Erasmus University Rotterdam

MSc Quantitative Finance
Thesis

Feasibility Test

Does it serve its purpose?

Author:
Sebastian van den Berg

Student Number:
324429

Supervisors:
M. Jaskowski
B. Diris
(Erasmus University)

M. Blankevoort
F. Hanlo
D. Siewe
(Kempen & Co)

ERASMUS UNIVERSITEIT ROTTERDAM
ERASMUS SCHOOL OF ECONOMICS

KEMPEN & CO
Merchant Bank

April 20, 2016

Preface

This thesis is the final proof of competence for obtaining the Master of Science (MSc) degree in Econometrics and Management Science, with a specialization in Quantitative Finance, from the Erasmus University Rotterdam. The research was conducted in cooperation with Kempen Capital Management (KCM), a subsidiary of Kempen & Co. Kempen & Co is a Dutch merchant bank providing financial services in Asset Management, Securities Broking and Corporate Finance. KCM is Kempen's Asset Management branch and manages portfolios for numerous large institutional investors, financial institutions, (semi)public institutions, foundations and high net-worth individual clients. This thesis was specifically written in collaboration with KCM's Fiduciary Management Team. The Fiduciary Management team acts as the alternative investment arm for pension funds and insurance companies from around the world. The outsourcing of tasks by the pension fund board enables them to focus on policy-related issues such as payments, investments, premiums, benefits and indexation policies.

I would like to take this opportunity to thank the following people for their help and support during the writing of my thesis. First of all, I would like to thank my supervisors Maarten Blankevoort, Florentine Hanlo, and Duncan Siewe at Kempen Capital Management for their constructive comments and guidance during this process. In addition, I want to thank Dr. Maya Beyhan (Fiduciary Manager at KCM) for her assistance on the quantitative part of my thesis. I also would like to thank my other colleagues at Kempen & Co for their wonderful cooperation during my graduate internship. Further, I want to thank Sander Muns (researcher at CPB) for providing the datasets and for helping me with the estimation and interpretation of the KNW model. Finally, I would like to thank my main supervisor Marcin Jaskowski and co-reader Bart Diris of the Erasmus University for a careful review of my work.

Amsterdam, April 20, 2016.

Sebastian van den Berg

Abstract

In this paper, we review the new feasibility test for pension funds, which is part of the updated Dutch regulatory framework known as FTK. The feasibility test involves a stochastic analysis of 60 years, which determines the fund's risk-profile. Furthermore, it provides insight into the effects that the pension policy has on the purchasing power conservation of the accrued pension benefits among different generations. This paper's contribution is threefold. First, we discuss the feasibility test in detail and highlight the results for three stylized pension funds. This study also presents a valuable tool for the implementation of the feasibility test, which is based on the generational accounting approach, as described by Chen et al. (2014). The paper's second part evaluates the set of scenarios, which is based on a model prescribed by Koijen et al (2010) (KNW model). In particular, the model parameters are re-estimated with updated data through the use of the simulated annealing procedure, in accordance with the work of Draper (2014). This paper discusses the results in detail and examines various stylized facts in order to assess the model's fit. In the final part, three alternative models are provided (based on three different types of interest rate models), including: (1) the CIR model (one-factor equilibrium model), (2) the G2++ model (a two-factor no-arbitrage model), and (3) the Libor Market Model. The models are estimated through the use of different estimation techniques. They are subsequently examined through an assessment of important features with respect to interest rate models, such as upward sloping average yields, a decreasing volatility of yields as well as a great variety of shapes over time.

Keywords: Economic Scenario Generator, Interest rate models, Asset-Liability Management, Financial Assessment Framework, Pension result

Contents

1 Introduction

The past decades have shown that Dutch pension funds are not immune to declining economic conditions. During the beginning of the 21st century, adverse equity markets and low interest rates caused funding ratios to plummet. The funding ratio is an important measure for pension funds, as it indicates whether a particular fund is able to pay for its participants' future pension benefits. In addition, during the 2008 financial crisis, the solvency position of most Dutch pension funds rapidly deteriorated. As a result of this economic turmoil, pension funds' stakeholders and regulators would like to acquire a better understanding of the risks associated with their investments and liabilities.

Economic conditions are not the only variable that affects pension funds. Demographic changes continue to reduce the ratio between active participants and retirees. Because of recent economic developments and an aging population, the regulation of pension funds has become a critical topic. Several regulatory reforms have been introduced, with the intention of ensuring pension funds' solvency and rendering pensions more robust to volatility in financial markets. One of the latest of a series of regulatory changes is the amendment of the Financial Assessment Framework (FTK), which came into force on January 1, 2015. The new FTK's objective is to reduce pensions' vulnerability to major shocks in financial markets. In addition, it is charged with contributing to a more balanced distribution of benefits and burdens across participants, especially between younger and older generations.

A new FTK element is the feasibility test, which replaced the current continuity analysis and the consistency test, as of the start of 2016. This feasibility test will be used to define a pension fund's long-term risk profile. This new regulatory instrument examines whether a fund's investment strategy and predetermined pension policy regarding premiums, indexations, and pension discounts are sufficiently realistic and feasible over the next 60 years. Furthermore, the feasibility test must provide insight into the effects of the pension policy on the purchasing power conservation of the accrued pension benefits. The metric 'pension result' measures pension's purchasing power conservation at the fund level as well as for different generations. The main objective of the feasibility test is to force pension funds to consider their ambitions and risk attitudes in advance and to actively communicate with their participants regarding these subjects. This transparency must improve the participant's understanding with respect to their pension situation.

The feasibility test entails a scenario analysis for the long-term financial position of a pension fund based on a uniform economic scenario set with a 60-year horizon. The Dutch Central Bank (DNB) publishes the uniform scenario set on a quarterly basis. The scenario set contains 2,000 scenarios for equity returns, price inflation, and interest rates. Based on these scenarios, pension results can be simulated over a period of 60 years. Stakeholders and regulators are especially interested in outcomes regarding the expected pension results and the pension results in a 'bad weather scenario'. The difference between these two scenarios may be used to describe a pension fund's risk profile. These results, therefore, play a major role in the communication directed towards individual fund participants.

The Dutch government asked the Commission Parameters (Langejan et al. (2014)) to advise a stochastic scenario set that will be used in the feasibility test. In the beginning of 2014, the commission recommended an economic scenario generator described by Koijen et al. (2010), also known as the 'KNW model'. The KNW model is relatively simple in comparison to more sophisticated ALM models, which ease the pension funds' implementation of the feasibility test. Since the model is comparatively basic, the Commission Parameters emphasizes that the KNW model is not a substitute for more sophisticated ALM models. This aspect immediately raises the question as to whether the KNW model produces scenarios that are realistic enough to yield useful results. Furthermore, in the pension sector, many people believe the feasibility test's design (and in particular, the scenario set) requires some improvements.[1] The KNW model contains two main drawbacks. First, it is comprised of only two asset classes (equities and bonds). Secondly, the

[1]Source: www.pensioenbestuurenmanagement.nl (Article 'Design feasibility test does not serve its purpose')

model may generate negative (long-term) interest rates. A further shortcoming is that the term structure scenarios must be used for liabilities as well as in fixed income portfolios. Finally, the implied initial yield curve does not match the prevailing DNB yield curve.

This research paper reviews the feasibility test and assesses whether this new regulatory instrument meets its objective. In particular, we examine whether the feasibility test can be used as an adequate risk-management tool and whether it provides sufficient insight into the effect that pension policy has on participants' pensions. Finally, we examine whether the KNW model is appropriate for the feasibility test and subsequently explore alternative models.

The contribution of this paper is threefold. First, we provide a detailed description of the feasibility test and examine the implementation of the scenario analysis. For the implementation, we present a tool that is specifically designed for the feasibility test. The tool is capable of modeling a fund's financial position over time. In addition, it records the cash flows of different generations according to the generational accounting approach described by Chen et al. (2014). The latter is essential to measure the purchasing power conservation of individual participants. We perform the feasibility test for a stylized Dutch pension fund. This fund is meant to be representative of an average fund in the Netherlands, as it takes into account the specific demographic characteristics of the Dutch population. Because the feasibility test results strongly depend on the fund's participant composition, we will also consider a relatively young (green) and a relatively old pension fund (grey).

The second part of this paper further focuses on the scenario set based on the KNW model. The KNW model was originally developed by Koijen et al. (2010) to describe the U.S capital market. The stock and bond market dynamics are based on a two-factor model, which accommodates time-varying interest rates, inflation rates, and bond risk premia. The CPB estimated the model parameters of the KNW model by utilizing relevant data for Dutch pension funds, first in 2012 (Draper (2012)) and later in 2014 with updated data (Draper (2014)). Draper (2012) noted that their results deviate in several aspects from those in Koijen et al. (2010) that were based on U.S. data. In particular, the coefficient estimates are less significant for Europe. This parameter uncertainty has an important implication for the evaluation of a pension fund's (future) financial position. In addition, the CPB indicated that the estimation in Draper (2014) did not lead to the maximum likelihood.

It is important for stakeholders to understand the assumptions that underlie the process of economic scenario generation. Only when managers and regulators possess sufficient insight regarding data and assumptions can they truly judge the results of policy choices. Hence, we first provide the methodology and estimation procedure of the KNW model. Subsequently, we re-estimate the two-factor model with an updated dataset. For the estimation procedure based on the method of simulated annealing, we convert the KNW model into a discrete version of the multivariate Ornstein-Uhlenbeck process. We discuss the results in detail and compare them to the results presented in Draper (2012) and Draper (2014) to evaluate the robustness of the estimates. To assess the goodness of fit, we examine the average yield curve, the volatility of the bonds, and the predictability of bond returns using the Campbell-Shiller regression. Furthermore, we explain the calibration and actualization of several model parameters to make scenarios consistent with predetermined expectations and current market conditions.

And, in the final part of this paper, we consider three alternative models to generate scenarios for the feasibility test. We compare the results from the KNW model with results from the alternative models to assess whether it is acceptable to use the KNW model to generate the set of scenarios. The alternative scenario sets focus mainly on modeling the term structure. Vlaar (2006) emphasizes the importance of correctly modeling interest-rate dynamics for pension funds since the introduction of market valuation for pension funds' liabilities. As pension funds' obligations stretch far into the future, the model should be a sufficient fit, both on the short end of the yield curve as well as the long end. In addition, the value of liabilities increases significantly if interest rates approach zero, so the probability of very low rates should be modeled correctly, according to Vlaar (2006).

We consider three different types of interest rate models. First alternative model is the CIR model, which is calibrated via maximum likelihood estimation on historical short rates. Second alternative model is the G2++ model, which is calibrated via Kalman filtering to capture the times-series dimension as well as the space dimension. Third alternative model is the Libor Market Model, which is calibrated via least squares estimation to historical forward volatilities and correlations. To assess the alternative interest rate models, we examine whether simulation results correspond to the stylized facts of the observed interest rate data and whether the model can produce a variety of shapes through time. And finally, we discuss the feasibility test results and measure whether outcomes are significantly different when an alternative model is utilized.

This paper shows that the feasibility test is able to demonstrate the uncertainty regarding the conservation of pension benefits' purchasing power for a long-term horizon. However, the study also indicates several shortcomings regarding the feasibility test's implementation and communication. Furthermore, the comprehensive report stresses the need for discussion on the scenario set's assumptions.

The paper also contributes to the existing academic literature in several ways. First, we provide proof that the estimation procedures in Draper (2012) and Draper (2014) did not lead to maximum likelihoods. Furthermore, we show that the KNW model provides a sufficient fit to the Dutch data and is able to match the cross-sectional moments of bond yields. In addition, this paper aims to contribute to the limited public research on interest rate models for risk-management purposes. We provide an overview on interest rate models and examine whether three different types of models are suitable for long-term stochastic analysis. And finally, we introduce two statistical tests to support the qualitative assessments among various feasibility test results.

The remainder of this paper is structured as follows. Chapter 2 provides background information on pension funds, including the new regulatory framework FTK. In addition, this chapter briefly overviews ALM models, economic scenario generators, and interest rate models. Chapter 3 discusses the feasibility test in detail and presents the results for an average Dutch pension fund, a green pension fund, and a grey pension fund. Chapter 4 provides a highly detailed description of the KNW model, including its methodology, estimation procedure, and results. Chapter 5 discusses the three alternative models. And finally, Chapter 6 concludes the paper.

2 Background Information

The Dutch pension sector has been a topic of great interest for several years, with the public, politicians, and (academic) researchers all growing increasingly concerned. This new consciousness is an important development, because it concerns our own pensions. In addition, Dutch pension funds play a major role in the Dutch economy. At the end of 2014, they managed more than €1.25 trillion in pension capital, a figure equal to 189% of Dutch GDP in that year.[2] Before I began this paper, my knowledge regarding this topic was limited, as is probably the case with most people. The next section provides a brief overview of pension funds intended to lay the foundation for understanding the topics covered in this paper. Sections 2.1 and 2.2 provide information on Dutch pension funds in general and on the (current) legislation. In Sections 2.3 and 2.4, we look more closely at ALM models and at economic scenario generators. In Section 2.5, we provide a brief review of former research regarding interest rate models.

2.1 Dutch Pension Funds

Pension funds ensure that employees save for their retirement. The pension benefits received at retirement age are based on contributions paid in the past and on the return on the investments of these contributions. Pension funds are responsible for investing contributions so that maximum returns are generated with minimum risk, so as to secure a stable retirement income for their members.

In the Netherlands, there are several types of pension schemes.[3] Nowadays, the majority of Dutch pension funds are based on the *average-salary Defined Benefit scheme*. In this scheme, pension funds guarantee a certain level of pension benefits after retirement, while contributions (premiums) can be adjusted to achieve the 'defined benefits'. Accrued pension benefits are based on the average salary earned during an employee's career. The main disadvantage of this scheme is that accrued benefits are not automatically adjusted for inflation, because they are based on past wages. However, in addition to the guaranteed nominal pension, most Dutch pension funds also strive to provide indexation, typically on a conditional basis. Indexation means that accrued pension benefits are adjusted for price inflation or for an increase in wages throughout a particular sector. Without regular interim indexation, the accrued nominal pension benefits of a 40-year-old participant will lose approximately half of their value by time of retirement. Actual indexation depends on a pension fund's financial position.

A fund's financial position can be expressed by the funding ratio, which is the value of the fund's financial assets over the present value of its pension liabilities, or:

$$FR = \frac{Present\ Value\ Assets}{Present\ Value\ Liabilities} \tag{1}$$

The funding ratio is an essential indicator for pension funds, as it signifies whether a particular fund can pay its members' future pensions. A funding ratio above 100% indicates that a fund has sufficient capital to meet all future obligations. The liabilities represent the accrued pension benefits of all participants that must be paid from the age of retirement. The present value of future payments is calculated as follows:

$$PV\ Liabilities = \sum_{t=0}^{n} \left(\frac{Liability\ Outflows_t}{(1+i)^t}\right) \tag{2}$$

where 'liability outflows' are estimated future payments based on actuarial factors, including accrued pension rights and life expectancy. In this equation, $1/(1+i)^t$ represents the discount factor, and 'n' is the horizon of the cash flows. The assets represent the value of the fund's

[2]Source: www.dnb.nl

[3]See Appendix A. for more information on the Dutch pension system and on different types of pension agreements.

investments, while the value depends on the investment return of participants' contributions and on pension payments to participants. To determine the optimal asset mix, pension funds make use of Asset-Liability Management (ALM) models to study possible developments regarding assets and liabilities in terms of their duration, returns, and inflation.

2.2 Financial Assessment Framework

As of January 1, 2007, a new pension fund legislation called the Financial Assessment Framework (FTK) has been in effect. The FTK was created to evaluate the pension funds' financial positions, along with the policies established by their boards. The FTK's most significant adjustment as compared to former legislation is its market-consistent valuation of liabilities, with the prevailing swap rate serving as the discount rate. Liabilities were initially valued using a market-inconsistent fixed rate of 4%, which led to underestimation of pension liabilities and to systematic camouflaged underfunding. Furthermore, the FTK prescribes strict solvency requirements. A pension fund is considered underfunded when its funding ratio drops below the minimum regulatory capital level (mVEV) of 105%. In addition, the FTK contains a risk-weighted capital requirement (VEV) that serves as a financial buffer in order to withstand financial shocks. The VEV level is fixed in such a manner so that a pension fund will remain fully funded ($FR \geq 100\%$) within a one-year timeframe at a 97.5% confidence level. If the funding ratio exceeds capital requirements, a pension fund can decide whether to provide indexation.

The introduction of the FTK in 2007 was a direct response to the economic and regulatory environment of the early 2000s. However, the 2008 financial crisis made clear that Dutch pension funds are still very vulnerable to economic developments and to the volatility of financial markets. In addition, the impact of demographic changes, along with the ongoing low interest-rate environment, has demonstrated that the pension industry needs to be reviewed once more. Since its inception in 2007, the FTK has been altered a number of times. One such modification was the 2012 introduction of the new Ultimate Forward Rate (UFR) for discounting liabilities.

The most recent legislation is the a*mendment Financial Assessment Framework* (nFTK), which was approved by parliament in December 2014 and put into effect on January 1, 2015. The objective of this law is to make pensions less vulnerable to major shocks in financial markets and to contribute to a more balanced distribution of benefits and burdens between involved participants (young vs. old). The amendments to the FTK also contain a number of new elements intended to make pension contracts more comprehensive and complete. For example, a pension fund must clearly define its investment policy in advance, including the degree of risk that it deems acceptable. The new feasibility test plays a significant role regarding the communication of the pension funds' risk-profile. Appendix B. provides a more detailed description of the FTK, including an overview of its most important amendments.

2.3 ALM models

Pension funds must strike the optimal balance between pursuing high returns and lowering the probability of not being able to pay promised pension benefits in the future. To determine the optimal asset allocation, pension funds perform ALM studies in which they evaluate their expected future financial position. An ALM study provides quantitative insight into the interaction between assets and liabilities over a certain evaluation period. ALM models evaluate a fund's current policy framework and explore alternative options and thus can be used to determine a fund's optimal strategic policy. To be more precise, ALM models identify the risks associated with different policy alternatives and illustrate the effect of various policy instruments, of which the most prominent are the contribution rate, indexation, and asset allocation. The integrated policy resulting from an ALM study is referred to as the pension deal. This pension deal must satisfy the objectives and risk appetite of all stakeholders (e.g. active members, pensioners, and regulators).

The future status of a pension fund's assets and liabilities is not solely dependent on policy decisions. Rather, exogenous actuarial factors and economic factors are also involved (see Appendix

D.). ALM studies use scenario analysis to forecast relevant factors, such as those related to participants' characteristics (e.g. life expectancy), as well as macro-economic variables. To predict possible future developments in the financial market, an economic scenario generator (ESG) model is developed. ESG models use stochastic simulations (i.e. Monte Carlo simulations) to project all relevant and uncertain macro-economic variables, like inflation, interest rates, and equity returns. These models produce many scenarios (e.g. 10,000) for a fixed time horizon, and each such scenario represents a possible future set of global conditions. This approach is very useful for modeling various uncertainties. However, in order to create plausible future scenarios, it is essential to use an appropriate model.

Based on the many economic scenarios generated by the model, all possible future developments of those variables relevant to pension funds (e.g. funding ratio) can be computed. As a result of the scenario approach, the projections of these variables can be seen as probability distributions. These probability distributions allow us to identify expectations and risks, calculate the probability of underfunding, or gauge the probability of benefit cuts in the future. A sensitivity analysis is often performed to explore assorted policy variants in terms of indexation ambition, premium policy, and the asset allocation mix. Such an analysis illustrates how different policy decisions affect a particular pension fund's financial performance. In addition to sensitivity analysis, funds can also apply optimization techniques to identify their optimal investment policy. Hence, with the help of ALM studies, pension funds can make well-informed decisions regarding their policy instruments.

2.4 Economic Scenario Generators

As mentioned above, scenario analysis has proven to be an essential method to monitor pension funds' financial risks. By generating economic scenarios, all possible future states of a pension fund can be simulated. When undertaking such an analysis, it is crucial that the scenarios for the different variables are plausible over time and that they fit the requirements of the ALM model. Hibbert et al. (2001) provide a detailed background of ESGs intended for risk-management purposes and highlight certain key changes that have taken place over the past several decades. The extraordinary innovations in computer technology represent a crucial development in risk management. Financial institutions such as pension funds have made substantial investments in this new technology to enhance their risk-management capabilities.

In addition, academic researchers and financial practitioners have published many papers on scenario modeling in recent years. However, these studies have been motivated by a variety of needs. Many papers describe improved techniques for pricing, trading, and hedging a range of financial instruments, including derivatives. The literature on models for interest rates with varying degrees of complexity is especially extensive. Also, economists have developed models for forecasting purposes, and much research has been done in the field of portfolio optimization. Furthermore, long-term financial planners (including pension funds) have examined models for multiple sources of uncertainty over long horizons in comparison to other short-term financial models.

Most academic research deals only with parts of the problem in which we are interested. For instance, there is a great deal of detailed research on equity price behavior, term-structure modeling and inflation modeling. However, there are very few ('open source') papers that put all of the components together within a consistent framework. Hoevenaars et al. (2003) provide a solid overview of several models that can be used to generate future scenarios. The main models that are described in the literature and used in practice include: the vector autoregressive (VAR) model, the cascade approach, and the stochastic differential approach.

The *VAR model* captures the linear interdependencies between multiple assets over time by modeling each variable as a linear function of past lags of itself and past lags of the other variables. The advantage of a VAR model is that it describes stochastic time series and also captures the conditional long-term dynamics, including the variances, autocorrelations, and cross-correlations between different variables. A major drawback of this model is that the accuracy of the estimated parameters rapidly decreases as the number of variables increases.

A second model for generating economic scenarios is the *cascade approach*, which was introduced by Wilkie (1987). Wilkie's model uses inflation as the 'driving' variable, with several other economic processes being driven off of inflation in a 'cascade'-type manner. In this model, a first-order autoregressive (AR) model is used to model inflation, as well as the other variables, thus limiting the parameters that must be estimated. However, in many papers on term-structure modeling, interest rates are modeled directly, and so using inflation for the (only) state variable is arbitrary.

In this paper, we focus on the *stochastic differential equations* (SDE) approach for generating economic scenarios. The SDE approach is a similar to the cascade approach. The main difference between the two is that SDEs are formulated in continuous time instead of in discrete time. SDE models have a variety of applications in many different fields, including modeling the underlying uncertainty of economic and financial variables. A SDE can be defined as an ordinary differential equation driven by one or more stochastic processes. The term 'stochastic' means 'random', and the random process in SDEs is often referred to as a Wiener process. The main drawbacks of SDE models are the theory uncertainty and the fact that their estimation methods are far from straightforward.

Hence, all of these models have advantages and disadvantages, as described by Hoevenaars et al. (2003). According to Hibbert et al. (2001) the scenarios must meet various requirements in order to be a 'good model'. An important obligation is that the model must provide an accurate representation of the financial assets that it contains. In other words, the model should 'mimic' the real-world behavior of financial assets, capturing their most essential characteristics. A further condition that scenarios must satisfy relates to the joint behavior of model variables. And, the behavior of assets should be consistent with generally accepted economic principles. Hibbert et al. (2001) emphasize that there are certain key properties of financial asset behavior on which economists have not arrived at a clear consensus. Furthermore, it is imperative to keep the model as simple as possible while retaining the most important features. Hibbert et al. (2001) concludes that no models exist that simultaneously meet all of these criteria. Therefore, users must understand the set of assumptions that underlie a particular model before analyzing its output. In other words, a pension fund's board and regulators can only truly judge the results of their policy choices if they possess sufficient insight into the ESG model and the assumptions used to generate the scenarios.

2.5 Interest Rate Models

One of the most challenging research topics in finance is identifying the factors that determine the behavior of interest rates (cross-section and time series). Previous research related to this subject is quite extensive. Many interest rate models have been developed, ranging from relatively simple (e.g., one-factor short-rate models) to extremely complex term-structure models (e.g. multi-factor short-rate models and market models). Most models are developed for the valuation of interest rate derivatives, such as caps and swaptions. Models developed for risk-management purposes have received comparatively little attention in the literature. However, many models that were originally developed for pricing derivatives are also applicable for long-term planning and actuarial work. In this section we provide a brief overview of the literature on interest rate models.

Interest rate models tend to fall into one of two categories: equilibrium models and no-arbitrage models. General equilibrium models derive bond yields from expectations regarding the economy. They typically start with the process for short-term rates, which is based on state variables that describe the overall economy. Subsequently, the entire term structure can be determined by looking at the expected path of interest rates until the bond's maturity. The two most famous models based the equilibrium approach are the models of Vasicek (1977) and Cox et al. (1985) (CIR). These models assume that the dynamics of the whole yield curve are driven by the instantaneous short rate, while the evolution of the short rate is described by a stochastic differential equation. Both models are considered to be one-factor short-rate models, because the term structure only depends on the short rate and not on any other state variables.

Short-rate models can be placed in the affine class and benefit from the analytical properties typical of this class. A large advantage of equilibrium models is that the bond prices often have closed-form analytic solutions, which makes these models fairly easy to use. A disadvantage is that equilibrium models cannot accurately replicate the existing term structure. Rather, they only have a finite number of free parameters, and it is not possible to specify the parameters so that the model exactly matches observed market yields. Equilibrium models focus more on the time-series properties of the term structure than on initial term structure's cross-sectional properties. Because of this, equilibrium models are not suitable for the valuation of derivatives.

No-arbitrage models were developed to avoid this drawback. These models take the observed term structure as given and subsequently generate the future dynamics of the yields. To precisely fit the initial term structure, one or more parameters are assumed to be deterministic functions of time (under the Q-measure). Popular no-arbitrage short-rate models are the models of Ho and Lee (1986), Black et al. (1990), Hull and White (1990) and Black and Karasinski (1991). A disadvantage of the no-arbitrage approach is that the incorporation of non-linearities into a model leads to more complexity. In particular, it often precludes a closed-form solution for bond prices. Another problem with no-arbitrage models is that they struggle to adequately describe interest rate dynamics over long periods of time.

The above-mentioned one-factor short-rate models insufficiently replicate observed term structures or are inadequate for explaining future movements. Another shortcoming of one-factor models is that bond yields are perfectly correlated. A solution to make the model more consistent with the data is to use multiple factors to determine the short rate. Multiple-factor models are more flexible than one-factor models and can provide a wider variety of term-structure movements. Extensions of one-factor affine models to the multi-factor case are relatively straightforward. Mathematically and conceptually speaking, there is barely any difference between one-factor and multi-factor models. However, estimation of a multi-factor short-rate model is significantly more complex.

Another key class of interest rate models is based on the Heath-Jarrow-Morton (HJM) framework. Heath et al. (1992) developed a model that starts with the current forward-rate curve and subsequently models stochastic changes in forward rates. The primary distinction between short-rate models and the HJM model is that the latter model captures the full dynamics of the entire forward-rate curve whereas the former models only capture the dynamics of a single point on the curve (the short rate). A disadvantage associated with the HJM model is that is its based on instantaneous forward rates, which are not directly observable in the market. Also, this class of models is, in most cases, non-Markovian (unlike the short-rate models), which makes the models computationally intensive and difficult to calibrate.

The LIBOR market model (LMM) enhances the HJM model by modeling the evolution of simple forward rates, which are directly observable in the market. This model can be interpreted as a collection of forward-rate dynamics with different tenors and maturities, and the forward rates are log-normally distributed. Advantages of the LMM are that the model exactly matches the initial yield curve and that it ensures positive interest rates. However, the model is originally designed for pricing purposes in a risk-neutral context and not for risk-management purposes.

The three-factor Nelson-Siegel model (Nelson and Siegel (1987)) and its variation, the four-factor Nelson-Siegel-Svensson model (Svensson (1994)) are based on a totally different mathematical approach. Although the models do not satisfy the no-arbitrage property, they are widely used to replicate the term structure.

There are many variations and extensions on the aforementioned models that attempt to incorporate all characteristics of interest rates. For example, the KNW model includes a two-factor short-rate model with time-varying bond risk premia. The overall consensus is that there is no single model that is best for all objectives, and the more complex the model, the more sensitive it will be to parameter misspecification. So, when selecting an interest rate model it is important to keep one's objective in mind. In our case, the goal is to produce realistic scenarios for the yield curve (and not to price derivatives).

3 Feasibility Test

As part of the new FTK, as of 2016, pension funds must conduct a feasibility test each year. Policymakers will use the feasibility test to assess a pension fund's risk profile and to examine whether the predetermined pension policy is sufficiently realistic and feasible for the long-term. The feasibility test involves a stochastic analysis of a pension fund's long-term financial position using a uniform economic scenario set with a 60-year horizon. As part of the new regulations, pension funds must predefine their ambition and risk profile in a so-called 'initial feasibility test'. This initial feasibility test must be submitted when pension funds introduce a new pension policy or after another such significant change, such as a new FTK. The test forces pension funds to think about their ambitions and risks in advance and to actively communicate about these factors to their social partners and regulators.

In addition to defining the risk profile on fund level, the feasibility test must indicate the potential consequences of the current pension policy for different generations. Here the key focus is on the degree to which the pension benefits' purchasing power is preserved over time. As mentioned before, most pension funds would ideally like to index accrued pension benefits each year to correct for inflation. When pension benefits are fully indexed for inflation, members' purchasing power remains the same. However, when pension funds are underfunded, they must abstain from indexing or even reduce accrued pension benefits. Before the financial crisis, most Dutch pension funds annually indexed pension benefits as a general rule. But, in current economic conditions, indexation for inflation has become more of an exception than a rule. Therefore, there is no guarantee that the purchasing power of the defined pension benefits will be preserved over time.

