\sigma} - \frac{1}{2}\right)\int_t^{t+\Delta t} V(u)du \\ &+ \frac{\sqrt{1-\rho^2}}{\sigma}\left(V(t+\Delta t) - V(t)\right) + \rho\int_t^{t+\Delta t}\sqrt{V(u)}dW_1(u) \end{aligned} \tag{9.70}$$

Now by making use of the drift interpolation method from equation (9.65) and observing that the quadratic variation equals $\rho^2\int_t^{t+\Delta t} V(u)du$, the conditional transition distribution follows:

$$\begin{aligned} \log S(t+\Delta t) \sim \mathcal{N}\Big(&\left[\mu - \frac{1}{2}\left(1+\theta\right)\right]\Delta t + \left[\frac{\kappa\sqrt{1-\rho^2}}{\sigma} - \frac{1}{2}\right]\Delta t\left[\gamma V(t) + (1-\gamma)V(t+\Delta t) - \theta\right], \\ &\rho^2\Delta t\left[\gamma V(t) + (1-\gamma)V(t+\Delta t)\right]\Big) \end{aligned} \tag{9.71}$$

Then, for a fixed endpoint T, increment Δt, initial asset price S_0, initial variance V_0, and parameter set ψ, the simulating algorithm is given by:

Algorithm 5 Simulation Continuous Heston Model

Set $c = \dfrac{\sigma^2\left(1 - e^{-\kappa\Delta t}\right)}{4\kappa}$

Set $d = \dfrac{4\kappa\theta}{\sigma^2}$

for $t = 1, 2, \ldots, T$ **do**

Calculate $\lambda(t) = \dfrac{4\kappa e^{-\kappa\Delta t}}{\sigma^2\left(1 - e^{-\kappa\Delta t}\right)} V(t-1)$

Draw $\dfrac{V(t)}{c} \sim \chi_d^{'2}\left(\lambda(t)\right)$

Draw $\log S(t) \sim \mathcal{N}\Big(\left[\mu - \frac{1}{2}\left(1+\theta\right)\right]\Delta t + \left[\frac{\kappa\sqrt{1-\rho^2}}{\sigma} - \frac{1}{2}\right]\Delta t\left[\gamma V(t) + (1-\gamma)V(t+\Delta t) - \theta\right],$
$\rho^2\Delta t\left[\gamma V(t) + (1-\gamma)V(t+\Delta t)\right]\Big)$

end for

9.4 Heston-Hull-White Model: a Continuous Time Solution?

The transition distributions of the short rate and log asset follow directly from Equation (3.18). Due to the Cholesky ordering, the transition distribution of the short rate has remained the same as under the Hull-White model. The conditional transition distribution of the log asset is Gaussian as we deal with a composition of a Riemann and an Itô integral.

$$\left(\log\left(S\left(t\right)\right)|\mathcal{F}\left(t\right),V\left(t\right),r\left(t\right)\right)\sim\mathcal{N}\left(\mu,\sigma^{2}\right)$$

where,

$$\mu=\mu\left(t-s\right)-\frac{1}{2}\int_{s}^{t}V(u)du-\frac{l_{21}}{\sigma_{r}}\kappa_{r}\left(\int_{s}^{t}\Theta\left(u\right)\sqrt{V(u)}du-\int_{s}^{t}r\left(u\right)\sqrt{V(u)}du\right)+\frac{l_{21}}{\sigma_{r}}\int_{s}^{t}\sqrt{V(u)}dr\left(t\right)$$

$$\sigma^{2}=l_{22}^{2}\int_{s}^{t}V(u)du$$

under drift interpolation, see Andersen (2007)

$$\mu=\mu\left(t-s\right)-\frac{1}{4}\left(V(s)+V(t)\right)-\frac{l_{21}}{2\sigma_{r}}\kappa_{r}\left(\left(\Theta\left(s\right)\sqrt{V(s)}+\Theta\left(t\right)\sqrt{V(t)}\right)-\left(r\left(u\right)\sqrt{V(u)}+r\left(t\right)\sqrt{V(t)}\right)\right)$$

$$+\frac{l_{21}}{2\sigma_{r}}\left(\sqrt{V(u)}+\sqrt{V(u)}\right)\left(r\left(t\right)-r\left(s\right)\right)$$

$$\sigma^{2}=\frac{l_{22}^{2}}{2}\left(V(s)+V(t)\right)\tag{9.72}$$

Before I turn to the transitional conditional density of the variance process, I propose the following Lemmata.

Lemma 9.3. *Given the dynamics of the Ornstein-Uhlenbeck process:*

$$dX(t)=\kappa\left(\theta-X\left(t\right)\right)dt+\sigma dW(t)$$

The dynamics of the squared Ornstein-Uhlenbeck process, $Y(t)=X(t)^{2}$, are defined by:

$$dY(t)=2\kappa\left(\frac{\sigma^{2}}{2\kappa}+\theta\sqrt{Y(t)}-Y\left(t\right)\right)dt+2\sigma\sqrt{Y(t)}dW(t)$$

Proof. The result follows directly by application of Itô's Lemma to $Y(t)=X^{2}(t)$. First note that:

$$\frac{\partial Y(t)}{\partial t}=0,\frac{\partial Y(t)}{\partial X(t)}=2X(t),\frac{\partial^{2}Y(t)}{\partial X(t)^{2}}=2,d\left[X,X\right]\left(t\right)=\sigma^{2}dt$$

Then:

$$\begin{aligned}dY(t)&=\frac{\partial Y(t)}{\partial t}dt+\frac{\partial Y(t)}{\partial X(t)}dX(t)+\frac{1}{2}\frac{\partial^{2}Y(t)}{\partial X(t)^{2}}d\left[X,X\right]\left(t\right)\\&=2X(t)\left(\kappa\left(\theta-X\left(t\right)\right)dt+\sigma dW(t)\right)+\frac{1}{2}2\sigma^{2}dt\\&=2\kappa\left(\frac{\sigma^{2}}{2\kappa}+\theta\sqrt{Y(t)}-Y\left(t\right)\right)dt+2\sigma\sqrt{Y(t)}dW(t)\end{aligned}$$

□

Lemma 9.4. *The conditional transition density of the Ornstein-Uhlenbeck process is given by:*

$$\left(X(t)|X(s)=x\right)\sim\mathcal{N}\left(xe^{-\kappa(t-s)}+\theta\left(1-e^{-\kappa(t-s)}\right),\frac{\sigma^{2}}{2\kappa}\left(1-e^{-2\kappa(t-s)}\right)\right)$$

Proof. Define $f(t, X(t)) = e^{\kappa t} X(t)$. First note that:

$$\frac{\partial f(t,X(t))}{\partial t} = \kappa e^{\kappa t} X(t), \frac{\partial f(t,X(t))}{\partial X(t)} = e^{\kappa t}, \frac{\partial^2 f(t,X(t))}{\partial X(t)^2} = 0, d\left[X,X\right](t) = \sigma^2 dt$$

Then by Itô's lemma:

$$\begin{aligned} df(t,X(t)) &= \frac{\partial f(t,X(t))}{\partial t} dt + \frac{\partial f(t,X(t))}{\partial X(t)} dX(t) + \frac{1}{2}\frac{\partial^2 f(t,X(t))}{\partial X(t)^2} d\left[X,X\right](t) \\ &= \kappa e^{\kappa t} X(t) dt + e^{\kappa t}\left[\kappa\left(\theta - X\left(t\right)\right) dt + \sigma dW(t)\right] \\ d\left(e^{\kappa t} X(t)\right) &= \kappa e^{\kappa t}\theta dt + e^{\kappa t}\sigma dW(t) \end{aligned}$$

Then by integration from s to t one obtains:

$$\begin{aligned} e^{\kappa t} X(t) &= e^{\kappa s} X(s) + \int_s^t \kappa e^{\kappa u}\theta du + \int_s^t e^{\kappa u}\sigma dWu \\ X(t) &= e^{-\kappa(t-s)} X(s) + \theta\left(1 - e^{-\kappa(t-s)}\right) + e^{-\kappa t}\int_s^t e^{\kappa u}\sigma dWu \end{aligned}$$

The latter integral is an Itô integral; therefore it's normally distributed with zero mean and variance:

$$e^{-2\kappa t}\int_s^t \sigma^2 e^{2\kappa u} dt = \frac{\sigma^2}{2\kappa}\left(1 - e^{-2\kappa(t-s)}\right)$$

Then, as the last integral is an Itô integral, it is clear that:

$$\left(X(t)|X(s) = x\right) \sim \mathcal{N}\left(xe^{-\kappa(t-s)} + \theta\left(1 - e^{-\kappa(t-s)}\right), \frac{\sigma^2}{2\kappa}\left(1 - e^{-2\kappa(t-s)}\right)\right)$$

