UNIVERSITY OF TWENTE.

Algorithmic investing within the philosophy of the Basis Zero experiment

An explorative study on the possibilities of implementing algorithmic investing within the scope of the Basis Zero experiment.

Public version

A thesis submitted in partial fulfillment of the requirements for the degree of

Master of Science in
Industrial Engineering and Management

Specialization:
Financial Engineering and Management

Department Industrial Engineering and Business Information Systems

February 2017

Author:
M.L. de Wit

University supervisors:
Dr. B. Roorda
Dr. R.A.M.G. Joosten

External Supervisor:
H. Terpoorten MSc

Management summary

The goal of the "Basis Zero: 10-10-10" experiment is the development of a new basic pension product that needs to meet three requirements. The concept should be explicable within 10 lines of text, with maximum asset management costs of 10 basis points a year (including transaction costs) and its execution costs should not be higher than 10 euro a year. In summary, a simple and inexpensive product that still offers appropriate quality features.

In the experiment phase three schools of thought underpin the development of the product. These schools of thought, or scenarios were based on the distance to current practices at APG – ranging from Heerlen to Groningen and Rottumerplaat. The focus of this thesis is on the scenario that is farthest removed from current practices: Rottumerplaat. On the desert island a complete new pension provider is built from scratch, not using any current systems or infrastructure of APG. As Rottumerplaat is uninhabited every action needs to be automated including asset management. The objective of this thesis is the development of an algorithm to automate the asset management process. The research question is therefore:

> *Is it possible for APG to create and implement an algorithm that automatically makes investment decisions - given a certain investment universe and the Basis Zero philosophy - with a pension objective that achieves adequate performance?*

Treleaven (2013) describes the trading process in five steps: data access/cleaning, pre-trade analysis, trading signal generation, trade execution and post-trade analysis. Our contribution focuses on the composition of portfolios over time where aspects of the pre-trade analysis as well as the trading signal generation will be included.

In line with the philosophy of Basis Zero the adequacy of the algorithms results will be based on a replacement rate. The (absolute) replacement rate generally refers to an indicator showing the level of pension income as a percentage of earnings before retirement and measures the extent to which a pension system enables continuing a certain standard of living. The minimum acceptable replacement rate level is set at 60% but the objective of the algorithm will be a 70% replacement rate. There is a widely supported belief, a 'rule-of-thumb', that 70% of final earnings will provide a good pension.

The original idea of the Basis Zero experiment was to arrange exchange connectivity to Euronext and be able to invest in every product traded on this exchange. For the purpose of this thesis the investment universe is limited to seven asset classes and ten liquid products on several corresponding indices. The tenth instrument is a(n approximated) 'risk-free' or cash instrument to which funds can be allocated. The asset classes included in the scope are equities in developed and emerging markets, real estate, commodities, credits, treasuries and cash. The corresponding equity indices are: MSCI North America, Europe, Asia Pacific ex Japan, Japan and MSCI Emerging Markets. The other indices are: FTSE EPRA/NAREIT Developed (Real Estate), S&P GSCI (Commodities), Barclays US Corporates: Investment Grade (Credits) and Barclays Euro Aggregated Treasury (Treasuries). The J.P. Morgan Euro Cash 1-month instrument is used as our cash option. The 10 instruments are chosen based on liquidity, so large volumes do not limit our trading possibilities.

From the selected indices historical time series are used as dataset. The set includes monthly closing values from the indices ranging from January 31st 2002 to April 29th 2016. This time frame is chosen based on the length of the period and actuality of the data.

The selected portfolio construction models can be divided in rule-based models using heuristics, and optimisation-based models using an objective function. We will use mean variance and minimum variance optimisation, momentum and mean reversion on the return rate, a 1/N model and a lifecycle. Some are rather simple, others more sophisticated.

By testing portfolio construction models on financial market data the performance of these models can be assessed in retrospect. Due to the pension horizon our available dataset is too short making it necessary to use simulated time series. A geometric Brownian motion and a basic historical simulation approach generate these series. Two extensions of the historical simulation approach are used. The first is an exponentially weighted moving average where recent observations get more weight. The second approach gives more weight to scenarios with a low interest rate. The historical data set will additionally be used to test the models on a short, but real, dataset.

The results of the simulations show that the mean variance model achieves the best performance across all simulation methods. The mean and median replacement rates exceed 70% in all historical simulations but underperform in the geometric Brownian motion simulation. It beats the performance of APG on these measures almost always. At the same time the dispersion of the results is large which is an important downside as it comes to pension investing. The Value at Risk is many times below satisfactory levels where APG shows decent numbers. By giving more weight to more recent scenarios or to scenarios with a low interest rate both mean/median results and the Values at Risk increase substantially while the variance numbers stay large. Nevertheless it is still APG that shows the preferred results.

According to our results the conclusion can be drawn that it is possible to develop an algorithm that automatically invests funds at low costs. The limit of ten 10 basis points should not be a problem. The best performing/mean variance model shows adequate performance in the majority of the results. However there is a substantial risk that an insufficient result will be obtained. The results of APG are on average a bit worse, but they are more stable making setbacks less severe. A conservative attitude is desirable when making decisions about (other) people's pension income. For participants or investors with a low risk appetite our solution is not yet an alternative to APG's current practice. However, a participant that is more risk tolerant could consider this alternative option.

As the number *10* is the foundation of the experiment we summarize our solution, for the professional reader, in *10* lines:

We created an algorithm that automatically invests pension capital with costs that stay under 10 basis points per year. When the participant retires, the total capital position is transferred to an annuity from which our performance can be measured by the replacement rate: the ratio between earnings before and after retirement. A limited investment universe is used including 10 indices on the following asset classes: equities[1], real estate, commodities, credits/treasuries and cash. The solution is based on a mean variance strategy that is tested on historical and simulated data. Our datasets are simulated by a geometric Brownian motion, general historical simulation and extensions based on an exponential weighted moving average where recent or low interest rate scenarios get more weight. Compared to our benchmark APG our results score well on mean and median values but fail on the dispersion of the results where low Values at Risk are the result. As the risk on insufficient results is substantial APG's own performance is still preferred due to the decent, stable results.

[1] Equities from both developed as well as emerging markets.

Table of Contents

List of figures

List of tables

Preface

To complete my master Industrial Engineering and Management I had the opportunity to write my master thesis at APG. I tried to explore the possibilities in the field of algorithmic pension investing as part of a new pension product originated from the Basis Zero experiment.

To make this project to a success I want to express my gratitude to a number of people. I want to thank my supervisor from APG, Hidde Terpoorten, for his help, feedback and positive views enabling me to complete this assignment. In addition to him also my colleagues of Group Risk and Compliance and my fellow interns (especially Jorgo Goossens) that were seated at the 17^{th} floor of the Symphony building. I also want to thank Eric de Rouw from Asset Management and Alwin Oerlemans, Cateautje Hijmans van den Bergh, Joep Beukers, Loes Frehen and Chris Veerkamp for their help, suggestions and feedback. I hope this report is helpful to you.

I would not have been able to write this thesis without the support of Berend Roorda. I experienced a very positive cooperation as you took plenty of time to help me during this assignment. I want to thank Reinoud Joosten for his comments that improved my final version. Most of all I like to thank them for their courses the past years that I really enjoyed.

Finally I would like to thank my mother and brother for their support during my studies.

Maarten de Wit
Hoogkarspel, February 2017

1. Research design

The research for this master thesis has been conducted at the All Pensions Group, APG. This report presents a proof of concept on the introduction of algorithmic investing within the scope of the Basis Zero experiment at APG.

In this chapter we will explain the research design. Verschuren and Doorewaard (2010) argue that the design of research activities involves two separate sets of activities. The first set involves determining everything that should be achieved through the research project. This is called the conceptual design of the research project. The second part, the technical research design, concerns the activities to realize this.

The chapter starts with the conceptual design and more specifically with a short introduction to the context of the assignment. In Section 1.2 we will describe the research objective after which the research questions are covered in Section 1.3. We will explain the technical design in Section 1.4. This section includes the research strategy, the method of data collection and the outline of this thesis.

1.1. Context of the study

The company APG

Financial services provider APG offers services for pension funds such as pension administration, pension communication, asset management and executive consultancy and offers individuals supplementary products on the pensions market. APG focuses on products in the second pension pillar and performs these activities on behalf of clients and their participants in the sectors of education, government, construction, cleaning and glass cleaning, housing associations, energy and utility companies, sheltered employment and medical specialists. APG manages about € 443 billion (August 2016) in pension assets for these sectors. APG provides pension for one in five families in the Netherlands, which is equal to approximately 4.5 million participants. APG has offices in Heerlen, Amsterdam, Brussels, New York and Hong Kong (APG, 2016).

APG Asset Management is an investor focused on pension investing. APG tries to leverage experience, expertise and innovative power to invest the pension assets of its clients and their participants cost effectively. APG wants to contribute to a qualitative, affordable pension for the clients' participants by achieving long-term stable returns while taking responsible risks. Last but not least APG wants to contribute to a sustainable world that is integral part of the investment process.

This research on algorithmic investing is not executed at APG Asset Management but at APG Group level, in particular at the department Group Risk and Compliance (GRC). The most important tasks of GRC are the development of risk management and compliance frameworks, giving independent advice and challenging the management on the implementation of these frameworks. Monitoring and reporting on the application of the frameworks also belongs to the responsibilities of GRC.

Innovation activities within APG

Based on the rapidly changing world in general, and the pension world in particular, APG believes that some of the products in their current form and context will lose their competitive edge and will be no longer relevant in the future. You might think of developments in the labour market (e.g. in robotics), in the field of health (e.g. changing life expectancies) and technology (e.g. artificial intelligence), but also in terms of changes in the current pension system. To remain a relevant player now and in the future APG works on innovation.

APG wants to make itself future-proof by responding to (expected) developments in technology, the labour market and society. This will be done by facilitating the development of innovative products within the innovation organization of APG. Expected results of innovation within APG are (radical) innovative products at the edge of the business of APG.

When it comes to innovation APG want to develop its current practice as well as create entirely new products and services. Ideas need exploration and development to grow to the so-called 'experiment phase'. When an idea is sufficiently developed, and enters the experiment phase, the quest for sponsors, team members, internal knowledge, external partners and new technologies will be started. After passing the 'delivery phase' the experiment will be absorbed within the organization or continues as an independent unit with funding from the innovation fund.

Problem description

The innovation activities within APG have led to multiple experiments to improve current or create new products and processes. One of these experiments is "Basis Zero: 10-10-10". Goal of this experiment is the development of a new basic pension product that needs to meet the three requirements that are presented in Figure 1. The concept should be explicable within 10 lines[2], with maximum asset management costs of 10 basis points a year and its execution costs should not be higher than 10 euro a year. In summary, a simple and inexpensive product that still offers appropriate quality features.

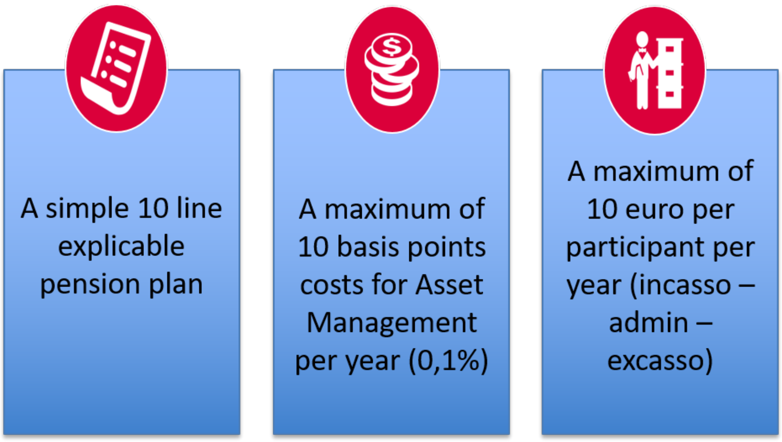


Figure 1: the Basis Zero principles.

Within the original experiment there were three schools of thought that were used to further develop the product. These scenarios are described by the following pseudonyms: 'Heerlen', 'Groningen' and 'Rottumerplaat'. The 'Heerlen' school of thought refers to execution of the new product by using current APG knowledge and infrastructure. A solution that remains close to home.

Groningen is already a bit more distant from current practices. In this scenario a new external basis will be built that will be used to create a new basic pension provider. The scenario's 'Heerlen' and 'Rottumerplaat' both describe the execution of asset management as well as the execution/administration problem of the experiment (i.e. block 2 and 3 in Figure 1). The Groningen scenario only describes the administration part but will not develop a solution on asset management.

Finally 'Rottumerplaat' represents a scenario that is furthest removed from the current situation. Here a whole new pension provider will be created: a pension provider 2.0. On the uninhabited island the new pension provider is built from scratch, not using any current APG systems or infrastructure. As Rottumerplaat is uninhabited every action needs to be automated. In this thesis the focus will be on the latter scenario.

The idea is a result of the desire to create a simple and inexpensive pension product. Simplicity is reflected in a collective product without individual choices that should be easy to understand. The inexpensive product is the result of automated administration and digital communication by using blockchain technology. However, the administration and communication part of the experiment are beyond the scope of this research. The focus will be on the last part of the proposition: automated asset management by using algorithms.

---

[2] The explanation of the concept within 10 lines is included in Appendix 1.

Contributions to APG

The bigger picture of the Basis Zero experiment is shown in Figure 2. The figure shows the two workstreams on asset management: workstream 1 on the Heerlen scenario and workstream 2 on the Rottumerplaat scenario.
In this report the focus is on the 'radical outside solution' without a clear link to current asset management practices. Within the Asset Management department itself, another part of the experiment is executed by examining if it is possible to invest at a maximum of 10 basis points by using current practices; the 'Heerlen' school of thought.
In other parts of the experiment a team is engaged in the development of a pension administration based on blockchain technology. This should give an answer on the question if 10 Euro administration costs are feasible.

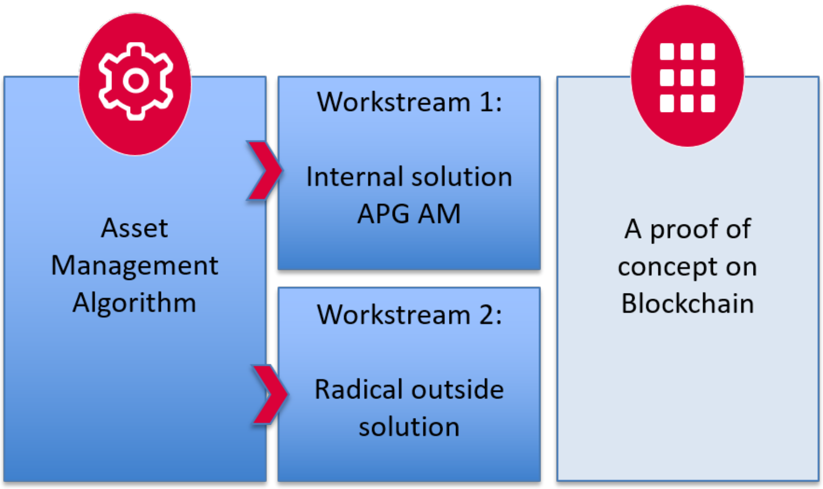


Figure 2: phase 2 of the Basis Zero experiment will focus on asset management algorithms and blockchain technology. The focus in this report is on workstream 2: the radical outside solution.

1.2. Research objective

The objective of this thesis is the development of an algorithm. In line with the Rottumerplaat philosophy the algorithm should be able to invest independently and automatically incoming contributions or divest funds when distributions exceed contributions - while keeping operating, transaction and management costs low. These investment decisions need to be optimised on a pension objective and should be executable within a certain investment universe.
As investments need to be made for participants' pensions, the decisions the algorithm makes need to be optimised on a pension objective. Furthermore, the investments have to be executable within a certain investment universe. At retirement data of the participant, the accumulated capital needs to be divested and distributed to the participant. The algorithm will be tested on several (simulated) scenarios to explore its feasibility and optimality in a pension context.

1.3. Research questions

To achieve the research objective several research questions are formulated. The main research question is:

> *Is it possible for APG to create and implement an algorithm that automatically makes investment decisions - given a certain investment universe and the Basis Zero philosophy - with a pension objective that achieves adequate performance?*

To answer the main research question several sub questions are formulated. These questions are selected and formulated in a way that the answers will yield information that is useful or necessary for accomplishing the research objective.

1. What is algorithmic trading and how can it be applied in the context of pension investing?

The answer on the first question provides a better understanding of the algorithmic trading environment from a literature perspective. It should introduce methods that can be used to apply at the problem in this thesis.

2. What is the investment objective?

The answer on the second question gives more information about investing with a pension objective. For the participant the pension capital will be part of the second pension pillar that is additive to the first pillar. This means the pension we are talking about is additive to the AOW; the general old age pension act that should provide elderly with a basic pension. A target needs to be set where the algorithm should focus on. In addition, it is debatable how far results may deviate from this target to remain solid.

3. How can an algorithm be used to automate the investment process when taking the pension objective into account?

When answering this question an algorithm has to be delivered that combines aspects of algorithmic trading and investing with a pension objective. The algorithm should be able to automatically make decisions on how to invest incoming premiums given a certain investment universe and the objective. These decisions will be based on different selection rules. Relevant restrictions need to be taken into account when developing the algorithm.

4. How does the proposed algorithm perform when it is tested with historical or simulated time series?

By answering this question, the performance of the proposed algorithm can be assessed. By using historical data or simulated time series several selection rules can be tested. This should provide an answer which selection criterion can be used to optimise the investment decisions with a pension objective. By comparing the results to a benchmark the conclusion can be drawn whether the proof of concept is successful or should be adjusted.

1.4. Research strategy

Methodology and data collection

Different research methods will be used depending on the sub-question that needs to be answered. These are described in this section and presented in Figure 3. The first sub question about algorithmic trading will be answered by a review of existing literature. By reviewing the literature, an overview of, and insight in algorithmic models and their applications will be obtained.

The approach to answer the second question is based on literature as well. In addition the principles found in the literature review will be assessed by a consultation of experts within APG. This combination gives the possibility to couple literature and the view of APG to answer the question.

The knowledge gathered in the first two questions provides the foundation of question three. In the third question the algorithm will be created. When developing the algorithm the investment universe has to be defined first to be able to determine the scope of our investment possibilities. Furthermore for practical usability several constraints need to be included. This will be at least transaction costs and rebalancing criteria as well as limitations on the short sale of securities. When assessing the selection criteria and the overall performance of the algorithm to answer question 4, (simulations based on) historical data will be used. This information is used to explore the performance of the algorithm and compares its performance to the performance of APG that will be used as a benchmark.

