Vrije Universiteit Amsterdam

KPMG

Master Thesis

Generational Effects of the Dutch Pension System Reformation: A Quantitative Assessment of Investment Risk Exposure - A Modelling Study

Author: Merlijn van der Lans (2671131)

1st supervisor: Prof. Dr. Ronald Meester
daily supervisor: Rick Bohte (KPMG)
2nd reader: Dr. Rene Bekker

A thesis submitted in fulfilment of the requirements for the VU Master of Science degree in Business Analytics (Financial Risk Management)

June 27, 2025

"In the end, the journey's brought joys that outweigh the pain"
from Journey of the Magi, *by Frank Turner*

Abstract

Since the reformation of the Dutch pension system was announced in July 2023, there have been numerous debates on the generational impact of this reformation. In the new system, investment risks move from the pension funds to participants, and the variable premium changes to a flat premium. The goal of this research is to investigate whether younger generations benefit more from the pension reformation than older generations. This is achieved by performing a historical data analysis to examine investment risk using Value at Risk, Expected Shortfall, and GARCH(1,1), and by simulating potential future portfolio trajectories to assess risk on a long-term basis using Monte Carlo Simulation with a Geometric Brownian Motion. A stress test analysis is conducted on these simulations. Lastly, a Practical Scenario Analysis is performed that combines the old and new systems. All models in this research are programmed in Python.

Beyond the expected finding that the risk increases with a longer time horizon, it was found that the elevated risk level following a financial crisis does not revert to the pre-crisis level within 40 years, the maximum time a participant spends collecting pension. This suggests that younger participants, who are encouraged to invest on a more risky basis according to the Dutch government, should still be cautious during periods of financial instability.

Furthermore, the Practical Scenario Analysis showed that the younger generation, spending their entire career in the new pension system, is likely to accumulate the highest pension payout after retirement. The generation that spent 30 years in the old system and had to switch to the new system for the last ten years of their working life accumulated the least amount of pension and is therefore expected to have the lowest pension payout. Generations that had already retired before the reformation was announced do not face significant disadvantages.

Contents

1

Introduction

On 1 July 2023, after years of debating, a new law was accepted in The Netherlands regarding the reformation of the pension system. This reformation is and will be a project of tremendous scale, which will take many years and involves each individual entitled to pension, and additionally every company, pension fund, administrator, asset manager, and insurer in the Netherlands. Newspapers and websites have been full of articles about this *Wet Toekomst Pensioenen (WTP)*, as it is called in Dutch. The WTP proposes that the current system that is known as the Defined Benefits (DB) system, switches to a Defined Contributions (DC) system. In this new system, instead of accruing a fixed amount of pension each month through your employer, the actual pension payout is influenced by the financial market.
Currently, the DB system consists of the following three pillars:

1. The first pillar is the General Old Age Pension, in Dutch the Algemene Ouderdomswet (AOW). To qualify for AOW, a person has to live in the Netherlands or have lived in the Netherlands. Upon retirement, they will receive the full AOW benefit monthly. The benefits are defined by the minimum wage and the number of years spent living in The Netherlands. The AOW is thus income independent and serves on a Pay As You Go (PAYG) basis. The current workers' contributions fund the current retirees. Its main purpose is to prevent poverty among retired people.

2. The second pillar is the work-related pension. This work-related pension is additional to the AOW and is built up through the employer. The work-related pension is income dependent. It is administered by a pension fund or an insurer. In the Netherlands there are over 200 different pension funds and contracts. An employee is obliged to join one of the three following pension funds: Industry-wide pension funds,

corporate pension funds, or pension funds for independent professionals. Industry-wide pension funds are the largest and hold up to 40% of total pension assets.

3. The third pillar consists of individual pension products, such as life insurance. This pillar can be used to collect extra pension savings or retire early.

The reform of the Dutch pension system is mostly concerned with the second pillar. This will be explained in more detail in Section 2.1.

1.1 Problem Statement

As mentioned before, in the New Dutch Pension Contract, often referred to as NPC, pensions are being invested into the financial market. In the old system, this was also the case, but with the key difference that only the pension funds bore the investment risks. They were obligated to ensure that each participant would receive their promised pension payout every month. In a DC scheme, where the Dutch pension system is moving towards, the exposure to investment risk shifts from the pension fund to the individual. Individual pension payouts are directly influenced by the returns on their invested pension. Therefore, individual pension payouts are now exposed to market fluctuations, recessions, changing interest rates, and political turmoil. Collectively, this known as *market risk*. In Section 2.2, market risk will be discussed in more detail.
In short, if the stocks in your investment portfolio, managed by a pension fund, perform well, there is the possibility of a higher return. On the other hand, when your portfolio is performing badly, it can be the case that the actual pension you receive is lower than it would have been in the DB system.
A factor that has a relatively large impact on the management of market risk is the investment horizon. The investment horizon is the period of time an individual will hold an asset or a portfolio. Clearly, for a generation that is close to retirement, this investment horizon will be significantly shorter than for younger generations, who have to work over 40 years before they will retire. This can have a significant impact on the outcomes of the invested pensions.

Additionally, the NPC uses a so-called flat premium (vlakke premie). In short, this means that the premium, the percentage of their (monthly or yearly) salary an employee contributes to a pension fund, stays the same for their entire career, while this was increasing over time in the old system. In Section 2.1.1, this is explained in more detail. The flat premium is mostly beneficial for younger generations as they do not switch from variable to flat premium midway through their career.

This research aims to investigate whether younger or older generations will benefit more from the pension system reformation. For this, it will be determined whether long-term investing is more profitable than short-term investing. On top of that, the effects of switching midway through a career will be investigated as well, taking into account moving from a variable to a flat premium. For this, first, risk measures and statistical models will be applied to historical data. Furthermore, a large number of potential future stock price paths will be simulated for different time horizons. This will help to understand the likelihood of certain portfolio outcomes. This model for the new system will then be connected to a simpler model for the old system, so that the different cases of switching midway can be investigated. In Section 1.4, the outline of the thesis will be discussed in detail.

Research Question: How does the Dutch pension reform affect the investment risk and retirement outcomes across generations?

Extending the research question, the following subquestions were formulated to construct an outline for this research. These subquestions will be implicitly answered throughout this report.

- What are the key differences between the old and new Dutch pension system, and how will various generations be affected by these differences?
- What factors contribute to the unequal performance of the new pension system and how do these factors influence investment strategies between generations?
- How can applying financial techniques such as VaR and GARCH(1,1) to historical data help compare and quantify the risks and benefits of long-term versus short-term pension portfolios?
- How will certain extreme market scenarios, such as financial crises, impact pension outcomes, and how do these effects differ between generations?
- How does Monte Carlo Simulation help simulate the potential long-term and short-term outcomes for pension portfolios, and does it provide insight into generational differences?

1.1.1 Scope of the Study

As the Dutch pension reformation is a broad topic, not all aspects of it can be included in this research. Therefore, the decision was made to focus on the impact of the reformation on different generations, focussing on investment risk and retirement outcomes. This will be done by analysing historical data and simulating portfolios in the future to investigate potential pension payouts. Forecasting using Machine Learning methods will not be part of this thesis.
In the literature, the research will focus on a comparison between the old and new system, a comparison to other reforming countries, and reasons for the reformation. These particular subjects are chosen because they help to substantiate the research. The literature could include subjects such as the perspective of the employer or of the pension funds, but these perspectives are outside the scope of this paper. This research has chosen to focus on the perspective of the individual. The risk measures and statistical models used will be Value at Risk, Expected Shortfall, GARCH(1,1) and Monte Carlo Simulation.
Furthermore, the historical data used for this study will range from around the year 2011 to 2023. It can be interesting to perform similar research on the same historical data for a different time frame like 1980-2000, but this is outside the scope of this research.
It should be noted that this thesis was written between January and June 2025. Therefore, the research focusses on the state of the pension system in January 2025. All potential changes and political inquiries after January 2025 will not be taken into account.

1.2 Goals of the Research

The goal of this research is not to determine whether the reformation of the pension system is justified, as the transition is already underway and irreversible. Instead, the focus is on investigating the implications of the reformation. Determining whether younger generations benefit more from the new system than older generations is one of the key goals of this research. To achieve this, this paper aims to investigate whether long-term investing is in fact more profitable than short-term investing under the new Dutch pension system and what the effect of the flat premium is. A comprehensive literature review should give insight into the key differences between the old and the new pension system and also reveal the reasons why the reformation is happening.
Based on all this, it should be possible to answer the research question of Section 1.1, taking into account the limitations of this research that will be discussed later.

1.3 Relevance for the Host Company

Pension reformation is a hot topic in the Netherlands right now. Every individual, but also every pension fund and company, must switch to the new pension system by the end of 2026. These companies often hire KPMG to help them with this reformation.
KPMG is a large consulting firm that operates in 143 different countries. They provide Audit, Advisory, Tax, and Legal services. It is a member of the so-called big four, together with Deloitte, EY, and PwC. They support companies, NGOs, governments, and communities in achieving goals such as limiting climate change and promoting sustainable growth and technological processes.
The Financial Risk Management Department consists of three different teams: Banking, Actuarial Insurance Risk (AIR), and Asset Management & Pensions.
The AIR Team (together with the Asset Management & Pensions Team), is a financial risk advisor in the Dutch pension market for pension funds, pension administrators, and employers. Specifically, the actuarial team helps with pension accounting, boarding, financial transformations, and insurance solutions for the pension industry.
Hence, investigating the pension reformation and specifically generational differences is very relevant with regard to KPMG's current projects.

1.4 Thesis Outline

This thesis will be structured as follows:
First, Section 2 will educate the reader further on the existing literature on the reformation and investment strategies of pension funds. Section 3 will discuss the risk measures and statistical models that will be implemented in this research in detail. Section 4 will explain data collection, portfolio construction, and data analysis. Furthermore, the risk measures and statistical models will be specifically tied to this research. An explanation of how they are used will be given. In Section 5, the results of the models will be visualised in graphs and tables, and notable results will be highlighted. Complementary, Section 6 will discuss the credibility and accuracy of these models, and how trustworthy the conclusions that are drawn will be. These conclusions are covered in Section 7. Here, the research question will be answered based on the results that are obtained. Lastly, Section 8 will encompass the limitations of this research and the possibilities of future research.

2

Background and Literature

This chapter will provide an in-depth explanation of the background knowledge required to understand the remaining parts of this thesis. First, background on the Dutch pension system will be given. Then, the key differences between the old and new system will be highlighted. Furthermore, a comparison will be made to pension systems in other countries. In addition, a definition of market risk will be provided. Lastly, the investment strategies of pension funds will be discussed.

2.1 The new Dutch pension system

According to the Global Pension Index of the Mercer CFA Institute (1), the Dutch pension system, after ranking first in 2023, maintained its position as the highest-ranked pension scheme worldwide in 2024, as can be seen in Figure 2.1. The ratings are based on three sub-indexes: Adequacy, Sustainability, and Integrity. The Netherlands came first on the Adequacy index: how much you actually receive. In the Sustainability and Integrity indexes, indicating whether the system is equipped to keep delivering in the future and whether the system can be trusted, respectively, the Netherlands fell outside the top three.

Figure 2.1: Highest Ranked Pension Schemes Worldwide in 2024 (Mercer, 2024)

Interestingly, the Netherlands had already ranked first in 2020, when the reformation was not yet underway. Back then, the strong second pillar, together with the strong PAYG scheme in the first pillar, contributed to the first place (2). In 2024, according to Mercer, despite the fact that the Dutch pension system is reforming from a DB to a DC scheme, it still provides good benefits, supported by strong regulations, and offers participants guidance regarding their pensions.
The similar rating of the old system in 2020 compared to the new system in 2024, raises some questions:

- Was the reformation actually necessary, as the Netherlands had the top-rated pension scheme in 2020 already?
- What is actually changing with the New Pension Contract?
- What is happening in other countries, causing their pension ratings to be less positive than in the Netherlands?

2.1.1 Key differences to the old system

In the paper *Transition to a new pension contract in the Netherlands: Lessons from abroad*, Benne van Popta and Otto Steenbeek (2021) (3) draw lessons from other pension schemes around the world, which will be discussed in Section 2.1.3. Before they discuss the pension reformations around the world, Van Popta and Steenbeek define the key differences between the old and new pension systems to be the following: the switch from DB to DC contracts, the disappearing funding ratio, the predefined distribution of collective investment returns across all plan members, the difference in risk aversion across age groups, and lastly, the possibility of lump-sum payments and other personal choice options. They also define a list of similarities to the old system, the main elements being mandatory participation, the pension funds being in charge of the pension arrangements, monthly payments, and collectivity.

The switch from DB contracts to DC pension contracts is arguably the most impactful part of the reformation. In Section 1 the current three-pillar pension system in the Netherlands was discussed. The key elements of the reformation will take place in the second pillar: the pension built up through employment. Instead of a fixed realised payout when retiring (DB), contributions are now invested in financial markets, and pensions are being paid from the return on investment (DC). The goal of the DC system is to move some of the weight of the risk from the employers towards the employees. In the old system, fluctuating markets, discount rates, and unexpected developments in longevity were risks borne primarily by pension funds. In the new DC system, individuals will have a portfolio based on their contributions. This portfolio of assets is exposed to market fluctuations. This increased reliance of individuals on the financial market is a development called *financialisation*, as Natascha van der Zwan (2017) explains in the paper *Financialisation and the Pension System: Lessons from the United States and the Netherlands* (4). According to her research, financialisation is an interconnected process that contributes to the increasing role of financial institutions, markets, and motives in today's political landscape. This is exactly what we see happening in the pension reformation. Van der Zwan compares the case of the United States with that of the Netherlands. She argues that, although the United States and the Netherlands both have highly financialised pension schemes, the process toward financialisation was very different. Where the United States eagerly started adopting investment practices, the Netherlands started investing in equities not sooner than the 1990's. Moreover, when Van der Zwan published her article in 2017, the Dutch pension system was still dominated by DB contracts, whereas the United States had already embraced DC contracts.

In Section 1.1, the flat premium was mentioned. In the old system, the premium was age dependent, meaning that the premium got higher as someone grew older. This is because the 'cost' of a pension increases with age. In the new system, the so-called flat premium means that the premium stays the same for everyone throughout their entire life. This will result in a smaller contribution later on and an extra contribution in the early stages of one's career. In Figure 2.2 on the next page, this effect is visualised.

As can be seen, this flat premium is mostly beneficial for younger generations. Combined with the longer investment horizon, where pensions will yield a return for a longer time, the portfolios of the younger generation are expected to result in a higher pension payout. For these reasons, the Dutch government is talking about compensating the elderly generation, among others, in the form of a 'solidarity reserve'. This solidarity reserve can be used to supplement pensions when the market performance is weak. There are a few restrictions regarding this reserve. The size of the fund is limited and the fund cannot become negative. This is done to reduce risk sharing between generations. However, the rules on how to apply the fund leave room for flexibility in the decision-making of pension funds (5).

Figure 2.2: Schematic View of Flat Premium (Werken Aan Ons Pensioen, 2024)

Van Popta and Steenbeek quickly mention the different levels of risk appetite across generations. Pension funds are now required to invest according to the risk preferences of the participants. However, this forms a problem, as the risk preferences of individual persons are usually unclear. The reason for this is that risk preferences are not directly observable and must be derived through surveys or preference models. Moreover, most people simply do not know what their risk preference is. Alserda, Dellaert, Swinkels, and van der Lecq (2016) investigate the impact of pension risk preferences on asset allocation in their paper *Pension risk preferences: A personalised elicitation method and its impact on asset allocation* (6). They conclude that determining risk attitude cannot be done based on population characteristics but requires measurements on a personal level. In her paper *Measurement and implementation of risk aversion by pension funds in the new Dutch pension law* by Femke Verkuijlen (2023), a Choice Sequence method or a Multiple Price List method are the best models to determine risk aversion of participants (7).

Lastly, the NPC allows for the possibility of a lump-sum payment at retirement. This payment can go up to a maximum of 10% of their accrued pension assets and is only accessible at retirement, not before or even after. The goal of this possibility is to individualise the pension contract. However, the limited amount that can be paid and the fact that this payout can only occur at retirement make it less attractive, and it is expected that many participants will not opt to use this option (5).

2.1.2 Reasons for the reformation

In their paper *Cost Saving and the Freezing of Corporate Pension Plans*, Rauh, Stefanescu and Zeldes (2020) (8) pose three categories of reasons for reforming the pension system. The first category is the change in the economic environment. Westerhout, Ponds and Zwaneveld (2024) argue in their paper *Reforming the Dutch pension system to ensure sustainability* (5) that population ageing is a contributing factor to the shift from DB to DC. According to the Dutch Central Bureau of Statistics (CBS), the Netherlands currently has more than 18 million inhabitants, of which more than 3.5 million people, or 20.5% of the population, are older than 65 years (9). CBS predicts that this grey pressure, the number of inhabitants over 65 relative to the number of 20 to 64-year-olds, will keep increasing. This can be seen in Figure 2.3 below. Westerhout, Ponds and Zwaneveld state that population ageing increases pension liabilities, as more retirees means more payout obligations.

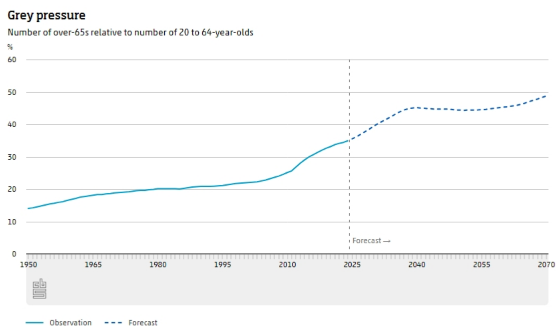


Figure 2.3: Grey Pressure in the Netherlands (CBS, 2024)

The second category is the institutional and legal changes that have occurred over the years. The DB scheme is regulated by the new Financial Assessment Framework, or 'nieuw Financieel Toetsingskader' (nFTK) in Dutch. Solvency rules in this nFTK require pension funds to hold up to 30% of the total liabilities. Pension liabilities represent the amount of money that a pension fund must pay out in the future to retirees. The ratio of total assets over liabilities is called the *funding ratio* (dekkingsgraad in Dutch), as is visible in Equation 2.1 below. The funding ratio is a key concept to understand the financial solidity of a fund. It has a direct connection to the amount of pension one receives when retired. The concept of pension indexation means that pension funds can increase payments at retirement to reflect increases in consumer goods. However, pension funds can only apply the full indexation when the funding ratio is higher than 130%, and no indexation when the funding ratio drops below 110%. When funding ratios drop below 104.6%, pension funds might even have to cut pensions (2). In recent years, these funding ratios have been declining, according to Metselaar, Zwaneveld, and van Ewijk (2022) (2), and Westerhout, Ponds, and Zwaneveld (2024)(5).

$$\text{Funding Ratio} = \frac{\text{Total value of assets}}{\text{Total value of liabilities}} \cdot 100\% \tag{2.1}$$

This drop in funding ratios can be explained by the fact that both nominal and real interest rates have also been declining. In addition, pension funds are required to use the risk-free rate to discount their liabilities. This risk-free rate has also dropped significantly, causing the present value (PV) of future liabilities to grow. This can be illustrated by a simple example. Suppose that we invest €100 today with an interest rate of 10%, then in the future (a day, a month, a year), this will be worth €110. Hence, the future value (FV), €110 discounted with a discount rate of 10% has a present value of €100 today. When this interest (or discount rate) increases, the present value today decreases, according to the formula $PV = \frac{FV}{1+r}$, where r is the discount rate. Similarly, when the interest (or discount rate) drops, the present value of future liabilities increases. When the discount rate becomes negative, more than €100 is needed to ensure a future value of €100 is met. This decrease in interest rates causes the funding ratio to drop. Similarly, when interest rates increase, the funding ratio increases as well, allowing pension funds to exercise higher indexation.

In the old system, when pension systems bore the investment risks, pensions were collectively invested. Periods of low interest rates caused pension funds to hedge these interest rate risks, especially in times of financial crisis. As explained before, younger people save their pensions on a very long-term basis, which should give them the opportunity to invest on a higher level of risk, for example, by investing more in risky equities instead of stable government bonds. In Section 2.3, this is explained in more detail. However, in the old pension system, pension funds hedging against interest rate risk caused them to hold more long-term fixed assets on behalf of all participants (collectivity), including younger participants, where other assets might have offered them a more attractive long-term prospect. This is mentioned as one of the reasons for changing the system, according to the paper *Navigating the road ahead: Tackling investment risks in the Dutch pension transition*, written by Northern Trust and True Partner Capital (2023) (10).

Furthermore, the nFTK framework is not able to solve the financial problems of pension funds caused by declining interest rates. The solvency regime was relaxed in 2019 after increasing pressure from stakeholders. A year later in 2020, interest rates had declined even further and the government needed to take action again by lowering the minimum funding ratio to 90%. These yearly needed measures gave a clear sign that a reform was necessary (5). Part of the reformation is the disappearing funding ratio.

Lastly, consumer preferences might have contributed to the reformation. In the current DB system, participants mostly obtain riskless benefits, but the downside of this is that there is no direct possibility to increase the net return upon retirement. DC schemes allow participants to choose their level of risk aversion and asset allocation (although this is different per country). It gives participants a higher level of freedom. However, this argument seems to be debunked by the Swedish pension reformation. Participants there were similarly allowed to make informed choices about their risk aversion, but practice shows that most members are not particularly engaged with this, as explained in Section 2.1.3

2.1.3 Comparison to other countries

The Netherlands is not the only country where the pension system is a hot topic. Steenbeek and Van Popta (2021) discuss pension reformations in nine different countries in their paper *Transition to a new pension contract in the Netherlands: Lessons from abroad* (3). Some of these countries will be discussed below.

Steenbeek and Van Popta start by analysing the pension reformation of the United Kingdom. The UK had a lot of reforms over the past years, most of which took many years before they were finally fully put into place. One of the most noteworthy reforms happened in 2012, when the UK introduced automatic enrolment for all qualified workers, significantly increasing the number of participants. According to Steenbeek and Van Popta, the key lessons the Netherlands can take from the UK are: Hang on to the strengths of your old system, make sure the bearers of new risk know what to expect, know where you want to go with the reformation, and lastly, that it is hard to turn back once the process has started. This last notion is quite relevant in today's political landscape in the Netherlands, as some parties are proposing to make the reformation undone.

The United States distinguishes between public and private pensions. Their public pensions are still structured as DB contracts, while their private pensions have switched to DC systems, as discussed in Section 2.1.1. The main motivations for the switch in the private sector were cost pressures, longevity risk, and a desire for flexibility. However, this raised some concerns, as employer liabilities were reduced, but employees were exposed to greater risks such as poor investment decisions. The question appeared to be whether the participants were capable of managing their pensions and whether the DC scheme would provide sufficient income after retirement. US reforms apply to new hires only but are implemented directly, where in the Netherlands the reform applies to all members but takes a longer time to have effect.

Denmark and Sweden have both switched from a DB to a DC system. Denmark laid a strong focus on transparency and consensus, with the argument of modernisation being the main reason for the switch. Addressing the necessary adaptation to the changing environment of pensions built participant trust. Denmark has switched to a fully funded DC system, which the Netherlands is trying to achieve in the future. Denmark could achieve this mainly due to the more favourable conditions, especially the broad consensus that seems to be lacking in the Netherlands.
In Sweden, transparency was also an important part of their reformation. They treated their participants as adults and offered individualised estimates of their pension outcomes. This allowed participants to make informed choices about, for example, their risk aversion. In the beginning of the reformation process, Sweden assumed that members would actively engage with their DC pensions and make trade-offs between risk and return. In reality, most of the members turned out not to be particularly engaged. A valuable lesson for the Netherlands is that they should not fully rely on member engagement in implementing DC systems.

Australia makes use of a so-called Superannuation system, where participation is mandatory for employees. The Australian pension system is shifting to a fully funded DC system, thus relying on market performance. This process had already started in 1992. The mandatory contribution consists of 11% of the salary of an employer, gradually increasing to 12% in July 2025. In the Netherlands, contribution is also mandatory, but makes use of collective asset management and risk-sharing. Australia has individual retirement accounts, focussing more on participants managing risks themselves. The clear policies on the contribution rates can enhance sustainability, which could be interesting for the Netherlands to adapt.

These five countries demonstrate lessons that the Netherlands can learn from regarding the reformation towards a DC system. Transparency and consensus turn out to be important, as has been established in both Denmark and Sweden. Building trust by treating participants as adults and allowing for personalised choices both contribute to making the switch easier. Automatic enrolment or mandatory participation in both the UK and Australia shows that this leads to more coverage, increasing the number of participants. These lessons can help navigate the complex transition to a fully funded DC system like in Denmark.

