Netspar

Network for Studies on Pensions, Aging and Retirement

Netspar THESES

Gerke Bouma

Sufficiency of the FTK Solvability Assessment

MSc Thesis 2009

Sufficiency of the FTK Solvability Assessment

Gerke Bouma
(Student Number: S1475797)

Master Thesis Econometrics
University of Groningen

May, 2009

Supervisor: L. Spierdijk

Abstract

After a brief discussion of the FTK and its background, I analyze the performance of the required reserve that results from the FTK solvability assessment. This is done by simulating the change in the value of the assets and liabilities of pension funds by means of a GARCH-BEKK model that is developed in the text, as well as an existing ALM model developed by Watson Wyatt. It is shown that in many cases the required reserve performs below its intended reliability level. This is caused by an underestimation of interest risk.

Foreword and thanks

I wrote this master thesis as a part of an internship at Watson Wyatt. Not only did this allow me to benefit from the extensive knowledge and expertise of the people at Watson Wyatt, but it also meant I was able to experience first hand how pension funds deal with the FTK in practice. During the writing of this thesis I learned far more than I ever expected. When I started my research, I knew little to nothing about pension regulations and I never even heard about the FTK. Now, I am approached on a regular basis by my colleagues at Watson Wyatt with questions about details of the FTK solvability assessment. To my own surprise, more often than not I am capable of answering those questions. I highly enjoyed my internship, and my thesis evolved far beyond my initial expectations. However, I doubt it would have been such a success without the support of the many people that lend their assistance during the process. I want to take this opportunity to thank them for their efforts.

I specifically want to thank my supervisors Laura Spierdijk and Michel van Ierssel for their guidance, input and support. My appreciation also goes out to Klaas Knot for taking the time and effort to read and assess this thesis. I also want to thank Saskia Donker, Ralph van Daalen, Pieter Wittekoek, Elmer van de Fliert, Josje Wijckmans, Esther Lamerikx, Djien Kwik, Morena Varga, Hugo Nieuwenhuijse, Ran Chen and all my other colleagues at Watson Wyatt for their time and helpfulness. My internship at Watson Wyatt was a great experience thanks to them. Last but certainly not least I want to thank my dear Marlou for her continued patience and support during the entire process.

Contents

1 Introduction

Pension funds are something that most people with a steady job in the Netherlands come into contact with. If you ask one of those people what the function of a pension fund is, he or she will probably give you an answer along the lines of “Each month I pay an amount of money, and in return I get an income when I retire”. While highly simplified, that is the core of what a pension fund does. It collects premiums payed by both employers and employees and provides those employees with an income in the form of a pension when they retire. The premiums are invested, and as such form the assets used to finance the pensions.

The Dutch government requires pension funds to maintain a reserve on top of the nominal value of the pension claims they guarantee. The size of this buffer depends among other things on the type of investments the pension fund makes, the demographics of their participants, and the type of pension scheme it offers. The buffer is determined by a solvability assessment which is part of a set of pension regulations called the *het Financieel Toetsings Kader* (FTK). Usually this buffer is reported in the form of a funding rate, which simply describes what the ratio between the assets and liabilities of the fund should be. The goal of this assessment is to determine a required reserve that should be sufficient to keep the funding rate of the pension fund above a minimum level of 105% within the next year with at least 97.5% certainty.

As has become painfully apparent in the last half of 2008 and start of 2009, investments are not without risks. Investments made by pension funds are not an exception to that. At the start of 2008 only nine out of the 416 Dutch pension funds had a funding rate below 105%, with 175 of the funds having a funding rate above the average required funding rate of 130%.[1] At the end of 2008 these numbers drastically changed; approximatly 300 funds were showing a funding rate below the minimal required 105%, with roughly 65 more finding themselves below their required funding rate. The average funding rate among Dutch pension funds at that time was 95%.[2]

This raises some questions about the sufficiency of the reserves required by the FTK assessments. The objective of this thesis is to investigate if the FTK required reserve is indeed sufficient to keep the funding rate of a pension fund above 105% with 97.5% certainty. This will give an indication of the general reliability of the required reserve.

[1] Source: *Kwartaalbericht december 2008*, De Nederlandsche Bank.
[2] Source: *Kwartaalbericht maart 2009*, De Nederlandsche Bank.

Moreover, it will serve as an indication of whether or not the current problems the Dutch pension funds experience are part of the 2.5% of scenarios that will lead to a funding deficit.

The reliability of the required reserve that the FTK solvability assessment yields is of great importance, not only to pension funds but also to their participants whom they promise an income after retirement. It is also relevant to actuarial firms like Watson Wyatt, who play a role in advising pension funds on the subject of their reserves, pension premiums and investments. Until now the discussion in the literature has been mainly about the methods used in the FTK. See for example Keating (2006) and the response of Siegman (2006). In this paper I will not focus the discussion on the methods of the FTK, but on the performance of the reserve it suggests. To my knowledge no other research into the quality of this reserve has been published.

To determine the reliability of the FTK required reserve, I will use historical data to develop models for the investment categories used in the FTK. Using these models, I will simulate how the funding rate suggested by the FTK solvability assessment changes in a one year timespan. I will try to determine if the funding rate of a pension fund stays above 105% with at least 97.5% certainty if it meets the required reserve/funding rate at the beginning of the year. I will do this for several different (artificial) pension funds, several investment portfolios and changing assumptions about the state of the economy at the start of the simulated year. Moreover, I will compare the results of my model to those of a model constructed by Watson Wyatt.

I will use many statistical and econometric techniques to estimate these models and analyze the results of these models. Due to the complexity of some of these methods and techniques, I cannot explain all the theory behind the analysis in detail without losing readability. While many concepts will be briefly explained in the text, I will assume that the reader has a significant background in econometrics or statistics, equivalent to that of a graduate student in econometrics. Some concepts that are not explained in the text will be explained in the appendices. I suggest that the reader who is less interested in the process but more in the results skips the more technical sections. This thesis is partially written on basis of an internship with Watson Wyatt, so where possible I took advantage of the knowledge and expertise of my colleagues at Watson Wyatt.

In section 2 I will discuss the history and basics behind pensions and the FTK. Section 3 contains a more detailed discussion of the FTK solvability assessment. Section 4 describes the data used for the model. In section 5 I will discuss the theory behind the model used in the analysis. The model itself and its estimation will be discussed in section 6. Section 7 describes the assumptions used for the simulation and analysis. Section 8 contains the analysis done with the estimated model, and its results. Section 9 considers a different model, namely the ALM model developed by Watson Wyatt and the results of the analysis performed with that model can be found in section 10. Section 11 lists my final conclusions. Section 12 contains several recommendations for further research. This is followed by the references and appendices.

2 Background of the FTK

This section contains a general reading about pensions, pension funds and how the rules surrounding pension funds came to be. I recommend that anyone who is relatively new to the field of actuarial science or has little knowledge about pension regulations reads this section before continuing with the rest of the paper. This should help the reader gain some insight in the issues surrounding the subject of this thesis. Readers already familiar with these concepts may want to skip this section in its entirety.

2.1 The Dutch pension system

To be able to fully understand this paper, it is important to have a basic understanding of the Dutch pension system. This system is based on three 'pillars', which together provide the Dutch population with an income after retirement. The pillars consist of a general state pension, a supplementary pension built up as a part of employment terms, and financial products serving as income which anyone can buy from insurers.

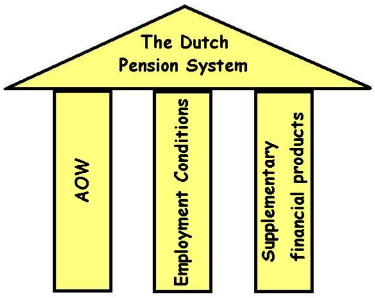


Figure 1: The three pillars of the Dutch pension system

The First Pillar

The first pillar is a state pension based on the *Algemene Ouderdoms Wet* (AOW), a law meant to provide all Dutch residents above the age of 65 a minimum income. The state pension is for anyone who lived and/or worked in the Netherlands between the age of 15 and 65. For each year a person between the age of 15 and 65 works or lives in the Netherlands, they accumulate a claim of 2% of this state pension. Exceptions might be applicable for people who work abroad during the years they live in the Netherlands or people who work in the Netherlands but live abroad.[3] The size of the pension received at age 65 depends on some additional factors, like whether or not the person in question is married or has young children. The default income is equivalent to 70% of the minimum wage for singles and 50% of the minimum wage for people with a partner.

Lately the state pension has been under some pressure, due to its 'pay-as-you-go' system[4]. Since the average age of the Dutch population is rising, this means the ratio of retirees versus the working population is going up. There are some concerns that the system might become too expensive over time, and research in the continued viability of the system is being done. Recently suggestions have been made to increase the age of retirement from an age of 65 up to an age of 67.

The Second Pillar

The second pillar consists of supplementary pensions as part of a person's employment conditions. While there is no legal obligation for Dutch employers to offer their employees a pension scheme, over 90%[5] of the Dutch working population participates in a pension plan. The pension schemes in the second pillar are almost always related to the first pillar. The yearly accumulated pension claim depends on the pensionable salary.[6] This pensionable salary is determined by lowering the salary with an offset. This offset has a fiscal minimum, which is related to the state pension. The reasoning behind the offset is that it is not necessary to accumulate pension in the second pillar over your entire income; you are already accumulating a part of your pension in the first pillar.

The Third Pillar

Any type of pension that is not part of the AOW or part of working conditions is part of the third pillar. It consists out of all financial products people can buy as a supplement to their pension. Life annuities are a common example. This pillar has nothing to do with pension funds and is of no interest to the subject of this thesis.

[3]More details about the AOW can be obtained from the *Sociale Verzekeringsbank*, see for example www.svb.nl.

[4]Dutch: *Omslagstelsel*

[5]Based on *Kwartaalbericht December 2008*, De Nederlandsche Bank and *Tabel Arbeidsdeelname 15 jaar of ouder*, Centraal Bureau voor de Statistiek.

[6]The exception to this are defined value schemes. However this kind of pension is extremely rare. See also section 2.2.

2.2 Types of Pension Schemes

The Dutch pension law distinguishes 3 different types of pension schemes.

Defined Contribution

Probably the simplest type is the defined contribution (DC) scheme. In this kind of pension scheme the pension premium is invested, and the capital accumulated at the age of retirement is then used to provide a pension, for example by using the capital to buy a life annuity. The size of the pension depends mainly on the capital that is available at the age of retirement, as well as factors such as the interest at that moment in time. This means that any risks involved in this type of pension scheme are a concern for the one who will receive the pension. The FTK only deals with the risks concerning pension funds. As such DC schemes are not of interest for this thesis. I will not go into any more detail about this type of pensions.[7]

Defined Benefit

Over 90%[8] of the pension schemes in the Netherlands are so called defined benefit (DB) schemes. Well known examples of DB schemes are the average-pay and final-pay schemes. Each year an employee participates in a DB scheme, he accumulates a certain amount of pension he is sure to receive at the age of retirement. Whether or not the accumulated claims are always compensated for inflation effects differs between schemes, but the currently accumulated pension is guaranteed. This means that unlike in the case of a DC scheme, there are barely any risks[9] involved for the participant in terms of the nominal value of the pension. The risks involved with a DB scheme are all on the account of the party who is supposed to execute the pension. It is these risks that are of interest in this paper and to the FTK in general. Later in this paper I will elaborate more on what these risks are and their relation with the FTK.

[7] For additional information about defined contribution schemes and the differences between this and other types of pension schemes, see for example Schulting (2008).

[8] Source: *Kwartaalbericht 2008*, De Nederlandsche Bank.

[9] Under exceptional circumstances like major financial problems on the side of the pension fund, claims may be cut as a last resort. Especially with the current state of the economy the possibility of such cuts has increased.

Defined Value

Finally there is a third type of pension scheme defined in the Dutch pension law: Defined value schemes. This type of contract is similar to a DB scheme in the sense that the pension claims accumulated are guaranteed. However, unlike the other two types of pension schemes, the size of the pension claim is independent of the salary of the participant. Anyone participating in a defined value scheme accumulates exactly the same pension each year, and also pays exactly the same premium. Even though this type of pension scheme has some advantages, especially in terms of simplicity, there are many obvious disadvantages. If the pensions are too high, employees with low incomes might have to give up a (too) substantial amount of income in order to participate. If the pensions are too low, they might not be sufficient to cover the income needs of employees with high incomes once they retire. In the past defined value schemes could be found in several industries in the Netherlands, but this type of pension scheme is becoming more and more rare. An example of a pension fund still implementing a defined value scheme is *Stichting Bedrijfstakpensioenfonds Herwinning Grondstoffen*, a pension fund that executes this scheme for 170 companies specialized in recycling or waste management. In terms of FTK regulations defined value schemes are treated similar to DB schemes, therefore I will not distinguish between these types of pension schemes in this thesis.

2.3 Pension Executants

Pensions are either executed by pension funds or by insurers. However, the FTK regulations are different for these types of providers. The reason for this is that pension funds and insurers have different characteristics. In this paper I only look at the risks concerning pension funds and how the FTK requires those pension funds how to handle those risks. Therefore I will not go into any detail on the regulations for insurers. For more information about the differences between insurers and pension funds see for example Hoekert (2007).

There are several reasons pension funds exist. For example it is determined by law that the assets allocated for pension schemes need to be separated from the assets of the employer. Pension funds offer a way to make this separation, since they are independent organizations. This prevents the employer from taking assets from the pension fund in times of financial trouble.

A second important reason for the existence of pension funds is the reduction of risks. This happens in several ways. Pension funds generally handle the pensions of several hundred up to millions of participants. While things like disability and mortality tables can never accurately describe what will happen to an individual, they are generally fairly accurate when looking at a group of individuals.[10] This is described in mathematics by the law of large numbers. Because of this convergence to averages, the cash flows insured by pensions become more predictable and therefore the insurance risk decreases.

Another way pension funds reduce risk is because of their continuity. Since pension funds generally have participants from many age groups, this offers the possibility for pension funds to use the high returns in good periods to compensate for bad returns in less fortunate times. Instead of significantly reducing premiums during good times, the surplus capital can help to guarantee that participants who retire during bad times are able to receive a pension as well. This is called the solidarity principle.

The solidarity principle also manifests itself in a different manner; all participants of a pension scheme usually pay an (relatively) equal part of their income as premium. However, the cost of obtaining €100 lifelong pension starting at age 65 is higher for someone aged 60 than it is for someone aged 30. The reason for this is that younger people have many years ahead of them before retirement, so the present value of the cash streams that form their pension is low due to a large discount factor. On top of that they have a smaller chance of actually reaching the retirement age. Probabilities of survival, disability and withdrawal are major factors in the costs of a pension. Similarly there is a difference between the costs of obtaining a pension for men and women. However, everyone pays a similar premium in terms of a fixed percentage of their income. This means that younger participants pay too much for the claims they accumulate compared to the actuarial value of those claims. In the meanwhile, older participants pay less than they should given the actuarial value of their claims. This is not a problem as long as the younger participants can safely assume the system will stay the same in the future. If so, they may be paying too much now, but their advantage will come when they are older. Again this means the fund relies on the solidarity of its participants.

2.4 Pension Financing

The financing of DB schemes is part of the core of the FTK. Pension schemes are financed by means of a premium, which consists of a premium payed by the employee and a premium payed by the employer. The size of these premiums depends on the pension scheme,

[10] Assuming the tables are derived from data that represents people with similar characteristics as the group in question.

but typically the employer pays a substantial larger sum than the employee. Employers pay about 80% of the premium[11] on average, but this varies heavily between pension schemes. Employer contributions of 50% are not unusual, but some employers even pay the full 100% of the premium, effectively making the pension plan free for their employees. Determining what the size of these premiums should be is not easy. Obviously, both employers and employees prefer to pay as little as possible. However, the premiums do need to be sufficiently large to cover the pension claims of all participants once they retire. On the other hand, there is little use in paying high premiums if a lower premium will suffice. A lot of actuarial and economic factors in the form of for example life expectancies, mortality trends, disability probabilities, discount rates and rates of return are used to determine the premium, but there are also laws and regulations that set conditions on the size of the premium. Any basic text in actuarial science will give a detailed description of the theory behind these factors.[12]

The premiums form the basis for the assets of the pension fund, which uses those assets to finance the pensions of the participants. The assets are usually invested in the form of stocks, bonds and other financial products. However, these investments are not without risk, which means that it is unsure what will happen to the assets of the pension fund over time. Also the value of the claims that the assets will actually have to cover is unsure; participants might never reach their age of retirements, or live much longer than expected. It is for these reasons that pension fund supervision exists.

2.5 Pension Fund Supervision

Pension premiums are collected now, to finance a pension claim that needs to be payed many years from now, and a lot can happen during that time. Moreover, a great deal of money is invested in pension funds. At the start of 2008 the Dutch pensions valued nearly 760 billion euro.[13] This explains why it is important for pension funds to handle their assets in a responsible manner. The Dutch government made financial regulations for pension funds to ensure this is done. The government also put supervision in place to make sure pension funds abide by these regulations.

Pensions can be traced back to medieval times when preachers, soldiers and certain employees of the government received a pension supplied by the government itself. In 1836

[11] See for example E.P. Davis et all, 2007.

[12] See for example Gerber, 1997.

[13] Source: *Kwartaalbericht december 2008*, De Nederlandsche Bank.

a pension fund[14] was founded to execute disability pensions for employees of the Dutch government. The first pension funds founded by firms followed around 1880. These and other pension funds went unsupervised for a long time, partly because the government did not yet value (self-)supervision as highly as it does now. The first rules pension funds had to abide by were listed in a royal decree from 1908. One of the most important implications of this decree was that the assets of the pension funds had to be separated from the assets of the employer. This ensured that a pension fund could continue to exist even if the employer went bankrupt.

It was not until 1936, when a pension fund did not have the financial assets to pay the claims that year, that the need for a full pension legislation became apparent. The result was the pension and savings law[15] that took effect in 1952. While this law included guidelines for pension funds, these were not very explicit. The guidelines contained statements like ‘prudent actuarial calculations’, ‘solid investment strategies’ and ‘reserves adequate to account for economic and actuarial risks’. The interpretation of these concepts was open to some discussion.

The supervision over the pension funds was from that point on in the hands of the *Verzekeringskamer* (VK) which was an independent organization founded by the Dutch government. The reason for this independence was twofold. First of all the goal was to make the gap between the supervisor and the insurance- and pension branches as small as possible. This was to lead to a better mutual trust and sense of authority between the supervisor and supervised. Second, the independence would prevent political agendas from affecting the supervision.

2.5.1 The Actuarial Principles for Pension funds (APP)

During the 90’s returns on the stock markets were high, which led a lot of pension funds to become more interested in these riskier types of investments. In order to keep these risks under control and to construct a more uniform way of safeguarding against low returns, the actuarial pension fund principles[16] (henceforth APP) were devised and put into effect starting in the fiscal year of 1997.

While relative to the current models the APP contained just some simple instructions, they contained the first explicit guidelines by which pension funds had to assess their

[14] This fund was named *het Algemeen Burgerlijk Pensioenfonds*, which is now by far the largest pension fund in the Netherlands, executing the pensions for all employees of the Dutch government.

[15] Dutch: Pensioen- en Spaarfondsenwet.

[16] Dutch: Actuariële principes voor pensioenfondsen.

financial status. The main importance of the APP was that pension funds had to perform a solvability assessment every year. I will list the most important requirements of that assessment, paraphrased from the original publication of these rules.

- **Only current liabilities and assets are relevant for the assessment. Claims that will be accumulated in the future will be disregarded, as well as any income from premiums.** In other words, the APP solvability assessment was a snapshot of the current investment portfolio, assets and liabilities.

- **The present value of the liabilities is to be determined on prudent actuarial bases, with a maximum discount rate of 4%.** Note that this discount rate was not directly linked to the interest rates on the market, and therefore by design lacked the volatility of the market rates. However, this also meant it was not neccesarily representative for the market value of the liabilities.

- **A reserve is to be kept for administration costs and any other operational costs involved in receiving premiums and paying out the claims.**

- **Investments are to be valued at present value, based on current market prices and yield curves.** This meant that expected returns on investments were irrelevant to the solvability assessment; only the assets a pension fund currently had were taken into count.

- **Funds will keep a resistance reserve[17] on top of the reserve equal to the value of the liabilities. This resistance reserve will be determined by estimating the effect of a drop in value of *all* current assets. This drop will be determined by reasonable historical estimates.** This resistance reserve was meant to keep the pension fund out of short term financial trouble.

- **Indexation risk is disregarded.** While officially indexation risk did not have to be part of the solvability assessment, some actuarial firms (e.g. Watson Wyatt) considered this to be of too great importance to leave it out, thereby making the assessment harsher than it legally needed to be. Obviously, these extra demands were taken into account when giving the final judgment about the financial status of pension funds.

[17] Dutch: Weerstandsvermogen.

The APP were updated near the end of 2002, by making parts of the guidelines above more explicit. For example, funds had to keep a 5% reserve for general risks in addition to the resistance reserve described above. Any fund without a funding rate of at least 105% (meaning they should have assets with a total value of at least 105% of the present value of the liabilities) was now said to have a funding-deficit. Funds with a funding-deficit would be put under close supervision by the VK. Indexation risks were now also taken into account if they were (legally) unconditional. Moreover, the way the resistance reserve over bonds and risky assets should be determined was specified in more detail.

In 2001 the VK was renamed *de Pensioen- en Verzekeringskamer* (PVK) to make the name and function of the organization clearer. In October 2004 the PVK and Dutch National Bank (DNB) merged since the supervisional role of both organizations over financial institutions more and more overlapped. The PVK and DNB combined continued as the Dutch National Bank.[18]

The solvability assessment resulting from the APP was regarded to be lacking in several different fields. First of all, assets were valued at market rates while liabilities were not. Also the difference between long term and short term risks were not clear in the APP regulations. One of the main goals of the FTK was improve risk management, both by increasing awareness of the different kinds of risks as well as giving better insight into the size of these risks.