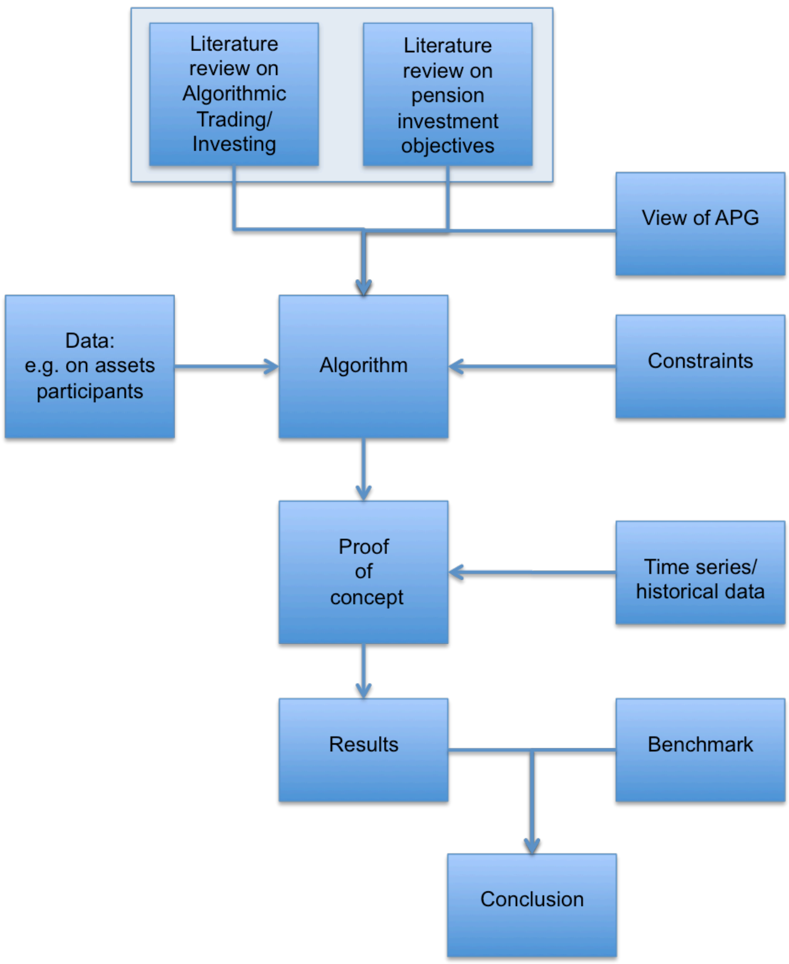


Figure 3: graphical representation of the research methods.

Outline thesis

This section gives an overview of the different chapters of this master thesis. Here, in the first chapter we introduced the company APG including the challenge underlying this report by illustrating the research design. The first part of Chapter 2 will be based on a literature review of relevant papers on the subject of algorithmic trading. In the second part of Chapter 2 we will focus on the investment objective that needs to be determined, keeping the pension setting in mind. The first two research questions should be answered in Chapter 2.
In Chapter 3 we will specify the investment universe followed by information about, for example, transaction costs and rebalancing requirements, that need to be kept in mind when developing the algorithm. In Chapter 4 the algorithm will be tested and evaluated by performing a case study based on real data and simulated time series. This will be the proof of concept of the algorithm and enables us to answer research question 4. Then we will present our results in Chapter 5. Chapter 6 will contain the conclusion and there we will answer the main research question. Limitations of this research and ideas for next steps/future research will be presented as well in Chapter 6. Chapter 2 and 3 provide the basis to create the algorithm, in Chapter 4 it will be tested and in the Chapters 5 and 6 the behaviour will be evaluated. This is shown schematically in Figure 4.

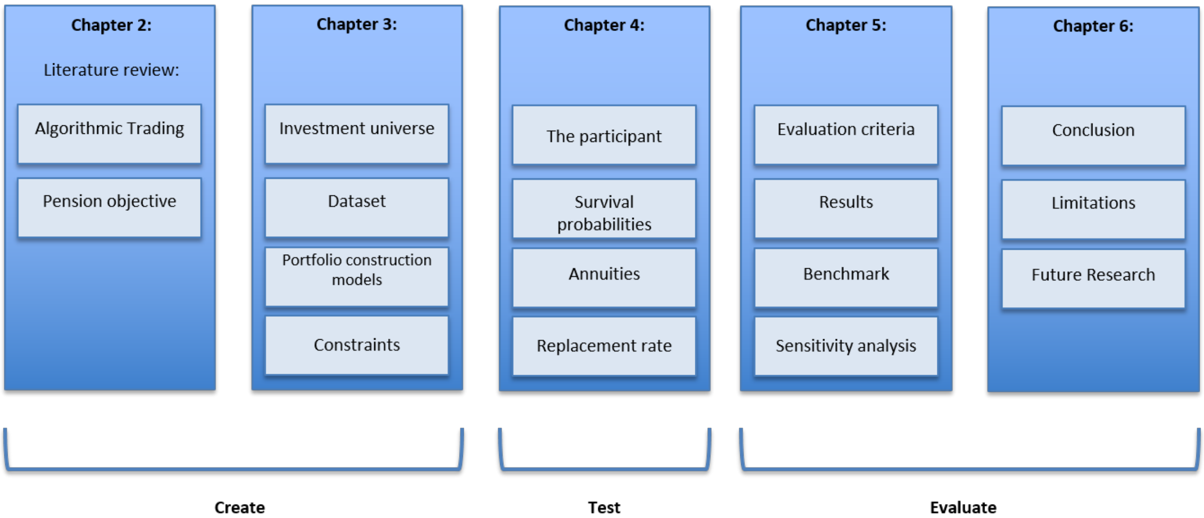


Figure 4: create, test and evaluate.

2. Literature review & theoretical framework

In this chapter we will answer the first and second research question. In the first part a literature review is conducted to introduce algorithmic trading. In the second part we will focus on the investment objective of the algorithm and therefore argue what characteristics a good pension should meet. This is done by a combination that starts with literature review, where after the results will be discussed with industry experts from APG. This assures the final algorithm of pursuing the right goals.
We will conclude both sections by a paragraph where the most important insights of each section will be summarized. These insights contribute to the development of the algorithm in subsequent chapters.

2.1. Algorithmic trading and investing

The first section of the literature review will focus on literature in the field of algorithmic trading and how it can be used in algorithmic investing. We will describe the five-step trading process by Treleaven (2013) and explain how it will be used in the remainder of this thesis. The trading process is very comprehensive, i.e. it includes many advanced steps, which enforces us to choose specific parts of the process and scope to those components that will be included in the algorithm.

Introducing Algorithmic Trading

Advances in telecommunications and computer technologies have created increasingly global, dynamic and complex financial markets (Nuti, 2011). This stimulated trading by computer programs and subsequently the rise of systems for Algorithmic Trading (AT) to automate one or more stages of the trading process. Treleaven (2013) defines algorithmic trading as any form of trading using sophisticated algorithms (programmed systems) to automate all or some part of the trade cycle. AT usually involves concepts of learning, dynamic planning, reasoning and decision taking.

Within the electronic trading environment there are several closely related terms that are sometimes confused. These include electronic trading, order management systems, automated trading, systematic trading, and algorithmic trading. Broadly, electronic trading is a method of exchanging securities, stocks, bonds, foreign exchange and derivatives. Within electronic trading, specialized programs bring together buyers and sellers through electronic media to create an exchange. Order management systems facilitate and manage order execution, generally connecting to one or more electronic exchanges. Automated trading systems usually refer to trade execution programs that automatically submit trades to an exchange. The distinguishing feature of algorithmic trading systems is the sophistication of their analysis and (speed of) decision-making. Broadly, these systems are deployed for highly liquid markets and high frequency trading, such as equities, futures, derivatives, bonds and foreign exchange (Nuti, 2011).

An algorithm can be described as a set of instructions for executing a specified task. A trading algorithm therefore is a computerised model that incorporates the steps required to trade an order in a specific way. Trading via algorithms requires investors to first specify their investing and/or trading goals in terms of mathematical instructions. Dependent upon investors' needs, customized instructions range from simple to highly sophisticated. After instructions are specified, computers implement those trades following the prescribed instructions (Kissell, 2014). The instructions can become quite complex for an algorithm to react to ever changing market conditions. These instructions determine the type, price and quantity for each order, often based on a mixture of historical and live market data (Johnson, 2010).
A computerized system is responsible for handling the algorithm's instructions, so its execution is fully automated. Developers use various types of simulations, including backtests and optimisations, to evaluate and improve utility of their algorithms (Treleaven, 2013).

Key stages in AT

Several papers (e.g. Johnson, 2010, Nuti, 2011, Treleaven, 2013) describe the trading process that can be roughly divided into diverse stages. We will continue with the process as described by Treleaven as he provides the most complete overview while others do not include all stages he covers. The trading process as described by Treleaven has a clear structure and describes five key stages of the investment process:

1. Data access/cleaning: the first stage obtains and cleans financial, economic and social data that will drive AT.
2. Pre-trade analysis: the second stage analyses properties of assets to identify trading opportunities using market data or financial news (data analysis).
3. Trading signal generation: identifies the portfolio of assets to be accumulated based on the pre-trade analysis (what and when to trade).
4. Trade execution: executing orders for the selected asset (how to trade).
5. Post-trade analysis: analyses the results of the trading activity, such as the difference between the price when a buy/sell decision was made and the final execution price (trade analysis).

Here the stages are presented as five separate phases. However, in practice several stages may overlap each other. Although one might think of AT as automating all stages, much of the activity in the industry is devoted to the data access/cleaning and pre-trade analysis stages, with the latter stages of the trading process being supervised by humans (Treleaven, 2013).

Data access/cleaning

Clean data is of major importance in AT for the analysis of a successful trading system. It may include financial or economic data, like price data on financial instruments but may also contain data from news sources or social media. Data can be gathered real time through data feeds (i.e. from exchanges) or can be accumulated and obtained as historical data. Unprocessed or raw data need to be cleaned by removing erroneous data that may be the source of errors in the subsequent process. Especially buying cleaned data is expensive and cleaning data is very time consuming. However it is essential due to the sensitivity of trading algorithms.

Pre-trade analysis

The next stage is the pre-trade analysis where an analysis is made whether several assets offer trading opportunities by using data analysis.
The pre-trade analysis comprises three main components: the alpha model, the risk model and the transaction cost model. The alpha model is a mathematical model designed to predict future behaviour of the financial instruments that the algorithmic system is intended to use. The risk model evaluates the levels of risk/exposure associated with a financial instrument. In the transaction cost model the potential costs related to trading the instruments are calculated. We will describe each component separately in the next sections.

Alpha model

In pre-trade analysis, and specifically the Alpha model, real-time and historic data will be analysed to identify potential trade opportunities. Nuti (2011) and Treleaven (2013) distinguish three principal techniques:

- Fundamental analysis: in a fundamental approach variables are evaluated that can affect a security's value. This can include macroeconomic factors (e.g. industry and overall economy conditions) as well as company specific factors (e.g. financial reports).
- Technical analysis: This approach is concerned with the analysis of trends and pattern recognition in charts.
- Quantitative analysis: this method uses a wide range of computational metrics based on statistics, physics or machine learning that are applied to capture, predict and exploit features of financial, economic (or other) data in trading. Quantitative analysis treats asset prices as random and uses mathematical and statistical analysis to find a suitable model for describing this randomness.

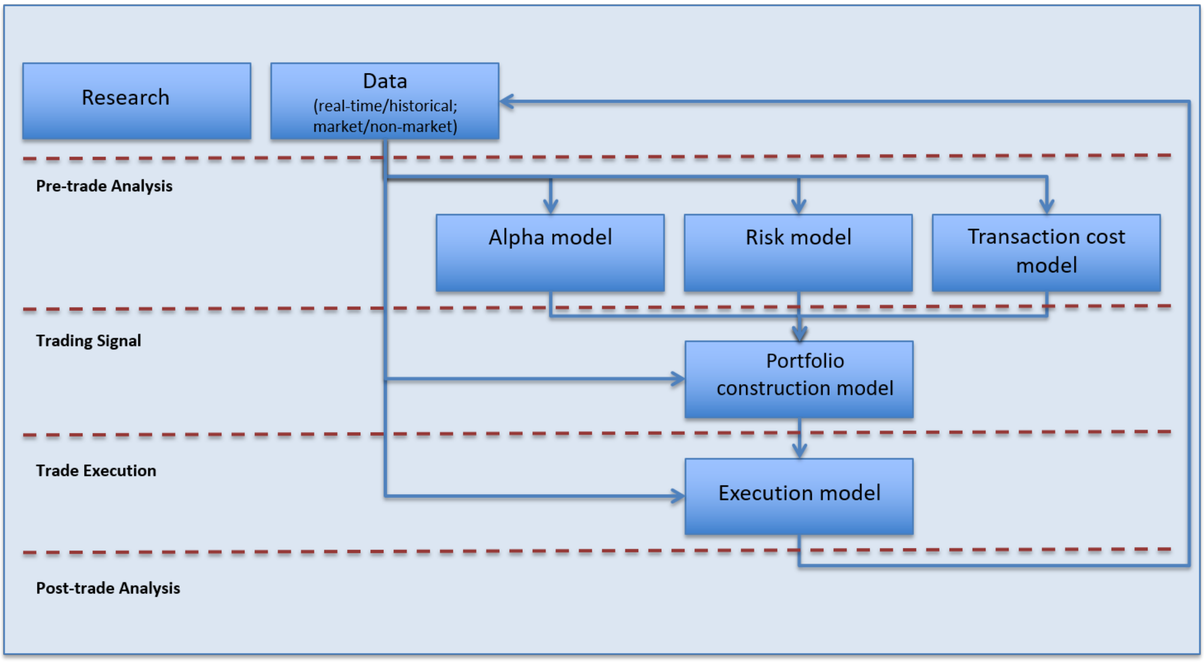


Figure 5: the major components of an AT system (Treleaven, 2013).

There are two strategic approaches within the alpha model: a theory driven and an empirical strategy. In the theory driven model a hypothesis is chosen that tries to describe the most likely behaviour of securities. By modelling this behaviour the hypothesis will be accepted or rejected. Examples are a momentum and a mean reversion approach that hypothesize market behaviour (theory-driven) based on price data. In an alternative empirical strategy, the algorithm will be used to identify a pattern in the underlying data of a security.

Constructing the Alpha model and more specifically setting the variables can be a highly complex task. Numerous factors influence the actual algorithm implementation: forecast goals (like direction, magnitude, duration and probability), forecast time horizon (such as millisecond, day, week, month), the mix of instruments, the data available, actual setting of the model's variables and the frequency of running the model (Treleaven, 2013).

Risk model

The risk model focuses on risks associated with a financial instrument and on the relevant factors that may affect the economic climate and so the future value of the (portfolio of) financial instruments. It tries to limit the amount of risk (e.g. volatility or leverage) and limits the type of risk by limiting whole types of exposure. Where the alpha model may propose many financial instruments in a particular industry the risk model is able to set a constraint that limits risks of the total exposure to that industry.

Transaction cost model

The last part of the pre-trade analysis is the transaction cost model. This model computes the possible transaction costs that could arise when different portfolios are constructed. Among those costs are commissions, slippage and market impact. Commissions are service charges assessed by (for example) exchanges in return for providing the purchase or sale of a security. Slippage is the difference between the expected and the actual price at which the trade is executed. Finally market impact is the effect of a market participant when it acts on the financial markets. When it buys or sells it is the quantity to which the price moves against the buyer or seller. The price goes up when buying and goes down when selling larger volumes of securities.

Trading Signal generation
In this stage of the trading process the portfolio construction model collects the results of the alpha, risk and transaction cost models. Based on this collection of data it selects the optimal portfolio (in terms securities and in what quantities they should be owned) for the next time step to maximize profit, limit risk and minimize transaction costs.

There are multiple portfolio construction models that can broadly be subdivided into two types of models: rule-based models and optimisation models. A rule-based model is a heuristic specification on how to assemble a portfolio with multiple instruments. An optimisation model uses an algorithm with an objective function. It iterates several portfolios until it finds the portfolio with, for example, a minimum variance level.
Both types of models are used, but in Chapter 4 we will provide a detailed description of the portfolio construction models that we will use.

Trade Execution
When the trading signal is generated and the optimal portfolio is constructed this model takes decisions regarding the execution of a transaction. Think about the trading venue (e.g. NYSE, NASDAQ), execution strategies (e.g. smart order routing) and order types (e.g. market or limit).

Post-trade analysis
After the trade is executed the results are evaluated during the post-trade analysis. An example is the evaluation of the difference between the expected price and the actual traded price.

From algorithmic trading to the Basis Zero algorithm
This section provided an overview of the concept algorithmic trading and covered the methods and stages that are part of the trading process. The next step is the application of this information in the development of the Basis Zero algorithm. Within the experiment/thesis not all phases of the trading process can be covered. As simplicity as well as feasibility is important, we decided to reduce the scope to a specific part of the process. Our contribution focuses on the composition of portfolios over time where aspects of the pre-trade analysis as well as the trading signal generation will be included. Focussing on those stages will result in an algorithm where the most essential components are included that should result in an algorithm within the philosophy of Basis Zero.

The consequence of our choice to focus on those stages is ignoring other parts of the trading process. An example is the cleaning of data, which is out of our scope. The data source that will be used is Thomson Reuters DataStream and is expected to deliver appropriate data for the trading algorithm. There will be no focus on trade execution and analysis. Therefore it is not the goal of the algorithm to be ready to include in operations immediately.

2.2. The investment objective

The second section of the literature review will focus on literature in the field of pensions. How is the pension system arranged and what is considered to be a good pension? In particular this last question is important when setting a target for our algorithm. The found objective is discussed and approved by APG experts and used in the successive chapters of this thesis.

Pension Pillars

Different types of pension schemes are usually grouped into multiple pillars of a pension system. Many systems distinguish between statutory, occupational and individual pension schemes. It is also common practice to distinguish on voluntary and mandatory schemes. The Dutch pension system consists of three pillars: the state pension, a collective second pillar and individual (sometimes additional) third pillar pension products. The focus of this thesis is on a second pillar product.

The first pillar is the state pension (AOW), a pay-as-you-go system that is the foundation of the old-age pension benefits. The AOW provides a basic income that is linked to the minimum wage level and the number of years that a person has resided in the Netherlands.
Single pensioners who have lived in the Netherlands between 15 and 67 receive 70% of the minimum wage; couples receive both 50% of the minimum wage. For people with a low or no pension income and (almost) no wealth the first pillar will be supplemented with social assistance to guarantee a social minimum.
The first pillar pension only provides a limited part of all old age benefits and can be supplemented with benefits from the second and third pillar.