2.2 Market Risk

The concept of market risk is essential to understand for the remainder of this research. Market risk can be defined as the risk of potential losses due to exposure to market fluctuations, recessions, or other factors that can affect the performance of the financial market. These market fluctuations can be caused by changing interest rates, exchange rates, or stock prices. When investing in the financial market, it is therefore useful to know how one can measure market risk and arm themselves against it. Quantifying market risk can be done using methods such as Value at Risk (VaR) and Expected Shortfall (ES). These models will be used in this research and will be explained in more detail in Section 3.1.
Market risk consists of two components: systematic risk and unsystematic risk. Unsystematic risk is the risk that is tied to a specific branch or company, whereas systematic risk is related to the entire market. Therefore, systematic risk is often just called market risk. Although unsystematic risk can be reduced or even eliminated by diversification, an investor should still take it into account.
Diversification is the concept of mixing various investments, such as stocks, bonds, or commodities, into a portfolio. For example, when you have a portfolio that contains only stocks based in the car industry, you create a high level of unsystematic risk. An event like an employee strike could cause stocks to plummet, and the portfolio to lose value. Spreading the assets in a portfolio across different industries and types of assets reduces this unsystematic risk, as a strike in the car industry does not necessarily cause stocks in the, for example, food industry to fall.
Systematic risk can be quantified by the Beta (β) of a stock. The Beta of a stock measures the volatility of a security compared to the volatility of the market. Beta's larger than 1

are interpreted to be more volatile. Adding stocks with higher Beta's, and therefore higher systematic risk will increase the risk on the portfolio, but possibly increase the return. These assumptions are supported by the paper *Portfolio Risk and Return Relationship - An Empirical Study* by Jeyachitra, Selvam and Gayathri (2010) (11). The authors calculated the stock and market returns, the market Beta, the Beta of the stocks, the total risk of a security, and the portfolio returns, to analyse the relation between returns and risk. They find that there is a significant correlation between systematic risk and the expected returns of a portfolio. Furthermore, they indeed find that unsystematic risk decreases as portfolios are more diversified.
Harry Markowitz introduced Modern Portfolio Theory (MPT) in 1952 in his famous paper *Portfolio Selection* (12). A key component of this paper was diversification. According to Markowitz, there are two types of stocks: high risk - high return, and low risk - low return. He argued that one should take into account risk aversion, and based on this create an optimal mix of the two types of assets. Complementary to Markowitz' research, William F. Sharpe quantified the balance between risk and return in his Sharpe Ratio (Equation 2.2).

$$\text{Sharpe Ratio} = \frac{R_p - R_f}{\sigma_p} \tag{2.2}$$

where:

R_p = Return of Portfolio

R_f = risk-free rate

σ_p = standard deviation (volatility) of the portfolio's excess return

A higher Sharpe Ratio usually indicates that the portfolio is better risk-adjusted. However, there are some pitfalls, as volatility can be kept artificially low by portfolio managers. Furthermore, the standard deviation in the denominator is calculated on the assumption of a normally distributed portfolio, which is often not the case.

2.3 Investment Strategies by Pension Funds

As mentioned in Section 1, there are three types of pension funds: Industry-wide, Corporate, and pension funds for independent professionals. In 2019, the assets held by all Dutch pension funds together amounted to 1900 billion euros, more than twice the Dutch GDP (5). The assets of Dutch pensions in terms of GDP are among the highest in the world. Industry-wide pension funds are the largest, the two largest funds hold about 40% of the assets (2). Bikker and de Dreu (2009) investigate the sophistication of investment policies by pension funds in their paper *Pension Fund Sophistication and Investment Policy* (13). They find that asset allocation policies of many pension funds vary widely but appear to be relatively simple. This raises the question whether these different allocation policies can all be optimal. Bikker and de Dreu distinguish between the following asset classes: equities, bonds, real estate, mortgages and loans, commodities, mixed mutual funds, money market instruments, and other investments. Equities and bonds are the largest asset classes in most pension portfolios. Equity instruments are investments that give an investor ownership of a share of a company. The most common equity instrument are stocks. Investing in the stock market is relatively simple and therefore stocks are the foundation of many individual investors' portfolios. Examples of companies that issue stocks are, for example, Apple, Tesla, or Philips.

Another common part of an investment portfolio are bonds. Bonds are considered debt instruments: the investor lends money to a company (corporate bonds) or government (government bonds), and that entity then promises to repay this money with interest at a certain time, the maturity date. Bonds are often used to diversify portfolios. There are many different kinds of bonds. Dutch pension funds usually invest in a few specific types, among others, government bonds, corporate bonds, emerging market bonds, and inflation-linked bonds.

While government bonds are usually stable and relatively safe, they yield low returns. Corporate bonds on the other hand, allow for higher returns and have become present in larger portions of Dutch pension portfolios. According to AXA Investment Managers, the expectation is that government bonds will decrease in proportion to corporate bonds in the future even more. In the UK, pension funds allocate nearly twice as much to corporate bonds, and allocate significantly less to government bonds, compared to the Netherlands (14). As discussed in Section 2.1.3, the UK offers a good guide to what may happen in the Netherlands in the future. Emerging Market Bonds have the potential to grow in the future and thus yield higher returns. However, they are vulnerable to credit and currency risk. Inflation-Linked Bonds help to balance the inflation risk, which is crucial for pension funds, since pension funds have long-term liabilities sensitive to inflation. The combination of low-risk and more volatile bonds can be useful for stress testing, as low-risk bonds remain stable during crises. In the financial crisis of 2008, and the COVID-19 crisis, investors sought low-risk bonds and massively invested in low-risk assets. Since emerging market bonds are more volatile, they expect to fall when a crisis occurs, as investors switch to safer assets.

Commodities represent goods taken from the earth; one can think of gold, oil, or wheat.

Commodity markets allow traders to agree on the future prices of these commodities through forward contracts.
Other investments can be infrastructure, private equity, and real estate. Private equity consists of stocks that are not traded on public markets, but rather offered to specialised investment funds. Infrastructure investments are investments in assets that are sensitive to GDP, such as airports, bridges, and tunnels. Investing in real estate can be done via Real Estate Investment Trusts (REITs). REITs are companies that own income-producing real estate such as office buildings, shops, hotels, and many others.

However, Bikker and de Dreu find that more than half of pension funds only invest in bonds and equities and do not consider the other asset classes. They also find that larger pension funds allocate more shares to equities than bonds, while for smaller funds, this is the other way around. Moreover, they find that most Dutch pension funds select 'attractive' numbers, meaning they allocate mostly based on multiples of 5%. This shows that assets are allocated based on human judgement rather than optimisation models. Small pension funds use attractive numbers more often than large pension funds. Similarly, Bikker and de Dreu also find that smaller pension funds invest less in alternative investments, but mostly in equities and bonds. Lastly, smaller funds appear to be more sensitive to home-bias, meaning they invest more in local stocks, rather than diversifying internationally.

Furthermore, Commodities, Cash and Alternative instruments make up the portfolio. Eleftheriadis and Anastasovitis (2023) investigate whether commodities are effective diversifiers for investment portfolios of pension funds (15). They find that integrating commodity futures into investment portfolios might be beneficial for diversification, since they have a low or negative correlation with traditional options and a positive hedging impact against inflation risk.
The Alternative asset class can consist of many different market instruments. Dixon (2008) concludes in his paper *The Rise of Pension Fund Capitalism in Europe: An Unseen Revolution?* (16), that investing in alternatives such as private equity, real estate, and infrastructure can be an important part of risk diversification. According to Dixon, infrastructure projects offer "stable and predictable long-term cash flows at attractive rates of return and low correlation with other asset classes". Cash is an asset class that usually exists as a small portion of a portfolio. It is mainly used to maintain liquidity.

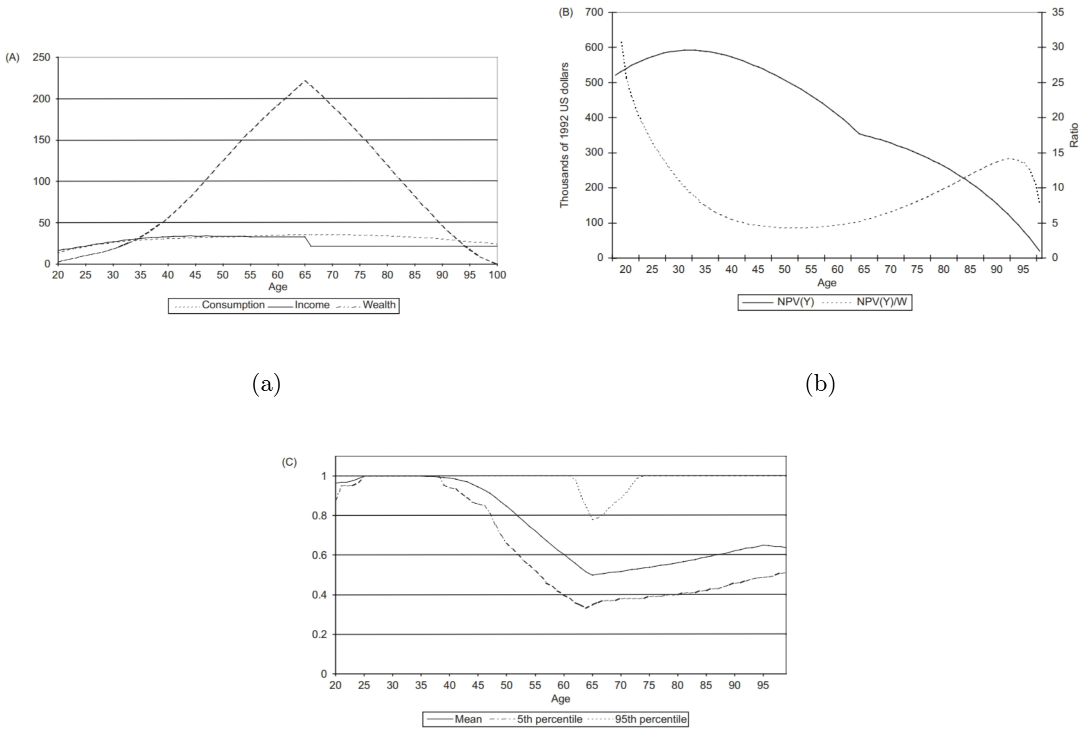
Besides asset allocation, the level of risk aversion is an important parameter in the investment strategy of a pension fund. As established in Section 2.1.3, most participants are not actively engaged with their level of risk aversion. Therefore, pension funds often apply a life cycle investment framework to investigate asset allocation and risk aversion across age cohorts. Cocco, Gomes and Maenhout (1998) augment a life cycle model of consumption to include portfolio choice between a safe asset and a risky asset in their paper *Consumption and Portfolio Choice over the Life-Cycle* (17). Their model is based on a few key assumptions:
- t denotes the adult age with a maximum age of T periods.
- An adult works for the first K years.
- p_t is the probability that the investor is alive on date $t+1$ conditional on being alive on date t.
- They use a time-separable power utility function, given in Equation 2.3 below. Here, $C_{i,t}$ is the level of date t consumption, $\gamma > 0$ is the coefficient of relative risk aversion, and $D_{i,t}$ is the amount of wealth the investor passes on to his descendants at death.

$$E_1 \sum_{t=1}^{T} \delta^{t-1} \left(\prod_{j=0}^{t-2} p_j \right) \left\{ p_{t-1} \frac{C_{i,t}^{1-\gamma}}{1-\gamma} + b(1-p_{t-1}) \frac{D_{i,t}^{1-\gamma}}{1-\gamma} \right\} \tag{2.3}$$

They define labour income $Y_{i,t}$ to be $log(Y_{i,t}) = f(t, Z_{i,t}) + \epsilon_{i,t}$ for $t \leq K$, and during retirement $log(Y_{i,t}) = log(\lambda) + f(K, Z_{i,K}) + \epsilon_{i,K}$ for $t > K$.
Furthermore, they assume a simple financial market with two trading assets: a risky asset and a riskless asset.

In Figure 2.4 on the next page, the results of the model can be seen. In Figure 2.4(c) it can be seen that in the very first years of life, agents invest in the riskless asset, but soon after that they switch to the risky asset and even hit the borrowing constraint. When their age increases, they tend to switch to the riskless asset again. A big reason for this will be saving for retirement. In retirement, wealth runs down quickly, as can be seen in Figure 2.4(a). This causes a slight upward trend of investing in more risky assets, seen in Figure 2.4(c).

The model of Cocco, Gomes, and Maenhout (1998) is a simplified model of real-world applications by pension funds, but it provides valuable insights. Younger participants will likely be given a more risky asset mix, containing more equities, to target a growth in capital, while older participants will be given an asset mix containing more riskless assets, to generate a stable income.



(a) (b)

(c)

Figure 2.4: (a) Simulated Income, Wealth and Consumption (b) Simulated NPV (c) Mean Simulated Allocation

3

Risk Measures and Statistical Models

3.1 Value at Risk & Expected Shortfall

As discussed in Section 2.2, Value at Risk (VaR) and Expected Shortfall (ES) are financial techniques that can measure market risk. In their paper *An Overview of Value at Risk*, Duffie and Pan (1997) (18) formally define Value at Risk as the loss in market value over the time horizon t that is exceeded with probability $1-p$, for a given confidence level $p \in (0,1)$. To present this mathematically, suppose the loss of a portfolio L follows the distribution $F_L(x) = \Pr(L \leq x)$. Given the confidence level $p \in (0,1)$, the VaR of the portfolio at the confidence level p is given by the smallest number l such that the probability that the loss L exceeds l is no larger than $1-p$: $\mathrm{VaR}_p := \inf\{I : \Pr(L > l) \leq 1-p\} = \inf\{l : F_L(l) \geq p\}$. Thus, a 95% VaR means that it is 95% certain that the Value at Risk will not be exceeded. One can also calculate VaR on a longer horizon than 1-day, for example a 10-Day VaR. This is the amount of expected loss over the next 10 days. Over a longer period of time, a portfolio has a higher chance of losing more money, as more market fluctuations can appear. Therefore, the VaR over longer time periods will always be higher than the VaR over one day. Hence, Value at Risk can inform a manager of the probability and quantity of possible losses occurring in the (near) future. In Figure 3.1 on the next page, VaR is visualised schematically.
It was decided to apply confidence levels of 95% and 99% for this research, as these are widely accepted standards and provide a more conservative confidence level (95%), as well as a more strict confidence level (99%).

Value at Risk is a benchmark for relative judgement, one can compare a portfolio to another portfolio. However, VaR by itself is not a solid risk measure. Artzner et al. (1999)

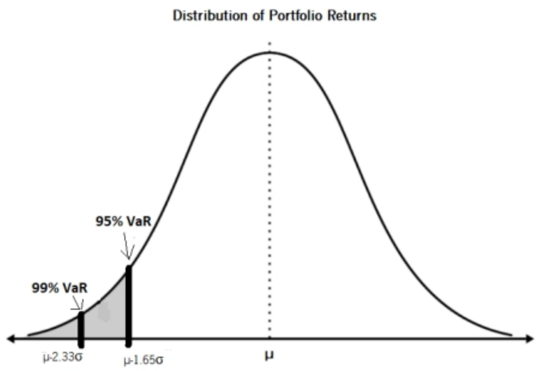


Figure 3.1: Value at Risk Schematic

conclude in their paper *Coherent Measures of Risk* (19) that the VaR does not behave nicely when adding risks (subadditivity), even independent ones. A risk measure is subadditive if $\rho(X_1 + X_2) \leq \rho(X_1) + \rho(X_2)$, however, VaR does not satisfy this property. Subadditivity is consistent with the principle of diversification: a diversified portfolio should generate lower or equal risk compared to a non-diversified portfolio. Value at Risk does not encourage diversification, as it does not take into account the economic consequences of events (19). Lastly, the VaR will never report losses greater than the highest recorded loss in the portfolio. However, losses bigger than this can still be present, only the VaR does not capture it. Including a stress period can eliminate this last disadvantage, as there is a higher frequency of larger losses. Despite these shortcomings, VaR is incorporated by many firms and institutions. The 1996 Risk Amendment of the Basel Accord even sanctioned banks to use VaR as standard approach for determining market risk.

To mitigate the shortcomings of the VaR model, Artzner et al. proposed a model called Expected Shortfall (ES), also known as Conditional VaR (CVaR). Expected Shortfall can be defined as the conditional expectation of the loss given that the loss exceeds the VaR (Equation 3.1).

$$ES(X) = E[X|X \geq VaR(X)] \tag{3.1}$$

Where VaR calculates the maximum loss that will not be exceeded with a certain confidence level, the Expected Shortfall calculates the average of the portfolio losses that do exceed the VaR, to capture large tail risks. For example, when the 1-Day 95%-VaR is €20.000, it is still possible that losses larger than €20.000 appear. The VaR does not capture the possible losses behind this threshold, while the ES does. It could be the case that behind the VaR-threshold, far larger losses appear. The ES calculates the average of the losses that exceed the VaR, to capture these large losses as well.

In 2016, the Basel Committee on Banking Supervision (BCBS) published renewed requirement policies with respect to trading book capital. One of the proposed changes was switching from VaR for Expected Shortfall as a quantitative risk metric. The main reason for this was the number of identified weaknesses in using VaR.

VaR and ES have been used in many research articles to identify exposure to losses of stocks, ETF's or portfolios. Pradhan and Tuwari (2021) (20) investigated the market risk of companies that use clean energy technologies using ES and VaR. They used data from 2001 to 2018 to assess the financial crisis in 2008. They ran a VaR and an ES model and backtested both of them. They concluded that ES forecasts can indeed measure market risk during a period of financial turmoil and during the overall sampling period. ES forecasts can validate the results of the VaR and identify inaccuracies in the characteristics of risk modelling. This research will also perform VaR and ES models on historical data and investigate periods of financial turmoil. It is therefore useful to know that it is indeed possible to obtain validated results.

The VaR of a portfolio can be calculated via a few different methods. One can distinguish between parametric and non-parametric VaR methods. Non-parametric methods have the benefit of not relying on many assumptions, for example data distributions, and are therefore easy to incorporate. The downside is that non-parametric methods are less powerful than a parametric VaR method, as (correct) assumptions on the data can lead to more accurate results. However, the assumptions made for a parametric VaR model are not necessarily correct assumptions. Running the model on incorrect assumptions will likely yield incorrect results. With non-parametric methods, the possibility of making a wrong assumption is eliminated.

3.1.1 Variance-Covariance Method

There are multiple parametric VaR methods that can be considered. A common method is the Variance-Covariance method. This method first looks at the mean and standard deviation of a portfolio. Furthermore, a parametric VaR method assumes that the portfolio returns follow a certain distribution, usually Normal or Student-t. Before performing the

model, it should be checked which distribution the data follows.
In order to perform any VaR method, it is useful to transform the prices of the assets and the portfolio to logarithmic (log) returns. It is hard to compare assets based on prices, as bigger companies tend to have higher stock prices than smaller companies, while this does not say anything about the stability or occurring losses in the asset. Therefore, it is more common to look at the returns. Returns are the difference between the price in the current period and the price in the previous period. For this research, instead of standard returns, logarithmic returns are used. Logarithmic returns have some useful properties. First, they scale linearly, which means that it is easy to compare different time periods of different lengths. Furthermore, they can be added up over time, meaning that you can add the log returns of a number of consecutive periods to get the log return of the whole period. Lastly, logarithmic returns are symmetric on gains and losses. This means that a loss of 5% mirrors a gain of 5%. Logarithmic returns are calculated using $r_t = log(\frac{X_t}{X_{t-1}})$, where X_t is the price of the asset at time t. Since logarithms cannot take negative values, the logarithmic returns of assets with negative values will be calculated using the Python function *.diff()*. This function calculates the difference between the previous return and the current return.

In order to perform the Variance-Covariance method, the covariance matrix of the returns needs to be computed. For this, the variances and covariances between all assets need to be calculated. A covariance matrix consists of all assets and their respective covariances with the other assets. On the diagonal, the covariance of the asset with itself is given: this is the variance of the asset. These covariances are calculated using Equations 3.3 and 3.4 below. The mean of an asset can be calculated using Equation 3.2. The mean of the portfolio can then be calculated by applying the portfolio weights to the means of the different assets.

$$\mu = \bar{X} = \sum_{i=1}^{T} \frac{X_t}{T} \tag{3.2}$$

$$\sigma_X^2 = \sum_{t=1}^{T} \frac{(X_t - \bar{X})^2}{T-1} \tag{3.3}$$

$$Cov_{XY} = \sum_{t=1}^{T} \frac{(X_t - \bar{X}) \cdot (Y_t - \bar{Y})}{T-1} \tag{3.4}$$

Then, the volatility and variance of the portfolio itself need to be calculated. For this, the weights applied to the assets are used and the so-called correlation coefficients ρ_{XY}, are calculated using Equation 3.5. Using Equation 3.6, the standard deviation of the portfolio is then calculated. This standard deviation can then be used to calculate the 1-day VaR using Equation 3.7, using the mean and variance of the portfolio.

$$\rho_{XY} = \frac{Cov_{XY}}{\sigma_X \sigma_Y} \tag{3.5}$$

$$\sigma_P = \sqrt{\sum_{i=1}^{N} w_i^2 \cdot \sigma_i^2 + \sum_{i=1,j=1,i\neq j}^{N} w_i w_j \rho_{ij} \sigma_i \sigma_j} \tag{3.6}$$

$$VaR(1,\alpha) = \mu_P + z_a \sigma_{P,tdist} \tag{3.7}$$

Here, z_a represents the *a-th* quantile of the student-t distribution. When calculating the VaR based on the student-t distribution, attention should be paid to calculating the standard deviation or volatility. The standard deviation of a Student-t random variable with parameters σ_P, the standard deviation of the historical returns, and the degrees of freedom v is given by $\sigma_P = \sigma_{P,tdist} \cdot \sqrt{\frac{v}{v-2}} \Rightarrow \sigma_{P,tdist} = \frac{\sigma_P}{\sqrt{\frac{v}{v-2}}}$. For returns that are student-t distributed, the portfolio variance must be scaled by the degrees of freedom to obtain the σ that must be inserted into Equation 3.7, to calculate the VaR for the student-t distribution. This logically results in the following formula for the variance: $Variance = \frac{\sigma^2_{P,tdist}}{\frac{v}{v-2}} = \sigma^2_{P,tdist} \cdot \frac{v-2}{v}$. Assuming independent and identically distributed (i.i.d.) returns, this variance can be scaled to an N-day variance by multiplying by N. Then taking the square root gives the variance over N days: $\sigma_N = \sqrt{N \cdot \frac{v-2}{v} \cdot \sigma^2_{P,tdist}} = \sigma_{P,tdist} \cdot \sqrt{N \cdot \frac{v-2}{v}}$. Since $VaR_N = VaR_1 \cdot \frac{\sigma_N}{\sigma_{P,tdist}}$, we get $VaR_N = VaR_1 \cdot \sqrt{\frac{N(v-2)}{v}}$, which allows for the calculation of long-term VaR's under the student-t distribution.

The Expected Shortfall calculated using Equation 3.8, specifically for student-t distributed returns. Here v is the degrees of freedom and $\tau^{-1}(\alpha)$ is the quantile at a level α of the student-t distribution.

$$ES_\alpha(X) = \mu + \sigma \cdot \frac{\nu + \tau^{-1}(\alpha)^2}{\nu - 1} \cdot \frac{\tau(\tau^{-1}(\alpha))}{1-\alpha} \tag{3.8}$$

3.1.2 Historical Simulation

The historical VaR method, often referred to as Historical Simulation, is a non-parametric VaR method, which means that it does not assume any distribution of the portfolio returns. Historical Simulation operates on the assumption that historical prices provide a reliable basis for investigating future risk. The premise of the method is that the asset or portfolio is re-valued based on changes in the historical data, as if those changes would happen again in the coming period. The historical method captures true market behaviour, as it uses actual historical returns. The method takes the logarithmic returns of the historical portfolio data over a time frame (the data window) and sorts these returns from the smallest to the largest. The smallest returns are the largest negative returns, representing the largest losses. Based on the confidence level of the VaR, the cut-off point is found by finding the quantile related to the chosen confidence level. The value exactly on the threshold is the VaR. A rolling window approach is used to find the Historical VaR for each time point in the historical data.

Historical Simulation is a method that is easy to implement, and moreover, easy to explain to executives without a mathematical background. The fact that the method does not assume any distribution is a strong suit, as not making assumptions provides more reliable results than making wrong assumptions. The Variance-Covariance method is for example largely dependent on the assumption of the distribution, which can be backed by QQ-plots, but the returns will never fully follow that distribution. Therefore, discrepancies between the actual VaR and the VaR provided by the Variance-Covariance method will appear. Historical Simulation avoids such issues. The Expected Shortfall under this method is calculated by selecting all the sorted returns that are lower than the VaR and taking the average of those. Since the losses are already sorted, no additional formula is needed for this.