□

Corollary 9.4.1. *The conditional transition distribution of the squared Ornstein-Uhlenbeck process is a non-central chi-squared distribution:*

$$\begin{aligned} &\left(Y(t)|Y(s) = y\right) \sim \frac{\sigma^2}{2\kappa}\left(1 - e^{-2\kappa(t-s)}\right)\chi_d^2(\lambda) \\ &\textit{where,} \\ &d = 1 \\ &\lambda = \frac{\left(ye^{-\kappa(t-s)} + \theta\left(1 - e^{-\kappa(t-s)}\right)\right)^2}{\frac{\sigma^2}{2\kappa}\left(1 - e^{-2\kappa(t-s)}\right)} \end{aligned}$$

Proof. Since we know that $\left(X(t)|X(s) = x\right)$ is Gaussian, the distribution of $\tilde{X}(t) = \frac{X^2(t)2\kappa}{\sigma^2\left(1-e^{-2\kappa(t-s)}\right)}$ conditional on $X(s) = x$ is non-central chi-squared distributed with one degree of freedom and non-centrality parameter $\lambda = \frac{\left(ye^{-\kappa(t-s)}+\theta\left(1-e^{-\kappa(t-s)}\right)\right)^2}{\frac{\sigma^2}{2\kappa}\left(1-e^{-2\kappa(t-s)}\right)}$. Therefore, conditional on $X(s) = x \equiv Y(s) = \sqrt{y}$:

$$\begin{aligned} &\left(\frac{Y(t)2\kappa}{\sigma^2\left(1 - e^{-2\kappa(t-s)}\right)}|Y(s) = \sqrt{y}\right) \sim \chi_d^2(\lambda) \\ &\text{where,} \\ &d = 1 \\ &\lambda = \frac{\left(ye^{-\kappa(t-s)} + \theta\left(1 - e^{-\kappa(t-s)}\right)\right)^2}{\frac{\sigma^2}{2\kappa}\left(1 - e^{-2\kappa(t-s)}\right)} \end{aligned}$$

□

Another proof to the same result:

Proof. Start by noting that:

$$\begin{aligned}P(Y_t \leq z|Y_s = y) &= P(X_t^2 \leq z|X_s = \sqrt{y} \vee X_s = -\sqrt{y}) \\ &= P(X_t^2 \leq z|X_s = \sqrt{y})p + P(X_t^2 \leq z|X_s = -\sqrt{y})(1-p)\end{aligned}$$

where $p = \dfrac{f_{X_s}(\sqrt{y})}{f_{X_s}(\sqrt{y}) + f_{X_s}(-\sqrt{y})}$ and f_{X_s} the unconditional probability density function of X_s. As $\lim_{x\to\infty} \mathbb{E}\left[X_s|X_0 = x\right] = \lim_{x\to\infty} xe^{-\kappa(t-s)} + \theta\left(1 - e^{-\kappa(t-s)}\right) = \theta$ and $\lim_{x\to\infty} \mathbb{V}\mathrm{ar}\left[X_s|X_0 = x\right] = \lim_{x\to\infty} \frac{\sigma^2}{2\kappa}\left(1 - e^{-2\kappa(t-s)}\right) = \frac{\sigma^2}{2\kappa}$, it follows that $X_s \sim \mathcal{N}\left(\theta, \frac{\sigma^2}{2\kappa}\right)$. Then:

$$p = \frac{\phi\left(\frac{\sqrt{y}-\theta}{\frac{\sigma}{2\kappa}}\right)}{\phi\left(\frac{\sqrt{y}-\theta}{\frac{\sigma}{2\kappa}}\right) + \phi\left(\frac{-\sqrt{y}-\theta}{\frac{\sigma}{2\kappa}}\right)} = \frac{e^{4\kappa\theta\sqrt{y}/\sigma^2}}{e^{4\kappa\theta\sqrt{y}/\sigma^2} + 1}$$

Define $a_1 \equiv \mathbb{E}\left[X_t|X_s = \sqrt{y}\right] = \sqrt{y}e^{-\kappa(t-s)} + \theta\left(1 - e^{-\kappa(t-s)}\right)$, $a_2 \equiv \mathbb{E}\left[X_t|X_s = -\sqrt{y}\right] = -\sqrt{y}e^{-\kappa(t-s)} + \theta\left(1 - e^{-\kappa(t-s)}\right)$, and $b = \mathbb{V}\mathrm{ar}\left[X_t|X_s = \sqrt{y}\right] = \mathbb{V}\mathrm{ar}\left[X_t|X_s = -\sqrt{y}\right] = \frac{\sigma^2}{2\kappa}\left(1 - e^{-2\kappa(t-s)}\right)$. Then, we can write:

$$\begin{aligned}&P(-\sqrt{z} \leq X_t \leq \sqrt{z}|X_s = \sqrt{y})p + P(-\sqrt{z} \leq X_t \leq \sqrt{z}|X_s = -\sqrt{y})(1-p) = \\ &p\left[\Phi\left(\frac{\sqrt{z}-a_1}{\sqrt{b}}\right) - \Phi\left(\frac{-\sqrt{z}-a_1}{\sqrt{b}}\right)\right] + (1-p)\left[\Phi\left(\frac{\sqrt{z}-a_2}{\sqrt{b}}\right) - \Phi\left(\frac{-\sqrt{z}-a_2}{\sqrt{b}}\right)\right] \\ &p\left[\frac{1}{2}\mathrm{erf}\left(\frac{\sqrt{z}-a_1}{\sqrt{2b}}\right) - \frac{1}{2}\mathrm{erf}\left(\frac{-\sqrt{z}-a_1}{\sqrt{2b}}\right)\right] + (1-p)\left[\frac{1}{2}\mathrm{erf}\left(\frac{\sqrt{z}-a_2}{\sqrt{2b}}\right) - \frac{1}{2}\mathrm{erf}\left(\frac{-\sqrt{z}-a_2}{\sqrt{2b}}\right)\right] \overset{\mathrm{erf}(-z)=-\mathrm{erf}(z)}{=} \\ &\frac{p}{2}\left[\mathrm{erf}\left(\frac{\sqrt{z}-a_1}{\sqrt{2b}}\right) + \mathrm{erf}\left(\frac{\sqrt{z}+a_1}{\sqrt{2b}}\right)\right] + \frac{(1-p)}{2}\left[\mathrm{erf}\left(\frac{\sqrt{z}-a_2}{\sqrt{2b}}\right) + \mathrm{erf}\left(\frac{\sqrt{z}+a_2}{\sqrt{2b}}\right)\right]\end{aligned}$$

The conditional transition probability density function is obtained by differentiation of the cumulative distribution function:

$$\begin{aligned}f_{Y_t|Y_s=y}(z) &= \frac{\partial}{\partial z}\left(\frac{p}{2}\left[\mathrm{erf}\left(\frac{\sqrt{z}-a_1}{\sqrt{2b}}\right) + \mathrm{erf}\left(\frac{\sqrt{z}+a_1}{\sqrt{2b}}\right)\right] + \frac{(1-p)}{2}\left[\mathrm{erf}\left(\frac{\sqrt{z}-a_2}{\sqrt{2b}}\right) + \mathrm{erf}\left(\frac{\sqrt{z}+a_2}{\sqrt{2b}}\right)\right]\right) \\ &= \frac{p}{2}\left[\frac{2}{\sqrt{\pi}}\frac{1}{2\sqrt{2bz}}\left(e^{-\left(\frac{\sqrt{z}-a_1}{\sqrt{2b}}\right)^2} + e^{-\left(\frac{\sqrt{z}+a_1}{\sqrt{2b}}\right)^2}\right)\right] + \frac{1-p}{2}\left[\frac{2}{\sqrt{\pi}}\frac{1}{2\sqrt{2bz}}\left(e^{-\left(\frac{\sqrt{z}-a_2}{\sqrt{2b}}\right)^2} + e^{-\left(\frac{\sqrt{z}+a_2}{\sqrt{2b}}\right)^2}\right)\right] \\ &= \frac{p}{2\sqrt{2\pi bz}}\left(e^{-\frac{(\sqrt{z}-a_1)^2}{2b}} + e^{-\frac{(\sqrt{z}+a_1)^2}{2b}}\right) + \frac{1-p}{2\sqrt{2\pi bz}}\left(e^{-\frac{(\sqrt{z}-a_2)^2}{2b}} + e^{-\frac{(\sqrt{z}+a_2)^2}{2b}}\right)\end{aligned}$$

Plugging p into the expression above yields:

$$\begin{aligned}f_{Y_t|Y_s=y}(z) &= \frac{e^{4\kappa\theta\sqrt{y}/\sigma^2}}{e^{4\kappa\theta\sqrt{y}/\sigma^2}+1}\frac{1}{2\sqrt{2\pi bz}}\left(e^{-\frac{(\sqrt{z}-a_1)^2}{2b}} + e^{-\frac{(\sqrt{z}+a_1)^2}{2b}}\right) + \frac{1}{e^{4\kappa\theta\sqrt{y}/\sigma^2}+1}\frac{1}{2\sqrt{2\pi bz}}\left(e^{-\frac{(\sqrt{z}-a_2)^2}{2b}} + e^{-\frac{(\sqrt{z}+a_2)^2}{2b}}\right) \\ &= \frac{1}{e^{4\kappa\theta\sqrt{y}/\sigma^2}+1}\frac{1}{2\sqrt{2\pi bz}}\left[e^{4\kappa\theta\sqrt{y}/\sigma^2}\left(e^{-\frac{(\sqrt{z}-a_1)^2}{2b}} + e^{-\frac{(\sqrt{z}+a_1)^2}{2b}}\right) + e^{-\frac{(\sqrt{z}-a_2)^2}{2b}} + e^{-\frac{(\sqrt{z}+a_2)^2}{2b}}\right]\end{aligned}$$