2.5.2 The FTK regulations

The FTK regulations were officially put into effect in the fiscal year of 2007. So while some pension funds and actuarial consultants were already using the FTK regulations in their preliminary form as early as the fiscal year of 2004, it is safe to say that the FTK is still a rather new concept. This especially holds true considering that based on experiences from all parties involved, parts of the preliminary FTK were changed before its implementation. The FTK assessment in its current form consists of three parts: The minimum assessment, the solvability assessment and the continuity analysis.

The Minimum Assessment

The minimum assessment is a familiar legacy of the APP assessment. Pension funds need

[18]Dutch: De Nederlandsche Bank.

to keep a minimal funding rate of about 105% of their liabilites.[19] Any fund that drops below this threshold is said to have a funding deficit. Funds with such a deficit are required to come up with a detailed short term strategy to regain their minimal funding rate within a maximum timespan of 36 months.[20] Since such strategies often require drastic measures, pension funds tend to do whatever they can to avoid this situation.

The Solvability Assessment

The solvability assessment has a similar goal as the APP assessment before. A required funding rate of the liabilities is determined based mostly on the properties and types of the investments, but also on the size and characteristics of the pension fund itself. If a fund is below the required threshold, it is said to have a reserve deficit. In this case funds need to come up with a long term strategy that will get them above the required funding rate within a timespan of 15 years. Since this strategy is long term, it might be that it does not have to be different from whatever strategy the fund is currently implementing. In order to determine what should (or should not) be changed, the continuity analysis is an important tool.

The Continuity Analysis

The continuity analysis gives an overview of what is likely to happen to the financial status of the pension fund over the course of the next 15 years, using the current strategies and policies. So while the minimum assessment and solvability assessment focus on the short term risks and financial position, the continuity analysis focuses on the long term. Due to this long term nature, under normal circumstances the continuity analysis only has to be performed once every three years, contrary to the other two parts of the FTK assessment that have to be performed on a yearly basis.

The three parts of the FTK provide a nice synergy. The current unrest on the stock markets and the changes in the interest rates have led the funding rates of pension funds to drop significantly. This means that the reserves of many pension funds are too small to maintain the required reserve, and reserve deficits (and even funding deficits) are not

[19] The exact percentage might differ slightly between pension funds.

[20] As of March of 2009 this timespan can be increased by DNB to a maximum of 60 months for funds that meet the conditions to be eligable for this exception.

uncommon. However, a recent continuity analysis may show that the strategies of the pension fund should be solid enough to regain a healthy financial position within 15 years. If so, this goes a long way to proving that there is no need for huge changes in the investment strategy, indexation policies or premiums.

Since the main subject of this thesis is the FTK solvability assessment, I will not discuss the minimum assessment or the continuity analysis any further. Therefore, whenever from this point on I refer to the 'FTK assessment' or 'FTK model', I always refer to the solvability assessment of the FTK unless explicitly stated otherwise. In the next section I will discuss the FTK solvability assessment in more detail.

3 The FTK Solvability Assessment

In this section I will explain the FTK solvability assessment in more detail. I will explain the risks taken into account in the assessment, and how the required reserve that results from the assessment is calculated. Since the FTK solvability assessment is the core of this thesis, I recommend that everyone who is not fully familiar with this subject reads this section carefully. Doing so should help both in understanding the remaining parts of this paper, as well as offering some insight into the results.

The logic behind the solvability assessment is straightforward. The assets of a pension fund should have a higher value than its liabilities, to assure that the participants can still receive their pensions in the future even if the returns on the funds' investments are bad during a short period. To be more precise: The assets should include a large enough reserve to keep the funding rate of the fund above 105% one year from now, with 97.5% certainty. Usually the level of assets required to achieve this goal is reported in terms of a required funding rate. The reason for reporting the required funding rate instead of the required reserve (while these terms obviously have a 1-to-1 relation at the moment of calculation) is the emphasis on the fact that not only changes in the assets are of importance, but also changes in the liabilities. If a drop in value of the assets is met by the same relative drop in value of the liabilities, no harm is done: The funding rate remains the same.

In order to determine the required funding rate, the solvability assessment looks at several typical risks a pension fund is susceptible to. These risks are labeled S_1 to S_6. In the remainder of this section I will discuss each of these risks in detail, starting out with a brief description of the risks followed by the calculations performed in the solvability assessment.

S_1: Interest Risk

Interest risk is caused by changes in the interest yield curve. Most pension funds invest a large part of their assets in bonds. Since bonds and pensions are cash flows with a long term nature, even a small change in the interest curve can lead to significant changes in the value of the assets and liabilities of a pension fund. The net result of these changes is called the interest risk. The amount of interest risk depends on the relative changes in the value of the bonds and liabilities. For example, an increase in the interest yield curve will lead to a decrease in the value of both bonds and liabilities. Whether or not this leads to a positive or negative change in the funding rate depends on the size of both these changes.

A way to measure the impact of a small change in the interest yield curve, is the duration of the underlying cash flows. The duration of a series of cash flows tell us approximately how severely the present value of those cash flows will react to a change in the interest.

$$\text{Factor change in present value} \approx \left(\frac{1+i}{1+i*F}\right)^D$$

Here i is the interest rate corresponding to the duration, F is the factor with which this interest rate changes, and D is the duration of the assets or liabilities. As can be seen in equation 1, the change in the present value of cash flows depends on the size of the duration. In the past it was common for pension funds to invest in bonds and similar assets with an average duration of less than 10 years. Since most funds have liabilities with significantly higher durations, this exposed them to severe interest risks. The FTK solvability assessment made funds aware of this fact, which has led several pension funds to partially or fully cover their interest risks.

As mentioned before, a change in interest can also lead to a positive result. In case interest rates go up, liabilities and bonds will decrease in value. If the decrease in value of the liabilities is much larger than that of the bonds, the net result will be an increase in the funding rate of the pension fund. For this reason pension funds may want to purposely subject themselves to interest risk during times interest rates are expected to go up.

The value of S_1 is calculated in a couple of steps. First the values ($V_L,\ V_{PB},\ V_{GB}$) and the durations ($D_L,\ D_{PB},\ D_{GB}$) of respectively the liabilities, private bonds[21] and government bonds of the pension fund are calculated. Then the interest rate corresponding to the durations ($i_{D_{\{L,PB,GB\}}}$) is determined from the interest yield curve. The duration is also used to select a set of factors ($f_{D_{\{L,PB,GB\}},p},\ f_{D_{\{L,PB,GB\}},n}$) from the parameters of the FTK assessment. These factors are used to represent changes in the interest rate.

Finally S_1 is calculated by first calculating the net effect (Δ_p, Δ_n) of a positive and a negative change of the interest rate, and then taking the maximum (worst case) of both results.

$$\begin{aligned}\Delta_p &= V_L*(1-(\frac{1+i_{D_L}}{1+f_{D_L,p}*i_{D_L}})^{D_L}) - V_{PB}*(1-(\frac{1+i_{D_{PB}}}{1+f_{D_{PB},p}*i_{D_{PB}}})^{D_{PB}}) \\ &\quad -V_{GB}*(1-(\frac{1+i_{D_{GB}}}{1+f_{D_{GB},p}*i_{D_{GB}}})^{D_{GB}})\end{aligned}$$

[21]Also known as credits.

$$\begin{aligned}
\Delta_n &= V_L * (1 - (\frac{1 + i_{D_L}}{1 + f_{D_L,n} * i_{D_L}})^{D_L}) - V_{PB} * (1 - (\frac{1 + i_{D_{PB}}}{1 + f_{D_{PB},n} * i_{D_{PB}}})^{D_{PB}}) \\
&\quad - V_{GB} * (1 - (\frac{1 + i_{D_{GB}}}{1 + f_{D_{GB},n} * i_{D_{GB}}})^{D_{GB}}) \\
S_1 &= \max(\Delta_p,\ \Delta_n)
\end{aligned}$$

It is worth noting that in some cases this calculation will yield a negative number, meaning that any interest change will yield an increase in the funding rate of the pension fund. If this is the case than S_1 will contribute a *reduction* to the required funding rate.

S_2: Risk on stocks and similarly risky assets

Especially with the latest developments on the stock markets it will not come as a surprise that the value of stocks and similar assets can be rather volatile. S_2 measures the drop in value that could occur in a worst case scenario event. Within S_2 investments are split up in several categories each with a different drop in value. A heavy positive correlation of 0.75 between these categories is assumed. The risky assets are divided in the following categories:

- **Mature markets and indirect real estate.** Mature markets consist of stocks from well-developed firms and industries, listed on the stock exchange. International examples of these stocks are the firms listed in the MSCI world index. Dutch examples would be the firms listed in the AEX index. Indirect real estate is investing in real estate by means of a third party, which invests your money on your behalf. Often these organizations are quoted on the stock market, which explains why these investments are treated similar to stocks.

  A drop in value of 25% (Δ_1) is taken into account for these investments.

- **Emerging Markets.** Unlike mature markets, these stocks consist of firms that are fairly new and/or part of a recently developed industry. Since the prospectives of these firms are often unsure, these stocks are typically more volatile (thus riskier) than those of more established firms and industries. This means that a high reserve is required for investments made in these markets.

  A drop in value of 35% (Δ_2) is taken into account for these investments.

- **Private Equity.** Private equity consists out of stocks that are not listed on any stock exchange. The firms behind these stocks do not necessarily have a high risk profile.

However private equity lacks the liquidity of stocks listed on the stock exchange, which can cause problems when the investor wants to liquidate these investments.

A drop in value of 30% (Δ_3) is taken into account for these investments.

- **Direct real estate.** These are investments in property or land. Compared to stocks real estate is generally viewed as a relatively safe investment. This is partly because the value of real estate tends not to be very volatile.

  A drop in value of 15% (Δ_4) is taken into account for these investments.

The Δ's are easily calculated by taking the value of the assets in the corresponding category, and multiplying that value with the percentage mentioned above. S_2 is not simply the sum of these four results. Instead a positive correlation ρ of 75% between every category is taken into account. This is done in the following way:

$$S_2 = \sqrt{\Delta_1^2 + \Delta_2^2 + \Delta_3^2 + \Delta_4^2 + 2\rho(\Delta_1\Delta_2 + \Delta_1\Delta_3 + \Delta_1\Delta_4 + \Delta_2\Delta_3 + \Delta_2\Delta_4 + \Delta_3\Delta_4)}$$

S_3: Currency risk.

Part of the investments may be valued in foreign currency, or the investment itself may simply *be* foreign currency. However, the pension claims are valued in euros. This means that if the value of the euro increase relative to the currency in which was invested, the value of those investments will drop. This is called currency risk. A drop Δ_{CU} in value of 20% (due to changes in exchange rates) on investments made in foreign currency is taken into account. The calculation of S_3 is simple:

$$S_3 = 20\% * (\text{Value assets invested in foreign currency})$$

S_4: Commodity risk.

Commodities are homogeneous products which are typically traded in large quantities. Wheat, oil, and gas are examples of commodities, but also less tangible products such as electricity are commodities. Pension funds do not have the infrastructure required to trade in the actual goods themselves. Therefore investments in commodities usually take place in the form of futures. Futures are financial contracts in which one party commits to sell a certain amount of a product to the second party at a specific date. In return the second party commits to paying a certain price for the product. In other words, the two parties agree on the details of a sale that will take place in the future.

Since the amount and price of the product in this sale is fixed, this means the sale has a certain value. As the date of the sale approaches, the price of the product on the market might change. Depending on the change of the price and whether you are the buyer or seller in the future-contract, this might mean a profit or a loss. What is common with future contracts is that the actual sale of the product never takes place; instead this profit or loss is simply settled in cash between the buyer and seller.

This makes trading in commodities (or rather, commodity futures) very similar to trading in stocks by means of put- and call options. The risk involved in these investments is the risk of an unfavorable price change in the commodity invested in. A drop Δ_{CO} of 15% is taken into account in the solvability assessment for these investments. S_4 is calculated in the following way:

$$S_4 = 15\% * (\text{Value investments in commodities})$$

S_5: Credit risk

Bonds offered by stable governments of countries with a solid economy are usually assumed to have an almost negligible chance on default. The rate of return offered on such bonds is often referred to as the *risk-free credit rate*. Private bonds (which under the FTK includes government bonds offered by less stable countries) usually give a higher return to compensate for a larger chance that the promised payments will not be made.

To determine S_5 the pension fund needs to do a few things. First of all, it needs to determine the rate of return that is offered on its portfolio of private bonds. Then it needs to determine the risk-free rate of return for that portfolio. The difference between the risk-free rate and the actual rate of return is called the credit spread.

Just like the risk-free rate of return can change (which was accounted for in S_1), the credit spread can change. This has effect similar to that of an interest change. However, a change in interest does not necessarily imply a change in the credit spread as well, which is why this risk is calculated separately. A pension fund needs to take an increase of 40% of the current credit spread (C) into account, in the following way:

$$S_5 = 40\% * C * D_{PB} * V_{PB}$$

Here the D_{PB} and V_{PB} are as defined earlier when discussing the calculation of S_1.

S_6: Technical insurance risk.

Technical insurance risk is the only risk factor that is totally unrelated to investments. It is a combination of the risk that stems from unexpected events among the participants, and more structural discrepancies caused by uncertainty about longevity trends and other assumptions.

Pension premiums are based on averages; if someone lives longer than expected, this means the pension fund will have to pay more than anticipated. But aside from mortality, the assumptions underlying the pension premium are unlikely to be correct for every single participant. An example of this is the assumed age difference between the participant and his partner. If the participant is male, it is often assumed that his partner is a three year younger female. If a participant dies and his partner turns out to be a 7 year younger female, it is likely that the widow-pension has to be payed longer than expected. Risks such as these are factored into S_6.

As mentioned before in section 2.3, pension funds are able to work with averages because of the law of large numbers; averages become a better approximation of reality when the number of observations becomes larger. This means that the technical insurance risk becomes relatively smaller as the number of participants of the pension scheme increases. This is also reflected in the calculation of S_6.

S_6 consists of three parts: Process-risk, long life risks[22] and negative stochastic discrepancies[23]. The process risk (PR) accounts for the risk of unforeseen deaths, and is determined in the following way:

$$PR(\%) = (\frac{c_1}{\sqrt{n}} + \frac{c_2}{\sqrt{n}})$$

Here n is the number of participants taking part in the pension scheme.[24] Clearly larger funds will have a relatively smaller process risk. This is in order to take the increased effect of the law of large numbers into account. c_1 and c_2 are percentages based on the type of pension scheme and average age of the participants. The values which a pension

[22] Dutch: *Langlevenrisico* or *trendsterfte onzekerheid* (TSO).
[23] Dutch: *Negatieve stochastische afwijkingen* (NSA).
[24] There are minimum values for n, depending on the type of pension scheme. If a fund has less participants than these minima, it is allowed to use these minima instead of the true value of n.

fund should apply can be determined from a table published by the DNB.[25]

The calculation of the long life risk (LLR) depends on the pension scheme and average age of the participants.

$$LLR(\%) \quad = \quad 2 + p_{LLR} * \max(\Pi - x, 0)$$

The value p_{LLR} is again determined with a table published by DNB. Π is the age of retirement in the pension scheme, while x is the average age of the participants.

The calculation of the risk of negative stochastic discrepancies (NSD) also depends on the type of pension scheme, as well as on the number of participants.

$$NSD(\%) \quad = \quad \frac{p_{NSD}}{\sqrt{n}}$$

Note that while the LLR is independent of the number of participants and depends on the average age, this is reversed for the NSD. When all three factors are calculated, S_6 can be determined.

$$S_6 = (PR + \sqrt{(LLR)^2 + (NSD)^2}) * (\text{value of the liabilities})$$

While the calculation of S_6 is rather complex, its impact on the required funding rate is usually small compared to the other parts, especially for large funds.

The Required Reserve

The required reserves is not simply the sum of these S_i components. First of all, a positive correlation of 0,5 between S_1 and S_2 is assumed. Also, it is unlikely that the worst case scenarios will all take place at once, so some sort of mitigation for this effect is in order. The following equation is used to determine the actual buffer that needs to be kept:

$$\text{Required Reserve} = \sqrt{(S_1)^2 + (S_2)^2 + 2 * 0,5 * S_1 * S_2 + (S_3)^2 + (S_4)^2 + (S_5)^2 + (S_6)^2}$$

This reserve should be sufficient to guarantee with at least 97.5% certainty that the pensionfund will still have a funding rate of 105% one year from the moment of calculation. Whether or not this is actually the case is what I will try to determine in this paper.

[25]See De Nederlandsche Bank (2006), "Advies inzake onderbouwing parameters FTK".

4 Choice of Data

The basis of any good model is reliable data, preferably a large amount of it. However, especially when it comes to macro-economic history, good data is not always easy to come by. Data is not always as reliable as it seems to be, and even if the data is of high quality, the amount of data available might be highly lacking. In this section I will discuss the data I use in the estimation of the model for my analysis. I will discuss data for all the asset classes I model, which are:

- Mature markets
- Emerging markets
- Private Equity
- (Direct) Real estate
- Commodities
- The interest yield curve as published by DNB.

I will use monthly data in order to maintain reasonable sample sizes. The only asset class for which I was not able to find monthly data is direct real estate, for which I could only obtain a mix of quarterly and yearly data.

4.1 The Asset Classes

Mature Markets

The asset class mature markets involves stocks of firms operating in well developed markets. These markets are mature in the sense that they have been around for a while, have stabilized, and the risks involved in these markets are more or less known. To some extent this means that investments in this category are less risky than investments in less transparent asset classes, which is also recognized in the FTK required reserve.[26]

An example of an index that represents a mature market is the Dutch AEX-index. Since pension funds are not required to limit their mature market investments to the Dutch stock market, and typically will not, I choose data from the MSCI world index to represent the mature market returns. This index, published by Morgan Stanley Capital International,

[26] For more details, see section 3.

contains stocks of approximately 1300 of the world's leading companies. It is generally considered to be one of the best indices for global mature market stocks.[27] DNB also used this index for the development of the FTK solvability assessment.

Emerging Markets

The term emerging markets speaks for itself. It concerns stocks that, contrary to those in the category mature markets, belong to firms in markets that are still under full development. Since the properties of these markets are not entirely clear yet, typically the risk of investing in this category is higher than investing in for example mature markets. Naturally, the same holds for the possible returns that can be made.

An example of an emerging market was the internet market during the internet hype near the end of the last century. Incredibly high returns on internet stocks were realized. However, after a while it became apparent that in some cases the promises that were made could not be fulfilled. Consequently the stock prices of some firms fell by a large amount and investors in these firms lost a lot of money.

Several indices describing the developments in emerging markets are available. In their development of the FTK solvability assessment, DNB used data from the MSCI Emerging market indices. Based on the amount of data available and indications[28] that Dutch pension show some preference in keeping their investments close to home, I chose to use the MSCI emerging market index for the Europe region.

Private equity

Private equity is a hard to measure investment type, and possibly even harder to find reliable and useful historical data for. The reason that it is so hard to measure is that the value of private equity only becomes apparent whenever it is sold. Even then, it is not always made public what the value of the sale was; private equity funds are more likely to report profitable sales than less profitable ones. This causes an upward push to the index which is called the self-selection bias. See for example Hoek (2007). Since private equity generally is not very liquid, this means that the number of opportunities to measure the value of the equity involved is limited as well. Moreover, this illiquidity (also known as economic market inefficiency) means private equity investments are risky in the sense that

[27] According to several investment experts at Watson Wyatt.

[28] Based on conversations I had with Watson Wyatt experts in the field of pension fund investments.

it might prove difficult to liquidate the investments at any given time. Again Hoek (2007) gives some more details about this.

Several private equity indices are available. Based on advice from the investment department of Watson Wyatt, I have chosen to use the worldwide LPX50 index. This index lists the performance of 50 private equity companies and is well diversified. Unfortunately it only dates back to December 1993, which limits me to roughly 15 years of observations for this asset class.

Direct Real Estate

Direct real estate suffers from the same kind of illiquidity as private equity does. Also any real estate index will suffer to some extent from the earlier described self-correcting properties: Brokers are far more likely to give publicity to profitable sales than less profitable ones. Moreover, Dutch pension funds often invest in direct real estate inside the Netherlands, so we need data that represents the development of Dutch real estate.

Such an index has been under development for quite some time, but earlier attempts at constructing an useful index failed. It was not until 1996 that a reliable (independent) index was established. This index is published by Stichting ROZ Vastgoedindex and contains data from 31 investors in direct real estate, among which several Dutch pension funds. Until 1999 the index was published on a yearly basis, but starting in march 1999 it was updated every 3 months. It is this quarterly data I will use for my model. This obviously is a very limited sample, which I took into account while constructing the model. However, this index is the same as used by DNB to develop the FTK solvability assessment.

Commodities

Commodities are products that are produced on a large scale by many suppliers and have a homogeneous nature. The products from the different suppliers are considered to be equivalent both in quality and price. The main products that are considered to be commodities are:

- Energy (Both literally and in terms of oil, gas, etc.)
- Metals (Both industrial and precious metals)
- Agricultural products (grains, livestock, softs (i.e. coffee, sugar, cotton))

A large part of the trade in commodities does not take place in terms of the actual goods, but by means of futures. Obviously trading in the actual goods would be rather cumbersome if not impossible for a pension fund.

Again there are several indices to choose from to represent the returns on commodities. The main difference between these indices is the mix of goods they contain. DNB used the Goldman Sachs Commodity Index (GSCI) for the developed of the FTK solvability assessment. I chose for the Dow Jones AIG Commodity Index (DAIG-CI). This index aims for diversification, liquidity and continuity, but my main reason for choosing the DAIG-CI was that it aims to have its weights between the different types of commodities based on economic significance, while the GSCI weighs according to world production.