This section describes the tool that we specifically designed for the feasibility test to calculate and analyze various pension policy alternatives. In contrast to a classical ALM model, our tool not only analyzes how different policies affect a pension fund's financial position, but also provides insight into the consequences for each generation in the fund. With the help of this tool, we perform the feasibility test for various types of pension funds. We afterwards examine whether the feasibility test met its objectives and look for any practical flaws in the test that need to be changed.

Before we provide the results for the feasibility test, we first discuss the fundamentals of the test itself and offer a more in-depth discussion of our tool. Section 3.1 describes the economic scenario set provided by the DNB. In Section 3.2, we provide the definition of the feasibility test's primary metric: the pension result. The assumptions underlying our model, with respect to pension structure and the policy framework, are discussed in Section 3.3. Regarding the pension fund structure, we utilize a stylized pension fund representative of a standard Dutch pension fund. This section includes assumptions on the fund's participant file, specific demographic changes, and financial characteristics. In Section 3.4, we provide a detailed description of our tool, which is partly based on a paper by Chen et al. (2014) on modeling the development of pension assets and liabilities over time. In the final part of this chapter, Section 3.5, we discuss the results of the feasibility test.

3.1 DNB Scenario Set

In the year 2013, the Dutch government asked the Commission Parameters led by T. Langejan to advise on a stochastic scenario set that can be used in the feasibility test. In particular, the commission was requested to develop a uniform set of stochastic scenarios so that results for different pension funds can easily be compared. In February 2014, the commission published the paper *Advice Commission Parameters* (Langejan et al. (2014)). In this paper they recommend an economic scenario generator that is based on the model proposed by Koijen et al. (2010), also known as the 'KNW-model'. The KNW-model generates scenarios for overall equity returns, the term structure of interest rates and the development of price inflation. The economic scenarios are based on both historical data and the current economic vision that is formulated by the Commission Parameters.

The KNW-model was selected because of its good balance between realism and applicability. The Commission Parameters claims that the generated scenarios are realistic and based on accepted economic principles. In addition, the model is relatively simple compared to more sophisticated ALM models what makes it easy for pension funds to implement the feasibility test. Pension funds only have to convert their investment portfolio into an equity portfolio and a fixed income portfolio.

The Commission Parameters states that the economic scenario generator, based on the KNW-model, is no substitute for ALM models. The primary objective of the feasibility test and that of a classical ALM study are different. The feasibility test evaluates the pension fund's performance over a 60-year period, and focuses on the expected pension result and corresponding risks on fund and generation level. The 60-year horizon was chosen to include pension payments of all generations in order to identify intergenerational differences arising from the pension policy. A classical ALM study evaluates a shorter (medium-term) horizon focusing on the allocation of different asset classes. The classical approach is used to identify strategies that lead to an efficient investment policy mix. In addition to the different objectives, both analyses differ in complexity. The 'simple' scenarioset used in the feasibility test cannot be considered as a realistic substitute for the scenarios used in classical ALM studies.

Each quarter, the DNB will publish the uniform economic scenario set that must be used for the feasibility test. The scenario set contains 2,000 scenarios with a horizon of 60 years for equities, price inflation and interest rates. Pension fund's liabilities and bonds prices have to be calculated using the simulated term structure. The interest rate curve for time 0 is added to calculate the initial funding ratio. Credit returns are based on a weighted average of equity returns and bond yields. Wage inflation scenarios are derived from price inflation scenarios plus a real wage growth premium of 0.5% points.

In the beginning of 2015, the DNB published an example scenarioset based on the KNW-model. Figure 1 displays the scenarios for equity returns, inflation, and 1-year and 20-year interest rates. Also the initial term structure is added to the scenario set to calculate the initial liabilities and initial funding ratio.

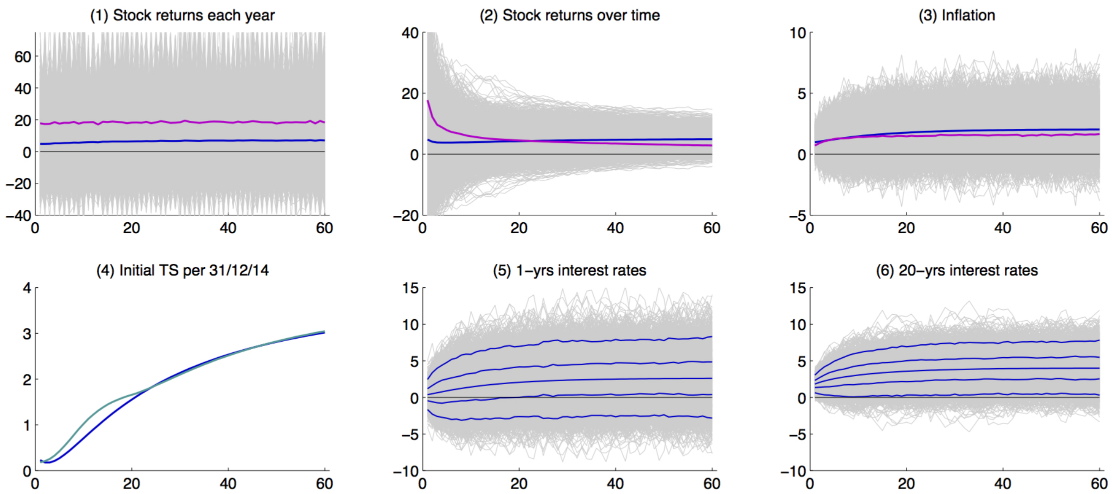

Figure 1: The DNB scenario set for Q1-2015. Graphs (1) and (2) display the arithmetic and geometric stock return scenarios plus the averages (blue) and the volatilities (pink). Graph (3) displays the inflation scenarios and the averages (blue) and the volatilities (pink) . Graph (4) displays the DNB term structure including the UFR component (green) and the initial term structure implied by the KNW-model (blue). Graphs (5) and (6) display the 1-year and 20-year interest rates including the 5th, 25th, 50th, 75th and 95th percentiles in blue.

If we analyse the scenario set, a few points stand out. First we see that the average return on equity is not equal over time. The average return increases from 5% to 7% in 30 years and then remains stable up to the horizon of 60 years; see (1). In addition, the volatility of the equity returns is on average 18%, which is lower than the 20% volatility proposed by the Commission Parameters. Secondly, the average price inflation grows to 2% in 30 to 40 years with a volatility of 1.5%; see (3). This does not correspond with the expected growth path of 5 years that is advised by the Commission Parameters. Furthermore, the initial interest rate curve given in the scenarioset is approximately, but not exactly, equal to the DNB interest rate curve on 31 December 2014 that is based on the UFR method; see (4). And finally, the interest rates can be (highly) negative, see the 1-years and 20-years interest rates. This is also in contradiction with the assumptions proposed by Commission Parameters who assume that (long-term) interest rates cannot be negative.

3.2 Pension Result

The feasibility test's objective is to offer insights on the development of the purchasing power of pension benefits. The development in purchasing power preservation over time can be expressed by the metric 'pension result'. Pension result is defined as the ratio of benefits that retirees actually receive over a particular time period versus received benefits that have been fully indexed for price inflation. For each participant, the pension result can be expressed as follows:

$$PR = 100\% \times \frac{\textit{Sum of received benefits}}{\textit{Sum of fully indexed benefits}} \tag{3}$$

The *sum of received benefits* depends on the simulations of the financial market variables in combination with the policy agreements on contributions, indexation, and pension cuts. The *sum of fully indexed benefits* also derives from the stochastic analysis but depends solely on price inflation. The pension result on fund level is calculated by summing the results for all current participants in the fund.

The two relevant outcomes of the feasibility test are: (1) the expected pension result, which is related to the pension fund's ambition, and (2) the pension result in a 'bad weather' scenario, which defines a pension fund's risk profile. One component of communicating the risk profile is to determine the lower bounds of these two outcomes. The feasibility test must then examine for each year whether the formulated pension ambition is realistic and if the 'bad weather' scenario remains within the boundaries specified by the pension fund board.

The stochastic analysis leads to 2,000 simulated outcomes for the pension result on fund level. These outcomes can be presented as a probability distribution of realized results, the so-called 'cloud'. To assess this cloud, scenarios are translated into orderly outcomes using quantiles. In the feasibility test, the expected pension result is equal to the median, and the pension result in a 'bad weather' scenario is set at the 5th percentile. To better understand the risk profile, the feasibility test also looks at the difference between the median pension result and the pension result in a 'bad weather' scenario.

In addition to the pension result on fund level, the pension result on individual level is also examined. Pension results for different generations are noteworthy, because the wealth distribution between generations is often unclear. The indexation policy and the chosen asset mix are of great importance for the wealth distribution across generations. For example, more indexation in the short term is attractive to older generations whose pension rights are relatively high. In contrast, high pension payments in the short term result in lower pension benefits in the long term, which is obviously detrimental to younger and future participants.

3.3 Stylized Pension Fund

In this section we define basic assumptions that we need to generate the results of the feasibility test. Section 3.3.1 discusses the participant file of our stylized pension fund. Section 3.3.2 describes the characteristics and policy framework of our stylized pension fund.

3.3.1 Participants File

To analyse the results for different cohorts we use a stylized pension fund that is based on realistic features. Following Commission Parameters and the paper Generation-effects Pension Agreement (Lever et al. (2012)) we use demographic data of the Dutch population to determine the pension fund's structure. The development of the population size and life expectancy are based on projections by the Central Bureau of Statistics (CBS) for 2014-2060, and extrapolated for the years beyond 2060.[4] The data for the population size and survival rates is gender specific and specified at a cohort level. Each specific cohort is defined as the group of people with the same birth year and the same accrued pension rights.

Our model will first use the initial population data as input, after which the future population's size for each cohort is generated using the survival rates. For example, the size of a male population of generation x at time t can be calculated as:

$$MalePop_t^x = MalePop_{t-1}^{x-1} \times p_x^{male}(t \mid t-1) \tag{4}$$

where $p_x^{male}(t \mid t-1)$ is a probability of a male person surviving to age x in period t conditional on this person having survived to age $x-1$ in $t-1$. New participants are based on the CBS predictions of future 25-year-olds. Note that future age distribution in real life will deviate from CBS predictions. However, in our model we do not take the uncertainty in demographic developments into account.

The top left graph in Figure 2 displays the age distribution of the total population (men + women) in 2015. Remarkable to see is the peek of the population between the ages 65 to 70. This is a result of the baby boom after World War II. This generation has just retired, which will put more pressure on the current workforce since smaller group of active participants will have to pay premium for a larger group of retirees.

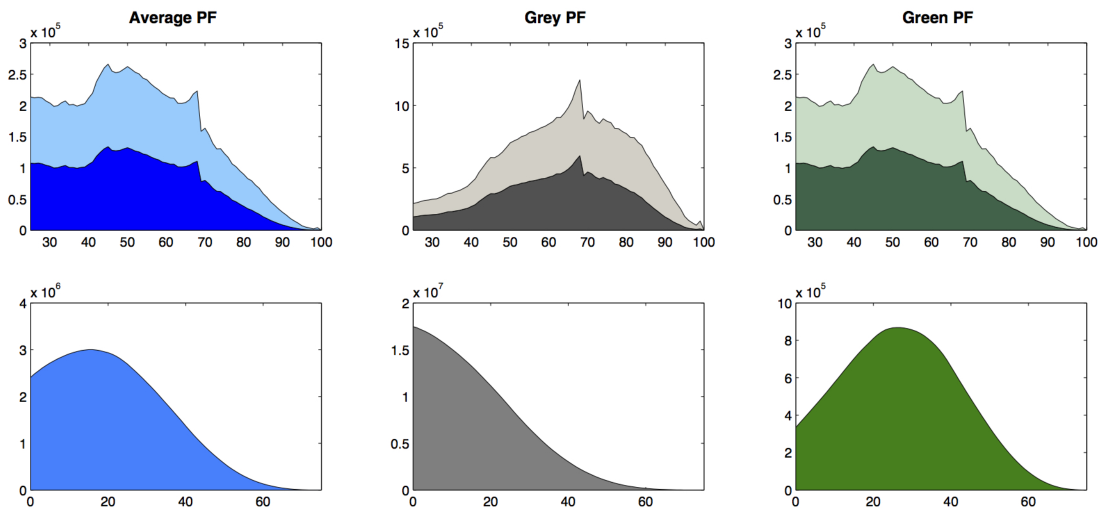

Figure 2: Initial population size for each age cohort. Upper left graph represents the total Dutch population per age cohort in 2015. The dark (light) blue area represents the men (women) population. The upper middle and right graphs represent the participant files of a grey and green pension fund, respectively. The bottom graphs display the related nominal cash flows.

The participant file based on the Dutch population will be consistent with an average Dutch pension fund in 2015. Because the model's output strongly depends on the fund's participant composition, we also consider a relatively young (green) and a relatively old (grey) pension fund.

[4]Source: www.cbs.nl and the paper Van Duin and Stoeldraijer (2014)

To obtain both participant files, we change the initial size of each cohort so that the green fund has relatively more young people, while the grey fund has relatively more elderly compared to the average Dutch pension fund. Hereby we assume that the survival rates and the number of new participants will remain the same as before.

3.3.2 Base Case Assumptions

Before we can use the tool we first have to determine some underlying assumptions for the characteristics of our stylized pension fund. Therefore we will use the same assumptions as in Langejan et al. (2014):

1. The pension scheme is an average-salary Defined Benefit scheme, where all active members have equal annual wages
2. The annual accrual rate for pension rights is 2% of the wage level
3. Full indexation has been granted up to now
4. The benefits are (conditionally) indexed to price level; Wages will increase each year based on wage inflation (price inflation + 0.5%)
5. The initial funding position is set to 105% in nominal terms
6. An individual participant is assumed to enter the pension fund at the age of 25, retire at the age of 65 and deceases at the age of 100 at maximum; The pension scheme contains only the 'active' participants and considers only the old-age pension
7. Investment policy: The asset allocation mix consists of 50% bonds and 50% equity and is rebalanced at each time period; The interest rate risk is additionally hedged with 40%
8. Premium policy: the contribution rate is 17% of the wage level; Premium discount when $FR > 145\%$ (equal to 10% of the surplus); Premium increase to 20% when $FR < 105\%$
9. Indexation policy: condition indexation with linear scale between 110% (no indexation) and 130% (full indexation); Backlogged indexation when FR above 130% in which maximum 1/5 of the difference between the funding ratio and the threshold may be used to offset missed indexation or pension cuts
10. Pension cuts policy: cuts to $FR = 105\%$ when $FR < 105\%$ for 5 consecutive years

Given the assumptions we can conclude that at any given time in the 60 year of simulations there are maximum 75 generations participating in the pension fund. A participant takes part in the pension scheme for a maximum period of 75 years. The size of each generation in time is determined by the initial population size and the survival rates. Furthermore, the active participants acquire annually 2% of their salary as accrued pension rights that will be paid each year from retirement. After 40 years the final pension benefit level will be 80% of the average wage during working period plus realized indexation.

Because we assume fully indexed pension rights in our initial situation, the pension benefits at time 0 are known for all cohorts. In combination with the life expectancy for each cohort, we can calculate the future cash flows for all participants. The bottom graphs of Figure 2 display the nominal cash flows for all three stylized pension funds. The duration of the initial liabilities of the average Dutch pension funds is equal to 16.9 years. This corresponds with DNB figures concerning the average duration of Dutch pension funds in 2010.[5] For the grey and green funds, the duration is equal to 13.2 years and 21.6 years, respectively. The grey fund's duration is similar to the average of the 5% most grey funds in the Netherlands, while the green fund's duration is similar to the 5% most green funds.

[5] Source: www.dnb.nl

3.4 Model Setup

In the previous subsections, we discussed the DNB scenario set and the assumptions for our stylized pension fund. Both are used as input for our model to evaluate the pension fund and to obtain feasibility test results. Our analysis makes use of a relatively 'simplified' ALM model, capable of fully modeling the pension fund with all its attendant policy instruments. Using a very sophisticated actuarial model would result in a more detailed and realistic picture of the pension fund's liabilities. However, the modeling process would then become very time-consuming and rather complicated. For that reason, it might be desirable to use a less sophisticated model for the liabilities. A less complex model is especially preferable for pension-fund asset managers, who often lack the necessary knowledge and time needed to employ the extensive actuarial model.

The feasibility test has two objectives: (1) to measure the effects of different policy measures on a particular pension fund's future financial position, and (2) to subsequently observe how these measures affect the preservation of purchasing power for participants on fund level and for different generations. Hence, pension benefits and contributions must be calculated not only at the aggregate level but also at the cohort level. This paper therefore adopts a generational modeling approach similar to that used in Chen et al. (2014). The 'Chen' model records the cash flows of different generations into separate accounts. This technique enables us to gain a comprehensive overview of benefits paid and contributions received for each specific generation throughout the model's simulation period.

In contrast to Chen et al. (2014), we focus on an open fund setting rather than on a closed pension scheme. In a closed pension scheme, participants do not pay contributions or accrue benefits from the beginning of the first year. In an open fund, current participants continue to pay contributions and accrue benefits over the whole simulation period. In addition, new participants join the fund each year.

In section 3.4.1 we first calculate the initial values for liabilities and assets. Then, in Section 3.4.2, we illustrate the simulation process, including cash flows in and out, the valuation of the liabilities, and the state of the fund at each point in time.

3.4.1 Initial Values

Before we start with the simulation procedure, we first have to determine the initial values of the pension fund's assets and liabilities. These initial values are based on assumptions we made for our stylized pension fund. We derive the initial assets by multiplying the initial liabilities with the predetermined initial funding ratio. This can be expressed as:

$$A_0 = L_0 \times FR_0 \tag{5}$$

To determine the liabilities at each point of time, we use two matrices: a *benefit matrix* for tracking the accrued pension rights and a *discount matrix* to derive the present value.

Benefit matrix

The benefit matrix represents the total accrued benefit claims for each generation as a percentage of the prevailing wage level in all 2,000 scenarios. Because we assume that the initial benefits are fully indexed up to now, we can calculate the relative benefits by multiplying the accrual rate of 2% with the number of active years in the pension scheme. The accrued benefits matrix at time zero can be displayed as follows:

$$B_0 = \begin{bmatrix} 0 & 0.02 & 0.04 & \dots & 0.8 & \dots & 0.8 \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \\ 0 & 0.02 & 0.04 & \dots & 0.8 & \dots & 0.8 \end{bmatrix} (2000 \times 75) \tag{6}$$

The columns represent the benefits for each generation beginning with the 25-year-olds (no accrued benefits yet) up to age of 99 (accrued 80% of the current wage level). The rows in the matrix represent the different scenarios. Note that the initial accrued benefits for each generation are the same in all scenarios. However when the simulation starts the accrued benefits will differ between scenarios due to the indexation policy, which is conditional on the funding ratio. To calculate the total accrued benefits, the benefit matrix needs to be multiplied by the total population size for each cohort and the wage level, which is assumed to be 1 in the initial situation.

Discount Matrix

To calculate the present value of the liabilities, the accrued benefit elements are multiplied by a specific discount element that depends on gender, age cohort, time period and scenario-path. The discount element is constructed by taking the sum of discount factors, retrieved from the simulated term structure in the current period in a particular scenario, times the survival probabilities for each maturity. So the discount element can be calculated as follows:

$$D_{t,s}^{x} = \sum_{i=max(65-x,0)}^{99-x} p_{x}^{g}(i \mid t)(1 + R_{t,s}^{(i)})^{-i} \tag{7}$$

where $p_x^g(i \mid t)$ is the probability that a male/female(g) aged x at time t will survive to year i. Hence the survival rates incorporated in the equation are conditional survival probabilities of surviving up to a particular year in the future given that a person has survived up to this year. $R_{t,s}^{(i)}$ denotes the interest rate for maturity i from the current nominal term structure in scenario s. The discount factors derived from the interest rates can be added together because all future claims (equal to the accrued benefits) are the same.

For extra clarification we present the case of a 65-year-old female who accrued 80% of the current average wage level. The discount element $D_{t,s}^{65}$ represent the sum of discount factors for the accrued benefit to be paid now and up to 35 years into the future multiplied by the probability she survives up to the given year.

To calculate the value of the liabilities, the discount matrix and the total accrued benefit matrix are multiplied element-wise. Subsequently, the values for the age cohorts are added together to obtain the present value per scenario.

3.4.2 Simulation

After the initial liabilities and assets are determined, the simulation process can be started. For the 2,000 constructed scenarios for equities, inflation and interest rates, we simulate the development of the stylized pension fund. The asset side of the balance sheet is influenced by pension payments, contributions and investment returns. The liabilities depend on the accrual rate of pension rights and the indexation of these rights throughout the simulation period. The simulation process for each year t is as follows:

At the beginning of each simulation period

1. First, we use the financial position of the fund at the beginning of the simulation period to calculate the premium and indexation levels
2. Second, we pay out pensions to all retired participants and receive premiums from all active participants. These cash flows affect the value of the asset side of the balance sheet
3. Subsequently the assets are reallocated to the different asset classes according to the predetermined investment policy

Jump to the end of the simulation period

4. We update the participant file for demographic changes and add new entitlements to the existing rights using the predetermined accrual rate. This additional right is a percentage of the prevailing wage for each age cohort

5. Next, all wages are indexed by the realized wage inflation as has been simulated in the scenarioset

6. All entitlements are increased by the indexation fraction multiplied with the realized price inflation as well as with recovery indexation as calculated in step 1. In case the price inflation is negative, the pensions rights are effectively reduced

7. Then a new term structure is determined to calculate the new present value of the liabilities by multiplying the updated benefit matrix and the discount matrix

8. We calculate the new value of the assets by multiplying the old assets (adjusted with contributions and pension payments) with the returns generated by the financial market model

9. Finally, given the updated asset and liabilities value, we calculate the new funding ratio. This funding ratio determines the premium and the levels for the next period

The simulation process ends after 60 years and gives output for the relevant performance measures, including the distribution of the funding ratio and the pension result over time. Note that new participants (future 25-year-olds), who join the participant file at a later moment, are not included in the pension result.

3.5 Feasibility Test Results

The DNB has published an instructional letter for pension funds regarding the submission of feasibility test.[6] The results that funds file must include several tabs, with a questionnaire, data series, and a recovery plan template. The questionnaire is a tool to determine whether the legal requirements have been met. The required data series include the following: formulated critical limits for the fund's ambition and 'bad weather' scenario, realized results of the feasibility test, and information regarding the composition of the participant file. The realized results must include the pension result on fund level for nine percentiles, as well as the pension results for six age cohorts for nine percentiles.[7]

In previous sections, we described the example scenario set, our stylized pension funds, and the model behind our tool. In this section, we present the results and check whether the feasibility test serves its objective. In Section 3.5.1, we first discuss pension results on fund level and on generation level for three stylized pension funds based on the Dutch population. Subsequently, in Section 3.5.2, we consider the solvency position of the pension fund in the stochastic analysis. In addition, we examine certain additional measures and provide a summary of the sensitivity analysis presented in Appendix E.

3.5.1 Pension Results

Table 1 presents pension results for the three stylized pension funds (average, green, and grey) on fund level and for six generations. Unlike the actual filing, we only show the most relevant percentiles, with the goal of illustrating our results in a clearer manner. The percentiles displayed include: the 5th percentile ('bad weather' scenario), the median (ambition), and the 95th percentile (upward potential). In addition, we present the difference between the median and the 5th percentile, an important figure to consider when determining a pension fund's risk profile.

[6]Source: http://www.toezicht.dnb.nl/binaries/50-233550.pdf

[7]Percentiles: 0, 5, 10, 25, 50, 75, 90, 95, 100; and Ages: 25, 35, 45, 55, 65, 75.

Pension Result	**5 perc.**	**50 perc.**	**95 perc.**	**50p - 5p**
Average Pension Fund				
Fund Level	51.2%	82.7%	99.9%	31.5%
25y	43.4%	87.2%	100.0%	43.8%
35y	38.1%	82.8%	100.0%	44.7%
45y	37.5%	79.5%	100.0%	42.0%
55y	41.1%	78.3%	100.0%	37.3%
65y	55.2%	82.8%	99.8%	27.7%
75y	63.1%	86.4%	99.8%	23.3%
Green Pension Fund				
Fund Level	46.2%	79.0%	99.7%	32.8%
25y	40.2%	80.3%	100.0%	40.2%
35y	35.8%	77.0%	100.0%	41.2%
45y	35.4%	75.3%	100.0%	39.9%
55y	41.7%	77.7%	99.9%	36.0%
65y	56.8%	83.0%	99.6%	26.2%
75y	65.1%	87.3%	99.7%	22.2%
Grey Pension Fund				
Fund Level	58.0%	88.7%	99.9%	30.6%
25y	49.4%	99.2%	100.0%	49.8%
35y	43.4%	96.5%	100.0%	53.2%
45y	39.7%	92.1%	100.0%	52.4%
55y	41.1%	86.8%	100.0%	45.7%
65y	54.3%	87.1%	99.9%	32.8%
75y	62.5%	88.3%	99.9%	25.8%

Table 1: The percentiles (5%, 50% and 95%) of the pension result outcomes for the three stylized pension funds (average, green, and grey). The outcomes are given at pension fund level and for six initial age-cohorts representing the generations that are present in the current participant file. The difference between the 50th and the 5th percentile is used as a risk measure.

Pension result on fund level

We first consider pension results on fund level. The median pension result of about 83% for the average pension fund indicates that the initial participant file is expected to receive 83% of target pension payments over time. For the 95th percentile, we see that the target of fully indexed pension payments is almost achieved over the 60-year simulation period. The pension result under the 'bad weather' scenario is only 51.2%.

If we compare the multiple population compositions, the grey pension fund outperforms the other two funds with an expected pension result of almost 90% and a 'bad weather' scenario outcome of 58%. In addition, it has the smallest difference between the median and the 5th percentile, which implies that it has the lowest risk profile of the three pension funds.

Pension result on generation level

The measure 'pension result' indicates the degree to which the purchasing power of accrued pension benefits is preserved. Thus, in theory, a fund-level pension result of 90% means a 10% overall loss in purchasing power. However, we must realize that this outcome has a different meaning for a 70-year-old retiree (who still has approximately 20 years to live) than for a 30-year old, for whom the loss in purchasing power is calculated over the next 60 years. An important component of the feasibility test is determining the generational effects of the current pension policy framework. Because the above study is conducted at the aggregate level, it is not very transparent regarding whether particular generations (or cohorts) are worse off than others. To analyze these generational effects, the feasibility test looks at the pension result for each generation.

Figure 3 displays pension results for all age cohorts after 60 years of simulation. For clarification, the pension result for each age cohort expresses the ratio of pension payments for that specific age cohort as compared to a fully indexed pension. Figure 3 and Table 1 both show that middle-aged participants have the lowest pension result for the average pension fund. Table 1 shows that the expected pension result for a 25-year-old is about 87%, about 78% for a 55-year-old, and about 86% for a 75-year-old retiree. The expected pension result line for the average fund (blue) in Figure 3 has a convex shape, with its lowest point at about 55 years. In a 'bad weather' scenario, we see that 35-year-olds have the lowest pension result. So, under such conditions, they would only receive about 38% of target pension payments.

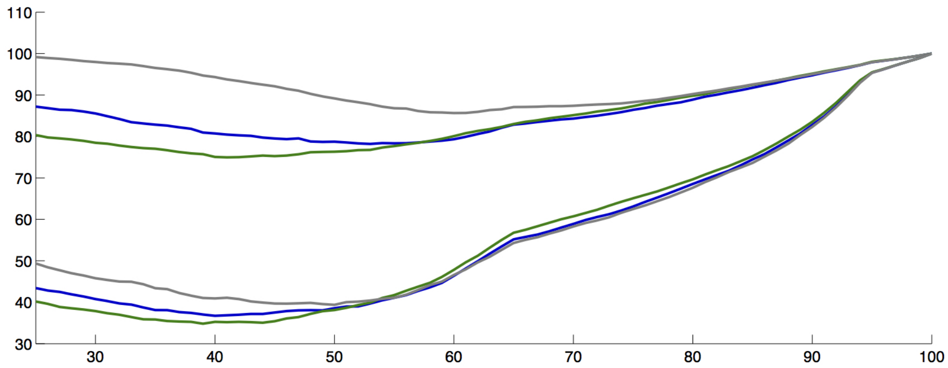

Figure 3: Pension results for all initial age-cohorts: 25-year-olds to 100-year-olds. The blue, green and grey lines present the 5th and 50th percentiles for the average, green and grey pension funds, respectively.

The spread between the median and the 'bad weather' scenario indicates the degree of risk, which is the highest among the younger generations approximately 44% and approximately 45% for the 25-year-olds and 35-year-olds, respectively. Explanation for the lower spread for a 25-year-old could be that, despite the long-term uncertainty, younger generations have more time to recover from missed indexation. For a 75-year-old only short-term indexation is important. We see that the upside for older generations is less than 100% as opposed to younger generations, which is probably due to the low initial funding ratio.

Next, we consider the impact of age distribution in the fund's participant file. The expected pension results for the different generations in the grey pension fund are higher as compared to results for the green pension fund. We see that the expected pension payments for younger generations are almost equal to fully indexed pension payments. For example, 25-year-olds have an expected pension result of approximately 99%. In contrast, the risk profiles for the different generations seem to be lower in the green pension fund. In a 'bad weather' scenario, older generations receive better outcomes from a green pension fund than from a grey pension fund. In a green pension fund there is a relatively large group of young participants who can absorb possible deficits created by the large group of elderly. The younger generations are better off in a grey pension fund. An explanation for this finding could be that, after some time, the relative size of elderly people will reduce so that the age-structure is more balanced leading to a better financial position.