Collective pension schemes are represented in the second pillar and are administered by pension funds or insurance companies. The second pillar accommodates capital-funded occupational pensions of which the primary responsibility lies at the level of employees and their employers. The pensions are financed by contributions of the participants, their employers and by the returns on investment over these contributions.
Occupational pensions in the Netherlands have a mandatory nature, such that 90% of all employees have pension schemes with their employer (Knoef, 2014).
Up to the beginning of the 21st century, most pension plans aimed to pay a pension income of 70% of the final gross salary starting at the age of 65. From 2003 onwards, pension funds lowered their ambition and now mostly aim to pay 70% of their average career salary, instead of 70% of the final gross salary (Knoef, 2014).
Recent economic instability also revealed the vulnerability to shocks in the financial markets. Many pension funds have difficulties achieving their ambitions and sometimes have too cut in pension payments.
In the arrangement of pension agreements social partners inevitably need to find the right balance between the aspired pension outcome, the degree of certainty of the pension outcome and the costs (the contributions). In particular, after the dotcom crises the real tradeoff between ambition, security and costs became clear.
For a long time there was a broad consensus that a target pension (including AOW) of 70% of the last salary was considered as a ‘good pension’. It should however be noted that there is a high degree of diversity between pension schemes and many people do not reach this target in practice (Goudswaard et al., 2010).

The third pillar consists of private individual pension products (like life annuities) that are mainly used by people in sectors without collective pension schemes or by self-employed people. It can also be used to purchase a product to meet additional requirements (not fulfilled in the first and second pillar), for example to save for extra pension.

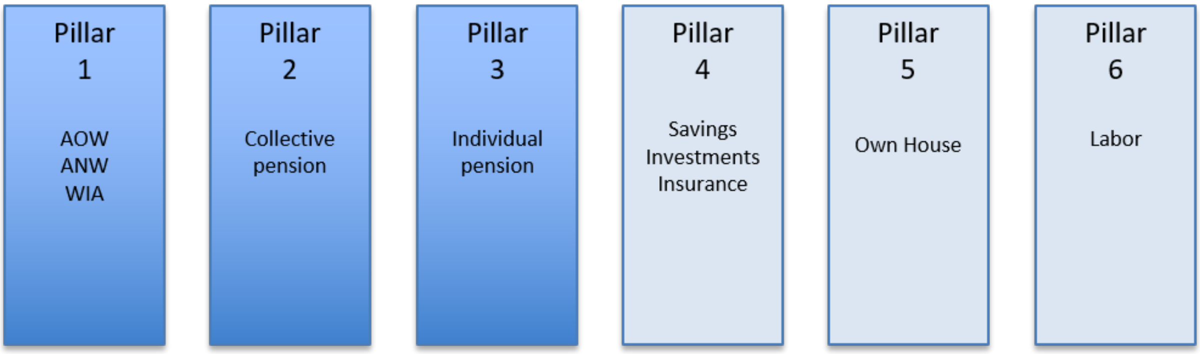


Figure 6: the three traditional pillars made up to six pillars (García-Huitrón, 2016).

Other pillars, like presented in Figure 6, can be found in housing wealth or an extension of working life. People who have paid off their mortgage can benefit from lower housing costs during their retirement. Although it is not commonly done by the current generation of elderly, people may move or use reverse mortgages to deplete housing wealth (Knoef, 2014). As already indicated the focus will be on the second pillar pension benefits: supplementary benefits to a basic income provided from the first pillar.

Introduction in retirement schemes

Retirement schemes may be classified according to how the benefits are determined. The two main categories are Defined Benefit (DB) and Defined Contribution (DC) schemes.

In a Defined Benefit scheme the benefits accrued are linked to earnings and the employment career of the participant. The future benefit is predefined and promised to the participant. Consequently it is the scheme sponsor who is bearing the investment risk and also longevity risk. If predictions about rates of returns or life expectancy are not met, the scheme sponsor must increase its contributions to pay for the resulting gap.

Opposed to Defined Benefit schemes, there are Defined Contribution schemes where the level of contributions is predefined. This means no future benefit is promised and the pension level will depend on the performance of the investments and the contributions made to the scheme. The individual participant therefore bears the risk and needs to decide how to mitigate the risks.

A Collective Defined Contribution (CDC) scheme is a hybrid scheme additional to the traditional DB and DC schemes. It combines the limited risks of fluctuating pension commitments for an employer with advantages of a collective pension scheme. In a CDC scheme the pension capital is based on the salary and the number of years someone is employed/participates in the scheme like in the DB scheme. The contributions are fixed for many years. That means when they are insufficient, the benefits will be lower than originally expected.

Basically a DC scheme will be the foundation of the algorithm that will be built. Extensions, like discussed in Chapter 5 on the results, can provide a collective touch to the product.

Goals of a pension system

According to Hinz and Hollman (2005), the primary goals of a pension system should be to provide an adequate, affordable, sustainable and robust retirement income.

Adequacy is reflected in a system that provides benefits to the population to prevent old-age poverty on an absolute level and in addition to provide reliable means to smooth lifetime consumption for the vast majority of the population, i.e. replacing sufficient lifetime earnings. This includes assurance that those individuals that live beyond the norms from the risk of longevity.

The system should also be affordable, one that is within the financing capacity of the individual participants and society. It should not unduly displace other social or economic imperatives or have untenable fiscal consequences.
World Bank experience (Hinz and Hollman, 2005) indicates that mandated contribution rates in excess of 20 percent are likely to be quite detrimental for middle- and high-income countries.

A sustainable pension is a pension that is financially sound, can be maintained over a foreseeable time horizon under a broad set of reasonable assumptions. A sustainable pension should be structured in a way that the financial situation does not require unannounced future hikes in contributions or unforeseen cuts in benefits. In other words, all adjustments that are needed to keep the pension system financially sound (i.e. changes in contributions, benefits or retirement ages) should be included in the design of the system. This includes mechanisms to adjust the program to periods of economic depression.

Finally, a robust pension is one that is able to withstand major shocks, like those from economic, demographic and political volatility. The system must have the capacity to remain viable when unforeseen circumstances arise. In this regard, the most important outcome is the ability to sustain income replacement targets in a predictable manner. A central element in meeting this goal is a credible analysis across the full range of likely scenarios, over the full term required to reach long-term stability. To fulfil this goal, we need to apply sophisticated modelling tools to present analyses that incorporate a significant range of variation in basic assumptions or scenarios to demonstrate the viability of our system over the long term.

In a publication of the Pensions Institute (2016), we find several criteria for a good DC pension scheme. The Pensions Institute set up a list and the most relevant criteria can be found in the list below. The pension scheme should:

- Deliver adequate and sustainable (stable) pensions
- Provide an investment strategy that reflects the scheme member's attitude to and capacity to take risk and generates a return at least as high as inflation.
- Provide value for money for every euro saved in the scheme
- Have transparent charges and costs
- Provide reliable and efficient administration
- Deliver effective communications to members

Now we gathered several qualitative criteria a good pension scheme should meet, but still a good pension is difficult to define; it is not defined in policy or regulation. Even between people the definition of a good pension may vary because individual needs may vary. It is also the question if good means if participants get what they need, rather than what they want. When developing a model and in particular to evaluate the performance we want to work towards a quantitative measure that enables us to assess the performance of the model in an objective manner.

How to measure a good pension?

What is the right target to aim for when it comes to retirement income? Knoef (2013) conducted a study to find out what an adequate level of resources for retirees should be. She concluded a variety of standards could be chosen against which to judge adequacy. The article is focused on two measures, the Life Cycle Hypothesis and an absolute or social replacement rate.

The Life Cycle Hypothesis is a theoretical framework that is able to assess the adequacy of savings (Banks et al., 1998). In the model, consumption is not based on current income, but by expected lifetime resources. It should be optimal for persons or households to save (or borrow) to the extent that, after discounting, the marginal utility of consumption is smoothed over the life cycle.

Another measure is the replacement rate. Knoef (2014) distinguishes between two different replacement rates: an absolute and a social replacement rate. In the second approach a social standard is set for adequacy. The retirement income is then considered adequate when it is equal or greater than poverty levels of income (Haveman et al., 2007).

The (absolute) replacement rate generally refers to an indicator showing the level of pension income as a percentage of earnings before retirement. The replacement rate measures the extent to which a pension system enables participants to continue their standard of living when moving from employment to retirement. This methodology assumes a gradual wage growth over time without for example a peak or an extraordinary pay rise shortly before retirement. The conventional replacement rate formula is given in Formula 2.1:

$$\text{Conventional replacement rate} = \frac{\text{gross income in first year of retirement (e.g.at age 67)}}{\text{gross preretirement final employment earnings (e.g.at age 66)}} \qquad (2.1)$$

The most commonly advocated benchmark in literature is having a retirement income of at least 70% of gross final annual employment earnings (e.g. Haveman et al., 2007, Goudswaard et al., 201 MacDonald et al., 2014). This amount is regarded as the income needed to sustain an individual's standard of living after retirement. This benchmark is used by financial planners, pensions plan advisors, academics, public policy makers and much of the research that predicts workers will be financially unprepared for an adequate retirement (MacDonald et al., 2014).
The 70% is based on the idea that retirees will pay lower taxes, will not be saving for retirement anymore, have often paid off their mortgage and no longer need to support children nor pay expenses related to their former jobs. Related to the Life Cycle Hypothesis, Boskin and Shoven (1987) conclude that a replacement rate of less than unity is consistent with the Life Cycle Theory.

The method of the replacement rate is simple, however there is also criticism on the measure and the 70% norm of a good pension. Several studies have been sceptical if the replacement rate indeed provides the benchmark for adequacy where we are looking for (e.g. Vanderhei, 2006, Scholz and Seshadri, 2009, MacDonald et al., 2014).

MacDonald et al. (2014) conclude that people who attain a replacement rate of 65-75% of gross final earnings will experience a large range of changes of their average living standards after retirement when the traditional replacement rate is used. As MacDonald et al. describe, it is clearly problematic when the living standard is further reduced while the opposite, an improved standard could be the result of an over-sacrifice of welfare before retirement.

According to MacDonald et al. (2014) the problem lies in the computation of the replacement rate itself. It criticizes the measurement period and disagrees with not (accurate) incorporating many components of living standards or even omitting them from the equation. The article provides an extensive list of factors that should be included from which the most important are stated below:
1. Household-level differences in consumption needs due to family size (and changes over time in household size and composition)
2. Imputed income from owner occupied housing
3. Taxes (specifically the differentials in taxation year by year pre and post retirement)
4. Transfers – e.g. unemployment insurance, child benefits and social assistance
5. The accumulation and drawdown of non-traditional forms of savings (non-registered financial wealth/debt, and home-ownership equity)

Knoef (2014) continues naming the replacement rate as the key indicator of savings adequacy and retirement readiness. It stays the most used concept for evaluating participants' likely living standard in terms of adequacy. In line with the philosophy of Basis Zero we also continue with this simple, easy explicable measure.

The investment objective in the Basis Zero algorithm

In this section the investment objective of the algorithm is set. The goal of the Basis Zero experiment is the development of a pension product that is easily explicable to its participants. Therefore it is a logical step to continue assessing the adequacy of the algorithm's results based on the replacement rate. It is a simple, widely used measure in theory and is supported by APG for the purpose of this thesis.
The minimum acceptable replacement rate level is set at 60% but the objective of the algorithm will be a 70% replacement rate. There is a widely supported belief, a 'rule-of-thumb', that 70% of final earnings will provide a good pension.

Now the 'adequacy question' is answered, the affordability, sustainability and robustness have to be checked. As a DC scheme is used a fixed percentage of salary will be invested in the pension product. An 18% premium percentage is considered common and can be seen as affordable. When contributions and benefits stay at the same level and the product can be maintained over a foreseeable time horizon the product is considered sustainable. The DC scheme is an individual product and therefore the investment risk is with the participant. It is uncertain what level of capital will finally be accumulated till the pensionable age is reached. Therefore the sustainability and robustness checks cannot be confirmed here. To test those properties the algorithm needs to be exposed to several (simulated) scenarios that will be illustrated in Chapter 4.

The last few criteria for a good pension scheme are reflected in the implementation of the various components of the Basis Zero experiment. By automating the investment process an attempt is made to provide value for money, and, at the same time, be transparent about charges and costs. The conclusion whether the algorithm succeeds sufficiently in this aspect cannot be drawn yet. Another part of the project, as described in the context of the study and Figure 2, will focus on reliable and efficient administration.

3. Characteristics of the Algorithm

In Chapter 3 we will combine the information gathered in the first two sub questions to create the algorithm. This chapter starts with an outline of the investment universe to scope the investment possibilities. Afterwards we will present the dataset that is used as input in the decision-making or investment process. The asset allocation will be based on several portfolio construction strategies that are described in the third paragraph. Important constraints that need to be taken into account are transaction cost and rebalancing criteria. We will discuss those criteria in Subsection 4. We will complete this chapter with a graphical representation of the succeeding steps that the algorithm takes to invest or divest funds every time step.

In this chapter we want to provide an answer to Question 3. The associated deliverable is an algorithm that combines aspects of algorithmic trading and investing with a pension objective. The algorithm itself will not be presented in this report as it consists of too many scripts to present a clear overview. However the different steps in the algorithm takes will be explained by us as clear as possible after which the obtained results will be presented in Chapter 5.

3.1. Scope of the investment universe

The original idea of the Basis Zero experiment was to arrange exchange connectivity to Euronext and be able to invest in every product traded on this exchange. For the purpose of this thesis the investment universe is limited to 7 asset classes and 10 liquid products on several corresponding indices. The tenth instrument is a(n approximated) 'risk-free' or cash instrument to which funds can be allocated. The indices are chosen based on geographical diversification in case of equities.

For the other asset classes, the most common instrument for each asset class will be used. The 10 instruments are chosen based on liquidity and therefore investment volumes do not form a restriction on our trading possibilities. The algorithm is designed in a way that makes expanding easy and the ten instruments are therefore only used as a scope for the purpose of this thesis.

We decided to use total return indices rather than price indices. While in a price index only the price movements are considered, in a total return index also cash distributions such as dividends are reinvested. Therefore the total return index is considered as a more accurate measure of performance compared to the price index where these distributions are ignored. Other criteria considered for the choice of the type of indices were the availability of sufficient historical data points and the need to comply with the general requirements that have to be met by a benchmark.

The asset classes and indices that will be used in this thesis are shown in Table 1 and will be explained hereafter.

Table 1: the asset classes and several corresponding indices that will be used.

Asset Class	Index
Equities Developed Markets	MSCI North America
	MSCI Europe
	MSCI Asia Pacific ex Japan
	MSCI Japan
Equities Emerging Markets	MSCI Emerging Markets
Real Estate	FTSE EPRA/NAREIT Developed
Commodities	S&P GSCI
Credits	Barclays US Corporates: Investment Grade
Treasuries	Barclays Euro Aggregated Treasury
Risk free / Cash	J.P. Morgan Euro Cash 1-Month

Equities Developed Markets

The MSCI World index offers a broad global diversified index and represents large and mid-cap equity performance across 23 developed markets. It is one of the best-diversified, transparent and replicable indices and therefore will provide a good representation of this asset class.

To include geographical diversification over the different developed markets the index is broken down in four categories representing North America, Europe, the Middle East-Pacific (ex-Japan) and Japan.

Equities Emerging Markets

For the emerging markets the MSCI index also offers a broad and diversified index that gives a representation of several countries in South America, Europe, the Middle East & Africa and the Asia region.

Real Estate

The FTSE EPRA/NAREIT Global Real Estate Index and the S&P Global Property Index are the most commonly used indices with the widest geographical coverage for global listed real estate. The FTSE Index is preferred as it sets some more stringent rules on minimum volumes, what better suits the investment portfolio of a pension investor like APG.

Commodities

Commodity futures are the most efficient way of investing in commodities compared to physical investing in commodities. The S&P GSCI Total return Index is the most well-known, prominent global commodities futures index. It uses a 5-year moving average value of each commodity as weighting factor that results in a large (60%) weight on the energy market. This makes the index not very diversified; however still representative for the commodity market.

Credits

For credits the Barclays US Corporates Investment Grade index is chosen as index. It reflects the benchmark of customer ABP and meets the general requirements that may be imposed for benchmarks whereby sufficient historical data is available.

Treasuries

The main goal of investing in treasuries is the preservation of capital and liquidity. The choice was made for treasuries in the local currency, Barclays Euro Treasuries Bond Index. APG benchmarks its performance for 50% on this index and for the other half on the Barclays Global Majors Bond index with fully hedged currency risk.

Risk free / cash

For the risk-free option we choose to use the J.P. Morgan Euro cash 1-month total return index.

Several asset classes (private equity, infrastructure, hedge funds and several debt instruments) are not included in our asset pool for various reasons. Some classes are not represented well by a corresponding index, like private equity and infrastructure investments (due to the deal-driven nature). Another reason is our self-imposed limit on 10 indices to scope the investment possibilities. Therefore the natural choice is made to prefer assets like commodities and real estate above an asset class like emerging markets debt.

The indices as specified before are noted in different currencies or focus on different geographical areas. An example is the credits index that is focused on US Corporates while the treasuries index is aimed at the euro zone. The benchmark ABP uses is leading in our choice when it comes to currency choice.

Correlations

Based on our complete dataset of returns we computed the correlations between the indices and present them in Table 2. One correlation matrix is used, that means no extra attention for tail correlations or changing correlations is taken into account.
Note that two variables are uncorrelated if $\rho = 0$. Otherwise they are correlated to a greater or lesser extent. When $\rho = 1$ the two variables are perfectly correlated and move in the same direction. If $\rho = -1$ the variables move exactly the same in the opposite direction. From the table can be concluded that some assets or asset classes correlate substantially with each other while others do not. As expected, there are some high correlations between the equities. The exception is the Asia Pacific index that has a somewhat lower correlation compared to the other equities. The index that covers real estate investments (FTSE EPRA/NAREIT) also significantly correlates with the equity indices.
The commodity index and the credits and treasuries show mainly negative correlations with other asset classes. The equity indices do not meet this criterion, but leaving them aside as individuals (and see them as one asset class) several low or negative correlations can be distinguished.

Table 2: correlation matrix.

	1.	2.	3.	4.	5.	6.	7.	8.	9.
1. MSCI North America	1,00								
2. MSCI Europe	0,93	1,00							
3. MSCI Emerging Markets	0,88	0,97	1,00						
4. MSCI Japan	0,71	0,85	0,86	1,00					
5. FTSE EPRA/NAREIT	0,63	0,76	0,85	0,78	1,00				
6. MSCI Asia Pacific ex Japan	0,20	0,45	0,55	0,63	0,82	1,00			
7. S&P GSCI	-0,15	-0,12	-0,01	-0,14	0,25	0,14	1,00		
8. BARCLAYS US CORP INVESTMENT GRADE	-0,36	-0,32	-0,27	-0,16	-0,19	0,22	-0,28	1,00	
9. BARCLAYS EURO AGG TREASURY	-0,74	-0,82	-0,79	-0,79	-0,77	-0,50	-0,05	0,58	1,00

The selected asset classes and their corresponding indices that will be used in the algorithm are described in this section. This provides the scope for the investment possibilities. In the next section we will specify our dataset more in depth.