Note that by substituting b back in:

$$\frac{1}{2\sqrt{2\pi bz}} = \frac{\sqrt{\kappa}}{2\sigma\sqrt{\pi(1-e^{-2\kappa(t-s)})z}}$$

Now note that plugging a_1, a_2, and b:

$$e^{-\frac{(\sqrt{z}-a_1)^2}{2b}} + e^{-\frac{(\sqrt{z}+a_1)^2}{2b}} = e^{-\frac{\kappa}{\sigma^2\left(1-e^{-2\kappa(t-s)}\right)}(z+ye^{-2\kappa(t-s)}+\theta^2(1-e^{-\kappa(t-s)})^2-2\sqrt{z}\theta(1-e^{-\kappa(t-s)}))}$$
$$\times\left[e^{-\frac{2\kappa\sqrt{y}e^{-\kappa(t-s)}}{\sigma^2\left(1-e^{-2\kappa(t-s)}\right)}(\theta(1-e^{-\kappa(t-s)})-\sqrt{z})} + e^{-\frac{2\kappa\sqrt{y}e^{-\kappa(t-s)}}{\sigma^2\left(1-e^{-2\kappa(t-s)}\right)}(-\theta(1-e^{-\kappa(t-s)})+\sqrt{z})}\right]$$

But also (!):

$$= e^{-\frac{(\sqrt{z}-a_2)^2}{2b}} + e^{-\frac{(\sqrt{z}+a_2)^2}{2b}}$$

Factoring this term out yields:

$$f_{Y_t|Y_s=y}(z) = \frac{1}{e^{4\kappa\theta\sqrt{y}/\sigma^2}+1}\frac{\sqrt{\kappa}}{2\sigma\sqrt{\pi(1-e^{-2\kappa(t-s)})z}}\left(e^{4\kappa\theta\sqrt{y}/\sigma^2}+1\right)$$
$$\times e^{-\frac{\kappa}{\sigma^2\left(1-e^{-2\kappa(t-s)}\right)}(z+ye^{-2\kappa(t-s)}+\theta^2(1-e^{-\kappa(t-s)})^2-2\sqrt{z}\theta(1-e^{-\kappa(t-s)}))}$$
$$\times\left[e^{-\frac{2\kappa\sqrt{y}e^{-\kappa(t-s)}}{\sigma^2\left(1-e^{-2\kappa(t-s)}\right)}(\theta(1-e^{-\kappa(t-s)})-\sqrt{z})} + e^{-\frac{2\kappa\sqrt{y}e^{-\kappa(t-s)}}{\sigma^2\left(1-e^{-2\kappa(t-s)}\right)}(-\theta(1-e^{-\kappa(t-s)})+\sqrt{z})}\right]$$
$$= \frac{\sqrt{\kappa}}{2\sigma\sqrt{\pi(1-e^{-2\kappa(t-s)})z}}e^{-\frac{\kappa}{\sigma^2\left(1-e^{-2\kappa(t-s)}\right)}(z+ye^{-2\kappa(t-s)}+\theta^2(1-e^{-\kappa(t-s)})^2-2\sqrt{z}\theta(1-e^{-\kappa(t-s)}))}$$
$$\times\left[e^{-\frac{2\kappa\sqrt{y}e^{-\kappa(t-s)}}{\sigma^2\left(1-e^{-2\kappa(t-s)}\right)}(\theta(1-e^{-\kappa(t-s)})-\sqrt{z})} + e^{-\frac{2\kappa\sqrt{y}e^{-\kappa(t-s)}}{\sigma^2\left(1-e^{-2\kappa(t-s)}\right)}(-\theta(1-e^{-\kappa(t-s)})+\sqrt{z})}\right]$$

The term in brackets can be written as a hyperbolic cosine function as $e^x + e^{-x} = 2\cosh(x)$:

$$f_{Y_t|Y_s=y}(z) = \frac{\sqrt{\kappa}}{2\sigma\sqrt{\pi(1-e^{-2\kappa(t-s)})z}}e^{-\frac{\kappa}{\sigma^2\left(1-e^{-2\kappa(t-s)}\right)}(z+ye^{-2\kappa(t-s)}+\theta^2(1-e^{-\kappa(t-s)})^2-2\sqrt{z}\theta(1-e^{-\kappa(t-s)}))}$$
$$\times 2\cosh\left(-\frac{2\kappa\sqrt{y}e^{-\kappa(t-s)}}{\sigma^2\left(1-e^{-2\kappa(t-s)}\right)}(\theta(1-e^{-\kappa(t-s)})-\sqrt{z})\right)$$
$$= \frac{\sqrt{\kappa}}{\sigma\sqrt{\pi(1-e^{-2\kappa(t-s)})z}}e^{-\frac{\kappa}{\sigma^2\left(1-e^{-2\kappa(t-s)}\right)}(z+ye^{-2\kappa(t-s)}+\theta^2(1-e^{-\kappa(t-s)})^2-2\sqrt{z}\theta(1-e^{-\kappa(t-s)}))}$$
$$\times\cosh\left(-\frac{2\kappa\sqrt{y}e^{-\kappa(t-s)}}{\sigma^2\left(1-e^{-2\kappa(t-s)}\right)}(\theta(1-e^{-\kappa(t-s)})-\sqrt{z})\right)$$

Write the part in brackets in the exponent as:

$$-\frac{\kappa}{\sigma^2\left(1-e^{-2\kappa(t-s)}\right)}(z+ye^{-2\kappa(t-s)}+\theta^2(1-e^{-\kappa(t-s)})^2-2\sqrt{z}\theta(1-e^{-\kappa(t-s)})) =$$
$$e^{-2\kappa(t-s)}\left[\left(-e^{\kappa(t-s)}\theta+e^{\kappa(t-s)}\sqrt{z}+\theta\right)^2+y\right]$$

Then:

$$
\begin{aligned}
f_{Y_t|Y_s=y}(z) &= \frac{\sqrt{\kappa}}{\sigma\sqrt{\pi(1-e^{-2\kappa(t-s)})z}} e^{-\frac{\kappa}{\sigma^2\left(1-e^{-2\kappa(t-s)}\right)}\left(e^{-2\kappa(t-s)}\left[\left(-e^{\kappa(t-s)}\theta+e^{\kappa(t-s)}\sqrt{z}+\theta\right)^2+y\right]\right)} \\
&\times \cosh\left(-\frac{2\kappa\sqrt{y}e^{-\kappa(t-s)}}{\sigma^2\left(1-e^{-2\kappa(t-s)}\right)}(\theta(1-e^{-\kappa(t-s)})-\sqrt{z})\right) \\
&= \frac{\sqrt{\kappa}}{\sigma\sqrt{\pi(1-e^{-2\kappa(t-s)})z}} e^{-\frac{\kappa\left(\left(\theta(1-e^{\kappa(t-s)})+e^{\kappa(t-s)}\sqrt{z}\right)^2+y\right)}{\sigma^2\left(e^{2\kappa(t-s)}-1\right)}} \\
&\times \cosh\left(-\frac{2\kappa\sqrt{y}e^{-\kappa(t-s)}}{\sigma^2\left(1-e^{-2\kappa(t-s)}\right)}(\theta(1-e^{-\kappa(t-s)})-\sqrt{z})\right) \\
&= \frac{\sqrt{\kappa}}{\sigma\sqrt{\pi(1-e^{-2\kappa(t-s)})z}} e^{-\frac{\kappa\left(\left(\theta(1-e^{\kappa(t-s)})+e^{\kappa(t-s)}\sqrt{z}\right)^2+y\right)}{\sigma^2\left(e^{2\kappa(t-s)}-1\right)}} \\
&\times \cosh\left(-\frac{2\kappa\sqrt{y}e^{-\kappa(t-s)}}{\sigma^2\left(1-e^{-2\kappa(t-s)}\right)}\frac{-1}{e^{\kappa(t-s)}}(\theta(1-e^{\kappa(t-s)})-e^{\kappa(t-s)}\sqrt{z})\right) \\
&= \frac{\sqrt{\kappa}}{\sigma\sqrt{\pi(1-e^{-2\kappa(t-s)})z}} e^{-\frac{\kappa\left(\left(\theta(1-e^{\kappa(t-s)})+e^{\kappa(t-s)}\sqrt{z}\right)^2+y\right)}{\sigma^2\left(e^{2\kappa(t-s)}-1\right)}} \\
&\times \cosh\left(-\frac{2\kappa\sqrt{y}e^{-\kappa(t-s)}}{\sigma^2\left(1-e^{-2\kappa(t-s)}\right)}\frac{-1}{e^{\kappa(t-s)}}(\theta(1-e^{\kappa(t-s)})-e^{\kappa(t-s)}\sqrt{z})\right) \\
&= \frac{\sqrt{\kappa}e^{\kappa(t-s)}}{\sigma\sqrt{(e^{2\kappa(t-s)}-1)\pi z}} e^{-\frac{\kappa\left(\left(\theta(1-e^{\kappa(t-s)})+e^{\kappa(t-s)}\sqrt{z}\right)^2+y\right)}{\sigma^2\left(e^{2\kappa(t-s)}-1\right)}} \\
&\times \cosh\left(\frac{2\kappa\sqrt{y}}{\sigma^2\left(e^{2\kappa(t-s)}-1\right)}(\theta(1-e^{\kappa(t-s)})-e^{\kappa(t-s)}\sqrt{z})\right)
\end{aligned}
$$