Pension result over time

Figure 4 illustrates how pension results develops over time, with the goal of providing a clearer understanding of how the stochastic analysis influences the final results shown in Table 1. The upper graphs provide pension results on fund level. The graphic display of the 'clouds' offers very useful insight into the metric during the 60-year simulation period. The grey lines display all 2,000 scenarios, and the blue lines represent the percentiles (5%, 50%, and 95%). In addition, we look at the quality of indexation over time on generation level, illustrated in the bottom six graphics representing the six generations. The blue, green, and grey lines in each graph present the percentiles (5% and 50%) for the average, green, and grey pension funds, respectively. In these graphs, pension payment dates for each age cohort are shown on the horizontal axes. For example, a 65-year-old receives his or her pension in the first year of the simulation (2015), whereas a 25-year-old will receive his or her first pension payment in 2055.

Figure 4: Pension result over time. Upper graphs display the pension result on pension fund level for three pension funds (average, green and grey) over the 60-year horizon. Blue lines indicate the 5th, 50th and 95th percentiles. Bottom graphs display the pension results over time for six generations. The blue, green and grey lines present the 5th and 50th percentiles for the average, green and grey pension funds, respectively.

According to the feasibility test, any missed indexations in the past are not taken into account at the beginning of the simulation period. Accordingly, pension results on January 1, 2015 are 100%. As a result, it is more difficult to compare the pension results for different generations. Because of missed indexation or pension cuts in the past, many pension funds have a lower pension result in reality. Not including these past events can have a significant impact on the test's outcome, especially for grey pension funds.

Because we start with a funding ratio of 105%, there will be no indexation in the first year, which immediately leads to a lower pension result. The pension results for the average and green pension funds steadily decline to approximately 80%, while pension results over time for the grey fund have a convex shape and return to the 90% level after an interval. We see that, in some undesirable economic scenarios, benefits paid to participants are equal to the initial nominal value of accrued rights. In the worst case, benefits paid are lower than the initially guaranteed figure, due to pension cuts. The upside is limited, as there is an effective cap on maximum pension benefits (i.e. full indexation is the most that a participant in this pension scheme can expect).

Nevertheless, for some scenarios we see that the pension result is higher than 100%. This outcome can be explained by how the feasibility test deals with pension rights in cases of negative inflation. Normally, pension funds do not index pension rights when inflation is negative. However, according to the feasibility test, we must assume that the indexation scheme does not differentiate between negative and positive inflation. Thus, under circumstances marked by negative inflation and a low funding ratio, fully indexed pension payments (equal to the denominator in the equation defining the pension result) would be reduced by the given inflation level. Actual pension rights (i.e. the numerator), however, would remain the same, because no indexation is applied. This would lead to a pension result above 100%.

Pension results would also be above 100% in situations where pension funds provided indexation based on wage inflation, which normally exceeds price inflation. In the feasibility-test design, as well as in the definition of the pension result, the ambition of a pension fund is always based on price-indexed pensions, to allow for the easier comparison of results across pension funds. This assumption is illogical, because in reality, various funds follow the wage index and therefore formulate their ambition and risks are based on wage inflation.

3.5.2 Solvency Position

To shed more light on previously described results, Table 2 presents several figures concerning the solvency positions of the three pension funds. The upper panel depicts funding ratios in the final year of the simulation. The expected funding ratios (mean and median) for the three pension funds increased relative to their starting positions of 105%. This result could be due in part to the indexation mechanism, which does not always adjust pension rights for inflation.

Solvency Position	**Average**	**Green**	**Grey**
Mean FR	172.2%	141.9%	373.8%
50th-percentile FR	134.2%	128.0%	161.2%
5th-percentile FR	92.2%	89.9%	95.5%
P(FR < 105)	21.8%	24.2%	16.8%
P(FR < 100)	15.0%	17.0%	11.4%
Pw(FR < 100)	87.8%	91.5%	77.2%
P(FR = 5y below 105)	5.2%	5.7%	4.0%
Max drawdown$_{t,t+1}$	24.3%	25.8%	22.8%

Table 2: Summary statistics representing the solvency position for the three pension funds (average, green and grey). The upper panel includes the average, 5th percentile and 50th percentile of the funding ratio (FR) in the year 2075. The lower panel includes the probability of funding shortfall in one of the years of simulation ($P(FR < 105)$), the probability of underfunding in one of the years of simulation ($P(FR < 100)$), probability of underfunding within the 60 years of simulation ($Pw(FR < 100)$), the probability of funding shortfall five years in a row *(P(FR = 5y below 105))* and the max drawdown after the first year of simulation.

In the bottom panel, we display the probabilities of underfunding ($FR < 100$) and shortfall ($FR < 105$) over the whole simulation horizon. The average probability that the average pension fund cannot meet the minimum required capital is 21.8%. The average probability that this fund is not able to pay future pension rights is 15%. The probability of underfunding for the grey pension fund is lower but still large with 11.4%. We also consider the 'within probability' as suggested by Kritzman and Rich (2002) for the funding ratio. This measure provides the probability that the funding ratio is below a certain value across a period of years at least once, instead of in a particular year only. The chances at ever becoming underfunded are equal to a probability of 88%, 92% and 77% for the average, the green and the grey pension fund, respectively. So for a substantial part of the scenarios there is a specific year during the 60-year period where the pension funds cannot meet their future obligations. The Dutch pension sector should be aware of this high probability.

Next we examine the probability of a funding shortfall five years in a row. If this situation were to occur, the pension fund in question would have to lower promised pension rights to improve its financial position. With a probability of 5%, pension cuts are not an exceptional situation. This high risk on cuts raises the question as to whether the present pension deal can guarantee nominal pension rights, in which pension cuts should only be used as an ultimatum remedy.

The max drawdown shows the largest fall in the funding ratio after the first simulation year. This largest drop for the three pension funds is created by scenario 1013, in which the stock return is -39%. The term structure for this scenario increases on average by 1%. The interest rates for the lowest maturities increased the most, but the increased rates for the long maturities had the greatest impact. This is reflected in the largest maximum drawdown for the green pension fund, the fund with the most long-term obligations.

Funding ratio over time
Figure 5 shows the funding ratio over time for the three pension funds. Looking at the average pension fund, the median slightly increases over time and remains close to 130%. That the expected funding ratio remains near this boundary is logical, because it is the threshold for backlogged indexation. So, if indexation was missed in the past, 20% of the surplus would be used for recovering the purchasing power of accrued pension rights. If we look at the effects of the participant composition, the green pension fund has a similar pattern over time. The median funding ratio of the grey pension fund grows more quickly and is substantially above the 130% funding ratio after 60 years.

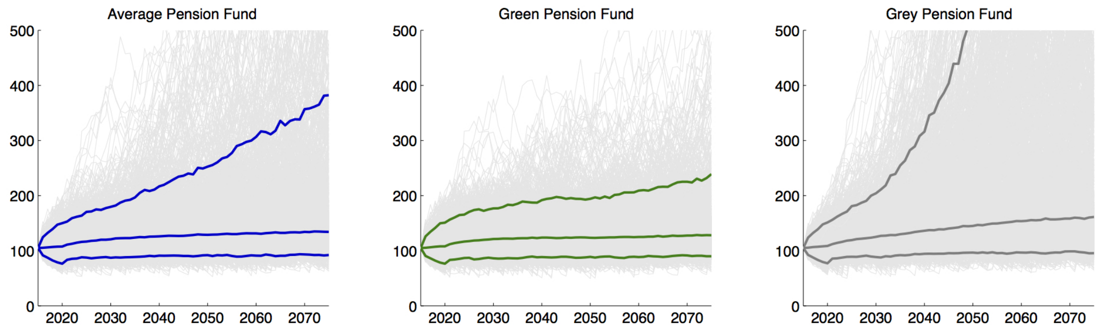

Figure 5: Funding ratio levels for three pension funds (average, green and grey) over the 60-year horizon. Blue lines indicate the 5th, 50th, and 95th percentiles

The development of the 5th-percentile funding ratio, representing the downside risk, is similar for all pension funds. The funding ratios remain above a certain level due to the chosen pension policy. In 2020, we note a clear contraction, which can be explained by the fact that that year represents the first time that pensions cuts might occur. After a fund experiences five years of shortfall, accrued pension rights must be reduced, so that its financial position recovers. The upward potential for the three pension funds varies. The long horizon of 60 years can lead to extreme outcomes on the upside. The 95th percentile for the grey pension fund's funding ratio seems to grow exponentially, whereas the upside for the average and green pension funds increases linearly.

Other figures
Figure 6 presents several other figures, measured over time so as to better understand the development of the funds' financial positions and to identify the underlying drivers that affect the quality of indexation. In the first graph, we see that as time passes, the probability of full indexation increases, and the probability of no indexation decreases. The second graph demonstrates the probability of pension cuts and the average size of such a pension cut, which fluctuates at around 12%. Finally, the third graph presents the expected premium level over time. Of particular note is that the grey pension fund's medium premium level goes to 5% as its solid financial position allows it to give premium discounts to its active participants.

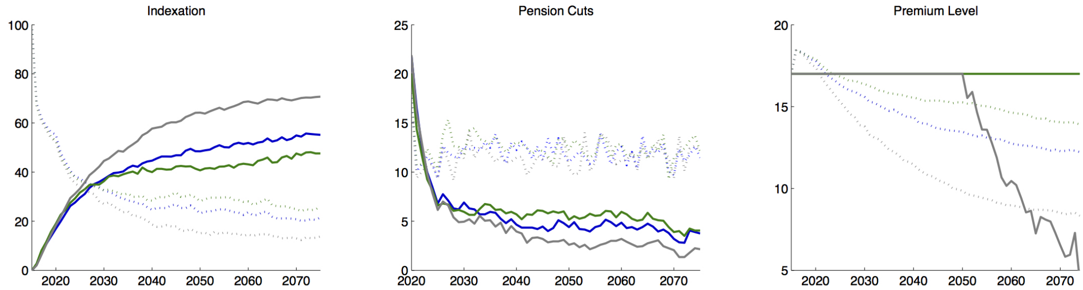

Figure 6: Other figures over time: (1) Probability of full indexation ($FR > 130$) (solid) and probability of no indexation ($FR < 110$) (dotted); (2) Probability of pension cuts (solid) and the average pension cut level (dotted); (3) Median premium level (solid) and average premium level (dotted).

Summary of the sensitivity analysis

The previous section suggested that it is challenging to compare results across generations. It is more interesting to weight the outcomes for each age cohort based on different input variables for stochastic analysis. In Appendix E, we present a detailed sensitivity analysis with respect to the initial funding ratio and investment allocation mix. Here, we provide a brief summary of the sensitivity analysis.

A high (low) initial funding ratio leads to a higher (lower) expected pension result and to a higher (lower) average funding ratio. The initial funding ratio has the greatest impact on the older generations. A high initial funding ratio results in more indexation at the beginning of the simulation, which is particularly attractive for older generations with large pension rights. The opposite is true regarding a low initial funding ratio. Younger generations have enough time to recover from a current pension deficit.

If we look at the asset allocation mix, we see that the more investment risk a pension fund takes on, the higher the expected results but also, the greater the dispersion around the expected results. It is worth mentioning that the likelihood of pension cuts is lower for the risky portfolio than for the riskless portfolio. The reason could be that the risky portfolio has a higher probability of recovery within the five-year recovery window. That said, more risk leads to, on average, higher pension cuts. And finally, the asset allocation mix has the greatest effect on the younger generations.

4 KNW model

As mentioned before, the Commission Parameters was asked to provide advice regarding the uniform scenario set for the feasibility test. In Langejan et al. (2014), the commission proposed the use of the capital market model described in Koijen et al. (2010) to generate economic scenarios. The proposed model is, in comparison with advanced capital market and ALM models, a relative simple model with a limited number of asset classes. The KNW model (only) generates scenarios for the overall equity index, the term structure of interest rates, and the development of price inflation over time.

The financial market model in Koijen et al. (2010) is closely related to Brennan and Xia (2002), Campbell and Viceira (2001), and Sangvinatsos and Wachter (2005). These papers focus on the optimal allocation to long-term bonds and show that it is ideal to hedge time variation in real interest rates. Important characteristic of the KNW model is that both nominal bond yields and equity returns depend on the inflation process. In addition, the model accommodates time-varying interest rates, inflation rates, and bond risk premia. The dynamics of the real interest rate and expected inflation are modeled using two state variables, which subsequently leads to (auto)-correlation between them. The state variables follow a mean-reverting process around zero, driven by a four-dimensional vector of independent Brownian motions representing the uncertainty in financial markets.

The model is formulated in continuous time, and the processes that generate the scenarios can be classified as stochastic differential equations, discussed in Section 2.4. However, to estimate the model parameters in the KNW model and to generate scenarios, a discretized version is required. Additional technical background information regarding the capital market model, including details related to the derivation and to the estimation procedure, are provided in Koijen et al. (2005), Koijen et al. (2006), Draper (2012) and Draper (2014).

Koijen et al. (2010) developed the model primarily for the U.S. financial market and estimated the model parameters with U.S. data. In 2012, the Dutch Bureau for Economic Policy Analysis (CBP) used the KNW model to evaluate the new Dutch pension agreement and to examine the resultant generation effects. For this purpose, Draper (2012) estimated the model parameters using historical data relevant to Dutch pension funds. The paper also compared the estimation results to the estimates for the United States. Draper (2012) observed that the results for the Netherlands deviated in several aspects from those to for the United States. In particular, the coefficient estimates were less significant for Europe. At the request of the Commission Parameters, the CPB re-estimated the KNW model based on more recent Dutch data in Draper (2014). The CPB has indicated that the estimation in Draper (2014) did not lead to the maximum likelihood. Most likely, the estimated set of parameters is a local optimum. This is a known problem for models with many dimensions (i.e. models with a large number of parameters to estimate).

Following the recommendation of the Commission Parameters, the DNB generates economic scenarios for the feasibility test, using the estimated parameters in Draper (2014) as the basis for the model. However, as mentioned in Draper (2012), the parameter uncertainty has an important implication for the evaluation of a pension fund's (future) financial position. Therefore, this paper tests the robustness of the results by estimating the model parameters over a longer period of time than in Draper (2014). In particular, we first estimate the model parameters using relevant data for the period from 1972 to 2014. We then compare the likelihood of our estimation with the likelihood value of the parameter sets presented in Draper (2012) and Draper (2014). In addition, we assess the fit of the models and examine the average yield curve, the volatility of the bonds, and the predictability of bond returns using the Campbell-Shiller regression.

The Commission Parameters and the DNB have largely adopted the methodology proposed by Koijen et al. (2010) for generating scenarios. No additional restrictions are imposed when estimating the model. However, the model does not immediately meet the expectations of the Commission Parameters and the DNB. Therefore, some parameters must be calibrated after estimation to make the model consistent with these expectations. In addition, to ensure that the scenario set remains representative of current economic conditions, economic scenarios should be periodically adapted to the prevailing interest-rate level. Also, the long-term average return on bonds must be set equal to the ultimate forward rate (UFR) of 3.9%, following the recommendations set forth by the Commission UFR. To that end, for each period, an adjustment is made to several variables of the KNW model.

This chapter provides technical documentation related to the KNW model, as well as specifics regarding the derivations, the estimation, and the calibration. Furthermore, new estimation results are compared with earlier results, and we examine the fit of the model. In Sections 4.1 and 4.2, we first describe the methodology and the estimation procedure based on papers by Draper (2014) and Koijen et al. (2010). Next, in Section 4.3, we present the estimated model and discuss several implications based on the model parameters. The results are also compared with results from Draper (2012) and Draper (2014). To gauge the fit of the model, in Section 4.4 we examine the average yield curve, the volatility of the bonds, and the predictability of bond returns. And finally, in Section 4.5, we explain the calibration and actualization of the model parameters as described by the Commission Parameters and the DNB.

4.1 Methodology

The uncertainty and dynamics of the KNW model are governed by two state variables $X = (X1, X2)^{'}$, which follow a mean-reverting process around zero

$$dX_t = -KX_t dt + \Sigma_x' dZt \tag{8}$$

where K is a 2×2 matrix, $\Sigma_x' = [I_{2\times 2} 0_{2\times 2}]$ and Z denotes a four dimensional vector of independent Brownian motions which drive the uncertainty in the financial market. Four sources of uncertainty can be identified:

- uncertainty about the real interest rate
- uncertainty about the instantaneous expected inflation
- uncertainty about unexpected inflation
- uncertainty about the stock return

Notice that, due to Σ_X, only the Brownian motions that drive uncertainty for real interest rates and expected inflation have impact on the state variables. The uncertainty and dynamics in the real interest rate and in the instantaneous expected inflation are modelled using the state variables. Any correlation between the interest rate and inflation is modelled using δ_{1r}' and $\delta_{1\pi}'$. More precisely, for the instantaneous real interest rate (r) holds

$$r_t = \delta_{0r} + \delta_{1r}' X_t \tag{9}$$

and for the instantaneous expected inflation (π)

$$\pi_t = \delta_{0\pi} + \delta_{1\pi}' X_t \tag{10}$$

The actual inflation or price index (Π) is equal to expected inflation (π) in combination with unexpected shocks:

$$\frac{d\Pi_t}{\Pi_t} = \pi_t dt + \sigma_\Pi' dZ_t, \quad with \quad \sigma_\Pi \in R^4 \quad and \quad \Pi_0 = 1 \tag{11}$$

The stock index S develops according to

$$\frac{dS_t}{S_t} = (R_t + \eta_s)dt + \sigma_S' dZ_t, \quad with \quad \sigma_S \in R^4 \quad and \quad S_0 = 1 \tag{12}$$

where R_T is the nominal instantaneous interest rate (which we derive below) and η_s the constant equity risk premium. Because the equity returns are based on the nominal interest rate and the equity risk premium, the KNW model has no fixed expected equity returns.

The model is completed with the specification of the nominal stochastic discount factor with the time-varying price of risk affine in the state variables. The stochastic discount factor is used to determine the value of all cash flows in all states of the world.[8] The stochastic discount factor may be interpreted as an indicator for the marginal utility between consumption today and in the future. In a good state of the world the value of a cash flow is lower than in a bad state of the world. This marginal utility ratio is for everyone the same in the case of complete markets (everyone can trade in all risks). The nominal stochastic discount factor (ϕ_t^N) is given by

$$\frac{d\phi_t^N}{\phi_t^N} = -R_t dt - \Lambda_t' dZ_t \tag{13}$$

with the time-varying price of risk Λ

$$\Lambda_t = \Lambda_0 + \Lambda_1 X_t, \quad with \quad \Lambda_t, \Lambda_0 \in R^4 \quad and \quad \Lambda_1 : 4 \times 2 \tag{14}$$

The price of risk will depend on the risk aversion of investors. Assume no risk premium for unexpected inflation, i.e. the third row Λ_1 contains zeros only. This restriction is imposed because unexpected inflation risk cannot be identified on the basis of data on the nominal side of the economy alone (see Koijen et al. (2010))

$$\Lambda_1 = \begin{bmatrix} \Lambda1_{(1,1)} & \Lambda1_{(1,2)} \\ \Lambda1_{(2,1)} & \Lambda1_{(2,2)} \\ 0 & 0 \\ \Lambda1_{(4,1)} & \Lambda1_{(4,2)} \end{bmatrix} \tag{15}$$

The stochastic discount rate can be used to determine the value of all assets in a complete market, for instance the fundamental valuation equation (see for instance Cochrane (2005)) of the equity index

$$Ed\phi^N S = 0 \tag{16}$$

implies the expected value of a discounted stock price does not change over time. This equation in combination with the constant equity risk premium implies a restriction

$$\eta_S = \Lambda_t' \sigma_S \tag{17}$$

which implies $\sigma_S' \Lambda_0 = \eta_S$ and $\sigma_S' \Lambda_1 = 0$. This restriction is imposed on the model.

Given the nominal stochastic discount factor we can define the real stochastic discount factor as $\phi^R = \phi^N \Pi$. For the real stochastic discount factor we find

$$\begin{aligned} \frac{d\phi^R}{\phi^R} &= -(R_t - \pi_t + \sigma_\Pi' \Lambda_t)dt - (\Lambda_t' - \sigma_\Pi')dZ_t & (18) \\ &= -r_t dt - (\Lambda_t' - \sigma_\Pi')dZ_t & (19) \end{aligned}$$

As a consequence we obtain the instantaneous nominal interest rate

$$\begin{aligned} R_t &= r_t + \pi - \sigma_\Pi' \Lambda_t \\ &= (\delta_{0r} + \delta_{0\pi} - \sigma_\Pi' \Lambda_0) + (\delta_{1r} + \delta_{1\pi} - \sigma_\Pi' \Lambda_1) X_t \\ &\equiv R_0 + R_1' X_t & (20) \end{aligned}$$

Next we determine the nominal term structure. In this economy, bond yields are affine in the state variables $X(t)$. As shown by Duffie and Kan (1996) the price of a nominal zero coupon bond has a single payout at a time $t + \tau$ which can be written as[9]

$$P(X_t, t, t+\tau) = exp\big(A(\tau) + B(\tau)' X_t\big) \tag{21}$$

[8] See Merton and Samuelson (1992) and Cochrane (2005) for theoretical justification of the stochastic discount factor.

[9] The corresponding yield is $y_t^\tau = -A(\tau)/\tau - B(\tau)' X_t/\tau$

where $A(\tau)$ and $B(\tau)$ solve a system of ordinary differential equations. The nominal zero coupon bond with duration $\tau = 0$ and payout 1 has a price $P(X_t, t, t) = 1$ which implies $A(0) = 0$ and $B(0) = 0$. The instantaneous (i.e. given the state of the economy) nominal yield of a bond with duration τ can be calculated using the following closed equations

$$B(\tau) = (K' + \Lambda_1' \Sigma_X)^{-1}[exp(-(K' + \Lambda_1' \Sigma_X)\tau) - I_{2\times 2}]R_1 \tag{22}$$

$$A(\tau) = \int_0^\tau \Big(- R_0 - (\Lambda_0' \Sigma_X) B(s) + \frac{1}{2} B'(s) \Sigma_X' \Sigma_X B(s) \Big) ds \tag{23}$$

The derivations of $A(\tau)$ and $B(\tau)$ are provided in Appendix F.

4.2 Estimation Procedure

The model will be estimated using historical stock returns, inflation rates and bond yields with six different maturities, namely the 3-months, 1-years, 2-years, 3-years, 5-years and 10-years bond yields. The data used for the estimation procedure is described in Appendix G. To estimate the model we first describe the bond price dynamics. Subsequently, we derive the discrete version of the model, which is very useful for the estimation procedure as well as for the simulation of the variables. And finally, we present the log likelihood function that needs to be maximized using the simulated annealing procedure.

4.2.1 Bond funds implementing constant duration

Following Shi and Werker (2012) and Bajeux-Besnainou et al. (2003), we introduce funds of bonds with a constant duration to model the bond yields. Assume a pension fund rebalances the bond portfolio permanently to hold the maturity τ constant, i.e. the fund invests only in bonds with maturity τ. The development of the bond portfolio is based on the following price dynamics equation

$$\frac{dP^{F\tau}}{P^{F\tau}} = \Big(R_t + B(\tau)' \Sigma_X' \Lambda_t \Big) dt + B(\tau)' \Sigma_X' dZ_t \tag{24}$$

with $P^{F\tau}$ the price of a bond portfolio with maturity τ. This expression has a clear cut interpretation: the dt term is the nominal instantaneous rate plus the risk exposure $B(\tau)'\Sigma_X'$ multiplied with the price of risk Λ_t. Note $B(0) = 0$ leading to $\frac{dP^{F0}}{P^{F0}} = R_t dt$.

4.2.2 Discretization

Exact discretization is possible by writing the whole model as a multivariate Ornstein-Uhlenbeck proces

$$dY_t = (\Theta_0 + \Theta_1 Y_t)dt + \Sigma_Y dZ_t \tag{25}$$

with

$$Y' = \Big[X \quad ln\Pi \quad lnS \quad lnP^{F0} \quad lnP^{F\tau} \Big]$$

in which X is the vector with the two state variables, Π the price index, S the stock index, P^{F0} the cash wealth index, $P^{F\tau}$ the bond wealth index with a duration τ, and Z the vector with the four independent Brownian motions extended with two zeros for cash and bond equations. Use Itô Doeblin thereom for log inflation

$$d\ln\Pi = \frac{\partial \ln\Pi}{\partial \Pi} d\Pi + \frac{1}{2}\Big(\frac{\partial^2 \ln\Pi}{\partial \Pi^2}\Big)(d\Pi)^2 \tag{26}$$

$$= (\pi_t dt + \sigma_\Pi' dZ_t) - \frac{1}{2}\Big[\pi_t dt + \sigma_\Pi' dZ_t\Big]^2 \tag{27}$$

$$= (\pi_t - \frac{1}{2}\sigma_\Pi' \sigma_\Pi)dt + \sigma_\Pi' dZ_t \tag{28}$$

and log equity

$$d\text{ln}S = \frac{\partial \text{ln}S}{\partial S}dS + \frac{1}{2}\Big(\frac{\partial^2 \text{ln}S}{\partial S^2}\Big)(dS)^2 \tag{29}$$

$$= (R_t + \eta_S)dt + \sigma_S' dZ_t) - \frac{1}{2}\Big[(r_t + \eta_S)dt + \sigma_S' dZ_t\Big]^2 \tag{30}$$

$$= (R_0 + R_1' X_t + \eta_S - \frac{1}{2}\sigma_S'\sigma_S)dt + \sigma_S' dZ_t \tag{31}$$

Log wealth invested in a constant duration fund develops according to

$$d\text{ln}P^{F\tau} = \frac{\partial \text{ln}P^{F\tau}}{\partial \text{ln}P^{F\tau}}dP^{F\tau} + \frac{1}{2}\Big(\frac{\partial^2 \text{ln}P^{F\tau}}{(\partial \text{ln}P^{F\tau})^2}\Big)(dP^{F\tau})^2 \tag{32}$$

$$= \Big(R_t + B^N(\tau)'\Sigma_X'\Lambda_t - \frac{1}{2}B'^N\Sigma_X'\Sigma_X B^N\Big)dt + B^N(\tau)'\Sigma_X' dZ_t \tag{33}$$

This implies for the multivariate Ornstein-Uhlenbeck process

$$d\begin{bmatrix} X \\ ln\Pi \\ lnS \\ lnP^{F0} \\ lnP^{F\tau} \end{bmatrix} = \left(\begin{bmatrix} 0 \\ \sigma_{0\pi} - \frac{1}{2}\sigma_\Pi'\sigma_\Pi \\ R_0 + \eta_S - \frac{1}{2}\sigma_S'\sigma_S \\ R_0 \\ R_0 + B(\tau)'\Sigma_X'\Lambda_0 - \frac{1}{2}B'\Sigma_X'\Sigma_X B \end{bmatrix} + \begin{bmatrix} -K & 0 \\ \sigma_{1\pi}' & 0 \\ R_1' & 0 \\ R_1' & 0 \\ R_1' + B(\tau)'\Sigma_X'\Lambda_1 & 0 \end{bmatrix}\begin{bmatrix} X \\ ln\Pi \\ lnS \\ lnP^{F0} \\ lnP^{F\tau} \end{bmatrix}\right)dt$$

$$+ \begin{bmatrix} \Sigma_X' \\ \sigma_\Pi' \\ \sigma_S' \\ 0 \\ B(\tau)'\Sigma_X' \end{bmatrix} dZ_t \tag{34}$$

After using the eigenvalue decomposition $\Theta_1 = UDU^{-1}$ the exact discretization reads as

$$Y_{t+h} = \mu^{(h)} + \Gamma^{(h)}Y_t + \epsilon_{t+h} \quad \text{and} \quad \epsilon_{t+h} \sim N(0, \Sigma^{(h)}) \tag{35}$$

in which:
(ι.) $\Gamma^{(h)}$ is defined as

$$\Gamma^{(h)} = exp(\Theta_1 h) = Uexp(Dh)U^{-1} \tag{36}$$

whereby the matrix exponential is defined as $exp(A) = I + \sum_{r=1}^{\infty}\frac{1}{r!}A^r$.

($\iota\iota$.) $\mu^{(h)}$ is defined as

$$\mu^{(h)} = UFU^{-1}\Theta_0 \tag{37}$$

where F is a diagonal matrix with elements $F_{ii} = h\alpha(D_{ii}h)$ with $\alpha(x) = \frac{exp(x)-1}{x}$ and $\alpha(0) = 1$.

($\iota\iota\iota$.) $\Sigma^{(h)}$ is defined as

$$V_{ij} = [U^{-1}\Sigma_Y\Sigma_Y'(U^{-1})']_{ij}h\alpha([D_{ii} + D_{jj}]h) \tag{38}$$

These relations are taken from Koijen et al. (2005) and Bergstrom (1984).