3.2 GARCH(1,1)

Generalized Autoregressive Conditional Heteroscedasticity (GARCH) is a statistical model used to predict the fluctuations of stock prices, also known as volatilities. It forecasts volatilities of stock prices based on past data, highlighting potential periods of higher risk. GARCH works under the assumption that the best predictor of variance (or volatility) is the average of long-run variances, the predicted variance over the current period, and new information. Bollerslev introduced the GARCH model already in 1986 (21), an adaptation of the in 1982 published ARCH model by Engle.

GARCH(1,1) is a specification of GARCH and is often used for its simplicity and effectiveness. The parameters (1,1) refer to the number of autoregressive lags and the number of specified moving average lags, respectively. GARCH(1,1) makes use of the so-called Markov property, which means that the future volatility depends only on the most recent volatility and not on any earlier volatilities. Mathematically, this can be written as $P(X_{t+1} = x|X_t, X_{t-1}, ..., X_{t-k}) = P(X_{t+1} = x|X_t)$. The GARCH(1,1) model is shown in Equation 3.9 below and shows that the volatility at time t depends only on the volatility and return at time $t-1$. Because of the Markov property, there is a strong correlation between the successive volatilities. This strong correlation allows the GARCH(1,1) model to find volatility clusters, which can be useful in risk analysis. Equation 3.9 contains the estimated parameters α, β, and ω. ω is the long-run average variance and is therefore constant. α captures the immediate effect of shocks in the returns. When α is high, this means that volatility reacts strongly to changes in returns. β accounts for the autoregressive behaviour or the persistence of volatility. Therefore, a high β suggests that once volatility increases, it tends to stay high for a longer period of time. This contributes to the model's ability to capture volatility clusters. All parameters (α, β, and ω) must be larger than or equal to 0 to ensure that there are no negative variances. Moreover, their sum needs to be strictly smaller than 1 to ensure stationarity. When the sum of α and β is less than 1, the effect of a shock decays over time. If $\alpha+\beta$ is close to 1, this implies high persistence, which is often observed in financial returns. When their sum is greater than 1, the shocks will have a compounding effect on volatility. This can cause volatility to grow unbounded and become an explosive process, suggesting permanent volatility shocks in the model. This is undesirable and unrealistic for the financial models used in this research.

$$\sigma_t^2 = \omega + \alpha_1 \cdot X_{t-1}^2 + \beta_1 \cdot \sigma_{t-1}^2, \omega > 0, \alpha_1 \geq 0, \beta_1 \geq 0 \tag{3.9}$$

where:

σ_t : the conditional volatility (variance) at time t

X_t : the return at time t

$\omega, \alpha_1, \beta_1$: coefficients that need to be estimated.

$\alpha_1 + \beta_1 < 1$

Hansen and Lunde (2005) investigated whether there is a model that beats GARCH(1,1) in predicting conditional volatility and concluded that there is no evidence against the claim

"Nothing beats GARCH(1,1)" (22). Therefore, it was decided not to use a more complicated GARCH model, as complicated models are thus not necessarily more accurate than GARCH(1,1).

GARCH(1,1) has been used in a lot of research regarding volatility forecasting, but it can also be used as a risk predictor. Laplante, Desrochers and Préfontaine (2008) (23) investigate whether the GARCH model can get more accurate estimates of portfolio risk than Random Walk (RW), Exponentially Weighted Moving Average (EWMA), and Historical Mean (HMM) in an international investing context. They do this by first estimating the covariance matrix of all these methods and then determining the proportions to invest in the portfolio. Lastly, they calculate the standard deviation of the Minimum Variance Portfolio (MVP) and compare the estimated and actual risk using the root mean squared error (RMSE). They find that the GARCH(1,1) model is good at estimating risk in the MVP, and better than HMM. Random Walk also performs well, but GARCH(1,1) seems to obtain a better covariance matrix. Hence, GARCH(1,1) enables pension funds to optimise their portfolios and navigate long-term market risk.

The VaR and ES methods described in Section 3.1 use the mean, variance, and volatility based on a rolling window approach. The time-varying nature of volatility is captured simply by recalculating the portfolio volatility for each time window. GARCH(1,1) allows us to model the time-varying volatilities by calculating them based on previous volatilities and returns. This allows for a more sophisticated approach to calculating the time-varying volatilities. Based on these estimates, the VaR and ES models can be performed again using Equation 3.7, with a rolling mean and the volatilities produced by the GARCH(1,1) model.
GARCH(1,1) can be performed directly on portfolio returns, or it can be performed on each asset separately; this is called Constant Conditional Correlation (CCC) GARCH. Both methods will be discussed in this research.

3.2.1 Univariate GARCH(1,1)

The Univariate GARCH(1,1) method applies a GARCH(1,1) model to a single asset, in this case the logarithmic returns of the portfolios, as these are a weighted combination of all assets in the portfolio. The Univariate model is quite straightforward, the volatilities of the portfolio are calculated based on Equation 3.9. GARCH(1,1) typically assumes normally distributed returns, but this can be changed when the data turns out to be student-t distributed, or follows any other distribution. The model outputs the volatility of the logarithmic returns for each time point. These volatilities can easily be inserted into Equation 3.7 to calculate VaR of the portfolio at each time point.

3.2.2 CCC-GARCH(1,1)

The CCC-GARCH(1,1) model is a multivariate model and can therefore be applied to a list of assets. A multivariate GARCH(1,1) model essentially means that multiple univariate GARCH(1,1) processes are running next to each other, each for an individual asset. The CCC stands for Constant Conditional Correlation, and implies that the model makes the assumption that the correlation between the assets is constant over time. Furthermore, the model outputs the volatilities of each asset per time point. Based on these time-varying volatilities and the constant correlation matrix, the covariance matrix of the assets can be calculated for each time step, using Equation 3.10 below. Here, D_t is a diagonal matrix with the volatilities of the assets of that time point on the diagonal, and 0's elsewhere. Furthermore, $\rho_{i,j}$ represents the elements of the constant correlation matrix.

$$C_{t,i,j} = D_t \cdot \rho_{i,j} \cdot D_t \tag{3.10}$$

These time-varying covariance matrices can then be used to extract the needed variances by applying the weights and hence the volatilities are obtained to calculate the VaR and the ES for each time point. However, this assumption of constant correlations can be unrealistic, especially in changing market conditions. Since the Univariate GARCH model does not make this assumption and is generally easier to implement and interpret, it was decided not to use the CCC GARCH(1,1) model, but only the Univariate GARCH(1,1) model.

3.3 Monte Carlo Simulation

The last tool that will be implemented in this research is Monte Carlo Simulation. Monte Carlo Simulation is a powerful tool that can be used in various contexts. In this case, it will be used to deal with uncertainty in the behaviour of the stock market, as it allows for the generation of multiple potential future portfolio paths based on historical data. In order to perform a Monte Carlo Simulation on stock prices, a model that captures the drift and randomness in returns needs to be assumed. A commonly used model is the Geometric Brownian Motion (GBM), as is the case in the paper *Monte Carlo Simulation Prediction of Stock Prices* by Xiang, Velu and Zygiaris (2021) (24). The GBM is a stochastic process that follows the Stochastic Differential Equation (SDE) given in Equation 3.11 below, where dS_t represents the change in the stock price at time t, calculated from the current stock price S_t, a drift term (μdt), and a volatility or shock term (σdW_t). W_t represents a Wiener Process, also known as a standard Brownian Motion, and thus a stochastic process, similar to the GBM. A Wiener process can be used as a stochastic process itself, but it can also be used as a building block for other processes. A Wiener Process has the following properties:

(i) $W_0 = 0$

(ii) the increment $W_t - W_s \sim N(0, t-s)$ for any $0 \leq s < t$

(iii) the increment $W_t - W_s$ is independent of $(W_u : u \leq s)$, for any $0 \leq s < t$

(iv) any sample path is a continuous function

In financial modelling, a Wiener process represents the randomness in price movements of assets. The sample path of a Wiener process is continuous almost everywhere. It can be seen as the continuous-time limit of a Random Walk with small increments. Then, in the limit the Random Walk converges to a Wiener process that is continuous, nowhere differentiable, and has independent normally distributed increments.

When simulating the GBM, the time is discretised into small intervals dt. During each time interval, the change in the Wiener process dW_t can be modelled as $\sqrt{dt} \cdot Z$, where Z is a standard normal random variable. Because a Wiener Process can take negative values, using it directly to model stock prices could result in negative prices, which are not realistic. Therefore, a Wiener Process by itself is not a good model for asset pricing, as stock prices are generally positive. However, the Wiener Process can be used to build a GBM that can describe stock prices in a more accurate way using Equation 3.11. The drift is the term that describes the general direction of the stock price, while the volatility represents the risk or uncertainty in the stochastic process. Equation 3.12 follows directly from Equation 3.11 using Ito's Lemma, which will not be explained here in detail as it is out of the scope of this research. Equation 3.12 describes the movement of the stock price. As the process is now exponential (e^x), since the solution to the GBM SDE involves the exponential of a normally distributed term, prices are ensured to be strictly positive. The GBM can be used to describe stock prices, since it only produces positive values, the relative changes always follow the same distribution, and the return does not depend on the current level of the stock price.

$$dS_t = \mu S_t dt + \sigma S_t dW_t \tag{3.11}$$

$$S_t = S_0 e^{(\mu - \frac{1}{2}\sigma^2)t + \sigma W_t} \tag{3.12}$$

Typically, 1000 to 10000 runs are done, as more runs can be very time-consuming, but will provide more thorough results. Xiang, Velu and Zygiaris performed 1000 runs due to the high time consumption. They run a Monte Carlo Simulation on both the Malaysian and the US stock market, to investigate which market would be better to invest in for retail traders. The simulations show that most of the predictions in the Malaysian stock ended up closing lower than the opening price, with a few exceptions, while the predictions in the US stock mostly ended up higher than the closing price. This suggests that it would be better to invest in the US stock market for retail traders. The research concludes that it thus is possible to model stock prices using Monte Carlo Simulation with an GBM model, as the majority of their simulations per market show similar (upward or downward) trends.

The benefit of using a Monte Carlo Simulation in the context of this research is that many simulations for different pension portfolios (long-term and short-term) can be run. As there is a lot of uncertainty in the market, predicting a single stock price in the future will be completely insignificant, but predicting a 1000 or 10000 future stock returns can allow for estimating the distribution of the portfolio returns, which is essential for investigating risks and returns of different investment strategies over time. Many papers conclude that running a Monte Carlo Simulation with a GBM to model the stock prices returns feasible solutions. Sonono and Mashele (2015) compare the GBM model to a Variance Gamma (VG) model in predicting stock prices using Monte Carlo Simulation in their paper *Prediction of Stock Price Movement Using Continuous Time Models* (25). Their empirical tests suggest that both models in any Monte Carlo method can be used to predict directions of stock movement. Brodd and Djerf (2018) model future stock returns via Monte Carlo Simulation in their paper *Monte Carlo Simulations of Stock Prices* (26). Although they mention that their methods tend to lack accuracy (the highest obtained accuracy for predicting trends was 90%, the lowest 50%), they suggest that altering weights could yield more accurate results than the study suggests. Overall, research shows that stock price prediction using Monte Carlo Simulation with a GBM model is feasible and effective.

4

Methodology

This chapter will provide a detailed outline of the application of the proposed models to the research question. Based on the literature reviewed in Section 2, and the risk measures and statistical models discussed in Section 3, the methodology for this research will be explained. First, data analysis will give insight into the data that was used, the allocation of assets, and the handling of missing values. Then, the application of the risk measures and statistical models on the historical data as well as the simulation of future data will be explained. In addition, a practical scenario analysis will be proposed, taking into account salary growth and the variable premium.

4.1 Data Analysis

In order to analyse the risk exposure of different generations caused by the new pension system, a dataset needs to be gathered. This dataset will mimic the portfolios of individuals in the NPC, managed by pension funds. To mimic individual portfolios in the NPC, the types of assets and their allocation need to be studied.

4.1.1 Historical Data

The first models will be run on portfolios constructed based on historical data. The historical data will range from 2011-01-01 to 2023-01-01, to incorporate the market shock of the corona crisis. Gathering data from before 2011 is very hard, as many of the market instruments that were preferred do not have data prior to this date. Data is collected from different sources, including Yahoo Finance, ECB Data Portal, and Investing.com. When downloading data from these sites columns with Price, Open, High, Low, Change %, Close, and Adjusted Close are obtained. Depending on the data source, some of these columns may not be present. The column with which this research concerns itself is the Adjusted Closed column. The Close column represents the raw closing price at the end of the day. One can think of this as the price of the last bought shares before the market closes. The Adjusted Close price incorporates factors that might affect the stock price after the market closes. It is often considered a more accurate representation of the stock value.

The historical data will contain a number of different asset classes, namely: Equities, Bonds, Commodities, Cash and Alternatives. In Section 2.3, the reasoning behind these specific classes is explained.
Dutch pension funds do not only invest in Dutch stocks, but also invest internationally to balance domestic exposure with global diversification. Therefore, the equities class will consist of stocks from both the Dutch and the international market, mainly stocks from the AEX and S&P 500, as these contain large market-representing stocks from different sectors, allowing for diversification. The stocks used in this research are ASML, Coca-Cola, ING, Philips, Apple, Shell, and NVIDIA. Investing in international markets as well as Dutch markets raises the issue of exchange rate risk, since the assets are not all traded in the same currency (euros, dollars). However, accounting for exchange rate risk brings about a new level of uncertainty, namely predicting future exchange rates. It is likely that predicting these exchange rates today will not be an accurate representation of exchange rates 40 years from now. Hence, the accuracy of the models will likely decrease. Since dollars and euros are relatively stable and lie close to each other (1:1.14), it was decided not to incorporate exchange rate risk in the models.
The bonds will consist of the 10Y Dutch Government Bonds, Core Euro Corp Bond UCITS ETF, J.P. Morgan Emerging Markets Bond ETF, and Vanguard Eurozone Inflation-Linked Bond Index. These bonds are chosen to represent the four main categories of bonds: Government, Corporate, Emerging Market, and Inflation-Linked. 10Y Dutch Government Bonds are issued by the Dutch government and have a maturity of 10 years. In general, they are quite safe bonds, due to the credit risk of the Netherlands being AAA. The Core Euro Corp Bond ETF tracks the performance of an index compased of euro-denominated investment-grade corporate bonds. Pension funds are known to hold investment-grade corporate bonds in their portfolios, and since this index is euro-denominated, it is representable for Dutch pension funds. Emerging Market bonds are securities issued by corporations within developing countries. Lastly, inflation-linked bonds are often used by pension funds to help mitigate inflation risk in long-run portfolios. For Commodities, the

S&P GSCI Index is chosen. This index is used to track commodity performance. It is designed to represent a broad range of commodities from different sectors, such as energy, agriculture, and gold. It is widely used as a reference for commodity investments.
As a cash instrument, the Euribor 3-Month is chosen. Euribor-3M is a Euro Interbank Offered Rate for 3-month deposits. European banks lend each other money at this interest rate level. Dutch Pension funds typically hold some low-risk cash reserve, and EURIBOR-3M is a good benchmark for this.
The last asset class is Alternatives, which contains every instrument that does not fall under the asset classes mentioned above. It contains the following assets: Private Equity, Real-Estate, Hedge Funds, and Infrastructure. These assets are represented by ETFs as it was hard to find daily data that went back far enough to be useful. In Table 4.1, an overview of the assets used can be found. The allocation of these asset classes will be discussed in Section 4.1.2.

Category	**Assets**
Stocks	AEX Stocks S&P 500 Stocks
Bonds	Dutch 10-year Government Bonds iShares Core Euro Corp Bond UCITS ETF EUR (Dist) (IEAC.L) iShares J.P. Morgan USD Emerging Markets Bond ETF (EMB) Eurozone Inflation-linked Bond Index
Cash	EURIBOR 3-month Rate
Commodities	S&P GSCI Index
Alternatives	Private Equity Index (PSP) Vanguard Real Estate Index Fund ETF Shares (VNQ) iShares Global Infrastructure ETF (IGF) Hedge Fund Proxy Index ETF (IXG)

Table 4.1: Historical Data: Asset Classes

4.1.2 Asset Allocation

Pension funds often publish their asset allocations in their annual report, although some go into more detail than others. The yearly report of Pensioenfonds Zorg & Welzijn (PFZW) was retrieved to look at their asset allocations (27). Combined with an asset allocation based on KPMG's research regarding investments of pension funds, a benchmark portfolio will be constructed. The insights of KPMG are based on experience in the market and a

generalisation of customer data from multiple clients. None of the mentioned percentages are directly connected to a certain client. From Table 8.1 and Table 8.2 in the Appendix, it can be seen that government bonds are a large part of the portfolios used by actual pension funds. Since the expectation is that this will decrease (14), less weight will be placed on government bonds in this research, and a little more on corporate bonds. Emerging market bonds and inflation-linked bonds both account for a smaller percentage. These smaller percentages allow for managing the risk, while maintaining the balance of inflation risk.
The largest portion of pension portfolios is often occupied by listed equities. It was chosen to allocate 25% to equity in the benchmark portfolio, as these are usually a large portion of the portfolio.
Cash is often a small part of a pension fund portfolio. KPMG's research does not even take cash into account. In the case of PFZW, cash is borrowed. In this research, the cash will not be borrowed, as it is desired to avoid taking on debt. A small percentage will be allocated to cash. This will allow maintaining a level of liquidity pension funds often desire, as they have short-term obligations such as pension payouts.
Commodities are considered relatively risky, as they are affected by supply, demand, and geopolitical events. For the benchmark portfolio of this research, 5% will be allocated to commodities.
Private Equity will contain 7.5% of the portfolio in this research, as private equity can provide higher returns than other equity indexes such as AEX and S&P 500. It also allows for diversification into private markets. Infrastructure will contain 5%, hedge funds will account for 2.5%, and real estate for 5% as well.

Based on these asset allocations, a benchmark portfolio was created that mimics a portfolio of a real pension fund. This portfolio will be used to establish a baseline of a risk profile of a Dutch pension portfolio, using Value at Risk and Expected Shortfall. These metrics will be performed on multiple time horizons, such as 1-Day VaR and 1-Week VaR to investigate long-term risk.

Then, it will be investigated what happens when the weights are altered. A higher proportion of high-risk assets can represent the portfolio of a younger person who prefers an offensive strategy, for they can still compensate potential losses later on. Similarly, a high proportion of low-risk stocks, such as government bonds, can represent the portfolio of someone of the older generation, more risk averse as they approach their retirement. These altered weights are likely to alter the VaR and show how the long- and short-term risk changes. Therefore, in addition to the benchmark portfolio, two additional portfolios were built, one with risky assets to represent a younger participant, and one with less risky assets to represent a member of an older generation. The specific allocations can be found in Table 4.2 below. As Equities and Commodities are considered high-risk instruments, a higher weight was allocated to these classes for the younger generations. The Corporate bonds and Inflation-Linked bonds are relatively stable and low risk, hence a higher percentage was allocated to these assets for the older generations. The benchmark portfolio represents a diversified approach without excessive allocations in a certain asset. In this way, three different levels of riskiness are represented. The VaR and ES of these portfolios will be compared.

Asset Class	**Benchmark Allocation**	**Younger Generation Allocation**	**Older Generation Allocation**
Equities	0.3	0.45	0.10
Corp Bonds	0.10	0.025	0.225
EM Bonds	0.20	0.025	0.20
Commodities	0.05	0.10	0.05
Infrastructure	0.05	0.025	0.10
Hedge	0.025	0.05	0.025
Private Equity	0.075	0.075	0.075
Real Estate	0.05	0.075	0.05
Gov Bonds	0.045	0.05	0.025
Inf Linked Bonds	0.10	0.10	0.10
Cash	0.005	0.025	0.05

Table 4.2: Asset Allocations for Different Generations

Afterwards, Monte Carlo simulation will be used to simulate long-run performances of a large batch of pension portfolios. The results of these simulations can demonstrate the likelihood of potential returns for short and long time-horizons. The output of the Monte Carlo simulation will be the portfolio paths, as discussed in Section 3.3.

4.1.3 Handling Negative and Missing Values

Once the dataset was downloaded, it was checked whether it contained missing values, negative values, or distinctive outliers. Missing values, often referred to as NaN's (Not a Number), can appear for a number of reasons. Bond markets often have different holidays than stock markets, causing bonds to have missing values on their holidays, while stock markets are still open. Another reason may be that some bonds do not trade daily and can have gaps in their data. Lastly, there may be differences in the data from different continents. America and Europe have different holidays, but also operate in different time

zones, possibly causing mismatches. The number of missing values in the dataset is visible in Table 4.3 on the next page. As the number of NaN's is limited, it is decided to impute them with numeric values.
These missing values can be imputed via a few different methods. The most common methods are forward filling and backward filling. The concept of these methods is quite straightforward: forward filling fills the NaN's with the value from the previous date, while backward filling fills the NaN with the value of the next date. Linear interpolation is another approach to remove NaN's. Linear interpolation draws a straight line between the last known point before the NaN and the first known point after the NaN and assumes that the missing value is on this line in between the two known rates. Since the slope of the line is the same everywhere, calculating the slope allows one to find the missing value. The problem with applying this method to stock prices is that it can reduce the impact of a sudden shock. When a shock occurs, the values of the stock price suddenly drop or rise, and interpolating a missing value will decrease the size of such a shock. It was briefly checked whether applying interpolation instead of filling yielded significantly different results, but due to the low number of missing values, there was no visible impact. Therefore, it was decided to use filling for the whole research. Since some missing values appeared in the first row of the dataset, the method of backward filling was chosen because forward filling does not work on the first row because there is no previous value to fill the NaN.

Although stock values generally do not turn negative, some bond values and cash instruments can. Negative values can cause problems when calculating log-returns, so it is important to check whether they are present or not. Since both Euribor-3M and 10Y Dutch Government Bonds have a significant number of negative values (Table 4.3), it does not make sense to remove all of these, as it would make the dataset unusable. Moreover, the negative values represent certain market states. Therefore, it was decided to leave them in the dataset, but to use a different way of calculating the logarithmic returns for these specific assets, as was explained in Section 3.1.1.

Asset	Negative Values	NaN's	Asset	Negative Values	NaN's
AAPL	0	0	10Y Dutch Gov Bond	641	6
ASML	0	0	Corporate Bond	0	56
CCEP	0	0	Emerging Market Bond	0	0
ING	0	0	Inflation Linked Bond	0	20
NVDA	0	0	Infrastructure	0	0
PHG	0	0	Hedge	0	0
SHEL	0	0	Private Equity	0	0
Commodities	0	0	Real Estate	0	0
Euribor 3M	1807	25			

Table 4.3: Negative Values and NaN's of Assets

4.2 Implementation of Risk Measures and Statistical Models

4.2.1 Historical Analysis

In order to compare the generational differences of an investment portfolio, the three different portfolios, benchmark, risky, and conservative, as explained in Section 4.1.2, are compared. First, the only difference between the portfolios is the allocation and thus the risk level. Later, other aspects, such as the time horizon, will be altered.

As explained in Section 3.1, the Value at Risk is calculated using the three different methods discussed in Section 3: Variance-Covariance, Historical Simulation, and Univariate GARCH(1,1). For each of these methods, a 1-Day VaR and a 1-Week VaR are calculated.

In order to perform the Variance-Covariance method, first, the distribution of the portfolios needs to be determined. For this, QQ-plots of the logarithmic returns are created. Figure 4.1 on page 39 shows the QQ-plots of the benchmark portfolio. It can be seen that the portfolio returns are not normally distributed, as indicated by large deviation of the fitted points from the red baseline. To double check, a Shapiro-Wilk test was used to test for normality. Razali and Yap (2011) (28) compare the power of the Shapiro-Wilk test, the Kolmogorov-Smirnov test, the Anderson-Darling test, and the Lilliefors test. The power of a normality test refers to the probability that the null hypothesis is correctly rejected. They find that the Shapiro-Wilk test is the most powerful test, followed by the Kolmogorov-Smirnov test. However, for a low sample size, the power of all tests is low. On the basis of this research, it was decided to use the Shapiro-Wilk test.

The null hypothesis H_0 of the test is *the data is normally distributed*. The Shapiro-Wilk test statistic W is defined in Equation 4.1 (28). Here, $y_{(i)}$ is the i-th order statistic, $\bar{y}$ is the sample mean, and a_i are constants computed from the expected values of the order statistics of a standard normal distribution. In Table 4.4, it can be seen that all p-values are less than 0.05, and thus the null hypothesis is rejected and it is likely that the data is not distributed normally.

$$W = \frac{\left(\sum_{i=1}^{n} a_i y_{(i)}\right)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \tag{4.1}$$

Portfolio	Shapiro-Wilk p-value
Benchmark	1.16×10^{-41}
Risky	1.49×10^{-36}
Conservative	8.18×10^{-44}

Table 4.4: Shapiro-Wilk p-values

Hence, running a Variance-Covariance method with a normality assumption will likely result in infeasible results. Therefore, it was decided to try a Student-t distribution, with different degrees of freedom, since Student-t distributions are known to capture tail risk, or fat tails, better compared to a normal distribution. The Student-t distribution with degrees of freedom 4 appears to be the best fit, according to the QQ-plots in Figure 8.2 in the Appendix. In Figure 4.1, it can be seen that even the degrees of freedom of 4 are not optimal. The same results appear when applying QQ-plots to the risky and conservative portfolios. These can be found in the Appendix as well. Despite the fit not being optimal, it is still decided to run the Variance-Covariance method on the t-distribution.
The Variance-Covariance method also makes use of the covariance matrix. This matrix consists of the 17 assets used and is therefore a matrix of 17×17, with the variances on the diagonal. Using this matrix, the VaR can be calculated, as explained in Section 3.1.1.