The interest yield curve

Every month DNB publishes a yield curve that is used in the FTK solvability assessments that are performed around the date of publishing. This yield curve is derived from a swap curve where a fixed interest rate is swapped for a 6 month EURIBOR interest rate. In other words, the two parties involved in the swap basically exchange cash flows of different durations. At the moment of the swap these cash flows usually have equal values.

Not all EURIBOR swap rates are used. To be exact, the 'London composite'-swap rates with a duration of 1-10, 12, 15, 20, 25, 30, 40, and 50 years are used. The reason the swap rates of intermediate durations are not used is that the trade in these durations is usually less liquid. Should the durations normally used also experience a period of reduced liquidity or otherwise irregular behavior, DNB can decide to make an exception and exclude that duration from the calculations of the yield curve.[29]

The swap curve is transformed into a curve consisting of zero (spot) rates for use in the FTK solvability assessment. These spot rates can be calculated from the swap curve by a bootstrapping procedure. This procedure is well explained in a publication by DNB.[30] Since the method is technical but straightforward I will not discuss it in detail here. However, a description of the bootstrap procedure used to generate the yield curve can be found in appendix A.

While DNB only published monthly yield curves starting January 2004, I was able to construct similar yield curves dating back to January 1999 using the bootstrapping procedure and historical data about the EURIBOR swap rates. Since the EURIBOR

[29] For example, in the yield curve of 31-12-2005 the swap rate with a duration of 9 years was excluded from the derivation, because it showed a strong outlier compared to the rest of the swap rates.

[30] De Nederlandsche Bank (2005), "Vaststelling methode rentetermijnstructuur".

swaps did not exist before that date, I had to find another way to construct earlier yield curves. Again several of the problems I described earlier arise: Quality and availability of data.

Swaps are fairly new financial products, and were not that popular (and hence not very liquid) in earlier years. Also, swaps with long durations (above 20 years) only recently became more common. On top of that, it is important to keep in mind that the swaps used should be representative for the Dutch interest rates.

After doing some research on the data available and consulting the experts at Watson Wyatt, I decided to use the German DEM swaps with durations of 1-10, 12, 15, 20, 25 and 30 years where possible to construct earlier yield curves. I filled in some gaps in this data by interpolation, and subsequently applied a similar bootstrapping procedure to this data as DNB uses for the EURIBOR rates.

4.2 The amount of data

The purpose of the FTK solvability assessment is to produce a required reserve that is enough to keep the funding rate of a pension fund above the minimum required level of 105% in 97.5% of the cases. In other words, the required reserve should keep the funding rate of a pension fund above 105% over a one year time period and should not fail to do so more often than once every 40 years on average. As DNB states:[31]

> "Determination of scenarios that occur once every 40 years is not easy. The main reason for this is that in many cases insufficient historical observations exist to make such estimations; even for stock- and interest markets for which a reasonable amount of historical data is available, this is not easy. Moreover, expected returns, volatility (standard deviation) and correlations are not constant measures over time."

This statement perfectly describes the dilemma I faced when deciding what data to use for analysis: Should I use as much data as possible arguing that this gives the best way to estimate events that happen once every 40 years, or should I use less data arguing that this gives a better description of the current volatilities and expected returns?

For my model, I limited my datasets to data of the last 15 years. While DNB tried to use as much data as possible in the development of the FTK model, I will use less (yet

[31]DNB (2006), "Advies inzake onderbouwing parameters FTK", page 7. Translated from Dutch to English. Any interpretation mistakes in the translation are mine.

more recent) data for the investment categories. This allows me to check whether the risks estimated by DNB based on long term data correspond to the risk that is implied by more short term data. However, this means I am using 15 years worth of data to analyze the FTK model, which is supposed to deal with events that happen on average once every 40 years. As evident from the statement made by DNB, this approach has both advantages and disadvantages. I will list the most important advantages, as well as the most important disadvantages.

Advantages of using 15 years worth of data

1. **The most recent data describes the most recent behavior of returns and volatility.** As mentioned in the statement by DNB, there is evidence that macro-economic parameters like volatility, expected returns and even correlations are not constants over time. See for example Whitelaw (1994). The theory behind using only recent data is that this best describes the current behavior of those parameters; a shorter observation period reduces the chance that somewhere in the data a structural change in the behavior took place.

2. **Correlations can be modeled.** The choice for using a time period of 15 years is also a modeling decision. As shown in section 3, the FTK solvability assessment assumes correlation between several asset classes. This means the assessment assumes that the results on those asset classes are related. Specifically, since a positive correlation is assumed, it implies that large losses (as well as large profits) in those asset classes tend to occur together. In order to model this correlation between the asset classes I constructed a multivariate model.[32] However, to estimate such a model I need to measure the correlation over time between the asset classes. Therefore I need observations for *each* asset class at the *same* moments in time. This limits the data I can use to the period in time of which I have observations for all asset classes. That means that my data is limited to data of the last 15 years, since beyond that I have no observations for the asset class private equity.

Disadvantage of using 15 years worth of data

1. **Estimation is subject to more sample variation.** This disadvantage is directly related to advantage 1 listed above. When using a small data set, the importance of

[32] More on this in section 5.

individual observations increases. This means that outliers (i.e. extreme results) will have a larger impact on the estimated model than they would have had in a larger data set. Part of this disadvantage can be avoided by excluding clear outliers from the used data.

Simply said, short term data might not describe long term risks accurately. This is basically what all the advantages and disadvantages revolve around. Long term risks are hard to model almost by definition. When using short term data, in- or excluding outliers in the data may lead the estimated model to over- or understate the risks it tries to model. These effects are also known as the *ghost features* of the data. However, I will not remove or smooth any extreme results, since it is specifically the extreme market outcomes that are of most interest to the FTK solvability assessment. Moreover, more data does not necessarily mean that long term risks are estimated better, due to possible trend breaks in the data. For example markets, governments, laws and regulations might change, changing the risk involved in the asset classes we describe. As Sir Alec Cairncross, a distinguished economist and former chief economic adviser to the British government, once put it:

> "A trend is a trend, is a trend. But the question is, will it bend? Will it alter its course through some unforeseen force and come to a premature end?"

In order to gain some insight into the decision of using only 15 years worth of data, in section 10 I will perform a analysis similar to that in section 8, but based on an ALM model developed by Watson Wyatt. This ALM model contains roughly 40 years worth of data, similar to the amount of data used by DNB to develop the FTK solvability assessment. A description of this model can be found in section 9.

5 Structure of the model

I will now continue with the discussion of the structure of the model I estimated. I will also discuss the theory behind it and the literature on which I based my model choice.

5.1 Heteroskedasticity and Autoregression

In the past 50 years a lot of empirical research on financial time series analysis has been done. During that time it became clear that the results of this research often shared similarities. Among the first to suggest that these common observations and inferences might hold true in general for most financial time series were Mandelbrot (1963) and Fama (1965). The commonalities have been proven to hold over and over again since then[33] and as such are now elevated to a status nearly equivalent to facts, known as the *stylized empirical facts of financial time series.* The main stylized facts are:

1. Return series are not independent although they show little autocorrelation.
2. Series of absolute or squared returns show profound serial autocorrelation.
3. Conditional expected returns are close to zero.
4. Volatility appears to vary over time.
5. Extreme returns appear in clusters.

These stylized facts are usually referred to in the context of daily time series, but often continue to hold if longer (e.g. weekly, monthly) or shorter intervals are used.[34]

Stylized facts 1 and 2 form an interesting duo. Fact one claims that financial time series usually show little evidence for *autocorrelation*[35]. Autocorrelation means that the value of current returns can be (partially) explained by a linear combination of past values of the returns. Taking R_t as the return in period t, in a mathematical equation autocorrelation looks like this:

$$R_t = c + (\sum_{i=1}^{p} a_i R_{t-i}) + \epsilon_t \tag{1}$$

[33]See for example Cont (2000).
[34]See McNeil, Frey and Embrechts, “Quantitative Risk Management”, page 117.
[35]Autocorrelation is also known as *serial correlation.*

Here c and the a_i's are constants, ϵ_t is the innovation (random shock) to the return at time t, and p is the number of past values that is of importance for the current value of the returns. A model like this is known as an autoregressive (AR) model of order p. Autocorrelation might also occur in the innovations, in which case the model will look like:

$$R_t = c + (\sum_{i=1}^{p} a_i R_{t-i}) + (\sum_{i=1}^{q} b_i \epsilon_{t-i}) + \epsilon_t \tag{2}$$

The constant q and the b_i's are similar to p and the a_i's in the autoregressive part of equation (1). A model like (2) is called an autoregressive moving average (ARMA) model of order (p, q).

According to facts 1 and 2, while a model like (2) does not seem to be useful for direct modeling of the returns, squared or absolute returns often do seem to show a relation like this. For most purposes a model of the squared or absolute returns is useless; it is impossible to infer whether the underlying returns were positive or negative. Moreover, in practice its often nearly impossible to estimate whether returns are more likely to be positive or negative in the future as can be seen from fact 3. I will not discuss the properties of ARMA modeling any further.[36] As fact 1 points out, while not autocorrelated in the sense of equation (1) or (2), returns are *not* independent of each other. The reason I still want to mention the ARMA models is that the model I will use in my analysis shows some strong similarities to the ARMA models in its structure, yet relates the subsequent values of the returns to each other in a different way.

The model I will use is related to facts 4 and 5. These facts describes what in the literature is referred to as *heteroskedasticity*. Heteroskedasticity means that periods of fairly normal returns are followed by periods of more extreme returns and vice versa. Related to this is fact 5, which tells us that extreme returns tend to occur close together. Note that nothing is said about the sign of those returns; extreme profits can be followed by either additional extreme profits, or extreme losses. This also leads to fact 3: Even if you know the current state of the volatility and returns of a certain stock, based on that information it is very hard to determine whether you are more like to make a profit or loss on that same stock in the next period.

[36] For readers that are interested to know more about the properties of ARMA models and their estimation, many excellent texts on this subject are available. See for example Johnston and DiNardo (4th edition, 1997) and McNeil, Frey and Embrechts (2005).

Heteroskedasticity is an interesting phenomena. Bohl and Siklos (2004) give a possible explanation for this in the way of feedback/trend chasers. They describe trend chasers as investors who invest based on what other investors do. If they see the value of a certain stock go up (normally a sign of increased demand for that stock), they will start investing in that stock as well. As a result, the price of that stock will go up even further. The risk in this is that the stock will become overpriced. The result is that investors will start selling the stock. When the trend chasers pick up on this, they will start selling their stocks as well, possibly causing a large drop in the value of the stock by doing so. This again might make the stock underpriced and attractive for investment again. All this means a high volatility in the stock price. However, it is likely that after a while the price of the stock will stabilize such that the returns of the stock are less excessive again.

Much research has been done on heteroskedasticity in the various asset classes considered in this paper. Aggarwal et all (1999) attempt to explain changes in volatility in emerging markets by means of social, political and economic events. De Santis and Imrohoroglu (1997) study heteroskedasticity in several emerging markets, while Karolyi (1995) does the same for mature markets in the U.S. and Canada. Michelfelder and Pandya (2005), Bohl and Siklos (2004) and Beirne et all (2008) are recent papers that study heteroskedasticity in both mature and emerging markets and the relation between those markets. Similarly Baillie and Myers (1991) study heteroskedasticity in the market of commodity futures. Susmel (2000) models heteroskedasticity in private equity markets. Interest rates are also shown to have heteroskedastic properties, see for example Avougi-Dovi and Jondeau (1999) and Lucchetti and Palomba (2008).

Moreover, all[37] papers mentioned above model the heteroskedasticity by means of the generalized autoregressive conditional heteroskedastic (GARCH) model developed by Bollerslev (1986) or a variation of that model. Specifically, Karolyi (1995) and Beirne et all (2008) use the multivariate GARCH-BEKK variation of this model, which was developed by Engle and Kroner (1995). I will adopt the same structure for the model I will use in my own analysis. Due to the complexity of the GARCH-BEKK model, I will discuss some of the theory behind this model in section 5.2.

[37]With the exception of Aggarwal et all (1999), which is more focused on explaining and identifying volatility changes than it is focused on modeling the volatility itself.

5.2 (G)ARCH modeling

In this section I will discuss some of the underlying theory behind the GARCH-BEKK model I will use in my analysis. This section does not contain any estimation or analytic results, but a summary of the most important theoretical properties of the GARCH-BEKK model and its underlying structures.

The ARCH-model

The autoregressive conditional heteroskedastic (ARCH) model developed by Engle (1982) was among the first models to explicitly model the relation between returns and the changes in their volatility. In his model, Engle let the volatility of the innovations in a period depend on previous innovations. The general definition of the ARCH model is given below.

$$
\begin{aligned}
r_t &= \mu + \epsilon_t \\
\epsilon_t &= \sigma_t u_t, \text{where} \\
u_t &\sim N(0,1) \\
\sigma_t^2 &= c + \sum_{i=1}^{n} \alpha_i \epsilon_{t-i}^2
\end{aligned}
$$

This is under the conditions that $c \in \Re^{++}$, $\alpha_i \in \Re^{+}$, $\mu \in \Re$ and $q \in \natural$. r_t represents the (log) returns at time t. While seemingly very complex due to its several layers, the interpretation of this model is rather straightforward. The returns r_t consist out of an average return μ and an innovation ϵ_t. This innovation is partially determined by a random shock u_t, which at the very least determines the sign of the innovation. However, unlike in the ARMA model mentioned in section 5.1, the *size* of the innovations is also determined by the factor σ_t. σ_t works as a scaling factor for the innovation at time t, and can either increase or decrease the volatility of the returns in that time period. In an ARCH(n) model an extreme innovation in one period will cause the volatility of the next innovations to increase, which increases the likelihood of additional extreme innovations in the next couple of periods.

In other words, this perfectly corresponds to the stylized empirical facts of financial time series. The returns are not autocorrelated, yet are not independent either (Stylized fact 1). Knowing the past returns and/or innovations does not tell you anything about which direction the next returns/innovations will take. Since past returns have an effect on the volatility of returns to come, volatility will vary over time (Stylized fact 4), and do so in such a way that extreme observations are likely to occur in subsequent periods (Stylized fact 5). Stationarity conditions and other properties of ARCH models are discussed in many papers, see for example Zaffaroni (2000).

GARCH models

In the standard ARCH model the volatility relies on the previous innovations. Since these previous innovations depend on the volatility in their corresponding period, the current volatility indirectly depends on previous volatilities as well. The generalized autoregressive conditional heteroskedastic (GARCH) model introduced by Bollerslev (1986) makes this relation between volatilities more explicit. It assumes a structure for σ_t similar to that of an ARMA model. The standard GARCH(n,m) structure is shown below.[38]

$$\begin{aligned} r_t &= \mu + \epsilon_t \\ \epsilon_t &= \sigma_t u_t, \quad \text{where} \\ u_t &\sim N(0,1) \\ \sigma_t^2 &= c + \sum_{i=1}^{n} \alpha_i \epsilon_{t-i}^2 + \sum_{j=1}^{m} \beta_j \sigma_{t-j}^2. \end{aligned} \tag{3}$$

Here the parameters are similar to those of the ARCH model, with the additional of the parameters $\beta_j \in \Re^+$. The estimation procedures for ARCH and GARCH models are well explained in the original papers by Engle and Bollerslev, yet far from trivial, especially for higher order models. However in practice it seems that the 'simple' GARCH(1,1) model often performs better for many purposes than higher order models.[39]

Many excellent texts have been written on (G)ARCH models. Readers that are interested in the application and general theory of (G)ARCH models, I can highly recommend Engle (2001), which offers a very readable introduction to the subject. GARCH models have been extensively applied in the literature, as for example the survey by Bollerslev et all (1992) shows. Many variations on the standard ARCH and GARCH models have been developed, for example the exponential GARCH (Nelson, 1991), Threshold GARCH (Rabamananjara and Zakoïan, 1993), and GARCH-in-mean (Engle et all, 1987) models.

[38] Bollerslev (1986) also suggests that using a student t-distribution instead of the normal distribution may improve the performance of the model for some time series.

[39] See for example Bollerslev (1986), as well as McCurdy and Morgan (1988).

While these variations may give better fits to specific data sets, according to Hansen and Lunde (2005) none of these variations clearly outperforms the standard GARCH(1,1) model for financial time series in general. For my own model I will also adopt a (multivariate) GARCH(1,1) structure. As explained in section 4 a multivariate model will allow me to model the (changing) correlation between different asset classes. Specifically, I will use the diagonal GARCH-BEKK structure proposed by Engle and Kroner (1995).

The diagonal GARCH-BEKK model

Since my model will be a multivariate variant of the GARCH model described in the previous section, it is most convenient to denote this model using matrix notation. I assume the reader has a sufficient understanding of matrix notation and calculation, those who have difficulties interpreting the notation below I refer to Simon and Blume (1994).

The basic notation for a n-variate GARCH(1,1) model is:

$$\begin{aligned} R_t &= \mu + \epsilon_t \\ \epsilon_t &\sim N(O, H_t) \\ H_t &= C + A\epsilon_{t-1}\epsilon'_{t-1} + BH_{t-1} \end{aligned} \tag{4}$$

Where R_t is the nx1-matrix containing the modeled returns at time t, and ϵ_t is the nx1-matrix with the corresponding innovations. H_t is the nxn variance-covariance matrix of the multivariate distribution of ϵ_t. C, A, and B are fixed, symmetric nxn matrices. μ and O are fixed nx1 matrices. Specifically, O is an nx1 matrix consisting of zeros.

While it is closest to the univariate model in terms of notation, this description of the multivariate GARCH(1,1) model is inconvenient compared to other notations. From the definition of a variance-covariance matrix it follows that any useful variance-covariance matrix needs to be *positive definite*. This also implies that in order for a multivariate GARCH model to be of any use, H_t needs to be positive definite for all possible values of ϵ_t. It is rather cumbersome to check if this condition holds in the formulation above after estimation, and even more so to impose this restriction during estimation. Therefore Engle and Kroner (1995) suggested an alternative formulation, which under very weak conditions guarantees that this condition will hold:

$$H_t \;=\; C'C + A'\epsilon_{t-1}\epsilon'_{t-1}A + B'H_{t-1}B \tag{5}$$

Where C, A and B are fixed nxn matrices with C triangular. This is the GARCH-BEKK model of Engle and Kroner. However, usually a restricted version of this model is used in which the matrices A and B are restricted to be *diagonal*. Unsurprisingly, this version of the model is called the *diagonal* GARCH-BEKK model. The restriction implies that variances only depend on past values of that variance and its own squared residuals, and similarly covariances only depend on past values of itself and the corresponding cross-products of residuals. This restricted model is intuitively very plausible and far easier to interpret than the unrestricted model. As such it is often preferred by researchers over the unrestricted model. Moreover, this formulation makes checking several conditions that need to be met considerably easier.

Most important of all, I need to check under what conditions the matrix H_t following from (5) is positive-definite. Engle and Kroner (1995) gives these conditions:

PROPOSITION 1 *If H_{t-1} is positive definite, then the parameterization of the GARCH equation given in (5) yields a positive definite H_t for all possible values of ϵ_t if the null space of C and the null space of B intersect only at the origin.*

For the proof of this proposition I refer to Engle and Kroner (1995). A sufficient condition for this to hold is that either C or B is of full rank, since the null space of a full rank matrix is equal to the origin. In the case of the diagonal GARCH-BEKK model, if the matrix B we estimate has all diagonal entries significantly different from 0, this condition is already satisfied.

The diagonal formulation also makes checking for (covariance) stationarity[40] trivially easy. From Engle and Kroner (1995) follows:

PROPOSITION 2 *Suppose the process $\{\epsilon_t\}$ is a doubly infinite sequence and equation (5) defines an n-variate diagonal GARCH-BEKK process. Let a_{ii} and b_{ii} denote the ith diagonal elements of A resp. B. Then, $\{\epsilon_t\}$ is covariance stationary if and only if $a_{ii}^2 + b_{ii}^2 < 1$ for all $i = 1, \ldots, n$.*

[40] An explanation of the importance of this property will follow in section 6.2

This concludes my discussion of the theory behind the model I used in my analysis. In the next section I will discuss the actual estimated models and the simulation results. The actual analysis of the FTK solvability analysis will be performed in section 7.

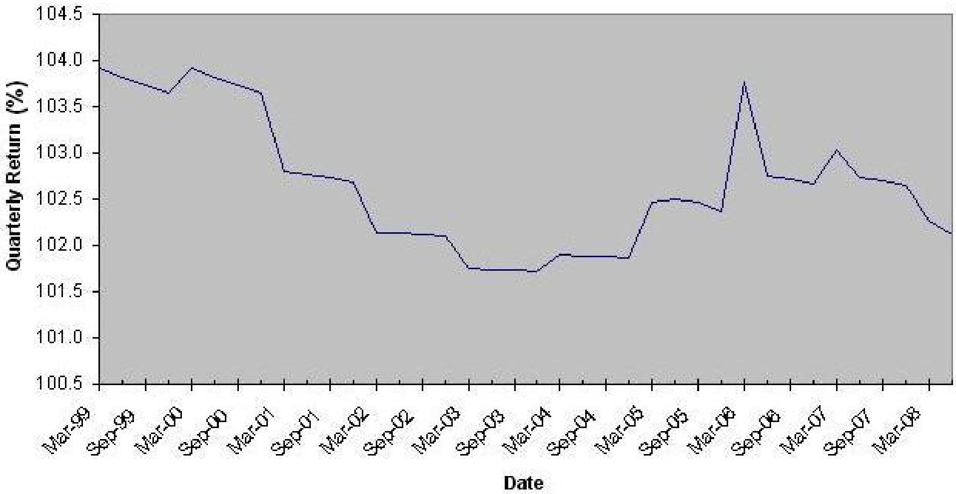
6 Model Estimation and Simulation

In this section I will discuss the estimated model. A discussion about the scenarios generated by the model under various assumptions will follow in the next section.