4.2.3 Likelihood

Assume, two yields are observed without measurement error. For those yields hold

$$y_t^\tau = \big(-A(\tau) - B(\tau)'X_t\big)/\tau \tag{39}$$

The observations can be used to determine the state vector X, given a set parameters which determine A and B. The other four yields are observed with a measurement error by assumption.

$$y_t^\tau = \big(-A(\tau) - B(\tau)'X_t\big)/\tau + \upsilon_t^\tau \quad \text{and} \quad \upsilon_t' \sim N(0, \Sigma^\tau) \tag{40}$$

with $\upsilon_t' = [\upsilon_t^{\tau 1}, \upsilon_t^{\tau 2}, \upsilon_t^{\tau 3}, \upsilon_t^{\tau 4}]$. Assume no correlation between the measurement errors. This system of measurement equations is extended with the equations from (35) for inflation and equity. The relevant part of the error term extended with zero's is $\tilde{\epsilon}$. The quasi log likelihood

$$\ln L = -0.5\Big(T\ln|\Sigma^\tau| - \sum_{t=1}^{T} \upsilon_t (\Sigma^\tau)^{-1} \upsilon_t'\Big) - 0.5\Big(T\ln|\Sigma| - \sum_{t=1}^{T} \tilde{\epsilon}_t (\Sigma)^{-1} \tilde{\epsilon}_t'\Big) - 0.5T\ln|B| \qquad (41)$$

with $B' = [B(\tau 5), B(\tau 6)]$ is maximized with respect to the parameters using the method of simulated annealing of Goffe et al. (1994) to find the global optimum. The algorithm begins with a random solution, makes a small change to that solution, tests it and accepts the new solution when it is an improvement. However, the simulated annealing procedure also accepts, with a certain probability, worse solutions. By accepting worse solutions, the algorithm avoids being trapped in a local optimum in early iterations and is able to explore globally for better solutions. Duffee (2002) details on the construction of this quasi log likelihood.

4.3 Estimation Results

Table 3 presents the parameter estimates and the standard errors for the KNW model based on data described in Appendix G. The parameters are expressed in annual terms. The two left columns present the results for the updated estimation period 1972-Q4 to 2014-Q4.[10] The four most rights columns show the estimated parameters given in the papers Draper (2014) and Draper (2012), which are presented for comparison reasons. In this section we first briefly summarize the relevant aspects of the estimation results. After that we show several implications of the estimates that are reported in Table 3.

The parameter $\delta_{0\pi}$ represents the long-term average inflation expectations and the estimation result for this parameter is equal to 1.98%. This is in line with the monetary policy of the European Central Bank (ECB), which aims to keep inflation rate below but close to 2.00%. The expected nominal long-term money market rate (R_0) is 1.98%. Note that the unconditional expected inflation and the unconditional nominal interest rate are (almost) equal. Draper (2014) show that the persistence of the aforementioned variables is high and even increased compared to the shorter sample period presented in Draper (2012). For our updated sample, the persistence remains high with first-order autocorrelations equal to 0.91 and 0.89 for the real rate and expected inflation, respectively. The parameter η_S represents the historical risk premium on equities and is equal to 4.20%. The equity risk premium seems the have decreased compared to the risk premium in the second sample (*ii*). Furthermore, we see that the volatility of stock returns ($\sigma_{S(4)}$) is approximately 16.39% and decreased a little just like the equity premium.

In the last two rows of Table 3 we present the log likelihood of the estimated parameters for the samples up to 2013-Q4 and 2014-Q4. Not surprisingly, the estimates based on the updated dataset have a higher log likelihood value for this sample (1972-Q4 to 2014-Q4) compared to the results presented in Draper (2014) and Draper (2012). In addition, these model parameters also outperform the other two sets for the sample up to the fourth quarter of 2013. Therefore we may conclude that the estimation procedure in Draper (2014) has not led to the maximum likelihood. The parameter set is probably a local optimum. As mentioned before, this is very common problem when estimating models with many parameters. Although our estimates resulted in improved likelihoods for both samples, this may also be a local optimum instead of a global optimum.

Draper (2012) shows that the significance of the estimates for the Netherlands is lower than for the US data. Here we find that, in general, the significance levels of the updated estimates are even lower when comparing them with the estimates based on the shorter Dutch sample. According to Draper (2012), the less significant estimates are partly explained by the large persistence in the state variables X1 and X2. We see that first-order autocorrelations for X1 and X2 increased to

[10] We have estimated the model parameters multiple times using the simulated annealing procedure. Although simulated annealing is proven to be very effective at finding a good solution, it still has a probability of getting stuck in a local optimum. In accordance with S. Muns (CPB) we present in Table 3 the estimated model parameters with the highest likelihood.

0.94 and 0.91 with respect to the prior first-order autocorrelations of 0.73 and 0.91. This confirms the partial explanation of Draper (2012). Note also that X1 is now more persistent then X2.

	(i) 1972.4 - 2014.4		(ii) 1972.4 - 2013.4		(iii) 1972.4 - 2011.3	
Parameter	Estimate	Std. error	Estimate	Std. error	Estimate	Std. error
Expected inflation $\pi_t = \delta_{0\pi} + \delta'_{1\pi} X_t$						
$\delta_{0\pi}$	1.98%	(4.05%)	1.81%	(2.79%)	2.24%	(1.45%)
$\delta_{1\pi(1)}$	-0.60%	(0.20%)	-0.63%	(0.10%)	0.49%	(0.27%)
$\delta_{1\pi(2)}$	0.27%	(0.41%)	0.14%	(0.24%)	0.49%	(0.24%)
Nominal interest rate $R_0 + R'_1 X_t$						
R_0	1.98%	(10.40%)	2.40%	(6.06%)	3.70%	(2.77%)
$R_{1(1)}$	-1.44%	(0.38%)	-1.48%	(0.22%)	1.40%	(0.43%)
$R_{1(2)}$	0.56%	(0.97%)	0.53%	(0.56%)	0.82%	(0.68%)
Process real interest rate and expected inflation $dX_t = -KX_t dt + \Sigma'_X dZ_t$						
κ_{11}	0.06	(0.17)	0.08	(0.11)	0.32	(0.23)
κ_{22}	0.32	(0.22)	0.35	(0.18)	0.13	(0.13)
κ_{21}	-0.22	(0.20)	-0.19	(0.08)	-0.23	(0.16)
Realized inflation process $\frac{d\Pi_t}{\Pi_t} = \pi_t dt + \sigma'_\Pi dZ_t$						
$\sigma_{\Pi(1)}$	0.02%	(0.08%)	0.02%	(0.07%)	-0.01%	(0.07%)
$\sigma_{\Pi(2)}$	-0.02%	(0.06%)	-0.01%	(0.06%)	-0.01%	(0.06%)
$\sigma_{\Pi(3)}$	0.61%	(0.04%)	0.61%	(0.04%)	0.60%	(0.04%)
Stock return process $\frac{dS_t}{S_t} = (r_t + \eta_S)dt + \sigma'_S dZ_t$						
η_S	4.20%	(3.77%)	4.52%	(3.73%)	3.52%	(3.88%)
$\sigma_{S(1)}$	-0.54%	(1.44%)	-0.53%	(1.44%)	-0.16%	(1.71%)
$\sigma_{S(2)}$	-0.78%	(1.54%)	-0.76%	(1.54%)	1.01%	(1.61%)
$\sigma_{S(3)}$	-2.23%	(1.46%)	-2.11%	(1.51%)	-2.65%	(1.56%)
$\sigma_{S(4)}$	16.39%	(0.93%)	16.59%	(0.96%)	16.71%	(0.98%)
Prices of risk $\Lambda_t = \Lambda_0 + \Lambda_1 X_t$						
$\Lambda_{0(1)}$	0.187	(0.513)	0.403	(0.333)	-0.271	(0.266)
$\Lambda_{0(2)}$	0.137	(0.624)	0.039	(0.270)	-0.279	(0.238)
$\Lambda_{1(1,1)}$	0.142	(0.184)	0.149	(0.156)	0.167	(0.252)
$\Lambda_{1(1,2)}$	-0.355	(0.037)	-0.381	(0.039)	-0.114	(0.239)
$\Lambda_{1(2,1)}$	0.144	(0.192)	0.089	(0.075)	0.395	(0.246)
$\Lambda_{1(2,2)}$	-0.100	(0.211)	-0.083	(0.129)	-0.126	(0.140)
$LL_{2013.4}$	6549.2		6525.6		6521.9	
$LL_{2014.4}$	6720.4		6696.7		6692.7	

Table 3: Parameters and standard errors of (i) own max. likelihood estimate using quarterly data from 1972-Q4 to 2014-Q4, (ii) max. likelihood estimate 2013 in Draper (2014), and (iii) max. likelihood estimate 2011 in Draper (2012). Standard errors are determined using the outer product gradient estimator of the likelihood function, which is only feasible at a max. likelihood estimate. LL_y refers to the loglikelihood with the sample ending at year y.

Next we look at the prices of risk and implied risk premia. First we find that the unconditional price of risk with respect to real interest rates is higher than for expected inflation risk, i.e. $\Lambda_{0(1)} > \Lambda_{0(2)}$. This is in line with recent literature such as Campbell and Viceira (2001) and Brennan and Xia (2002). Note that $\Lambda_{0(1)}$ is the parameter which is most decisive for the long-term risk premium on bonds. Table 4 shows the risk premia on nominal bonds and the associated volatilities with maturities of 1, 5 and 10 years. With these values we can calculate the Sharpe ratio as well. The risk premium is derived in equation (24) and is equal to $B(\tau)'\Sigma'_X\Lambda_t$. The time-varying price of risk factors (Λ_t) are set equal to their unconditional expectation (Λ_0). In our model the risk premia range from 20 basis points (bps) for a 1-year nominal bond to 209 basis points for a 10-year bond. The Sharpe ratio of the 1-year nominal bond is slightly lower than for the 5-year and 10-year nominal bonds. Comparing our results with Draper (2014) and Draper (2012), the risk premia and Sharpe ratios are decreased considerably. The risk premium for bonds with a maturity of 10 years is even a staggering 102 basis points lower than when data up to the year 2013 is used.

	1972.4 - 2014.4		1972.4 - 2013.4		1972.4 - 2011.3	
Maturities	Risk premium	Volatility	Risk premium	Volatility	Risk premium	Volatility
One-year	0.20%	1.32%	0.52%	1.33%	0.53%	1.37%
Five-year	1.08%	4.89%	1.94%	4.99%	1.79%	5.12%
Ten-year	2.09%	9.01%	3.11%	9.10%	2.69%	9.38%

	1972.4 - 2014.4	1972.4 - 2013.4	1972.4 - 2011.3
Maturities	Sharpe ratio	Sharpe ratio	Sharpe ratio
One-year	0.15	0.39	0.39
Five-year	0.22	0.39	0.35
Ten-year	0.23	0.34	0.29

Table 4: The risk premia and return volatility on one-year, five-year and ten-year nominal bonds using the estimation results of table 3 (Upper table). Sharpe ratios for one-year, five-year and ten-year nominal bonds (Lower table). The price of risk factors equal their unconditional expectation ($X_t = 0_{2\times1}$). The risk premia and volatilities are expressed in annual terms.

In addition to the unconditional risk premium, we are interested in the impact of the time-varying prices of risk on bond risk premia. The time-variation in prices of risk is governed by Λ_1. Figure 7 presents the five-year nominal bond risk premia for a realistic range X. For the updated model parameters, the state variable X1 has a positive effect on the risk premium and X2 has a negative effect. We see also that the bond risk premium is more sensitive to shifts in X2 than in X1. Furthermore, we find that the bond risk premia based on the updated model parameters are more sensitive towards both state variables than the other two models. This can be explained by the high persistence of the state variables, which we mentioned earlier.

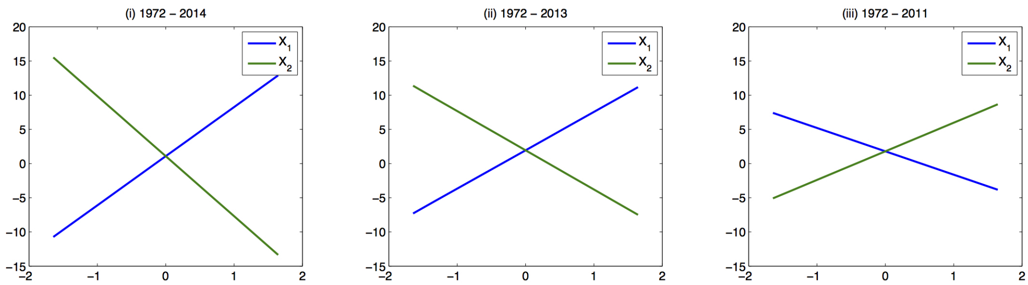

Figure 7: Time-variation in risk premia. Value 5-year nominal bond risk premium for different values of state variables X1 and X2. The horizontal axis depicts the value of the state variable expressed in unconditional standard deviations around their means. The risk premia are expressed in annual terms.

Note that for the shortest sample the risk premia are decreasing in X1 and increasing in X2. This is due to the opposite signs of some of the estimated parameters, see for example $\delta_{1\pi(1)}$, $R_{1(1)}$ and $\Lambda_{0(1)}$ in Table 3. This results in an opposite sign for the value $B(\tau)$, which is derived in equation (22). Draper (2014) shows that model estimates based on the same data can have different signs and still have the same maximum likelihood. This points to indeterminacy of the sign, which can be explained by the sign switch in the state variables. In other words, a new state variable with an opposite sign can lead to the same maximum likelihood with a different sign for some of the model parameters.

Following Sangvinatsos and Wachter (2005) and Koijen et al. (2010), we also present the correlation between stock returns, nominal bonds and the risk premia on nominal bonds in Table 5. In addition, we provide the correlations between the assets classes based on data.

A. KNW model	**Stocks**	**3m Bond**	**1y bond**	**5y Bond**	**10y bond**
Stocks	1.00	-0.02	-0.02	-0.05	-0.05
3m Bond		1.00	0.99	0.79	0.55
1y bond			1.00	0.86	0.64
5y Bond				1.00	0.95
10y bond					1.00
5y risk premium	0.03	0.71	0.62	0.13	-0.20
10y risk premium	0.02	0.76	0.68	0.20	-0.13
B. Data	**Stocks**	**3m Bond**	**1y bond**	**5y Bond**	**10y bond**
Stocks	1.00	-0.15	-0.14	-0.10	-0.09
3m Bond		1.00	0.98	0.91	0.86
1y bond			1.00	0.95	0.90
5y Bond				1.00	0.99
10y bond					1.00

Table 5: Correlation between asset returns and risk premia. Panel A reports the implied correlations between stock returns, 3-months nominal bonds, 1-year nominal bonds, 5-year nominal bonds, 10-year bonds and the risk premia on 5-year nominal bonds and 10-year nominal bonds on the basis of the updated parameter estimates (i) in Table 3. Panel B shows unconditional correlations from the data.

First, we notice that the stock returns and nominal bond yields are (slightly) negatively correlated. This is not consistent with Koijen et al. (2005) and Sangvinatsos and Wachter (2005). These two paper report correlations between stock returns and bond returns with different maturities that vary between 19% and 21%. This corresponds with the correlations based on the US data that they used. Nevertheless, the negative correlation is consistent with the correlations based on our data that vary between -9% and -15%. The low correlation or even lack of correlation may partly explain the less significant coefficients for our model compared to other papers like Koijen et al. (2010). The correlations between the different nominal bond returns are strongly positive obviously. Though we see stronger correlations between short-term bonds and long-term bonds for the data compared to the model. Furthermore, we notice that the long-term bonds are negatively correlated with bond risk premia, which is comparable to the results in Koijen et al. (2010).

4.4 Fit of the Model

Koijen et al. (2010) conclude that their tractable two-factor model provides an overall good fit to the data and that the model does a reasonable job of fitting the cross-sectional moments of bond yields. However the estimates presented in Draper (2014) show different results than Koijen et al. (2010), including less significant estimates and low (negative) to zero correlations between stock returns and the term structure. And as Campbell and Viceira (2001) also emphasized, "It is not guaranteed that models such as this one fit time series". So, we cannot immediately assume that the KNW-model is well suited for replicating the Dutch financial market. Following Koijen et al. (2010), we examine the fit of the model parameters and compare results with previous estimations. In this section, we first examine the average yield curve and the volatility of bond returns. We then turn to the predictability of bond returns to check whether the model replicates the empirical findings of Campbell and Shiller (1991).

4.4.1 Simulations term structure

Following Dai and Singleton (2002) and Sangvinatsos and Wachter (2005), we illustrate the average and the volatility of bond yields implied by the estimated model parameters. Thus, we first simulate 5,000 sample paths for different bond yields of the same length as our data. Then we compute the average yield for each of the sample paths. Afterwards, we plot the averages and the 95% confidence intervals of these sample means. And lastly, we plot the average yield based on the data to test whether the model fits the data. We repeat these steps for the volatility of the various bond yields in our sample. In addition to the six maturities used in the estimation procedure (3m, 1y, 2y, 3y, 5y, and 10y), we also simulate all other maturities available in our dataset (see Appendix G). Hence, we also include long-term maturities (20-60 years) critical for the feasibility test, as these are needed to discount long-term pension-fund liabilities.

The implied average bond yield and the volatility of bond yields are presented in Figure 8. We also display the implied sample values for the estimates presented in Draper (2012) and Draper (2014). We see that the KNW model provides a reasonably good fit to the Dutch data. All implied bond yield averages are comfortably within the 95% confidence interval. It is remarkable that the model generates long-term bond yields that fit the data very well. In contrast, the implied means for the model parameters presented in Draper (2012) and Draper (2014) are closer to the data for short-term bonds. This observation can be explained by the data sample on which the model parameters have been estimated. According to this data, bond rates have been decreasing over time since 1980. For long-term bonds, we have no data available before 1980. Also, the ECB's monetary policy in response to the European credit crisis has led to extremely low interest rates in recent years. And, in contrast to Draper (2012), our updated estimates take into account these recent low interest rates.

The implied standard deviations are very close to those found in the data. In addition, the smaller confidence interval shows that the volatilities are estimated more precisely than are the averages. Following Koijen et al. (2010), we can conclude that the KNW model estimated with Dutch data is able to match the cross-sectional moments of bond yields.

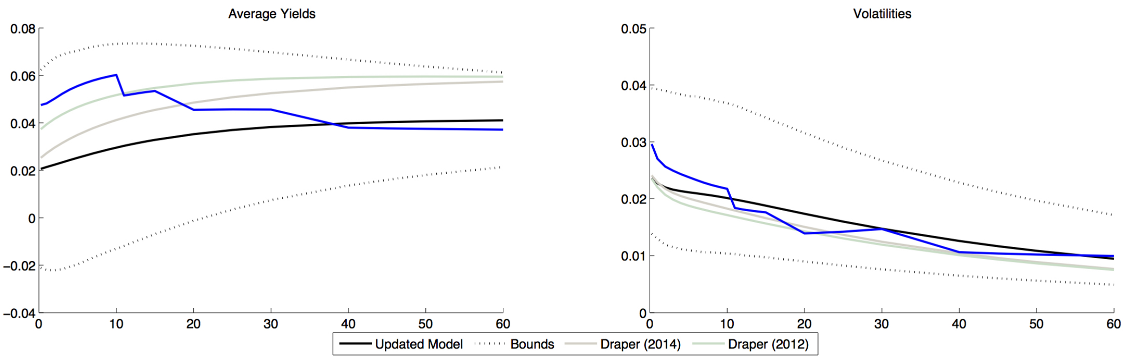

Figure 8: Average yields and volatility of yields implied by KNW model and data. Left graph displays average yields for different maturities and right graph displays volatility of yields for the different maturities. Black lines (solid, dotted) represent the mean and 95% confidence interval for our updated model (estimates (i) in Table 3). The light grey and green lines display the average moments for estimates of Draper (2014) and Draper (2012). Blue lines represent the data sample.

4.4.2 Campbell-Shiller Regressions

One technique for assessing the goodness of fit of a term-structure model is to examine whether the features of the model match the features described in the existing literature. In order to understand bond returns and their yields, researchers developed several hypotheses that are associated with the term 'expectations hypothesis' (EH). The EH includes numerous statements that link the yields, bond returns, and forward rates of different maturities and periods. Sangvinatsos (2008) summarized the main statements of the EH, including: (1) the expected excess returns are constant over time, (2) yield term premia are constant, and (3) forward term premia are constant over time.

Macaulay et al. (1938) wrote one of the first papers to question the accuracy of this hypothesis. Since then, the EH has been further examined many times, and almost all papers statistically reject the expectations theory. In addition, Campbell and Shiller (1991) showed that, for almost any pair of maturities between one month and ten years, a high yield spread between a longer-term and a shorter-term interest rate forecasts a declining yield on the longer-term bond over the life of the shorter-term bond. In other words, when the spread is high, the long-term bond rate tends to fall. The negative relationship between bond returns and the slope of the yield curve is inconsistent with the EH.

Dai and Singleton (2002) focused on several key stylized facts about the excess returns on bonds to test different dynamic term-structure models. For testing these models, their paper considered the following empirical observations:

(i) Linear projections of $R_{t+1}^{n-1} - R_t^n$ onto $\frac{1}{n-1}(R_t^n - r_t)$ give negative, often statistically significant slope coefficients β_1.

(ii) Moreover, β_1 typically becomes increasingly negative with maturity.

(iii) On average, the term structure of treasury bond yields is upward sloping

In the previous section, we have already seen that the model generates, on average, upward sloping term structures. To test whether a term-structure model replicates the other empirical findings of Dai and Singleton (2002), we make use of the following Campbell-Shiller regression (n>m):

$$y_{t+m}^{n-m} - y_t^n = \beta_0 + \beta_1 \frac{m(y_t^n - y_t^m)}{n-m} + \epsilon_{t+m} \tag{42}$$

a more general form of the regressions in Dai and Singleton (2002). Under the EH, the coefficient β_1 would be equal to 1. However, the estimates coefficients in Campbell and Shiller (1991) and Dai and Singleton (2002) not only differ significantly from 1 but are also often significantly negative, particularly for large n. These results indicate that the EH fails more dramatically for long-term bonds. It is important to note the empirical evidence against the EH is primarily based on data from the United States. Sangvinatsos (2008) reviewed studies that examine the validity of the EH outside of the United States. The paper concluded that for bond yields outside of the United States, while the Campbell-Shiller coefficients are less than zero, they are less negative as compared to the results based on U.S. data.

Following Koijen et al. (2010), we simulate 5,000 sample paths of the same length as our sample and run the predictive regression as in Equation (42). For n, we assume the same maturities as in Koijen et al. (2010) and Sangvinatsos and Wachter (2005). For m, we use three months, which means that we regress the quarterly changes in yield ($y(t,s)$ - $y(t+\frac{1}{4},s)$) on the spread between the n-year bond and the 3-month bond, scaled by $1/(4 \times n - 1)$. Next, for each of the samples, we compute the regression coefficient. In addition, we calculate the regression coefficients for the data and compare them.

Figure 9 shows that our term-structure model is able to match the empirical observations described by Dai and Singleton (2002). The model captures the negative coefficients (i) and the downward sloping maturity structure of the coefficients (ii). In addition, the regression coefficients based on the data fall within the 95% confidence bands implied by the model. Furthermore, the confidence interval shows that the model rejects the EH, as the coefficients differ significantly from 1. However, the coefficients are not statistically significant negative, which is consistent with papers based on non-U.S. data.

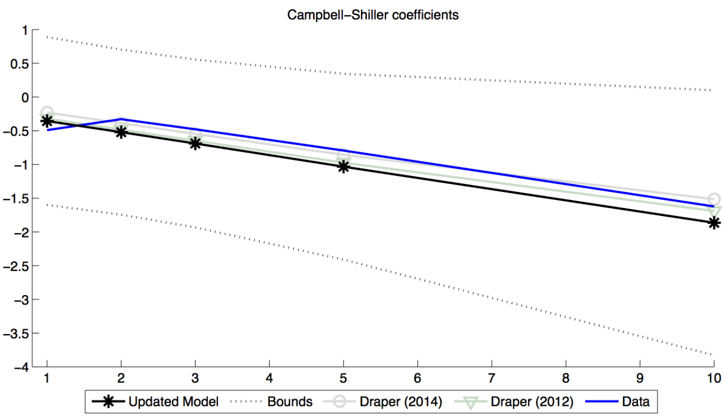

Figure 9: Campbell-Shiller regression coefficients. Solid black lines represent the average coefficients and 95% confidence interval for the updated model. Green dashed and red dotted lines represent the coefficients implied by estimates of Draper (2014) and Draper (2012). The circles (stars, plusses) correspond to the maturities used in estimation. The blue line represents the average coefficients implied by the data.

4.5 Summary of Calibration and Actualization

In line with the recommendations of the Commission Parameters, the DNB uses the estimated model parameters of Draper (2014) as the basis for generating the scenario set. In Langejan et al. (2014), the Commission Parameters also recommend a number of modifications to certain model parameters to render the scenario set's expected values more consistent with the commission's expectations. Despite extensive research by the Commission Parameters, the DNB has not adopted all of its recommendations and applies several other assumptions for generating the scenarios. In Appendix H., we examine the calibrated model parameters using both the assumptions made by the Commission Parameters and the assumptions made by the DNB and then comparing the results. Not surprisingly, we see clear differences in the results of the model calibrated according to the Commission Parameters approach and the model calibrated according to the DNB approach.

For the DNB scenario set, we measured a 5% geometric average return-on-equity over the lifetime of the scenario analysis. However, the Commission Parameters recommends an expected geometric return of 7% on listed equity. So, the DNB scenario set is very conservative with respect to equity returns. The reverse is true for the inflation. The Commission Parameters considers 2% to be a realistic figure for expected inflation, given the ECB's inflation target, and the Commission Parameters assumes inflation will reach that level within five years. The DNB's expectations regarding average inflation are the same, but in the scenarios we see that it takes 30 to 40 years until the inflation level is back at 2%.

Not only did the regulators incorporate their own expectations into certain variables but they also want the scenario set stay current. To ensure that the interest rates in the scenario set are sufficiently correlated with prevailing economic conditions, the model must be updated quarterly. The Commission Parameters proposed a method based up the prevailing yield curve and the 10-year forward rate curve. The DNB, however, has chosen another method for updating the model, in which only the prevailing yield curve is employed. Appendix I. describes both actualization methods and compares the results.

The assumptions made for the calibration and actualization of the model parameters have a large impact on a pension fund's future financial position and therefore also on the results of the feasibility test. In a stochastic analysis over 60 years, any small change can have a sizeable influence on final results. If the outlook is overly optimistic, there is a risk that necessary recovery measures could be delayed for too long. In contrast, an excessively conservative outlook could lead to unnecessary recovery measures, which could be detrimental to the purchasing power of some generations. Therefore, the motivation behind the calibration and actualization of the model parameters should be clear.

5 Alternative Models

As mentioned before, the pension industry has criticized the KWN model. First of all, the KNW model only utilizes one equity asset class rather than multiple classes. Second, the scenarios generated for the term structure must be used for discounting liabilities, as well as for the valuation of the fixed-income portfolio and the interest-rate hedge. This creates a discrepancy in the valuation of these elements, as normally happens outside the context of the feasibility test. The DNB recommends scaling the fixed-income portfolio and the hedge so that values correspond to the real book value. According to critics, such measures would have an unduly strong influence on the underlying interest-rate sensitivity and on funding-ratio development. Third, the initial term structure in the scenario set takes the UFR component into account. However, it does not precisely match the prevailing term structure for discounting pension liabilities that the DNB publishes. Fourth, the model is only estimated using six short-term bond yields, and it does not focus on long-term yields, which are very important for pension funds.

And last, but certainly not least, the KNW model can lead to extremely low or even negative (long-term) interest rates for some scenarios. Negative interest rates occur, because the model does not have any restrictions concerning the term structure. As current low interest rates serve as input parameters, the probability on negative interest rate scenarios is relatively large. In theory, nominal interest rates can become negative due to strong deflationary expectations, negative real interest rates, or a combination of the two factors. Nevertheless, in practice there are no examples of negative long-term rates (>5yr). In addition, the current recovery plan scheme excludes negative interest rates. For a clear regulatory regime, it is important to match assumptions for all legislative instruments, including the feasibility test and the recovery plan scheme.

This chapter compares results from the KNW model with results from three alternative models to assess whether it is acceptable to use the KNW model as the ESG. In other words, we examine how different models and model assumptions affect feasibility test outcomes. Hereby, we focus on the pension result, which is the feasibility test's primary metric that pension funds use in communication with their members.

For the construction of alternative ESGs, we first need to select the asset classes that will be modeled. In addition, we need to choose which stochastic models to use and how to calibrate them. Many variables play a role in pension funds' risk management. However, to limit the complexity of the alternative models, we focus on modeling interest rates. The term structure is the most important risk factor, because it has greatest impact on a pension fund's (future) financial position. So, it is necessary for our alternative model to incorporate term-structure features as well as possible. Section 2.5 provides some general background information on different types of interest rate models. We propose three different interest-rate models that are based on different assumptions and that are estimated using various techniques. In addition, we model the equity returns and the inflation level.

In the previous section, we have seen how difficult it is to find robust estimates for the KNW model. Also for our alternative models it is important to find a calibration procedure that fits our objective. There are three basic approaches to calibrate interest rate models. The first approach corresponds to calibration over the space dimension. Model parameters are obtained by calibrating the model to current market prices. In other words, it only considers various spot rate maturities at a certain point in time. This approach is typically used for pricing options and other derivatives in order to avoid arbitrage opportunities. The second approach calibrates parameters over the time dimension. Model parameters are calibrated using time-series for a single maturity, for example historical short rates. This approach is interesting for risk management applications or long-term scenario analysis as it takes past behavior into account. The third approach considers a combined space-time dimension and uses historical spot rates for various maturities as input. A possible method is to express the interest rate model as a state space formulation and then to use the Kalman Filter to estimate the parameters. In addition to the three approaches, users can also just specify their own views on how the market will behave in the future.

In the feasibility test, the approaches are combined. The KNW model is first estimated using historical data, after which some model parameters are calibrated to the prevailing yield curve and the expectations of the Commission Parameters. For the alternative models, our focus is primarily on calibration approaches based on historical yield data, but we also try to fit the current term structure as good as possible.