To assess the 1-Day VaR and ES throughout the portfolio time period, a rolling window approach is used. The VaR and ES are then recalculated at each point in time using a fixed-size window. A window size of 250 days was chosen, as this represents one year in trading days. For each day t, the VaR and ES are calculated on the returns in $[t-250, t-1]$, $t > 250$. Hence, for a window size of 250, the first VaR and ES are calculated on day 251, based on values from day 1 ($t - 250 = 251 - 250$) to 250 ($t - 1 = 251 - 1$). Then, the window moves one day to the right, and the VaR and ES for day 252 are calculated from

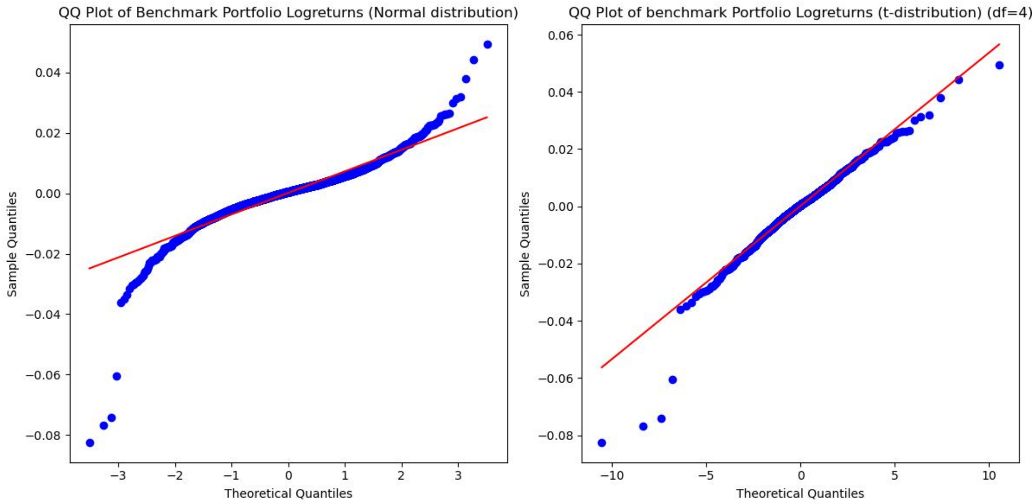


Figure 4.1: QQ-Plots of Benchmark Portfolio Returns

day 2 to day 251, and so on.
Inside the rolling window, the VaR is calculated by calculating the means and covariances of the returns for each window. Then, the covariance matrix is calculated using Equation 3.4. Applying the portfolio weights then provides the variances for each window. These variances can then be used to obtain the volatility for each time window. Inserting these into Equation 3.7 yields the VaR for each window.
Hence, for daily data, the VaR can be computed for every day from day 251 onwards. This allows for the investigation of periods of higher risk. The same methodology is possible for the Expected Shortfall. There, Equation 3.8 uses the volatilities per window to calculate the rolling ES.

In Section 3.1.1, the $\sqrt{N}$ rule was discussed to assess the 1-Week VaR and compare it to the 1-Day VaR. However, the $\sqrt{N}$ rule multiplies every VaR by exactly the same factor, so the VaR will increase by the same amount at every time point. Therefore, it was chosen to calculate the N-Day VaR in another way, namely using a rolling window. This approach is very straightforward: An N-day rolling window is used, depending on what VaR needs to be calculated (1-Week, 1-Month, etc.). This rolling window is then looped over the historical logarithmic returns, and then the VaR is calculated exactly the same way as for the 1-Day VaR.

Where both the Variance-Covariance method and the Historical Simulation method find the volatility of the portfolio for each time point using a rolling window, GARCH(1,1) can be used to obtain more accurate volatility values. The Univariate GARCH(1,1) method, discussed in Section 3.2, was implemented. In order to perform a GARCH(1,1) analysis, the logarithmic returns are scaled by multiplying by a factor 100, as the GARCH model works best on values between 1 and 100. The logarithmic returns are clearly not in this range and were therefore scaled. Furthermore, GARCH(1,1) assumes that the data is normally distributed, but since Figure 4.1 showed that a Student-t distribution is more plausible, this was implemented. In addition, the weekly VaR is also being calculated. Since the GARCH volatilities are daily, these need to be converted to weekly volatilities. This was performed via a rolling window with size 5, as there are 5 trading days in a week.

All the methods mentioned above are performed on the benchmark, risky, and conservative portfolio in order to compare the VaR and ES values and determine whether riskier investments indeed yield higher VaR's.

4.2.2 Simulations

After the analysis of historical data, the focus shifts towards simulating future price paths of the portfolio, to analyse potential risks in the future. As mentioned in Section 3.3, Monte Carlo simulation will be used for this. The Geometric Brownian Motion used to simulate the stock prices is based on the same historical data mentioned in Section 4.1.1. In order to compare younger and older generations via Monte Carlo Simulation, the asset allocation weights and the time horizon are chosen as variable input factors. It is decided to start with the three asset allocations, benchmark, risky, and conservative, identical to those in Section 4.1.1. For the time horizon, it is chosen to take 10, 15, 20, 25, 30, 35, and 40 years. These values are chosen based on how long participants are expected to be employed in the new system. It is assumed that the younger generations will spend around 40 years in the new system, from graduating university to retirement. Ten years was chosen as a starting point, as participants spending less time will likely experience less impact of the reformation, as they are already close to retirement.

It was decided to take relatively small intervals of five years to be able to create a curve of the VaR with respect to these time horizons. In that way, it can be investigated whether there is a turnover point when the VaR increases significantly, or whether this goes in a graduate fashion. A similar approach is taken for the Expected Shortfall.

A starting value of €1000 was taken for each simulation. In Section 4.3, actual starting

salaries, experience, and indexation are taken into account. In this first part, it will be assumed that everyone starts with the same value and that there is no salary increase over time. This method will be used particularly to assess risk levels over certain horizons, while Section 4.3 is concerned with calculating actual pension payouts per generation.
Lastly, it is decided to perform 5000 runs, as 1000 to 10000 runs are typically done, although there is no clear argument as to why this is the case. As 10000 runs can be very time-consuming, it is chosen to take 5000 simulations in this research.

Historical data can be used in multiple ways to build a GBM for future stock prices. The first approach uses a static mean and volatility. Based on the historical data, the mean and volatility of each asset are calculated. These means and volatilities are kept constant for all simulated time points. This is clearly not a realistic model, as volatilities fluctuate over time, but it provides a solid baseline for investigation. Combined with the GBM, these means and volatilities are used to simulate 5000 potential stock paths per asset. Eventually, the portfolio weights are applied to these simulated stock paths, and 5000 potential portfolio stock paths are constructed, over multiple time horizons. The final values of these paths are particularly interesting as they represent the possible values of the investment after a certain period of time. For the final values of these different time horizons, the mean of the simulation is calculated, as well as the 95% and 99% VaR. These final values will be plotted into histograms to assess the distribution of the outcomes. Lastly, the VaR will be calculated at the end of each time horizon to investigate how much it will increase over time. This approach will be taken for all simulation methods.

Complementary to the aforementioned approach, the Monte Carlo Simulation was run using GARCH volatilities as well. Identical to Section 4.2.1, the GARCH model is used to calculate the volatility of each data point of the historical data. Then, instead of taking a constant volatility for all future data points, the GARCH volatilities will be used. Since the historical data contains around 2000 data points, and the simulation will continue up to 10000 future data points, the GARCH volatilities are being repeated in a cycle. However, since the GARCH volatilities of the historical data range from 2011-2023, the Covid-19 spike is present in the data, which can be considered as an extreme event. Because the model loops through the volatilities in a cycle, this extreme event is repeated every 2000 data points, or roughly every 10 years. Since such frequent extreme events might not be very realistic, this model was discarded.
As an alternative, the GARCH volatilities over a relatively quiet period in the historical

data were taken to loop through. In that way, the volatilities are not static, but do not prompt multiple unrealistic volatility spikes either. Hence, this model was implemented in the research.
Another approach that was considered, instead of looping through the GARCH volatilities, is random sampling with replacement. In this way, the Covid peak can incidentally be incorporated in future price paths. The method samples a random value for each time point from the array containing all the GARCH volatilities over the historical data. However, a key concept of the GARCH(1,1) model is the strong correlation between volatilities due to the Markov property, as explained in Section 3.2. Random sampling of the historical volatilities eliminates this property, making the model significantly less strong. Therefore, it was decided not to implement this approach.

4.2.3 Stress Testing

Since both implemented volatility models discussed in Section 4.2.2 do not incorporate any volatility spikes, it was decided to manually perform a stress testing analysis. A spike is created by significantly increasing the volatility by a certain factor for a short period of time (around 100 days). Spikes are incorporated in the model with the GARCH volatilities and not in the constant-volatility model, as volatility is constant there anyway. The spikes will be introduced at four different time points: after about 8 years, 16 years, 24 years, and 32 years. These were chosen as stepwise increments between 0 and 40 years. Clearly, peaks that occur later do not have an effect on portfolios that are only simulated for shorter horizons. Therefore, younger generations will experience a higher risk of the occurrence of stress events. Moreover, it will be interesting to see to what degree the VaR increases when a stress event occurs and whether this increase differs when introducing the spikes at different time points. Therefore, the difference between the regular VaR and the VaR with spikes is calculated for every time horizon. In this way, it can be investigated whether the impact of a spike differs depending on when it occurs.

4.3 Practical Scenario Analysis

In addition to the previous method, this section will present a simulation to model practical scenarios for different generations. In this model, salaries, experience, and indexation will be dynamic variables. Different career paths will be taken into account so that this model can predict possible monthly pension payouts for an individual participant after retirement. This model will combine the old and the new pension systems together, so that the effect of switching midway through one's career becomes clear as well. The modelling of the new system is largely based on the model in Section 4.2.2, with a few added assumptions.

4.3.1 Assumptions

First, it was decided to make the assumption that everyone starts with the same starting salary. Of course, this is not fully realistic, but starting salaries in most branches will lie relatively close to each other, as no individual will start at the top of the career ladder. Therefore, it was decided to assume that all individuals start on the same base salary but experience different growth rates. In certain branches, progress can be made regarding salaries faster than in other branches. Similarly, there are career paths that generate little to no salary increase over time. It was decided to take three levels of growth rates to work with: low, mid, and high, chosen to be 1%, 2.5% and 4%. These growth rates then evolve over time, following a curve. Since there is no available data of an average salary growth curve over the lifecycle, a curve was created from scratch. The new salary will be calculated using $current\_salary = previous\_salary \cdot (1 + growth\_rate)$. Depending on whether the low, mid, or high growth rate is inputted, the growth rates in the other life phases will then be calculated. This growth rate will be an input variable that can be changed for different industries. Different salary growths will be used to resemble a fast-growing salary, a slowly increasing salary, and one in between. The assumption was made that the salary increases with a certain growth rate for the first ten years. Then, after ten years, this growth rate will increase by 50%, so if the growth rate is 0.02, it will become 0.03. The assumption is made that after ten years one starts to build their career and will likely increase in function, and thus in salary. Lastly, after 25 years, the starting growth rate (0.02) will decrease by 50%, becoming 0.01, as in the last years before pension, it is less likely that one's salary will increase significantly. In this way, when the growth rate is applied to the premium, the premium follows an S-curve over time.

Furthermore, as explained in Section 2.1.1, the new pension system uses a flat premium, compared to the variable premium in the old system. Since this new premium is flat, it is relatively easy to model. As a value for the flat premium 27% (0.27) was chosen, as this is the premium used by the largest pension fund in the Netherlands, ABP (29).
Although there is no clear data available on the variable premium, it is known that the premium is low for younger participants and increases over time once you become older, as is visible in Figure 2.2. Since this curve closely represents an exponential curve, it was

chosen to model the premium in the old system as an exponential curve as well, starting from 0.1 and increasing to 0.4 over a period of 40 years, since this is the maximum period one is in either the old or new pension system.

Lastly, this model will be run without crisis events, to see the general effect of the pension reformation. Introducing volatility spikes into the model will likely cause the results of the new pension system to get very high or very low. The aim of this simulation is to provide insight on the average impact of switching midway through one's career. Implementing volatility spikes will skew the results and not allow for general conclusions.

4.3.2 Modelling

First, salaries and pension contributions are calculated under the old system. The salary tracks how much money an individual makes year-round over time. The pension contribution, or pension premium, is the percentage of the salary an individual contributes to their pension portfolio, which will be saved for retirement. In real life, the pension premium is divided between the employer and the employee. The 27% premium used by ABP is divided into 18.9% paid by the employer and 8.1% paid by the employee (29). This 8.1% is what can be seen on the monthly pay slip. In order to simplify the calculations, the premium in this research is calculated yearly as a percentage of the yearly salary, which results in the same amount as when calculating it monthly and summing every month. Furthermore, in this research, no distinction will be made between the premium paid by the employer and the employee. As the goal is to calculate the total gathered pension contribution at the end of an employee's career, only the total premium set aside on a yearly basis is being taken into account. The employer's contribution is only relevant when one wants to calculate the actual monthly salary payout of an employer before retirement, which is not relevant for this investigation.

The salary and pension contribution over time are tracked separately. Every year, the salary growth rate is applied to the salary, after which the pension premium corresponding to that year is being calculated based on the premium curve discussed above. Both salary and contribution are tracked for 10, 20, 30 and 40 years in the old system, all starting from the same initial salary, chosen to be €30.000. Salaries are chosen to be gross salaries, taxes will not be taken into account. After a certain number of years in the old system, an employee either has to switch to the new pension system or has retired before the new system was introduced. In the last case, the employee will have spent 40 years in the old system. In the case of switching midway through one's career, it is relevant to know how high the employer's salary is when the switch is happening. This salary is then taken as a starting salary for the simulation of the new system. For example, when an individual has to switch to the new system after spending 10 years in the old system, the salary after ten years will be the starting point for simulating the new system. Logically, when one has spent 10 years in the old system, they will spend 30 years in the new system to get the total of 40 years. This works similarly for switching at other points in time. This research will distinguish between five cases:

1. 40 Years in the Old System
2. 30 Years in Old System + 10 Years in New System
3. 20 Years in Old System + 20 Years in New System
4. 10 Years in Old System + 30 Years in New System
5. 40 Years in New System

The contribution is also tracked since this is relevant later on, when the total contribution an individual builds in their lifetime needs to be calculated to be able to determine pension payouts after retirement.

In the old system, the contribution was tracked by taking the value of the variable premium for every year and multiplying this by the salary in that year. In that way, both the salary growth and the growth of the variable premium will contribute to an increase in pension contribution over the years. When switching to the new system, the variable premium disappears and will be replaced with the flat premium of 0.27. The starting point for modelling the contribution in the new system is therefore the ending salary of the old system, multiplied by 0.27. In the new system, only the salary growth will contribute to an increase in the contribution. Simulating the effect of the financial market on the pension contribution in the new pension system is done in a similar way to the model proposed in Section 4.2.2. Monte Carlo Simulation is applied to simulate again 5000 pension contribution paths for 10, 20, 30 and 40 years in the new system. Here, 40 years account for a full career spent in the new system.

4.3.3 Pension Payouts

In order to quantitatively compare the impact of switching from the old to the new system at a certain time, the total pension contribution for each generation is calculated. Based on this total contribution gathered over an individual's working lifetime, the expected monthly pension payouts can be calculated based on indexation and life expectancy. The total contribution is calculated by summing the contributions made each year in the old system plus the contributions made each year in the new system, depending on when the switch happened (or not). Since the new system works with a Monte Carlo Simulation and returns 5000 potential total contributions, it was chosen to sum the mean value of all paths each year to obtain the collected contribution.

These results will already give a solid indication of which generation will profit more from the reformation. This total pension contribution will then be divided over the months someone will live to give a clear indication on the monthly pension payouts. Indexation will be applied on a yearly basis. Historical indexation levels of pension funds can be found online. Since it is outside the scope of this research to predict pension indexation for the upcoming years, it was decided to use the most recent indexation level applied by the pension fund ABP. ABP was chosen because it is the biggest pension fund, and their percentage for the flat premium was used in this research as well. On the ABP website, it can be found that the current indexation level for 2025 is 1.84% (30). This will be compared with the scenario in which no indexation is applied. This has namely happened in the past, as can be seen in the historical indexation data of ABP where between 2011 and 2022, no indexation has occurred. As discussed in Section 2.1.2, a pension fund can only provide indexation when the funding ratio allows it. Since interest rates continued to decline over time, the funding ratio dropped, and ABP was not allowed to execute indexation. Pension funds that are expected to reform can take advantage of the benefits of the new pension system. One of these benefits is a relaxation of the indexation rules. Indexation is then possible from a funding ratio of 105% instead of 110%. As ABP is transferring to the new system as of 2027, these relaxed rules apply to them, and they can start to apply indexation again. When they fully convert to the new pension system, the funding ratio will be completely eliminated.

To decide how the total contribution of a participant should be divided over their life as a retiree, pension funds use life expectancy tables and recalculate the life expectancy of an individual again every year. Since this approach is outside the scope of this research, it was decided to use the life expectancy of a 65-year-old in 2024, which can be found in the so-called AG2024 Prognosetafel (31). These life expectancies differ for men and women, as women are known to live longer than men. The AG2024 Prognosetafel provides estimated life expectancies for 0- and 65-year-olds. The life expectancy of a 0-year-old is lower as life expectancy at birth includes all the risks of dying at younger ages. Since this research focusses on retirees, the probabilities of dying at younger ages are not relevant. Therefore, it was decided to take the values of the 65-year-olds. Since the historical data used in this investigation ranged from 2011-2023, it was decided to take the prognosis of 2023, which is the same as for 2024 when rounded up to whole years. According to the AG2024 Prognosetafel, 65-year-old men will on average live another 19 years, while 65-year-old women will on average live another 21 years (31). Since the current retirement age is 67 and not 65, it was decided to apply indexation for 17 years for men and 19 years for women. In this way, women receive a smaller payout per month, since they will need more money later as they have a higher life expectancy. This is the same method that actual pension funds use to calculate their pension liabilities.

5

Results

In this chapter, the results obtained by the methods presented in Section 4 will be presented. First, the most important findings will be summarised. Then, the results of the models on the historical data will be discussed. A comparison will be drawn between the benchmark, risky, and conservative portfolios. Then, the outcomes of the simulations will be addressed. These results will highlight potential portfolio price paths for certain time horizons. Lastly, the results of the combination of the old and new systems will be discussed. These results will demonstrate which generation benefits the most from the pension reformation.

5.1 Summary of most important findings

First, the rolling Value at Risk models show that the different asset allocations of risky, benchmark, and conservative indeed behave in the desired way. The risky portfolio shows higher returns but greater downside risk. The conservative portfolio demonstrates lower volatility and, therefore, lower downside risk. The benchmark portfolio operates in the middle. The Variance-Covariance method seems to be able to give good risk estimates but underestimates the risk in crisis events. The Historical Simulation method seems to be poor at capturing tail risk, as the VaR level is lower and underestimates crisis events as well. GARCH(1,1) is the only model capable of capturing volatility clustering, showing higher VaR and ES in crisis periods, and following the pattern of returns more closely. All models demonstrate that a longer time horizon (1-Week compared to 1-Day) leads to a higher VaR and ES.

The Monte Carlo simulation predicting future portfolio values for different time horizons confirms that riskier portfolios demonstrate a larger spread of possible outcomes, meaning

that the returns can be much lower or much higher than those of other asset allocations. Conservative portfolios exhibit a smaller spread, which means that they are less affected by financial turmoil. Furthermore, the simulation shows that the longer the simulated time horizon is, the higher the VaR and ES get. However, the mean returns of the risky portfolios are not necessarily always higher than the benchmark returns. This suggests that taking on more risk does not guarantee a higher profit, even over a time horizon of 40 years. Introducing spikes shows a direct increase in VaR and ES. Notable is that the effect of the spike does not necessarily diminish over time, not even for a horizon of 40 years.
The Practical Scenario Analysis exhibits that the switch from variable to fixed premium has a clear impact on the yearly gathered contribution. Younger generations spending their entire life in the new system collect the highest yearly contribution considering average investments. The second highest contribution is gathered by the older generation, spending their entire working life in the old system. The age cohort who spent 30 years in the old system and has to switch for the last ten years gathers the lowest total contribution, and therefore the reformation will affect them the most. Lastly, assuming investments that do not perform well, the model shows that their is almost no difference between spending 40 years in the new system or in the old system. It should be noted that financial crises are not included in the model.

5.2 Historical Data Analysis

The results of the Historical Analysis will show how financial data can change over a long period of time. The different risk measures and statistical models will show how the different portfolios will react to market changes. All results will be shown for a 95% confidence level. The results for the 99% confidence level can be found in the appendix.

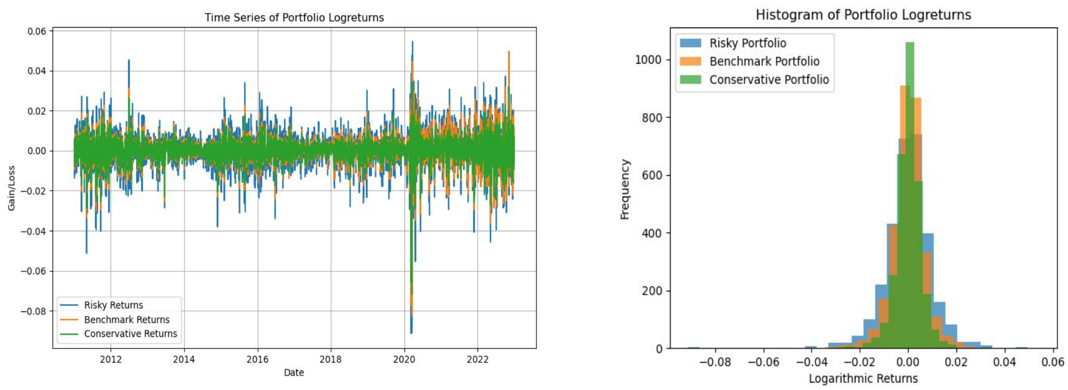
5.2.1 Data Visualisation

Visualising the portfolios with the different risk profiles allows one to check whether the first instincts are right: namely, that a riskier portfolio should yield higher returns, but similarly higher losses. Both Figure 5.1(a) and Figure 5.1(b) on the next page show that these expectations are indeed met. Figure 5.1(a) shows the time series of the returns on the historical data. It can be seen that the returns of the risky portfolio are higher at every point, including in the peaks. Similarly, the losses are also greater at each time point. The conservative portfolio tends to be the most stable over time, but therefore the gains are also lower than those of the benchmark and risky portfolios for all time points. Around

2020, a clear peak is visible; this is caused by the Covid-19 crisis. In this peak, losses of more than 8% were recorded in one day, which is remarkably high. These excessive losses will likely be reflected in the results of the Value at Risk.

Figure 5.1(b) shows the distribution of the returns. From the figure, it is clear that for the conservative portfolio, most returns are concentrated around zero. The risky portfolio, on the other hand, has a higher frequency in extremes compared to the other portfolios, both in losses and in gains.

This aligns with expectations and demonstrates that the asset allocations implemented behave as desired. This allows for a comparison between these asset allocations for the upcoming models.