$\lambda \equiv \frac{\sigma^2}{2\kappa}(e^{2\kappa(t-s)}-1)$

$$
f_{Y_t|Y_s=y}(z) = \frac{e^{\kappa(t-s)}}{\sqrt{2\lambda}\sqrt{\pi z}} e^{-\frac{\left(\left(\theta(1-e^{\kappa(t-s)})+e^{\kappa(t-s)}\sqrt{z}\right)^2+y\right)}{2\lambda}} \cosh\left(\frac{\sqrt{y}}{\lambda}(\theta(1-e^{\kappa(t-s)})-e^{\kappa(t-s)}\sqrt{z})\right)
$$

The hyperbolic cosine is directly related with a specific modified Bessel function of the first kind. Namely:

$$
\sqrt{\frac{2}{\pi x}}\cosh(x) = I_{-\frac{1}{2}}(x)
$$

Then:

$$
\begin{aligned}
f_{Y_t|Y_s=y}(z) &= \frac{e^{\kappa(t-s)}}{\sqrt{2\lambda}\sqrt{\pi z}} e^{-\frac{\left(\left(\theta(1-e^{\kappa(t-s)})+e^{\kappa(t-s)}\sqrt{z}\right)^2+y\right)}{2\lambda}} \sqrt{\frac{\pi\frac{\sqrt{y}}{\lambda}(\theta(1-e^{\kappa(t-s)})-e^{\kappa(t-s)}\sqrt{z})}{2}} \\
&\times I_{-\frac{1}{2}}\left(\frac{\sqrt{y}}{\lambda}(\theta(1-e^{\kappa(t-s)})-e^{\kappa(t-s)}\sqrt{z})\right) \\
&= \frac{1}{2}\frac{e^{\kappa(t-s)}}{\lambda\sqrt{z}} e^{-\frac{\left(\left(\theta(1-e^{\kappa(t-s)})+e^{\kappa(t-s)}\sqrt{z}\right)^2+y\right)}{2\lambda}} \sqrt{\sqrt{y}(\theta(1-e^{\kappa(t-s)})-e^{\kappa(t-s)}\sqrt{z})} \\
&\times I_{-\frac{1}{2}}\left(\frac{\sqrt{y}}{\lambda}(\theta(1-e^{\kappa(t-s)})-e^{\kappa(t-s)}\sqrt{z})\right)
\end{aligned}
$$

From which we see that Y is a scaled non-central chi squared distribution with one degree of freedom. More specifically, $\frac{Y(t)2\kappa}{\sigma^2\left(1-e^{-2\kappa(t-s)}\right)} \sim \chi_1^2(\nu)$, with $\nu = \frac{\left(\sqrt{y}e^{-\kappa(t-s)}+\theta\left(1-e^{-\kappa(t-s)}\right)\right)^2}{\frac{\sigma^2}{2\kappa}\left(1-e^{-2\kappa(t-s)}\right)}$ □

The dynamics of the volatility process in equation (3.18) have the form of a squared Ornstein-Uhlenbeck process. Therefore, one can deduct that the conditional volatility transition distribution is a non-central chi squared distribution. More specifically:

$$
\begin{aligned}
V(t)|V(s) &\sim \chi_1^2(\lambda) \\
\lambda &= \frac{2\tilde{\kappa}\left(V(s)e^{-\tilde{\kappa}(t-s)}+\tilde{\theta}\left(1-e^{-\tilde{\kappa}(t-s)}\right)\right)^2}{\tilde{\sigma}^2\left(1-e^{-2\tilde{\kappa}(t-s)}\right)} \\
\tilde{\kappa} &= \frac{\kappa_V}{2}-\frac{l_{32}\sigma_V}{4l_{22}} \\
\tilde{\sigma} &= \frac{l_{33}\sigma_V}{2} \\
\tilde{\theta} &= \frac{\left(l_{31}-\frac{l_{21}l_{32}}{l_{22}}\right)\frac{\sigma_V}{\sigma_r}\left(\frac{dr(t)}{dt}-\kappa_r\left(\Theta\left(t\right)-r\left(t\right)\right)\right)}{2\tilde{\kappa}}
\end{aligned}
\tag{9.73}
$$

$$
\begin{aligned}
dV(t) = &\left(\kappa_V\theta+\frac{l_{32}\sigma_V}{l_{22}}\left(\frac{d\log\left(S\left(t\right)\right)}{dt}-\mu\right)-\left(\kappa_V-\frac{l_{32}\sigma_V}{2l_{22}}\right)V(t)\right. \\
&\left.+\left(l_{31}-\frac{l_{21}l_{32}}{l_{22}}\right)\frac{\sigma_V}{\sigma_r}\left(\frac{dr(t)}{dt}-\kappa_r\left(\Theta\left(t\right)-r\left(t\right)\right)\right)\sqrt{V(t)}\right)dt+l_{33}\sigma_V\sqrt{V(t)}dW_3\left(t\right)
\end{aligned}
$$

$dXt = \kappa(\theta - Xt)dt + \sigma dWt = dXt = (\kappa\theta - \kappa Xt)dt + \sigma dWt$
$dYt = 2\kappa\left(\frac{\sigma^2}{2\kappa}+\theta\sqrt{Y(t)}-Y\left(t\right)\right)dt + 2\sigma\sqrt{Y(t)}dW(t)$

If true, it must also hold that $\frac{l_{33}^2\sigma_V^2}{4} = \kappa_V\theta + \frac{l_{32}\sigma_V}{l_{22}}\left(\frac{d\log(S(t))}{dt}-\mu\right) \equiv \tilde{\sigma}^2$
$l_{33}^2 = 1-\rho_{13}^2-\frac{(\rho_{23}-\rho_{13}\rho_{12})^2}{1-\rho_{12}^2}$
$\frac{l_{32}}{l_{22}} = \frac{\rho_{23}-\rho_{12}\rho_{13}}{1-\rho_{12}^2}$
When one allows for full correlation one sees that there is some indeterminacy in the model. This is not surprising. Once two processes are set, the third is for a part dictated through the correlations. A solution is to set a correlation $\rho_{1,3}$ to zero, then The hybrid model is dictated by:

$$
\begin{aligned}
dr\left(t\right) &= \kappa_r\left(\Theta\left(t\right)-r\left(t\right)\right)dt+\sigma_r dW_r\left(t\right) \\
d\log\left(S\left(t\right)\right) &= \left(\mu-\frac{1}{2}V(t)\right)dt+\sqrt{V(t)}dW_S(t) \\
dV(t) &= \kappa_V\left(\theta-V(t)\right)dt+\sigma_V\sqrt{V(t)}dW_V(t)
\end{aligned}
\tag{9.74}
$$

where the Wiener processes are correlated by covariance matrix $\Sigma = \begin{bmatrix} 1 & \rho_{12} & 0 \\ \rho_{12} & 1 & \rho_{23} \\ 0 & \rho_{23} & 1 \end{bmatrix}$. A Cholesky decomposition on $\Sigma = LL'$, yields $L = \begin{bmatrix} l_{11} & 0 & 0 \\ l_{21} & l_{22} & 0 \\ l_{31} & l_{32} & l_{33} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ \rho_{12} & \sqrt{1-\rho_{12}^2} & 0 \\ 0 & \frac{\rho_{23}}{\sqrt{1-\rho_{12}^2}} & \sqrt{1-\frac{\rho_{23}^2}{1-\rho_{12}^2}} \end{bmatrix}$. Then, by using the general notation of the lower triangular matrix L for compactness, the model can be written as:

$$\begin{aligned} dr(t) &= \kappa_r(\Theta(t) - r(t))\,dt + \sigma_r dW_1(t) \\ d\log(S(t)) &= \left(\mu - \frac{1}{2}V(t)\right)dt + \sqrt{V(t)}\,(l_{21}dW_1(t) + l_{22}dW_2(t)) \\ dV(t) &= \kappa_V(\theta - V(t))\,dt + \sigma_V\sqrt{V(t)}\,(l_{32}dW_2(t) + l_{33}dW_3(t)) \end{aligned} \tag{9.75}$$

$$\begin{aligned} dr(t) &= \kappa_r(\Theta(t) - r(t))\,dt + \sigma_r dW_1(t) \\ d\log(S(t)) &= \left(\mu - \frac{1}{2}V(t) - \frac{l_{21}}{\sigma_r}\kappa_r(\Theta(t) - r(t))\sqrt{V(t)}\right)dt + \frac{l_{21}}{\sigma_r}\sqrt{V(t)}dr(t) + l_{22}\sqrt{V(t)}dW_2(t) \\ dV(t) &= \left(\kappa_V\theta + \frac{l_{32}\sigma_V}{l_{22}}\left(\frac{d\log(S(t))}{dt} - \mu\right) - \left(\kappa_V - \frac{l_{32}\sigma_V}{2l_{22}}\right)V(t)\right. \\ &\quad \left. - \left(\frac{l_{21}l_{32}}{l_{22}}\right)\frac{\sigma_V}{\sigma_r}\left(\frac{dr(t)}{dt} - \kappa_r(\Theta(t) - r(t))\right)\sqrt{V(t)}\right)dt + l_{33}\sigma_V\sqrt{V(t)}dW_3(t) \end{aligned} \tag{9.76}$$