The remainder of this section is structured as follows. In Section 5.1, we discuss the methodology of the three alternative interest rate models and assess the fit of the model. For each model we also discuss the advantages and the disadvantages. The models for equity returns and inflation levels over time are presented in Section 5.2. And finally, in Section 5.3, the pension results based on the three alternative models and the KNW model are presented and compared.

5.1 Alternative Interest Rate Models

For our alternative model, we seek an interest rate model that produces realistic yield-curve movements, does not allow for negative interest rates, and matches the initial yield curve to the observed yield curve. To that end, we examine three different types of interest rate models: (1) the CIR model (one-factor equilibrium model), (2) the G2++ model (a two-factor no-arbitrage model), and (3) the Libor Market Model. Sections 5.1.1 to 5.1.3 describe the methodologies of these alternative models. In addition, we also discuss some of the strengths and weaknesses of each model and explain how the parameters are estimated using historical data.

To assess the alternative interest rate models, we test whether simulation results correspond with several stylized facts. Similar to Section 4.4, we use 5,000 simulation paths with a 60-year horizon to check if the average yield curve is upward sloping and concave, as well as if the short end of the yield curve is more volatile than the long end. In addition, to gauge whether the models are suited for a long-term risk-management context, we also present several scenario paths for each model. We check if the models can produce a variety of shapes through time, including upward sloping, downward sloping, humped, and inverted humped.

We want to emphasize that we consider that this paper's main contribution is to review various models that can be used for economic scenario simulation purposes. The formulation and derivation of the different models - which is, after all described in other, more in-depth papers is not our primary focus. Therefore, this paper provides only the necessary details for the implementation of these models, and their theoretical background is not touched upon. For the theoretical justifications and calibration procedures of each model, we refer to the original papers and books.

5.1.1 Cox-Ingersoll-Ross model

For our first alternative interest rate model we use the classical one-factor equilibrium model proposed by Cox et al. (1985). The CIR model involves a mean-reverting process with volatility proportional to the square root of the current interest rate level. The continuous model can be expressed as

$$dr(t) = \kappa(\theta - r)dt + \sigma\sqrt{r}dW(t) \tag{43}$$

where r is the instantaneous short rate, κ is the speed of mean reversion, θ is the long-term average to which r tends to revert over time, σ is a volatility parameter and dW is a simple Brownian motion process.

The drift term ensures that the interest rates are pulled back towards the long-term average over time. In addition, the inclusion of the square root in the volatility term ensures that the volatility is high when interest rates are high and vice versa. This is in line with empirical evidence, see for example Ahlgrim et al. (2004). An additional advantage of relating the volatility to interest rates is that negative interest rates are ruled out. When the short rate declines, the volatility term approaches zero. In this case the short rate process will only be affected by the drift term, causing the short rate to revert to the mean.

Cox et al. (1985) provide a closed-form solution for the price of zero-coupon bonds $P(t, T)$, whereby bond prices depend on the short rate as follows

$$P(t, T) = A(t, T)e^{-B(t,T)r} \tag{44}$$

where

$$A(t,T) = [\frac{2\gamma e^{(\kappa+\gamma)(T-t)/2}}{(\gamma+\kappa)(e^{\gamma(T-t)}-1)+2\gamma}]^{\frac{2\kappa\theta}{\sigma}^2} \tag{45}$$

and

$$B(t,T) = [\frac{(2e^{\gamma(T-t)}-1)}{(\gamma+\kappa)(e^{\gamma(T-t)}-1)+2\gamma)}] \tag{46}$$

where $\gamma = \sqrt{\kappa^2 + 2\sigma^2}$. This shows that the bond price is an affine function of the short rate, which serves as the only underlying uncertainty. The bond yields are perfectly correlated, which implies that all rates move in the same direction. However in reality it is quite common to see interest rates for different maturities move in different directions. Nonetheless, the shape of the yield curve can change over time because perfect correlation does not imply changes of the same amount.

There are two basic approaches to estimate the parameters of the CIR model, namely the cross-section approach and the time-series approach. In the cross-section approach, the model is calibrated using only the yield curve at a certain point in time. Disadvantage is that this approach ignores all historical data. The time-series approach captures the dynamics of the short rate over time. Disadvantage of this approach is that it ignores the cross-sectional information. Because our goal is to construct a scenario set over a long horizon, we choose the latter approach.

The parameters in Equation (43) are estimated using the maximum likelihood estimation method based on the historical 3-month interest rates that are described in Appendix G. The estimated parameter values for the CIR model are

$$\kappa = 0.062, \quad \theta = 0.014, \quad \sigma = 0.058$$

For the initial short rate we use the 3-month interest rate of the first quarter in 2015, which is equal to 0.078%. Figure 10 shows that the average yield curve implied by the CIR model is almost flat and lower than the average yield curve based on observed data. The volatility decreases with maturity, which is line with the general trend, but is underestimated for most maturities. The yield-curve graphs for the CIR model in Figure 13 show that upward and downward sloping shapes occur. However, the variation in shapes is limited because all yields (gradually) converge to the mean reversion factor. These results indicate that the CIR model is not very suitable for a long-term risk management context.

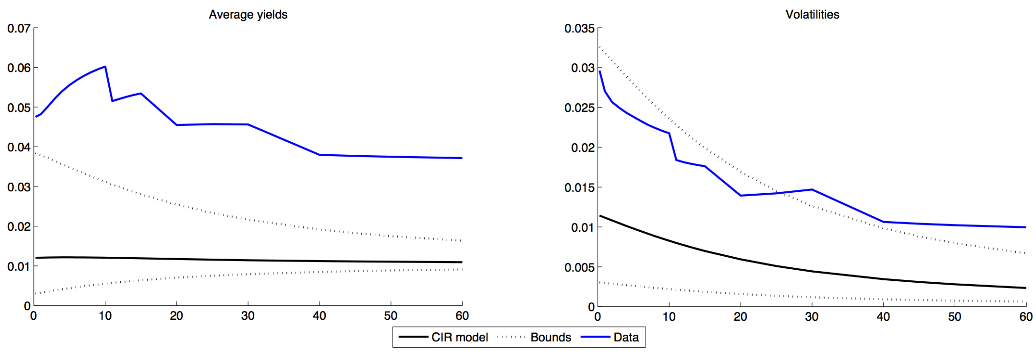

Figure 10: Average yields and volatilities with 95% confidence interval implied by the CIR model and the corresponding moments of the historical yield data.

5.1.2 G2++ model

For our second alternative interest rate model we use the two-factor additive Gaussian model, referred to as G2++ by Brigo and Mercurio (2007). The G2++ model is a popular short rate model due to a number of features that make it attractive for analysis and implementation purposes. For example, it gives an analytical expression for bond prices. In addition, the initial yield curve is consistent with the current term structure because of its no-arbitrage approach. On the other hand, it has the unpleasant feature of theoretical possibility of negative rates. Interesting to note is that the G2++ model is completely equivalent to the two-factor Hull-White model following a coordinate transformation, see Brigo and Mercurio (2007). It is easier to interpret the parameters of the Hull-White model but the derivation and estimation of the model is more complex than for the G2++ model.

The dynamic equations for the G2++ model are formulated as follows

$$\begin{aligned} r(t) &= x(t) + y(t) + \delta(t) &(47)\\ dx(t) &= -\alpha x(t)dt + \gamma dW_1(t) &(48)\\ dy(t) &= -\beta y(t)dt + \eta dW_2(t) &(49)\\ \delta(t) &= f(0,t) + \frac{\gamma^2}{2\alpha^2}(1-e^{-2\alpha t})^2 + \frac{\eta^2}{2\beta^2}(1-e^{-2\beta t})^2 + \kappa\frac{\gamma\eta}{\alpha\beta}(1-e^{-\alpha t})(1-e^{-\beta t}) &(50) \end{aligned}$$

with

$$dW_1(t)dW_2(t) = \kappa dt, \quad x(0) = 0, \quad y(0) = 0 \tag{51}$$

where $x(t)$ and $y(t)$ represent the two stochastic dynamics and $\delta(t)$ the deterministic function. α and β are constants reflecting the rate of mean reversion, γ and η are the volatility constants of the factors x and y, respectively. κ denotes the correlation of the 2-dimensional Brownian motion. $f(0,t)$ is the instantaneous forward rate, which can be defined in terms of the spot rate curve $R(0,t)$ as $f(0,t) = t\frac{d}{dt}R(0,t) + R(0,t)$. This shows that $\delta(t)$ is defined to match the observed current term structure.

Park (2004) provides the closed-form formula for bond prices. The price at time t of a zero-coupon bond maturing at time T and with unit face value is

$$\begin{aligned} P(t,T) &= \frac{P(0,T)}{P(0,t)} exp\{A(t,T)\} &(52)\\ A(t,T) &= \frac{1}{2}\big[V(t,T) - V(0,T) + V(0,t)\big] - \frac{1-e^{-\alpha(T-t)}}{\alpha}x(t) - \frac{1-e^{-\beta(T-t)}}{\beta}y(t) &(53) \end{aligned}$$

where

$$\begin{aligned} V(t,T) = {} & \frac{\gamma^2}{\alpha^2}\big[(T-t) + \frac{2}{\alpha}e^{-\alpha(T-t)} - \frac{1}{2\alpha}e^{-2\alpha(T-t)} - \frac{3}{2\alpha}\big] \\ & + \frac{\eta^2}{\beta^2}\big[(T-t) + \frac{2}{\beta}e^{-\beta(T-t)} - \frac{1}{2\beta}e^{-2\beta(T-t)} - \frac{3}{2\beta}\big] \\ & + 2\kappa\frac{\gamma\eta}{\alpha\beta}\big[(T-t) + \frac{e^{-\alpha(T-t)}-1}{\alpha} + \frac{e^{-\beta(T-t)}-1}{\beta} - \frac{e^{-(\alpha+\beta)(T-t)}-1}{\alpha+\beta}\big] \end{aligned} \tag{54}$$

Park (2004) also describes the discrete version of the model for generating forward paths and for the calibration procedure. For calibration of the G2++ model, Park (2004) describes two calibration methods, one based on volatility data and one using historical yield data. We follow the latter approach, where the calibration problem can be framed as a Kalman filtering problem with suitable parameter update law. In this approach the parameters are calibrated over the space dimension as well as over the time dimension, i.e. the parameters are based on historical spot rate time series for various maturities. In Appendix J we provide a brief description of the calibration via Kalman Filtering. For a detailed description of the Kalman filtering algorithm and the quasi-maximum likelihood function we refer to Park (2004).

For the estimation procedure we use the historical interest rates for all maturities available, ranging from the 3-month interest rate up to the 60-year interest rate. Note that the long-term interest rates are only available from Q3-2001. For the initial yield curve we use the DNB curve in the first quarter of 2015 as input of the model, which ensures that the model is consistent with the prevailing term structure including the UFR component. The estimated parameter values for the G2++ model are

$$\alpha = 0.0100, \quad \gamma = 0.0054, \quad \beta = 0.2210, \quad \eta = 0.0115, \quad \kappa = -0.3178$$

Figure 11 shows that the average yield curve is upward sloping and that the volatility decreases with maturity. The model replicates the actual data reasonably well as the observed moments fall within the confidence intervals. In addition, the yield curve graphs in Figure 13 show that this two-factor model provides a richer pattern of yield curve movements than the previously considered one-factor model. The interest rate curves are not as smooth and have many different shapes, reflecting actual observed term structures. Another good feature is that the implied initial yield curve exactly matches the current term structure. Unfortunately, the model generates a significant number of negative interest rates throughout the 60-year simulation period. This is partly due to the current low interest rate environment. Despite of this, the two-factor model does a good job at reproducing the stylized facts of the yield data.

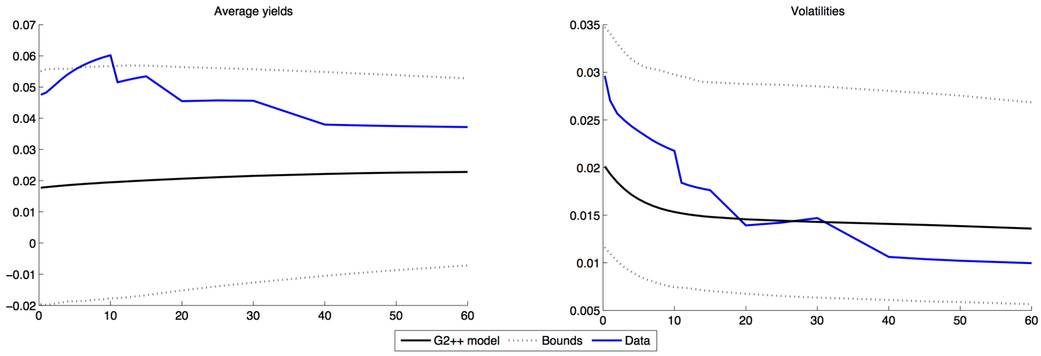

Figure 11: Average yields and volatilities with 95% confidence interval implied by the G2++ model and the corresponding moments of the historical yield data.

5.1.3 LIBOR Market Model

For our final alternative interest rate model we use the LIBOR Market Model (LMM). The LMM is based on evolving LIBOR market forward rates. Unlike the short rate models, it captures the dynamics of the entire yield curve using a set of market-observable forward rates. Other benefits are that the model ensures positive interest rates and exactly matches the current yield curve. Disadvantage is that LMM is based on a risk neutral context.

Following Brigo and Mercurio (2007), we model the forward rate dynamics using the lognormal LMM. The evolution of each forward rate is described by the stochastic differential equation

$$\frac{dF_i(t)}{F_i} = -\mu_i dt + \sigma_i(t) dW_i \tag{55}$$

where dW is an N-dimensional geometric Brownian motion with

$$dW_i(t) dW_j(t) = \rho_{ij} dt \tag{56}$$

where ρ_{ij} is the instantaneous correlation between the forward rates.

The LMM relates the drifts of the forward rates based on no-arbitrage arguments. Specifically, under the Spot LIBOR measure, the drifts are expressed as the following:

$$\mu_i(t) = -\sigma_i(t) \sum_{j=q(t)}^{i} \frac{\tau_j \rho_{ij} \sigma_j(t) F_j(t)}{1 + \tau_j F_j(t)} \tag{57}$$

where τ_i is the time fraction associated with the ith forward rate. $q(t)$ is an index function defined by the relation $T_{q(t)-1} < t < T_{q(t)}$ and the Spot LIBOR numeraire is defined as the following:

$$B(t) = P(t, T_{q(t)}) \prod_{n=0}^{q(t)-1} (1 + \tau_n F_n(T_n)) \tag{58}$$

Note that the drift is based on other forward rates and their instantaneous volatilities and correlations. Hence the evolution of the forward rates depends completely on the volatilities and the correlations. For this paper we assume that both are based on deterministic functions as follows

$$\sigma_i(t) = \phi_i(a(T_i - t) + b)e^{c(T_i - t)} + d \tag{59}$$

$$\rho_{i,j} = e^{-\beta|i-j|} \tag{60}$$

where ϕ adjusts the curve to match the volatility for the i^{th} forward rate.

Next we need to calibrate the parameters of both functions. There are three main approaches for estimating the volatilities and correlations: they can be estimated from historical data; they can be obtained using current market data, typically swaptions or caps/floors; or they can be based on own expectations of the future evolution. Also a combination of these approaches can be applied.

In this paper we estimate a, b, c, d and β using historical data. In order to derive volatilities and correlation from historical data, we first need to derive the yield curves from the data. Here we estimate the full historical yield curves using the Nelson-Siegel model based on the data described in Appendix G. This allows use to calculate the quarterly log-returns of one-year forward rates ($\tau = 1$), which we use to estimate the volatilities and the correlation matrix. We specify ϕ_i so that the volatility declines for higher maturities. The estimated parameter values for the LMM model are

$$a = 0.0081, \quad b = 0.0030, \quad c = 0.4277, \quad d = 0.0211, \quad \beta = 0.0086$$

Figure 12 shows that the average yield curve is upward sloping for the first 30-year interest rates. Subsequently, the average yield declines for higher maturities. The volatility of the yields simulated by the LMM model decrease for higher maturities. However, the volatilities based on the observed data are higher for most maturities. The evolutions of three yield curve scenarios presented in Figure 13 show us that the shape of the current yield curve strongly influences the behavior of future interest rates. As the initial yield curve is strongly upward sloping, the short-term interest rates tend to increase over time. The simulated yield curves are not unrealistic but for risk management purposes there is too little variation in the shape of the curves.

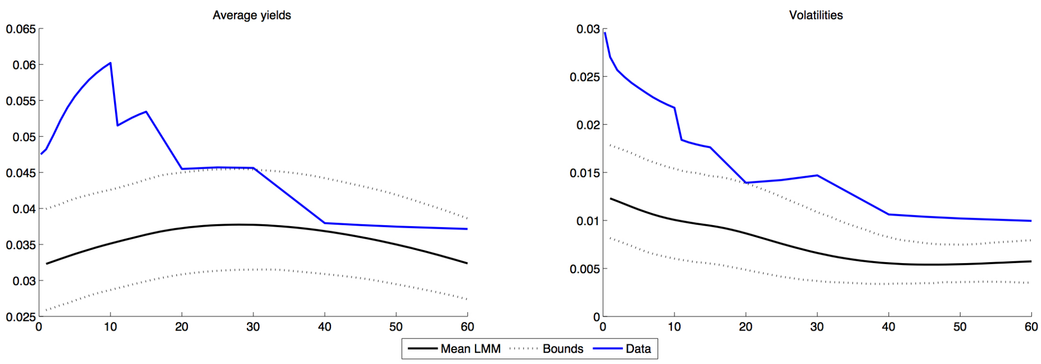

Figure 12: Average yields and volatilities with 95% confidence interval implied by the Libor Market Model and the corresponding moments of the historical yield data.

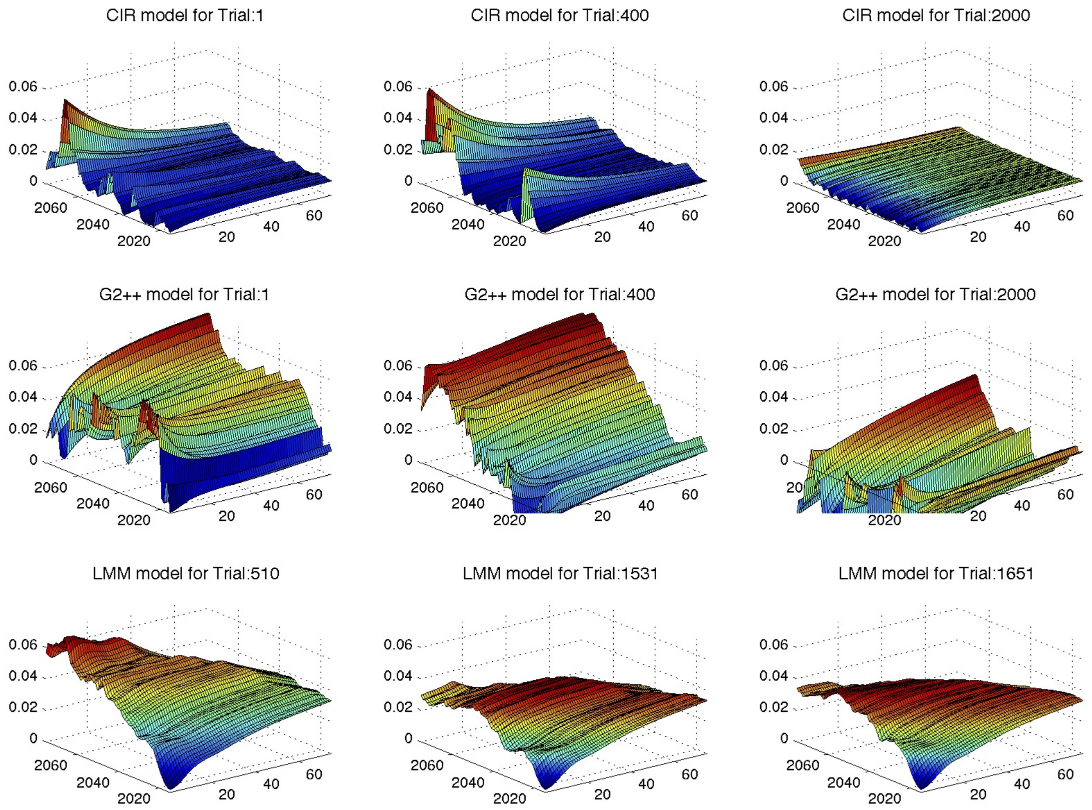

Figure 13: Three scenarios for the evolution of the yield curve with respect to the CIR model (upper graphs), the G2++ model (middle graphs) and LMM model (bottom graphs)

5.2 Stock returns and Inflation

The interest rate models, discussed in the previous section, provide only a part of the alternative scenariosets that we need for the sensitivity analysis. To complete the economic scenario generator we need to model the stock returns and the inflation as well. For our alternative models we estimate and simulate the interest rates, inflation and stock returns separately. In other words, we do not incorporate any (auto)-correlation between the three variables. Reason is that correlations significantly complicate the model. It is difficult to assign correlations to the multiple processes in the interest rate models G2++ and LMM. Another problem is that several papers, including Van den Goorbergh et al. (2011), show that there is no clear relation between short rates, price inflation and stocks. It is difficult to estimate the correlations, because they are not constant over time. Therefore we will model the economic scenario generator without any correlations between the defined processes.

In the literature, among the papers of Baillie et al. (1996) and Lee and Wu (2001), it has been shown that inflation rates follow a mean reverting process. Therefore we propose the Ornstein-Uhlenbeck model to simulate the inflation rates as this model exhibits mean reverting behaviour

$$dI_t = \lambda(\mu - I_t)dt + \sigma_I dW_t \tag{61}$$

where $\lambda(\mu - I_t)$ is the mean reverting instantaneous drift with λ representing the mean reversion speed and μ the equilibrium rate. σ_I represents the standard devation of the inflation rate. To estimate the Ornstein-Uhlenbeck model the discrete form of the process is derived and subsequently the ordinary least squares (OLS) method is used to estimate the parameter of the inflation model. The estimated parameter values for the inflation model are

$$\lambda = 0.2564, \quad \mu = 0.0219, \quad \sigma_I = 0.0124$$

To be consistent with the formulated expectations of the Commission Parameters we modify μ to get an inflation expectation of 2%.

To simulate the overall equity returns in our alternative economic scenario generator we use of the Geometric Brownian Motion model (GBM). The GBM is a well established and widely used descriptive model for equity return and is described in many papers and books, for example in Baxter and Rennie (1996) and Mitra et al. (2009). With GBM, the price of a stock S_t follows the stochastic differential equation:

$$dS_t = \mu S_t dt + \sigma S_t dW_t \tag{62}$$

where μ and σ are constants. Assuming an initial price S_0 then equation (62) has the analytical solution

$$S_t = S_0 exp\big[(\mu - \frac{\sigma^2}{2})t + \sigma W_t\big] \tag{63}$$

which shows that the asset price in the GBM follows a log-normal distribution, while the logarithmic returns $log(S_{t+\Delta t}/S_t)$ are normally distributed. We use the method of maximum likelihood estimation to find the best fit to the historical data. The estimated parameter values for the equity return model are

$$\mu = 0.0762, \quad \sigma = 0.1626$$

To be consistent with the formulated expectations of the Commission Parameters we modify μ to get expected equity returns equal to 7%.

5.3 Results

For each of the three interest rate models, we generated 2,000 scenarios with a 60-year horizon, which is similar to the feasibility test. In addition, we generated 2,000 scenarios for the equity model and the inflation model that will be combined with each interest rate model. For all three ESG models, we simulated the development of an average Dutch pension fund and subsequently calculated the pension result. Regarding pension fund characteristics, we applied the same base-case assumptions discussed in Section 3.3. For the analysis, we assumed that the fund invested 50% in stocks and 50% in fixed income (10-year zero-coupon bonds). In addition, 40% of the interest-rate risk was hedged. The initial funding ratio was set at 105%. Figure 14 displays pension results on fund level (top figure), as well as pension results for the initial age cohorts of 25 and 75 (lower figures) for the KNW model and the three alternative models. The graphs show the 5th percentile pension result, the median pension result, and the difference between these two percentiles. We see some interesting differences between the results of the KNW model and the three alternative models.

On fund level, the KNW model (83%) leads to an expected pension result similar to that predicted by the G2++ model (85%), and lies well within the range of the CIR model (78%) and the LMM model (89%). The CIR model leads to overall lower pension results, whereas the LMM produces higher pension results in general. The KNW model shows a relatively small bandwidth between the median and the 5th percentile, which indicates a more conservative view of the future than the three alternative models. In terms of pension results on generation level, we see that the results across the models differ the most for the younger generations. This disparity shows that it is not evident to assume that the KNW model is appropriate for creating scenarios for the long-term analysis required by law.

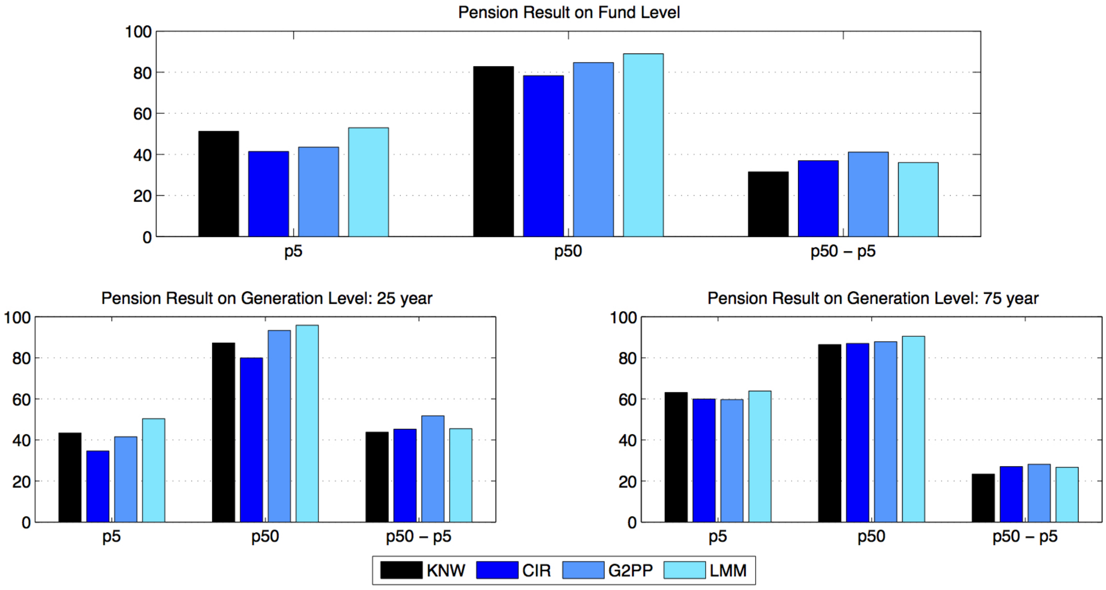

Figure 14: Pension results using base-case assumptions. Figures show various percentiles (P5, P50, and P50 P5) of the pension result on fund level and on generation level (25-year and 75-year initial age cohorts) for the KNW model (black), the CIR model (dark blue), the G2++ model (blue), and the LMM model (light blue).

In addition to these qualitative assessments, we want to support our findings with statistical proof. We can use the Kolmogorov-Smirnov (K-S) test to check whether two datasets in this case, the output of the KNW model and the output of one of the alternative models are significantly different. This test does not depend on the distribution of the underlying data, and therefore it is not as sensitive as tests based on specific distributions. The K-S test metric 'D' is the maximum vertical difference between the cumulative distribution of the KNW model and cumulative distribution of each alternative model. For all three alternative models, the K-S test rejects the hypotheses that their outcomes regarding pension results have a similar distribution as the outcomes of the KNW model.

Another technique for examining whether pension results are significantly different from each other is to use the Wilcoxon signed-rank test. The test examines the null hypothesis, which states that the median of the difference between both samples is zero. A major advantage of the Wilcoxon signed-rank test is that it is distribution-free. On fund level, pension results from the CIR model and LMM are significantly different from those yielded by the KNW model, according to the Wilcoxon test. The median of the difference between the KNW model and G2++ model is not significantly different from zero (prob. = 0.396). For the 25-year-old initial age cohort, the hypothesis of the Wilcoxon test is rejected for all three alternative models. This is again an indication that younger generations have to deal with much uncertainty. For the 75-year-old initial age cohort, only the LMM is rejected.

Given the above-mentioned tests, it is only possible to conclude whether there is a significant difference between the models. It is not possible, however, to determine whether one model is significantly better suited for the feasibility test than another model. Therefore, it important to understand how the scenarios of a particular model are created and which variables influence a pension fund's financial development.

6 Conclusion and Discussion

6.1 Conclusion

In the aftermath of the 2008 financial crisis, additional legislation on Dutch pension funds was introduced as an attempt to make pensions less vulnerable to economic variables and demographic changes. As part of the new Dutch regulatory framework, known as FTK, pension funds are subject to a feasibility test each year. The feasibility test entails a stochastic analysis of over a 60-year period to determine the risk-profile of a pension fund, as well as provide insight into the pension's purchasing power, both on the fund level and individual level. The feasibility test's objective is to force pension funds to reflect upon their ambitions and risk attitude in advance, as well as to force them to actively communicate these aspects to their participants. In this paper, we review the feasibility test through an examination of the test and further analyze the underlying set of scenarios. For the latter portion, we also investigate alternative models to generate scenarios.