3.2. Dataset asset classes

We will use historical time series from the 10 indices presented in the previous section as our dataset. It includes monthly closing values from the indices ranging from January 31st 2002 to April 29th 2016. This results in a data set of 172 data points for each index retrieved from Thomson Reuters DataStream. In consultation with APG Asset Management this time frame is chosen based on the length of the period (to provide sufficient data) and actuality of the data (too old data does not provide a realistic picture of current markets anymore). The transition to the Euro currency is also excluded.

In Figure 7 we show a graphical performance representation of the indices in the chosen time frame. This time frame includes periods of economic prosperity as well as periods of economic distress. The drop in the graph around 2008 corresponds to the collapse of the stock markets at that moment.

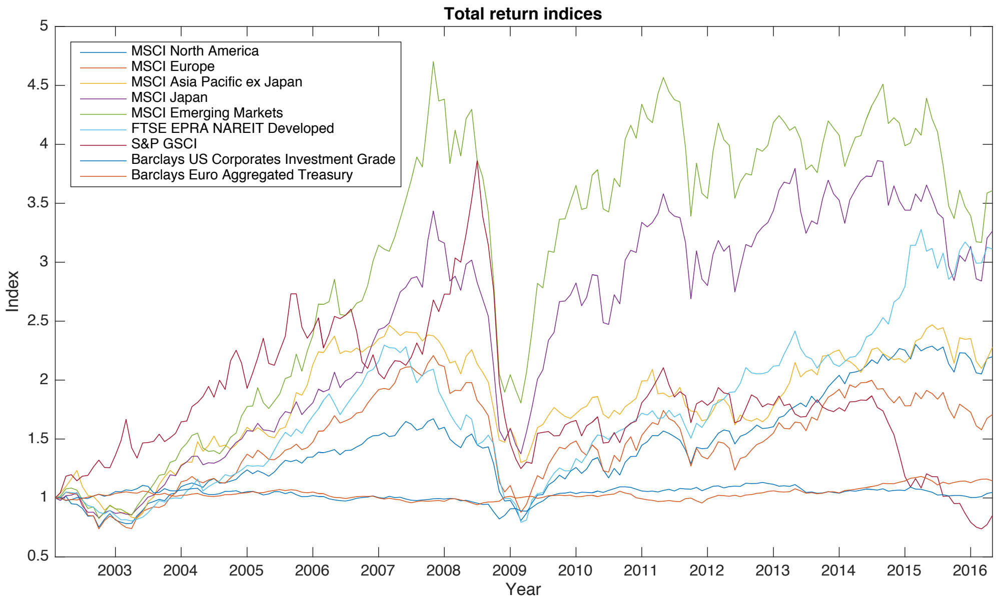


Figure 7: Total return indices over the specified period (source: Thomson Reuters DataStream).

In the algorithm price data are turned into return data that will be used as input for the investment models. To evaluate the performance of the portfolio construction models applied in the algorithm, time series will be created on which the algorithm is tested. Due to the period of financial distress both calm and volatile periods on the financial markets are included in the dataset. This may provide us several stress scenarios in the simulations. Chapter four will take a closer look on this topic.

3.3. Portfolio construction models

Having described the investment instruments, we will now introduce the method to determine the asset weights within the portfolio during each month *t*. In Chapter 3 already two types of portfolio construction models were discussed: rule-based (heuristic) and optimisation models (objective function). For the algorithm we will introduce 3 rule based models and 2 optimisation models and evaluate them in the remainder of the thesis. Besides these models a life cycle model is included whereby the number of models equals 6.

Selection of the model types

When considering which models we should evaluate for the algorithm the paper of Treleaven (2013) provides several options that are widely used in an algorithmic trading environment. Various models focus on individual products (i.e. single stocks) in contrast to indices in the manner and scale they are used in our situation. An example is a market-neutral strategy. A strategy that takes a long position in certain instruments while shorting others. The result should be a portfolio with no net exposure to market moves. As our focus is different, several strategies were considered not usable when keeping the Basis Zero vision in mind.

We selected several models out of the paper of Treleaven (2013). In the category 'rule-based models', momentum, mean reversion and equal position-weighting (*or 1/N*) meet the requirements. Opposite to the rule-based models there are optimisation models. Examples that are considered are Black-Litterman-, mean variance- and minimum variance optimisation. Treleaven (2013) mentions Black-Litterman optimisation as a sophisticated optimizer popular with AT and it looks like a nice extension to the traditional Markowitz models. However, the inclusion of the 'view of the investor' makes the model less applicable in the Basis Zero – Rottumerplaat philosophy where complete automation is important. Our algorithm should make its investments without a 'view' from any analyst or portfolio manager. Although there might be hundreds of other models, the chosen models meet the requirements and provide a good representation of the different possibilities proposed in literature. Some are rather simple; others use a more sophisticated approach but all of them are not too complex and explicable, in line with the Basis Zero philosophy. A schematic overview of the models is given in Figure 8 and each model will be discussed in detail starting from the subsequent section.

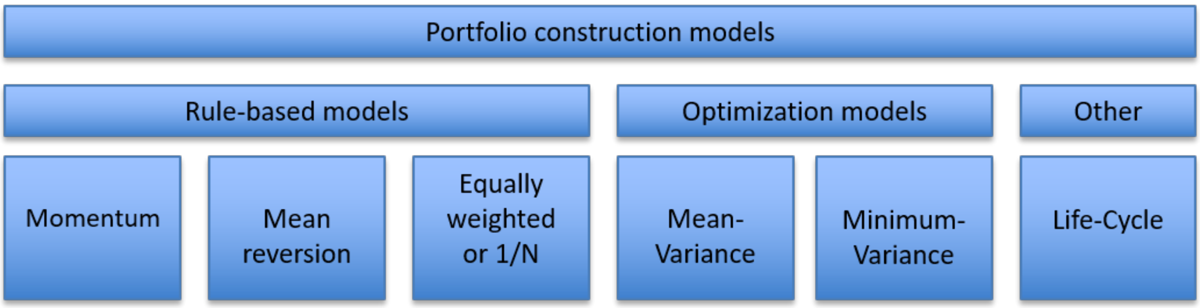


Figure 8: the portfolio construction models.

General model assumptions

Every asset allocation model uses several assumptions. One of them is a long only constraint, i.e. the weights of each asset must be greater than or equal to zero; short positions are not allowed. Subsequently a budget constraint will be used, i.e. all weights must sum to one in order to allocate all funds to an asset class (risky as well as risk free assets). The risk free asset is included as part of the mean variance model.

All portfolio construction models use (simulated) time series with return data as input parameters. Section 4.2 will explain how these time series will be created. Based on these time series and the chosen model a recommended asset allocation is determined. Assets are invested according to the recommended allocation. However, the model may deviate from the recommended allocation to save on transaction costs. This is further described in the section on rebalancing costs starting from page 32.

Description of the model characteristics

The selected models, increasing in complexity, are described in the subsequent sections. It is the purpose of the sections to introduce our implementation of the portfolio composition techniques from literature. The first model is a simple equal position-weighting model that does not need much explanation. Passing the rule based models we arrive at Markowitz' mean- and minimum-variance portfolio models.

Equal Position Weighting or 1/N

The naïve portfolio diversification rule is defined as allocation a fraction *1/N* of your wealth to each of the *N* assets available for investing. DeMiguel et al. (2009) mention two reasons for using this naïve rule as a benchmark. First, it is easy to implement. You do not have to rely on estimation of the moments (e.g. mean/variance) of asset returns or on any form of optimisation. Second, despite the development of sophisticated theoretical models and the advances in methods for estimating the parameters of these models, investors continue to use such simple allocation rules for allocating their wealth across the assets. Therefore it qualifies itself as a benchmark to evaluate the performance of the other models compared to this model.

Momentum

The momentum strategy looks for trends in stock prices. It is based on the conviction that stocks with an upward momentum keep rising and should therefore be bought. Stocks with downward momentum are expected to decline in value and should be sold. The two categories are classified as 'winners' for the former and 'losers' for the latter.

Jegadeesh and Titman (1993) show that strategies that buy stocks that have performed well in the past generate significant positive returns over 3- to 12-month holding periods. Also by selling stocks that performed poorly and selling them based on the same time horizons generate significant positive returns. The strategy that they examined in most detail selects stocks based on their past 6-months returns and holds them for the upcoming 6 months. When holding the stocks for a longer period a part of the abnormal returns generated in the first year after the portfolio formation will dissipate in the following two years according to the authors.

In our application of the momentum approach we want to replicate the method of Jegadeesh and Titman. Therefore indices are selected on their prior 6-month performance and will be held for the upcoming six months. Since it is unclear how each index would perform in the next period, a simple portfolio construction rule will be used: invest equal amounts in each index. Investments will be made in the top three based on equal amounts.

In this thesis the momentum approach of Jegadeesh and Titman (1993) is followed while using a simple and interpretable implementation. However, in literature there are more sophisticated models that incorporate a momentum factor. An example is the Carhart Four Factor Model. It is an extension of the more commonly known Fama-French Three Factor Model. The Carhart model adds an additional factor to the three-factor model: momentum. As the Four Factor model is much more complicated compared to the philosophy of Basis Zero we do not continue with these more sophisticated models, but rather continue with the basic momentum approach of Jegadeesh and Titman.

One of the constraints that need to be taken into account when implementing 'momentum type' models is a limitation on short sales. As the momentum model uses short sales it cannot be fully replicated. This does not need to be a problem, however, when adopting the model users should be aware, as the results will be influenced by this constraint. Half of the effect may disappear. The same holds for the mean reversion model that we will illustrate in the next section.

Mean Reversion

The mean reverting model assumes returns eventually move back toward their mean or average. Balvers and Wu (2005) reviewed several studies and conclude that there is considerable evidence that both momentum and mean reverting models produce excess returns. Having discussed the momentum model in the previous section here a deep delve into the mean reversion assumption will be made.

Balvers and Wu (2005) indicate that positive excess returns may be generated when sorting firms by previous returns and holding those with the worst prior performance and shorting those with the best performance. On the contrary there is the momentum strategy where firms are sorted by previous returns and those with the best prior performance are bought and those with the worst performance sold.

Balvers and Wu (2005) already conclude that it might seem inconsistent that both models generate excess returns. They argue that both strategies are effective since a different time span is used. The mean reverting model is powerful for a sorting period ranging from three to five years and a holding period of again three to five years. Momentum models typically work for a sorting and holding period of three to twelve months.

In our application of the model the performance of the indices will be monitored over the past 48 months. The indices showing the worst performance will be bought for a period of four years after which they are sold again.

Mean Variance Optimisation

In mean variance optimisation a portfolio will be assembled that maximizes the expected return for a given amount of risk; defined as variance. The portfolio that meets this goal is called the efficient portfolio. In mean variance optimisation risky assets as well as the risk free asset will be used.

The mean variance portfolio of risky assets and the risk free asset can be found by solving the maximization problem that is defined by DeMiguel et al. (2009). It can be solved by quadratic programming – a special type of mathematical optimisation where the optimisation problem is a quadratic function of variables subject to linear constraints on the same variables:

$$\begin{aligned} &\max_w w^T \hat{r} - \frac{\gamma}{2} w^T \Sigma\, w \\ &subject\ to\colon w^T \iota_{n+1} = 1 \\ &\qquad\qquad w_i \geq 0\,, \forall i \in N \end{aligned} \tag{3.1}$$

In the maximization problem as defined in Formula 3.1 w^T represents the transposed vector of weights invested in each asset. The vector includes risky as well as risk-free assets. $\hat{r}$ is the sample vector of expected returns based on a moving average. Sigma (Σ) represents the covariance matrix of risky assets, supplemented with the risk free asset. Gamma represents a risk aversion parameter that will be discussed below.

The first constraint ensures the sum of the weights is 1. The second constraint represents the long only assumption, basically the same as in the minimum variance optimisation problem.

Intuitively in mean variance optimisation we want to maximize the returns (in the left part of the maximisation problem). However we are inhibited by the variance multiplied by a risk aversion parameter γ, in the right part of the problem.

Gamma (γ) represents the risk aversion of the investor. Numbers like $\gamma = 2, 5, 10$ are considered common in literature (e.g. Fugazza et al., 2015). A risk-averse investor would invest more in the risk free asset and therefore has a higher risk aversion parameter. A risk-tolerant investor has a lower risk aversion parameter and would therefore invest more in risky assets. In the remainder of this report we consider a gamma equal to 5 as appropriate: not too risk averse nor too risk-tolerant.

This model requires mean returns and a covariance matrix of asset returns. Those are determined based on our sample of historical data (and simulations).

Minimum Variance Optimisation

Suppose an investor wants to invest its capital in a portfolio with as little risk as possible. Moreover, he does not care about possible missed returns; the only important criterion is the least amount of risk.
The focus needs to be on minimizing the variance and so minimum variance optimisation will be the problem that needs to be solved.

The minimum variance portfolio of risky assets only (without the risk free asset) can be found by solving the quadratic programming problem as defined below (DeMiguel et al., 2009). The risk free asset is not included here because the minimum variance portfolio always seeks the least amount of risk and would always fully exploit the possibilities to invest in the risk free asset, as the variance is considered very low. As we use 9 assets here we rename our vector w as we used in the mean variance optimisation to w_r. Our covariance matrix changes from 10x10 to a matrix of 9x9 and is therefore renamed by Σ_r. Iota is a vector of ones with length *n*, (n=9).

$$\begin{aligned} &\min_{w_r} w_r{}^T \Sigma_r\, w_r \\ &subject\ to\colon w_r{}^T \iota_n = 1 \end{aligned} \tag{3.2}$$

The solution should result in a portfolio where the variance of the risky asset returns is minimized. A minimum variance investor is only interested in a variance that is as low as possible and does not take returns into account. The vector that represents the weights of the minimum variance portfolio can be found by:

$$w_r^{min} = \frac{1}{(\iota_n)^T \Sigma_r^{-1} \iota_n} \Sigma_r^{-1} \iota_n \tag{3.3}$$

It is the solution of the minimization problem as defined in Formula 3.2 and its derivation is illustrated below.

The minimum variance portfolio is the portfolio determined by Formula 3.2. We can find the conditions for a solution to this problem using Lagrange multiplier λ. Now the Lagrangian can be formed:

$$L = w_r^T \, \Sigma_{\rm r} \, w_r - \lambda \, (w_r^T \, \iota_n - 1) \tag{3.4}$$

The Lagrangian can be differentiated with respect to w_r and set this derivative to zero.

$$\frac{\delta}{\delta w_r} L = 2 \, \Sigma_{\rm r} \, w_r - \lambda \, \iota_n = 0 \Rightarrow w_r = \lambda (2 \, \Sigma_{\rm r})^{-1} \, \iota_n \tag{3.5}$$

$$w_r^T \, \iota_n = 1 \;\Leftrightarrow w_r^T \, \iota_n = \lambda \, \iota_n^T (2 \, \Sigma_{\rm r})^{-1} \, \iota_n = 1 \;\Rightarrow \lambda = \frac{1}{\iota_n^T (2 \, \Sigma_{\rm r})^{-1} \, \iota_n} \tag{3.6}$$

Now we can conclude that

$$w_r = \frac{1}{\iota_n^T (2 \, \Sigma_{\rm r})^{-1} \iota_n} (2 \, \Sigma_{\rm r})^{-1} \, \iota_n = \frac{1}{\iota_n^T \, \Sigma_{\rm r}^{-1} \, \iota_n} \, \Sigma_{\rm r}^{-1} \, \iota_n \tag{3.7}$$

The left hand side or the denominator of Formula 3.3 represents the normalisation and ensures all capital will be invested. The right hand side ensures that the indices with a high variance get a low weight and those with a low variance get a high weight (by inverting them).

The explicit solution of the optimal portfolio as defined in Formula 3.3 does not take the long only assumption into account. In the quadratic programming problem as programmed in MATLAB a constraint is included to prevent for short positions. Besides preventing for short positions, upper bounds are set regarding the maximum level of the long positions of each asset. This constraint prevents the model allocating all capital to the least risky asset only, for example in treasuries, resulting in very low returns.

Life cycle investing

The concept of life cycle investing or a target-date retirement fund is the method of automatically changing the risk profile of the investment portfolio based on the age of the investor/participant. The fund is diversified across asset classes with the feature that the proportion invested in stocks, or higher risk assets, automatically declines as time passes. By contrast, the proportion invested in (lower risk) bonds increases over time. The goal of the life cycle is a decreasing risk of the portfolio when the participant approaches its retirement age.

The rationale behind life cycle investing is based on the age of the participant. As the participant gets older, the remaining time when possible investment losses can be recovered declines. The investment risk therefore declines as a person approaches its retirement age. Besides, during the last years of someone's working life, the risk of a low pension due to a low interest rate is also reduced.

Within life cycle investing one could think of many risk profiles reflected in the asset mix, from conservative to aggressive, at each period in time. The asset mix of an aggressive life cycle has a higher risk compared to a conservative mix, however both of them use the concept of reducing risk as the retirement age approaches.

In this thesis we use one life cycle, based on the six asset classes that are defined in Section 3.1. The capital allocated to equities developed markets will be equally distributed over the four indices as mentioned before. The life cycle that will be used is indicated in Figure 9. In the so-called area graph a representation of the asset allocation will be presented. An example of the percentage of assets invested in treasuries is marked yellow. This starts around 10% but at retirement date this number increased to over 50%. By contrast, the riskier asset class equities decreased by almost 50% over the years. An overview of the exact numbers can be found in Appendix 2 – lifecycle data.

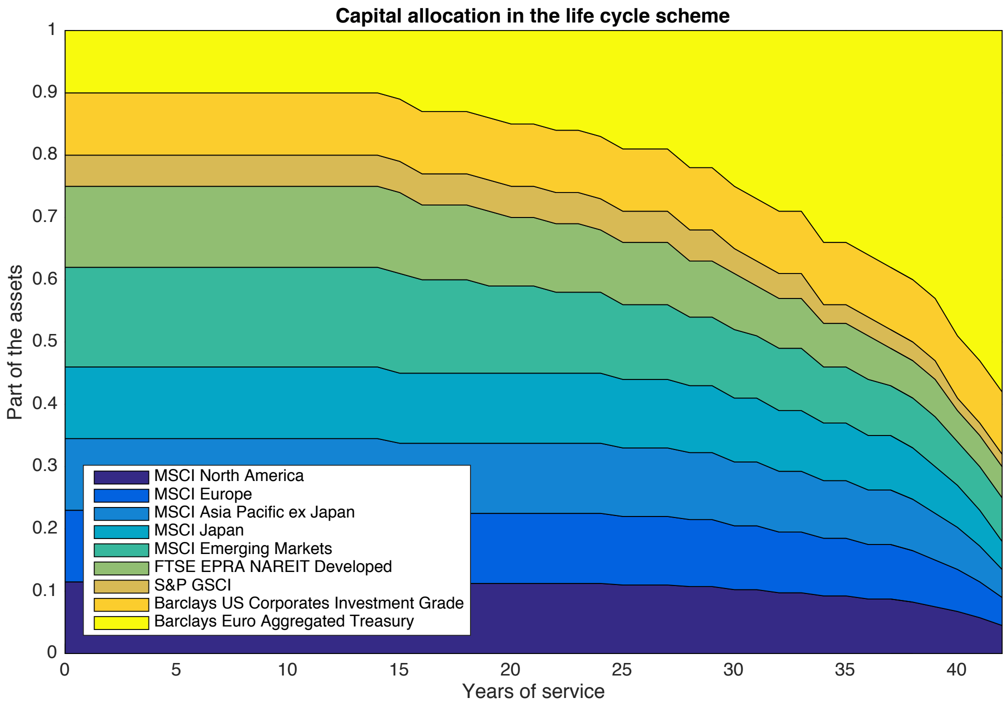


Figure 9: lifecycle capital allocation over time.