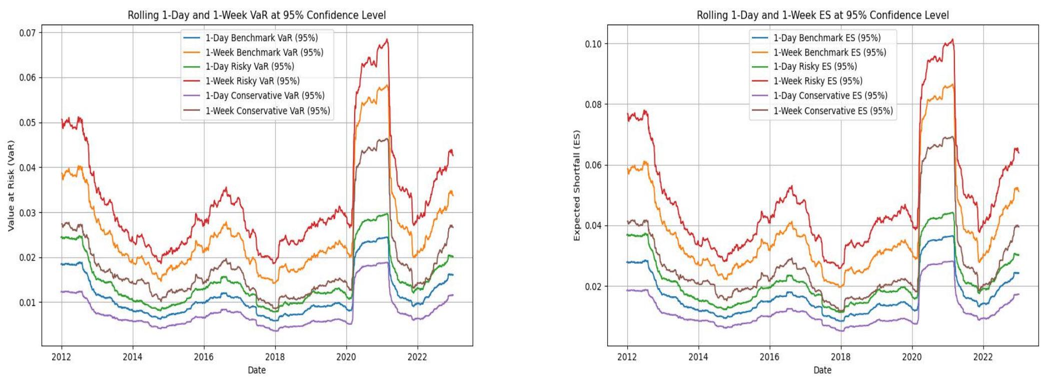
(a) (b)

Figure 5.1: (a) Time Series of Portfolio Returns (b) Histogram of Portfolio Returns

5.2.2 Variance-Covariance Method

In Figures 5.2(a) and 5.2(b) on the next page, the Value at Risk and Expected Shortfall for a 95% confidence level, according to the Variance-Covariance method, are plotted over time. It is clear from the plots that the VaR and ES of the risky portfolio, the portfolio of the younger generation, is at all times higher than for the other portfolios. Logically, the 1-Week VaR and ES are higher than the 1-Day VaR and ES, as a portfolio can lose more money over a longer period of time. When comparing Figure 5.2(a) to the graph of the time series of the returns, it is noticeable that the VaR follows the pattern of the losses relatively well. There is a clear peak around 2020, which represents the Covid-19 crisis. For all portfolios, these peaks are in the same place, but the magnitudes differ. The peak of the risky portfolio is far more extreme than the peak of the conservative portfolio. For the conservative portfolio, the 1-Day VaR seems to be fairly stable, even in the crisis period.

The peaks of the 1-Day VaR for all portfolios lie around 2% to 3%, which is lower than the daily losses of the time series of returns around this time. The VaR levels over the stable periods seem to follow the magnitude of the losses in the time series more accurately. However, the model seems to capture the volatility of larger losses less accurately, as small incidental spikes in the returns is not covered by Figure 5.2(a). From Figure 5.2(b) it is also visible that the ES seems to follow a pattern similar to that of the VaR. Clearly, its values are higher as they represent the losses behind the VaR threshold.



(a) (b)

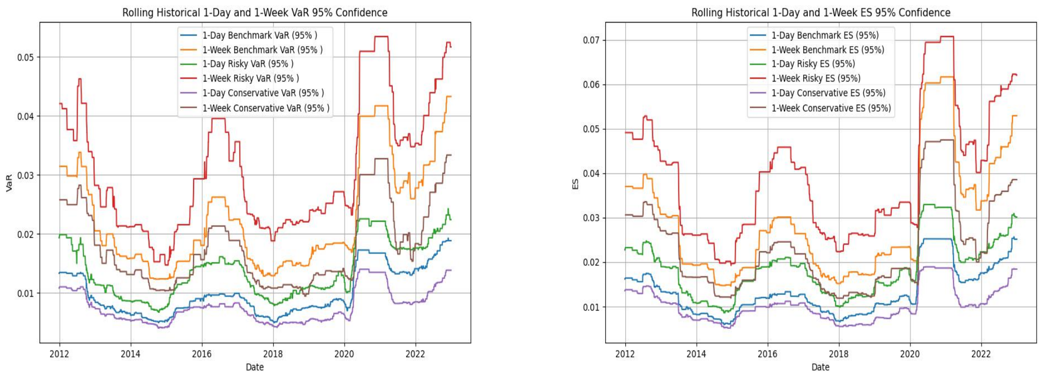
Figure 5.2: (a) Rolling 1-Day and 1-Week VaR 95% (b) Rolling 1-Day and 1-Week ES 95%

5.2.3 Historical Simulation

The Historical Simulation produced graphs of the VaR and ES over time that are quite similar to those of the Variance-Covariance method, as can be seen from Figures 5.3(a) and 5.3(b) on the next page. However, these graphs tend to be a bit more "blocky" and less smooth, which makes sense since a rolling-window approach is used on sorted losses. Therefore, moving the window one day further gives exactly the same portfolio losses, apart from the newly added loss, which is the loss directly next to the previous loss.

Again, from the graphs it is clear the assumption that the risky portfolio yields a higher

VaR is met. This also holds for the Expected Shortfall, visible in Figure 5.3(b). The VaR values resulting from this method are overall lower than those of the variation-covariance method, which would indicate that the method does not capture the extreme losses well either. This will be tested in more detail in Section 6.



(a) (b)

Figure 5.3: (a) Rolling Historical 1-Day and 1-Week VaR 95% (b) Rolling Historical 1-Day and 1-Week ES 95%

5.2.4 GARCH(1,1)

The GARCH(1,1) model calculated the volatilities of the portfolios on the historical data. These can be seen in Figure 5.4 on the next page. Notable is that the volatilities of the portfolios lie relatively close to each other. The risky portfolio has slightly higher volatilities overall, but the benchmark and conservative portfolios are quite similar to one another. The model seems to be able to identify volatility clustering: around 2020 a clear volatility spike is visible, and periods of higher volatility, for example from 2015 to 2017, seem to follow each other. Similarly, a low volatility period appears to exist between the end of 2013 and the beginning of 2015. This shows that periods of high volatility follow periods of high volatility and periods of low volatility follow low volatility.

In Figures 5.5(a) and 5.5(b) on the next page, the VaR and ES with GARCH(1,1) volatilities are visible. It is immediately clear from these figures that this model seems to be far more sensitive to changing volatilities. The values of the VaR seem to follow the daily losses significantly better than the previous methods. It should be noted that the

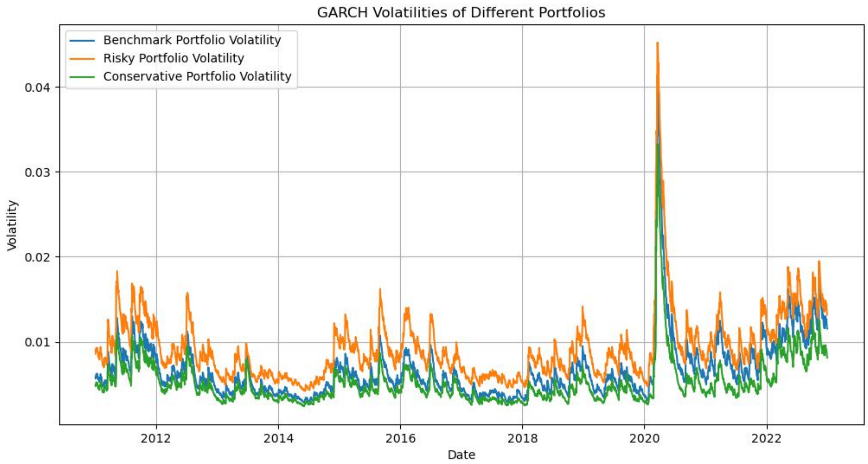
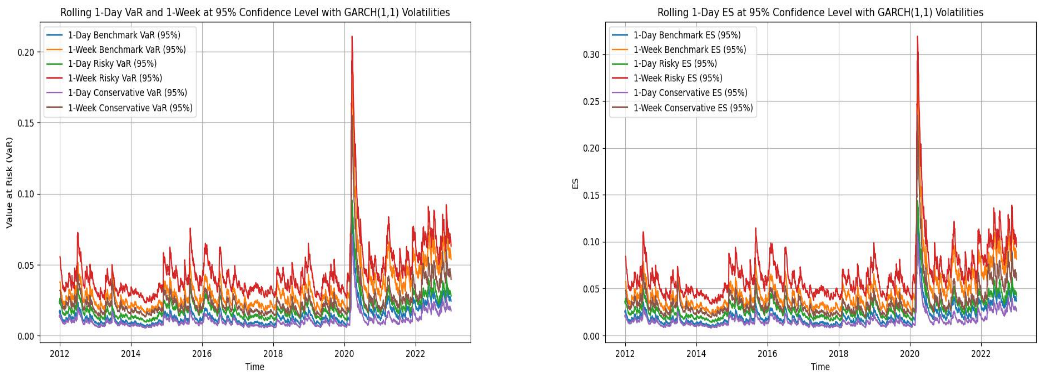


Figure 5.4: GARCH(1,1) Volatilities

1-week VaR levels in the Covid peak are exceptionally high, more than 20% for the 95% VaR (Figure 5.5(a) on the next page), and even more than 35% for the 99% VaR (Figure 8.3(e) in the Appendix). Although this high VaR appears only for a short period of time, it is valuable to note that portfolios with riskier assets can lose such excessive amounts of money over a short period of time during a financial crisis. The graphs of the Expected Shortfall show that there are even higher losses behind the VaR, indicating that such a financial crisis can result in an enormous shift in portfolio value.

It appears as if the 1-Week VaR values are not much higher than the values of the 1-Day VaR, but this is mostly due to the scale of these plots, as the large spikes cause the lower values to be more squeezed together at the bottom of the graph.



(a) (b)

Figure 5.5: (a) Rolling GARCH(1,1) 1-Day and 1-Week VaR 95% (b) Rolling GARCH(1,1) 1-Day and 1-Week ES 95%

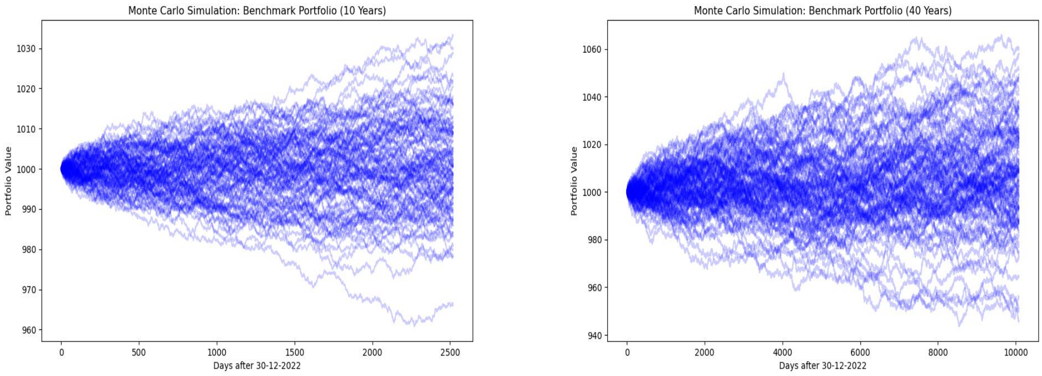
5.3 Monte Carlo Simulation

The results of the Monte Carlo simulation will indicate how the length of a portfolio and the occasional financial crisis influence the risks of a portfolio over different time horizons.

5.3.1 Constant Volatility

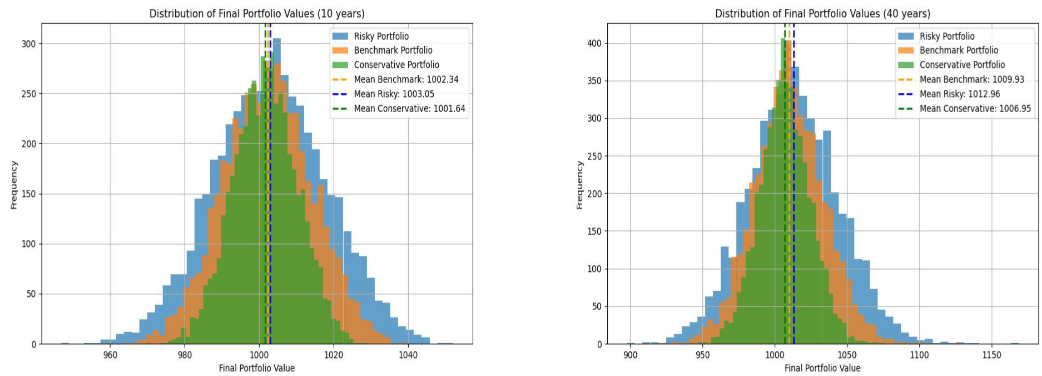
A Monte Carlo simulation with constant volatilities produced the paths visible in Figure 5.6 on the next page. The figure demonstrates 100 of the 5000 paths the simulation produced. For clarity, it was decided only to plot 100 of the paths, as visualising 5000 paths would produce a large blur. Therefore, the actual spread of all 5000 paths could be greater than is visible in Figure 5.6. In Figure 5.7 on the next page, the final values of all portfolio paths are visualised, and therefore the actual spread is visible here.

It can be seen from Figure 5.6 that the paths behave as expected, without anomalies. It was decided to display only the paths of the benchmark portfolio, as the risky and conservative portfolio paths behave in similar ways, but with a larger spread for the risky portfolio and a less broad spread for the conservative portfolio. These spreads are illustrated by Figure 5.7, which shows the distribution of portfolio values on the final day for each risk level. The risky portfolio reports both higher losses and gains, as one would expect, indicating that the simulation provides realistic results. The mean of the portfolio values on the final day is also provided. It is interesting to note that the risky portfolio, despite also recording large losses, shows the highest mean value, both after 10 and 40 years. The conservative portfolio records the lowest mean, both after 10 and 40 years. Furthermore, it can be seen that the mean increases over time, as after 40 years the mean value of all portfolios is higher than it was after 10 years. This would suggest that investing over a longer horizon will yield a higher mean profit.



(a) (b)

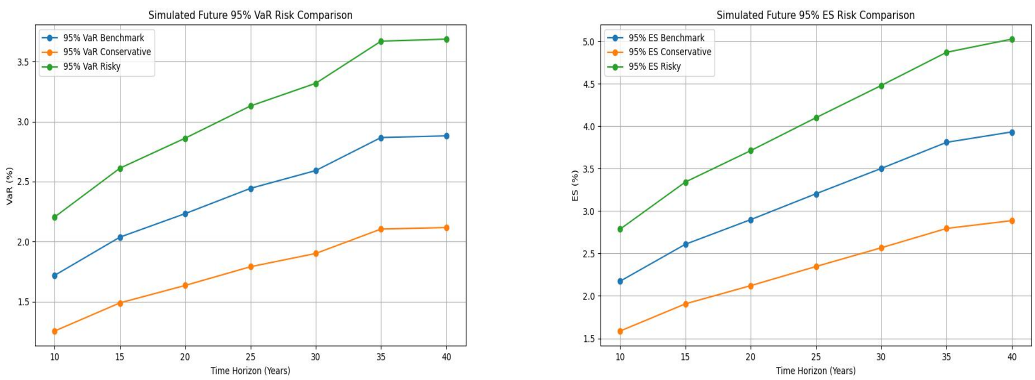
Figure 5.6: (a) 100 Simulated Paths Constant Volatility 10Y (b) 100 Simulated Paths Constant Volatility 40Y



(a) (b)

Figure 5.7: (a) Final Values Constant Volatility 10Y (b) Final Values Constant Volatility 40Y

In Figures 5.8(a) and 5.8(b) below, the VaR and ES are plotted for different asset allocations over the time horizon. As expected, it can be seen that the VaR increases as the time horizon increases. Furthermore, the VaR for the risky portfolio of young participants is higher at all times, compared to the benchmark and conservative portfolios. In addition, note that the differences between the three portfolios seem to be consistent over time. The difference between the conservative and benchmark portfolios is similar to the difference between the benchmark and risky portfolios over the entire horizon. Similarly to the results of Section 5.2, the Expected Shortfall in Figure 5.8(b) follows the same pattern as the VaR values in Figure 5.8(a).

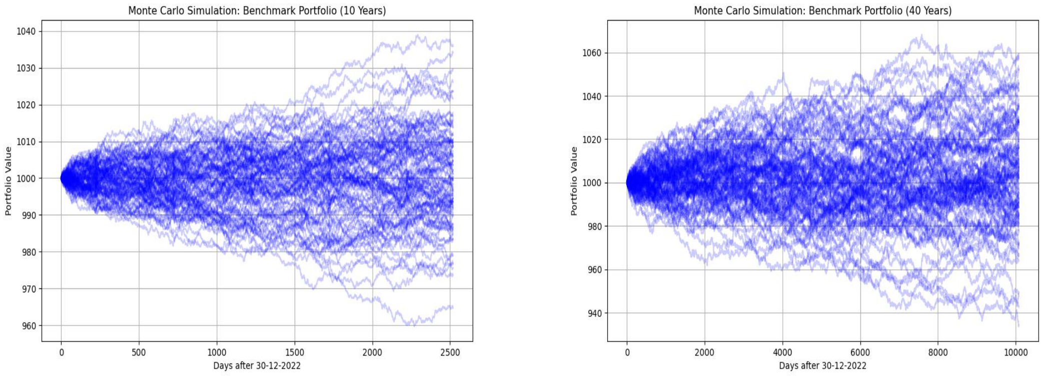


(a) (b)

Figure 5.8: (a) Simulated future 95% VaR with Constant Volatility (b) Simulated future 95% ES with Constant Volatility

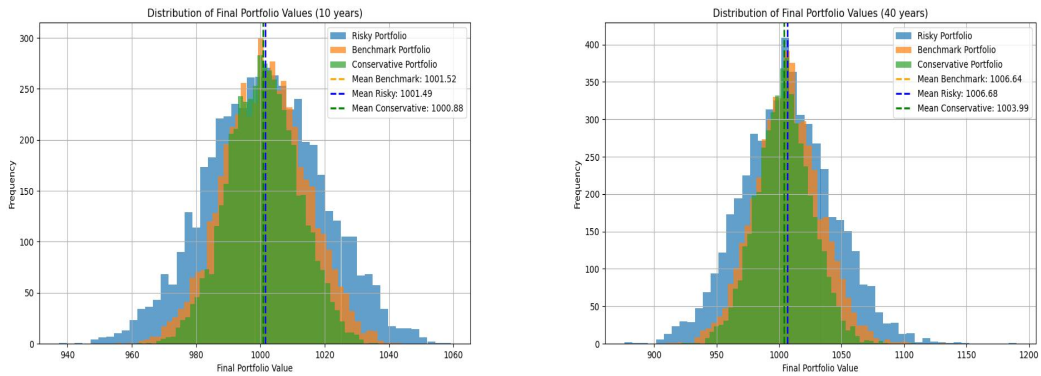
5.3.2 GARCH(1,1) Volatilities

The Monte Carlo simulation with GARCH(1,1) volatilities produced simulated paths similar to the constant volatility method. This is shown in Figure 5.9 on the next page. The distribution appears to be relatively similar, and again no anomalies seem to occur. Clearly, the spread after 40 years is larger than after 10 years, as there is more time for paths to end up in higher or lower outcomes. Figure 5.10 on the next page shows the distribution of the final values of the portfolios. The pattern is largely the same as that of the constant volatility method. Remarkable is the fact that the mean of the risky portfolio is extremely close to the mean of the benchmark portfolio, while for the constant volatility method, the risky portfolio clearly had a higher mean. Here, after 10 years, the mean of the risky portfolio is even lower than the mean of the benchmark portfolio. This would suggest that riskier investments do not necessarily yield higher returns after longer periods of time. Furthermore, the mean final value of all portfolios is higher than the initial value of 1000. This suggests that on average profits are made with this particular portfolio, regardless of the asset allocation.



(a) (b)

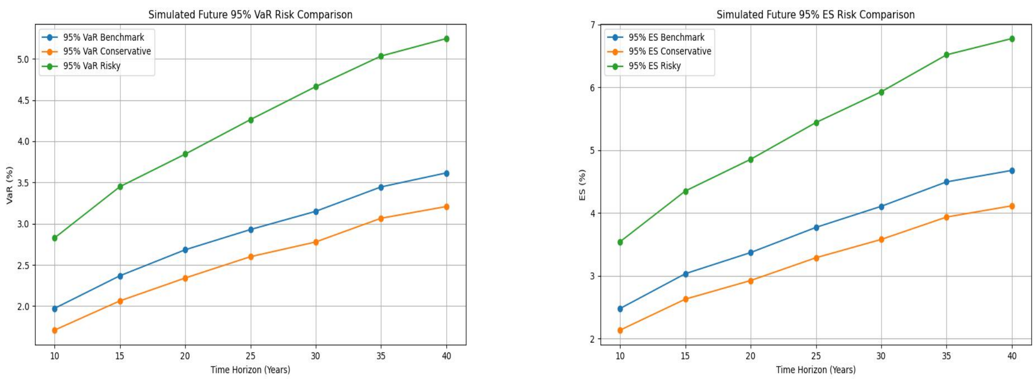
Figure 5.9: (a) 100 Simulated Paths GARCH(1,1) Volatilities 10Y (b) 100 Simulated Paths GARCH(1,1) Volatilities 40Y



(a) (b)

Figure 5.10: (a) Final Values GARCH(1,1) Volatilities 10Y (b) Final Values GARCH(1,1) Volatilities 40Y

From Figure 5.11 below it is clear that introducing GARCH volatilities in the Monte Carlo simulation yields similar results to keeping the volatility constant. The increasing pattern of the VaR and ES with time resembles that of Figure 5.8. However, a clear difference is the fact that the VaR of the risky portfolio here is distinctly higher than the other two portfolios. The VaR's of the benchmark and conservative portfolios lie significantly closer to each other. This is in line with the volatility pattern that can be seen in Figure 5.4. There, the risky portfolio yields higher volatilities compared to the benchmark and conservative portfolios as well. This is due to the way the allocations are selected. Therefore, this suggests that GARCH volatilities have a significant impact on the riskiness of the portfolio. Furthermore, the VaR values outputted by the Monte Carlo simulation with GARCH volatilities lie a little higher than the values outputted by the constant volatility model. Again, the Expected Shortfall, shown in Figure 5.11(b), follows the same pattern as the VaR.

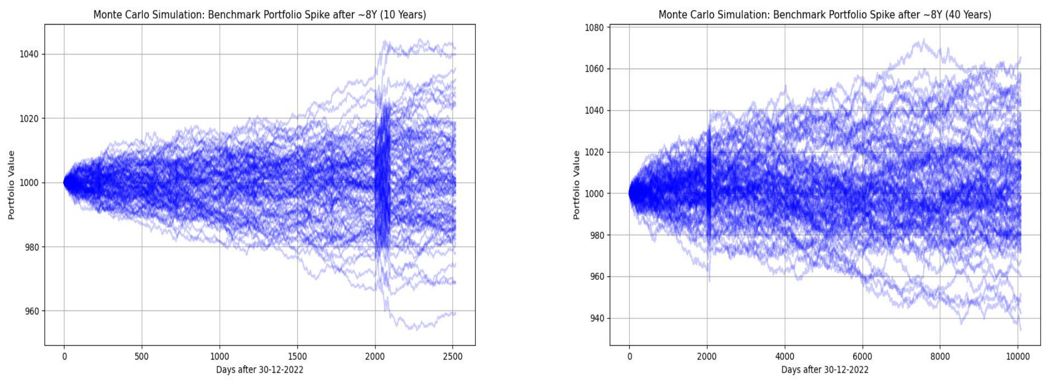


(a) (b)

Figure 5.11: (a) Simulated future 95% VaR with GARCH Volatilities (b) Simulated future 95% ES with GARCH Volatilities

5.3.3 Stress Testing

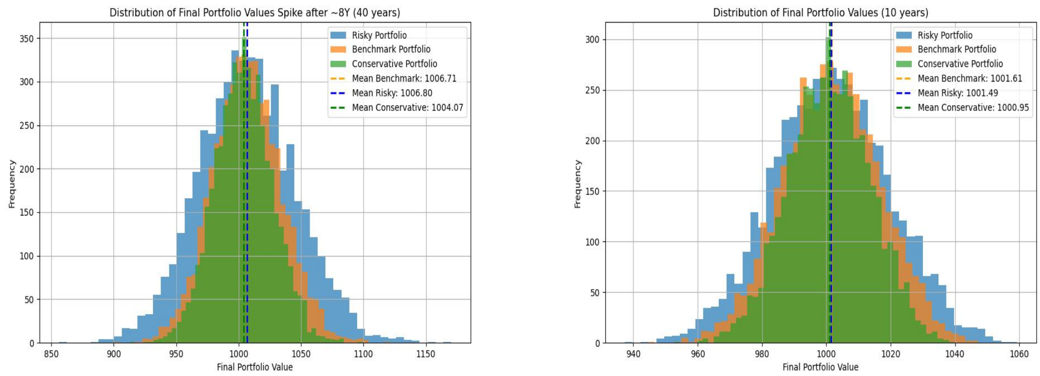
Introducing volatility spikes in a simulation can cause portfolio paths to behave differently, as shown in Figure 5.12. This effect is particularly evident in Figure 5.12(a) on the next page, where the spike occurs after approximately 8 years. The results of spikes happening at other time points can be found in the Appendix. This causes the portfolio paths to either skyrocket or plummet, which is noticeable right after day 2000, when the graph becomes more erratic. In Figure 5.12(b), the effect is less evident as the portfolio paths are plotted for 40 years, but the effect is the same. Figure 5.13 on the next page demonstrates the distribution of the final values of the portfolios. It can be seen that the distribution is less smooth compared to the results without spikes. Again, the mean of the risky and benchmark portfolios are quite close to one another. The risky portfolio yields a higher mean return, but only barely. Furthermore, Figure 5.13 demonstrates a higher spread among the final values compared to the spikeless portfolio paths. This makes sense as the spikes cause both bigger gains and bigger losses.