This paper's first step was to perform a feasibility test, examine the results for different types of pension funds and check whether the feasibility test served its objective. For the implementation of the feasibility test, we presented a valuable tool based on the generational accounting approach, as described by Chen et al. (2014). The results of the feasibility test clearly demonstrated that, in current economic conditions, it is not evident that the purchasing power conservation of accrued pension benefits is being conserved. The feasibility test indicated that the uncertainty is much higher for younger generations in comparison to their older counterparts. It is difficult, however, to compare results across generations, because missed indexations in the past are not included in the pension result. The assumption that pension benefits are fully indexed at the beginning of the simulation period can also give participants false impressions regarding outcomes, and older generations are especially vulnerable here. Additionally, the feasibility test contains several standard assumptions that do not always correspond with reality. The feasibility test assumes that pension funds always emphasize on price inflation, although in fact various funds use wage inflation. Furthermore, the assumption that pension funds index pension benefits in cases of negative inflation is often not true.

The second part of this paper focused on the set of scenarios for the feasibility test. The Commission Parameters (Langejan et al. (2014)) advised the DNB to use the KNW model to generate a uniform scenario set, which enables feasibility test results to be compared across pension funds. Pension fund managers must understand the model and its underlying assumptions in order to interpret its results. We therefore review the methodology, evaluate the robustness of the estimated model parameters and provide an overview of the assumptions underpinning the calibration and the actualization methods. We re-estimated the two-factor model through the employment of the simulated annealing procedure, as is proposed in Draper (2014), with an updated dataset. The estimated results are described in detail and subsequently compared to the earlier results presented in Draper (2012) and Draper (2014).

Our estimates resulted in improved likelihoods, which proves that the estimation procedures in Draper (2012) and Draper (2014) did not lead to maximum likelihoods. However, we must state that our parameter set is also likely a local optimum. In addition, we find that the significance levels of the updated estimates are lower in comparison to two CPB papers. Furthermore, we found that the KNW model provides a sufficient fit to the Dutch data and is able to match the cross-sectional moments of bond yields. Additionally, we demonstrate that the KNW model is able to match the empirical observations of Campbell and Shiller (1991) as well as Dai and Singleton (2002).

The DNB has largely adopted the recommendations of the Commission Parameters. Following its advice, the DNB uses the estimated parameters in Draper (2014) as the basis for the KNW model. However, the DNB uses different assumptions and a different approach to calibrate and actualize the model parameters of the KNW-model in comparison to the approach recommended by the Commission Parameters. We found several values in the DNB set of scenarios that differ from the proposed values of the Commission Parameters. The most striking difference is that the DNB assumes a risk premium for equity returns that is approximately 2% lower than what the

Commission Parameters proposes. This discrepancy can have a great impact on the results of the feasibility test.

In the final part of the thesis, we compare the results of the KNW-model with three alternative models and examined the impact of alternative scenario sets on feasibility test results. Since the term structure is the most important risk factor, our focus was to model the interest rates. The three alternative interest rate models are based on different assumptions, which are estimated through the use of different techniques. The G2++ model, which is calibrated via Kalman filtering, is most suitable for a simulation-based context such as a feasibility test. The model effectively reproduces the key stylized facts and provides a rich pattern of yield curve shapes over time. In addition, the implied initial yield curve exactly matches the current term structure. The main drawback of the G2++ model is that it may yield negative interest rates. The other two alternative interest rate models, the CIR model and the Libor Market Model, both rule out negative interest rates. The Libor Market Model also matches the current yield curve. Apart from this aspect, both models are not suitable for the context of long-term risk management.

After conducting the feasibility test with base case assumptions for the stylized pension fund, we found clear distinctions between the pension result based on the KNW model and the pension results based on the three alternative models. The most noteworthy finding is the fact that the KNW model has the smallest bandwidth between the expected pension result and the pension result in a 'bad weather' scenario, both on the on fund level as well as for all generations. This finding indicates that the KNW-model produces a more conservative view of the future than the three alternative models. Furthermore, we noticed that the KNW model leads to similar pension results as the G2++ model. This conclusion is supported by the fact that the Wilcoxon sign rank test does not reject the result on fund level and the results that concern the elderly. The pension results based on the different models differ most for the younger generations. For all three of the alternative models, the results differ significantly according to both the Kolmogorov-Smirnov Test and the Wilcoxon Sign Rank Test. This finding emphasizes the fact that an adequate scenario set is typically essential for younger generations to be able to interpret the amount of uncertainty regarding their future pensions. This emphasizes the fact that an adequate scenario set is especially important for younger generations.

6.2 Discussion

The feasibility test is less sophisticated than classical ALM-studies and is based on a limited set of scenarios for a (unrealistic) long evaluation period. In addition, we found several small shortcomings regarding the test's implementation and communication. The feasibility test uses complex measures. For a layman, such as a pension fund board member or participant, it is difficult to understand what a 'bad weather' scenario actually means after 60 years. It is more comprehendible to visualize the development of pensions, including the risk-profile, and to discuss other variables, such as the probability of pension cuts or the critical values of the funding ratios.

The Commission Parameters proposed the KNW model from the perspective of practicality and comparability. In addition, the Commission Parameter claims that the generated scenarios are realistic and based on accepted economic principles. Our research's main concerns the calibration approach by the DNB, which does not correspond to the recommendations of the Commission Parameters in terms of the expectations of certain variables. Following this study, we have no reason to disagree with the selection of the KNW-model to generate the uniform scenario set. We also conclude that it is difficult to compare the performance of different interest rate models and to determine which model is best suited for the feasibility test. Especially in the case of very long-term horizons, the results are subjected to high uncertainty. For that matter, there is not such a thing as the 'right' model.

Despite its simplicity and several inadequacies, we find that the feasibility test can still serve as an indicator, especially because it will be performed annually. Since the test is based on constant assumptions, the stakeholders will know something has changed when the results are different in comparison to the prior year. In order for the objective of the feasibility test to succeed, however, further communication is necessary, first between pension funds and the general public. More insight ultimately leads to greater awareness and more confidence. For example, it is interesting to communicate how different pension policies affect each generation. Secondly, communication is necessary between regulators and pension funds, with respect to the set of scenarios and assumptions that form the basis for the scenarios.

This research's limitation is that it does not consider the effect of multiple asset classes in alternative models. Similar to the KNW model, we only used a single process to model the equity returns. A further study could examine the impact that greater diversity of asset classes have on the results of the feasibility test, perhaps through the use of different investment strategies. An interesting approach to model multiple equity classes could be the use of a multivariate Geometric Brownian Motion. The KWN model was selected partly because it is available to everyone ('open source'). More public research may be conducted on developing new economic scenario generator models for long-term stochastic analysis. Another possible area of future research could entail the exploration of more interest rate models for risk-management purposes. This direction is of interest given the practical importance of generating reliable term structures. Examples of possible interest rate models for Dutch pension funds include the widely used Nelson-Siegel model and the extended Libor Market Model proposed by Norman (2009), who developed a new calibration procedure to obtain parameter estimates from historical data. Furthermore, it would be interesting to examine different estimation procedures for the KNW model in order to find the global optimum.

References

Ahlgrim, K. C., DArcy, S. P., Gorvett, R. W., ARM, F., 2004. Modeling of economic series coordinated with interest rate scenarios. Proyecto de investigación. Disponible en http://casact.org/research/econ/[consultado el 1 de Febrero de 2009].

Allianz, 2015. Retirement income adequacy indicator. International Pension Papers.

Baillie, R. T., Chung, C.-F., Tieslau, M. A., 1996. Analysing inflation by the fractionally integrated arfima-garch model. Journal of applied econometrics 11 (1), 23–40.

Bajeux-Besnainou, I., Jordan, J. V., Portait, R., 2003. Dynamic asset allocation for stocks, bonds, and cash*. The Journal of Business 76 (2), 263–287.

Bauer, R., Hoevenaars, R., Steenkamp, T., 2006. Asset liability management. Oxford Handbook of Pensions and Retirement Income, 417–440.

Baxter, M., Rennie, A., 1996. Financial Calculus An introduction to derivative pricing. Cambridge university press.

Bergstrom, A. R., 1984. Continuous time stochastic models and issues of aggregation over time. Handbook of econometrics 2, 1145–1212.

Black, F., Derman, E., Toy, W., 1990. A one-factor model of interest rates and its application to treasury bond options. Financial analysts journal 46 (1), 33–39.

Black, F., Karasinski, P., 1991. Bond and option pricing when short rates are lognormal. Financial Analysts Journal 47 (4), 52–59.

Brennan, M. J., Xia, Y., 2002. Dynamic asset allocation under inflation. The journal of finance 57 (3), 1201–1238.

Brigo, D., Mercurio, F., 2007. Interest rate models-theory and practice: with smile, inflation and credit. Springer Science & Business Media.

Campbell, J. Y., 2008. Viewpoint: Estimating the equity premium. Canadian Journal of Economics/Revue canadienne d'économique 41 (1), 1–21.

Campbell, J. Y., Shiller, R. J., 1991. Yield spreads and interest rate movements: A bird's eye view. The Review of Economic Studies 58 (3), 495–514.

Campbell, J. Y., Viceira, L. M., 2001. Strategic asset allocation: Portfolio choice for long term investors.

Chen, Z., Pelsser, A., Ponds, E., 2014. Evaluating the uk and dutch defined-benefit pension policies using the holistic balance sheet framework. Insurance: Mathematics and Economics 58, 89–102.

Cochrane, J. H., 2005. Asset pricing. Vol. 1. Princeton university press Princeton, NJ.

Cox, J. C., Ingersoll Jr, J. E., Ross, S. A., 1985. A theory of the term structure of interest rates. Econometrica: Journal of the Econometric Society, 385–407.

Dai, Q., Singleton, K. J., 2002. Expectation puzzles, time-varying risk premia, and affine models of the term structure. Journal of financial Economics 63 (3), 415–441.

Dimson, E., Marsh, P., Staunton, M., Holland, D., Matthews, B., Rath, P., 2014. Credit suisse global investment returns yearbook 2014. Credit Suisse, February.

Draper, N., 2012. A financial market model for the us and the netherlands.

Draper, N., 2014. A financial market model for the netherlands.

Duffee, G. R., 2002. Term premia and interest rate forecasts in affine models. The Journal of Finance 57 (1), 405–443.

Duffie, D., Kan, R., 1996. A yield-factor model of interest rates. Mathematical finance 6 (4), 379–406.

Goffe, W. L., Ferrier, G. D., Rogers, J., 1994. Global optimization of statistical functions with simulated annealing. Journal of Econometrics 60 (1), 65–99.

Heath, D., Jarrow, R., Morton, A., 1992. Bond pricing and the term structure of interest rates: A new methodology for contingent claims valuation. Econometrica: Journal of the Econometric Society, 77–105.

Hibbert, J., Mowbray, P., Turnbull, C., 2001. A stochastic asset model & calibration for long-term financial planning purposes. In: Finance and Investment Conference.

Ho, T. S., Lee, S.-B., 1986. Term structure movements and pricing interest rate contingent claims. The Journal of Finance 41 (5), 1011–1029.

Hoevenaars, R., Molenaar, R. D., Steenkamp, T. B., 2003. Asset and liability management tools.

Hoevenaars, R. P. M. M., 2008. Strategic asset allocation & asset liability management. Universiteit Maastricht.

Hull, J., White, A., 1990. Pricing interest-rate-derivative securities. Review of financial studies 3 (4), 573–592.

Koijen, R. S., Nijman, T. E., Werker, B. J., 2005. Labor income and the demand for long-term bonds. Tilburg University.

Koijen, R. S., Nijman, T. E., Werker, B. J., 2010. When can life cycle investors benefit from time-varying bond risk premia? Review of Financial Studies 23 (2), 741–780.

Koijen, R. S., Nijman, T. E., Werker, B. J., et al., 2006. Optimal portfolio choice with annuitization. Pension Institute, City University.

Kritzman, M., Rich, D., 2002. The mismeasurement of risk. Financial Analysts Journal 58 (3), 91–99.

Langejan, T., Gelauff, G., Nijman, T., Sleijpen, O., Steenbeek, O., 2014. Advies commissie parameters. Technical report, SZW.

Lee, H.-Y., Wu, J.-L., 2001. Mean reversion of inflation rates: evidence from 13 oecd countries. Journal of Macroeconomics 23 (3), 477–487.

Lever, M., Mehlkopf, R., Ewijk, C., et al., 2012. Generatie-effecten pensioenakkoord.

Macaulay, F. R., et al., 1938. Some theoretical problems suggested by the movements of interest rates, bond yields and stock prices in the united states since 1856. NBER Books.

Mercer, 2015. Mercer melbourne global pension index. Melbourne: Australian Center for Financial Studies.

Merton, R. C., Samuelson, P. A., 1992. Continuous-time finance.

Mitra, L., Mitra, G., Roman, D., 2009. Scenario generation for financial modelling: Desirable properties and a case study. OptiRisk Systems: White Paper Series, OPT 10, 1–20.

Nelson, C. R., Siegel, A. F., 1987. Parsimonious modeling of yield curves. Journal of business, 473–489.

Norman, J. P., 2009. Real world interest rate modelling with the bgm model. Available at SSRN 1480174.

Park, F., 2004. Implementing interest rate models: A practical guide. Capital Markets and Portfolio Research (CMPR).

Pensioenregelingen, C. T. A., 2010. Een sterke tweede pijler. Naar een toekomstbestendig stelsel van aanvullende pensioenen.

Sangvinatsos, A., 2008. Expectations hypothesis. Encyclopedia of Quantitative Finance.

Sangvinatsos, A., Wachter, J. A., 2005. Does the failure of the expectations hypothesis matter for long-term investors? The Journal of Finance 60 (1), 179–230.

Shackman, J. D., 2006. The equity premium and market integration: Evidence from international data. Journal of International Financial Markets, Institutions and Money 16 (2), 155–179.

Shi, Z., Werker, B. J., 2012. Short-horizon regulation for long-term investors. Journal of Banking & Finance 36 (12), 3227–3238.

Svensson, L. E., 1994. Estimating and interpreting forward interest rates: Sweden 1992-1994. Tech. rep., National Bureau of Economic Research.

Van de Grift, M., 2009. Pensioenaanspraken en vergrijzing. Den Haag: Centraal Bureau voor de Statistiek.

Van den Goorbergh, R., Molenaar, R., Steenbeek, O. W., Vlaar, P., 2011. Risk models with jumps and time-varying second moments.

Van Duin, C., Stoeldraijer, L., 2014. Bevolkingsprognose 2014–2060: groei door migratie.

Vasicek, O., 1977. An equilibrium characterization of the term structure. Journal of financial economics 5 (2), 177–188.

Vlaar, P. J., 2006. Term structure modeling for pension funds: What to do in practice? Available at SSRN 965363.

Wilkie, A., 1987. Stochastic investment modelstheory and applications. Insurance: Mathematics and Economics 6 (1), 65–83.

Appendices

A Dutch Pension System

The Dutch pension system is regarded as one of the best in the world; see Mercer (2015) and Allianz (2015). Retirees are ensured of a relatively high standard of living. In addition, both reports state that our pension system is sustainable in the long run. As in most European countries, the Dutch pension system relies on three pillars: public pension, occupational pension and private savings. Together these three pillars determine the amount of pension a person will receive when retiring.

- The *first pillar* consists of the state pension also known as AOW pension. It is a tax-funded basic retirement income for all citizens regardless of past earnings. So all people who have lived and/or worked in the Netherlands between the age of 15 and 65 are entitled to receive a state pension when reaching retirement age. This pillar can be seen as a base security scheme for all retirees.
- The *second pillar* consists of (mandatory) occupational pension schemes. Employment in a particular sector or company determines enrolment in the accompanying pension fund or insurer. Pension rights are accrued over the participant's career and are related to salary earned and length of employment.
- The *third pillar* concerns voluntary individual retirement savings including individual investments and life insurance products. This pillar is relevant for the self-employed who do not accrue an occupational pension or for individuals who want to save for an additional pension. Savings in this pillar are tax-deductible to ensure the same advantage of tax treatment as in the second pillar.

The first pillar accounts for approximately 50% of the pension income for a retiree in the Netherlands (Pensioenregelingen (2010)). Van de Grift (2009) shows that this is a very small component relative to many other European countries. The supplementary pension income is mainly collected in the second pillar (approx. 45%). Where other European countries (e.g. Spain and Italy) have a small second pillar, the Dutch pension system is characterized by a large occupational pension scheme. This emphasizes the importance of the second pillar in the total pension income in the Netherlands.

Pension Funds

Most pension money in the second pillar is managed by pension funds. In the Netherlands there are three different types of pension funds:

1. **Industry pension funds** for people working across an entire sector such as the hotel, catering, retail, construction industries or the civil service
2. **Corporate pension funds** for a single company or corporation such as Akzo Nobel, Phillips, Shell and Unilever
3. **Pension funds for independent professionals** such as medical specialists and dentists

The system consist of approximately 80 industry pension funds, 300 pension schemes of individual companies and 12 pension funds for certain type of profession. About 75% of the members in Dutch pension funds are with an industry pension fund.

Van de Grift (2009) shows that more than 90% of all employees in the Netherlands have a pension contract. This high participation rate results from the compulsory nature of the system. In order to ensure a good pension scheme for all employees, the Dutch government has made pension schemes mandatory in most industries. This has led to industry-wide pension funds with sufficient economies of scale, enabling cost efficient management of the schemes.

End 2014, Dutch pension funds managed more than €1.25 trillion of pension capital. This is equal to 189% of the Dutch Gross Domestic Product (GDP) in 2014.[11] The Dutch pension funds vary considerably across number of members as well as the pension capital they manage. The largest pension fund in the Netherlands is APB (Government and Education), which has more than 1 million active members and has an invested capital of circa €345 billion.[12] However, there are also pension funds with less than 100 members and an invested pension capital of only a few million Euros.

Different Types of Pension Agreements

As mentioned before, pension funds ensure that employees save for their retirement in the second pillar. Basically, the pensions paid by pension funds are financed from contributions paid by its members in the past and from the return on the investments of these contributions. So pension funds are responsible for investing the contributions so that maximum returns are generated with minimum risk in order to secure a stable retirement income for its members. Employers and employees together determine the precise features of the collective pension scheme. There are different types of pension schemes. The Defined Contribution (DC) scheme and the Defined Benefit (DB) scheme are the most common.

In a *Defined Contribution* scheme, employees pay a fixed premium or 'defined contribution' each year. At retirement the participant receives a benefit that depends on the investment returns throughout the years. Hence the contributions are predetermined, but the future benefits are not. In this scheme the members bear the risk of a low pension. In a *Defined Benefit* scheme, pension benefits are determined beforehand, while contributions can be adjusted to achieve the 'defined benefits'. In other words, pension funds guarantee a certain level of benefit after retirement. This scheme transfers all the (financial) risk to the fund.

Defined Benefit schemes can be divided into a final-salary scheme and an average-salary scheme. In a *final-salary* scheme, the final accrued pension benefits depend on the number of years worked and the salary earned in the last year of service. This scheme aims at a percentage (often 70%) of the last earned salary. In an *average-salary* scheme, the accrued pension benefits are based on the average salary earned throughout the employee's career. The latter scheme does not automatically compensate for inflation, since it is based on wages from the past. The final-salary scheme does compensate for inflation since wages grow with inflation every year.

Most pension funds with an average-salary scheme only guarantee a nominal pension but also have the ambition to provide indexation, typically on a conditional basis. Indexation means that the accrued pension rights of a member are adjusted for price inflation or for the increase in wages throughout the sector. Without provisioning annual indexation, the value of the accrued pension rights is merely nominal and will substantially fall short of the benefits earned in a final-salary scheme. The actual indexation depends on the financial position of the pension fund.

In the Netherlands, the majority of the pension schemes are based on the Defined Benefit scheme. In recent years there has been a gradual move from final-salary DB schemes to average-salary DB schemes, see Figure 15. Because of deteriorating economic conditions and demographic developments, pension funds have switched to an average-salary DB scheme as the accrued benefits are expected to be less than in a final-salary DB scheme. Up until the beginning of the twenty-first century most members had a final-salary DB scheme (66.5% in 1998). In 2015, only 0.1% of the active members have a final-salary DB scheme and 90.7% have an average-salary DB scheme. In last couple of years, members with a DC scheme steadily increased up to a 6.5% share in 2015. The other schemes are a mixture of different types of schemes. In this thesis we will focus on pension funds with an average-salary Defined Benefit scheme.

[11] Source: www.dnb.nl
[12] Source: www.abp.nl

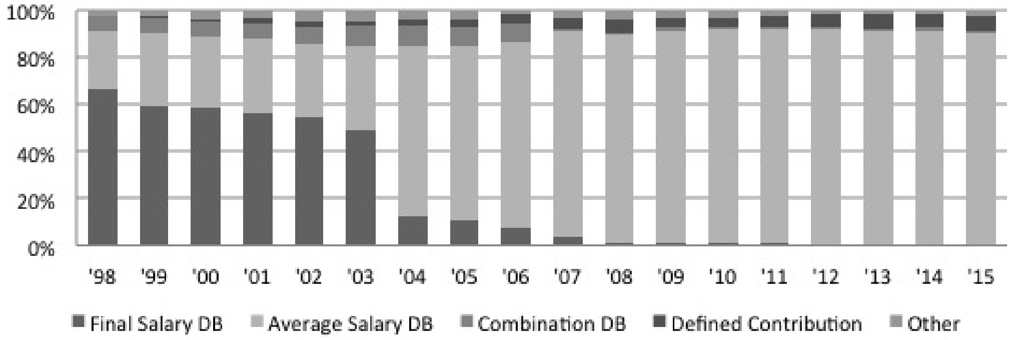

Figure 15: Number of members for the different pension schemes in percentages.

B Regulation Dutch Pension Funds

In the Netherlands, the Dutch Central Bank (DNB) and the Dutch Authority for Financial Markers (AFM) are responsible for the regulation for Dutch pension funds. Pension fund legislation is set out in the Pensions Act. The Financial Assessment Framework, which is part of the Pensions Act, defines the requirements regarding the financial position of pension funds. The framework came into law on January 1, 2007 but has undergone some intermediate alternations in the meantime; see the timeline in Figure 16.

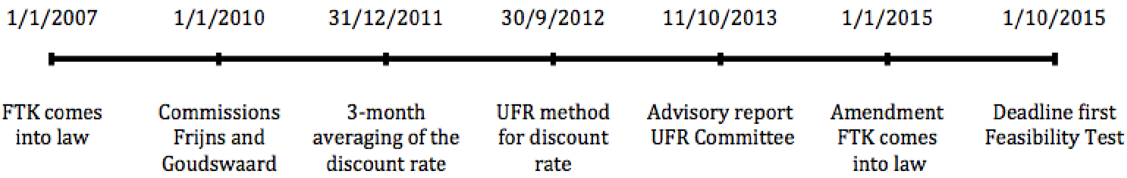

Figure 16: Timeline of regulations and important events concerning regulations

Financial Assessment Framework (2007)

The Dutch regulators developed the Financial Assessment Framework as a response to the threatening low funding ratios amongst pension funds in the beginning of the 21st century. The FTK is created to evaluate the current financial position and the policy set by the pension fund's board. It is based on the principles of market valuation, risk-based financial requirements and transparency. The two main components of the framework are the *liability discount rate* and the *required capital buffers* to ensure that pensions can be paid, complemented by a recovery plan in case a pension fund fails to meet the capital requirements.

The most significant adjustment compared to former legislation is the market-consistent valuation of liabilities. Liabilities were initially valued using a fixed rate of 4% regardless market conditions. This fixed rate was considered to be a prudent estimate of long-term interest rates on sovereign bonds. However, the low interest environment of the early 21st century differed considerably from this fixed rate, which led to an underestimation of the pension liabilities. As of 2007, pension funds had to value their liabilities using the 6-month EURIBOR swap rate. This resulted in a better representation of the financial position of pension funds.

The market-valuation approach for liabilities also entails disadvantages for pension funds. The volatility in capital markets and financial shocks have now a direct impact on the value of the liabilities making the funding ratio much more volatile. Also decreasing interest rates lead to a higher value for the liabilities, which subsequently lowers the funding ratio. Both are detrimental because pension funds prefer a stable and high funding ratio. In December 2011, the DNB decided to use a 3-months moving-average of the swap rate as the statutory discount curve because of the ongoing volatile market conditions. Another motive was that the market for long-term interest rate swaps is assumed to be insufficiently liquid for proper market valuation, which leads to extra volatility in the value of the liabilities.

The second main component in the FTK concerns two important capital requirements for Dutch pension funds. According to the first requirement, the funding ratio has to be above the minimum required capital level (mVEV), which is set equal to 105%. A pension fund is considered underfunded when it is unable to meet this threshold. In addition, the FTK contains a risk-weighted capital requirement (VEV) that serves as a financial buffer to ensure a pension fund will remain solvent. The VEV level is set so that a pension fund remains fully funded ($FR \geq 100\%$) at a 97.5% confidence level within a one-year timeframe. This capital requirement level depends on the composition of the investment portfolio, i.e. the riskier the investments, the higher the buffer requirement. For an average pension fund the VEV has been approximately 120% to 125% in the past.

When pension funds fail to meet these two conditions, they are granted a timeframe to meet the minimum requirements. If the funding ratio drops below the required capital level (VEV), the pension fund has to make a long-term recovery plan (max 15 years) that addresses the reserve deficit. If the pension fund is underfunded relative to the minimum required capital level (mVEV), it must submit a short-term recovery plan (max 3 years) to the DNB. In case the pension fund is unable to recover within these 3 years, it is obliged to cut the accrued pension rights of the members as a last resort. If the funding ratio is between mVEV (105%) and VEV (e.g. 120%), a pension fund is allowed to partially compensate the pension rights for inflation. If the funding ratio exceeds the capital requirements, the fund may provide full indexation. This is also known as conditional indexation.

UFR-method (2012)

In September 2012, the DNB introduced the UFR-method for determining the present value of liabilities for Dutch pension funds. Until then, the valuation of pension liabilities was based only on market interest rates. The UFR valuation methodology includes customized discount rates after the 'last liquid point' in the swap market. The main motivation for the UFR-method is to make liabilities less sensitive to fluctuations and disruptions in the financial markets. Especially, this method tackles the problem of limited liquidity in long-term swap rates, which leads to volatile and unreliable prices and therefore an unreliable discount rate.

In the UFR-method, the longest maturity for which the market is assumed to be fully liquid is 20 years. Therefore, the discount rates with maturities below 20 years are directly derived from the observed swap rates. From the 20-year point onward, the discount rates are calculated using a weighted-average of market forward rates and a fixed stability factor, called the ultimate forward rate (UFR). As maturity increases, the weight of the UFR component increases linearly up to the 60-years forward rate. In this way, the forward rates converge to a predetermined level, which subsequently leads to more stable long-term interest rates. The DNB has set the ultimate forward rate at 4.2%, which is based on an expected inflation rate of 2% and an expected long-term real interest rate of 2.2%. The technical description of the UFR methodology is given in Appendix C.

The UFR-method affects the stability and the level of the discount curve, which subsequently leads to a less volatile funding ratio. This is important for pension funds because the funding ratio plays a central role in the financial management.Figure 17 illustrates the impact of the UFR-method on the discount curve at time of the introduction, September 2012, and in March 2015. In the current low-interest rate environment, the long-term discount rates will increase due to the UFR-method. As a result, the financial position of pension funds improves significantly because future liabilities become cheaper. In case long-term market interest rates exceed the UFR, the long-term discount rates will decrease. Hence, the UFR method mitigates the effect of low and high long-term interest rates on the valuation of pension liabilities.

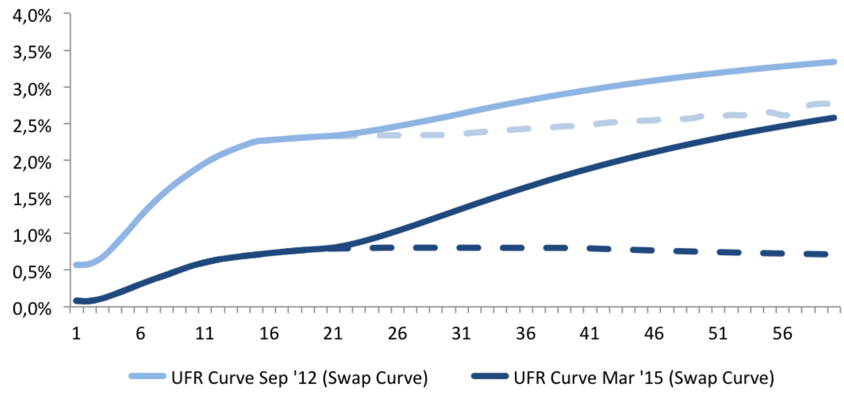

Figure 17: Swap curve and UFR curve in September 2012 and March 2015

Amendment FTK (2015)

The Financial Assessment Framework introduced in 2007 was a direct response to the economic and regulatory environment in the early 2000s. However, the financial crisis in 2008 made clear that Dutch pension funds are still very vulnerable to economic developments and the volatility of financial markets. Also the impact of demographic changes and the ongoing low interest rate environment demonstrated the need to review the pension industry once more.

In May 2009, Minister Donner announced in the letter '*Brede aanpak pensioenvraagstukken*' to thoroughly assess the framework. Subsequently the Dutch government assigned two committees to investigate the durability and possible shortcomings of the Dutch pension system. Commission Frijns (2010) examined the pension fund governance and investment policy. Commission Goudswaard (2010) investigated pension schemes and the long-term durability of the pension system. The findings of both committees resulted in a proposal for a new pension agreement.