An advantage but at the same time a disadvantage of investing by the life cycle method is the reduction of riskier investments as time passes. Theoretically, higher returns are possible when taking higher risks. At the moment your capital position is optimal, and you are able to generate the largest capital growth, the allocation will be less risky aiming on lower returns and less capital growth.

3.4. Constraints

When developing the algorithm APG indicated that at least transaction costs and rebalancing criteria needed to be taken into account.

Transaction costs

Trading does not come for free. When buying or selling securities expenses need to be incurred. To implement realistic transaction costs entry and exit fees are included in the algorithm. APG uses these fees when computing expected costs of a transaction. Entry and exit costs are the sum of the costs like commissions, market impact etc. In **Fout! Verwijzingsbron niet gevonden.** (removed from the public version) the fees are displayed in basis points for several asset classes as each asset class has its own cost structure. These fees can be multiplied by the volume of the transaction to get an expectation of the costs associated with the transaction. Some classes use one fee for entry and exit, other categories face higher entry fees compared the costs related to the sale of securities. The 1-month cash option has no entry or exit fees in the computation of the transaction costs; they are negligible.

Portfolio rebalancing

Asset allocation itself is one of the major strategic decisions in the investment process. However, the decision of how to achieve this allocation in a cost-effective manner is no less important in obtaining good and consistent performance (Fabozzi, 2007). Given the current holdings, the portfolio managers need to decide how to rebalance their portfolio in an efficient way incorporating updated views on risk and return as the asset mix and the environment change over time.

There are two important aspects in portfolio rebalancing. The first aspect is robust management of the trading and transaction costs when starting the rebalancing process. The second is combining short- and long-term views. The latter is particularly important when taxes or liabilities need to be taken into account. When portfolio managers incorporate a long-term view overall transaction costs may be reduced, as portfolios do not have to be rebalanced as often.

In our algorithm we decided not to rebalance the whole portfolio each time step according to the recommended portfolio. This approach is chosen to avoid massive shifts in portfolio. This makes the algorithm more realistic but certainly also saves on transaction fees. It may be a drawback that the recommended portfolio will (almost) never be completely replicated; however the arguments against this statement prevail.

In the example below the operational functioning of the algorithm is described when it comes to rebalancing for all strategies.

Suppose there are four instruments to allocate our funds to: stocks, bonds, real estate and commodities. According to the recommended portfolio, given by a portfolio construction model, our funds need to be invested equally among the categories (i.e. 25% per asset class). Suppose our capital position is €10 then the portfolio is displayed in the table below:

Stocks	Bonds	Real Estate	Commodities	Sum
€25	€25	€25	€25	€100

Funds will be invested and for simplicity we assume a return of 5% here on each asset class. During the next time step premiums are collected and need to be invested. In this example the sum of incoming premiums will be 10.

For the new period the recommended allocation changes:

Stocks	Bonds	Real Estate	Commodities	Sum
30% (+5%)	26% (+1%)	24% (-1%)	20% (-5%)	100%

In monetary terms this is reflected in the following portfolio, now based on a total capital position of €115 (€100 initial capital, €5 return and €10 premium):

Stocks	Bonds	Real Estate	Commodities	Sum
€34,50	€29,90	€27,60	€23	€115

The actual positions:

Stocks	Bonds	Real Estate	Commodities	Sum
€26,25	€26,25	€26,25	€26,25	€105

Compared to the actual positions there is a difference with the recommended allocation. According to the recommended allocation we should invest more in stocks (+€8), bonds (+€3.6) and real estate (+€1.4) and less in commodities (-€3).

Stocks	Bonds	Real Estate	Commodities	Incoming premiums
€8,25	€3.65	€1.35	- €3,25	€10

The incoming capital will be allocated to the asset class that has the biggest difference with its recommended level. That means, when €10 capital needs to be invested, €8,25 will be invested in stocks. Then €1,75 capital remains and will be invested in bonds during this period. This means no exact replication will be conducted, but an approach will be used instead.

This example shows the allocation for a positive inflow of money during an accumulation phase of the fund. In a distribution or pay out phase, outgoing benefits will be larger compared to incoming contributions and divestments will be needed. This will occur using the same method and signs become negative.

4. Data generation and Testing

In Chapter 4 we will show how the time series were created that form the input for the portfolio construction models out of Chapter 3. First, we take a closer look at the participants of the pension fund. What parameters do we take into account, how long do they live and how will their pension levels be determined? As making a statement on the performance is the objective, these numbers are relevant when starting our test phase.

4.1. The participant

This section focuses on the participant of the pension scheme, how its life expectancy is computed and how contributed pension capital is transformed to an annuity.

When simulating certain developments of a complete pension fund a reflection of the associated population can be used, for example a sample population provided by Statistics Netherlands (CBS). However since in this thesis the focus is on individual participants also an individual approach is chosen. That means no complete population is simulated. This gives the most pure results in terms of replacement rates since the replacement rate of participants entering the scheme at any point in time after the start of their career will only create noise and will bias the results.

Over time, participants in a pension scheme change. Therefore it is important to share the assumptions that were made according to the characteristics of the participants. The focus will be on active participants (i.e. employed people). The participant will start working at the age of 25 and will retire at 67. He or she can decease every year and has a life expectancy that is determined according to the survival tables explained more in depth in the next section. Premium payments are made from the start of someone's career until retirement date. It is assumed that the participant stays employed full time up to retirement.

People in a pension scheme do not only become older as time passes. On an average level they also face raising wages due to a career path and by Collective Agreement (CAO) increments. These numbers significantly influence the deposits people make for their pensions. First of all the salary of the participant needs to be determined. According to the CPB Netherlands Bureau for Economic Policy Analysis (CPB, 2016), the average gross income is set at €36.500 for the year 2016. As this might be too high as an average initial salary this number is set at €25.000 for our participants. A yearly 2% Collective Agreement increment is considered as well as an age-dependent salary increase. From 25 to 35 a yearly salary increase of 3% is considered, from 35 to 45 it is 2% and from 45 to 55 1%. Above this age no age-dependent increases in salary are expected anymore. The Collective Agreement increments continue until retirement.

As the product in which people invest is a DC product, a fixed percentage of annual wages can be used in the determination of each person's premium level. In this thesis a number equal to 18% is used. These numbers were obtained from the actuarial business unit of APG.

Survival tables

The development of survival probabilities, and therefore the life expectancy, is an important factor in actuarial calculations. The Actuarial Association (Actuarieel Genootschap, AG) is the professional association of actuarial professionals in the Netherlands and presents estimates of survival probabilities once every two years.

Over the past 50 years life expectancy has increased by about two years each decade (AG, 2016). This means that each generation will live five years more than the previous. This trend has an impact on society and of course on pension funds and life insurers. They need a continuous insight on this development to fulfil their responsibility towards society as well as possible. The AG provides the sector of statistical information through the publication of these forecast tables. An example of a part of the table is shown in Table 3.

Table 3: part of a mortality table for males (source: Actuarieel Genootschap).

Age	2014	2015	2016	2017	2018	2019	2020	2021
67	0,014654	0,014300	0,013956	0,013619	0,013291	0,012970	0,012657	0,012352
68	0,016266	0,015875	0,015494	0,015123	0,014760	0,014405	0,014060	0,013722
69	0,017964	0,017532	0,017111	0,016700	0,016298	0,015906	0,015524	0,015151
70	0,019655	0,019178	0,018713	0,018259	0,017816	0,017384	0,016962	0,016551
71	0,022078	0,021556	0,021046	0,020548	0,020062	0,019587	0,019124	0,018671
72	0,024710	0,024131	0,023566	0,023014	0,022475	0,021948	0,021433	0,020931

Table 3 has two dimensions. The age can be found on the rows while the year is found at the columns. The numbers in the table represent mortality probabilities. It may not be a surprise that the probabilities increase while age increases. However, a person aged 70 today will have a lower survival probability compared to a person aged 70 next year. This reflects the increasing life expectancy of each generation.

By using the survival tables we computed the remaining life expectancy of the participants in our population. We find that the remaining life expectancy of males that retire this year is about 18 years compared to 21 years for females. This gradually increases to 24 and 26 years for males and females in 40 years from now.

At retirement date the participant of a pension scheme will buy an annuity that will pay monthly amounts during the rest of his/her life. To determine the height of these payments we use the remaining life expectancy.

From annuity to replacement rate

An annuity is an asset that pays a fixed sum each period for a specified number of periods. Mortgages and pensions are common examples of annuities. Brealy et al. (2010) describe a general formula for the value of an annuity that pays €1 a year, for *t* years with interest rate *r*, starting in year 1 in Formula 4.1:

$$\frac{1}{\mathrm{r}} - \frac{1}{r(1+r)^t} \tag{4.1}$$

This expression is also known as the *t*-year annuity factor. The level of the yearly pension payments can be computed by dividing the accumulated pension capital by this annuity factor.

The two important variables are the interest rate and the duration of the pension payments. The way to determine the duration has already been explained in the previous section. Since the interest rate is assumed to be constant, one value needs to be found. A yield curve will be used to find this value.

The determination of the interest rate provides an interesting question, as the horizon for pension investing is 40 plus years. That means the, for example, 18 year interest rate for a male retiring this year will be known. However, for a female retiring 40 years from now the yield curve of that moment is not known yet.

Therefore some assumptions need to be made. One option is using an interest rate model that estimates a yield curve and its projections of the future. However, it is debatable what the reliability of these values will be in this case. After all, how well can someone estimate an 18-year interest rate in (more than) 40 years from now?

Keeping this remark in mind any representative value needs to be used. Simulating interest rates can be an option, however we prefer a fixed interest rate. The interest rate is a very important factor and by using a fixed rate all results are affected by the same fixed rate. For the interest rates the nominal interest rate term structure (zero coupon) for the Financial Assessment Framework (FTK) will be used. This interest rate term structure is published by the Dutch Central bank and based on the swap curve with the intention to determine the actual value of pension liabilities. The interest rate term structure present interest rates for durations from 1 to 100 years. In this thesis the interest rates published on July 31st 2016 will be used and are included in Appendix 3 – Interest rate.

For a female retiring 40 years from now the 66 year rate will be used. These 66 years are the combination of 40 (years from now) and the 26 years remaining life expectancy at that time. Now the question remains how reliable these predictions are. The interest rates change each month and are therefor subject to significant change. An illustration of the term structure that will be used here is shown in Figure 10. For the purpose of comparison also a term structure of a year ago will be presented. Significant changes can be observed in these rates, and, as will be seen later, these interest rates will have a major impact on the value of the pension benefits and replacement ratios. The probability of this changing interest rate is an important limitation of the research. To be able to estimate the size of the effect of a changing interest rate a sensitivity analysis is included in Section 5.7 where the level of the interest rate is varied.

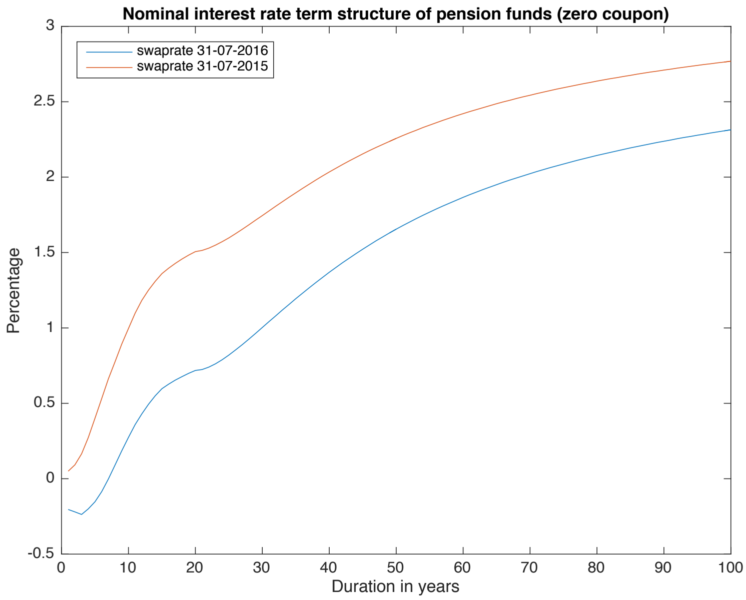


Figure 10: Nominal interest rate term structure of pension funds (zero coupon) (source: DNB).

By dividing the final level of the investments, the accumulated capital, by the annuity factor we find our annual cash flow. We find the replacement rate by adding this amount to the annualized first pillar pension (AOW) and dividing it by the final annual salary. This is shown in Formula 4.2 and 4.3.

$$\text{Annual cash flow} = \frac{\text{final investments}}{\text{annuity factor}} \tag{4.2}$$

$$\text{Replacement Rate} = \frac{\text{Annual cash flow} + \text{annualized first pillar pension (AOW)}}{\text{final annual salary}} \tag{4.3}$$

4.2. Financial markets

In this section we provide an explanation how the financial markets will be simulated. When testing the algorithm on financial market data the performance of the portfolio construction models can be assessed. In the ideal situation the algorithm should be tested on real historical data. This method cannot be used when a pension horizon of over 40 years needs to be taken into account, as the time span of representative historical data is too short. Therefore we will use two other methods to create new, longer time series based on our historical data set. We will use a geometric Brownian motion supplemented with several historical simulation approaches to simulate our indices. Separately, we will use the Vašíček method to generate simulations for the interest rate.

Finally we will test the models on the available historical dataset. Simulating over a long time horizon has the disadvantage that the corresponding results have to deal with increasing uncertainties. Simulated time series may vary over long periods. By exposing our models to a shorter period of real data we cannot draw a conclusion on replacement rates but yearly returns can be assessed. As our historical data set contains data of about 14 years it is still possible to some reasoned judgements. As the true performance of APG is available a genuine test of our models can be carried out. Testing on a shorter period of time will also deliver reliable results. The Quantitative Equities department of APG also use it in their evaluation of new developed strategies.

Historical simulation

The key assumption underlying the historical simulation approach is that history is, in some sense, a good guide to the future (Hull, 2012a). More precisely it is that the empirical probability distribution estimated for market variables over the last few years is a good guide to the behaviour of the market variables over the next day. Therefore this approach will be used to simulate time series that can be used as input when testing the algorithm. The behaviour of market variables is not stationary. Sometimes the volatility of a market variable is high; sometimes it is low. That is why extensions on this model will be used that will be explained in the subsequent sections.

In Chapter 3 we introduced our data sample. The movements in the market variables deliver 172 alternative scenarios for what can happen between today and tomorrow. By randomly drawing a scenario (with replacement) and repeating this for a certain number of times several new time series can be generated. An example of a simulated price path is shown in Figure 11.

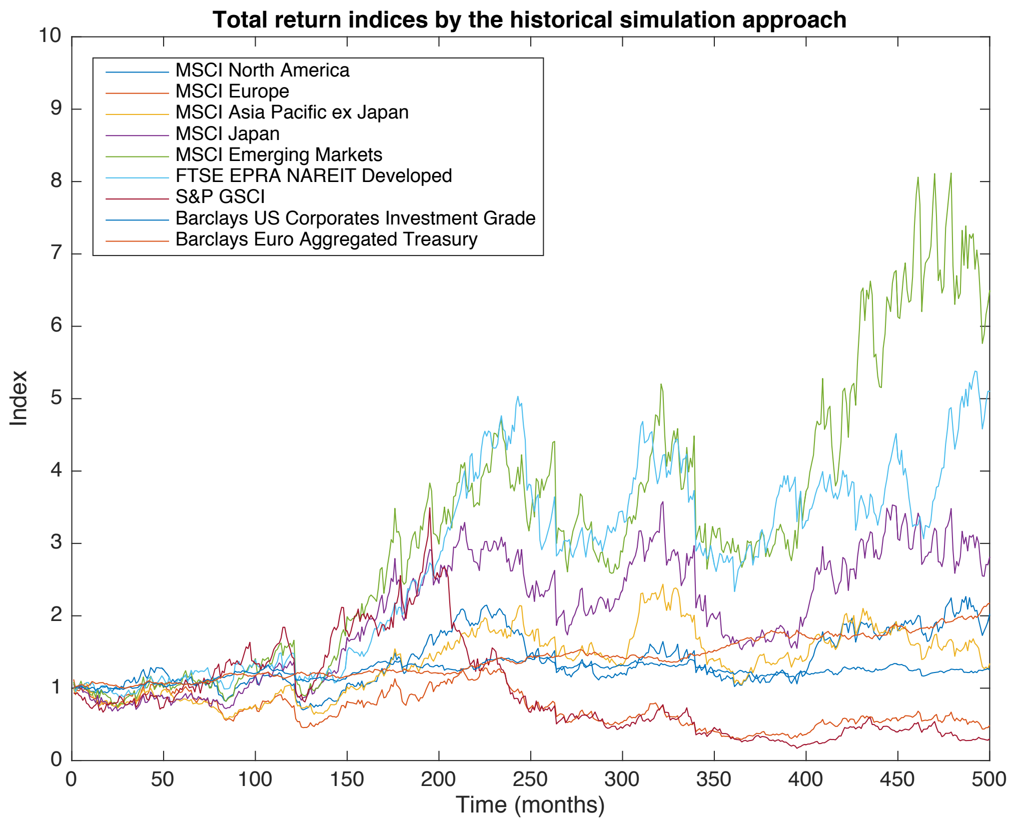


Figure 11: an example of simulated total return indices by the historical simulation approach.

The variance in returns can be quite large over time when applying historical simulation as we included the 2008 financial crisis in our dataset. When using historical simulation it is possible to randomly draw several times return figures out of that period. It is a helpful method to test the algorithm on its robustness however large drops in the value of the indices can be observed frequently, probably more frequently than in practice. This should be taken into account when the results are evaluated.

Extensions of the historical simulation approach

Exponentially weighted moving average

When looking at the basic historical simulation approach, no particular probability distribution is assumed regarding the asset returns. The assumption is made that every day in the past is weighted equally. Boudoukh et al. (1998) show that more recent observations should be given a higher weight because these observations are more reflective of current volatilities and current macro-economic conditions. Here an exponential weighted moving average (EWMA) approach can be used.