6.1 The exception: Direct Real Estate

As mentioned I will estimate a multivariate model that takes the dependence between all asset classes into account. The exception to this is the asset class direct real estate, which I will model separately. I have two reasons for doing so: The amount of data available and properties of that data. This is best illustrated by a graph of the data.

Figure 2: Quarterly returns on direct real estate, based on the ROZ-IPD index



Note that the returns have been strictly positive in the entire existence of the index. Many real estate indices suffer from a strong bias due to self reporting; real estate funds have more interest in showing their profits than losses. Since one of the main goals of the ROZ-IPD index was to construct an independent benchmark free of those effects I will assume that those effects are minor here. However, a model based on this data will not reflect any risk one might take by investing in real estate.

This data does raise an interesting question: Is the assumed risk of a loss of 15% on direct real estate in the FTK solvability assessment too extreme? While the available data is limited, it still describes a timespan of roughly ten years without any negative return.

Another problem with this sample is that since it is quarterly data from the past 10 years, the sample is too small to be able to estimate autoregressive or heteroskedastic effects with reasonable accuracy. The estimates obtained from such a small sample would be subject to considerable sample variation.

Due to these reasons I decided to model the returns on direct real estate simply by taking the average return. The average quarterly return on direct real estate in the data is equal to 2.66%, which translates to an average return of 11.1% per year. This means that that in my model reserves for direct real estate investment risks will always be sufficient to compensate for those risks, which seems questionable at best from a practical standpoint, but is consistent with my data.

6.2 General Modeling Conditions and Conventions

During the development of my model I will take some general conventions into consideration. Also, there are some general conditions my model should meet in order to be useful.

Log-Returns

The first convention I will follow is the use of the log-returns of my data to develop my model. The log-returns are obtained by simply taking the (natural) logarithm of the normal returns. The reason for using log returns is partly one of convenience. As shown below, when using log returns one can simply take the sum of the returns during a period in order to determine the total return in that period, instead of the product. R_{t_1,t_2} denotes the total return in the period starting at time t_1 and ending at time t_2.

$$\begin{aligned} R_{0,T} &= \prod_{i=0}^{T-1} R_{i,i+1} \\ \ln(R_{0,T}) &= \ln(\prod_{i=0}^{T-1} R_{i,i+1}) \\ &= \sum_{i=0}^{T-1} \ln(R_{i,i+1}) \end{aligned}$$

Because of their multiplicative nature, returns are usually modeled by means of multiplicative models. However, using log-returns allows me to model these same relations in a linear manner. In practice, linear models are usually a lot easier to estimate than non-linear models which gives another reason for the use of the log-returns.

Stationarity

Stationarity is one of the most important concepts in time series modeling. In most cases, for a model to be useful for forecasting or simulation one needs to ensure that the series generated by such a model will not become unrealistically excessive. For example, if the variance of ϵ_t in an GARCH model would on average continue to grow over time, the returns generated by such a model would become more and more extreme and simulation outcomes from the model would stop making any sense after a while. In order to be able to estimate a stationary model, usually the underlying time series should be stationary as well.[41]

However, there is no logic in trying to estimate a stationary model if the underlying time series is non-stationary. In order to check stationarity of the underlying series I use the augmented Dickey Fuller (ADF) test.[42] Each of the log return series of the data described in section 4 is stationary according to the ADF test at a 1% significance level. The same holds for the first difference series of each of the interest yield curve durations I use for the model.

6.3 Estimation Results

I estimate a diagonal GARCH-BEKK model of order (1,1) for the asset classes mature markets, emerging markets, private equity and commodities, along with interest rates from the interest yield curve of durations of 5, 10, 15 and 25 years. I briefly discuss each of the estimated elements of the model structure in appendix C, but this discussion is omitted here because it does not give much additional insight into the structure of the model. The main purpose of using a GARCH-BEKK structure to model the returns, is that this model tries to incorporate changes in both volatility and correlation between asset classes over time. Therefore in this section I will show and discuss the modeled changing correlation over time for each of the asset class combinations.

[41] This does not hold true in general. Some models, like for example the vector error correction (VEC) model can be estimated from non-stationary time series under certain conditions. However, such models are not of interested for this thesis and any further details about them will be omitted.

[42] For more details about this test, see Said and Dickey (1984).

Correlation between returns on risky assets

Figure 3 shows the correlation between the various risky asset classes over time.

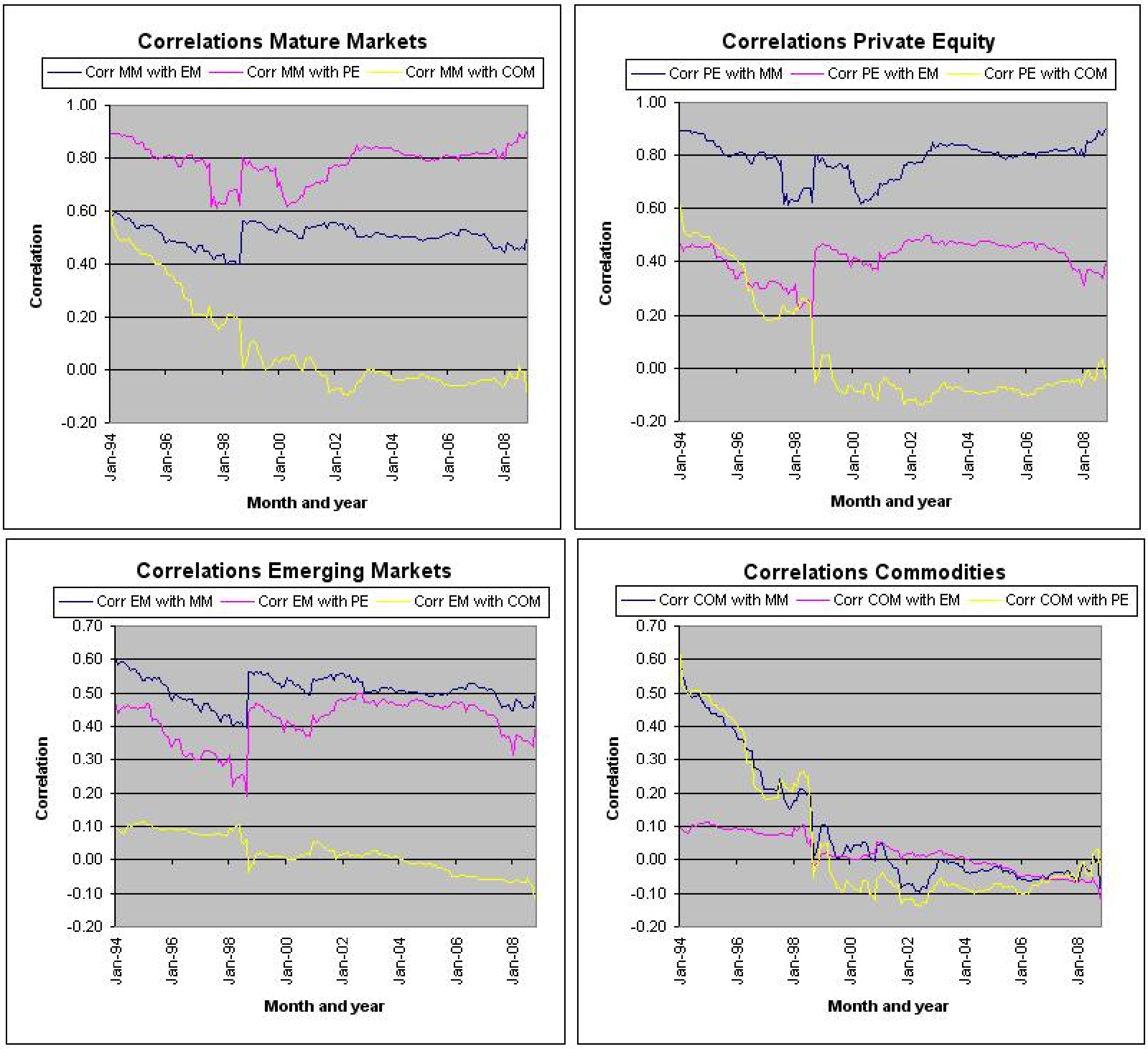


Figure 3: Correlations between the returns on various risky assets

There are several interesting results shown in figure 3. First of all, there is a significant jump in each of the correlation graphs near the end of 1998. These jumps occur around August of that year, the month in which the Russian financial crisis took place.[43]

[43] While usually referred to as a 'financial' crisis, this crisis was strongly related to the declining commodity prices around that time. The downfall of these prices caused much of the financial problems for countries like Russia, for which the exports consist largely out of commodities.

Not only is the correlation between commodities and the other asset classes consistently close to zero, especially during times of crisis the correlation seems to be gone all together instead of strongly increasing like with the other asset classes. This is a very interesting result in the sense that it confirms the decision that was made in the FTK to not include the asset class commodities in S_2.

The only other large jump that occurs in the correlation between the various asset classes is the one that occurs in the correlation between emerging/mature markets and private equity in the third quarter of 1997. In July 1997 there was a major financial crisis among many Asian countries. Unlike the Russian financial crisis of 1998, this event was purely financial in nature which explains why it mostly had an effect on the stock markets, but not as much on private equity and the market of commodities.

Something else worth noting is that the average correlation between the risky assets is quite a bit lower than the assumed 75% in the FTK solvability assessment. Only the correlation between private equity and mature markets seems be close to this level. Since the risk due to correlation increases with the amount of (positive) correlation, this means that the FTK solvability assessment *overestimates* the correlation risk. Obviously, this is under the assumption that my model is a good representation of reality. Assuming all other risks are measured accuratly, this implies that the FTK solvability assessment is likely to perform better than required.

Correlation between interest rates of different durations

Figure 4 displays the modeled correlations over time between the various interest rates used in the model. It is immediately clear from this figure that the correlation between changes in interest is very high. Moreover the figure shows that the correlation between interest rates is most volatile for the correlations with the interest rate of duration 5. This is not unexpected. Investments experts at Watson Wyatt indicated to me that the short term interest rates tend to be more volatile than the long term interest rates. The reason for this is simple: Good or bad news might have a significant impact on the short term expectations on the market, but it takes a lot more for the long term expectations to be affected. A change in the long term interest rate means that investors belief something structural has changed on the market that goes beyond a temporary effect. The correlations between the higher duration interest rates (10, 15 and 25) are more or less constant over time.

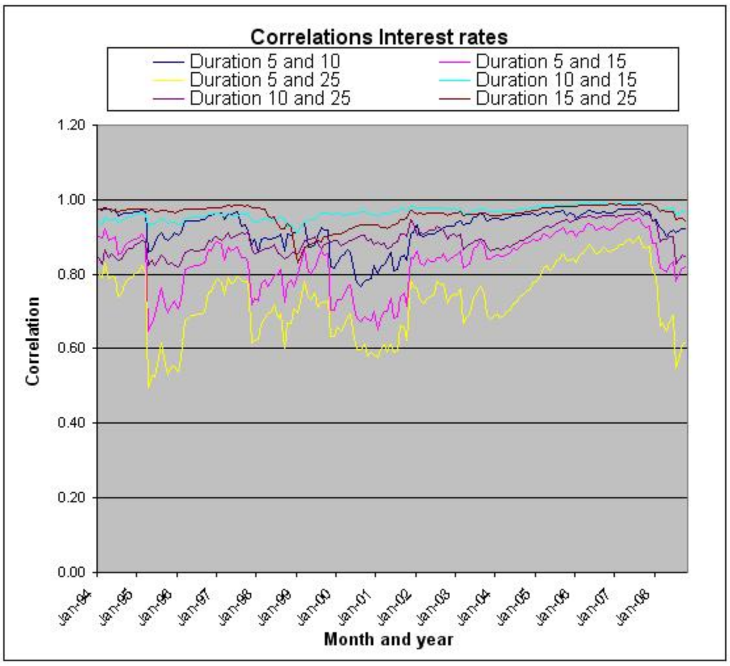


Figure 4: Correlations between the various interest rates

Correlation between interest rate changes and returns on risky assets

Figure 5 shows the correlation between the changes in interest rates with duration 5 and 25 years, and the returns on various asset classes. I omitted the graphs of the correlations between the other interest rates and the risky assets, since they did not give any additional insight into the model that was not already included in the displayed graphs.
When comparing the graphs in figure 5, it immediately becomes clear that the short term interest rate shows a larger correlation with the risky assets than the long term interest rates. This is consistent with my earlier observations; Like the stock prices, the short term interest rate is more dependent on (the same) short term expectations of the economy. That said, the sign and changes in the correlation of the long and short term interest rates seem to follow a similar pattern.

However this pattern is far from what I expected to find. The FTK solvability assessment assumes a positive correlation between assets and interest rates of 50%. However, based on this model the correlations in the first 5 years of my observations are negative. Moreover, even in the periods where the correlation is positive the level of 50% is only reached for a few small periods of time. Again this suggests that the risk due to corelation

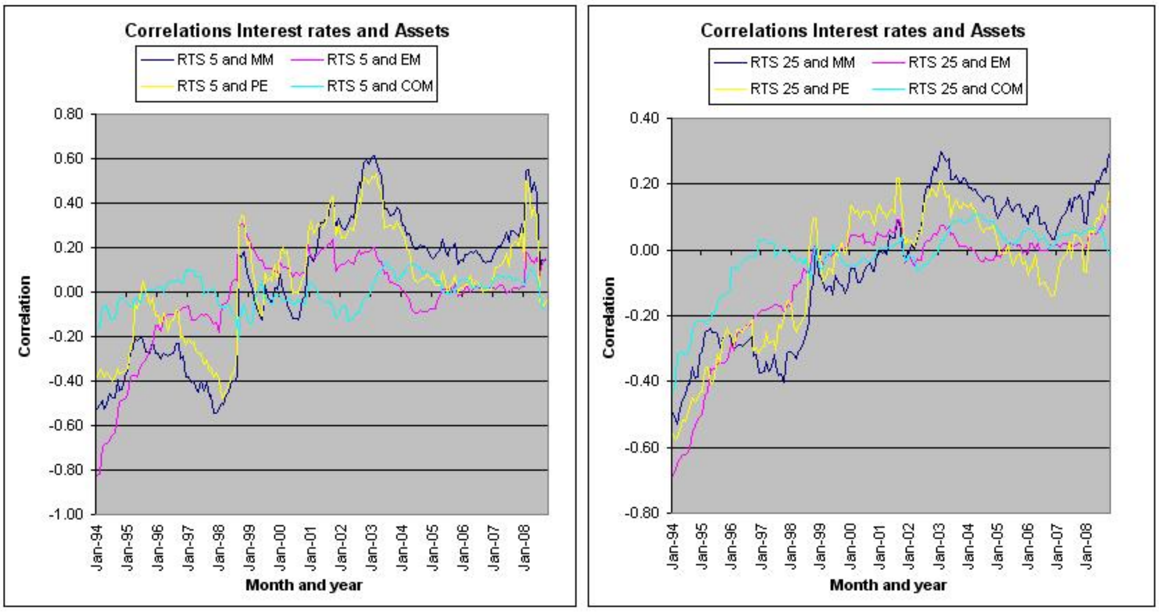


Figure 5: Correlations between the changes in interest rates and returns on risky assets

is overestimated in the FTK solvability assessment, which in turn suggests that the FTK required reserve is likely to perform better than required.

Stationarity

As explained in section 6.2 it is preferable if not necessary for simulation purposes that the model is covariance stationary. For the diagonal GARCH-BEKK model the condition for stationarity is that $a_{ii}^2 + b_{ii}^2 < 1$ for all i. In table 1 the values corresponding to the various models are listed.

Elements & Asset class	value
$a_{11}^2 + b_{11}^2$ (MM)	0.95
$a_{22}^2 + b_{22}^2$ (EM)	0.67
$a_{33}^2 + b_{33}^2$ (PE)	1.01
$a_{44}^2 + b_{44}^2$ (COM)	1.01
$a_{55}^2 + b_{55}^2$ (interest dur. 5)	1.00
$a_{66}^2 + b_{66}^2$ (interest dur. 10)	1.01
$a_{77}^2 + b_{77}^2$ (interest dur. 15)	1.02
$a_{88}^2 + b_{88}^2$ (interest dur. 25)	1.04

Table 1: Stationarity statistics

From the table it is clear that parts of the model hover on the verge of non-stationarity. The stationarity condition is not met when using the estimated parameters, but taking into account the sampling variance of the estimators the hypothesis of stationarity cannot be rejected either. Since my simulation runs will be rather short term (twelf periods, one year) this should not lead to any stationarity problems.

7 Simulation Assumptions

In this section I describe the process I will use to to analyze the performance of the FTK required reserve, and the assumptions I need to make in order to be able to perform this analysis. I will start with a short description of the process.

Using the model discussed in section 6 I simulate the returns on the investment portfolio during a one year period, as well as the changes in interest rates during that time. Specifically, I simulate the returns on investments in mature markets, emerging markets, private equity, commodities and direct real estate, as well as the changes in the interest rates of the DNB interest yield curve corresponding to durations of 5, 10, 15 and 25 years. I call one set of such outcomes a scenario. Since the generated scenarios depend on the assumptions about the volatility on the various markets at the start of the year, I will use different sets of economic assumptions.

I obtain the required reserve and corresponding funding rate for pension funds from the FTK solvability assessment. In order to be able to do this, assumptions are needed about the investment portfolio of the pension fund in question, and the duration of its liabilities and assets. I will check the performance of the FTK solvability assessment for different types of pension funds, by using several different investment portfolios and assumptions about the durations of the assets and liabilities.

For each set of scenarios I will then analyze the performance of the required reserve. Most importantly, I will check in how many scenarios the funding rate of the pension fund remains above the minimum level of 105% after a one year period. This is done under the assumption that the fund starts out with a funding rate equal to the required funding rate.

This process gives opportunities at several points to make decisions about what kind of situation to analyze. In order to make a thorough analysis of the performance of the required reserve, I repeat this process several times using different combinations of assumptions each time. I will simulate a sample of 25000 scenarios for each of these sets of assumptions. In the rest of this section I will describe what those different sets of assumptions are. Section 8 contains the actual analysis of the FTK solvability assessment.

7.1 Assumptions about the economy

In order to be able to simulate the returns on an investment portfolio in a given month, I need to know the variances and correlations of the underlying assets in that month. Once

I know these values for a particular month, the variances and correlations for the next months follow from the model. For my simulation of the returns this means I need to make some assumptions about the initial volatility and correlations at the start of the year.

These assumptions basically give a description of the state of the economy at the start of the year, something that is almost completely ignored in the FTK solvability assessment.[44] I will consider four different starting situations for amount of market volatility. Each of these situations are based on real variances and correlations measured in the data. One of the main reasons I choose to use real covariance matrices as starting situations is that this ensures that the relative values of all the correlations are both feasible and realistic. Another reason for doing this is that the matrices now also give valuable insights into the recent developments of the funding rates of pension funds. Due to the size of the resulting covariance matrices they are placed in appendix D. I will explain my choice for each particular matrix.

Starting situation A: December 2007

My choice for the covariance matrix of this period has several reasons. First and foremost, this period serves as a average starting situation. Compared to all the other months included in my data, the average variance of the assets and interest rates, as well as the average correlation between all these variables are close to the overall average. This means that the results obtained while using the covariance matrix of December 2007 should be able to serve as a good reference point for comparison when determining the effect of using more extreme starting situations.

Moreover, there is a high practical value in using this month's covariance matrix as a starting situation. As I discussed in the introduction, at the start of 2008 there was no real hint of the impending problems pension funds would face later that year. Since the FTK solvability assessment has only been in official use for less than 5 years and pension funds already find themselves in major trouble now, this has raised questions about the reliability of the FTK. Was the major downfall of stock markets and interest rates during 2008 one of those scenarios that belongs in the 2.5% of the cases the FTK required reserve would not be sufficient? Or was the reliability of the reserve less than 97.5% to begin with? In other words, could we have seen the problems coming? By analysing the performance

[44] The value of S_1 does scale slightly with the current value of the interest rates. However, this says nothing about the volatility on the market. Instead it implies that if the interest rates are lower, interest risk is lower as well. Whether or not this is actually true is questionable, but is beyond the scope of this paper.

of the reserve based on the data of december 2007, it is possible to obtain at least a part of the answers to these questions. The values of the covariance matrix corresponding to this starting situation can be found in appendix D, table 16.

Starting situation B: June 2008

This situation serves as a follow up to situation A in two ways. First of all, it is the second step in answering the question ‘could we have seen the problems coming?’. The events that led to the eventual economic crisis in the end of 2008 already started in the summer of 2008, when it turned out that many mortgages in the United States were defaulted upon due to the house owners not being able to pay the banks their money back.

The interesting part about the covariance matrix of this particular month is that while the variances were only slightly higher than average, the correlations in this period were among the highest recorded in the entire 15 years of the data. Using this period as one of the starting situations gives another indication on how likely the problems near the end of 2008 would have been, given this model. Moreover, this starting situations gives me the opportunity to see what the effect of high correlation on the performance of the FTK is. The values of the covariance matrix corresponding to this starting situation can be found in appendix D, table 17.

Starting situation C: September 1998

As discussed before, this is the period the Russian financial crisis occurred, which has a clear effect on the correlations between the asset classes I use in my analysis. More importantly, it also happens to be the period with the highest average variance recorded in my data. Interestingly enough, the level of the average correlation between asset classes in this period is roughly equal to the average of the whole data. Therefore, this starting situation allows me to check the effect of extreme variance on the performance of the FTK. The values of the covariance matrix corresponding to this starting situation can be found in appendix D, table 18.

Starting situation D: October 2008

This period is a combination of starting situation B and C in the sense that it combines high variances with high volatility. Of the four starting situations, this one is the most extreme. It serves as a major stress test on the performance of the FTK. Moreover, in combination with starting situation B and C it shows the combined effect of high variances and high correlations. From a practical point of view, it shows how reliable the FTK solvability reserve is under highly volatile market situations as we have seen in the very recent past. The values of the covariance matrix corresponding to this starting situation can be found in appendix D, table 19.