Reforming the framework caused a lot of turmoil between stakeholders (e.g. pension funds, participants) resulting in a long period of time before adoption of the new legislation. In the meanwhile, a number of alternations had already been made such as a new discount rate including the ultimate forward rate. On 25 June 2014, the State Secretary for Social Affairs and Employment Klijnsma published the legislative proposal 'Amendment Financial Assessment Framework' which was approved by the parliament in December 2014 and put into effect on 1 January 2015.

The objective of this 'new FTK' (nFTK) is to make pensions less vulnerable to major shocks in financial markets and to contribute to a more balanced distribution of benefits and burdens between involved participants (e.g. young vs. old). The adjusted recovery plan and the new rules regarding indexation are intended to spread financial windfalls and setbacks more evenly over time. Also the nFTK contains a number of elements to make the pension contract more complete. For example, pension funds have to determine in advance how to handle financial shocks. In addition, the financial policy must be clearly defined in advance, including the degree of risk that is acceptable. For the latter, the new feasibility test plays a significant role. Figure 18 displays the most important changes between the former FTK and the new FTK.

Overview of most important elements of the new Framework	**FTK (2007)**	**nFTK (2015)**
A. Policy funding ratio	Prevailing funding ratio	12-month moving-average of funding ratio
B. Discount rate for the valuation of pension liabilities	3-month average swap rate + UFR	Prevailing swap rate + (revised) UFR
C. Capital requirements	Min. capital requirement (mVEV) 105% Risk-weighted capital requirement (VEV) ca. 20% (based on 97.5% CI)	Min. capital requirement (mVEV) 105% Risk-weighted capital requirement (VEV) ca. 25% (based on 97.5% CI)
D. Recovey plans	FR under mVEV: max. 3 years to recovery term mVEV FR between mVEV and VEV: max. 15 years torecovery term VEV	Adj. recovery plan with a (rolling) 10-year recovery period
E. Pension cuts	Unconditional pension cuts after 3 successive years underfunding	Unconditional pension cuts after 5 successive years underfunding
F. Indexation	Conditional indexation from FR 105%	Conditional indexation from FR 110%
G. Premium	Multiple options to smooth premium	Only two options to smooth premiums Discount threshold above VEV
H. Policy tests	Continuity analysis (every 3 years) Consistency test	Feasibility test (annually)

Figure 18: Overview of most impactful changes in the nFTK relative to the old framework

The funding ratio is a key measure for pension boards and regulators to measure the financial solvency of pension funds and is therefore an important metric for the financial policy. After the introduction of market-based valuation of pension liabilities, the funding ratio became vulnerable to capital market volatility and financial shocks. This enhanced volatility in funding ratios causing policy implications. Pending the new legislation, several intermediate adjustments were implemented to reduce the impact of short-term market volatility. To smooth the funding ratio, the 3-month average of the swap rate was used to discount pension funds' liabilities. Nevertheless, this smoothing mechanism proved to be relatively ineffective.

In the new FTK, pension funds will need to report the 'policy funding ratio', which is based on the 12-month moving-average of the funding ratio (see A). The policy funding ratio will be the basis on which the regulator will evaluate the financial position of pension funds. The three-months averaging in the term structure will expire (see B). In contrast to the averaging of the interest rate term structure, the policy funding ratio does not distort the relationship between assets and liabilities as it impacts both equally.

In late 2012, the Dutch government asked the Commission UFR to review the UFR methodology and to provide advice on the height of the UFR, the starting point of convergence and the method of extrapolation. According to the Commission UFR, the fixed UFR of 4.2% was not sufficiently substantiated and they proposed an UFR based on the 120-month average of the 20-years 1-year forward rates. On 15 July 2015, the DNB adopted the new UFR-method following the recommendations of Commission UFR. At time of introduction, the new UFR was equal to 3.3%, which differs significantly from the former fixed UFR of 4.2%.[13]

The FTK imposes strict rules on the (minimum) capital buffers that pension funds should hold. The standard model used by the Dutch Central Bank (DNB) to calculate the required capital remains intact. However, the underlying parameters (e.g. interest rate risk, equity shock scenarios) and underlying asset class correlations have been reassessed. According to Commission Parameters (2014), the change in the parameters will increase the risk-weighted capital requirement (VEV) on average by c. 5% (see C).

The old FTK consisted of a long-term recovery plan concerning the capital requirement (VEV) and a short-term recovery plan for pension funds that are regulatory underfunded (i.e. below mVEV). The given time is a strict requirement and shocks have to be absorbed within the recovery plan. Because of this, it is possible that very severe measures must be taken at the end of the recovery period. This phenomenon is called 'drifting ice'.

In the new FTK, both recovery plans are replaced by a rolling 10-year recovery plan. In case of underfunding, pension funds have to submit a plan to reduce their current shortfall with respect to VEV by a tenth annually. Each year the recovery plan is based on the current shortfall and substitutes the previous recovery plan, i.e. the recovery plan has no memory. By working with a rolling 'memory-less' recovery period instead of a 'drifting ice' principle the chance of severe measures significantly reduces (see D). In case of five successive years of underfunding, pension funds must institute a series of unconditional pension cuts to regain the minimum regulatory buffer immediately (see E).

In the old framework, pension funds start with (partially) indexation when its funding ratio exceeds the threshold of 105%. In case of sufficiently high surplus backlogged indexation can be provided to compensate for missed indexation in the past. The new FTK specifies a more stringent indexation policy to ensure durable indexation (see F). The threshold for indexation increases to 110% and the maximum height of indexation is determined by the scheme of 'future-proof indexing'. Pension funds must be able to provide indexation for many years to come. Backlogged indexation is allowed when the capital requirement level (VEV) is exceeded and the full wage or price indexation for the current year can be provided. Compared to the old FTK, the new policy has considerable less space to restore previously forgone indexation.

The FTK states that the premium rate must be cost-effective, i.e. cover the costs of the pension plan. So the premium rate is determined by the pension plan and the risk attitude of the

[13]Note that this paper does not apply the new UFR-methodology since it was not yet approved at time of research.

pension fund. Because stakeholders and regulators prefer a stable premium rate, pension funds are allowed to adjust the 'normal' premium rate to prevent volatile premiums over time. In the old FTK, pension funds had many options to smooth their premium rate. In the new FTK, there are only two ways allowed: through the use of 10-year interest rates averaging or through the use of expected returns minus surcharge ambition. The fund's premium policy must be defined in advance. It must also include the funding ratio threshold at which a premium discount may be given to members.

The DNB wants to have a good view on the possibility that future coverage will be at risk. To check for consistency between the indexation ambition and the expected future realizations, pension funds were required to carry out a continuity analysis and a consistency test. The continuity analysis provides insight in the financial position of pension funds for the coming 15 years. Pension funds were obliged to perform the continuity analysis every 3 years. In the new FTK the continuity analysis and the consistency test will be replaced by the feasibility test. This test must be performed on annual basis (see H).

C UFR - Technical Description

This section describes the methodology introduced on September 20, 2012 by the Dutch Central Bank (DNB) to construct the term structure for pension funds.[14] The constructed term structure is used to discount future pension liabilities. The adjustment of the interest rate term structure only concerns a change in the zero interest rates for maturities longer than 20 years. For maturities of 21 years and longer, the zero interest rate is determined by extrapolating the underlying 1-year forward rates towards the so-called Ultimate Forward Rate (UFR). As of September 30, 2012 the UFR has been set at 4.2%. So the extrapolation of the forward interest rates consists of a weighted average of the observed forward rate and the UFR.

When constructing the term structure the following assumptions are made:

- For maturities up to 60 years, the zero interest rates are calculated using the interest rates for swaps published by Bloomberg in which a fixed rate is exchanged for the 6-month EURIBOR.
- For these maturities the 1-year forward rates are derived from the swapcurve.
- The customized 1-year forward rates up to 20 year are the same as the 1-year forward rates from the swap curve.
- The weight of the UFR-component in the customized 1-year forward rates for maturities 20 to 60 year increases with maturity. The weighs used for each maturity are constant over time.
- The customized 1-year forward rates with a maturity over 60 year are constant over time and equal to the UFR.

At the end of each month the construction of the term structure will take place according to the following routine:

1. Derive the 1-year forward rates $F_{t-1,t}$ using the swapcurve for maturities $t = 1$ up to $t = 60$ year. Here the following applies:

$$F_{t-1,t} = \frac{(1+R_t)^t}{(1+R_{t-1})^{t-1}} - 1 \quad \text{for} \quad t = 1, 2, \ldots, 60 \quad \text{and} \quad R_0 = 0; \tag{64}$$

2. Determine the adjusted 1-year forward rates $F^*_{t-1,t}$ as follows

$$F^*_{t-1,t} = \begin{cases} F_{t-1,t} & \text{if} \quad 1 \leq t \leq 20 \\ (1-w_t) \cdot F_{t-1,t} + w_t \cdot UFR & \text{if} \quad 21 \leq t \leq 60 \\ UFR & \text{if} \quad 61 \leq tt \end{cases} \tag{65}$$

   where $F_{t-1,t}$ is the 1-year forward rate for maturity t as calculated in step 1, w_t is the weight for maturity t and UFR is the ultimate forward rate equal to 4.2%.

3. Calculate the zero interest rate R^*_t for maturity t using the following equation

$$(1+R^*_t)^t = \prod_{j=1}^{t} (1+F^*_{j-1,j}) \quad \text{for} \quad t = 1, 2, \ldots \tag{66}$$

The weights used in step 3 of the technical description are determined based the market data when the methodology was introduced. The weights are based on the extrapolation method Smith-Wilson proposed in Solvency II with a small modification. This modification is applied

[14] Source: www.dnb.nl

to counter the concentration of the interest rate sensitivity on the 20-year point. This approach results in fixed weights for each maturity that are specified in table 6.

Maturity	Weight	Maturity	Weight	Maturity	Weight	Maturity	Weight
21	0.086	31	0.701	41	0.903	51	0.974
22	0.186	32	0.732	42	0.914	52	0.978
23	0.274	33	0.760	43	0.923	53	0.982
24	0.351	34	0.785	44	0.932	54	0.985
25	0.420	35	0.808	45	0.940	55	0.988
26	0.481	36	0.828	46	0.947	56	0.990
27	0.536	37	0.846	47	0.954	57	0.993
28	0.584	38	0.863	48	0.960	58	0.995
29	0.628	39	0.878	49	0.965	59	0.997
30	0.666	40	0.891	50	0.970	60	0.998

Table 6: The fixed weights for each of the maturities to calculate the UFR forward curve

D Asset Liability Management

Pension funds transfer cash flows between active members who pay premium and retirees who receive pension. Needles to explain that all members want a system in which premiums are as low as possible and pension benefits are as high as possible. In addition, pension funds look for the optimal balance between pursuing high returns and lowering the probability of not being able to pay promised pension benefits in the future. To determine the pension deal, it is important for board members to evaluate the financial health of the pension fund, both now and in the (distant) future. This section will shed some light on the balance sheet of an average-salary Defined Benefit pension fund in order to understand the factors that influence the financial position of a pension fund.

Figure 19 displays a simple balance sheet for a pension fund. The asset (A) side represents the investment portfolio, including the strategic asset allocation mix for an average Dutch pension fund.[15] The right-hand side consists of the liabilities (L) and the Surplus (S). The liabilities represent the accrued pension rights of members that have to be paid in the future. The surplus (S) can be calculated by $(A - L)$ and indicates the financial health of the fund. Note that the funding ratio is equal to 1+(S/L).

Assets (A)		**Liabilities (L)**
Fixed Income	55%	Pension rights
Equities	30%	
Real Estate	10%	
Alternative Inv.	5%	**Surplus (S)**

Figure 19: Typical balance sheet of a Dutch pension fund.

Based on the balance sheet, pension funds perform an ALM study to gain insight into future developments of assets and liabilities, and therefore the future financial position of the fund. To model possible developments we need to understand the different factors that influence the value of the assets as well as the value of the liabilities. Following Bauer et al. (2006) we present a short overview of the key factors in Figure 20, including economic variables and policy decisions. We will first discuss the liabilities and subsequently the asset side of the balance sheet.

Liabilities

The liabilities consist of the accrued pension rights, which will be paid out from the age of retirement. According to Bauer et al. (2006) the value of the liabilities is determined by three factors: actuarial factors, interest rates and inflation.

Actuarial factors determine the level and length of future payments to participants. These factors entail properties of the participants including gender, age, life expectancy, retirement age, job promotion and discharge. Also assumptions regarding demographic trends like an aging population (ratio active members vs. retirees), labour participation and wage growth have to be taken into account. Underlying transition probabilities (e.g. mortality rates) determine the length of the outgoing cash flows. When for instance the life expectancy of the population increases, the horizon of payments (n) will be affected. This subsequently increases the value of the liabilities.

[15] Source percentages strategic asset allocation mix: www.pensionthermometer.nl

Δ Surplus =	Δ Assets	-/- Δ Liabilities
Contribution policy	Contributions (+)	
Indexation policy	Pension payments (-)	
Investment policy	Gross investment return (+/-) Costs (-)	
Actuarial factors		Longevity risk (+)
Economic factors		Interest rate (+/-) Inflation (+)

Figure 20: Factors influencing the Surplus.

Interest rates have a direct impact on the valuation of pension liabilities because the discount rates are based on the term structure of nominal yields, for example government bond yields or swap rates. In the Netherlands, the discount factors are based on the interest rate curve determined by the Dutch Central Bank (DNB). Section C discusses the methodology of constructing this DNB interest rate curve in more detail.

Inflation has impact on the purchasing power of the accrued pension rights. Pension funds aim to index the accrued pension rights by wage or price inflation. This would ensure the purchasing power level of (future) pension recipients. Without regular interim indexation, the accrued nominal pension benefits of a participant at age 40 will lose approximately half of their value by time of retirement. On the other hand, not or partially implementing indexation has a strong positive effect on the pension funds financial position. Therefore the level of indexation is often conditional on the financial position of the fund, which is measured by the funding ratio.

Assets

The value of the asset side of the balance sheet is influenced by pension payments to participants (cash outflow), contributions from participants (cash inflow) and investment returns (cash inflow/outflow).

The *pension payments* each year depend on the accrued pension rights for each retired member. In most pension schemes, the annual accrual rate for pension benefits is between 1.75% and 2.00% of the pensionable salary. For example, an employee accrues pension rights equal to 1.75% of his salary in each service year. After 40 working years, the annual pension received will be 70% of his average salary. These pension rights are predetermined. In addition, if possible, indexation of the accrued rights can be provided as explained above.

The *contributions* of the active members depend on the premium rate, which is defined as a percentage of the (pensionable) salary. In Hoevenaars (2008), for example, they assume the average premium is c. 20% of the salary. The Pension Act requires that the premium must al least cover the costs of the pension fund. Therefore, the premium rate consists of three components: the actuarially necessary contribution for the (un)conditional part of the pension agreement, a surcharge for handling costs and a surcharge for capital buffers. The premium rate is calculated each year using an ALM-study.

In many pension schemes the premium rate as well as the level of indexation depend on the funding ratio. It can happen that a pension fund gets into financial difficulties and that the funding ratio drops below 100%. The underfunded pension fund will not have enough assets to meet all its future obligations towards its members. In this case all stakeholders involved have to contribute to the recovery: the premium level must be increased and the indexation will be limited. When pension funds fail to retain solvency from their underfunded position, the board can choose to reduce the accrued pension rights. This is also known as *pension cuts.* It is an extreme measure and is therefore considered as an ultimate remedy.

The *investment policy* selects the strategic asset mix based on macro-economic and financial developments in the past and expectations for the future. Additionally, the investment policy also contains the decisions made by the board regarding the rebalancing strategy of the strategic mix, for example buy-and-hold versus dynamic rebalancing. When determining the investment policy a balance must be found between the needs of those who are (almost) retired and the needs of the younger contributor, who probably wants to take more risk to achieve a good return on investment.

The main asset classes for pension funds are fixed income, equities, real estate and alternative investments, see Figure 19. To minimize the risk, pension funds allocate a large part in relatively save bonds, which return annually a fixed interest rate. The return on equity yields more risk but also higher returns than bonds. Naturally, the strategic asset mix differs over time. For example, when bonds are giving a relatively high return, the allocation towards bonds will increase. Comparably it would be unfair towards the younger generation if a pension fund would invest fully in bonds when the stock markets are bullish. In that sense, pension funds have to find the optimal trade-off between risk and return.

Interest Rate Risk

The investment policy reviews all downside risks carefully. Both assets and liabilities are valued according to the market, which makes the funding ratio sensitive to market shocks. This is unbeneficial for pension funds since they prefer to have a stable and high funding ratio. Therefore, the pension fund has to incorporate various types of risks in the investment strategy. The most explicit risk a pension fund faces is interest rate risk.

Because future pension payments are discounted using the term structure, the value of the liabilities is sensitive to changes in the interest rates. The sensitivity of the value to interest rates changes is indicated by duration. Duration depicts the percentage change in value to a 1%-point change in the interest rate. Due to the long horizon of promised pension benefits, the value of the liabilities has a high duration. According to the DNB, the average duration of pension liabilities is circa 20 year. So if interest rates decline with 1%, the value of the liabilities increases with 20%. In this case, the funding ratio will very likely decrease because the duration of the fixed income portfolio is normally lower than the duration of the liabilities.

To decrease the volatility of the funding ratio, pension funds will primarily attempt to match the duration of the assets to the duration of the liabilities so that changes in interest rates will not significantly influence the funding ratio. As shown on the balance sheet in Figure 19, pension funds have on average a substantial fixed income portfolio, which partly offsets the sensitivity of the liabilities. To hedge the remaining risk, pension funds can make use of derivatives like interest rate swaps.

To incorporate the various risks and the fund's view on the market, pension funds often work with a matching portfolio and a return portfolio. Investments in the matching portfolio are acquired to meet obligations, while the return portfolio has to generate high returns for the shareholders. In order to get the optimal asset mix, pension funds make use of Asset Liability Management (ALM) models to study possible developments of assets and liabilities in terms of duration, returns and inflation.

E Feasibility Test Sensitivity Analysis

In Section 3 the analyses are based on three stylized pension funds and various assumptions including the initial financial position and the policy measures. Changing the initial assumptions can have a significant impact on the values of the fund's performance measures. Therefore it is important to understand how a pension fund will react once the assumptions are adjusted to various situations. By modifying one of the assumptions and keeping the others unchanged, we can have a good overview of how sensitive the output is to specific variables. In this section we present the sensitivity of a fund regarding the initial funding ratio and the asset allocation mix. In addition, we analyse the impact of the different policies on the well being of different generations.

Initial Funding Ratio

In Table 7 we first examine how the impact of the initial funding ratio. In the benchmark case the initial funding ratio is 105%. Now we also include the initial funding ratios of 80% and of 105% to represent a bad and a good initial funding position, respectively.

Initial Funding Ratio	**IFR = 80%**	**IFR = 105%**	**IFR = 130%**
Mean FR	153.0%	172.2%	206.6%
50th-percentile FR	131.0%	134.2%	141.0%
5th-percentile FR	90.4%	92.2%	93.6%
P(FR < 105)	28.7%	21.8%	14.8%
P(FR < 100)	21.8%	15.0%	10.0%
Pw(FR < 100)	100.0%	87.8%	67.0%
P(FR = 5y below 105)	7.5%	5.2%	3.5%
50th-percentile PR	70.7%	82.7%	92.7%
5th-percentile PR	44.1%	51.2%	58.2%
p50 - p5	26.6%	31.5%	34.5%

Generation level	**50 perc.**	**% Δ**	**5 perc.**	**% Δ**
Initial FR = 80%				
25y	81.6%	(-6%)	41.0%	(-6%)
45y	65.6%	(-18%)	31.6%	(-16%)
75y	67.8%	(-22%)	50.1%	(-21%)
Initial FR = 130%				
25y	94.1%	(8%)	45.8%	(6%)
45y	91.5%	(15%)	43.5%	(16%)
75y	97.8%	(13%)	74.8%	(19%)

Table 7: Summary statistics representing the solvency position and indexation quality for the sensitivity anaysis concerning the Initial Funding Ratio for the average pension fund. The upper panel includes finacial statistics on fund level and the lower panel includes the pension results for different generations. %Δ stands for the percentage change compared to the benchmark.

As one would expect, an improved initial funding ratio leads to a higher average funding ratio after 60 years as well as a higher lower bound. It also reduces the chance on a funding shortfall and decreases the number of pension cuts over time considerably. Inherently, the opposite is true for the very low initial funding ratio. A higher initial funding ratio is also very positive for the participants. Their accrued benefits are more likely to be indexed and less likely to be cut. The median of the pension result increases by as much as 10%-points to 93%. On the contrary when the current financial position is poor the median pension result decreases to 71%.

If we look at the pension result on generation level we notice the impact of the initial funding ratio is largest for the older generations. On the one hand, a high initial funding ratio results in more indexation at the beginning of the simulation, which is of course extra attractive for the older generations with high pension rights. And on the other hand, a low initial funding ratio leads to less indexation and an increased risk of pension cuts. This will affect the elderly especially because of their high pension rights and less time to recover from any setbacks. Note that the extended recovery period of the new FTK (pension cuts after 5 years instead after 3 years) is beneficial for the elderly.

A possible reason why the impact for young participants is lower is that a good financial position can lead to high pension payments in the short term, which can lead to lower pension benefits in the long term. Also the younger generations are less affected by an initial funding ratio of 80% because the pension fund has time enough to recover from the pension deficit.

Asset Allocation Mix

The second assumption we examine is the pension fund's asset portfolio. Note that the asset portfolio consists of only equities and bonds. The returns on both asset classes are different, i.e. equity returns are highly volatile whereas the bond returns are less so. Hence, the asset allocation mix will directly impact the development of the pension fund's assets over time and therefore will also have a direct effect on the final output variables. Here we assume that the benchmark invests 50% in equities and 50% in bonds and rebalances the portfolio each year. In addition, 40% of the interest rate risk is hedged by interest rate swaps. For the sensitivity analysis we will create a risky portfolio and a riskless portfolio. The risky portfolio invests 75% in equities and 25% in bonds. The riskless portfolio invests only 25% in equities and the rest in safe bonds. The additional hedge will remain the same. The results are shown in Table 8.

Looking at the risky portfolio, we see that the median funding ratio increases compared to the benchmark but the downside risk increases as well. So the more investment risk a pension fund takes, the greater the dispersion around the expected results. A riskless portfolio results in lower returns but also in a more stable funding ratio. This leads to a lower probability of underfunding.

Remarkably to see is that the chance on pension cuts, which is equal to 5 consecutive years of underfunding, is higher for the riskless portfolio than for the risky portfolio. It seems that the risky portfolio has a higher probability on recovery within this time period. Though, higher risk leads to higher pension cuts on average.

For a less risky investment policy the downside risk for the pension result decreases. However, the average pension results decreases as well with more than 10%-points to 72%. On the contrary, more risk leads to a higher pension result on average.

If we look at the pension result for each on generation, the impact of the investment policy is largest for the younger generations. In particular, the riskless portfolio leads to substantial lower median pension results for these particular generations. For example, for the initial 45-year-olds the median pension result falls from 80% to 64%.

Asset Allocation Mix	Equity = 25%	Equity = 50%	Equity = 75%
Mean FR	122.1%	172.2%	344.2%
50th-percentile FR	120.7%	134.2%	156.8%
5th-percentile FR	96.9%	92.2%	85.1%
P(FR < 105)	24.3%	21.8%	22.1%
P(FR < 100)	12.2%	15.0%	17.3%
Pw(FR < 100)	86.6%	87.8%	86.9%
P(FR = 5y below 105)	6.8%	5.2%	5.0%
50th-percentile PR	71.5%	82.7%	86.1%
5th-percentile PR	55.5%	51.2%	45.7%
p50 - p5	16.0%	31.5%	40.4%

Generation level	50 perc.	% Δ	5 perc.	% Δ
Equity = 25%				
25y	72.0%	(-17%)	48.5%	(12%)
45y	63.7%	(-20%)	43.4%	(16%)
75y	81.1%	(-6%)	68.3%	(8%)
Equity = 75%				
25y	93.3%	(7%)	36.1%	(-17%)
45y	86.1%	(8%)	29.2%	(-22%)
75y	89.5%	(4%)	56.5%	(-10%)

Table 8: Summary statistics representing the solvency position and indexation quality for the sensitivity anaysis concerning the Asset Allocation Mix for the average pension fund. The upper panel includes finacial statistics on fund level and the lower panel includes the pension results for different generations. %Δ stands for the percentage change compared to the benchmark.

F Derivation Affine Term Structure in KNW model

Koijen et al. (2010) derive nominal bonds prices in the financial market, described in Section 4, following the results of affine term structure models in papers Duffie and Kan (1996) and Sangvinatsos and Wachter (2005). Following the literature, bond prices are assumed to be smooth functions of time and the term structure factor X. The price of a nominal bond at time t that matures at time T is indicated by $P(X_t, t, T)$. Koijen et al. (2010) show that $\phi_t P(X_t, t, T)$ is a martingale, where ϕ is defined in equation (13).

Draper (2014) defines the fundamental pricing equation for a nominal zero coupons as

$$E(d\phi^N P^N) = 0 \tag{67}$$

This equation implies that the expected discounted value of the price of a nominal bond does not change over time. The condition also implies for inflation linked bonds

$$E(d\phi^N P^R \Pi) = 0 \tag{68}$$

which indicates that the discounted value of the inflation corrected price of real bonds doesn't change over time. A second-order approximation of the fundamental pricing equation (67) of a nominal zero coupon bond is

$$E[d\phi^N.P^N + \phi^N.dP^N + d\phi^N.dP^N] = 0 \tag{69}$$

Using the Itô Doeblin theorem we obtain

$$\begin{aligned} dP^N &= P_X'^N dX + P_t^N dt + \frac{1}{2} dX' P_{XX'}^N dX + dX' P_{Xt}^N dt + \frac{1}{2} dt P_{tt}^N dt & (70) \\ &= P_X'^N (-KX_t dt + \Sigma_X' dZ_t) + P_t^N dt + \frac{1}{2}(dZ_t)\Sigma_X P_{XX'}^N \Sigma_X' dZ_t & (71) \end{aligned}$$

because in the limit dt tends to 0, the dt^2 and $dtdZ$ terms disappear and the dZ^2 term tends to dt. Substitution of this equation for the price changes and the nominal stochastic discount factor (13) intro the the fundamental valuation equation (67) brings about

$$0 = P_X'^N(-KX_t) + P_t^N + \frac{1}{2} tr(\Sigma_X P_{XX'}^N \Sigma_X') - P^N R_t - P_X'^N \Sigma_X' \Lambda_t \tag{72}$$

Note, the trace term (see Cochrane (2005)) appears because only quadratic terms remain due to independence of the error terms. This partial differential equation has a solution of the form

$$P^N(X_t, t, t+\tau) = exp\big(A^N(\tau) + B^N(\tau)' X_t\big) \tag{73}$$

In case of a single pay-off at time T, duration is defined as $\tau = T - t$. Substitute the derivatives

$$\begin{aligned} \frac{1}{p^N} P_X^N &= B^N & (74) \\ \frac{1}{p^N} P_t^N &= -\frac{1}{P^N} P_\tau^N = -\dot{A}^N - \dot{B}'^N X_t & (75) \\ \frac{1}{p^N} P_{XX'}^N &= B^N B'^N & (76) \end{aligned}$$

into the partial differential equation (72)

$$0 = B'^N(-KX_t) + (-\dot{A}^N - \dot{B}'^N X_t) + \frac{1}{2} tr(\Sigma_X B^N B'^N \Sigma_X') - R_0 - R_1' X_t - B'^N \Sigma_X' (\Lambda_0 + \Lambda_1 X_t) \tag{77}$$

to obtain explicit expressions for A^N and B^N. Note: $tr(\Sigma_X B^N B'^N \Sigma_X') = tr(B'^N \Sigma_X' \Sigma_X B^N) = B'^N \Sigma_X' \Sigma_X B^N$ because $tr(AB) = tr(BA)$. Both the stochastic term and the non-stochastic term have to be zero, leading to

$$\begin{aligned} \dot{A}^N(\tau) &= -R_0 - (\Lambda_0' \Sigma_X) B^N(\tau) + \frac{1}{2} B'^N(\tau) \Sigma_X' \Sigma_X B^N(\tau) & (78) \\ \dot{B}^N(\tau) &= -R_1 - (K' + \Lambda_1' \Sigma_X) B^N(\tau) & (79) \end{aligned}$$

The nominal zero coupon bon with duration $\tau = 0$ and payout 1 had a price $P^N(X_t, t, t) = 1$, which implies $A^N(0) = 0$ and $B^N(0) = 0$. The instantaneous nominal yield of a bond with duration zero is defined as $-d\ln P(X_t, t, t) = -(\dot{A}^N(0) + \dot{B}^N(0)'X_t) = R_0 + R_1'X_t \equiv R_t$. The instantaneous nominal yield of a bond with duration τ is $-d\ln P(X_t, t, t+\tau) = -(\dot{A}^N(\tau) + \dot{B}^N(\tau)'X_t)$. The differential equations can be solved in closed form

$$B(\tau) = (K' + \Lambda_1'\Sigma_X)^{-1}[exp(-(K' + \Lambda_1'\Sigma_X)\tau) - I_{2\times2}]R_1 \tag{80}$$

$$A(\tau) = \int_0^\tau \dot{A}^N(s)ds \tag{81}$$

with $I_{2\times2}$ the two by two identity matrix.

G Data

Following Draper (2014), this paper uses the same data as in Van den Goorbergh et al. (2011) to ensure that a comparison between the results of the corresponding papers is possible. Van den Goorbergh et al. (2011) have provided us quarterly data for historical term structures, price levels and an global equity index for the period 1973 to 2014. For estimating the model parameters all returns are geometric defined.