The weighting scheme as used in Hull (2012a) is one where weights decline exponentially: the weight assigned to scenario 1 (which is the one calculated from the most distant data) is λ times that assigned to scenario 2. This in turn is λ times that given to scenario 3 and so on. The weight given to scenario *i* is equal to:

$$\frac{\lambda^{n-i}(1-\lambda)}{1-\lambda^{n}} \tag{4.4}$$

When lambda approaches 1 the basic historical simulation is reached again, where all observations have a weight *1/N*. For lambda the value 0.97 is taken. By choosing this value more weight is given to recent observations, while still all older scenarios will be included in our exponential weighted moving average historical simulation.

Exponentially weighted moving average based on the interest rate.

Interest rates are an important factor in the pension industry. As an alternative EWMA approach we include the interest rates in a historical simulation approach. We do not rank returns based on time as used in the previous paragraph, but rank asset returns based on the corresponding interest rate at each time step. Returns in a period of low interest rate are expected to be more reflective of current market circumstances. Here lambda is set at 0.97 again. A simulation is executed by giving more weight to recent, low interest rate scenario's compared to scenarios with higher interest rates.

Geometric Brownian motion

As mentioned in the previous section stock price behaviour has to be modelled. The stochastic process usually assumed for a stock price is the geometric Brownian motion (Hull, 2012b). Under this process the return to the holder of the stock in a small period of time is normally distributed and the returns in two non-overlapping periods are independent.

The model is widely used in academics however it has some disadvantages as well. In practice, stock prices show jumps but in the GBM the stock path is assumed to be continuous. As the method is used for large time steps (of a month) this effect is significantly less important compared to a situation when the time step would be smaller. Another disadvantage is the volatility that is assumed to be constant, while it changes over time in practice. Compared to the historical simulation method, the variance of the returns is much smaller when using a geometric Brownian motion. This result may not be a surprise as one of the input parameters of this method is the standard deviation.

The discrete-time version where the change in the stock price ΔS in a small time interval Δt is given by (Hull, 2012b):

$$\Delta S = \mu S \Delta t + \sigma S \varepsilon \Delta t \quad (4.5)$$

Here μ is the expected rate of return per unit of time for the stock and σ is the volatility or standard deviation of the stock price. For μ and sigma the values explained in the previous section are implemented in the model. The stochastic variable ε follows a standardized normal distribution ϕ $(0,1)$. The parameters are assumed to be constant over time.

The process as described above models the stock price of one individual stock. In our simulation multiple instruments will be used that are correlated with each other. The random number epsilon therefore is not equal for each instrument. The correlation between the instruments will be included in the computation of the value of ε. This is reflected in the MATLAB function '*mvnrnd*' that delivers (a matrix of) random vectors chosen from the multivariate normal distribution with mean vector μ and covariance matrix sigma. μ is a vector of zeros (having a length which is equal to the number of instruments) and σ is the variance covariance matrix of our assets. This results in a matrix with a ε for each instrument at each time step.

The discrete-time version of the change in stock price that can now be used is:

$$\Delta S = \mu S \Delta t + S \varepsilon \Delta t \quad (4.6)$$

Interest rate simulation

Several paths for the various indices are simulated by the approaches as described above. For the mean variance optimisation model the risk free interest rate should also be known and hence interest rate paths need to be simulated. Hull (2012b) considers several one-factor equilibrium models that can be used to simulate interest rates. In this section three of them will be evaluated:

- Rendleman and Bartter model (1980)
- Vašíček model (1977)
- Cox, Ingersoll and Ross (CIR) model (1985)

Hull (2012b) describes that in a one-factor equilibrium model, the process for the interest rate involves only one source of uncertainty. An Itô process represented by means of the following expression describes the risk-neutral process for the interest rate:

$$dr = m(r)dt + s(r)dz \quad (4.7)$$

The drift, represented by *m*, and the standard deviation, *s*, are assumed to be functions of *r* but are independent of time. A one-factor model implies that all rates move in the same direction over any short time interval, but not that they all move by the same amount. The shape of the zero curve can therefore change with the passage of time (Hull, 2012b).

One important difference between interest rates and stock prices is that interest rates appear to be pulled back to some long-run average level over time (Hull, 2012b). This effect is known as mean reversion and implies a negative drift for high interest rates and a positive drift for low interest rates. The Rendleman and Bartter model (1980) does not incorporate mean reversion and therefore it is not chosen to implement here.
Both the Vašíček (1977) and the CIR (1985) model seem to be appropriate models for the purpose we want to use them. However, under the CIR model interest rates cannot become negative where under Vašíček this is possible. Keeping the current interest rates in mind the Vašíček model is chosen to simulate interest rate paths.

The Vašíček model describes the risk neutral process for *r*. Actually; this process should only be applicable in the risk neutral world where we use it as a real world simulator. The process for *r* is described by:

$$dr = a(b - r)dt + \sigma dz \quad (4.8)$$

In the formula, a, b and σ are constants. The interest rate is mean reverting to level b at rate a. These variables indicate that the interest rate reverts to b% with a reversion rate of a%. Typical values for a are in the range 15 to 20%, where the reversion rate a is assumed to be 15%. The assumption is made to revert to 0%. Typically interest rate assumptions are 2% or 4%. This assumption is based on the mean value of the time series of interest rates that is introduced before and it is the most realistic assumption at the moment. From our data sample the standard deviation is calculated and used as the sigma in the Vašíček model. Figure 12 shows an interest rate path as simulated by the Vašíček model.

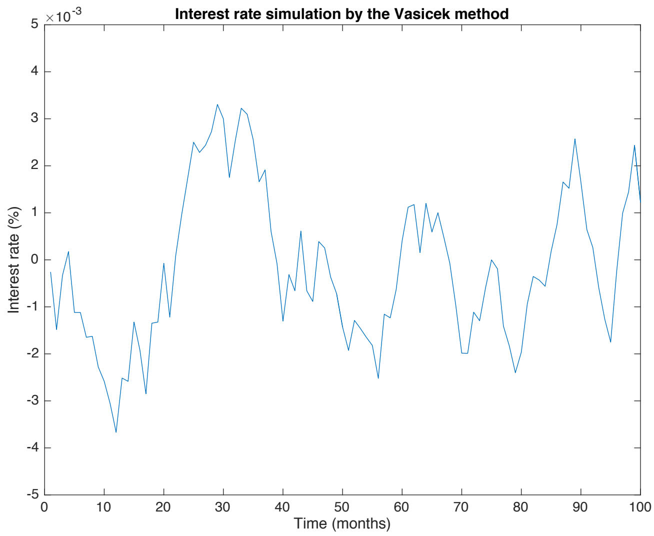


Figure 12: an example of an interest rate simulation by the Vašíček method over the first 100 months.

4.3. Testing

By running a number of simulations we will test the performance of the algorithm. Each simulation starts with the creation of a new time series. These time series will be created by the two methods as explained in Section 4.2. This will be the input for the portfolio construction model. Based on the movements of the financial markets the model will select several investments. However, in case of the life cycle model it is the age of the participant that determines the asset allocation. As it is the goal to determine which portfolio construction model delivers the best results all models will be run on the same data sets.
Based on the performance of the models and the characteristics of the participant the algorithm delivers a certain capital position at the participants' retirement date, which will be transferred to an annuity.

When determining the number of simulations an assessment is made based on the computing time versus the difference in results when increasing the number of simulations. When performing 1000 simulations (i.e. 1000 time series are generated), each with three participants delivered acceptable results.
Three participants per simulation are used to decrease the probability of losing the value of a time series when one or two participants do not reach their retirement date. This is relevant as more than 10% of the participants do not reach their retirement age and therefore many simulations would be useless.

5. Results

In Chapter 5 we will present the results and we will provide an interpretation of the findings of the simulations. We answer the fourth research question on the performance of the algorithm in this chapter. As the algorithm computes expected annuity levels in several scenarios the results will (mostly) be presented in terms of the replacement rate: the variable of our interest. The results of the simulations based on real historical data will be presented in terms of returns.

5.1. The accumulation of pension capital

Before being able to determine the value of a participant's annuity we first need to compute its final capital position. Each participant contributes a certain amount of money to its pension scheme every month after which these funds will be invested. In Figure 13 the development of a participant's capital position is shown. The exponential increasing line is the cumulative amount of the contributions; the sum of each months contributions. It is exponential because the participant's contribution increases each year. The line above is the sum of the contributions and the investment returns. At retirement date the total capital position of a participant will be withdrawn and invested in an annuity.

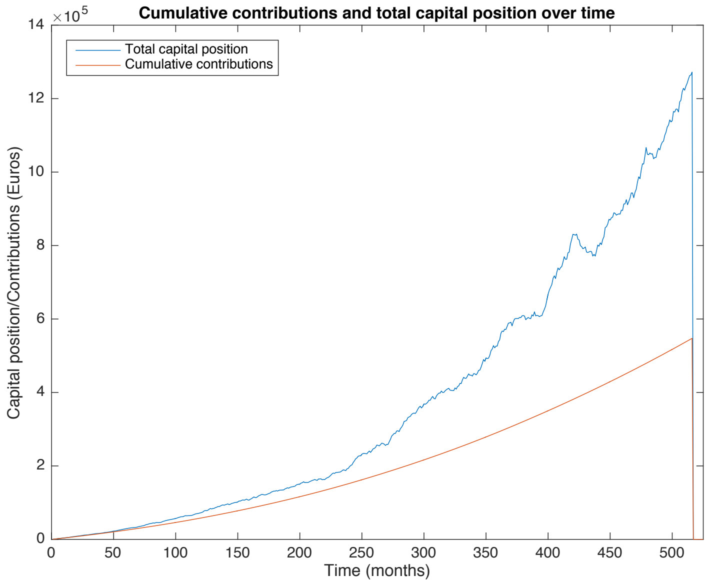


Figure 13: the total capital position increases faster compared to the contributions.

5.2. Evaluation criteria and benchmark

The result of our simulation is a distribution of replacement rates. Those distributions will be presented in histograms in the next paragraph. Next to the visual presentation, several statistics will be presented to assess the numbers. These will be the mean value, the median, variance and a 95% Value at Risk. The expected value is the most frequently used measure of a variable's central tendency. If one or several values of the variable are either much smaller or larger than all others the value of the true mean value can be distorted. It no longer reflects the center of the distribution in a meaningful way. Another measure of central tendency that is not sensitive to outliers is the median, the midpoint of a distribution of values in a way that there is an equal probability of falling above or below the value. As simulations are made based on two methods of which at least one, historical simulation, is very sensitive to outliers we both use the mean and median value when presenting the results.

A second feature of a distribution that needs further investigation is its dispersion. That is, the extent to which its values are spread out. When describing our results the variance will be presented.

The best result will be a mean value at or above 70% with a low variance. That means the results are stable and at the right level. Of course, replacement rates of above 70% are always appreciated but the downside must be limited. To measure this downside a 95% Value at Risk is included to monitor the tails of the distribution. This measure enables making an estimation of the risk of the chosen portfolio construction model.

To assess the results of the algorithm we will use several methods to compare our results with a benchmark. The most important benchmark is the performance of ABP, as it is the largest customer of APG it is a good indicator of the performance of APG Asset Management. A valuable solution should at least approach APG's performance to some extent. Another benchmark described in literature is one of the portfolio construction models, the 1/N model. This model naively distributes capital over all assets and therefore gives a reference level of a selection criterion compared to 'doing nothing'.

Again we will need to create time series in order to be able to include the entire future time span. We will use the historical simulation and the geometric Brownian motion technique as well for this purpose.
Starting with historical simulation, now the monthly net returns over the almost 15 year period of APG are ranked (e.g. on interest rate). Where the individual returns of the instruments were selected to create new time series, now the corresponding APG returns will be administered simultaneously as well. This gives the possibility to monitor APG's performance in the same period as the return data that will be used as input for the algorithm.

To compare the performance when using Geometric Brownian Motion as a technique to create new time series the way APG's results are included needs to be changed. With the parameters extracted from the APG return data the mean and variance can be used as input variables when creating new time series according to the Geometric Brownian motion methodology.

5.3. Histograms and descriptive statistics

In Section 5.3 we will present the results of the simulations. In the figures several histograms and the related statistics are presented. They will be discussed per simulation method as four different techniques are used. To keep focus we go through the results by comparing the best performing model and compare it to our benchmark.

To facilitate reading the statistics are colored, from green, yellow to red, as an indication of the positive or negative value of the outcome. The corresponding histograms of each model are presented after these tables. When assessing the histograms it is important to take the scale of the different axis into account as they may differ from figure to figure.

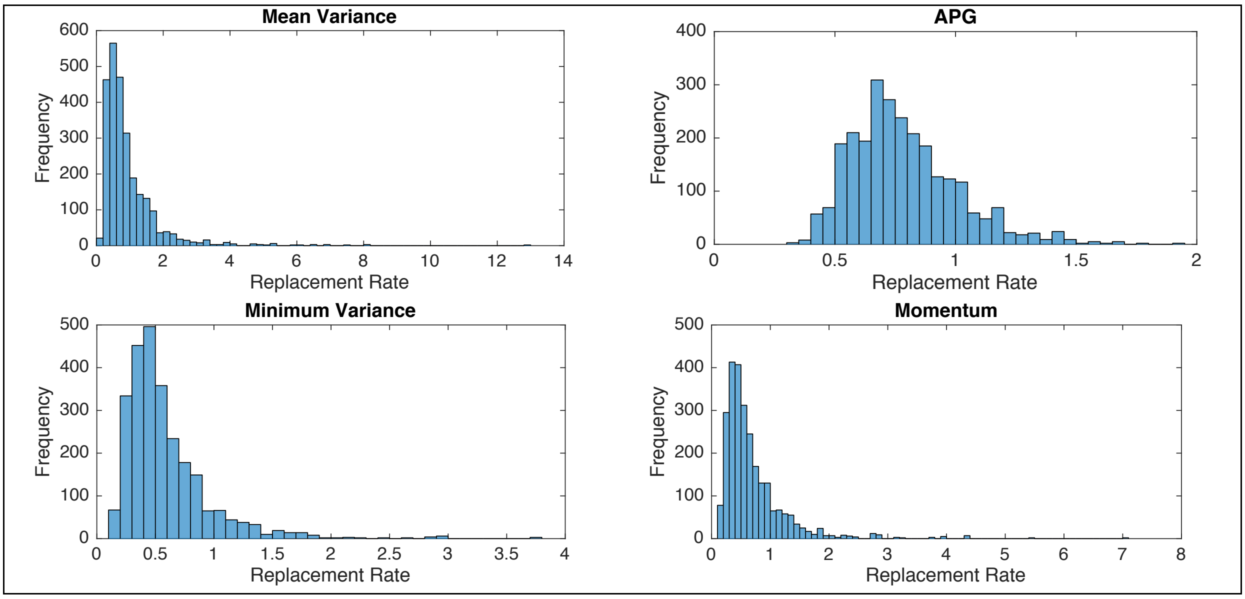
Historical simulation

The historical simulation approach immediately shows the solid performance of APG, our benchmark. The replacement rate is on the right level supplemented by a low variance and an appropriate Value at Risk is reflected in the completely green column in Table 4. Very stable results, meeting the objective of a pension investor. Considering our results the mean variance model shows the best mean and median values, however the variance is quite large and Value at Risk are not satisfactory. The large variance can be observed in the corresponding histogram where several outliers can be recognized. Compared to APG, it seems that all models fail due to the substantial probability on a (very) low replacement rate.

The high variance and low Value at Risk can at least partly be attributed to the simulation method and the data set that includes the collapse of the financial markets as well as the recovery afterwards. This results in the possibility of extreme scenarios with large variances as a result. The low Value at Risk can be the result of a lot of pessimism in a scenario due to the years of economic crisis. This needs to be taken into account when these numbers are assessed.

Table 4: statistics of the historical simulation approach.

Model	Mean Variance	Minimum Variance	Momentum	Lifecycle	Mean Reversion	1/N	APG
Mean	95,0%	58,9%	69,4%	59,2%	60,7%	55,2%	78,6%
Median	80,7%	49,1%	53,5%	46,4%	47,7%	44,8%	74,6%
Variance	80,7%	14,0%	32,0%	18,8%	19,8%	14,6%	5,0%
95% VaR	27,4%	22,6%	22,1%	22,5%	20,8%	20,1%	49,2%



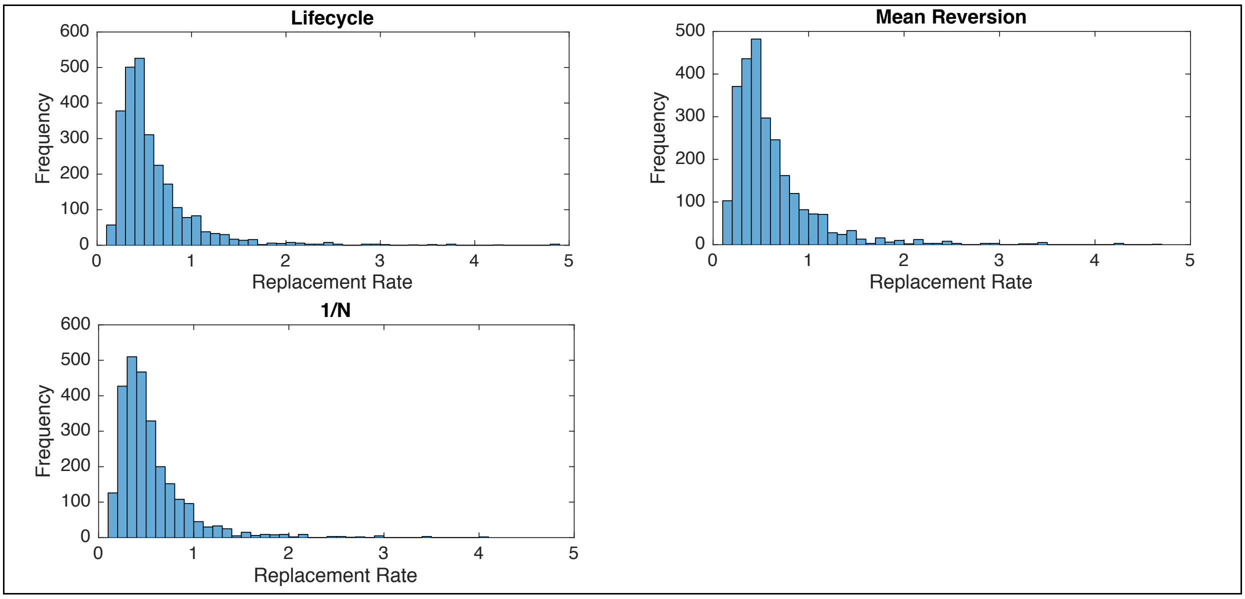


Figure 14: histograms presenting the replacement rates by using the historical simulation approach.