However, this specification does not reflect the true dynamics. Hence, we can conclude the dynamics are not a squared-Ornstein-Uhlenbeck process and thus follows a different distribution, which is a point for further research.

9.5 Tables

9.5.1 Transaction Costs

Table 12: Cumulative Relative Transaction costs in bps During QE

	E =0.2	0.4	0.6	0.8	$\hat{\sigma}_{0.2}$	$\hat{\sigma}_{0.4}$	$\hat{\sigma}_{0.6}$	$\hat{\sigma}_{0.8}$
H=0.0	0.615	0.533	0.454	0.357	0.144	0.330	0.582	0.929
0.2	1.470	1.664	2.502	2.936	0.611	1.398	11.440	6.315
0.4	2.661	3.040	3.607	4.937	1.116	2.777	4.162	13.782
0.6	3.911	4.561	5.294	6.005	1.654	8.164	6.484	7.921
0.8	5.159	5.638	7.749	8.545	2.372	4.295	29.425	45.963
1.0	6.889	7.156	8.226	8.835	15.124	5.917	10.628	24.343

E denotes the equity-to-total-asset-value ratio, and H defines the hedge ratio.
The left panel shows the mean of the simulation outcomes, the right panel the standard deviation.

Table 13: Cumulative Relative Transaction costs in bps in EE

	E =0.2	0.4	0.6	0.8	$\hat{\sigma}_{0.2}$	$\hat{\sigma}_{0.4}$	$\hat{\sigma}_{0.6}$	$\hat{\sigma}_{0.8}$
H=0.0	0.458	0.333	0.214	0.103	0.065	0.079	0.082	0.067
0.2	0.476	0.374	0.299	0.252	0.113	0.155	0.202	0.296
0.4	0.699	0.557	0.466	0.441	0.190	0.258	0.339	0.532
0.6	0.948	0.750	0.642	0.634	0.270	0.353	0.484	0.747
0.8	1.200	0.946	0.815	0.831	0.342	0.443	0.615	0.944
1.0	1.449	1.135	0.992	1.037	0.406	0.526	0.737	1.154

E denotes the equity-to-total-asset-value ratio, and H defines the hedge ratio.
The left panel shows the mean of the simulation outcomes, the right panel the standard deviation.

9.5.2 Fund Performance

Table 14: Fund Value corrected for initial size during QE

	E =0.2	0.4	0.6	0.8	$\hat{\sigma}_{0.2}$	$\hat{\sigma}_{0.4}$	$\hat{\sigma}_{0.6}$	$\hat{\sigma}_{0.8}$
5 Years								
H =0.0	1.017	1.049	1.061	1.082	0.133	0.265	0.406	0.544
0.2	1.022	1.051	1.066	1.085	0.126	0.260	0.405	0.541
0.4	1.027	1.054	1.070	1.081	0.125	0.257	0.404	0.538
0.6	1.025	1.053	1.065	1.078	0.127	0.256	0.399	0.537
0.8	1.022	1.046	1.056	1.066	0.131	0.256	0.394	0.526
1.0	1.016	1.040	1.047	1.063	0.138	0.258	0.389	0.524
10 Years								
0.0	1.071	1.130	1.190	1.244	0.216	0.433	0.662	0.909
0.2	1.083	1.143	1.201	1.250	0.207	0.430	0.660	0.907
0.4	1.089	1.148	1.206	1.245	0.206	0.426	0.661	0.900
0.6	1.090	1.147	1.202	1.239	0.205	0.423	0.657	0.890
0.8	1.086	1.145	1.194	1.232	0.208	0.425	0.651	0.884
1.0	1.080	1.131	1.179	1.229	0.217	0.420	0.639	0.880
15 Years								
0.0	1.168	1.272	1.382	1.457	0.314	0.625	0.987	1.357
0.2	1.174	1.284	1.386	1.468	0.295	0.619	0.977	1.366
0.4	1.178	1.289	1.390	1.464	0.286	0.613	0.968	1.345
0.6	1.179	1.291	1.385	1.468	0.284	0.606	0.958	1.340
0.8	1.182	1.283	1.381	1.460	0.286	0.594	0.948	1.321
1.0	1.183	1.274	1.379	1.464	0.286	0.587	0.933	1.306
20 Years								
0.0	1.247	1.369	1.476	1.567	0.398	0.774	1.220	1.747
0.2	1.264	1.389	1.492	1.590	0.375	0.763	1.209	1.765
0.4	1.280	1.416	1.515	1.612	0.372	0.769	1.200	1.773
0.6	1.296	1.445	1.548	1.646	0.369	0.767	1.224	1.776
0.8	1.319	1.472	1.571	1.670	0.373	0.774	1.207	1.751
1.0	1.343	1.504	1.624	1.713	0.383	0.772	1.205	1.738

E denotes the equity-to-total-asset-value ratio, and H defines the hedge ratio.
The left panel shows the mean of the simulation outcomes, the right panel the standard deviation.

Table 15: Fund Value corrected for initial size in EE

	E =0.2	0.4	0.6	0.8	$\hat{\sigma}_{0.2}$	$\hat{\sigma}_{0.4}$	$\hat{\sigma}_{0.6}$	$\hat{\sigma}_{0.8}$
5 Years								
H =0.0	1.289	1.390	1.488	1.591	0.125	0.260	0.424	0.596
0.2	1.327	1.419	1.506	1.593	0.124	0.261	0.424	0.593
0.4	1.313	1.405	1.494	1.594	0.120	0.256	0.418	0.594
0.6	1.303	1.395	1.491	1.592	0.118	0.251	0.413	0.591
0.8	1.295	1.392	1.490	1.593	0.115	0.248	0.410	0.588
1.0	1.291	1.391	1.490	1.592	0.115	0.247	0.406	0.586
10 Years								
0.0	1.694	1.967	2.244	2.576	0.246	0.545	0.904	1.366
0.2	1.761	2.011	2.281	2.579	0.245	0.542	0.914	1.356
0.4	1.745	1.998	2.270	2.573	0.238	0.530	0.904	1.348
0.6	1.730	1.986	2.263	2.572	0.230	0.524	0.896	1.339
0.8	1.718	1.983	2.263	2.574	0.227	0.521	0.890	1.337
1.0	1.713	1.983	2.263	2.570	0.228	0.519	0.887	1.329
15 Years								
0.0	2.271	2.866	3.579	4.414	0.430	1.011	1.817	2.911
0.2	2.386	2.963	3.649	4.429	0.428	1.014	1.824	2.901
0.4	2.376	2.962	3.641	4.441	0.413	0.999	1.800	2.909
0.6	2.375	2.948	3.627	4.444	0.407	0.980	1.782	2.894
0.8	2.359	2.950	3.649	4.448	0.398	0.976	1.787	2.886
1.0	2.356	2.970	3.652	4.442	0.396	0.977	1.772	2.864
20 Years								
0.0	3.042	4.149	5.599	7.459	0.700	1.710	3.338	5.834
0.2	3.260	4.366	5.742	7.535	0.698	1.721	3.320	5.842
0.4	3.265	4.388	5.757	7.547	0.666	1.688	3.277	5.823
0.6	3.282	4.422	5.788	7.558	0.655	1.676	3.280	5.777
0.8	3.298	4.452	5.840	7.558	0.634	1.658	3.266	5.721
1.0	3.329	4.524	5.884	7.635	0.636	1.653	3.246	5.768

E denotes the equity-to-total-asset-value ratio, and H defines the hedge ratio.
The left panel shows the mean of the simulation outcomes, the right panel the standard deviation.