7.2 Analysis of the Risks and Returns

I have nearly arrived at the point where I can discuss the results of the analysis. However, it is crucial that I first verify whether or not the returns and changes in interest rate that the model generates (based on the given assumptions) are reasonable. If the model does not seem capable of generating plausible scenarios for the returns, obviously the analysis done with such scenarios would be completely irrelevant in practice.I perform an analysis of the generated returns by generating a sample of 25000 scenarios for each set of volatility assumptions. By looking at some statistics based on these samples, I should be able to determine the general plausibility of the scenarios generated by the model.

First I look at the average returns and changes in interest rates generated by the model. For comparison I also list the average yearly returns and changes from the data I used to estimate the model.

	Starting Situation				
Investment Category	A	B	C	D	Data
Mature markets	1.04	1.05	1.05	1.05	1.05
Emerging markets	1.26	1.27	1.30	1.27	1.24
Private Equity	1.05	1.06	1.08	1.08	1.10
Commodities	1.07	1.07	1.05	1.07	1.05
Δ Interest (5y)	0.0%	0.0%	0.0%	0.0%	0.0%
Δ Interest (10y)	-0.1%	-0.1%	-0.1%	0.0%	0.0%
Δ Interest (15y)	-0.1%	-0.1%	-0.1%	-0.1%	0.0%
Δ Interest (25y)	-0.1%	-0.1%	-0.1%	-0.1%	0.0%

Table 2: Average yearly returns and interest changes

I am not looking for an exact match between the data and the model, since I am comparing 15 observed yearly returns with 25000 simulated yearly returns. I take 25000 scenarios from my model because I want to obtain a detailed picture of the scenarios it generates. However, 15 observed yearly returns form a very limited sample for comparison; one extreme result in the sample will already have a strong effect on the sample average, even if the distribution of the real returns is exactly the one modeled. Still, the comparison may serve to give some intuition about the model. Judging from table 2 the starting situations have barely any effect on the average returns. The average returns of the data seem similar to the model averages, both in absolute and relative size.

I am not satisfied by just looking at the averages of the returns and interest changes. It is also important to know if the extreme values generated by the model are sensible, as well as the general distribution of the returns between those extremes.

	Starting Situation									
	A		B		C		D		Data	
	Min	Max	Min	Max	Min	Max	Min	Max	Min	Max
MM	0.61	1.85	0.50	1.81	0.55	2.01	0.49	2.20	0.63	1.26
EM	0.37	3.86	0.34	4.29	0.22	6.86	0.29	4.97	0.42	2.24
PE	0.34	3.97	0.17	4.27	0.12	6.35	0.16	6.78	0.46	2.64
COM	0.34	2.63	0.40	2.38	0.61	1.75	0.38	2.82	0.70	1.24
Δ I5	-3.2%	+3.2%	-3.8%	+3.6%	-3.9%	+3.5%	-3.6%	+3.5%	-2.2%	+2.5%
Δ I10	-3.5%	+2.9%	-3.2%	+3.1%	-3.2%	+3.6%	-2.9%	+3.3%	-2.2%	+2.5%
Δ I15	-3.4%	+3.4%	-3.1%	+3.1%	-3.8%	+3.8%	-3.3%	+3.2%	-1.2%	+1.4%
Δ I25	-3.8%	+5.1%	-3.7%	+5.3%	-6.0%	+4.8%	-3.8%	+4.0%	-2.2%	+2.5%

Table 3: Most extreme yearly returns

The extremes from the data are included in table 3 along with the extremes from the simulations. Due to the size of the table I was forced to abbreviate the names of the asset classes, but the ordering of the asset classes in table 3 is identical to table 2. The reason for including the extremes of the data is that they are a basic test to check if the model is capable of generating extreme enough results, while at the same time it is meant as a check to see if the near non-stationarity of the model leads to any practical problems in the sense of excessively extreme returns. As pointed out before, the sample constructed from the real life data is very limited, so it is unlikely that it will contain the most extreme results that can occur on the markets. However, this means that we would expect the extreme outcomes of the model to be more extreme that those of the data. If this is not the case, something is wrong.

Looking at the minimal returns, the models seems to be capable of generating results at least as extreme as those in the data, but not so extreme that they seem intuitively improbable to ever occur. Also it seems that if I assume more extreme starting situations, the yearly returns that are generated tend to be more extreme as well. This is exactly what I would expect based on the stylized facts of financial time series that I described in section 5.

At first glance the maxima of the returns seem rather high. However, remember that the model generates monthly data. Even the most extreme result, namely a return of 6.78 on private equity, can be realized with an average return of about 1.17 per month. A return of 1.17 in a month on private equity is very high, but not extremely excessive. However it is highly unlikely that this will occur for 12 months in a row in reality, so a return like this from the model should be extremely rare. The same holds for the extreme changes in interest in table 3.

For this reason I also made a histogram of the simulated returns, in order to get an idea about their spread. In appendix F a few of those graphs are shown. The histograms give the final reassurance that the model produces sensible scenarios. For example, the excessive maximum return of 6.78 for private equity was clearly an outlier, since the great majority of the simulated returns lie between 0.5 and 2.2. The same goes for the excessive changes in the interest rates, since the simulated changes are mostly smaller than 2% in absolute value. Based on the observed returns in the data, the model indeed seems to produce reasonable economic scenarios.

7.3 Assumptions about the pension fund

In order to analyze the performance of the required reserve that results from the FTK solvability assessment, I first need to determine the required reserve. As explained in section 3, the size of the required reserve depends on the investment portfolio of the pension fund, as well as the duration of its liabilities. In order to be able to make a well founded statement about the general performance of the required reserve I need to know how it performs for various pension funds with different investment portfolios.

First of all, the duration of the liabilities plays an important role in the required and actual funding rate. DNB assumed an average duration of 16 for the liabilities of its stylized pension fund. Funds with a high average age among its participants are likely to have a duration that is slightly lower. However, the difference is likely to be largest for pension funds in young industries. It is not unrealistic for such funds to have liabilities with an average duration of 25 years or more. This means that changes in interest cause large changes in the value of the liabilities for such funds.

In order to check the effect of the duration of the liabilities on the performance of the required reserve, I will conduct the analysis for pension funds with a duration of the liabil-

ities of 15 and 25 years. Obviously results for a pension fund with liabilities with a higher or lower duration will be slightly different, but based on the results for these two funds an indication can be given on how the reserve will perform for different funds.
The investment portfolio of the pension fund is the other factor that determines the FTK required reserve for a pension fund. Investment strategies differ quite a lot between pension funds, which also means the risk involved in the investments is different for each pension fund. I will distinguish two different assumptions in the investment portfolio: Asset allocation and duration mismatching. The asset allocation is simply the way the assets are divided among the investment categories, while the duration mismatching is related to the difference between the duration of the bonds invested in and the duration of the liabilities of the pension fund. I will distinguish three different investment portfolios, which are shown in table 4.

	Portfolios		
Investment Category	Risky	Neutral	Safe
Mature markets	60%	40%	25%
Emerging markets	5%	3%	0%
Private Equity	5%	2%	0%
Direct real estate	2.5%	5%	5%
Commodities	2.5%	5%	10%
Bonds	25%	45%	60%

Table 4: Investment portfolios (in % allocation of total assets)

I want to emphasize that the names of the portfolios only reference to investment risk. Generally stocks and equity are considered to be more risky investments (and therefore also referred under the collective name risky assets) than bonds. The names of the investment portfolios only indicate the spread between risky assets and bonds, and are in no way intended to refer to the performance and/or risk concerning the funding rate of the pension fund. The relation between the portfolios and the risk for the pension fund in terms of the funding rate is to be determined in section 8.

The observant reader may have noticed that if I were to use all possible combinations of the assumptions I have described so far, this would result in a very extensive amount of results. However, not all combinations of these assumptions are equally realistic and interesting. I will therefore focus the discussion of my results on the most interesting cases, but I will try to include as much results as is possible without loosing readability or giving redundant results.

Additional assumptions

I made some additional assumptions about the other specifics of the pension funds that concern the FTK solvability assessment. The first of these assumptions is that all the pension funds' investments are valued in euros. This effectively makes the pension fund exposure to currency risk equal to zero, which means $S_3 = 0$. Many pension funds in reality have investments in currencies other than the euro. The percentage of assets invested in foreign currencies differs a lot between pension funds. Many pension funds with high exposure to foreign currencies take precautions to protect themselves from the currency risks involved. What makes the currency risk in the solvability assessment so hard to analyze, is that it not only matters how much of the assets is invested in foreign currency, but also how this is done. Pension funds might invest directly in foreign currency, or indirectly by means of buying stocks, bonds or equity valued in those currencies. Two pension funds who are otherwise identical, will still obtain different portfolio returns if their foreign currency investments are spread differently between the various asset classes. Moreover, the type of currency or currencies that is invested in will also play a big role in determining the returns and risks involved in these investments. To summarize, in order to model currency risk I would need assumptions about the total amount invested in foreign currencies, the spread of these investments among asset classes and about the spread among specific currencies. The amount of research needed to make sensible sets of assumptions about this is beyond the scope of this paper. However, the assumption of $S_3 = 0$ can also be interpreted as the assumption that any contribution of S_3 to the required reserve is always exactly enough to cover the risk involved in this asset class.

I make the same assumption for the technical insurance risk, S_6. In order to correctly measure this risk, a full model simulating the demography of a pension fund over the course of a year would be needed. For example, realistic assumptions about the marital status, age and mortality rates of the participants would be needed. Again, this extensive amount of research is beyond the scope of this paper. The assumption of $S_6 = 0$ is equivalent to assuming that the contribution of S_6 to the required funding rate is always exactly enough to cover the risks.

The final assumption I make is that all the investments in bonds are made in the highest rated type of bonds. This assumption, unlike the previous two, was not made because of model complexity issues. In fact, given an average spread among the different ratings of bonds, the effect on the assets of an interest/credit spread change would be quite easy to model. However, I was not able to find enough historical data of credit spreads relative to the DNB interest yield curve in order to make such a model.

It is possible that the total effect of these additional assumptions has a noticeable effect on the outcomes of my analysis. However if so, I expect this effect to be small; S_1, S_2, their correlation along with S_4 together constitute by far the largest part of the FTK required reserve. The benefit of the assumptions above is that they allow me to really focus specifically on these major parts of the FTK solvability assessment. Moreover, should I find that the performance of the required reserve under these assumptions is lacking, I will give an indication how much risk should be 'compensated' by the reserves resulting from S_3, S_5 and S_6 and to what degree this seems reasonable.

Now the FTK and its background are explained, the model is developed, its output analyzed and the assumptions for the analysis are set, I have arrived at the core of this thesis: The analysis of the performance of the FTK solvability assessment.

8 Analysis of the FTK solvability assessment (I)

8.1 Simulation Results

In this section I will discuss the observations I made during the simulation runs, as well as the conclusions that can be drawn from those observations. I will include some of the simulation data to illustrate the results. Note that the results included in the main text of this section are only a small sample of the simulation runs I performed. However, the conclusions drawn hold for the other simulation results as well. The full results of the simulation runs can be found in appendix G. For readability I will refer to a pension fund with a liability duration of X and an asset duration of Y as an X/Y-pension fund.

The goal of this analysis is to determine if the required reserve resulting from the FTK solvability assessment meets its goal of keeping the funding rate of a pension fund above 105% with a success rate of at least 97.5%. In order to test this, I will use the null hypothesis that the success rate is in fact exactly equal to 97.5%. I will test this hypothesis against the alternative hypothesis that the success rate is less than 97.5%. I will use critical values corresponding to a 99% confidence level. This means that I will only reject the null hypothesis if I observe simulation outcomes that have a probability of less than 1% of occuring if the null hypothesis were true.[45] Since I assume the concepts of critical values and significance levels are familiar to the reader, I will not explain these concepts further here. However, appendix E contains a more detailed explanation about how these critical values are determined.

After each simulation run, I determine in how many of the simulated scenarios a pension fund starting with the required funding rate would find itself below a funding rate of 105% after a one year period. I also determine the number and percentage of scenarios where the funding rate would drop below the required funding rate and those below a funding rate of 100%. Moreover, I determine the number of scenarios where the value of the assets decreased in value during a one year period. Note that a drop in the value of the assets does not necessarily imply that the funding rate drops too; interest changes may lower the value of the liabilities, possibly in such a way that the resulting funding rate actually goes up.

[45] To avoid confusion: The *success rate* (97.5% in the null hypothesis) is the reliability of the required reserve resulting from the FTK solvability assessment. The *confidence level* (99%) is the reliability of the decision to *reject* the null hypothesis. While the two values are similar, they refer to two completely different things.

General observations

First of all I consider the effect of changing starting situations on the performance of the required reserve. Since the FTK solvability assessment does not take recent economic volatility into account, extreme starting situations are likely to result in the threshold of 105% to be crossed more often. As table 5 shows[46], this indeed turns out to be the case.

Starting situation:	A	B	C	D
> 100%	(24839) 99.356%	99.092%	99.208%	99.018%
> 105%	(24373) 97.492%	96.704%	96.944%	96.560%
< Req. Funding rate	(10533) 42.132%	42.808%	43.144%	43.312%
Drop in asset value	(7358) 29.432%	28.568%	33.080%	31.544%
Required funding rate: 122.1%				

Table 5: Simulation results for a 15/10-pension fund with neutral portfolio

As expected, the FTK required reserve performs better under a typical starting situation (A) than it does under the more extreme starting situations. The success rate decreases when correlations between asset classes or the volatility of individual asset classes increases. While this holds without exception when comparing situation A to situations C and D, there are some exceptions when comparing the results found under situations A and B. These exceptions occur for the 25/5- and 25/10-pension fund under the safe investment portfolio as well as the neutral portfolio. In these cases, as well as for the 25/15-pension fund under the safe investment portfolio, the performance of the FTK required reserve is actually better under situation B than it is under situation A. In other words, in the simulations with the pension funds that are exposed to the largest (relative) amount of interest risk, the performance of the reserve actually seems to benefit from the increased correlations. Moreover, this benefit decreases as the durations of the liabilities and assets are better matched. Unfortunatly I was not able to find an explanation for these exceptions. However, it means that aparantly increased correlations do not always negatively influence the performance of the FTK required reserve.

Judging from the results in table 5 the FTK solvability assessment performs worst under the starting situations B and D. This holds not true in general though, since under

[46]To emphasize that the percentages are determined from the number of scenarios that meets a certain condition, I listed both the number of scenarios as well as the percentage of the total number of scenarios in table 5 for starting situation A. These figures obviously have a one-to-one relation and are direct substitutes of eachother. However, since the percentage notation is a lot easier to interpret in terms of the null hypothesis, I will omit the underlying number of scenarios in further tables.

which of the extreme starting situations the required reserve performs best or worst does not seem to be constant at all between simulation sets.

Table 5 shows two additional trends that are also visible in the rest of my simulation results. First of all, notice that the number of scenarios in which the funding rate has dropped below the required level after a one year period is larger than the number of scenarios in which the total assets drop in value. This means that there have to be scenarios in which the assets increase in value, but the liabilities increase in value even more (due to a change in interest) resulting in a drop in the funding rate. This clearly shows the importance of interest risk for a pension fund; even when the asset portfolio increases in value, the financial position of a pension fund may deteriorate.

Second, a change in the starting situation does not seem to have much (if any) effect on the probability a pension fund will see its funding rate drop over the course of a year. This means that the decrease in performance I mentioned before is not a result of more frequent drops, but of more severe drops of the funding rate. The combination of these observations leads to the following conclusion:

CONCLUSION 1 *As volatility on the market increases, the success rate of the reserve required by the FTK solvability assessment decreases. Increased correlations on between asset classes can also lead to a reduced performance of the FTK solvability assessment, but this does not hold in general. Finally, the increase in volatility and/or correlations does not cause the funding rate of a pension fund to drop more frequently, but does increase the overall severity of the drops which take place.*

Critical values

So far I only discussed the relative performance of the required reserve. The time has come to analyze the actual performance of the FTK required reserve in various conditions. As discussed before I will need critical values to test the null hypothesis that the required reserve achieves its goal of a 97.5% success rate. I will be testing at a 1% confidence level. The corresponding critical value is listed in table 6. I also noted some additional critical values for other signficance levels to give the reader some idea about the sensitivity of the analysis to the choice of significance level. As mentioned before, appendix E gives a detailed explanation of the derivations of these critical values.

Sign. level	Critical value
5%	97.34%
1%	97.27%
0.1%	97.19%

Table 6: Critical values for a success rate of at least 97.5%

The effect of duration (mis)matching

First I want to determine if the changes in the FTK required reserve that result from duration (mis)matching correspond well to the difference in interest risk that results from the (mis)matching. This can be analyzed by checking the performance of pension funds with fixed investment portfolio, fixed liability duration but varying asset duration. Table 7 shows such an analysis for two pension funds. The first is a pension fund with liabilities with a duration of 15 years, the second is a pension fund with liabilities with a duration of 25 years. Both funds use a neutral investment portfolio.

Pension fund assumptions			Starting situation			
Duration Liabilities	Duration Assets	Required Funding rate	A	B	C	D
15	5	124.9%	97.61%	96.78%	97.07%	97.04%
15	10	122.1%	97.49%	96.70%	96.94%	96.56%
15	15	119.5%	97.58%	96.39%	96.68%	96.29%
25	5	132.1%	94.19%	95.01%	92.14%	93.54%
25	15	125.8%	93.53%	94.06%	89.69%	91.33%
25	25	122.1%	95.23%	95.07%	93.61%	93.82%

Table 7: Success rate of the FTK required reserve sorted by duration matching

Note that only for the old pension fund, under the least extreme starting conditions, the null hypothesis of a success rate of at least 97.5% cannot be rejected when comparing the outcomes with the critical values in table 6. In each other case the results suggest that it is highly unlikely that the success rate of the FTK meets its required goal.

The results in table 7 show a difference between the pension fund with liability duration of 25 years and the pension fund that has liabilities with an average duration of 15 years. For simplicity I will refer to these different pension funds as the young fund respectively old fund. For the old fund the success rate of the FTK solvability assessment seems to steadily decrease as the duration matching between the liabilities and assets becomes better. For the young fund the success rate also decreases but increases again when the duration matching increases further. This effect is curious, since intuitively one would expect to see the same effect for both funds. However, it is possible the results show two separate effects at the same time.

First of all, consider the steadily declining success rate shown for the old fund. Obviously, as the duration of liabilities and assets are better matched, the interest risk the pension fund is exposed to is reduced, and this should lead to a drop in the required reserve/funding rate. However, as the required funding rate in table 7 drops, so does its success rate. This can indicate one or more of three things: Either the reduction in the interest risk due to a better duration matching is overestimated, the investment risk is underestimated, or the correlation between these risks is underestimated.

This last option can be ruled out quite easily, since I already determined that the correlation between interest changes and asset returns based on my model and data is comfortably below the assumed 50% that DNB assumes in the FTK solvability assessment. The first option is selfexplanatory, but the second option requires some explanation.

When the durations of the assets are better matched with those of the liabilities, interest risk is reduced. Meanwhile, the amount of investment risk due to investments in risky assets stays the same. That means that relatively speaking, the interest risk decreases while the investment risk increases, since investment risk now becomes a larger (relative) part of the risks the pension fund is exposed to. In other words, because interest risk becomes less important (since it is reduced), investment risk becomes more important. Now, if the risks measured in S_2 are underestimated, this will contribute to a weak performance of the required reserve. The effect of this underestimation will have a stronger effect on the lacking performance of the required reserve if its relative importance increases. Therefore, a possible explanation of the reduced performance of the required reserve for the old fund when its durations are better matched, might be that the risks in S_2 are underestimated.

If the duration of the assets of the young fund is increased from 5 to 15, so a similar step as for the old fund, we observe the same overestimation of the reduction in interest risk. However, when the duration of those assets is increased even further, suddenly the FTK required reserve starts to perform better again. This contradicts the effect of S_2 I decribed, leaving an underestimation of interest risk as the most likely explanation for the shifting performance of the reserve. The question that remains now, is why the performance increases again in the last step for the young fund.

An explanation for this might be that the FTK solvability assessment underestimates the interest risk for low durations, but approximates it better for higher durations, possibly even overestimating it in those cases. In fact, this explanation does not seem unreasonable.

I already discussed earlier that short term interest rates are typically more volatile than long term interest rates. This difference in volatility needs to be taken into account when determining the interest risk a pension fund is exposed to. The FTK solvability assessment actually attempts to do so in the calculation of S_1: The factors for the lower duration interest rates are more extreme than those used for the higher durations. If the volatility of the interest rates is underestimated in the FTK solvability assessment, this is likely to play a bigger role for lower durations than for higher durations, since those interest rates tend to be less volatile to begin with.

Table 7 provides another indication that the lacking performance of the required reserve is caused by an underestimation of interest risk: The success rate for young pension fund is substantially lower than that of the old fund. The young fund obviously is exposed to a greater interest risk than the old fund, both in absolute and relative sense. Using the same line of reasoning I used above, this confirms the other indications that there is an underestimation of interest risk.

CONCLUSION 2 *According to my simulation, interest risk is underestimated in the FTK solvability assessment. As a result, the target success rate of at least 97.5% is not attained and the actual success rate may be considerably lower, depending on the pension fund in question and its investment portfolio. The estimation of interest risk appears to be better, yet still insufficient, for high durations.*

While the results I discussed sofar clearly suggest that it is an underestimation of interest risk that leads to the lacking performance of the FTK required reserve, it is difficult to say anything about the sufficiency of the buffer that is taken on investment risk (by means of S_2) at this point.