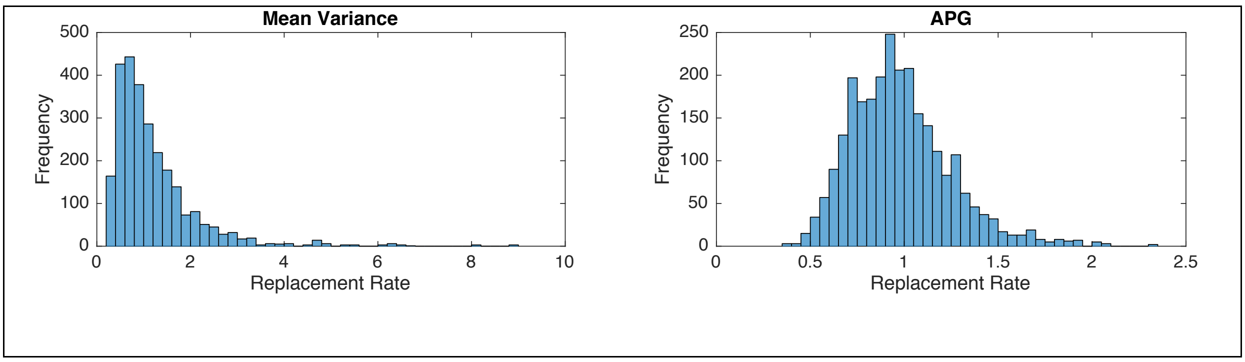
Historical simulation – Exponential Weighted Moving Average

By giving additional weight to more recent observation the results of the historical simulation change. Again, the mean variance model shows the best results. Compared to the basic historical simulation the mean/median values increased significantly while the variance only increased a bit. The fact remains that the variance is large. The Value at Risk increases but is still not at an appropriate level. The results of the benchmark APG are excellent: high expected values supplemented with a high Value at Risk and a low variance.

The increased results could be the result of the 2008 crisis that is less represented in the data by using the exponential weighted moving average.

Table 5: statistics of the historical simulation approach by using an exponential weighted moving average.

Model	Mean Variance	Minimum Variance	Momentum	Lifecycle	Mean Reversion	1/N	APG
Mean	120,5%	64,3%	47,9%	41,0%	37,4%	31,9%	98,0%
Median	95,0%	53,4%	41,0%	35,9%	32,6%	27,9%	94,8%
Variance	85,9%	16,9%	7,0%	4,0%	3,9%	2,6%	7,1%
95% VaR	38,3%	24,9%	21,3%	20,2%	17,6%	15,3%	61,3%



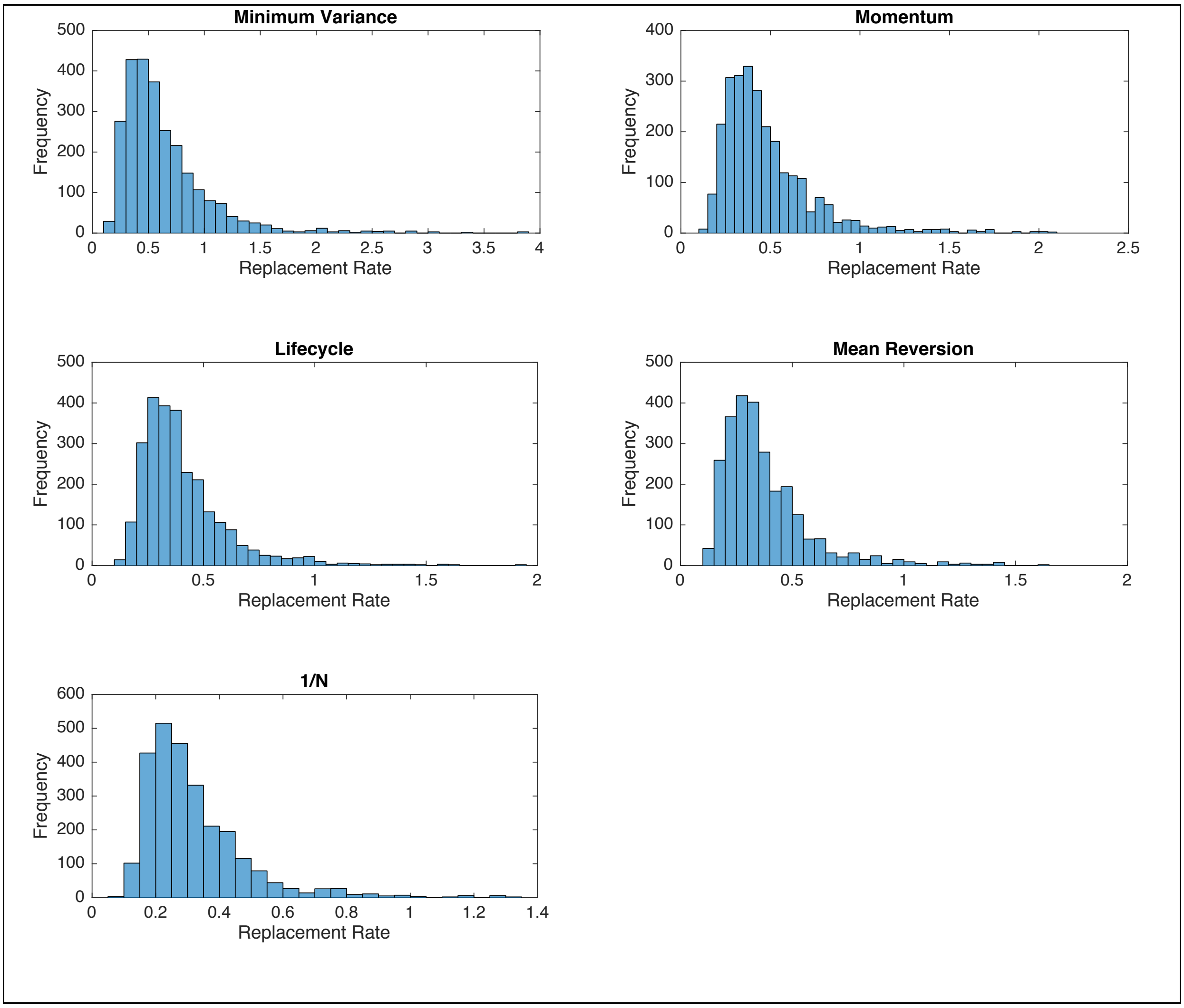


Figure 15: histograms presenting the replacement rates by using the exponential weighted moving average historical simulation approach.

Historical simulation - Exponential Weighted Moving Average on the Interest Rate

When giving more weight to scenarios with a low interest rate even higher numbers are the result. The data is sorted on increasing interest rates with most emphasis on low interest rates. It turns out that the crisis is moved backwards, i.e. more weight is given to data points after the crisis. That will be the reason for the higher results.

The mean and median value of the mean variance model increase to an extreme level. It is the high Value at Risk that is the reason to choose this model as our favourite. The others also show good results, with lower variances, however it is the Value at Risk that falls a bit behind. Also here, APG presents good results. Very adequate expected values combined with a really high Value at Risk and a low variance rate.

Table 6: statistics of the historical simulation approach by using an exponential weighted moving average on the interest rate.

Model	Mean Variance	Minimum Variance	Momentum	Lifecycle	Mean Reversion	1/N	APG
Mean	171,0%	97,0%	81,1%	65,4%	63,1%	53,9%	109,4%
Median	138,2%	81,1%	66,8%	55,7%	52,8%	46,3%	68,6%
Variance	147,3%	34,9%	25,2%	12,3%	13,7%	8,3%	9,2%
95% VaR	55,7%	38,6%	31,7%	30,3%	27,3%	24,1%	68,6%

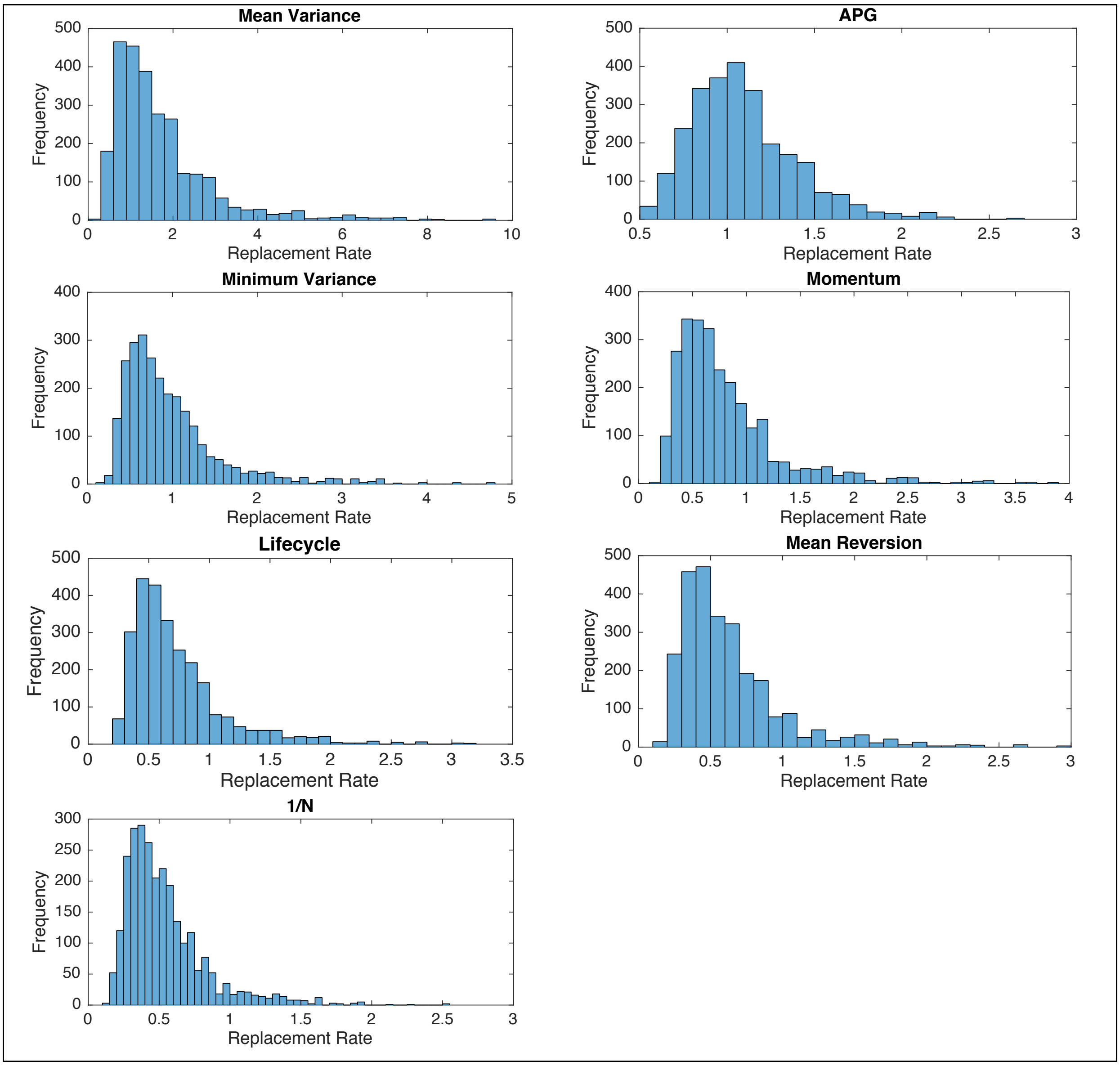


Figure 16: histograms presenting the replacement rates by using the exponential weighted moving average historical simulation approach based on the interest rate.

Geometric Brownian motion

The geometric Brownian motion is a complete different simulation method but in general the results are similar in terms of best performing models. The mean variance model shows the best numbers but they are disappointing compared to the previous presented results. APG does well again, but does not convince anymore. Our model as well as APG shows larger variance numbers and an insufficiently low Value at Risk. The underlying problem may lie in, once again, the crisis years. This makes the average yield, which is an important factor regarding the drift rate, beneath expectations. Consequently this effect can be seen in the replacement rates.

Table 7: statistics of the geometric Brownian motion simulation.

Model	Mean Variance	Minimum Variance	Momentum	Lifecycle	Mean Reversion	1/N	APG
Mean	59,9%	40,7%	45,5%	41,1%	40,7%	37,5%	59,5%
Median	46,1%	35,8%	36,2%	34,0%	33,3%	31,6%	47,8%
Variance	24,5%	4,2%	9,8%	5,8%	6,3%	4,5%	17,6%
95% VaR	21,3%	19,6%	18,9%	19,9%	17,9%	17,0%	19,9%

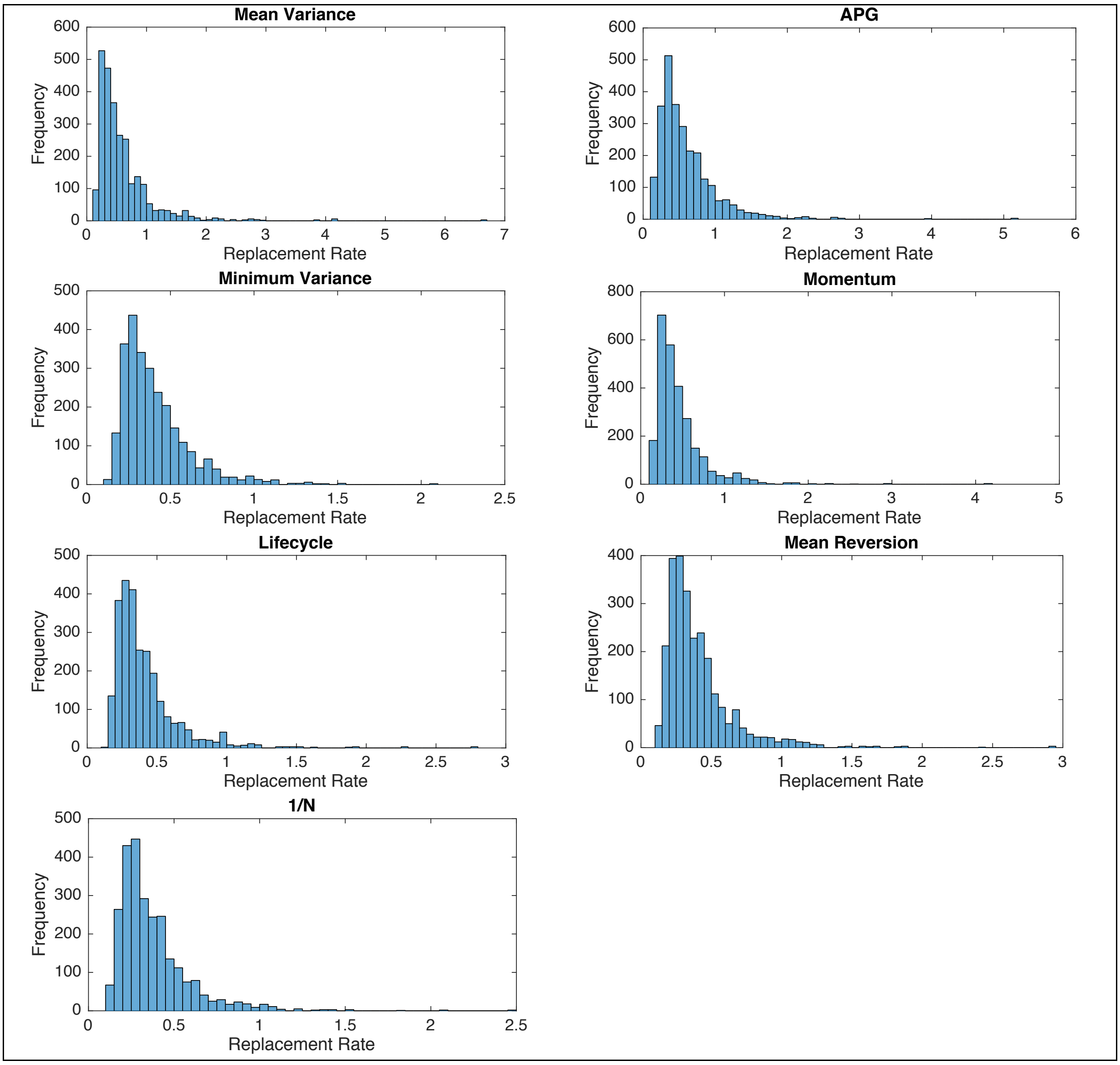


Figure 17: histograms presenting the replacement rates by using the geometric Brownian motion.

5.4. Simulation on real data

APG uses an average long-term return expectation of about 7% a year. In the 12-year period as used here a return of 6.1% is obtained. A bit below the long-term average, however the financial crisis of 2008 is included in the data set. The performance of our investment models, including the performance of APG is presented in Table 8. Again the mean variance model shows the best scores in terms of returns. It outperforms APG by a significant difference. By using this data set the momentum model also demonstrates good results as its returns are slightly above those of APG. This shows that a momentum strategy is more effective in the real markets than on simulated data. The minimum variance and 1/N model can be found at the lower part of the table as their returns remain behind with respect to the other models. This is a logical result, as those models do not explicitly focus on generating returns.

Table 8: performance of the investment models and APG over the 12-year period.

Mean Variance	12,0%
Momentum	6,8%
APG	6,1%
Mean Reversion	5,6%
Lifecycle	5,3%
1/N	4,9%
Minimum Variance	4,5%

The results of our mean variance model are superior to APG's performance in this test on 12-year historical data. When assessing the performance of the simulations we had the possibility to judge more numbers like the variance of the results or a Value at Risk.

5.5. Smoothing of the results

The results presented in Section 5.3 can be smoothed to some extent. Participants receiving a large pension in terms of replacement rates could compensate people whose investments have performed worse. This theme is currently under discussion in politics. The SER (2016) uses the metaphor 'good' weather and 'bad' weather to describe the financial markets. In the bad weather scenario the investment returns are low. Consequently the risk of low pension benefits increase. In the exploration of the SER the committee investigated if sharing the results of good and bad periods ensures more stable pension benefits. The findings of the committee show that a buffer, a joint reservation for young and old, is the most effective way of sharing the investment risk. In times of well performing financial markets the buffer is filled, while a part of the excess returns are excluded from the individual pension capital in favor of the buffer. If returns are low, the individual capital is supplemented with funds from the buffer. The buffer cannot become negative, so no shortfall can be passed on to next generations.

Furthermore the results only include pension capital that is contributed by the participant. In practice, when someone does not reach its retirement date the contributed capital is divided over the fund. This effect is not taken into account here but is significant.

5.6. Asset management costs

One important element in the Basis Zero experiment is the maximum asset management cost of 10 basis points a year. In the results in terms of replacement rates the transaction costs are already included. When separating the transaction costs of the best performing portfolio construction model (mean variance) the costs meet this requirement. In the mean variance model the transaction costs are on average 3,64 basis points a year over the capital position of a participant.

Table 9: transaction costs of the models.

Method	Transaction costs (basis points)
Historical Simulation	3,87
Historical Simulation – EWMA	3,40
Historical Simulation – EWMA Interest Rate	3,35
Geometric Brownian motion	3,95
Average	**3,64**

The transaction costs are low due to investing small amounts of money each period over a continuous increasing total capital position. Therefore the costs are relatively high in the beginning and decline as time passes.