Table 16: Sharpe Ratio during QE

	E =0.2	E =0.4	E =0.6	E =0.8
5 Years				
H =0.0	0.022	0.001	-0.046	-0.079
0.2	0.058	0.014	-0.037	-0.078
0.4	0.075	0.012	-0.042	-0.086
0.6	0.047	0.003	-0.051	-0.090
0.8	0.007	-0.013	-0.065	-0.103
1.0	-0.022	-0.028	-0.080	-0.103
10 Years				
0.0	0.019	-0.012	-0.041	-0.071
0.2	0.006	-0.010	-0.041	-0.068
0.4	0.014	-0.010	-0.037	-0.066
0.6	0.011	-0.006	-0.030	-0.064
0.8	0.022	0.003	-0.027	-0.065
1.0	0.030	0.003	-0.025	-0.055
15 Years				
0.0	0.093	0.062	0.027	-0.012
0.2	0.077	0.058	0.025	-0.012
0.4	0.078	0.061	0.028	-0.016
0.6	0.071	0.062	0.029	-0.000
0.8	0.074	0.065	0.033	0.010
1.0	0.093	0.075	0.031	0.001
20 Years				
0.0	-0.035	-0.068	-0.109	-0.142
0.2	0.006	-0.047	-0.099	-0.124
0.4	0.038	-0.034	-0.065	-0.097
0.6	0.055	-0.015	-0.040	-0.069
0.8	0.086	0.032	-0.007	-0.046
1.0	0.124	0.068	0.019	-0.016

E denotes the equity-to-total-asset-value ratio, and H defines the hedge ratio.

Table 17: Sharpe Ratio in EE

	E =0.2	E =0.4	E =0.6	E =0.8
5 Years				
H =0.0	0.326	0.295	0.250	0.206
0.2	0.451	0.338	0.266	0.209
0.4	0.399	0.322	0.260	0.210
0.6	0.380	0.314	0.270	0.211
0.8	0.359	0.311	0.268	0.210
1.0	0.360	0.312	0.264	0.210
10 Years				
0.0	0.307	0.275	0.245	0.217
0.2	0.357	0.298	0.258	0.219
0.4	0.376	0.307	0.261	0.224
0.6	0.395	0.308	0.266	0.226
0.8	0.382	0.316	0.272	0.229
1.0	0.384	0.332	0.275	0.230
15 Years				
0.0	0.335	0.320	0.297	0.272
0.2	0.423	0.371	0.327	0.280
0.4	0.473	0.393	0.336	0.289
0.6	0.506	0.401	0.337	0.294
0.8	0.506	0.407	0.351	0.301
1.0	0.515	0.424	0.364	0.307
20 Years				
0.0	0.308	0.287	0.261	0.230
0.2	0.391	0.331	0.283	0.239
0.4	0.464	0.354	0.292	0.242
0.6	0.518	0.378	0.301	0.247
0.8	0.573	0.400	0.313	0.255
1.0	0.618	0.426	0.330	0.266

E denotes the equity-to-total-asset-value ratio, and H defines the hedge ratio.

9.5.3 Funding Ratio

Table 18: Funding Ratio Volatility corrected for initial size during QE

	E =0.2	0.4	0.6	0.8	$\hat{\sigma}_{0.2}$	$\hat{\sigma}_{0.4}$	$\hat{\sigma}_{0.6}$	$\hat{\sigma}_{0.8}$
5 Years								
H =0.0	1.011	1.047	1.073	1.095	0.185	0.307	0.447	0.582
0.2	1.017	1.051	1.077	1.098	0.174	0.295	0.439	0.576
0.4	1.024	1.057	1.077	1.097	0.160	0.287	0.432	0.573
0.6	1.025	1.058	1.074	1.088	0.147	0.280	0.426	0.564
0.8	1.025	1.056	1.067	1.083	0.135	0.275	0.418	0.557
1.0	1.021	1.049	1.061	1.078	0.125	0.264	0.410	0.550
10 Years								
0.0	1.043	1.115	1.178	1.243	0.279	0.474	0.692	0.944
0.2	1.053	1.123	1.184	1.239	0.256	0.460	0.682	0.925
0.4	1.054	1.118	1.189	1.222	0.238	0.440	0.680	0.902
0.6	1.052	1.116	1.183	1.216	0.220	0.429	0.671	0.893
0.8	1.049	1.109	1.170	1.206	0.200	0.417	0.656	0.879
1.0	1.044	1.096	1.158	1.193	0.188	0.403	0.644	0.859
15 Years								
0.0	1.070	1.175	1.269	1.377	0.365	0.622	0.931	1.325
0.2	1.064	1.169	1.269	1.377	0.325	0.590	0.914	1.313
0.4	1.065	1.170	1.273	1.373	0.305	0.577	0.906	1.295
0.6	1.064	1.169	1.272	1.367	0.282	0.561	0.896	1.276
0.8	1.063	1.173	1.268	1.361	0.261	0.555	0.884	1.256
1.0	1.067	1.167	1.267	1.357	0.250	0.542	0.869	1.230
20 Years								
0.0	1.079	1.195	1.293	1.407	0.434	0.730	1.098	1.628
0.2	1.081	1.208	1.308	1.431	0.396	0.709	1.091	1.657
0.4	1.093	1.222	1.329	1.429	0.386	0.699	1.083	1.617
0.6	1.101	1.246	1.344	1.450	0.372	0.696	1.072	1.599
0.8	1.116	1.267	1.373	1.482	0.359	0.689	1.072	1.579
1.0	1.135	1.301	1.420	1.531	0.361	0.698	1.072	1.589

E denotes the equity-to-total-asset-value ratio, and H defines the hedge ratio.
The left panel shows the mean of the simulation outcomes, the right panel the standard deviation.

Table 19: Funding Ratio Volatility corrected for initial size in EE

	E =0.2	0.4	0.6	0.8	$\hat{\sigma}_{0.2}$	$\hat{\sigma}_{0.4}$	$\hat{\sigma}_{0.6}$	$\hat{\sigma}_{0.8}$
5 Years								
H =0.0	1.206	1.307	1.407	1.504	0.154	0.277	0.426	0.582
0.2	1.241	1.333	1.421	1.508	0.151	0.276	0.422	0.579
0.4	1.228	1.320	1.409	1.505	0.142	0.268	0.413	0.574
0.6	1.218	1.312	1.404	1.505	0.134	0.258	0.406	0.571
0.8	1.210	1.307	1.403	1.505	0.126	0.251	0.402	0.567
1.0	1.207	1.307	1.402	1.504	0.120	0.246	0.400	0.565
10 Years								
0.0	1.466	1.715	1.980	2.281	0.278	0.522	0.842	1.245
0.2	1.519	1.752	2.008	2.284	0.274	0.517	0.839	1.235
0.4	1.503	1.739	1.998	2.276	0.260	0.501	0.830	1.223
0.6	1.491	1.726	1.991	2.280	0.247	0.487	0.819	1.223
0.8	1.479	1.723	1.991	2.280	0.233	0.482	0.813	1.218
1.0	1.474	1.725	1.990	2.272	0.226	0.474	0.807	1.206
15 Years								
0.0	1.806	2.292	2.889	3.557	0.449	0.873	1.527	2.384
0.2	1.890	2.368	2.949	3.579	0.439	0.870	1.531	2.380
0.4	1.879	2.363	2.943	3.578	0.417	0.853	1.513	2.372
0.6	1.873	2.359	2.939	3.583	0.397	0.841	1.503	2.367
0.8	1.860	2.360	2.952	3.582	0.383	0.835	1.500	2.354
1.0	1.858	2.372	2.957	3.598	0.373	0.829	1.491	2.367
20 Years								
0.0	2.227	3.077	4.189	5.658	0.665	1.394	2.627	4.613
0.2	2.378	3.231	4.301	5.706	0.664	1.402	2.625	4.608
0.4	2.382	3.252	4.316	5.727	0.637	1.386	2.612	4.615
0.6	2.397	3.275	4.321	5.759	0.619	1.373	2.584	4.629
0.8	2.399	3.300	4.364	5.751	0.592	1.366	2.583	4.583
1.0	2.416	3.326	4.397	5.787	0.584	1.328	2.569	4.588

E denotes the equity-to-total-asset-value ratio, and H defines the hedge ratio.
The left panel shows the mean of the simulation outcomes, the right panel the standard deviation.