(a) (b)

Figure 5.12: (a) 100 Simulated Paths Volatility Spike after $\sim$ 8Y (10Y) (b) 100 Simulated Paths Volatility Spike after $\sim$ 8Y (40Y)

Introducing volatility spikes into the Monte Carlo simulation leads to the results visible in Figure 5.14 on page 61. Results are provided for the introduction of a volatility spike at different time points. It is apparent that the introduction of a spike leads to a sudden increase in the VaR, as one would expect. It is interesting to note from Figure 5.14(a) that a spike appearing after 8 years still results in a higher VaR after 40 years. The difference between the two decreases over time, but the VaR never gets back to the level it was before the spike in the 40 years that are simulated. This would suggest that a portfolio does not easily recover from a financial crisis, even over a longer period of time. Furthermore, it is observable that the timing of the spike does not seem to affect the size of the impact. All risky portfolios show a VaR of around 5.5% after 40 years, regardless of when the crisis happened. Similar results appear for the benchmark and conservative portfolios. Lastly,



(a) (b)

Figure 5.13: (a) Final Values GARCH(1,1) Volatility Spike after $\sim$ 8Y (10Y)(b) Final Values GARCH(1,1) Volatility Spike after $\sim$ 8Y (40Y)

it can be seen that a more risky portfolio has a higher increase in VaR than a more conservative portfolio. This also makes sense intuitively, since the volatility is increased by a certain factor. Larger numbers logically increase more than smaller numbers when the multiplication factor is the same. Therefore, the VaR will likely increase more as well.

Since 40 years does not seem long enough for a portfolio to recover from a spike, the VaR is also calculated for 100 years and 200 years, once without a spike and once with a spike after 8 years. The results of this can be seen in Figure 5.15 on page 62. The figure shows that, while after 40 years the VaR of the portfolio with a spike is still visibly higher, this effect diminishes over time. After 100 years, the portfolio still has a slightly higher VaR, but after 200 years this difference is negligible. This is an interesting effect, as one of the key arguments for investing in riskier stocks for younger generations is that they have more time to recover from potential losses. However, this model demonstrates that the 40 years a younger participant spends in the new system is not enough to fully recover. Therefore, after 40 years, their VaR after a crisis event is still higher than the VaR without a crisis event.

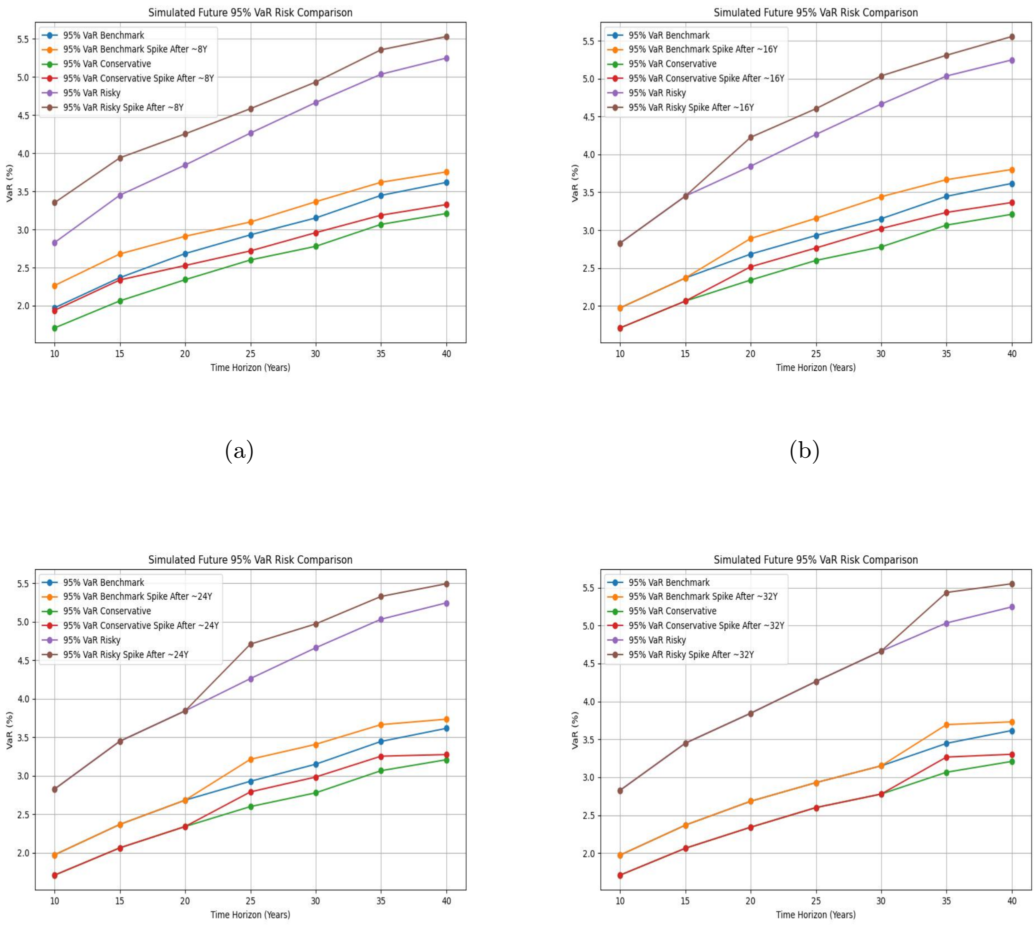


Figure 5.14: (a) Simulated Spike 95% VaR after 8 Years (b) Simulated Spike 95% VaR after 16 Years (c) Simulated Spike 95% VaR after 24 Years (d) Simulated Spike 95% VaR after 32 Years

5.3.4 Practical Scenario Analysis

As explained in Section 4.3, this research distinguishes between five scenarios in which a mid-life change is made or not. It should be noted that all the plots provided in this section are based on low-salary growth. Since the results for high- and mid-salary growth provide the same pattern of results only with higher or lower values, it was decided to omit them from this section.

Figure 5.16 on the next page shows the contribution collected over a lifetime in case of spending 40 years in the old pension system. The variable premium following an expo-

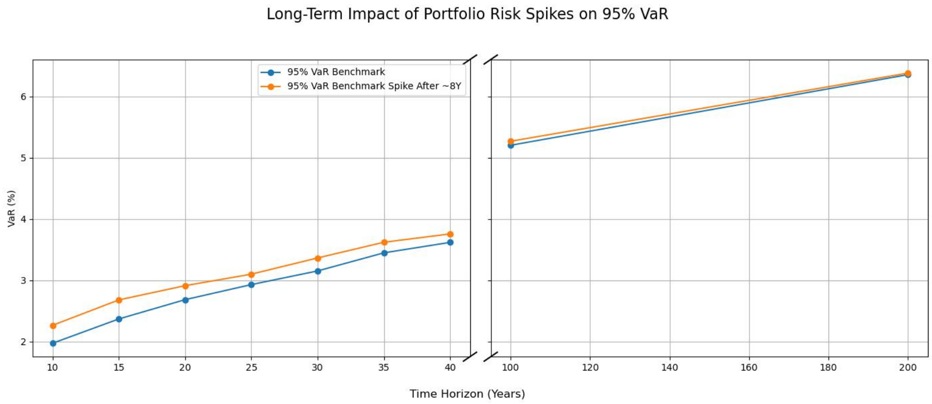


Figure 5.15: Long-Term Spike Impact VaR 95%

nential distribution can be seen here quite clearly. Furthermore, the stepwise increases represent the salary growth, which is different for different life phases. This is hard to see due to the strong effect of the variable premium, but the growth rates do follow an S-curve. The contribution of each year will be summed to obtain the total contribution gathered over the life cycle.

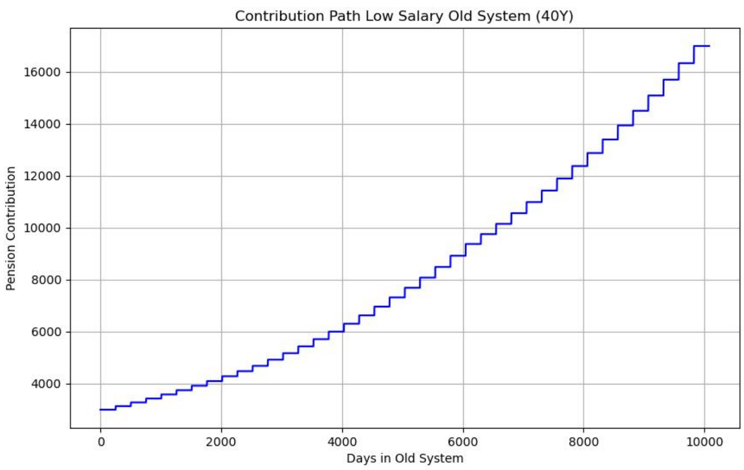
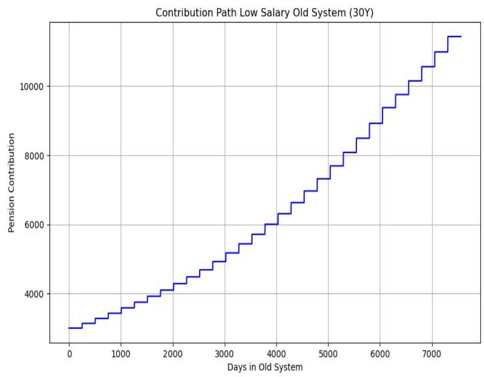
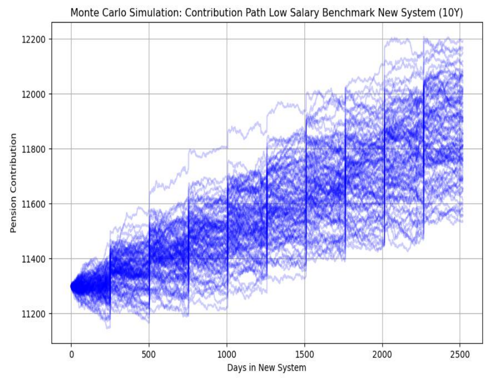


Figure 5.16: Contribution Path Old System 40Y

In Figure 5.17 on the next page, the transition to the new system after 30 years in the old system is visualised. Figure 5.17(a) shows the contribution path in the old system, where Figure 5.17(b) represents the 10 years in the new system. A distinct difference is clearly the dependence on the financial market in the new system, as is demonstrated by the 100 visualised paths out of the 5000 simulations. The contribution in the old system still follows the shape of the exponential distribution with the stepwise salary growth. It is cut off at 30 years since that is the moment the switch to the new system is made. It can be seen that the contribution after 30 years is close to €12.000. This is because the variable premium increases from 0.1 to 0.4, but is cut off before it reaches 0.4 as only 30 years are being spent in the old system. Since the fixed premium in the new system is set to 0.27, Figure 5.17(b) shows that the contribution starts at €11.300. Hence, the premium has gone down due to the switch from the old to the new system, and thus from the variable to the fixed premium. Clearly, there is a loss of contribution due to this switch. Furthermore, Figure 5.17(b) shows that the flat premium causes a smaller increase in contribution. The only factor contributing to the increase in contribution is salary growth. To calculate the total contribution, the yearly contributions in the old system are summed. To get an idea of the contribution collected in the new system, it was decided to take the mean value of the paths for each year. Summing the old and new system together then gives the total contribution over the life cycle.



(a)

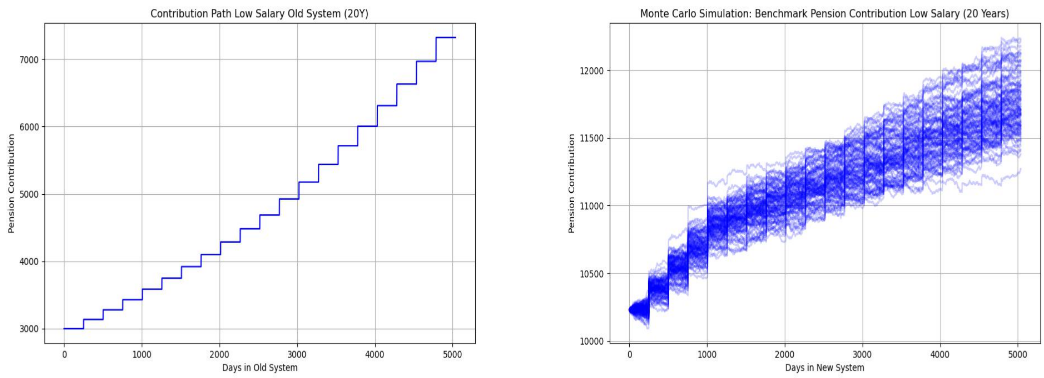


(b)

Figure 5.17: (a) Contribution Path Old System 30Y (b) Contribution Path New System 10Y

Figure 5.18 on the next page is similar to Figure 5.19 on page 65, only now the switch is executed after 20 years instead of 30. Therefore, Figure 5.18(a) is cut after 20 years. The exponential shape is still visible in the graph. The contribution level after 20 years is just above €7000. The starting contribution in the new system is around €10.250, which is significantly higher than €7000. This is because the variable pension was cut off before

it got a chance to increase a lot. Therefore, a jump to the flat premium of 0.27 is made, and the contribution is higher. Had the individual fully started in the new system, the contribution would have been higher in the first 20 years of the life cycle as well.
Since 20 years will be spent in the new system, the contribution is increasing more in the beginning of Figure 5.18(b), since the individual is still in a time in his career when the salary increase is higher. Again, the total contribution is calculated by summing the contribution made for each year.

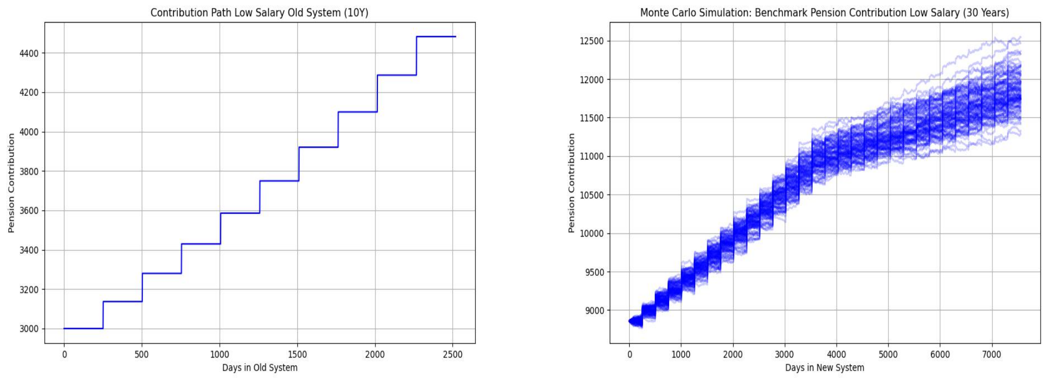


(a) (b)

Figure 5.18: (a) Contribution Path Old System 20Y (b) Contribution Path New System 20Y

In Figure 5.19, the effect of switching after 30 years in the old system is visualised. It can be seen that 10 years in the old system is not enough to see the effect of the exponentially distributed variable premium. The contribution after 10 years is only slightly above €4400, as is visible in Figure 5.19(a). The jump to the flat premium in the new system is then significant, as the contribution there starts around €8900. Similarly to Figure 5.18(b), the salary increase in the beginning of Figure 5.19(b) is high, as after 10 years, the individual is in a life phase with a strongly increasing salary. Again, the total contributions of this life cycle are calculated.

Figure 5.20 demonstrates the contribution when spending 40 years in the new system. The S-curve followed by the salary growth can be seen here. Since the new system uses the flat premium, the simulation starts with the starting salary (€30.000) multiplied by



(a) (b)

Figure 5.19: (a) Contribution Path Old System 10Y (b) Contribution Path New System 30Y

0.27, giving €8100. To calculate the total contribution over the life cycle, the mean of each portfolio is again taken every year and then summed.

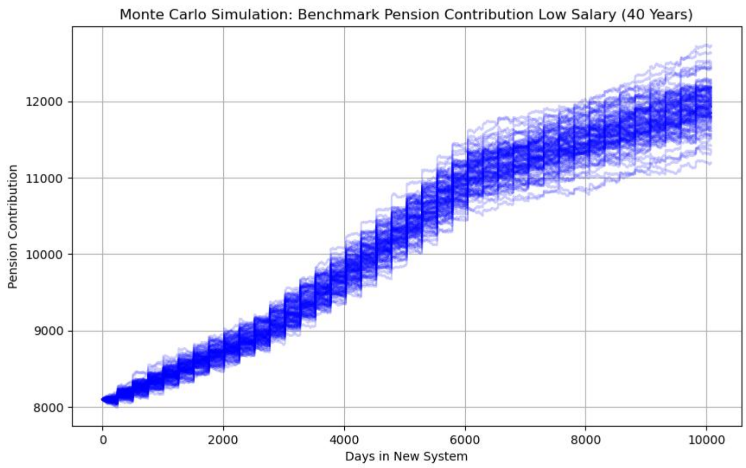
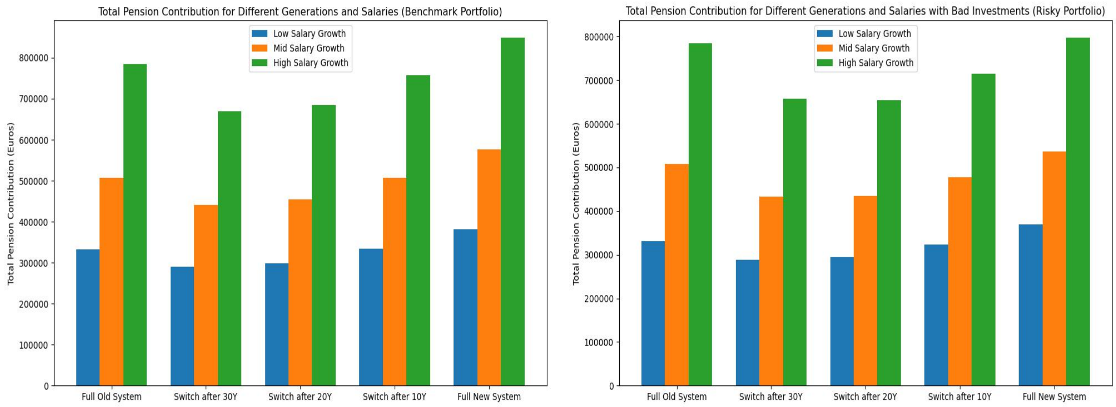


Figure 5.20: Contribution Path New System 40Y

Clearly, the interesting part of assessing all these different life-cycles is how the total contributions in the end compare. Figure 5.21(a) on page 67 shows the results of this. The low salary growth represents an industry in which it is difficult to climb the career ladder and the salary does not increase much over time. In contrast, high salary growth

represents a career in which it is easy to make career progress, and thus the salary increases significantly faster over time. Therefore, the total contributions for high-salary growth are significantly higher than for low-salary growth.
It is interesting to see that all levels of salary growth seem to follow the same pattern regarding the total contribution of different life cycles. It can be seen that spending the full life cycle in the new pension system is most beneficial, as the total contribution is the highest. This suggests that the younger generations of around 25 years of age at the time of writing this thesis and all generations in the future will benefit the most from the pension reformation. Spending 40 years in the old system still yields a relatively high total contribution, but less than it would have been in the new system. Hence, all generations that were already retired before the reformation happened will have collected less contribution than the future generations will collect on average.
However, the key insight here is that switching mid-way in a career always yields a lower collected contribution than when spending 40 years in either the old or the new system. Individuals who switch after 30 years in the old system will realise the lowest total contribution according to the model. This represents the generation that is almost at retirement age but still has to fulfil their last employed years in the new system. They built up most of their pension in the old system under the variable premium, which is low at the beginning of a career. Right before the variable premium gets to a high level, the individual switches to the new system, and the contribution 'drops' back to 0.27, causing them to miss out on contribution. The same logic applies to switching after 20 years and after 30 years.

To determine whether the switch from a variable to a flat premium is the main reason for these discrepancies, or if the financial market also has a significant impact, it was decided to run the new salary model again, but instead of using the mean of each portfolio path, the minimum value was taken. This was performed on the risky portfolio instead of the benchmark as well, to obtain the worst-case scenario for younger generations. The result of this is visible in Figure 5.21(b). It can be seen that even in the worst-case scenario, the younger generations still outperform all the other age cohorts. The differences are clearly much smaller, and the full old pension system is almost as profitable as the full new system. However, it is remarkable that the new system still outperforms all the other scenarios. This suggests that the main factor contributing to the discrepancies is the flat premium. It should be noted that this model was run without adding volatility spikes to mimic a financial crisis. It is likely that a financial crisis will result in different outcomes.



(a) (b)

Figure 5.21: (a) Total Contribution over Lifetime (b) Total Contribution over Lifetime (Bad Investments)

Figure 5.22 on the next page demonstrates the effect of indexation over the years after retirement. It was chosen to show the indexation on the payouts for individuals switching after 30 years, as these are the lowest payouts. For all other scenarios, the payouts, and thus the indexation will only be higher. These can be found in the Appendix.

Clearly, there is a big difference between receiving indexation and not receiving indexation. For participants with a higher monthly payout, indexation has a greater impact. After around 15 years, this impact can even be more than €1000 per month gross. However, even for a low salary, the difference can be around €500 per month after around 10 years. This shows that indexation, even when the monthly payout is not that high, can still have a strong effect. Therefore, a low funding ratio that causes pension funds to not be able to provide indexation can be costly for many participants.

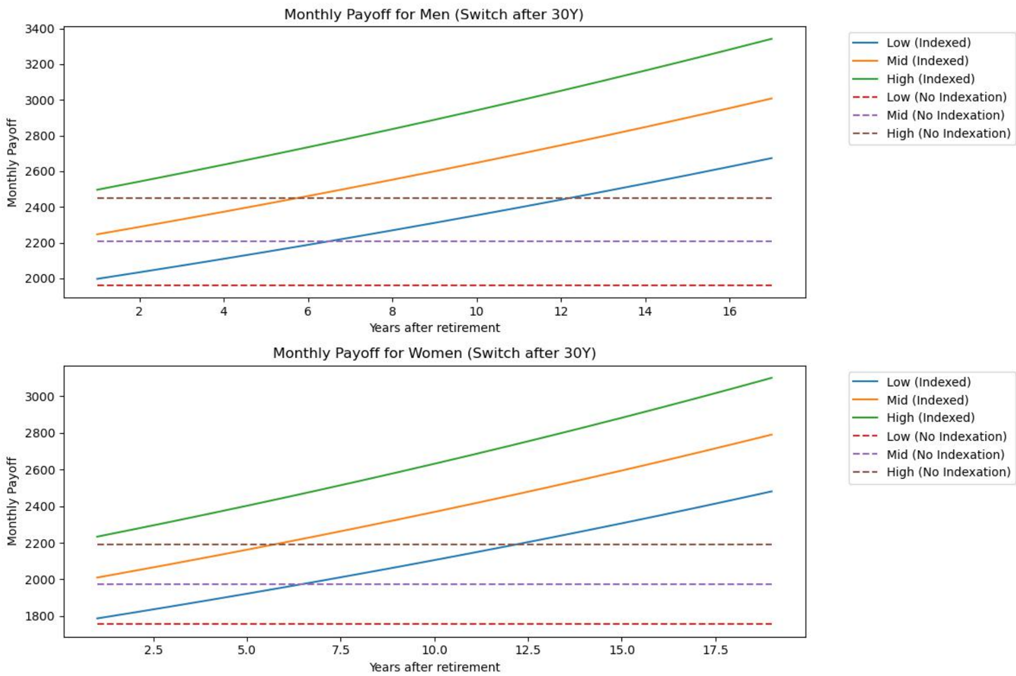


Figure 5.22: Indexation Switch after 30 Years

6

Model Evaluation

This chapter will provide a thorough evaluation of the models used in this research. This evaluation will consider the reliability and accuracy of the models, as well as the behaviour of the models when altering variables.

6.1 Rolling Window Models on Historical Data

The results of the models presented in Section 5 should not be taken for granted without critical evaluation. The accuracy of the Value at Risk models can be tested using a method called backtesting. Backtesting the Value at Risk is done by calculating the expected number of violations and comparing it to the actual number of violations. A violation represents the case where the actual return recorded exceeds the predicted Value at Risk level at that point in time. The expected number of violations can be calculated via $(1 - \alpha) \cdot N$, where α is the confidence level of the VaR and N is the length of the backtesting period. For example, when one calculates the 1-Day 95% VaR, in 5% of the cases one expects a violation. Hence, multiplying $(1 - 0.95)$ times the length of the backtesting period gives the expected number of violations. Comparing this number with the actual counted violations gives an indication of the performance of the method.