The effects of asset allocation on the performance of the required reserve

In order to get a more clear picture of how investment risk affects the performance of the FTK I also performed several simulations with different investment portfolios for each fund. In table 8 I listed some simulation results for both a 25/15-pension fund and a 15/10-pension fund, using different investment portfolios.
Table 8 shows a very clear result: As the investment portfolio of the pension funds shifts more into stocks and less into bonds, the performance of the required reserve improves. This is very interesting in the sense that when a larger portion of the assets is invested in

Pension fund assumptions				Starting situation			
Investment Portfolio	Duration Liabilities	Duration Assets	Required Funding rate	A	B	C	D
Safe	15	10	115.2%	95.66%	95.13%	95.44%	95.84%
Neutral	15	10	122.1%	97.49%	96.70%	96.94%	96.56%
Risky	15	10	132.3%	98.12%	97.12%	97.28%	96.97%
Safe	25	15	118.2%	88.66%	90.37%	83.02%	87.22%
Neutral	25	15	125.8%	93.53%	94.06%	89.69%	91.33%
Risky	25	15	137.5%	96.11%	95.79%	93.62%	94.24%

Table 8: Success rate of the FTK required reserve sorted by duration matching

stocks, which is typically regarded as riskier than investing in bonds, the risk of a funding deficit for the pension fund is actually reduced! Apparently the increase in the required funding rate due to the reallocation of assets to stocks and equity is enough to compensate for the additional investment risk.

More importantly, the simulation results suggest that the risks in S_2 are estimated at a sufficiently high level. When portfolios with a larger portion of risky assets are used, obviously invesmtent risk increases, not just in absolute sense but also as a relative part of the total risk a pension fund is exposed to. As I explained before, as the relative importance of a specific type of risk increases, the degree to which this risk is correctly estimated in the FTK required reserve will have a larger impact on the performance of this reserve. As table 8 shows, when the pension funds use portfolios with a larger investment risk, the performance of the reserve increases, in some cases even to an performance level that is high enough to not reject the hypothesis that the succes rate of the FTK required reserve is at least 97.5%. This also suggests that the reserve kept on invesmtent risk may actually *compensate* partially for the underestimation of interest risk.

CONCLUSION 3 *According to my simulation results the FTK solvability assessment, specifically the calculated value of S_2, yields a sufficiently high reserve to compensate for investment risk. There is also some evidence that the reserves kept on investment risk are able to push the success rate of the FTK required reserve above the lower bound of 97.5% in cases where other risks are of limited importance.*

This conclusion can also be read as “investment risks are overestimated in S_2”, if the goal of a pension fund would be to keep the performance of its reserve as close to 97.5% as possible. However, since the underestimation of interest risk reduces the performance of the reserve well below that lowerbound of 97.5% in the majority of the cases, an overestimation of investment risk is not of much interest at this point. Should the performance of the other parts of the solvability assessment be improved in the future, then this overestimation may become of more interest to study.

The results in table 8 also confirm my conclusion that the main issue that causes the performance of the FTK required reserve to fail is the underestimation of interest risk. Even though the absolute interest risk is reduced when a larger portion of the assets is invested in bonds[47], the relative importance of the interest risk increases, causing the performance of the required reserve to drop.

Sufficient reserves

One of the questions that now comes to mind is, if the success rate is indeed less than 97.5%, how much less is it? The answer to this question depends strongly on the assumptions of the pension fund and the investment portfolio used. The critical value belonging to a hypothesis that the success rate of the FTK required reserve is at least 97.0% is equal to 96.75%. Clearly, this hypothesis would still be rejected for many of the observed results in table 8. In fact, a large part of the results in table 8 would also lead to rejecting a hypothesis that the success rate is at least 96%, which has a corresponding critical value of 95.71%.

While a drop of the success rate with 1.5% might not seem like much on such a high percentage, it is *exactly* because the success rate required is so high that this change is huge. This is because the interest of the FTK is not as much in its success rate, as it is in what happens in the remaining part of the scenarios. A drop in the success rate from 97.5% to 96.0% means a increase of 60% in the number of scenarios in which the reserve is insufficient. It is very likely that such a risk for a pension fund will be considered unacceptable by both the participants of such a pension fund, not to mention the pension fund itself and its supervisor, DNB. A success rate of 96% would imply that a pension fund which exactly meets the FTK solvability requirements would still find itself in a funding deficit every 25 years on average. To make matters worse, the hypothesis that this will occur with a frequency of at most 25 years is rejected in several of the cases in table 8.

The simulation results suggest that the required reserve is insufficient to meet the goal of a 97.5% succes rate. However, how much higher should the funding rate be to make it

[47]That is, assuming the assets have a lower average duration than the liabilities.

sufficient? In order to get an estimation of this, I ran several of the simulations again, but this time with several different initial funding rates. This way I determined an indication on how far off the required funding rate is compared to what would be sufficient for the average starting situation A as well as sufficient to even be sufficient in the extreme starting situations B through D. The results are shown in table 9.

Pension fund assumptions				Sufficient for starting situation	
Investment Portfolio	Duration Liabilities	Duration Assets	Required Funding rate	A	B, C and D
Safe	15	5	118.8%	120.8%	121.2%
Neutral	15	10	122.1%	122.1%	123.3%
Neutral	25	15	125.8%	132.2%	139.5%
Risky	25	15	137.5%	141.4%	148.3%

Table 9: Sufficient required reserves.

The sufficient funding rates were determined on basis of one simulation of 25000 scenarios. Since any such estimation will be subject to (limited) sample variation, these funding rates should merely be viewed as an indication of the error made by the FTK solvability assessment according to my model.

Sample variation or not, it is clear from table 9 that the difference between sufficiency and insufficiency is marginal for funds that have only a limited exposure to interest risk, at least in terms of base points of the required funding rate. However, even for the 15/10-pension fund for which the FTK required reserve performs decent under average conditions, this still means an increase of about 5.4% of the required reserve in order to be sufficient under extreme circumstances.

So far I have assumed that the target success rate of 97.5% implies that the FTK required reserve that is produced by the FTK solvability assessment at some point in time should always have a success rate of at least 97.5%. In my opinion this is also the way it should be interpreted, especially since the addition of the words 'at least' suggest that higher is acceptable, but under no circumstances the success rate should be allowed to drop below 97.5%. However, one could also interpret the target success rate of at least 97.5% in other ways. For example, it could be interpreted as a target for the *average* success rate. Under such an interpretation the insufficient performance of the required reserve under extreme circumstances would not neccesarily be a problem, as long as the performance under less extreme circumstances is good enough. This would imply that during times the reserves of a pension fund are under the most stress, the reliability of the reserve the pension fund is allowed to be below 97.5%. I highly doubt this is the way the target of a 97.5% success rate should be interpreted, since that would mean that the reliability of the reserves matters least when the reserves themselves are needed the most.

8.2 Simulation results in the light of recent events

As mentioned before I also want to look at the simulation results from the perspective of recent developments in the average financial situation of Dutch pension funds. As mentioned in the introduction, at the start of 2008 nearly all pension funds were well above the minimal funding rate of 105%, with a large part of them meeting or surpassing the required reserve demanded by the FTK, while at the end of the same year many pension fund were in a situation of funding deficits.

Based on the performance of the required reserve under the various starting situations, I have to conclude that the required reserve of the FTK might have given several types of pension funds a misplaced amount of confidence in their financial situation. Especially funds particularly vulnerable to interest risk may have underestimated the size of this risk. The simulation results clearly show that pension funds with a relatively high exposure to interest risk face a significantly higher probability of a funding deficit on the short term. The corresponding success rate of the FTK required reserve seems to lie well below the intended 97.5% in many cases, even under average economic conditions like those in starting situation A. Moreover, the success rate of the FTK severely worsens when the economic situations become less stable, especially in cases of high volatility. This means that in situations like those of the summer of 2008, when correlations were high and volatility on the markets started to increase but most pension funds were not yet in major financial trouble, pension funds might think themselves to be relatively safe from risk while in fact the FTK required reserve is no longer enough to ensure the 97.5% success rate.

8.3 Summary

The required reserve that results from the FTK solvability assessment seems to be sufficient only in cases where the interest risk a pension fund is exposed to is limited. The analysis done by means of simulation with my model shows that the percentage of scenarios in which the FTK required reserve is enough to keep the funding rate of a pension fund above 105% after a one year period, can drop considerably below the by DNB stated goal of 97.5%. This specifically holds true for pension funds with a relatively high exposure to interest risk, be it due to duration mismatching between liabilities and assets, due to liabilities with high durations, or simply because a large part of the assets is invested in bonds making interest risk relatively more important for the performance of the reserves.

This effect gets amplified even more in case of extreme variance and/or correlation between the different asset classes. Especially the effect of high variances causes can severely influence the success rate of the reserve. A high exposure to investment risk due to investments in stocks and equity seems to be adequately covered by the required reserve.

The difference between a sufficiently high reserve and the actual required reserve that results from the FTK solvability assessment generally will tend to increase as the interest risk a pension fund is exposed to increases. Moreover, a reserve that performs sufficient under extreme conditions needs to be considerably higher than a reserve that only needs to be sufficient in average situations. In other words, high correlations and/or volatility on the economic markets severely increase the stress on the performance of the reserve, which may drastically reduce its performance.

9 Alternative model: The Watson Wyatt ALM model

In the previous sections I analyzed the performance of the reserve resulting from the FTK solvability assessment by means of my own model, that simulates changes in the assets and liabilities of a pension fund. Watson Wyatt Worldwide also developed a model that can be used to perform a similar analysis. In practice this model is used as an asset and liability management (ALM) model and can be applied for example as part of a continuity analysis. The model is capable of generating up to 1000 scenarios that simulate the financial status of a pension fund up to several decades into the future. While these scenarios can contain many details like premiums payed, various types of indexation and asset management strategies and many, many other details that will affect a pension fund's financial status in the long run, for this analysis I will only need the basic functions of the model to simulate changes in the assets and liabilities of the pension fund in the course of a year. Future cash flows other than those resulting from the current assets and liabilities are not of interest since they are disregarded in the FTK solvability assessment. Because of this, I will only discuss the characteristics of the Watson Wyatt ALM model that are of importance to my analysis. While I am not at liberty to disclose the exact values of the parameters of the model, I will discuss the general structure and assumptions used and indicate any important differences between my GARCH-BEKK model and the Watson Wyatt ALM model in terms of the structure used.

Characteristics

The Watson Wyatt ALM model is based on a multivariate time series model, which in turn finds its roots in the capital asset pricing model (CAPM). Specifically, the ALM model is determined globally through a blend of economic theory, historical analysis and the views of investment managers. Moreover, region-specific corrections to the model are developed in the corresponding countries. This is done according to the CAPM model:

$$R_Y = R_{fy} + B_Y(R_m - R_{fw}) + e_y$$

Where:

R_Y = the return for country Y.
R_{fy} = the risk free rate of return in a local market.
R_m = the (hedged) return on the global equity market.

R_{fw} = the global risk free rate of return.
B_Y = the beta value for country Y.
e_y = a random risk component with zero mean.

This means that the returns are determined by three parts. First of all there is the local risk free rate of return. To this a factor is added that is based on the difference between the global risk free rate of return and the return on the global equity market. While this difference is fixed, in what way this difference contributes to the actual returns is country specific through B_Y. This is called the *systematic risk* of a country and describes how the market in this country structurally differs from the global market. Finally there is a factor similar to the error terms which I discussed several times throughout this paper. This factor represents all the effects on the returns that are not yet included in the other two factors that play a role in the returns. The distribution of e_y can be either a normal distribution or a student-t distribution, depending on the asset class in question.

The error terms of the different asset classes are correlated. Moreover, these correlations are redetermined on a quarterly basis, along with the expected returns. This ensures that the ALM model stays up to date with the current state of the economy. For my analysis, I use the version of the ALM model that is specifically tailored to be used in the Netherlands. Moreover, this version of the model contains the parameter values as determined for the first quarter of 2009.

Differences from the GARCH-BEKK model

There are several differences between this model and the GARCH-BEKK model I used in the previous section to analyze the FTK required reserve. One of the key differences is that while the GARCH-BEKK model allows the correlations to change over time, the Watson Wyatt ALM model keeps those correlations fixed. This means that the ALM model assumes that the inserted correlations are constant over time, while periods of high and low volatility may alternate in the GARCH-BEKK model.

While the GARCH-BEKK model I used is purely constructed based on economic theory and historical analysis, the Watson Wyatt ALM model also uses the knowledge and expertise of the investment department of Watson Wyatt, along with the general expectations on the investment markets. This means that it is unavoidable that the ALM model contains subjective elements. The goal of this is to use the expertise available to fine tune the model.

Like the GARCH-BEKK model, the ALM model is not autoregressive in the sense that current returns directly depend on previous returns. However the GARCH-BEKK model does include this relation indirectly through the changing volatility over time. The ALM model therefore contains less autoregressive properties than the GARCH-BEKK model. However, this only holds for the equity returns. Other factors like for example inflation rate, interest changes and bond yields are modeled as first or second order AR-models, in some cases including exogenous variables as well.

An important reason for the differences between the GARCH-BEKK model I use and the Watson Wyatt ALM model is that the latter was developed with long term analysis in mind, while the former is designed to determine what happens during a one year period. As such, the ALM model gives yearly returns while I generate yearly returns based on twelf monthly returns. Properties like heteroskedasticity are more smoothened out as the lengthy of the period increases. Therefore, it is very common in practice to use models like the Watson Wyatt ALM model for long term analysis. Also, the differences between the models do not necessarily imply that they will yield very different results. In fact, under the conditions of the analysis I perform in this paper I would expect both models to yield similar results, assuming both are estimated in a theoretically sound manner.

10 Analysis of the FTK solvability assessment (II)

10.1 Simulation results

I used the Watson Wyatt ALM model to perform a analysis similar to the one in section 8. However, since the Watson Wyatt ALM model is only capable of producing 1000 scenarios, I will need appropriate critical values to the hypothesis testing. These values are listed in table 10.

Sign. level	Critical value
5%	96.70%
1%	96.30%
0.1%	95.80%

Table 10: Critical values for a success rate of at least 97.5%

While I listed a few additional critical values to give insight into the effect of the significance level on the analysis, I will still test the null hypothesis of a success rate of at least 97.5% against the alternative hypothesis that the success rate is less than 97.5% based on a significance level of 1%. I will limit the discussion of the simulation results to the most interesting cases. Note that because of the nature of the ALM model, I cannot distinguish between different starting situations like I did during the analysis with the GARCH-BEKK model.

The effect of duration (mis)matching

Pension fund assumptions				Analysis results		
Investment Portfolio	Duration Liabilities	Duration Assets	Required Funding rate	< Req. Res.	> 105%	> 100%
Safe	15	5	118.8%	54.90%	86.70%	92.00%
Safe	15	15	111.9%	68.80%	77.10%	85.50%
Neutral	25	5	132.1%	25.80%	97.10%	98.70%
Neutral	25	15	125.8%	38.90%	92.90%	97.10%

Table 11: Success rate of the FTK required reserve

The results in table 11 confirm the results I found using the GARCH-BEKK model. For both the young and old fund the performance of the required reserve decreases as durations are better matched. Again, this indicates that the reduction of the risk as a result of the duration matching is overestimated in the required reserve. Moreover, the performance of the reserve is significantly worse for the old fund compared to the results in table 7. The results for the young fund seem comparable.

The effects of asset allocation on the performance of the required reserve

In my analysis with the GARCH-BEKK model I discussed the effect of using a different asset allocation on the performance of the FTK required reserve. I did the same with the Watson Wyatt ALM model, some results are shown in table 12.

Pension fund assumptions				Analysis results		
Investment Portfolio	Duration Liabilities	Duration Assets	Required Funding rate	$<$ Req. Res.	$>$ 105%	$>$ 100%
Safe	15	5	118.8%	54.9%	86.7%	92.0%
Neutral	15	5	124.9%	40.8%	92.0%	96.6%
Risky	15	5	134.0%	22.9%	97.5%	99.3%
Neutral	25	15	125.8%	38.9%	92.9%	97.1%
Risky	25	15	137.5%	17.2%	98.3%	99.4%

Table 12: Success rate of the FTK required reserve

Again these results confirm those found during the analysis with the GARCH-BEKK model. As the investment portfolio changes from containing mostly bonds to containing mostly risky assets, the performance of the required reserve increases. This means that both my own model and the Watson Wyatt ALM model indicate that the required reserve handles investment risk better than interest risk. Moreover, both models indicate that the reserve kept for investment risk results in a success rate above 97.5% if investment risk is of relatively large enough importance, while the opposite holds for the reserve required for interest risks.

10.2 Summary

There is very little to discuss about the results found during the analysis with the Watson Wyatt ALM model, mainly because they seem to confirm each of the results previously observed and discussed for the analysis done with the GARCH-BEKK model. Therefore I will safe the actual summary of these results for the final conclusions. The fact that the results of both models are consistent with each other greatly increases the reliability of the drawn from those results conclusions. This is in itself an important result, making the analysis done with this model certainly worthwhile.

11 Summary and Conclusion

An important part of the Dutch pension regulations is the *Financieel Toetsingskader*, or FTK. In this paper I focus on a specific part of these regulations, specifically the solvability assessment that determines a reserve that pension funds need to keep on top of the value of their liabilities. This required reserve is intended to ensure that the pension fund is able to keep its funding rate (the ratio between the assets of the pension funds and its liabilities) above a minimal level of 105% in most cases even if things go wrong. The goal that was set for this required reserve is very explicit: it is supposed to keep the funding rate of pension funds above 105% in *at least* 97.5% of the cases after a one year period.

According to the analysis I describe in this paper, which is based not only on a GARCH-BEKK model estimated from historical data of several different asset classes from a period of 15 years but also on the Asset Liability Model developed by Watson Wyatt, this goal is not met in many cases. While the two models are based on different type and amounts of data and use different theoretical structures, the analyses done with each model yield similar results. Several observations can be made from which the following conclusions can be drawn.

The FTK required reserve mainly seems to perform below the intended success rate of 97.5% due to an underestimation of interest risk. The underestimation occurs in two ways. First of all, there seems to be a structural underestimation of the interest risk. The performance of the FTK required reserve significantly decreases when a larger portion of a pension fund's investment portfolio becomes invested in bonds. The same holds for when the durations of the liabilities of a pension fund increase. Therefore when interest risk becomes a relatively larger factor for a pension fund, the underestimation of the interest risk also affects the performance of the required reserve more. This translates in many cases to a success rate of the FTK required reserve well below the goal of 97.5%.

The FTK solvability assessment also seems incapable of correctly estimating relative interest risks when pension funds change the durations of their assets. As pension funds change the durations of their assets to better match the durations of the liabilities, in order to reduce interest risk, the FTK required reserve decreases in an attempt to reflect this reduction in risk. However, my analysis indicates that the reduction in the required reserve is too large; the performance of the FTK required reserve tends to decrease as the durations are matched more closely. This especially seems to be the case for relatively durations below 15 years. There is some evidence that the performance of the FTK required reserve increases again when the matching takes place with higher average durations, suggesting that the interest risk for higher durations is estimated better than the interest risk for lower durations.

While not unexpected, my analysis also shows that the success rate of the FTK required reserve decreases when the volatility and/or the correlations between asset classes are relatively high at the start of a period. This means that in addition to the issues described above, the reliability of the required reserves diminishes even more in times of economic unrest, like for example during the economic events and subsequent drops in equity values and interest rates that took place in the second half of 2008.

I also tried to determine how much higher the required reserve should be to be sufficient to meet the 97.5% success rate requirement under average circumstances, as well as the additional increase that would be needed to make the reserve reliable under extreme conditions as well. It seems that this highly depends on the pension fund in question. Funds for which the funding rate is less depending on changes in the interest rate, either due to the liabilities having a low duration and/or the investment portfolio containing a relatively small porton of bonds, an increase of 5-10% of the required reserve (in terms of the original reserve) may suffice to bring the success rate of the FTK required reserve in the intended range above 97.5%. However, increases up to 50% may be needed for funds that are more vulnerable to interest risk.

12 Recommendations for further research

Based on the results in this paper, I suggest that research into a better estimation of interest risk for the FTK required reserve is in order. The results of my analysis indicate that the required reserve that results from the calculation of S_1 is not sufficient. This in turn means that it is likely that the factors that are used to determine the interest risk in S_1 are not extreme enough, especially for the lower durations. A more detailed study into the correct values for those factors might give valuable insight into the actual interest risks a pension fund in exposed to, along with offering a better approximation of a sufficiently performing required reserve.

It may also be of interest to include an analysis of the parts of the required reserve I did not analyze in this paper: credit risk, currency risk and technical insurance risk. Given the results in this paper concerning interest risk, especially additional research in credit risk seems of interest.

I did not give much attention to the risks involved in commodity investments. As is obvious from the stylized pension fund DNB used to calibrate the FTK assessments, for the large majority of pension funds this asset class barely plays a roll at all in their investment portfolio. While I have not encountered any pension funds that do have a significant portion of their investments in commodities, this might change in the future or there may be pension funds who already currently have a large position in commodities. If so, a more detailed study into the sufficiency of this particular part of the FTK solvability assessment can be of interest.

Similarly, I only discussed the performance of S_2 as a whole. However this factor consists out of seperate calculations for four different asset classes. While the portfolios I used in my analysis are designed to be close to common types of portfolios pension funds have, not all pension funds have similar portfolios. Therefore the performance of the FTK required reserve may differ from my analysis for individual pension funds. Obviously by inputting the specifics of individual funds in my model, an analysis can be performed that is more tailored to that specific fund. Doing so will give a more detailed indication of how the FTK required reserve performs for specific pension funds.

A final suggestion I have for further research is related to my earlier discussion about the interpretation of the 97.5% success rate target. By using the full extent of the historical data I collected, the performance of the FTK required reserve could be determined based on the economic starting situation of every month in the past 15 years. By doing so a detailed analysis can be done to estimate the average performance of the FTK required reserve.

While the analysis done in this paper clearly suggests that there are many situations in which the required reserve is insufficient, I did not discuss nor research how the FTK performs on average. Depending on the spread of the types of starting situations that can occur, it is possilbe that the average performance of the FTK is different from the performance discussed in this paper.