9.5.4 Insolvency Probabilities

Table 20: Insolvency Relative to MVEV probability estimates during QE

	E =0.2	0.4	0.6	0.8	$\hat{\sigma}_{0.2}$	$\hat{\sigma}_{0.4}$	$\hat{\sigma}_{0.6}$	$\hat{\sigma}_{0.8}$
5 Years								
H =0.0	0.095	0.173	0.224	0.256	0.293	0.378	0.417	0.437
0.2	0.090	0.171	0.230	0.266	0.286	0.377	0.421	0.442
0.4	0.090	0.175	0.238	0.273	0.286	0.380	0.426	0.446
0.6	0.096	0.187	0.244	0.291	0.295	0.390	0.430	0.454
0.8	0.101	0.205	0.258	0.305	0.301	0.404	0.438	0.461
1.0	0.105	0.219	0.274	0.310	0.307	0.414	0.446	0.463
10 Years								
0.0	0.162	0.220	0.289	0.333	0.369	0.414	0.454	0.472
0.2	0.146	0.224	0.283	0.335	0.353	0.417	0.451	0.472
0.4	0.150	0.229	0.295	0.338	0.357	0.420	0.456	0.473
0.6	0.152	0.242	0.302	0.342	0.359	0.429	0.459	0.475
0.8	0.158	0.250	0.309	0.348	0.365	0.433	0.462	0.477
1.0	0.163	0.272	0.317	0.361	0.370	0.445	0.466	0.481
15 Years								
0.0	0.195	0.255	0.317	0.386	0.396	0.436	0.466	0.487
0.2	0.181	0.258	0.320	0.385	0.385	0.438	0.467	0.487
0.4	0.181	0.259	0.327	0.390	0.385	0.438	0.469	0.488
0.6	0.183	0.260	0.337	0.392	0.387	0.439	0.473	0.488
0.8	0.186	0.267	0.343	0.397	0.389	0.443	0.475	0.490
1.0	0.186	0.277	0.348	0.394	0.389	0.448	0.477	0.489
20 Years								
0.0	0.237	0.299	0.365	0.430	0.425	0.458	0.482	0.495
0.2	0.227	0.295	0.363	0.432	0.419	0.456	0.481	0.496
0.4	0.219	0.296	0.357	0.434	0.414	0.457	0.479	0.496
0.6	0.222	0.285	0.361	0.424	0.416	0.452	0.481	0.494
0.8	0.213	0.271	0.353	0.403	0.410	0.445	0.478	0.491
1.0	0.208	0.268	0.326	0.395	0.406	0.443	0.469	0.489

E denotes the equity-to-total-asset-value ratio, and H defines the hedge ratio.
The left panel shows the mean of the simulation outcomes, the right panel the standard deviation.

Table 21: Insolvency Relative to VEV probability estimates during QE

	E =0.2	0.4	0.6	0.8	$\hat{\sigma}_{0.2}$	$\hat{\sigma}_{0.4}$	$\hat{\sigma}_{0.6}$	$\hat{\sigma}_{0.8}$
5 Years								
H =0.0	0.492	0.467	0.472	0.482	0.500	0.499	0.499	0.500
0.2	0.480	0.456	0.462	0.479	0.500	0.498	0.499	0.500
0.4	0.462	0.449	0.460	0.482	0.499	0.498	0.499	0.500
0.6	0.449	0.446	0.465	0.484	0.498	0.497	0.499	0.500
0.8	0.447	0.454	0.468	0.484	0.497	0.498	0.499	0.500
1.0	0.452	0.453	0.474	0.488	0.498	0.498	0.500	0.500
10 Years								
0.0	0.464	0.442	0.455	0.478	0.499	0.497	0.498	0.500
0.2	0.436	0.424	0.450	0.478	0.496	0.494	0.498	0.500
0.4	0.423	0.419	0.448	0.480	0.494	0.494	0.498	0.500
0.6	0.423	0.415	0.443	0.489	0.494	0.493	0.497	0.500
0.8	0.423	0.418	0.454	0.492	0.494	0.493	0.498	0.500
1.0	0.418	0.428	0.462	0.501	0.493	0.495	0.499	0.500
15 Years								
0.0	0.451	0.450	0.478	0.508	0.498	0.498	0.500	0.500
0.2	0.444	0.446	0.468	0.505	0.497	0.497	0.499	0.500
0.4	0.435	0.436	0.466	0.506	0.496	0.496	0.499	0.500
0.6	0.439	0.438	0.464	0.508	0.497	0.496	0.499	0.500
0.8	0.430	0.428	0.469	0.508	0.495	0.495	0.499	0.500
1.0	0.421	0.421	0.462	0.513	0.494	0.494	0.499	0.500
20 Years								
0.0	0.472	0.470	0.504	0.556	0.499	0.499	0.500	0.497
0.2	0.456	0.451	0.491	0.553	0.498	0.498	0.500	0.497
0.4	0.444	0.439	0.482	0.552	0.497	0.497	0.500	0.498
0.6	0.425	0.429	0.483	0.542	0.495	0.495	0.500	0.498
0.8	0.408	0.414	0.457	0.519	0.492	0.493	0.498	0.500
1.0	0.389	0.379	0.431	0.500	0.488	0.485	0.495	0.500

E denotes the equity-to-total-asset-value ratio, and H defines the hedge ratio.
The left panel shows the mean of the simulation outcomes, the right panel the standard deviation.

Table 22: Insolvency Relative to MVEV probability estimates in EE

	E =0.2	0.4	0.6	0.8	$\hat{\sigma}_{0.2}$	$\hat{\sigma}_{0.4}$	$\hat{\sigma}_{0.6}$	$\hat{\sigma}_{0.8}$
5 Years								
H =0.0	0.002	0.023	0.054	0.084	0.045	0.150	0.226	0.278
0.2	0.002	0.020	0.053	0.084	0.045	0.140	0.224	0.278
0.4	0.002	0.023	0.059	0.087	0.045	0.150	0.236	0.282
0.6	0.003	0.027	0.061	0.089	0.055	0.162	0.239	0.285
0.8	0.003	0.027	0.063	0.089	0.055	0.162	0.243	0.285
1.0	0.002	0.029	0.068	0.089	0.045	0.168	0.252	0.285
0.0	0.003	0.011	0.031	0.053	0.055	0.104	0.173	0.224
10 Years								
0.2	0.002	0.011	0.028	0.053	0.045	0.104	0.165	0.224
0.4	0.002	0.011	0.027	0.056	0.045	0.104	0.162	0.230
0.6	0.002	0.010	0.028	0.057	0.045	0.100	0.165	0.232
0.8	0.002	0.010	0.030	0.055	0.045	0.100	0.171	0.228
1.0	0.002	0.009	0.031	0.055	0.045	0.094	0.173	0.228
15 Years								
0.0	0.000	0.010	0.021	0.041	0.000	0.100	0.143	0.198
0.2	0.000	0.005	0.019	0.040	0.000	0.071	0.137	0.196
0.4	0.000	0.004	0.018	0.038	0.000	0.063	0.133	0.191
0.6	0.000	0.001	0.018	0.038	0.000	0.032	0.133	0.191
0.8	0.000	0.001	0.017	0.038	0.000	0.032	0.129	0.191
1.0	0.000	0.003	0.015	0.037	0.000	0.055	0.122	0.189
20 Years								
0.0	0.000	0.005	0.016	0.028	0.000	0.071	0.126	0.165
0.2	0.000	0.003	0.010	0.026	0.000	0.055	0.100	0.159
0.4	0.000	0.002	0.009	0.026	0.000	0.045	0.094	0.159
0.6	0.000	0.003	0.010	0.024	0.000	0.055	0.100	0.153
0.8	0.000	0.002	0.008	0.023	0.000	0.045	0.089	0.150
1.0	0.000	0.001	0.005	0.018	0.000	0.032	0.071	0.133

E denotes the equity-to-total-asset-value ratio, and H defines the hedge ratio.
The left panel shows the mean of the simulation outcomes, the right panel the standard deviation.

Table 23: Insolvency Relative to VEV probability estimates in EE

	E =0.2	0.4	0.6	0.8	$\hat{\sigma}_{0.2}$	$\hat{\sigma}_{0.4}$	$\hat{\sigma}_{0.6}$	$\hat{\sigma}_{0.8}$
5 Years								
H =0.0	0.092	0.142	0.169	0.200	0.289	0.349	0.375	0.400
0.2	0.061	0.129	0.159	0.195	0.239	0.335	0.366	0.396
0.4	0.058	0.134	0.162	0.192	0.234	0.341	0.369	0.394
0.6	0.056	0.130	0.161	0.187	0.230	0.336	0.368	0.390
0.8	0.047	0.125	0.157	0.186	0.212	0.331	0.364	0.389
1.0	0.043	0.119	0.154	0.185	0.203	0.324	0.361	0.388
10 Years								
0.0	0.033	0.067	0.105	0.141	0.179	0.250	0.307	0.348
0.2	0.014	0.055	0.099	0.133	0.118	0.228	0.299	0.340
0.4	0.012	0.048	0.099	0.132	0.109	0.214	0.299	0.339
0.6	0.010	0.045	0.096	0.131	0.100	0.207	0.295	0.338
0.8	0.008	0.044	0.087	0.131	0.089	0.205	0.282	0.338
1.0	0.006	0.041	0.083	0.127	0.077	0.198	0.276	0.333
15 Years								
0.0	0.014	0.036	0.069	0.102	0.118	0.186	0.254	0.303
0.2	0.008	0.025	0.059	0.096	0.089	0.156	0.236	0.295
0.4	0.004	0.022	0.050	0.092	0.063	0.147	0.218	0.289
0.6	0.003	0.015	0.047	0.089	0.055	0.122	0.212	0.285
0.8	0.002	0.016	0.041	0.086	0.045	0.126	0.198	0.281
1.0	0.000	0.014	0.039	0.085	0.000	0.118	0.194	0.279
20 Years								
0.0	0.008	0.018	0.040	0.062	0.089	0.133	0.196	0.241
0.2	0.003	0.011	0.028	0.053	0.055	0.104	0.165	0.224
0.4	0.001	0.008	0.022	0.052	0.032	0.089	0.147	0.222
0.6	0.000	0.008	0.018	0.047	0.000	0.089	0.133	0.212
0.8	0.000	0.006	0.016	0.045	0.000	0.077	0.126	0.207
1.0	0.000	0.004	0.014	0.041	0.000	0.063	0.118	0.198

E denotes the equity-to-total-asset-value ratio, and H defines the hedge ratio.
The left panel shows the mean of the simulation outcomes, the right panel the standard deviation.