An initial thought might be that a low number of violations indicates that the model is performing well. However, this is not necessarily the case. For example, when the expected number of violations is 100, but the actual number of violations turns out to be 30, the model is too conservative. This means that the VaR values returned by the model are higher than necessary, leading to an overestimation of risk. The model then suggests that the potential losses are higher than they actually are. This can cause a company to

over-allocate capital to cover for the potential losses. Setting aside capital unnecessarily means that this capital cannot be used for other purposes.
Similarly, when the expected number of violations is 100, but the actual number of violations is 150, the model is too aggressive. The VaR levels outputted by the model are then lower than necessary, causing an underestimation of risk. This can cause a company to under-allocate their capital, which results in not being able to cover potential losses. The model then gives a false sense of security, which might lead to riskier behaviour by a company.
Therefore, it is important that the number of actual violations is relatively close to the number of expected violations. A statistical test can help determine whether a VaR model is accurate. Zhang and Nadarajah (2017) propose several methods for backtesting VaR models in their paper *A review of backtesting for value at risk* (32). It was decided to run both the Binomial Distribution Test and Kupiec's Proportion of Failure (POF) test. Equation 6.1 gives the test statistic of the Binomial Distribution Test, where H is the number of violations, n is the length of the backtesting period and $\alpha = 1 - confidence\_level$.

$$T = \frac{H - n\alpha}{\sqrt{n\alpha(1-\alpha)}} \tag{6.1}$$

Equation 6.2 below shows the test statistic of Kupiec's POF Test. In Equation 6.4, $I_t(\alpha)$ is an indicator function that is 1 when VaR is violated at time t, and 0 when it is not violated. Therefore, $I(\alpha)$ is the number of observed violations. The $\hat{\alpha}$ in Equation 6.3 is the observed failure rate. Again, α is the confidence level, so for a 95% VaR, α is 0.05. Hence, the number of observed violations is compared with the number of expected violations, just like the Binomial Distribution Test. Both tests assume that the number of VaR violations follows a binomial distribution, where a return can either be a violation or not. They furthermore assume that violations are i.i.d. distributed.

$$POF = 2ln\left[\left(\frac{1-\hat{\alpha}}{1-\alpha}\right)^{n-I(\alpha)} \cdot \left(\frac{\hat{\alpha}}{\alpha}\right)^{I(\alpha)}\right] \tag{6.2}$$

$$\hat{\alpha} = \frac{1}{n} I(\alpha) \tag{6.3}$$

$$I(\alpha) = \sum_{t=1}^{n} I_t(\alpha) \tag{6.4}$$

The null hypothesis (H_0) is *The number of observed violations is statistically consistent with the number of expected violations in the specified VaR model.* For a p-value less than

0.05, this null hypothesis will be rejected, meaning the observed violations deviate strongly from the expected violations. This means that the model either underestimates or overestimates the risk. Similarly, when the p-value is greater than 0.05, the null hypothesis will not be rejected, and the claim that the observed violations are statistically consistent with the expected violations cannot be rejected.

The Expected Shortfall will also be backtested. This is done by taking the average of the losses where the VaR is exceeded and comparing it to the average of the rolling 1-Day ES. The closer the values are, the more likely it is that the results from the ES model can be trusted.

6.1.1 Results

The VaR models performed on the historical data give insight into the level of risk of a historical portfolio. In this section, the results of the 1-Day VaR models with 95% and 99% confidence level will be presented. The visualisations for the 1-Day 99% confidence level VaR models will be provided in the Appendix. Furthermore, the backtesting will only be done on benchmark portfolios, not on the risky and conservative portfolios, as the models operate in the same way for each risk level.

Table 6.1 and Figure 6.1 on the next page show the results of backtesting the Variance-Covariance VaR model. Some key observations can be made. First, it is clear that the observed and expected violations are relatively close to each other. This is reflected in the p-values of the Binomial Test and Kupiec's POF test as well, as both are above 0.05. Therefore, the null hypothesis *The number of observed violations is statistically consistent with the number of expected violations in the specified VaR model* cannot be rejected and it can be assumed that the VaR model behaves as desired.
However, looking at Figure 6.1, it can be seen that while the number of violations is looking good, the magnitude of the violations tells a different story. From the graph, it can be seen that the model does not pick up sudden volatility spikes like the Covid crisis around 2020. The model detects that there is a period of high losses, but the spike in the VaR level is not nearly as high as the actual losses. Furthermore, sudden individual large losses are not picked up at all. Hence, while the statistical tests claim that the model is accurate, the visualisations show that the magnitude of the losses is not reflected in the model. Expected Shortfall should account for these losses behind the VaR. In Table 6.1, it can be seen that the average of the losses where the VaR is exceeded for a 95% VaR is 0.0161. It can be

seen that the estimated ES, the average of the rolling 1-Day ES visible in Figure 5.2(b) is 0.0169, which is in strong alignment with the average violation loss. For 99% VaR, the average violation loss and the estimated ES are even close to each other. Therefore, it is safe to say that the ES model can be interpreted as to behave as desired.

Metric	95% VaR	99% VaR
Backtesting Period (days)	2767	2767
Observed Violations	151	34
Expected Violations	138.35	27.67
Binomial p-value	0.2698	0.2265
Kupiec's p-value	0.2766	0.2429
Average Violation Loss	0.0161	0.0268
Estimated ES	0.0169	0.0274

Table 6.1: Backtesting Results for 95% and 99% VaR-Cov Method

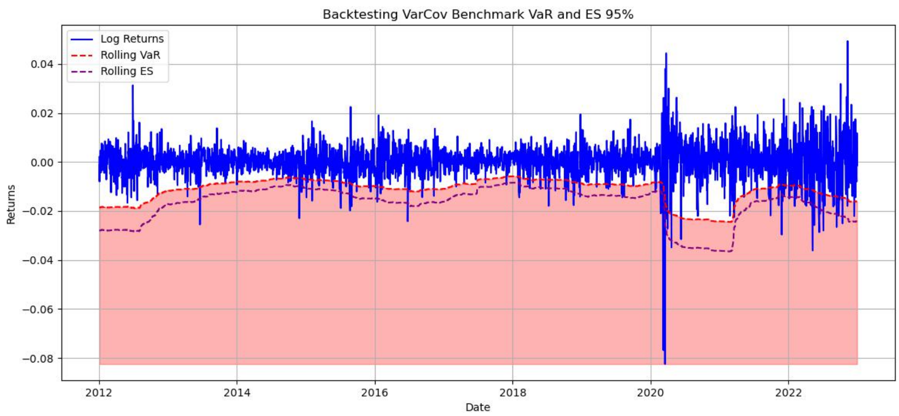


Figure 6.1: Backtesting Variance-Covariance method 95%

Table 6.2 and Figure 6.2 on the next page show the results of backtesting the Historical Simulation model. From Table 6.2 it is clear that there are more observed violations than expected violations. Therefore, the p-values of both the Binomial Test and Koupiec's POF Test are very small, and thus the null hypothesis is rejected. Hence, the statistical tests claim that the Historical Simulation model is not performing as desired. This claim is supported by Figure 6.2, where it is clearly observable that the model does not detect any large losses. Similarly to the Variance Covariance method, individual large losses are not represented by the VaR model. The Historical Simulation method is even worse at picking up longer periods of large losses. The Covid spike is only slightly visible in the model, but not even near the actual observed losses. It can be seen that the Expected Shortfall is barely greater than the VaR, which would suggest that even the Expected Shortfall does not account for these large losses. The average violation loss for 95% VaR is 0.0148, while the estimated ES is 0.0133. Similarly to the VaR, the observed results are greater than the estimated results. For VaR, the observed violations were greater than the expected violations; for ES, the average observed violation loss is larger than the estimated ES. This contributes to the claim of the statistical tests that the model is too conservative. Since both the VaR and the ES are not deemed accurate, it would be recommended not to take the results of the Historical Simulation model for granted.

Metric	**95% VaR**	**99% VaR**
Backtesting Period	2767	2767
Observed Violations	180	67
Expected Violations	138.35	27.67
Two-sided p-value	0.000280	5.7065e-14
Kupiec's p-value	0.000503	2.0594e-10
Average Violation Loss	0.0148	0.0200
Estimated ES	0.0133	0.0183

Table 6.2: Backtesting Results for 95% and 99% VaR Historical Simulation

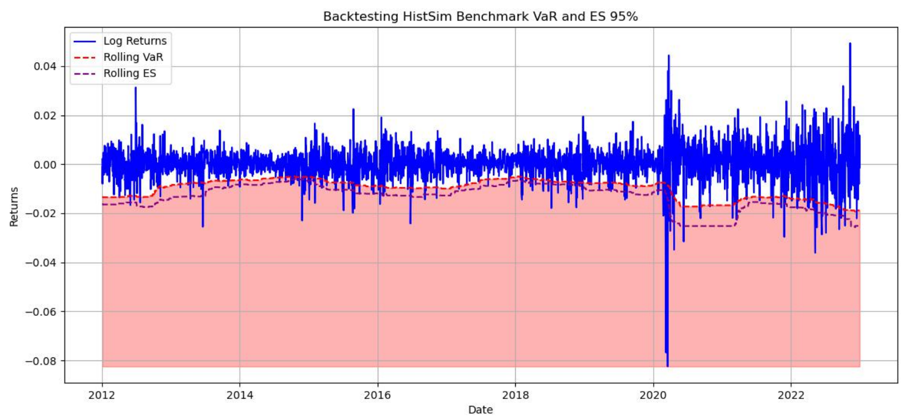


Figure 6.2: Backtesting Historical Simulation method 95%

Lastly, the GARCH(1,1) VaR model is being backtested. In Table 6.3 below, it can be seen that the observed violations are lower than the expected violations. Since the observed violations are much lower than the expected violations, the p-value of both tests is very small, and therefore the null hypothesis is rejected. The tests claim that this model is not accurate. In Figure 6.3 on the next page, it can be seen that the VaR follows the loss pattern very closely. The statistical tests suggest that the GARCH(1,1) model overestimates the risk. The Estimated ES is for both 95% and 99% VaR higher than the Average Violation Loss, which is in line with the backtesting of the VaR.
Despite there being fewer violations than expected, this is the only model that seems to capture large volatility spikes. GARCH(1,1) is known to be able to capture volatility clustering, which is proven by these results. Therefore, it would be recommended to use this model to estimate potential losses in both crisis and non-crisis periods. However, since the model overestimates risk, it is possible that the model will perform poorly on new data.

Metric	**95% VaR**	**99% VaR**
Backtesting Period	2767	2767
Observed Violations	76	10
Expected Violations	138.35	27.67
Two-sided p-value	5.3709e-08	0.000735
Kupiec's p-value	3.1140e-09	0.000102
Average Violation Loss	0.0188	0.0323
Estimated ES	0.0213	0.0345

Table 6.3: Backtesting Results for 95% and 99% VaR Univariate GARCH(1,1)

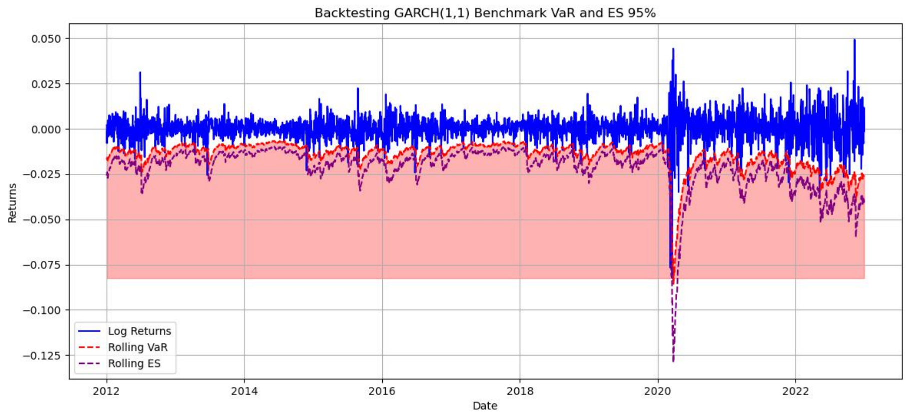


Figure 6.3: Backtesting GARCH(1,1) method 95%

6.2 Monte Carlo Simulation

The Monte Carlo simulation provides insight into the VaR over a longer period of time. In Section 5.3, the benchmark, risky, and conservative portfolios are compared. The question is whether these models provide useful insights or if the results simply reflect what was put into the model. Both asset allocation and volatility contribute to the risk levels of the different portfolios. In the constant volatility model, the only factor that influences risk is asset allocation. Obviously, constant volatility is not a realistic assumption, so the results of this model are not in line with the real world. However, it will provide information on the impact of asset allocation, as this is the only factor that influences risk. In Figure 5.8, it can be seen that the VaR of the risky portfolio is in fact the highest. This verifies that the portfolio that was constructed as risky is indeed the most risky portfolio. Furthermore, it verifies that the asset allocation has a large impact on the risk level, since all other things are the same across the three different portfolios. However, it does not give any new insights. Riskier allocations lead to higher VaR is quite a trivial conclusion. The same holds for the fact that the VaR increases as the time horizon gets longer.

Figure 5.14 demonstrates how the model responds when spikes are added to the simulation. As expected, the VaR increases directly after a spike appears. On the one hand, this suggests that the model is working correctly. However, the results of the model are quite predictable. Implementing spikes at different points in time results in higher VaR values from those same time points onwards. However, it is impossible to predict when a volatility spike will actually occur. Modelling these crisis events as if they can be predicted leads to

results based on guessing rather than informed decisions. It is clear from the models that a crisis event will result in a higher chance of losing more money, but since the occurrence of the events is unpredictable, it is almost impossible to arm against it.
An interesting result is that Figure 5.15 suggests that after a crisis event, it will take more than 40 years for a portfolio to recover. This suggests that the long-term portfolios of younger generations are not long enough to mitigate the risks that arise when investing on a more risky basis. In contrast to the results presented earlier, this result was not expected based on the input of the model. Moreover, the result goes against the general consensus that younger generations should invest on a more risky basis.

In the end, it can be concluded that the results of the Monte Carlo simulation can be interpreted as accurate but influenced by the manually selected input they were given, except for the long-term distinguishing impact of the VaR spikes. Apart from this notion, these simulations will not provide any spectacular new insights, but rather confirm that the model and assumptions are correctly implemented. This is valuable information, especially for the real world simulation, in which a very similar model is used to simulate salaries in the new pension system. Then it can be assumed that this model is also implemented correctly.

6.3 Practical Scenario Analysis

The Practical Scenario Analysis builds on the Monte Carlo simulations to estimate monthly pension payouts for individuals with different salary and indexation levels. For this investigation, a starting salary that is the same for each participant is chosen. In practice, most participants have a distinct starting salary based on education, job industry, and job level. Similarly, the salary growth rates are selected based on an estimated guess. It was chosen to work with different levels of growth, but the increasing and decreasing pattern of the rates is the same for all participants. In reality, these patterns are likely to differ between industries. Lastly, it was chosen to use a static indexation level of 1.84% for all future years. As indexation depends on interest rates, which are likely to change over time, this assumption is not realistic. These assumptions affect the output of the model, meaning the results cannot be directly applied to individual cases without adjustments.

However, even with these artificial input values, the model seems to be able to predict output values that are reasonable in this context. The yearly contribution matches possible real-life yearly contributions for participants, meaning the parameters of the model are chosen well enough to resemble real-life scenarios. This would suggest that when participants know their starting salary and can estimate their salary growth based on their respective industry, the model should be able to provide insightful results. Of course, the total contribution in the new system is calculated using the mean of the simulated paths. This means that on average, a participant contributes that amount of money to their pension. However, this amount can be much higher or lower, depending on the behaviour of the financial market. Therefore, a simulation was done with the lowest portfolio values as well. However, the occurrence of a financial crisis was not implemented.

7

Conclusion

The goal of this research is to analyse the effects of the Dutch pension reformation on different generations. This was achieved by simulating pension portfolios over different time horizons to investigate the risk levels associated with long-term and short-term investments. Complementary, a practical scenario analysis was completed to examine the impact of switching from the old system to the new system at different points in one's career. Together, these models contribute to answering the research question proposed in Section 1:

How does the Dutch pension reform affect investment risk and retirement outcomes across generations?

Pension portfolios were constructed based on assets and allocations used by actual pension funds in the Netherlands. Different allocations were chosen to mimic benchmark, risky, and conservative portfolios. Historical data analysis showed that out of the three proposed Value at Risk models, the Variance-Covariance method is the most reliable method as the expected violations and the observed violations are in alignment with each other. The model showed that during financial turmoil like the Covid-19 crisis, the VaR increases. Furthermore, it showed that the portfolios constructed as risky indeed turned out to have the highest VaR, demonstrating that the asset allocations are chosen in a way that allows for using them as representable for younger and older generations. Although Historical Simulation and GARCH(1,1) demonstrate similar results, they are considered less accurate models as they underestimate and overestimate risk, respectively.

Monte Carlo simulations demonstrated that riskier investments lead to a broader spread of possible portfolio outcomes. Risky portfolios can generate higher gains or higher losses. These high gains and large losses mainly average each other out, causing the mean of the risky portfolio to be only slightly higher than the mean of the benchmark portfolio in the case of constant volatility. In the case of the GARCH(1,1) volatilities, the mean of the benchmark portfolio is sometimes even higher. Crucially, the VaR of the risky portfolio is higher than that of the benchmark portfolio at all times. The models show that the VaR increases as the time horizon increases. This suggests that taking on a lot more risk will not guarantee a higher payout, even after 40 years. Both the constant volatility model and the GARCH(1,1) volatilities model show this result.

Introducing spikes to the simulations results in a direct increase in VaR right after the spike. It is interesting to note that, when a spike occurs in a risky portfolio, after 40 years, the impact of this spike is still visible in the VaR level. The Dutch government suggested that younger generations can invest more risky as they have time to compensate for potential losses along the way. However, this model suggests that the 40 years an individual spends in the new system are not enough to recover from large volatility spikes. The impact of the spike returned to its original level only after 100 years. For more conservative asset allocations, the impact of the spike is less intense and, therefore, it also takes less time to recover.

The Practical Scenario Analysis shows that younger generations, who will spend their whole life in the new system, are able to collect the highest total contribution over their working life. Interestingly, older generations who have spent their entire life in the old system seem to be in the second-best scenario. The generation that has spent 30 years in the old system and now has to spend the last 10 years of their life in the new system, collects the lowest total contribution over their lifetime. These differences are mainly caused by the change from a variable to a flat premium, as explained in Section 5.3.4. To obtain the total contribution in the new pension system, each year the mean of the portfolio paths was taken. The fact that the total contribution in the new system is so much higher than the total contribution in the other scenarios suggests that on average, investing contribution in the financial market does not impact the outcome overall that much. The missed contribution resulting from the change from variable to the flat premium is the largest contributing factor to the differences between the chosen generations. This is supported by Figure 5.21(b), where the results are shown when investments perform poorly and yield lower returns. Still, the younger generation obtains the highest pension contribution after 40 years. Therefore, it can be said that the change from variable to flat premium contributes the most to the discrepancies between age cohorts.

In conclusion, investing on a riskier basis increases the Value at Risk over time but does not necessarily increase the payouts. Younger generations should be careful with investing on a high-risk basis as it is not guaranteed that their time horizon is long enough to diminish the effects of potential crisis events. Without any crisis events, younger generations seem to benefit the most from the reformation, as they do not have to switch from a variable to a fixed premium. Generations who spent 30 years in the old system and have had to switch to the new system for the last 10 years of their working career will experience the biggest downsides of the reformation. The generation that has already retired does not face clear disadvantages compared to younger generations. Therefore, the compensation for disadvantaged generations should mainly be provided to age cohorts who spent 20 to 30 years in the old system and had to switch afterwards.

8

Discussion

8.1 Limitations

In Section 6, the performance of the models was evaluated. Beyond the limitations of these models, this research has several other limitations. First, pension portfolios were constructed using only 17 assets. Although these assets represent the main asset classes utilised by pension funds, in reality, pension funds invest in a much broader array of assets across numerous countries and currencies. This study only included Dutch and American stocks. Furthermore, the portfolio lacked sufficient diversification, being predominantly composed of stocks in the technology sector. Expanding portfolios to include a wider range of stocks and a higher level of diversification would likely yield more realistic results. Moreover, some of the implemented assets are ETFs, since collecting historical data of actual indexes was difficult without a permit. Using actual indexes or stocks instead of ETFs will make the portfolio more realistic, as ETFs hold multiple underlying assets and are therefore not specific to a certain type of asset.

Furthermore, the Practical Scenario Analysis is performed without stress testing, due to time constraints. Implementing volatility spikes in the Practical Scenario Analysis will likely alter the results of the younger generations as their investments in the new system will likely decrease in value. The total contributions will then likely be much closer to one another. It will be interesting to see whether spending 40 years in the new system will still result in the highest total contribution, or whether spending 40 years in the new system will then be the most beneficial scenario.

Lastly, as mentioned earlier, the risk aversion of the age cohorts is modelled constant over time. This implies that a younger participant, who initially invests more aggressively due to their longer investment horizon, continues to do so for their entire life, including periods of extreme market stress. For future research, it would be interesting to incorporate behavioural responses into the model, allowing risk preferences to evolve in response to market events and increasing age. This would help simulate more realistic investor behaviour and may significantly impact outcomes across generations.

8.2 Future Research

As discussed, the models to assess the risk level are relatively simple models. Both Value at Risk and Expected Shortfall are easy to implement and interpret, which is why they were used in this research. For future research, the implementation of machine learning models such as Random Forest or Gradient Boosting could be used to predict risk levels. In addition, in this research, stress testing is manually implemented. Furthermore, volatility is modelled based on the GARCH(1,1) volatilities of the historical data. An alternative approach could be to implement a Heston Stochastic Volatility model to incorporate random volatility over time.

The conclusion of the Practical Scenario Analysis is that the switch from a variable to a fixed premium contributes the most to the differences between generations. Of course, when one contributes less of their salary to their pension, the monthly salary when still working will be higher, as less money is transferred to the pension portfolio. Therefore, it would be interesting for future research to investigate what happens when we take all this extra salary collected over the working life and invest it in a risk-free rate. This money can then be added to the total contribution and it is likely that the differences between the age cohorts then decrease a lot. It will be interesting to see whether younger generations still profit the most or whether other generations will end up with a higher portfolio.

Lastly, some pension funds in the Netherlands have already switched to the new pension system. In the future, data from these funds might be able to help funds that still need to make the switch. However, since these funds have just recently transferred, the available data is very short-term. Therefore, it is not likely that they will be able to provide useful information before the next firms are transferring. In a few years, research could be conducted on the impact on participants of the new pension system.