If this is indeed the case, an adjustment on the required reserve based on the current state of the economy may prevent pension funds from keeping excessively high reserves in times of stability, as well as ensuring that the reserves are high enough during times of instability. Whether or not such an adjustment is feasible in practice is definitly not an easy question to answer but offers very interesting research opportunities.

References

Aggarwal, R., Inclan, C. and Leal, R. (1999) ‘Volatility in Emerging Stock Markets’, *The Journal of Financial and Quantitative Analysis*, vol. 34, no. 1, pp. 33-55.

Baillie, R.T. and Myers, R.J. (1991) ‘Bivariate Garch Estimation of the Optimal Commodity Futures Hedge’, *Journal of Applied Econometrics*, vol. 6, no. 2, pp. 109-124.

Beirne, J., Caporale, G. M., Schulz-Ghatlas, M. and Spagnolo, N. (2008) ‘Volatility Spillovers and Contagion from Mature to Emerging Stock Markets’, *IMF Working Paper*, no. WP/08/286.

Bohl, M. and Siklos, P. (2004) ‘Empirical Evidence on Feedback Trading in Mature and Emerging Stock Markets’, *Quantitative Finance Research Centre Research Paper*, no 137.

Bollerslev, T. (1986) ‘Generalized Autoregressive Conditional Heteroskedasticity’, *Journal of Econometrics*, vol. 31, no. 3, pp. 307-327.

Bollerslev, T., Chou, R.Y. and Kroner, K.F. (1992) ‘ARCH modeling in finance: A review of the theory and empirical evidence’, *Journal of Econometrics*, vol. 52, pp. 5-59.

Broeders, D., Chen, A. (2008) ‘Pension regulation and the market value of pension liabilities - a contingent claims analysis using Parisian options’, *DNB Working Paper*, no. 183, Amsterdam: De Nederlandsche Bank.

Brouwer, R. (2005), ‘Financieel Toetsingskader: afweging en mogelijkheden verlengen looptijd obligatieportefeuille’, *Review magazine*, vol. 5.

Cont, R. (2001) ‘Empirical properties of asset returns: Stylized facts and statistical issues’, *Quantitative Finance*, vol. 1, pp. 223-236.

Davies N. and Newbold P. (1979) ‘Some power studies of a portmanteau test of time series model specification’, *Biometrika*, no. 66, pp. 153–156.

Davis, E.P., Grob, S. and Haan, L. de (2007) ‘Pension fund finance and sponsoring companies; empirical evidence on theoretical hypothesis’, *DNB Working Paper*, no. 158, Amsterdam: De Nederlandsche Bank.

De Nederlandsche Bank (2005), ‘De invloed van het Financieel Toetsingskader op de financiële marketen’, *Kwartaalbericht Juni*, Amsterdam/Apeldoorn: De Nederlandsche Bank

De Nederlandsche Bank (2005), *Vaststelling methode rentetermijnstructuur*, Amsterdam/Apeldoorn: De Nederlandsche Bank.

De Nederlandsche Bank (2006), *Advies inzake onderbouwing parameters FTK*, Amsterdam/Apeldoorn: De Nederlandsche Bank.

De Nederlandsche Bank (2008), *Kwartaalbericht December*, Amsterdam/Apeldoorn: De Nederlandsche Bank

De Nederlandsche Bank (2009), *Kwartaalbericht Maart*, Amsterdam/Apeldoorn: De Nederlandsche Bank
De Santis, G. and Imrohoroglu, S. (1997) 'Stock returns and volatility in emerging financial markets', *Journal of International Money and Finance*, vol. 16. no. 4, pp. 561-579.
Engle, R.F. (1982) 'Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of UK inflation', *Econometrica*, no. 50, pp. 987-1008.
Engle, R.F., Lilien, D.M. and Robins, R.P. (1987) 'Estimating Time Varying Risk Premia in the Term Structure: The Arch-M Model', *Econometrica*, vol. 55, no. 2, pp. 391-407.
Engle, R.F. (2001) 'GARCH 101: The Use of ARCH/GARCH Models in Applied Econometrics', *Journal of Economic Perspectives*, vol. 15, no. 4, pp. 157-168.
Fama, E.F. (1965) 'The Behaviour of Stock Market Prices', *Journal of Business*, vol. 38, pp 34-105.
Forbes, K. and Rigobon, R. (1999) 'No contagion, only interdependence: Measuring stock market co-movements', *NBER Working Paper*, no. 7267.
Gerber, H.U. (1997), 'Life Insurance Mathematics', 3rd edition, Berlin: Springer.
Hansen, P.R. and Lunde, A. (2005) 'A forecast comparison of volatility models: does anything beat a GARCH(1,1)?', *Journal of Applied Econometrics*, vol. 20, no. 7, pp. 873-889.
Hoek, H. (2007) 'An ALM Analysis of Private Equity', *OCFR Working Paper Series*, Applied Paper No. 2007-01.
Hoekert, W. (2007) "Het FTK als exportprodukt", *Actuariaatcongres Magazine*, December edition, pp 17-19.
Johansen, S. (1988) 'Statistical Analysis of Cointegrating Vectors', *Journal of Economic Dynamics and Control*, no. 12, pp. 231-254.
Johnston, J. and DiNardo, J. (1997), *Econometric Methods*, 4th edition, Singapore: McGraw-Hill Book Co.
Kakes, J. (2006) 'Financial behavior of Dutch pension funds: a disaggregated approach', *DNB Working Paper*, no. 108, Amsterdam: De Nederlandsche Bank.
Karolyi, G.A. (1995) 'A Multivariate Garch Model of international transmissions of Stock Returns and Volatility: The Case of the United States and Canada', *Journal of Business and Economic Statistics*, vol. 13, no. 1, pp. 11-25.
Keating, C. (2006a) 'The True Cost of the FTK: Part One', *SEI Research Paper*, London: SEI Investments (Europe) Limited.
Keating, C. (2006b) 'The True Cost of the FTK: Part Two', *SEI Research Paper*, London: SEI Investments (Europe) Limited.

Koehler, A.B. and Murphree, E.S. (1988) ‘A comparison of the Akaike and Schwarz Criteria for Selecting Model Order’, *Applied Statistics*, vol. 37, no. 2, pp. 187-195.
Ljung, G. M. and Box G.E.P. (1976) ‘On a measure of lack of fit in time series models’, *Biometrika*, no. 65, pp. 297–303.
Loretan, M. and English, W.B. (2000) ‘Evaluating “correlation breakdowns” during periods of market volatility’, *International Finance Discussion Paper*, no. 658.
Mandelbrot, B.B. (1963) ‘The Variation of Certain Speculative Prices’, *Journal of Business*, vol. 36, pp. 394-419.
McCurdy, T.H. and Morgan, I.G. (1988) ‘Testing the martingale hypothesis in Deutschmark futures with models specifying the form of heteroskedasticity’, *Journal of Applied Econometrics*, vol. 3, pp 187-202.
McNeil, A.J., Frey, R. and Embrechts, P. (2005), *Quantitative Risk Management*, Princeton: Princeton University Press.
Michelfelder, R.A. and Pandya, S. (2005) ‘Volatility of stock returns: emerging and mature markets’, *Managerial Finance*, vol. 31, no. 2, pp. 66-86.
Ministerie van Sociale Zaken en Werkgelegendheid (2006a), ‘Regeling parameters pensioenfondsen’, *Staatscourant*, no. 250, p. 38.
Ministerie van Sociale Zaken en Werkgelegendheid (2006b), ‘Regeling Pensioenwet en Wet verplichte beroepspensioenregeling’, *Staatscourant*, no. 253, p. 14.
Nelson, D. (1991) ‘Conditional heteroskedasticity in asset returns: A new approach’, *Econometrica*, no. 59, pp. 347-370.
Newbold P., Carlson, W.L. and Thorne B. (2003), *Statistics for Business and Economics*, 5th edition, Upper Saddle River: Pearson Education Inc.
Pensioen- en Verzekeringskamer (2001), *De uitgangspunten voor een financieel toetsingskader*, Apeldoorn: Pensioen- en Verzekeringskamer.
Pensioen- en Verzekeringskamer (2002), *De planning voor een financieel toetsingskader*, Apeldoorn: Pensioen- en Verzekeringskamer.
Pensioen- en Verzekeringskamer (2004), *Consultatie document financieel toetsingskader*, Apeldoorn: Pensioen- en Verzekeringskamer.
Pensioen- en Verzekeringskamer (2004), *Beleidsregel Toepassing Hoofdlijnennota*, Apeldoorn: Pensioen- en Verzekeringskamer.
Rabemananjara, R. and Zakoïan, J.M. (1993) ‘Threshold Arch Models and Asymmetries in Volatility Threshold Arch Models and Asymmetries in Volatility’, *Journal of Applied Econometrics*, vol. 8, no. 1, pp. 31-49.

Rice, J.A. (1995), *Mathematical Statistics and Data Analysis*, 2nd edition, Belmont: Duxbury Press.

Said, S.E. and Dickey, D.A. (1984) 'Testing for Unit Roots in Autoregressive-Moving Average Models of Unknown Order', *Biometrika*, no. 71, pp. 599-607.

Siegman, A. (2006) *Bespreking 'The true cost of the FTK' (SEI Research), Con Keating*, [Online], Available: `http://staff.feweb.vu.nl/asiegmann`, May 19 2009.

Siegmann, A. (2008) 'Minimum Funding Ratios for Defined-Benefit Pension Funds', *DNB Working Paper*, no. 180, Amsterdam: De Nederlandsche Bank.

Simon, C.P. and Blume, L. (1994), *Mathematics for Economists*, 1st edition, New York: W.W. Norton & Company Inc.

Susmel, R. (2000) 'Switching Volatility in Private International Equity Markets', *International Journal of Finance and Economics*, vol. 5, no. 4, pp. 265-283.

Whitelaw, R.F. (1994) 'Time Variations and Covariations in the Expectation and Volatility of Stock Market Returns', *The Journal of Finance*, vol. 49, no. 2, pp. 515-541.

Zaffaroni, P. (2000) 'Stationarity and memory of ARCH models(∞)', *London School of Economics and Political Science STICERD Research Paper Series*, no. EM383.

A Calculation of the DNB Interest Yield Curve

In this section I will briefly explain the techniques use by DNB to construct the interest yield curve. I used the exact same methods and assumptions in the derivation of yield curves predating those published by DNB. This section is largely based on the DNB publication 'Vaststelling methode rentetermijnstructuur' on March 10, 2005.

The data used in the construction of the yield curves are the interest rates of the 6 month EURIBOR swap curve, specifically the 'London composite bid rates' with a duration of 1 to 10, 12, 15, 20, 25, 30, 40 and 50 years obtained from the data source Bloomberg[48]. Before continuing I will explain what interest swaps are and why they exist.

Interest swaps are financial contracts in which a long position with a fixed interest rate is exchanged for a short position with varying interest rate, or vice versa. In the case of the swaps used in the yield curve calculation, a long term interest rate is swapped for a interest rate matching a 6 month duration. Usually this is done in such a way that at the moment the 'swap' takes place, the value of the underlying cash flows is equal so no payment is needed from either party.

The DNB interest yield curve is partly determined as a yearly curve consisting of forward rates which are then recalculated into spot rates. Forward interest rates are interest rates that will be payed in the future over a certain period. This concept is best explained with an example. Define f_{t_1,t_2} as the forward interest from time t_1 until time t_2. Suppose $f_{0,1}$, $f_{1,2}$ and $f_{2,3}$ denote yearly forward rates and your starting capital is 100 euro.

- After one year you will have a capital of 100*(1+$f_{0,1}$)
- After two years you will have a capital of 100*(1+$f_{0,1}$)*(1+$f_{1,2}$)
- After three years you will have a capital of 100*(1+$f_{0,1}$)*(1+$f_{1,2}$)*(1+$f_{2,3}$)

Spot rates are interest rates that will give you constant returns each year. Define z_t as the spot rate with duration t years. Assuming the market is efficient such that there are no arbitrage opportunities, the three year spot rate z_3 should give the same return after three years as the three combined forward rates in the example above. In other words, we should have

$$(1 + z_3)^3 = (1 + f_{0,1}) * (1 + f_{1,2}) * (1 + f_{2,3})$$

[48]Bloomberg is a financial computer system that provides real time financial data. Access to this system is a financial service offered by the company Bloomberg L.P. an is used on a daily basis by a wide range of companies and organizations.

Spot rates are typically more convenient to work with. In the FTK solvability assessment, the average duration of the assets and liabilities is calculated and linked to the spot rate with that duration in order to estimate the effect of an interest change. While this is only an approximation, it requires considerably less calculation than calculating the exact effect of interest changes on the underlying cash flows.

Define r_t as the swap rate with duration t. The cash flows of the underlying fixed interest obligation (assuming a swap value of 1) are:

Time (t)	1	2	...	$t-1$	t
Cash flow	r_t	r_t	...	r_t	r_t+1

The one year spot rate should give the same return as the one year swap rate, so:

$$(1+z_1) = (1+r_1)$$

This simply means that $z_1 = r_1$. The two year spot rate is calculated by calculating the present value of the two year swap using the spot rates, and setting it equal to 1:

$$\frac{r_2}{1+z_1} + \frac{1+r_2}{(1+z_2)^2} = 1$$

Since z_2 is the only unknown in this equation it can be solved to determine the value of z_2. Once this value is known, the value of z_3 can be calculated by means of the three year swap rate. This process continues up to a duration of 10 years. In order to calculate the 11 and 12 year duration spot rates, the forward rates during that time are assumed to be constant:

$$\begin{aligned} (1+z_11)^{11} &= (1+z_10)^{10} * 1 + f_{10,11}) && (6) \\ (1+z_12)^{12} &= (1+z_10)^{10} * (1+f_{10,11}) * (1+f_{11,12}) \\ (1+z_12)^{12} &= (1+z_10)^{10} * (1+f_{10,11})^2 && (7) \end{aligned}$$

(6) and (7) can be substituted in the equation used to calculate the present value of the 12 years swap:

$$\begin{aligned} 1 &= \frac{r_{12}}{1+z_1} + \frac{r_{12}}{(1+z_2)^2} + \ldots + \frac{r_{12}}{(1+z_{11})^{11}} + \frac{1+r_{12}}{(1+z_{12})^{12}} \\ &= \frac{r_{12}}{1+z_1} + \frac{r_{12}}{(1+z_2)^2} + \ldots + \frac{r_{12}}{(1+z_{11})^{10}(1+f_{10,11})} + \frac{1+r_{12}}{(1+z_{12})^{10}(1+f_{10,11})^2} \end{aligned}$$

This equation can now be solved numerically for $f_{10,11}$ which then can be used to establish the value of z_{11} and z_{12}. A similar procedure is used to calculate the subsequent spot rates.

B Stationarity

Several types of stationarity exist. The most commonly used type of stationarity is called *covariance stationarity* or *weak stationarity*. As the name weak stationarity suggests, it only puts some very general restrictions on a model or time series. The name 'covariance stationarity' gives an indication about the goal of these restrictions. Definitions of covariance stationarity can be found in any basic text of time series modeling. A time series $X(t)$ is stationary if:

$$
\begin{aligned}
E(X(t)) &= \mu \;\forall\; t \in \Re && (8)\\
E((X(t)-\mu)(X(s)-\mu)') &= E((X(t+\tau)-\mu)(X(s+\tau)-\mu)') \;\forall\; \tau \in \Re && (9)
\end{aligned}
$$

In other words, the *unconditional* (co)variances and mean of the time series needs to be constant over time. Note that nothing is said about the *conditional* (co)variances and mean, those are still allowed to vary over time. For example, by design in any GARCH model the conditional variance (volatility) varies over time. However, it can still be covariance stationary under certain conditions. I will illustrate this with an example. In the univariate GARCH(1,1) model[49] the unconditional variance of the returns in period t is:

$$
\begin{aligned}
\mathrm{Var}(r_t) &= \mathrm{Var}(\mu + \epsilon_t)\\
&= \mathrm{Var}(\epsilon_t)\\
&= E(\epsilon_t^2) - E(\epsilon_t)^2\\
&= E(\epsilon_t^2) - 0\\
&= E(\sigma_t^2 u_t^2)\\
&= E(\sigma_t^2)\\
&= E(c + \alpha\epsilon_{t-1}^2 + \beta\sigma_{t-1}^2\\
&= c + \alpha E(\epsilon_{t-1}^2) + \beta E(\sigma_{t-1}^2)\\
&= c + \alpha \mathrm{Var}(\epsilon_{t-1}^2) + \beta \mathrm{Var}(\epsilon_{t-1}^2) && (10)
\end{aligned}
$$

In order for (9) to hold, we need to have that $\mathrm{Var}(r_t) = \mathrm{Var}(r_{t-1}) = \sigma \;\forall\; t \in \Re$. Using (10) and substituting σ this means we should have:

[49]See equation (3)

$$\sigma = c + \alpha\sigma + \beta\sigma \quad (11)$$

which implies

$$\sigma = \frac{c}{1 - \alpha - \beta} \quad (12)$$

In order for the model to make any sense, we need that $\sigma > 0$. Therefore in order for the univariate GARCH(1,1) model to satisfy (9) we should have $\alpha + \beta < 1$. It is easy to show that this is the only condition that needs to be satisfied in order for a GARCH(1,1) model to be stationary. Every GARCH model satisfies (8):

$$\begin{aligned} E(r_t) &= E(\mu + \epsilon_t) \\ &= \mu + E(\epsilon_t) \\ &= \mu + E(\sigma_t u_t) \\ &= \mu + E(\sigma_t)E(u_t) \\ &= \mu + E(\sigma_t) * 0 \\ &= \mu. \end{aligned}$$

.

The stationarity conditions for the GARCH-BEKK model were given in section 5.2, and the stationarity of the estimated model is discussed in section 6.

C Estimation Results

In this appendix I will present and discuss the parts of the estimated model that were not discussed in the main text. First of all I consider the constant μ in the returns equation. Its estimated values are given in table 13.

Asset type	estimated value	p-value
Mature markets	0.003	39.3%
Emerging markets	0.016	4.6%
Private equity	0.002	67.47%
Commodities	0.003	19.51%
Interest (dur. 5)	0.000	97.3%
Interest (dur. 10)	0.000	65.3%
Interest (dur. 15)	0.000	35.9%
Interest (dur. 25)	0.000	26.2%

Table 13: Estimation results for the constant matrix μ

Clearly the constants are not significantly different from zero in most cases. The large p-values for the constants of the interest rates indicate that they play no significant role in the outcome of the model, and might as well be considered to be zero. It is interesting to see that for each of the risky assets the constants are positive. This indicates that the model will slightly lean to positive returns. Moreover, this effect seems to be largest for the as riskiest perceived asset types emerging markets. Intuitively this seems consistent with reality, since for each asset class we normally expect to make a profit in the long term, and riskier investments tend to give a higher return on average.

A large part of the estimated values for the constant matrix C are nearly zero and not significant. Moreover, because of the very small values the effect of the matrix C on the outcome of the model is not of much interest to discuss any further, nor does listing the value for each individual element of C yield any additional insights.

The values for the diagonal elements of the matrix A are listed in table 14. The coefficients are all significant at a 1% confidence level, with the exception of the coefficient corresponding to the asset class commodities. The most interesting remark about these outcomes is that there is only one negative value in the matrix A, the element corresponding to the asset class commodities. While for the volatility of the asset class commodities itself the sign of this element does not make any difference, it does for the correlation

Element (Corresp. Asset)	value	p-value
a_{11} (Mature markets)	0.164	0.0%
a_{22} (Emerging markets)	0.298	0.0%
a_{33} (Private equity)	0.342	0.0%
a_{44} (Commodities)	-0.077	5.5%
a_{55} (interest dur. 5)	0.240	0.0%
a_{66} (interest dur. 10)	0.307	0.0%
a_{77} (interest dur. 15)	0.352	0.0%
a_{88} (interest dur. 25)	0.424	0.0%

Table 14: Estimation results for the constant matrix A

between commodities and the other asset classes. Since this is the only negative element in the matrix A, large innovations will have a reducing effect on the correlation between commodities and the other asset classes, which is the opposite of what we observe between the other asset classes. Still, this negative value is also the smallest (in absolute value) and the only one that is insignificant.

The values for the diagonal elements of the matrix B are listed in table 15.

Element (Corresp. Asset)	value	p-value
b_{11} (MM)	0.959	0.0%
b_{22} (EM)	0.760	0.0%
b_{33} (PE)	0.944	0.0%
b_{44} (COM)	1.000	0.0%
b_{55} (interest dur. 5)	0.973	0.0%
b_{66} (interest dur. 10)	0.957	0.0%
b_{55} (interest dur. 15)	0.946	0.0%
b_{66} (interest dur. 25)	0.928	0.0%

Table 15: Estimation results for the constant matrix B

Obviously these parameters are all significant at any sensible significance level. Note that most values are rather close to one. This indicates that the volatility of past periods is rather persistent and as such plays a major role in the returns.

D Volatility assumptions

This appendix contains the values of the (co)variance matrices used in the analysis as the initial volatility at the start of the year for each scenario. Abbreviations mean the following:

- MM = Mature markets.
- EM = Emerging markets.
- PE = Private Equity.
- COM = Commodities.
- Δ I5 = Interest 5 year duration.
- Δ I10 = Interest 10 year duration.
- Δ I15 = Interest 15 year duration.
- Δ I25 = Interest 25 year duration.

Due to formatting considerations and readability, all values in the tables in this appendix have been multiplied by a factor of 10^3 and subsequently rounded to three decimals. Therefore, the actual values used in the simulation are the unrounded version of those in the tables, divided by one thousand. Note that the (co)variances in the tables are the (co)variances of the log returns or in the case of the interest rates, the (co)variances of the first difference of the interest rates.