9.5.5 Indexation

Table 24: Indexation levels during QE, percentages

	E =0.2	0.4	0.6	0.8	$\hat{\sigma}_{0.2}$	$\hat{\sigma}_{0.4}$	$\hat{\sigma}_{0.6}$	$\hat{\sigma}_{0.8}$
5 Years								
H =0.0	1.319	1.292	1.246	1.198	0.834	0.878	0.902	0.921
0.2	1.329	1.295	1.244	1.193	0.842	0.885	0.907	0.926
0.4	1.344	1.291	1.235	1.181	0.851	0.891	0.914	0.932
0.6	1.323	1.270	1.215	1.167	0.871	0.902	0.925	0.937
0.8	1.291	1.242	1.194	1.151	0.893	0.918	0.934	0.943
1.0	1.236	1.208	1.172	1.137	0.922	0.935	0.943	0.950
10 Years								
0.0	1.271	1.261	1.216	1.153	0.891	0.916	0.932	0.947
0.2	1.302	1.276	1.219	1.153	0.885	0.916	0.935	0.950
0.4	1.305	1.276	1.213	1.141	0.892	0.921	0.938	0.952
0.6	1.295	1.266	1.202	1.127	0.906	0.928	0.943	0.956
0.8	1.285	1.252	1.184	1.113	0.915	0.936	0.947	0.960
1.0	1.266	1.231	1.170	1.100	0.928	0.942	0.950	0.964
15 Years								
0.0	1.272	1.237	1.157	1.077	0.899	0.917	0.949	0.968
0.2	1.280	1.242	1.155	1.073	0.897	0.920	0.955	0.968
0.4	1.276	1.237	1.152	1.069	0.903	0.929	0.957	0.969
0.6	1.267	1.234	1.146	1.064	0.911	0.936	0.961	0.970
0.8	1.261	1.232	1.142	1.057	0.921	0.943	0.962	0.973
1.0	1.265	1.232	1.152	1.054	0.929	0.949	0.961	0.973
20 Years								
0.0	1.213	1.187	1.088	0.988	0.918	0.943	0.964	0.968
0.2	1.239	1.200	1.096	0.989	0.916	0.943	0.966	0.968
0.4	1.245	1.215	1.107	0.997	0.922	0.943	0.967	0.972
0.6	1.264	1.229	1.125	1.015	0.922	0.943	0.963	0.969
0.8	1.276	1.263	1.164	1.051	0.929	0.937	0.964	0.970
1.0	1.291	1.307	1.207	1.094	0.932	0.931	0.958	0.970

E denotes the equity-to-total-asset-value ratio, and H defines the hedge ratio.
The left panel shows the mean of the simulation outcomes, the right panel the standard deviation.

Table 25: Indexation levels in EE

	E =0.2	0.4	0.6	0.8	$\hat{\sigma}_{0.2}$	$\hat{\sigma}_{0.4}$	$\hat{\sigma}_{0.6}$	$\hat{\sigma}_{0.8}$
5 Years								
H =0.0	1.972	1.928	1.752	1.707	0.134	0.277	0.616	0.662
0.2	1.987	1.938	1.757	1.708	0.088	0.255	0.614	0.664
0.4	1.988	1.946	1.752	1.706	0.087	0.239	0.623	0.668
0.6	1.990	1.952	1.750	1.705	0.078	0.223	0.628	0.671
0.8	1.992	1.954	1.750	1.705	0.071	0.220	0.632	0.673
1.0	1.992	1.963	1.750	1.705	0.076	0.193	0.634	0.675
10 Years								
0.0	1.996	1.983	1.959	1.927	0.037	0.108	0.197	0.299
0.2	2.000	1.989	1.967	1.931	0.008	0.086	0.167	0.289
0.4	1.999	1.992	1.971	1.935	0.014	0.076	0.152	0.280
0.6	2.000	1.989	1.975	1.937	0.006	0.088	0.141	0.273
0.8	2.000	1.990	1.981	1.941	0.011	0.083	0.128	0.262
1.0	2.000	1.991	1.980	1.940	0.002	0.077	0.135	0.271
15 Years								
0.0	2.000	1.996	1.982	1.963	0.007	0.046	0.111	0.205
0.2	2.000	1.998	1.987	1.968	0.002	0.035	0.092	0.185
0.4	2.000	1.999	1.990	1.970	0.000	0.018	0.085	0.178
0.6	2.000	2.000	1.990	1.972	0.000	0.000	0.084	0.173
0.8	2.000	2.000	1.994	1.970	0.000	0.006	0.065	0.181
1.0	2.000	1.999	1.995	1.974	0.000	0.016	0.052	0.166
20 Years								
0.0	1.999	1.999	1.994	1.986	0.012	0.026	0.061	0.117
0.2	2.000	2.000	1.997	1.990	0.000	0.007	0.039	0.095
0.4	2.000	2.000	1.999	1.995	0.000	0.000	0.024	0.055
0.6	2.000	2.000	1.999	1.993	0.000	0.003	0.016	0.082
0.8	2.000	1.999	2.000	1.992	0.000	0.014	0.004	0.086
1.0	2.000	2.000	2.000	1.995	0.000	0.000	0.000	0.060

E denotes the equity-to-total-asset-value ratio, and H defines the hedge ratio.
The left panel shows the mean of the simulation outcomes, the right panel the standard deviation.

9.6 Figures

9.6.1 Funding Ratio

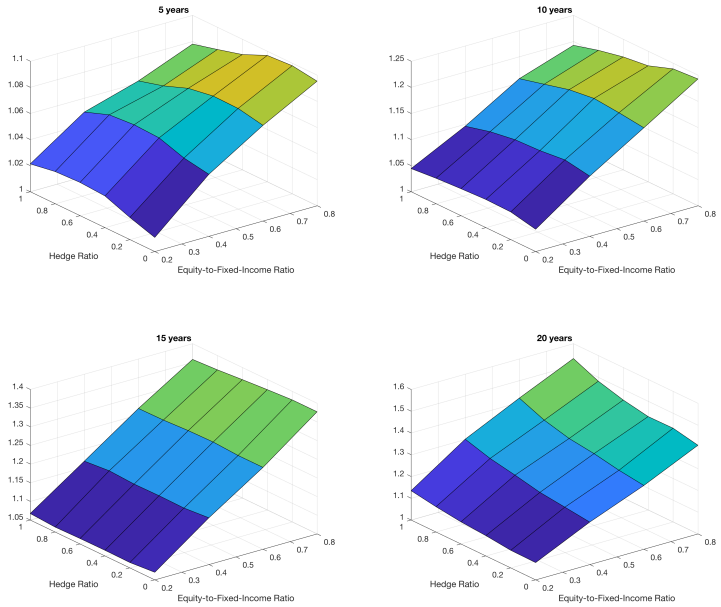


Figure 36: Funding Ratio During QE

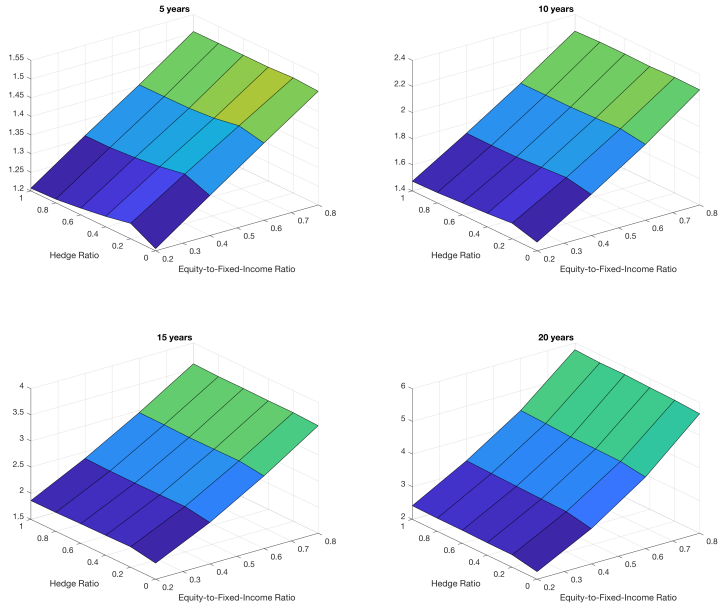


Figure 37: Funding Ratio in EE