- *Inflation*: From 1973 to 1999, the German (Western German until 1990) consumer price index figures published by the International Financial Statistics of the International Monetary Fund are used. From 1999 on, the Harmonized Index of Consumer Prices for the euro area from the European Central Bank data website is included, see figure (22).
- *Yields*: Six yields are used in estimation of the KNW model: 3-month, 1-year, 2-year, 3-year, 5-year, and 10-year maturities, respectively. For the estimation of the alternative models (G2++ and LMM) more bond maturities are used, including: maturities 11-year to 15-year (available from Q4-1986), marurities: 20-year, 25-year, and 30-year (available from Q2-1996), maturities: 40-year, 45-year, 50-year, and 60-year (available from Q3-2001). 3-month money market rates are taken from the Bundesbank. For long-term nominal yields, the zero-coupon rates are constructed from swap rates published by De Nederlandsche Bank (www.dnb.nl), see figure (21).
- *Stock Market*: MSCI World index from FactSet. Returns are in euros (Deutschmark before 1999) and hedged for US dollar exposure, see figure (22).

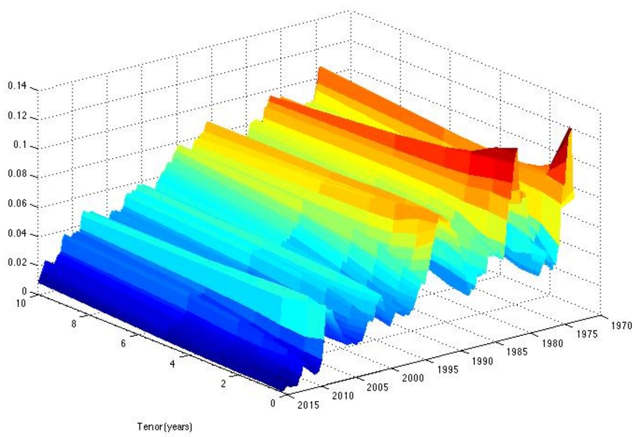

Figure 21: Nominal yields

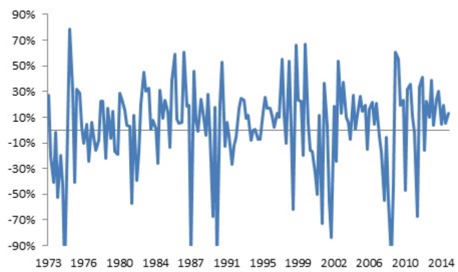

Figure 22: Inflation (left) and equity return (right)

H Calibration Scenarioset

In line with the findings of the Commission Parameters, the DNB publishes each quarter a scenarioset for the feasibility test that is based on the model developed by Koijen et al. (2010). The basis for the model parameters are the estimates presented in Draper (2014), which are estimated using historical time series from 1972 to 2013. In Langejan et al. (2014), the Commission Parameters recommends a number of modifications to certain model parameters to make the expected values of the scenarioset more consistent with the expectations of the Commission Parameters. However, the DNB does not adopt these recommendations and applies other modifications for generating the scenarios. In this section, we examine the calibrated model parameters for both the Commission Parameters and the DNB and compare the results.

The Commission Parameters has several tasks, including advising the DNB on the stochastic scenarioset for the feasibility test. In addition, the commission advises on maximum limits for financial variables that must be used in the recovery plan to calculate expected investment returns. The DNB requires pension funds with a funding shortfall to conduct a deterministic analysis with the formulated expectations in order to generate realistic expectations regarding the recovery power of the pension fund.

For both the feasibility test and the recovery plan, the Commission Parameters examined extensively the expected returns and volatilities of various investment classes. The commission has carried out a large number of qualitative and quantitative analyses and has also conducted a survey among market participants. In Langejan et al. (2014), the commission provides a detailed report on the various methods as well as their findings regarding the expected values.

Table 9 summarizes the expectations for the key financial variables formulated by the Commission Parameters, including the expected equity and bond returns, the volatility of stocks and the expected inflation. The commission recommends using geometric expected returns and inflation rates when evaluating the expected future financial position of pension funds. The geometric mean is most relevant as it expresses the actual long-term growth of both the assets and pension liabilities.

Expected equity return	7.0%
Expected price inflation	2.0%
Expected wage inflation	2.5%
Standard deviation equity return	20.0%
Ultimate forward rate	3.9%

Table 9: Expectations Commission Parameters. The expected returns are defined using geometric discounting.

The Commission Parameters recommends an expected geometric return on listed equity of 7%. They use two common methods to determine the expected return on equity. Firstly, the commission considers past realized stock returns as the main measure. The historical overview in Langejan et al. (2014) shows that equity returns fluctuate strongly over the years. According to Dimson et al. (2014), the average real geometric return for the period 1900-2013 is equal to 5.2%. Assuming an inflation of 2%, the nominal expected return is 7.2%.

Secondly, the expected return on equities is regarded as the sum of the risk-free rate and a risk premium. This method is often the starting point in the academic literature. The estimates for the risk premium range broadly from 3.0% (Campbell (2008)) to 9.5% (Shackman (2006)). The wide range reflects the inherent uncertainty when estimating the future equity risk premium. Dimson et al. (2014) presents a geometric value of 4.3% for the risk premium relative to the yield on short-term Treasury notes. With current low interest rates, the average total return will not exceed 5%. However, the Commission Parameters denotes a degree of negative correlations between the level of risk premium and the risk-free rate. Given the current low interest rate environment, they deem a surcharge on the historical risk premium justified and therefore a 7% total return defensible.

Langejan et al. (2014) shows that for the period 1900 - 2012 the volatility on equity was equal to c. 17%. In addition, they conclude that the volatility fluctuates very much over time. For example, the high volatility in the past 10 years is caused by the recent financial crisis. The commission proposes a standard deviation of 20%, which is higher than the historical average. Applying an above average volatility for scenarios implies that extreme events will occur more often in the future.

The Commission Parameters considers 2% for the expected inflation realistic given the ECB's inflation target (close, but below 2%) and prior inflation figures since 1983. In addition, the Commission Parameters recommends setting the expected increase in contractual wages at 2.5%. This based on an inflation rate of 2% and an expected real wage increase of 0.5% each year. For long-term interest rates they apply the UFR methodology, which is described in Section B.

Despite the extensive research of the Commission Parameters, the DNB has not adopted the recommendations.[16] The DNB has only adjusted the historical estimates so that the long-term inflation tends to 2% and the long-term interest rates tend to the UFR level. Table 10 illustrates the model parameters from Draper (2014) that are changed to capture the expectations of the Commission Parameters and the DNB.

Parameters	Estimated	CP Calibrated	DNB Calibrated
δ_0	1.81%	1.98%	2.00%
η_S	4.52%	6.57%	4.52%
$\sigma_{S(4)}$	16.59%	17.69%	16.59%
$\Lambda_{0(1)}$	0.403	0.242	0.280
$\Lambda_{0(2)}$	0.039	0.039	0.027

Table 10: Estimates for the KNW model given in Draper (2014) and the calibrated values for both the Commission Parameters (Langejan et al. (2014)) and the DNB. The returns are geometrically defined.

For the 'Commission Parameter'-calibration four model parameters are altered. There is a small increase in parameter $\delta_{0\pi}$ so that the expected inflation equals 2.0%. Furthermore, the risk premium on bonds ($\Lambda_{0(1)}$) is increased so that the long-term expected return on bonds is equal to the ultimate forward rate of 3.9%.[17] Finally, the parameters for the risk and volatility of equity returns are adjusted. The equity risk premium (η_S) has increased by more than 2%-points and the volatility of equity ($\sigma_{S(4)}$) by more than 1%-point. This ensures that the expectation and volatility of equity returns tend to 7% and 20% as determined by the Commission Parameters. For the 'DNB'-calibration only three parameters are altered. First the parameter $\delta_{0\pi}$ is set equal to 2.0%. And secondly, both variables that determine the unconditional price of risk (($\Lambda_{0(1)}$) and ($\Lambda_{0(2)}$) are altered to ensure that the long-term interest correspond with the UFR methodology.

It is interesting to see what the results are for the different calibration approaches. Hence we compute the moments for the stock returns and inflation using the KNW model with the adjusted model parameters. To calculate the geometric average and volatility of stock returns for the different model parameters we simulate 10.000 sample paths with each path covering 60 years, which is the same length as the simulation period for the feasibility test. For each simulation we assume the economy begins in equilibrium, which corresponds to initial state variables equal to zero. We repeat this for inflation.

In Table 11 we present the results for the model estimates presented in Draper (2014) and for both calibrated parameter sets according to the Commission Parameters and the DNB. To be complete, we add the results for the model parameters based on the updated data (Berg (2016)) and Draper (2012). Not surprisingly, we see a clear difference for the average stock returns between the model calibrated according to the Commission Parameters approach and the model calibrated according to the DNB approach. The average for the CP calibrated parameter set exceeds the predetermined expectation of 7% for equity returns. However, the reason for the Commission Parameter to increase the risk premium parameter is to compensate for the current low interest rate environment. The simulation, however, is based on an initial economy in equilibrium. In

[16] Source: www.toezicht.dnb.nl/3/50-233690.jsp

[17] Note that the parameter Λ_0 will be adjusted again in the next section when the model is actualized for the current term structure.

the next section we will discuss the actualization of the model. Furthermore we notice that the expected average of this paper is substantial lower than the other models. And also, the average inflation for all models is close to the predetermined 2%.

	Stock Returns		*Inflation*	
Source	Average	Std. Dev.	Average	Std. Dev.
Draper (2014)	5.74%	17.42%	1.84%	1.67%
Calibrated CP	7.73%	19.03%	2.01%	1.67%
Calibrated DNB	5.74%	17.42%	2.03%	1.67%
Berg (2016)	4.99%	16.99%	2.01%	1.53%
Draper (2012)	6.00%	17.16%	2.26%	1.35%

Table 11: The two left columns present the expected average and volatility of stock returns according to the KNW model based on parameter sets Draper (2014), Calibrated CP, Calibrated DNB, Berg (2016) and Draper (2012). The moments are based on 10.000 simulations with each a path equal to 60 years. The two most right columns present the average inflation and its volatility according to the aforementioned sources and simulation method.

In Figure 23, the initial term structures are given for the different models. Again we assume that the economy is in equilibrium. Unsurprisingly, the calibrated interest rates are much lower than the original term structure based on Draper (2014) as the the long-term interest rates tend to the UFR-level of 3.9%.

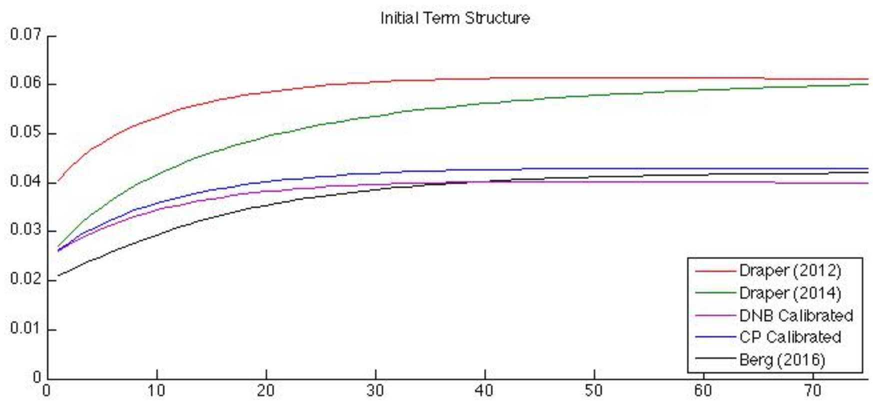

Figure 23: The estimated and calibrated term structures with state variables equal to zero.

I Actualization Scenarioset

To ensure that the interest rates in the scenarioset correlate sufficiently with the prevailing term structure, the model must be updated each quarter. The Commission Parameters proposes a method in which the initial interest rate curve of the KNW model and the average simulated curve at $t = 10$ should match as good as possible with the prevailing yield curve and the 10-years forward rate curve. To this end, two model parameters and the initial state variables are adjusted each period to fit the model to the prevailing yield curves. The DNB has, however, chosen another method in which only the initial yield curve in the scenarioset should correspond as closely as possible to the prevailing yield curve. In order to fit the model, the DNB only adjusts the initial values for the state variables. In this section we describe both actualization methods and compare the results.

The Commission Parameters explored a number of ways on how to update the model each period. They considered different 'weighting functions' (the function that needs to be optimized to obtain the best fit) and various parameters that need to be optimized. Based on their analysis, the Commission Parameters selected two 'target values' that are important when fitting the curve

1. The present term structure including the UFR component. The 'KNW'-interest rate curve including UFR on $t = 0$ should match as good as possible.
2. The expected interest rate curve over 10 years, which is determined by the forward interest rate scheme using the present term structure including UFR. The 'KNW'-interest rate curve including UFR on $t = 10$ should match as good as possible.

These target values are implemented in a weighting function to retrieve the best fit. The weighting function is equal to the sum of the squares of the differences between the two target values and two yield curves based on the KNW model. So the following function must be minimized:

$$W = \sum_{t=1}^{50} (TS_t^m - TS_t^{ext})^2 + (TL_t^m - TL_t^{ext})^2 \tag{82}$$

where TS is de initial term structure and TL the term structure after 10 years, both including de UFR component. The superscript m indicates that the value is calculated in the KNW model and ext indicates that it is imposed externally, i.e. the current term structure including the UFR component. The weighting function is determined using all maturities between 0 and 50 years with all maturities carrying the same weight. TLm is the yield curve that is determined on the average interest rates after 10 years using 3000 scenarios. The variables selected by the Commission Parameters that have to be optimized are

- $X_1(0)$ en $X_2(0)$: the start state variables at time $t = 0$
- $\Lambda_{0(1)}$ and $\Lambda_{0(2)}$: the constant part of the 'price of risk'

Hence these two model parameters and the initial state variables will be determined by an optimization process in order to fit model as good as possible. In the optimization process, no restrictions are imposed on the state variable or on the risk parameters. The Commission Parameters states that including less fit-parameters (e.g. only initial state variables) would lead to a less good fit of the initial term structure. Also including more fit-parameters from the KNW model (e.g. also Λ_1) would lead to lower stability and convergence of the fit and increases the quality of the fit only a little bit.

Using forward rates to determine the expected yield over 10 years is in line with the recovery plan methodology. However as mentioned before, this is not in line the academic literature and empirical observations. The forward rate methodology is derived from the expectations hypothesis, which is the proposition that the long-term rate is determined purely by current and future expected short-term rates. However under this assumption the presence of a risk premium is neglected. And from empirical research we know that (long-term) bonds have positive risk premia.

Forward rates are therefore not equal to market expectations of future interest rates, but rather reflect the sum of market expectations, risk premia and convexity effects. So the forward rate scheme overestimates expected future interest rates, which may result in too low expectations for the pension fund's liabilities.

In contrast to the actualization method of the Commission Parameters, the DNB only determines to initial state variables so that the model corresponds to the prevailing yield curve. The initial values for the optimization procedure are the previous discussed calibrated parameters of the KNW model. Note that the constant parameters for the 'price of risk' ($\Lambda_{0(1)}$ and $\Lambda_{0(2)}$) are already adjusted to ensure that the long-term average return on bonds tends to the ultimate forward rate of 3.9%.

Figure 24 shows the actualized interest rate curves based on the method recommended by the Commission Parameters (solid lines) and the method used by the DNB (dotted lines). Both methods are applied for the yield curves (including the UFR-component) of two periods, namely for 31 December 2013 and 30 June 2015. Note that the solid red and black lines overlap.

For the term structures at $t = 0$, we see that the DNB method fits best. For example, the initial curve based on the CP method is slightly higher for shorter maturities than the prevailing yield curve. The disadvantage of the CP method is that it is difficult to fit the initial interest rate curve as well as the 10-years forward rate curve at the same time.

For the term structures at $t = 10$ it is not surprisingly that the CP method corresponds better with the forward rate curve than the DNB method. For the curves based on 31 December 2013, the DNB actualization method generates interest rates that are on average below the forward rate curve. The interest rates of the CP method are above the interest rates of the DNB method. This may the result of the absence of a risk premium in the forward rate scheme.

For the average term structures at $t = 10$ based on the yield curve at 30 June 2015, we see something very interesting. The average interest rates after 10 years based on the DNB method are higher than the 10-years forward rates. This suggests that the expected interest rates for the KNW model exceed the market expectations. Furthermore it is important to note that the long-term interest rates we use for 30 June 2015 are already higher than normal rates due to the UFR-methodology. Hence the presented forward rate curve is higher compared to a forward rate curve derived from a normal yield curve without the UFR-component.

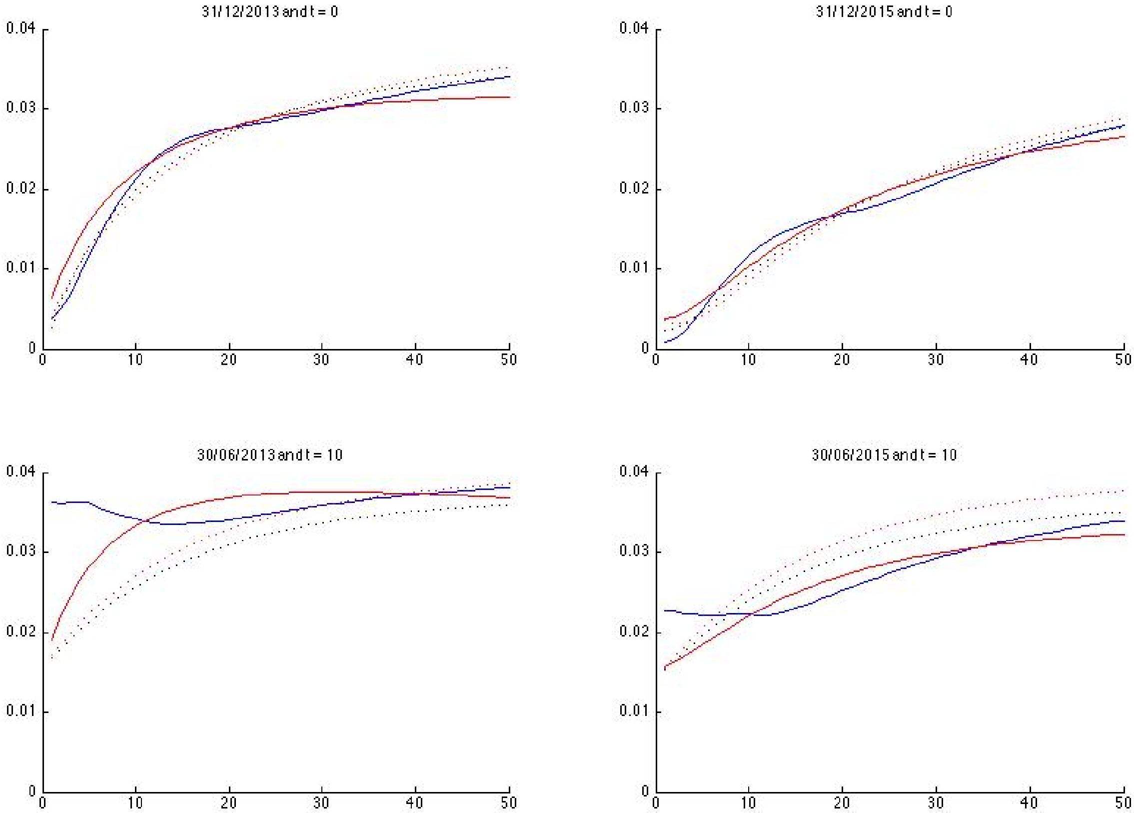

Figure 24: Graphs display the initial interest rate curves (upper graphs) and the (average) interest rate curves after 10 years (bottom graphs) for two dates, 31 December 2013 (left) and 30 June 2015 (right). For the initial values of the actualization process we use the calibrated set according to the Commission Parameters (red lines) and the calibrated set according to the DNB (black lines). The solid lines represent the actualization results based on the Commission Parameter method and the dotted lines represent the actualization results based on the DNB method. The blue lines represent the input curves for the actualization process including the initial term structure with UFR (t=0) and its implied forward rate curve (t=10).

The assumptions we make for future expected interest rates, including the calibration and actualization methods, have a large impact on the future financial position of a pension fund. In case of a too optimistic interest rate outlook, there is a risk that necessary recovery measures are delayed too long. This is particularly undesirable for pension funds with a funding shortfall. The implementation of the forward rate scheme can lead to an overestimation of the financial position of a pension fund. On the other hand, introducing market expectations to the actualization procedure can ensure that the average future interest rates for the KNW model will not become too high. An alternative may be to adjust the forward rate scheme by including the risk premium. So for example a method in which the expected yield curve is a combination of the forward yield curve and the spot yield curve.

The actualization of the model parameters has also an impact on the average stock returns and expected inflation, see Figure 25. The state variables have a direct impact on the expected inflation. For the actualization at 30 June 2015, the average inflation decreases c. 21 basis points. The equity returns in the KNW model are based on the instantaneous nominal interest rate and a constant equity risk premium. Because the premium is assumed constant, the estimated equity returns are very sensitive to the level of the short rate. Due to the low interest rates at 30 June 2015 the average stock returns in the beginning of the simulation period are lower than when the economy starts in equilibrium. For this date, both actualization methods ultimately result in an average decrease of 50 basis points for the average stock returns. Clearly this will have large implications for the future financial position of the pension fund and therefore also for the end results in the feasibility test. The motivation for the calibration and actualization of the model parameters is clear, however, the feasibility test provides an analysis over 60 years and any small change can have a big impact on the final results.

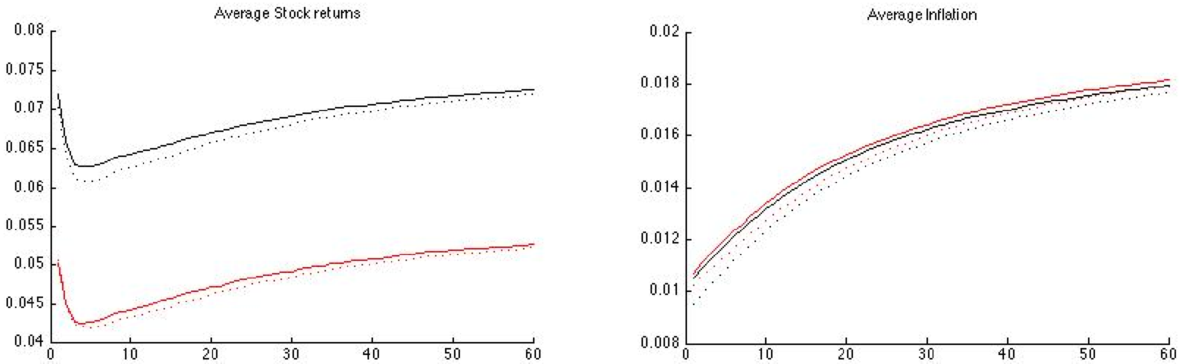

Figure 25: The average stock returns and average inflation over time after the parameters are actualised for the yield curve at 30 June 2015. The averages are geometrical defined. For the initial values of the actualisation process we use the calibrated set according to the Commission Parameters (red lines) and the calibrated set according to the DNB (black lines). The solid lines represent the actualisation results based on the Commission Parameter method and the dotted lines represent the actualisation results based on the DNB method.

J Calibration G2++ model via Kalman Filtering

In this section, we provide implementation details on the calibration of the G2++ model following the paper Park (2004). The G2++ model has five parameters that need to be estimated: α, β, γ, η and κ. Here $\alpha, \beta > 0, -1 \leqslant \kappa \leqslant 1$, and typically both α and β are greater than zero. Because the governing equations are linear, the calibration problem can naturally be framed as a Kalman filtering problem with suitable parameter update law. Therefore, we apply the Kalman filtering algorithm for model calibration, using the traditional linear Kalman filter for state propagation, and the quasi-maximum likelihood estimate to update the model parameters at each iteration.

Consider first the general discrete-time formulation. Let $x_k \in \Re^n$ denote the state vector and $y_k \in \Re^p$ the observation vector (not to be confused with x and y in the G2++ model. Denote the model parameters to be identified by the vector θ. The state and observation equations are then respectively given by

$$x_{k+1} = A_k(\theta)x_k + B_k(\theta)\omega_k + c_k(\theta) \tag{83}$$
$$y_k = D_k(\theta)x_k + e_k(\theta) + \eta_k \tag{84}$$

where $\omega_k \sim N(0, Q_k)$ with $Q_k \in \Re^{m\times m}$, $\eta_k \sim N(0, H_k)$ with $H_k \in \Re^{q\times q}$, and $A_k(\theta) \in \Re^{n\times n}$, $B_k(\theta) \in \Re^{m\times m}$, $D_k(\theta) \in \Re^{p\times n}$, $c_k(\theta) \in \Re^n$, $e_k(\theta) \in \Re^p$ are given time-varying vectors and matrices of given dimension. Assuming the time series ranges over the index $t = 1, ..., N$, the filtering procedure is given as follows (the θ argument is suppressed for notational convenience):

- **Prediction**

$$x_{k|k-1} = A_k x_{k-1} + c_k \tag{85}$$
$$P_{k|k-1} = A_{k-1}P_{k-1}A_{k-1}^T + B_{k-1}Q_{k-1}B_{k-1}^T \tag{86}$$

- **Update**

$$\upsilon_k = y_k - D_k x_{k|k-1} - e_k \tag{87}$$
$$F_k = D_k P_{k|k-1} D_k^T + H_k \tag{88}$$
$$x_k = x_{k|k-1} + P_{k|k-1} D_k^T F_k^{-1} \upsilon_k \tag{89}$$
$$P_k = P_{k|k-1} - P_{k|k-1} D_k^T F_k^{-1} D_k P_{k|k-1} \tag{90}$$

- **Parameter estimation**

$$\theta = \arg\max \left\{ -\tfrac{pN}{2}\log 2\pi - \tfrac{1}{2}\textstyle\sum_{k=1}^N \log|F_k| - \tfrac{1}{2}\sum_{k=1}^N \upsilon_k^T F_k^{-1} \upsilon_k \right\} \tag{91}$$

  where $|\cdot|$ denotes determinant. Note that both υ_k and F_k are functions of θ in the above equation.

To begin the iteration, initial values for x_k and P_k are required. One populair choice is

$$x_0 = (I - A_0)^{-1}c_0 \tag{92}$$
$$P_0 = (I - A_0A_0^T)^{-1}Q_0 \tag{93}$$

The value for x_0 corresponds to the conditional mean, while P_0 corresponds to an approximation of the conditional variance.

In the event that matrix inversion is a significant computational burden, the inverse F_k^{-1} can be computed more easily using the Sherman-Morrison-Woodbury formula:

$$F_k^{-1} = H_k^{-1} - H_k^{-1}D_k\left(D_k^T H_k^{-1} D_k + P_{k|k-1}^{-1}\right)^{-1} D_k^T H_k^{-1} \tag{94}$$

With the above formula F_k^{-1} can effectively be obtained by simple inversion of the 2×2 matrix enclosed in parentheses. In practice H_k will also be assumed diagonal, further simplifying the above calculation.

For the G2++ formulation we have

$$A_k = \begin{bmatrix} e^{-\alpha \Delta t} & 0 \\ 0 & e^{-\beta \Delta t} \end{bmatrix} \tag{95}$$

$$B_k = \begin{bmatrix} \gamma \sqrt{\frac{1-e^{-2\alpha\Delta t}}{2\alpha}} & 0 \\ 0 & \eta \sqrt{\frac{1-e^{-2\beta\Delta t}}{2\beta}} \end{bmatrix} \tag{96}$$

and $c_k = 0$. The matrix D_k is given by

$$D_k = \begin{bmatrix} \frac{1-e^{-\alpha\tau_1}}{\alpha\tau_1} & \frac{1-e^{-\beta\tau_1}}{\beta\tau_1} \\ \vdots & \vdots \\ \frac{1-e^{-\alpha\tau_p}}{\alpha\tau_p} & \frac{1-e^{-\beta\tau_p}}{\beta\tau_p} \end{bmatrix} \tag{97}$$

while the i-th element of $e_k \in \Re^p$ is given by

$$\begin{aligned} e_{k,i} = &\frac{t_k + \tau_i}{\tau_i} R(0, t_k + \tau_i) - \frac{t_k}{\tau_i} R(0, t_k) \\ &- \frac{1}{2\tau_i} \left[V(t_k, t_k + \tau_i) - V(0, t_k + \tau_i) + V(0, t_k) \right] \end{aligned} \tag{98}$$

The covariance matrix $Q_t \in \Re^{2\times 2}$ is given by

$$Q_t = \begin{bmatrix} 1 & \kappa \\ \kappa & 1 \end{bmatrix} \tag{99}$$

while $H_t \in \Re^{p\times p}$ is assumed to be of the form

$$H_t = \text{Diag} \begin{bmatrix} \sigma_h^2 \lambda & \dots & \sigma_h^2 \lambda^p \end{bmatrix} \tag{100}$$

where λ is a user-specific weighting factor between 0 and 1, and σ_h is a scalar constant to be identified. The λ factor reflects the difference in observed spot rates for various maturities: spot rates of longer maturities, for example, may exhibit lower volatility than those of shorter maturities. A value of λ close to one implies that the volatility difference is minimal, whereas setting λ close to zero indicates a large difference in short-term and long-term spot rate volatilities. Setting λ greater than one has the effect of placing greater longer maturity spot rate observations. Finally, the initial value of the covariance matrix P_0 is set to a user-specified positive definite diagonal matrix.