	MM	EM	PE	COM	Δ I5	Δ I10	Δ I15	Δ I25
MM	1.367	1.889	1.767	0.034	0.029	0.015	0.009	0.005
EM	1.889	6.274	2.397	-0.040	0.027	0.012	0.004	-0.001
PE	1.767	2.397	3.506	-0.009	0.042	0.018	0.005	-0.010
COM	0.034	-0.040	-0.009	4.100	0.006	0.003	0.003	0.004
Δ I5	0.029	0.027	0.042	0.006	0.003	0.002	0.002	0.002
Δ I10	0.015	0.012	0.018	0.003	0.002	0.002	0.002	0.002
Δ I15	0.009	0.004	0.005	0.003	0.002	0.002	0.002	0.002
Δ I25	0.005	-0.001	-0.010	0.004	0.002	0.002	0.002	0.003

Table 16: Initial (co)variance matrix under starting situation A (December 2007).

	MM	EM	PE	COM	Δ I5	Δ I10	Δ I15	Δ I25
MM	1.545	2.105	2.350	0.070	0.043	0.028	0.020	0.016
EM	2.105	6.898	3.144	-0.020	0.064	0.047	0.036	0.034
PE	2.350	3.144	5.602	-0.010	0.072	0.052	0.036	0.028
COM	0.070	-0.020	-0.010	4.400	0.010	0.003	0.003	0.002
Δ I5	0.043	0.064	0.072	0.010	0.005	0.004	0.003	0.002
Δ I10	0.028	0.047	0.052	0.003	0.004	0.003	0.003	0.002
Δ I15	0.020	0.036	0.036	0.003	0.003	0.003	0.003	0.002
Δ I25	0.016	0.034	0.028	0.002	0.002	0.002	0.002	0.002

Table 17: Initial (co)variance matrix under starting situation B (June 2008).

	MM	EM	PE	COM	Δ I5	Δ I10	Δ I15	Δ I25
MM	2.160	5.025	3.372	0.031	0.019	-0.002	-0.009	-0.002
EM	5.025	23.346	10.395	-0.664	0.138	0.080	0.045	0.020
PE	3.372	10.395	7.859	-0.272	0.045	0.015	0.000	0.028
COM	0.031	-0.664	-0.272	1.119	-0.009	-0.008	-0.007	-0.010
Δ I5	0.019	0.138	0.045	-0.009	0.005	0.003	0.002	0.002
Δ I10	-0.002	0.080	0.015	-0.008	0.003	0.003	0.003	0.002
Δ I15	-0.009	0.045	0.000	-0.007	0.002	0.003	0.003	0.002
Δ I25	-0.002	0.020	0.028	-0.010	0.002	0.002	0.002	0.004

Table 18: Initial (co)variance matrix under starting situation C (September 1998).

	MM	EM	PE	COM	Δ I5	Δ I10	Δ I15	Δ I25
MM	1.995	3.321	3.554	-0.061	0.029	0.016	0.013	0.027
EM	3.321	11.394	5.885	-0.852	0.050	0.037	0.039	0.078
PE	3.554	5.885	8.984	-0.049	0.009	0.004	0.007	0.052
COM	-0.061	-0.852	-0.049	4.799	-0.006	-0.008	-0.008	-0.011
Δ I5	0.029	0.050	0.009	-0.006	0.005	0.004	0.003	0.002
Δ I10	0.016	0.037	0.004	-0.008	0.004	0.003	0.002	0.002
Δ I15	0.013	0.039	0.007	-0.008	0.003	0.002	0.002	0.002
Δ I25	0.027	0.078	0.052	-0.011	0.002	0.002	0.002	0.002

Table 19: Initial (co)variance matrix under starting situation D (October 2008).

E Significance levels and critical values

Often real life processes are simulated by means of statistical models, which are estimated based on real life data. However, real life datasets are typically limited in their size. This means that even if the structure of the estimated model is an exact representation of reality, this does not necessarily imply that the estimated parameters will be exactly the same as the real values of those parameters.

To illustrate this, consider a coin which has exactly 50% probability of landing as tails, and 50% probability to land as heads. Obviously this can be modeled by means of a Bernoulli distribution with probability parameter $P = 0.5$. Now suppose the parameter P is unknown and it is estimated based on a sample of 10 coin tosses. The chance of finding exactly 5 heads and 5 tails in the sample is only 24.6%. This means that its is more likely that the value of P will be estimated based on a sample which does not consist for 50% out of heads and 50% tails. More importantly, if that happens the estimated value of p will not be equal to the 'true' value of 0.5.

This needs to be taken into account when testing a hypothesis about a parameter value. For example, in this paper I want to determine if the required reserve that follows from the FTK solvability assessment is enough to keep the funding rate of a pension fund above 105% after a one year period with a 97.5% success rate. Such a theory about the true value of a parameter is called a (null) hypothesis. However, as the coin example shows, if I find a success rate of 98% or 97% during my simulations I cannot immediately conclude whether or not the hypothesis of a 97.5% success rate is met based solely on that outcome.

This is where the concepts of *p-values* and *critical values* enter the analysis. A p-value tells us what the probability is that we would get an estimated parameter value deviating *at least as much* from the hypothesis value as the currently estimated value. A p-value can be either one-sided or two-sided, depending on what the hypothesis is that is being tested. The difference between testing one or two-sided is whether or not deviations in both directions (positive and negative) are taken into account. For my thesis, I am only interested in testing whether or not a success rate of *at least* 97.5% is attained. This means that if I find an *estimated* success rate of more than 97.5%, I have no reason to assume that the *actual* success rate is less than 97.5%. If I find an estimated success rate lower than 97.5%, I will have to determine how big the probability of geting such an outcome is, given my hypothesis of a minimal success rate of 97.5% is true. This is equivalent to determining the p-value of the estimated success rate, given that the hypothesis that the success rate is exactly equal to 97.5% is true; if I reject this hypothesis, I will certainly reject the hypothesis of a larger success rate. The same thing holds the other way around: If I cannot reject this tested hypothesis, I cannot reject the hypothesis of a success rate of at least 97.5% either since this includes the tested hypothesis.

In the case of my paper the p-value is fairly easy to determine. I will use 25000 scenarios to obtain an estimate for the success rate. Given that the null hypothesis of a 97.5% success rate is true, the amount of scenarios below a funding rate of 105% I will observe is binomially distributed with $n = 25000$ trials and probability of success $p = 2.5\%$.[50] This means that the p-value corresponding to observing at least N scenarios with a funding rate below 105% is equal to:

$$p - value = \sum_{k=N}^{25000} \frac{25000!}{(25000-k)!k!} 0.025^k (0.975)^{25000-k}$$

Once the p-value is determined, a rather subjective decision has to be made: Is the observed value *unlikely enough* to occur to reject the null hypothesis? Whatever I decide to do, there are two possibilities that I make a mistake:

1. I may reject the null hypothesis while it is true. This is called a type I error, and the probability of this mistake occurring is called the significance level of the test denoted by α.

2. I may accept the null hypothesis while it is false. This is called a type II error, and its probability is denoted by β.

Obviously, if I set the acceptance region such that α decreases, β in turn will increase. It is conventional to set α in advance. Typical values are 10%, 5%, 1% or 0.1% depending on what is being tested and the importance of not making type I or type II errors. For example, suppose someone is testing the null hypothesis that a certain medicine is safe versus the hypothesis that is is not, based on a certain number of fatalities among the test subjects. Obviously, making a type I mistake is a lot less of a problem than making a type II mistake in this case. Still, if β is set so low that α becomes too high, the medicine will never be accepted for general use. Therefore it is up to the researcher to find acceptance regions such that both values are reasonable. The boundaries of the acceptance region are called the *critical values* of the test.

[50] 100%-97.5% = 2.5% chance of observing a scenario with a funding rate of less than 105%.

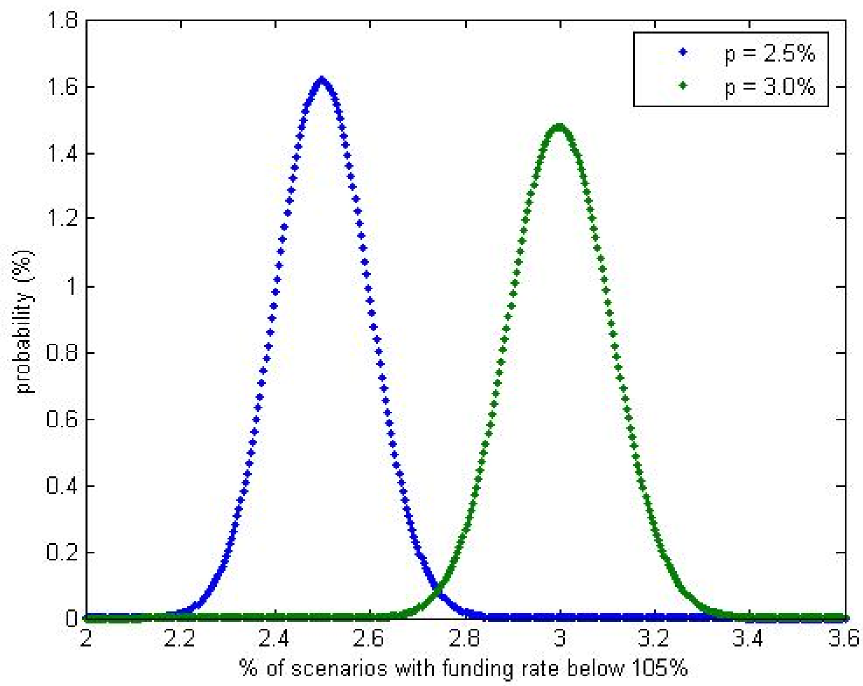
When performing an analysis like in this thesis, where the null hypothesis has to be repeatedly tested (once for every set of assumptions) it is convenient to work with critical values instead of calculating the p-value for every outcome. By simply comparing the outcome with the critical values it is easy to see whether or not the outcome of the simulation will lead the null hypothesis to be rejected or not.

As I mentioned before, the number of scenarios that will fall below 105% is binomially distributed with $n = 25000$ and probability of a funding deficit p. I want to test the null hypothesis that $p \leq 2.5\%$ versus the alternative hypothesis that $p > 2.5\%$. Suppose $\bar{p}$ is the estimated value. To be exact:

$$\bar{p} = \frac{\#\text{ scenarios below }105\%}{25000}$$

If the null hypothesis is true, small values of $\bar{p}$ are more likely than large values of $\bar{p}$. To see this I displayed the relevant parts of the probability mass function of the estimate $\bar{p}$ for the value of p under the null hypothesis and a value of p under the alternative hypothesis in figure 6.

Figure 6: Probability distributions



As figure 6 indicates, small values of $\bar{p}$ are more likely to occur if the true value of p is small. Therefore, I will only reject the hypothesis that $p \leq 2.5\%$ if the outcome of $\bar{p}$ is large enough. This is equivalent to rejecting the null hypothesis if we find 'too much' scenarios below 105%. In order to determine what 'too much' means, I determine the critical values. I will use the distribution of $\bar{p}$ under the hypothesis that $p = 2.5\%$ and use that to determine what the largest value may occur given a specific signficance level. Figure 6 shows clearly that if the true value of p deviates even slightly from the null hypothesis, the probability of the null hypothesis to get rejected is quite large. The critical values used for the testing are not of any use in this appendix, but are listed in the main text in table 6.

Readers who are interested in more information about p-values, critical values and significance levels can find additional information in any basic text about statistical testing.

F Histograms section 7.2

All graphs in this appendix were made based on the model results under the economy starting situation D, which means the covariance matrix of October 2008 was used to generate these graphs.

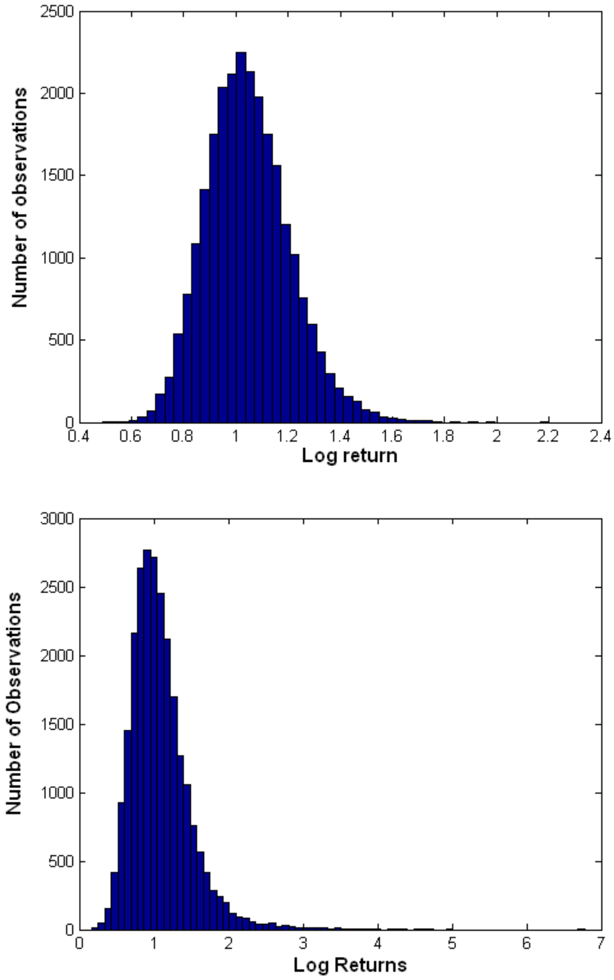


Figure 7: Histogram of the modeled returns for mature markets (above) and private equity (below)

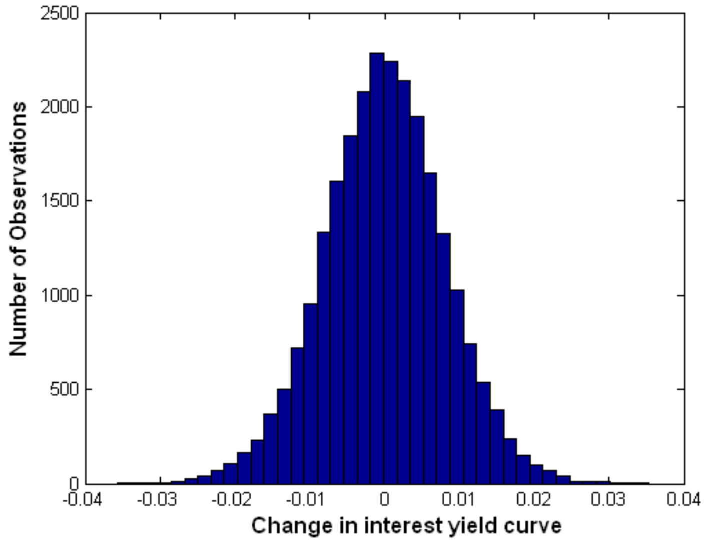


Figure 8: Histograms of the modeled changes in the 5 year duration interest

G Simulation results

This appendix contains the full extent of the simulation output I obtained from the simulation runs performed with the estimated GARCH-BEKK model, including the results already mentioned in the main text. The output is ordered first by the duration of the liabilities (D_L), then by investment portfolio, then by the duration of the bonds (D_B) and finally by the starting situation.

Noted in the table are (from left to right): The portfolio used, the duration of the liabilities, the duration of the bonds, the required reserve according to the FTK, the starting situation, the percentage of scenarios that resulted in a funding rate less than 100%, the percentage of scenarios that resulted in a funding rate less than 105%, the percentage of scenarios that resulted in a funding rate below the required funding rate, and finally the percentage of scenarios that resulted in a drop in the value of the assets (V_t). Some of these numbers were represented in a different manner in the main text in a late stadium of the writing of this thesis. Obviously this has no effect on any of the conclusions or outcomes. For readability, the percentage signs were omitted in the third and last four columns of table 20.

Table 20: Simulation results of the GARCH-BEKK model

Portfolio	D_L	D_B	**Req. F.R.**	**St. Sit.**	**$<$ 100%**	**$<$105%**	**$<$ Req. res.**	$V_1 < V_0$
Safe	15	5	118.8	A	0.968	4.104	46.68	22.552
Safe	15	5	118.8	B	1.228	4.504	46.984	23.196
Safe	15	5	118.8	C	1.188	4.240	47.156	28.888
Safe	15	5	118.8	D	0.752	3.536	46.724	27.052
Safe	15	10	115.2	A	0.72	4.34	43.60	23.83
Safe	15	10	115.2	B	0.83	4.87	43.95	23.64
Safe	15	10	115.2	C	0.78	4.56	43.93	30.09

Continued on the next page

Table 20 – continued from the previous page.								
Portfolio	D_L	D_B	**Req. F.R.**	**St. Sit.**	$<$ **100%**	**<105%**	$<$ **Req. F.R.**	$V_1 < V_0$
Safe	15	10	115.2	D	0.60	4.16	43.08	26.37
Safe	15	15	111.9	A	0.32	4.59	36.43	24.80
Safe	15	15	111.9	B	0.53	5.82	38.02	24.90
Safe	15	15	111.9	C	0.34	4.49	37.29	30.08
Safe	15	15	111.9	D	0.45	5.82	38.11	25.62
Neutral	15	5	124.9	A	0.80	2.39	43.96	30.73
Neutral	15	5	124.9	B	1.07	3.22	44.86	30.47
Neutral	15	5	124.9	C	0.93	2.93	43.66	33.66
Neutral	15	5	124.9	D	0.87	2.96	44.38	32.86
Neutral	15	10	122.1	A	0.64	2.51	42.13	29.43
Neutral	15	10	122.1	B	0.91	3.30	42.81	28.57
Neutral	15	10	122.1	C	0.79	3.06	43.14	33.08
Neutral	15	10	122.1	D	0.89	3.44	43.31	31.54
Neutral	15	15	119.5	A	0.40	2.42	39.66	28.08
Neutral	15	15	119.5	B	0.89	3.61	40.66	27.15
Neutral	15	15	119.5	C	0.79	3.32	40.19	32.52
Neutral	15	15	119.5	D	0.80	3.71	40.38	29.64
Risky	15	5	134	A	0.74	1.78	43.52	34.40
Risky	15	5	134	B	1.21	2.76	43.42	35.29
Risky	15	5	134	C	1.07	2.40	43.13	37.20
Risky	15	5	134	D	1.25	2.79	43.42	35.98
Risky	15	10	132.3	A	0.72	1.88	42.36	33.66
Risky	15	10	132.3	B	1.21	2.88	43.35	34.51
Risky	15	10	132.3	C	1.18	2.72	42.76	36.70
Risky	15	10	132.3	D	1.49	3.07	43.43	36.29
Continued on the next page								

Table 20 – continued from the previous page.

Portfolio	D_L	D_B	**Req. F.R.**	**St. Sit.**	< **100%**	<**105%**	< **Req. F.R.**	$V_1 < V_0$
Risky	15	15	130.7	A	0.75	2.02	41.66	33.02
Risky	15	15	130.7	B	1.10	2.91	42.07	32.43
Risky	15	15	130.7	C	1.08	2.78	42.43	36.61
Risky	15	15	130.7	D	1.33	3.17	41.92	34.20
Safe	25	5	126.6	A	4.32	8.59	53.92	22.18
Safe	25	5	126.6	B	3.09	6.64	53.85	23.14
Safe	25	5	126.6	C	6.52	11.48	53.19	28.77
Safe	25	5	126.6	D	3.77	7.86	53.55	27.23
Safe	25	15	118.2	A	4.68	11.34	49.96	25.68
Safe	25	15	118.2	B	3.39	9.63	50.68	24.24
Safe	25	15	118.2	C	9.00	16.98	50.18	30.03
Safe	25	15	118.2	D	5.66	12.78	50.47	25.07
Safe	25	25	113.9	A	1.27	7.60	44.92	29.33
Safe	25	25	113.9	B	1.15	6.92	44.15	26.93
Safe	25	25	113.9	C	1.79	9.38	45.17	32.86
Safe	25	25	113.9	D	1.52	8.19	44.55	25.19
Neutral	25	5	132.1	A	3.05	5.81	50.88	29.85
Neutral	25	5	132.1	B	2.50	4.99	50.65	30.27
Neutral	25	5	132.1	C	4.48	7.86	50.55	34.57
Neutral	25	5	132.1	D	3.34	6.46	50.68	32.84
Neutral	25	15	125.8	A	2.99	6.47	48.82	28.16
Neutral	25	15	125.8	B	2.49	5.94	48.54	27.12
Neutral	25	15	125.8	C	5.66	10.31	49.00	31.75
Neutral	25	15	125.8	D	1.22	4.27	45.12	27.88
Neutral	25	25	122.1	A	1.22	4.27	45.12	27.88

Continued on the next page

Table 20 – continued from the previous page.

Portfolio	D_L	D_B	**Req. F.R.**	**St. Sit.**	< **100%**	<**105%**	< **Req. F.R.**	$V_1 < V_0$
Neutral	25	25	122.1	B	1.42	4.93	45.67	25.92
Neutral	25	25	122.1	C	2.50	6.39	45.75	31.52
Neutral	25	25	122.1	D	2.09	6.18	45.79	25.66
Risky	25	5	141.3	A	1.86	3.56	48.92	34.70
Risky	25	5	141.3	B	2.17	3.84	48.85	34.90
Risky	25	5	141.3	C	3.29	5.73	48.73	37.00
Risky	25	5	141.3	D	3.22	5.41	48.79	36.31
Risky	25	15	137.5	A	2.06	3.89	47.65	32.65
Risky	25	15	137.5	B	2.13	4.21	48.36	33.11
Risky	25	15	137.5	C	3.76	6.38	47.75	34.40
Risky	25	15	137.5	D	3.40	5.76	47.75	34.40
Risky	25	25	135.1	A	1.35	3.06	46.65	31.58
Risky	25	25	135.1	B	1.69	3.68	46.62	31.17
Risky	25	25	135.1	C	2.68	5.04	45.76	34.27
Risky	25	25	135.1	D	2.84	5.07	46.62	32.00