References

[1] Mercer. **Mercer CFA Institute Global Pension Index**. Technical report, Mercer, CFA Institute, Monash University, 2024. 6

[2] L. Metselaar, P. Zwaneveld, and C. van Ewijk. **Reforming Occupational Pensions in the Netherlands: Contract and Intergenerational Aspects**. *De Economist*, 2022. 7, 11, 15

[3] B. Van Popta and O Steenbeek. **Transition to a new pension contract in the Netherlands: Lessons from abroad**. *Netspar*, 2021. 7, 12

[4] Natascha van der Zwan. **Financialisation and the Pension System: Lessons from the United States and the Netherlands**. 2017. 8

[5] E. Westerhout, E. Ponds, and P. Zwaneveld. **Reforming the Dutch pension system to ensure sustainability**. *International Social Security Review*, 2024. 8, 9, 10, 11, 12, 15

[6] G. Alserda, B. Dellaert, L. Swinkels, and F. van der Lecq. **Pension risk preferences: A personalized elicitation method and its impact on asset allocation**. *Netspar*, 2016. 9

[7] F. Verkuijlen. *Measurement and implementation of the risk aversion by pension funds in the new Dutch pension law*. Master's thesis, Tilburg University, 2023. 9

[8] J. D. Rauh, I. Stefanescu, and S. P. Zeldes. **Cost Saving and the Freezing of Corporate Pension Plans**. *NBER National Bureau of Economic Research*, 2020. 10

[9] CBS. **Elderly People**. Technical report, Central Bureau Statistics, 2024. 10

[10] Northern Trust and True Partner Capital. **Tackling Investment Risks in the Dutch Pension Transition**, 2023. 11

[11] A Jeyachitra, M. Selvam, and J. Gayathri. **Portfolio Risk and Return Relationship - An Empirical Study**. *Asia-Pacific Business Review*, 2010. 15

[12] H. Markowitz. **Portfolio Selection**. *The Journal of Finance*, 1952. 15

[13] J. A. Bikker and J. de Dreu. **Pension Fund Sophistication and Investment Policy**. *Netspar*, 2009. 16

[14] AXA Investment Managers. **AXA IM Barometer**, 2021. 16, 34

[15] I. Eleftheriadis and I. Anastasovitis. **Can the commodity futures be used as effective diversifiers for pension funds investment portfolios?** *International Conference on Business and Economics - Hellenic Open University*, 2023. 17

[16] A. D. Dixon. **The Rise of Pension Fund Capitalism in Europe: An Unseen Revolution?** *Oxford University*, 2008. 17

[17] J. F. Cocco, F. J. Gomes, and P. J. Maenhout. **Consumption and Portfolio Choice over the Life-Cycle**. *The Review of Financial Studies*, 1998. 18

[18] D. Duffie and J. Pan. **An Overview of Value at Risk**. *Journal of Derivatives*, 1997. 20

[19] P. Artzner, F. Delbaen, E. Jean-Marc, and D. D. Heath. **Coherent Measures of Risk**. *Mathematical Finance*, 1999. 21

[20] A. K. Pradhan and A. K. Tiwari. **Estimating the market risk of clean energy technologies companies using the expected shortfall approach**. *Renewable Energy*, 2021. 22

[21] T. Bollerslev. **Generalized Autoregressive Conditional Heteroskedasticity**. *Journal of Econometrics*, 1986. 25

[22] P. R. Hansen and A.. Lunde. **A forecast comparison of volatility models: does anything beat a GARCH(1,1)**. *Journal of Applied Econometrics*, 2005. 27

[23] J. Laplante, J. Desrochers, and J. Préfontaine. **The GARCH (1, 1) Model As A Risk Predictor For International Portfolios**. *International Business & Economics Research Journal*, 2008. 27

[24] J. N. P. Xiang, S. Velu, and S. Zygiaris. **Monte Carlo Simulation Prediction of Stocks**. *International Conference on Developments in eSystems Engineering (DeSE)*, 2021. 28

[25] M. E. Sonono and H. P. Mashele. **Prediction of Stock Price Movement Using Continuous Time Models**. *Journal Of Mathematical Finance*, 2015. 30

[26] T. Brodd and Djerf A. **Monte Carlo Simulations of Stock Prices**. *School of Electrical Engineering and Computer Science Stockholm*, 2018. 30

[27] Pensioenfonds Zorg & Welzijn. **Jaarbericht Pensioenfonds Zorg en Welzijn 2024**, 2024. 33

[28] N. M. Razali and B. W. Yap. **Power comparisons of Shapiro-Wilk, Kolmogorov-Smirnov, Lilliefors and Anderson-Darling tests**. *Journal of Statistical Modeling and Analytics*, 2011. 37, 38

[29] ABP. **Premiebericht 2025**, 2025. 43, 44

[30] ABP. **Indexatie: verhoging pensioenen ABP**, 2025. 46

[31] Koninklijk Actuarieel Genootschap. **Prognosetafel AG2024**, 2024. 47

[32] Y. Zhang and Nadarajah S. **A review of backtesting for value at risk**. *Communications in Statistics - Theory and Methods*, 2017. 70

Appendix

Investment Category	**Allocation (%)**
Equities and Alternatives	51.7
Listed Equities	21.3
Private Equity	8.8
Listed Real Estate	5.4
Private Real Estate	7.2
Infrastructure	5.6
Insurance	3.2
Other	0.2
Credit	19.7
Corporate Bonds	10.6
Emerging Markets Debt (Local Currency)	4.5
Credit Risk Sharing	2.6
Mortgages	2.0
Fixed Income	25.4
Government Bonds	25.4
Legacy Fixed Income	0.0
Overlay	3.1
Interest Rate Overlay	3.5
Currency Overlay	0.0
Cash	-0.3
Total	100.0

Table 8.1: Investment Portfolio Allocation PFZW (December 31, 2024)

Asset Class	Allocation (%)
Real Estate Investments	10.0
Equities	25.0
Private Equity	5.0
Infrastructure	5.0
Fixed Income – Government Bonds (non-inflation-linked)	15.0
Fixed Income – Inflation-Linked Bonds	5.0
Fixed Income – Mortgages	5.0
Fixed Income – Corporate Bonds/Credits	15.0
Fixed Income – Short-Term Receivables and Liquid Assets	0.0
Hedge Funds	2.5
Commodities	2.5
Other Investments	10.0
Strategic Interest Rate Hedging	50.0

Table 8.2: Possible Asset Allocation based on KPMG insights

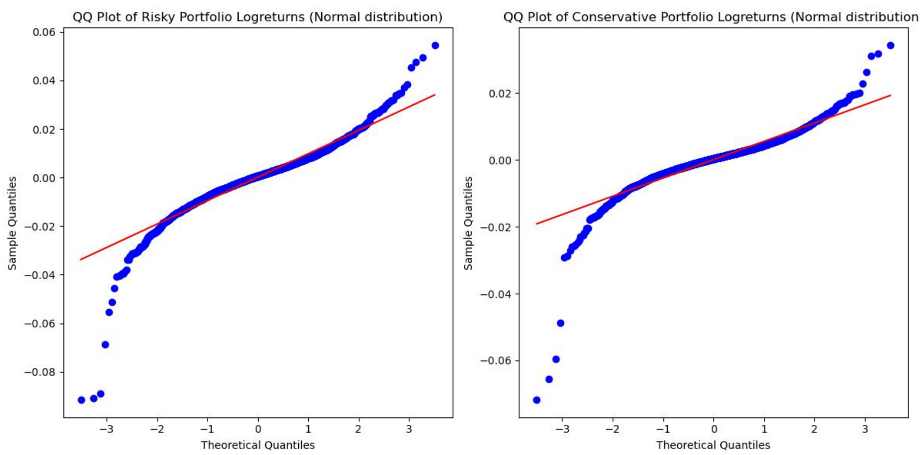
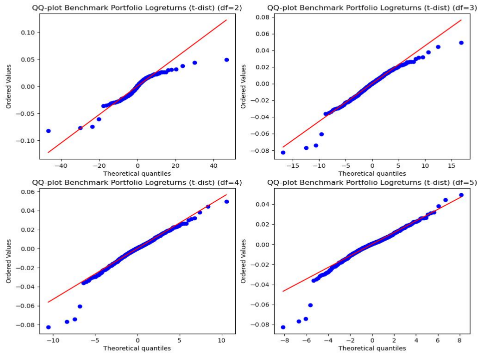
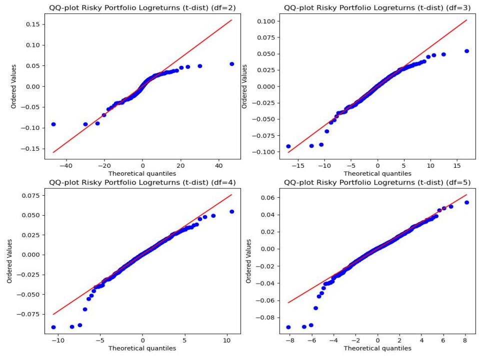


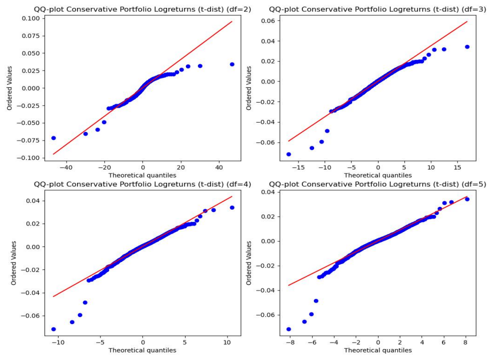
Figure 8.1: Normal QQ-Plots Risky and Conservative



(a)

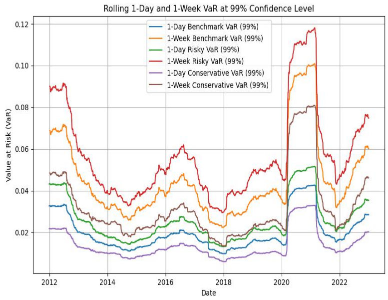


(b)

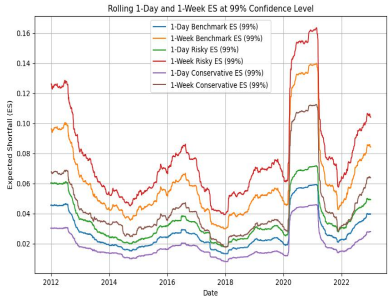


(c)

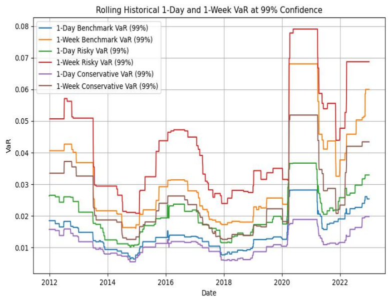
Figure 8.2: (a) Benchmark Portfolio: Different degrees of freedom (t-distribution) (b) Risky Portfolio: Different degrees of freedom (t-distribution) (c) Conservative Portfolio: Different degrees of freedom (t-distribution)



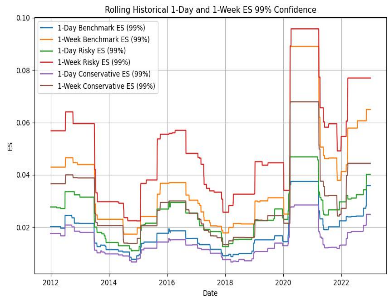
(a)



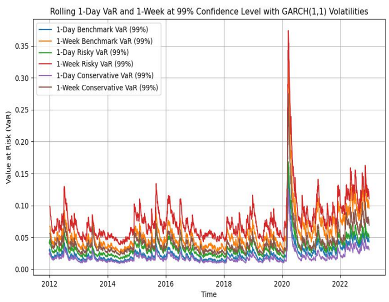
(b)



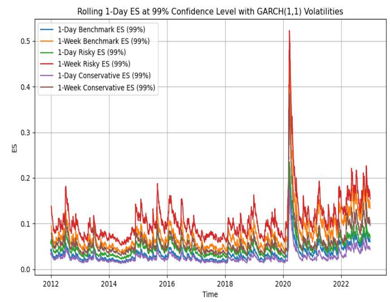
(c)



(d)

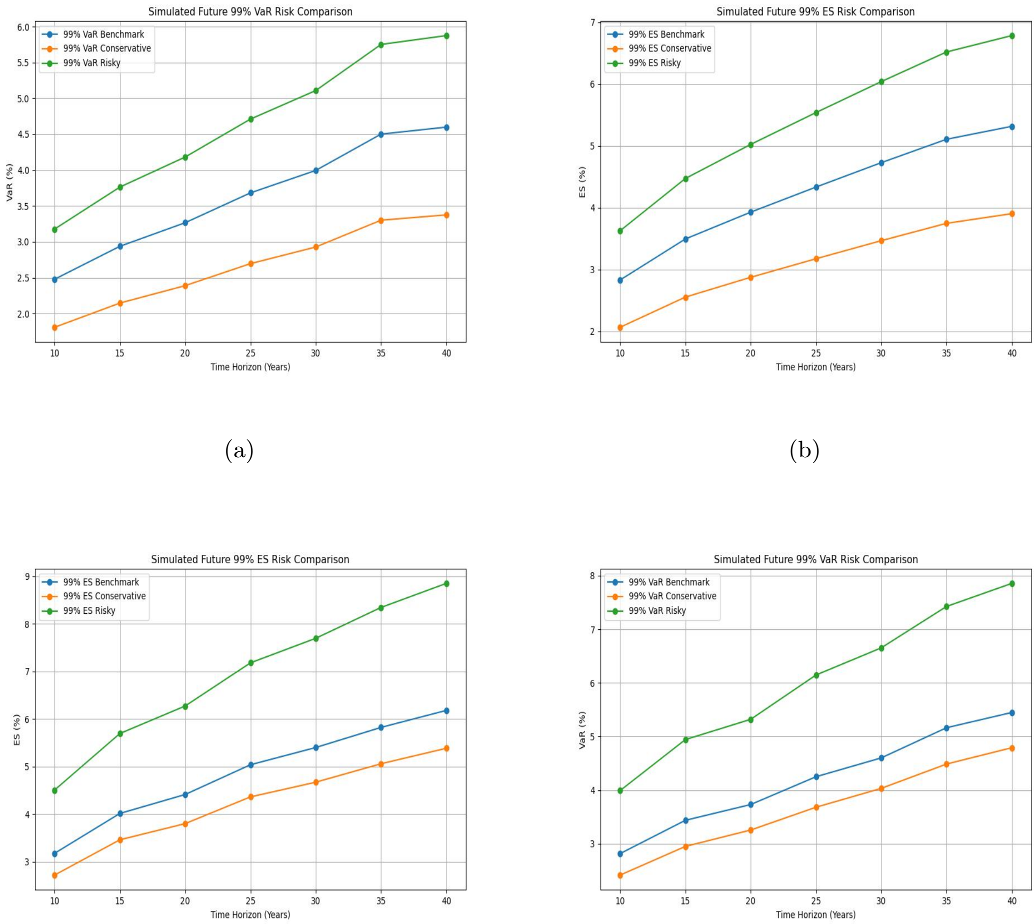


(e)



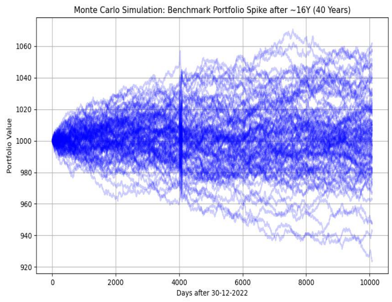
(f)

Figure 8.3: (a) Rolling VaR 99% (b) Rolling ES 99% (c) Rolling Historical VaR 99% (d) Rolling Historical ES 99% (e) Rolling GARCH(1,1) VaR 99% (f) Rolling GARCH(1,1) ES 99%

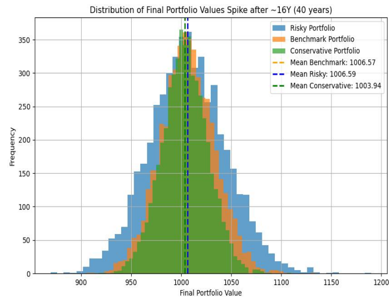


(a) (b) (c) (d)

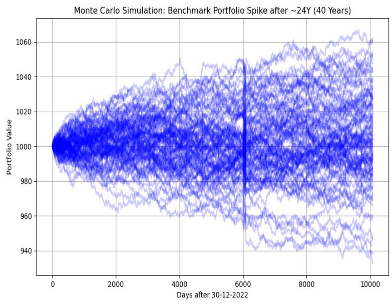
Figure 8.4: (a) Simulated future 99% VaR with Constant Volatility (b) Simulated future 99% ES with Constant Volatility (c) Simulated future 99% VaR with GARCH Volatilities (d) Simulated future 99% ES with GARCH Volatilities



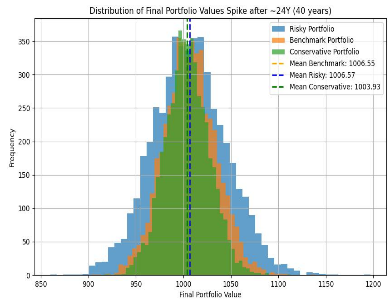
(a)



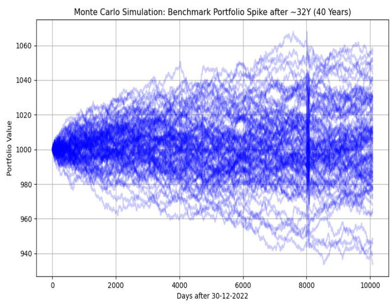
(b)



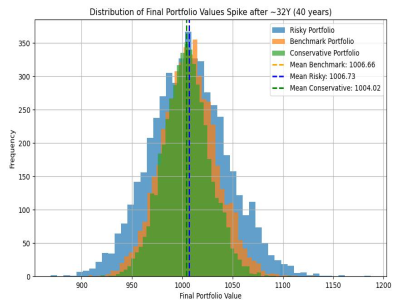
(c)



(d)

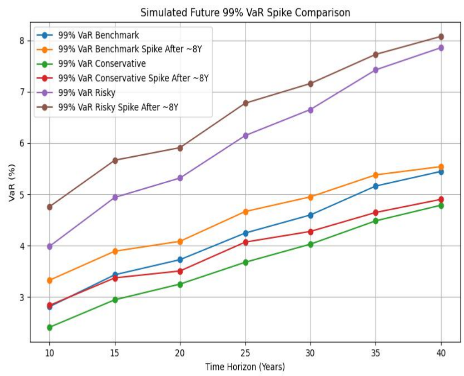


(e)

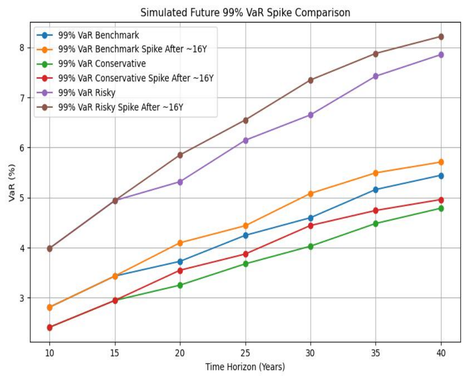


(f)

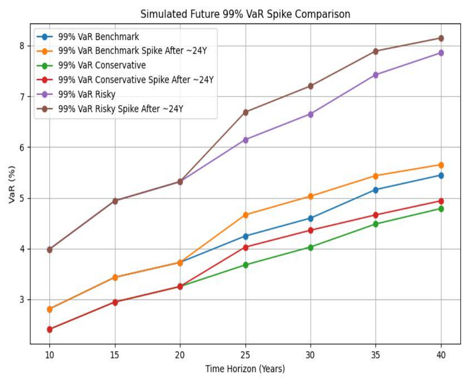
Figure 8.5: Impact Spikes at different Time Points



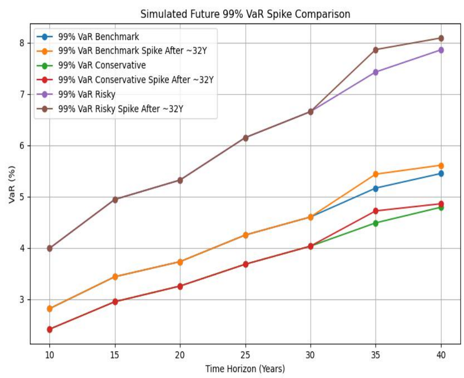
(a)



(b)

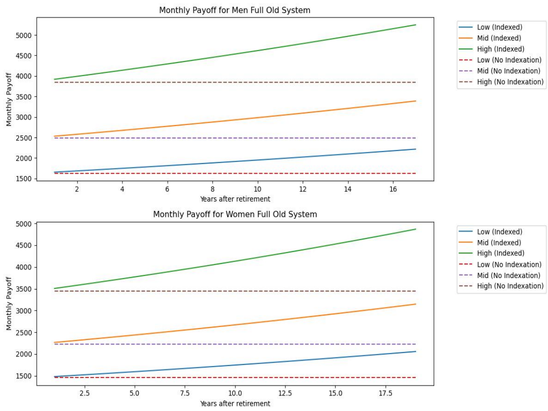


(c)

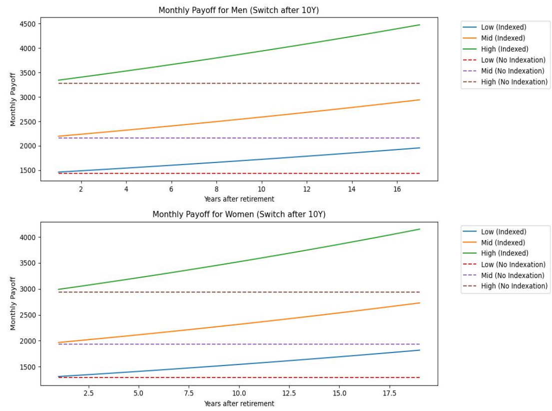


(d)

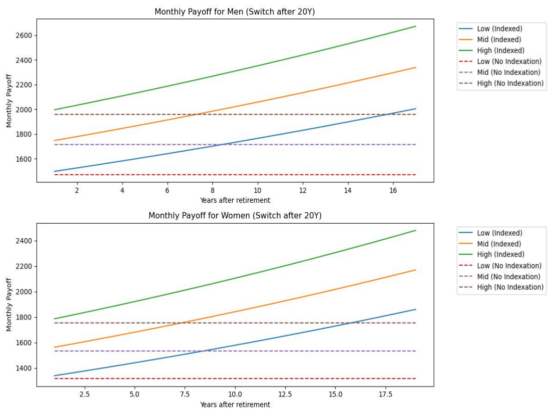
Figure 8.6: (a) Simulated Spike 99% VaR after 8 Years (b) Simulated Spike 99% VaR after 16 Years (c) Simulated Spike 99% VaR after 24 Years (d) Simulated Spike 99% VaR after 32 Years



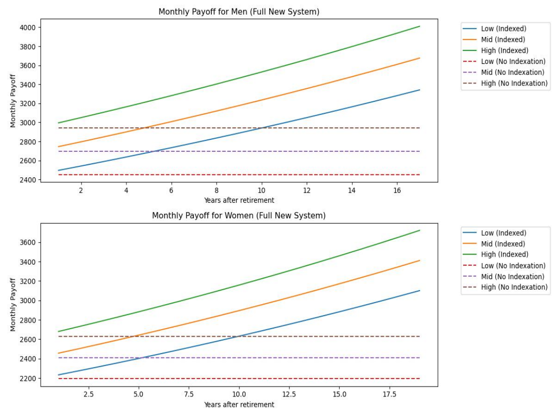
(a)



(b)



(c)



(d)

Figure 8.7: (a) Indexation Full Old System (b) Indexation Switch after 10Y (c) Indexation Switch after 20Y (d) Indexation Full New System

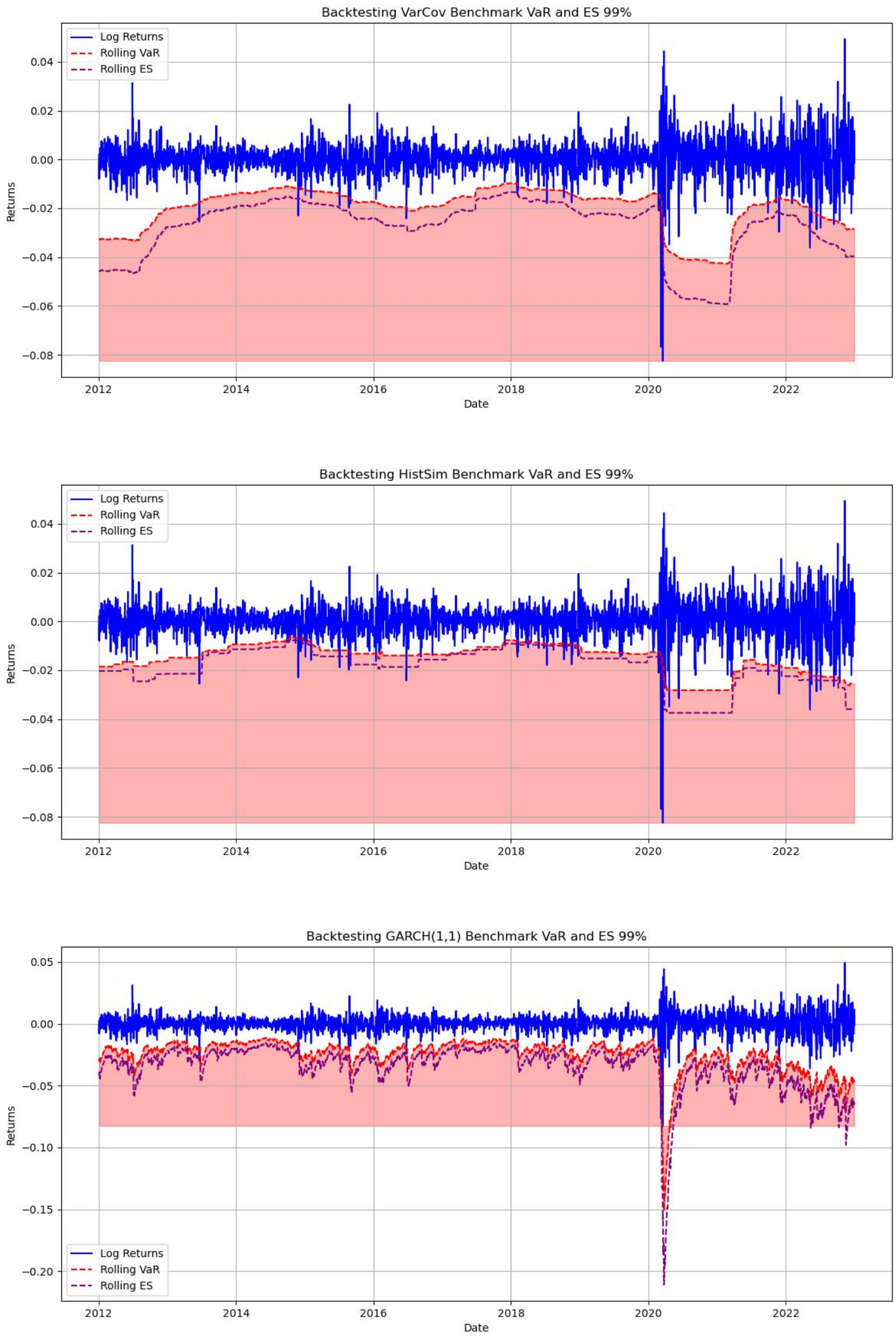


Figure 8.8: Backtesting VaR models 99%