Costs of Asset Management are not only in transaction fees. Therefore total costs are not covered with (on average) 3,64 basis points. However it is not easy to give a final number of total asset management costs in the Rottumerplaat scenario of the Basis Zero algorithm. Based on the assumption that transaction costs are the largest cost item in an automated product, costs should be able to stay within 10 basis points.

5.7. Sensitivity analysis

Pensions are very sensitive to the interest rate. This becomes clear when the value of accumulated capital is converted to an annuity. As the execution of a sensitivity analysis on all simulation methods was too time consuming and the results would be comparable, the geometric Brownian motion will be used as an example of the effect of changing interest rates on the results.

Our standard situation is computed with the interest rates of July, 31st 2016. As an indication the 10-year interest rate is provided in the table.

Table 10: results of a sensitivity analysis by decreasing or increasing the nominal interest rate.

	-1%	Current results	+1%	+2%
10 year interest rate	-0,73%	0,27%	1,27%	2,27%
Mean	52,1%	59,9%	67,4%	75,0%
Median	40,4%	46,1%	51,4%	57,3%
Variance	16,7%	24,5%	31,5%	36,5%
SD	40,9%	49,5%	56,2%	60,4%
95% VaR	19,1%	21,3%	23,7%	26,8%

It is different for each variable, but an interest rate increase creates significant rises in replacement rates. The median replacement rates increase by about 5 percent point if interest rates increase by 1%. The effect is smaller for the Value at Risk that increases by over 2 percent point.

6. Conclusions, limitations and future research

In this chapter we will present the conclusions of the research. We will answer the main research question and give recommendations based on our findings. We will also discuss limitations of our research and provide ideas for future research and the future of this project at APG.

6.1. Conclusion

The goal of the Basis Zero experiment is the development of a new basic pension product that needs to meet three requirements. It should be explicable within 10 lines; with maximum asset management costs of 10 basis points a year and its execution costs may not exceed 10 euros a year. The focus of this thesis is on the automation of asset management by developing an algorithm. The algorithm should make investments within 10 basis points costs while still achieving adequate performance. Therefore the research question was formulated in the following way:

> *Is it possible for APG to create and implement an algorithm that automatically makes investment decisions - given a certain investment universe and the Basis Zero philosophy - with a pension objective that achieves adequate performance?*

We reviewed literature on algorithmic trading after which our focus moved to the development and evaluation of portfolio construction models. We simulated the financial markets by a basic historical simulation approach and two extensions. One uses an exponential weighted moving average that gives more weight to scenarios with a low interest rate. The other gives more weight to recent data as both the low interest rate as recent data are expected to better reflect current economic conditions. Additionally we used the geometric Brownian motion approach to simulate the financial markets through another method.

We evaluated a combination of rule-based and optimisation models to assess what model could be most appropriate in a pension environment. Their performance is measured against a benchmark that is the result of APG Asset Management. The absolute performance can be compared with replacement rates that are considered adequate in literature. Achieving a 70% replacement rate is the objective, while 60% can be considered as the absolute minimum level.

By using several simulation methods we obtained different results. However, the same pattern could be recognized across the full spectrum of our results. All results show that the mean variance model achieves the best performance. The mean and median replacement rates often exceed the 70% and repeatedly are at or above the level of our benchmark APG. Although this might look attractive there is a downside. The main problem is the high dispersion of the results while a pension investor desires stable results.

'Positive' outliers cause a part of the large variance but at the same time we see that the 95% Value at Risk results are several times below satisfactory levels. A comparison to the Values at Risk of APG shows that APG is doing better in the historical simulation and its extensions.
A reason of the large dispersion could be the influence of the financial crisis in the data. This may be a cause of very volatile scenarios and Values at Risk at low levels. By giving more weight to more recent observations more stable markets are simulated. The variance remains large but the Values at Risk increase substantially. Nevertheless the results of APG can still be called very decent.

The results could be improved such that the pension levels of different generations are more balanced. Adopting a risk sharing mechanism could provide pensions that stay closer together in terms of replacement rates. When certain financial conditions are met pre-determined rules could be applied to adjust benefits and contributions to increase the Value at Risk.

When returning to our main research question we could say that it is possible, according to our results, to develop an algorithm that automatically invests funds at low costs. Investing below 10 basis points a year must be achievable. Our best performing model shows adequate performance in the majority of the results, however the risk of achieving critical results is substantial. The results of APG are on average a bit worse, but they are more stable making setbacks less severe. A conservative attitude is desirable when it comes to pensions. For participants or investors with a low risk appetite our solution is not yet ready as an alternative to APG's current practice. However, a participant that is more risk tolerant could consider this alternative option.

6.2. Limitations

The most important limitation in our research is the large set of assumptions in the development and the testing of the algorithm. Making these assumptions ensured the possibility to develop the algorithm but they could also have a significant impact on our results. Assumptions vary in shapes and sizes. From assumptions on the interest rate, the period of time that is taken into account when determining the moving averages in the models to the selection of the representative asset classes. Results also heavily depend on the methods by which new time series are created. We try to discuss some important limitations in this section.

We assumed a system based on averages. An example is the way salaries and premiums are modelled based on a gradual development till the final wage is reached. In reality some smaller or larger jumps would be included. Another example is a participant in the scheme that is expected to live as long as its life expectancy while during that time pensions are provided. In reality other risks, like longevity risk, are important but as the simplification was not detrimental for the research goal it was made.

The interest rate is a very important parameter in the pension industry. In the sensitivity analysis we have already seen that a change in interest rate of 1% could increase or decrease the expected value of the replacement rate by more than 7%. We used a yield curve to determine the interest rate, but it is debatable how reliable these projections are over a period of over forty years.

Another important input parameter is the dataset that we used as the input of our simulations. The data allowed executing the simulations as we did, but the results could significantly change when using other data points. The exponential weighted moving average already indicated large changes in replacement rates when other scenarios get more weight. Furthermore results could show a large dispersion due to the long simulation horizon.

To compute time series a historical simulation and a geometric Brownian motion were used. These models make some assumptions that are not always completely realistic. An example is the stock price in the geometric Brownian motion. We assume that the price does not make jumps; however in practice the path may include jumps. Another example is the volatility that is assumed to be constant over time, while it changes in practice. We used one correlation matrix and that means no extra attention is given to tail correlations or correlations that changed over time. In practice correlations change over time and especially in crisis situations correlations may change significantly.

We already referred to the Vašíček interest rate model in the corresponding section. To simulate the paths a model is used to simulate interest rates in the risk neutral world while we use the results in the real world. It is not exactly clear what the effect is of this difference.

When developing an algorithm in a thesis trajectory, as was the case here, it is logical to make several assumptions. Sometimes they are argued well, like several actuarial assumptions due to actuarial help of APG. By a lack of involvement or cooperation from the asset management department several assumptions on that side will be less justified. Assumptions and choices could have been more soundly based or better-connected to practice if AM had been more involved in this project. Therefore some decisions had to be made without a background in asset management. This is also a reason why the research should be seen as indicative on this topic rather than as an algorithm that can be adopted in practice tomorrow. Further involvement of asset management would be required when further steps would be taken.

6.3. Future research

In this section about the future research possibilities we want to explore the options to improve our current solution but we also want to discuss some ongoing developments in the wealth management sector related to the automation of investments.

As demonstrated by the results the mean variance model performs consistently better than the other models. Simultaneously, the high variance is also a concern. Indeed, the more stable the pension expectation the better it is. It would be unacceptable if pensions suddenly fall short. We need to bring down the variance and / or increase our Value at Risk.

We want to make a move towards a minimum variance model. However, this has the disadvantage that the focus on obtaining returns is less important and this would have the effect that yields will decline sharply. An alternative is to make use of the properties of the lifecycle. In retirement products it is used extensively and we can apply the underlying principle on mean variance model. The idea of the lifecycle is risk reduction as the retirement age approaches. This can be linked to the risk parameter that is implemented in the mean variance model. As retirement gets closer, the risks can be continuously reduced to finally reduce the variance of our results.

In order to save transaction costs we have chosen to keep invested funds invested. In general this leads to low transaction costs as confirmed by our results. By reallocating more funds we will incur higher costs, but it can be advantageous for the results as well as we might be able to better respond to adverse market circumstances. This can also be partly achieved by changing the way in which annuities are bought. At this moment one annuity is bought at the retirement date, but buying several smaller annuities over a longer period of time maybe helps to reduce the risk of disappointing results at retirement date.

A useful addition to the models could be a change from the static models as it is now to a more dynamic model. This would provide new insights regarding the effects between generations and within cohorts on which our inputs (like risk parameters or investment strategies) could be adapted.

The second part of this section is reserved for the developments in the wealth management industry regarding the automation of products in this sector. Robo-advisors become increasingly popular and are trying to take over business from traditional financial institutions. Robo-advisors are online, automated portfolio management services that are able to provide their services against low costs. They can offer their services for very low costs compared to traditional financial advisors due to the management of client investments by computer algorithms. The goal of this part of the Basis Zero experiment was to create a kind of robo-advisor in the field of pension investing.
A small but fast growing Dutch example is the company Pritle. But also large wealth management firms like BlackRock (acquiring FutureAdvisor), Vanguard (starting personal advisor services) and Fidelity (launching Fidelity Go) already acquired robo-advisors or developed in house solutions. The low-cost management, and other features like automatic portfolio rebalancing could result in higher net returns for investors compared to traditional asset managers. It is a sign that there are developments in the wealth management industry focusing on more automation and artificial intelligence. This does not mean a pension investor like APG should change its proposition and become a robo-advisor tomorrow, but it should stay up to date regarding the developments in this area.

7. References

Actuarieel Genootschap (2014). *Prognosetafel AG2014*. Retrieved from http://www.ag-ai.nl/view.php?Pagina_Id=535

APG (2016). Retrieved from https://www.apg.nl/en/apg-as-asset-manager

Balvers, R.J. & Wu, Y. (2005). Momentum and mean reversion across national equity markets. *Journal of Empirical Finance*, 13, 24–48.

Banks, J., Blundell, R. & Tanner, S. (1998). Is there a retirement-savings puzzle? *American Economic Review*, 88(4), 769-788.

Boskin, M. Shoven, J. (1987). *Concept and measurement of earnings replacements during retirement*. NBER Working Paper. Cambridge MA.

Boudoukh, J., Richardson, M. & Whitelaw, R. (1998). The best of both worlds: a hybrid approach to calculating value at risk. *Risk* (11): 64-67

Brealy, R.A., Myers, S.C. & Allen, F. (2010). *Principles of Corporate Finance*. New York: McGraw-Hill Irwin.

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

CPB (2016). *Kortetermijnraming maart 2016*. Retrieved from https://www.cpb.nl/cijfer/kortetermijnraming-maart-2016

DeMiguel, V., Garlappi, G. & Uppal, R. (2009). Optimal versus naïve diversification: how inefficient is the 1/N portfolio strategy? *The Review of Financial Studies*, 22(5), 1915-1953.

Fabozzi, F. K. (2007). *Robust Portfolio Optimization and Management.* Hoboken, New Jersey: John Wiley & Sons Inc.

Fugazza, C., Guidolin, M. & Nicodano, G. (2015). Equally Weighted vs. Long-Run Optimal Portfolios. *European Financial Management*, 21(4), 742-789.

García-Huitrón, M., Van Leuvenstein, M. & Oerlemans, A. (2016). Proactive planning as the answer to consumer and labor market uncertainty in the Netherlands. *Pension doc.*, 2, 8-18.

Goudswaard, K.P., Beetsma, R.M.W.J., Nijman, Th. E. & Schnabel, P. (2010). Een tweede sterke pijler: naar een toekomstbestendig stelsel van aanvullende pensioenen. Retrieved from https://www.rijksoverheid.nl/documenten/rapporten/2010/01/01/rapport-een-sterke-tweede-pijler-naar-een-toekomstbestendig-stelsel-van-aanvullende-pensioenen

Haveman, R., Holden, K., Romanov, A. & Wolfe, B. (2007). Assessing the maintenance of savings sufficiency over the first decade of retirement. *International Tax and Public Finance*, 14, 481-502.

Holzmann, R. & Hinz, R. (2005). *Old-Age Income Support in the 21st Century. An international perspective on pension systems and reform*. Washington: The World Bank.

Hull, J.C. (2012a). *Risk Management and Financial Institutions*. Hoboken, New Jersey: John Wiley & Sons Inc.

Hull, J.C. (2012b). *Options, futures, and other derivatives*. Hoboken, New Jersey: John Wiley & Sons Inc.

Jegadeesh, N. & Titman, S. (1993). Returns to buying winners and selling losers: implications for stock market efficiency. *Journal of Finance*, 48(1), 65-91.

Johnson, B. (2010). *Algorithmic Trading & DMA*. London: 4Myeloma Press.

Kissell, R. (2014). *Algorithmic Trading and Portfolio Management*. Oxford: Elsevier.

Knoef, M., Been, J., Alessie, R., Caminada, K., Goudswaard, K. & Kalwij, A. (2014). Measuring retirement savings adequacy: developing a multi-pillar approach in the Netherlands. *Journal of Pension Economics and Finance*, 15(1), 55–89.

MacDonald, B.J., Osberg, L. & More, K.D. (2014). How accurately does 70% final earnings replacement measure retirement income (in)adequacy? *Astin Bulletin*, 46(3), 627–676.

Nuti, G., Mirghaemi, M., Treleaven, P. & Yingsaeree, C. (2011). Algorithmic trading. *IEEE Computer*, 44(11), 61–69.

Pensions institute. (2016). *We Need a National Narrative: Building a Consensus around Retirement Income*. Independent Review of Retirement Income. Retrieved from http://www.pensionsinstitute.org/IRRIReport.pdf

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

Scholz, J. & Seshadri, A. (2009). *What Replacement Rates Should Households Use?* Working Paper. Wisconsin: Michigan Retirement Research Center.

SER. (2016). *Persoonlijk pensioenvermogen met collectieve risicodeling*. Retrieved from https://www.ser.nl/~/media/db_adviezen/2010_2019/2016/persoonlijk-pensioenvermogen.ashx

Treleaven, P. G. (2013). Algorithmic Trading Review. *Communications of the ACM*, 76-85.

VanDerhei, J. (2006). *Measuring Retirement Income Adequacy: Calculating Realistic Income Replacement Rates*. Washington DC: Employee Benefit Research Institute.

Vašíček, O.A. (1977). An Equilibrium Characterization of the Term Structure. *Journal of Financial Economics*, 5, 177–88.

Verschuren, P. & Doorewaard, H. (2010). *Designing a Research Project*. The Hague: Eleven International Publishing.

Appendices

Appendix 1 – The Basis Zero concept

The Basis Zero concept is explicable within 10 lines:

- During employment you and / or your employer pay a fixed amount per month.
- The premiums are invested collectively – you cannot make individual choices.
- For all participants investment risks and results will be divided.
- If you retire, the accumulated capital is converted into a lifelong annuity.
- When you die during employment there is a guaranteed compensation for survivors.
- You can watch, online, 24/7, what your accumulated capital position is.
- All communication is interactive and digital (through website, email and mobile)
- You get an online digital archive.

Appendix 2 – lifecycle data

The table below shows the percentages of total investments that are invested in each asset class at each time step. The years to the pension date (YTP) are given in the first column. The asset classes are represented on the first row. The same order of the thesis is used, so we start with Equities Developed Markets (North America) and end with Treasuries.

YTP	EDM1	EDM2	EDM3	EDM4	EEM	RE	COM	CRD	TRS	Sum
42	0,115	0,115	0,115	0,115	0,16	0,13	0,05	0,1	0,1	1
41	0,115	0,115	0,115	0,115	0,16	0,13	0,05	0,1	0,1	1
40	0,115	0,115	0,115	0,115	0,16	0,13	0,05	0,1	0,1	1
39	0,115	0,115	0,115	0,115	0,16	0,13	0,05	0,1	0,1	1
38	0,115	0,115	0,115	0,115	0,16	0,13	0,05	0,1	0,1	1
37	0,115	0,115	0,115	0,115	0,16	0,13	0,05	0,1	0,1	1
36	0,115	0,115	0,115	0,115	0,16	0,13	0,05	0,1	0,1	1
35	0,115	0,115	0,115	0,115	0,16	0,13	0,05	0,1	0,1	1
34	0,115	0,115	0,115	0,115	0,16	0,13	0,05	0,1	0,1	1
33	0,115	0,115	0,115	0,115	0,16	0,13	0,05	0,1	0,1	1
32	0,115	0,115	0,115	0,115	0,16	0,13	0,05	0,1	0,1	1
31	0,115	0,115	0,115	0,115	0,16	0,13	0,05	0,1	0,1	1
30	0,115	0,115	0,115	0,115	0,16	0,13	0,05	0,1	0,1	1
29	0,115	0,115	0,115	0,115	0,16	0,13	0,05	0,1	0,1	1
28	0,115	0,115	0,115	0,115	0,16	0,13	0,05	0,1	0,1	1
27	0,1125	0,1125	0,1125	0,1125	0,16	0,13	0,05	0,1	0,11	1
26	0,1125	0,1125	0,1125	0,1125	0,15	0,12	0,05	0,1	0,13	1
25	0,1125	0,1125	0,1125	0,1125	0,15	0,12	0,05	0,1	0,13	1
24	0,1125	0,1125	0,1125	0,1125	0,15	0,12	0,05	0,1	0,13	1
23	0,1125	0,1125	0,1125	0,1125	0,14	0,12	0,05	0,1	0,14	1
22	0,1125	0,1125	0,1125	0,1125	0,14	0,11	0,05	0,1	0,15	1
21	0,1125	0,1125	0,1125	0,1125	0,14	0,11	0,05	0,1	0,15	1
20	0,1125	0,1125	0,1125	0,1125	0,13	0,11	0,05	0,1	0,16	1
19	0,1125	0,1125	0,1125	0,1125	0,13	0,11	0,05	0,1	0,16	1
18	0,1125	0,1125	0,1125	0,1125	0,13	0,1	0,05	0,1	0,17	1
17	0,11	0,11	0,11	0,11	0,12	0,1	0,05	0,1	0,19	1
16	0,11	0,11	0,11	0,11	0,12	0,1	0,05	0,1	0,19	1
15	0,11	0,11	0,11	0,11	0,12	0,1	0,05	0,1	0,19	1
14	0,1075	0,1075	0,1075	0,1075	0,11	0,09	0,05	0,1	0,22	1
13	0,1075	0,1075	0,1075	0,1075	0,11	0,09	0,05	0,1	0,22	1
12	0,1025	0,1025	0,1025	0,1025	0,11	0,09	0,04	0,1	0,25	1
11	0,1025	0,1025	0,1025	0,1025	0,1	0,08	0,04	0,1	0,27	1
10	0,0975	0,0975	0,0975	0,0975	0,1	0,08	0,04	0,1	0,29	1
9	0,0975	0,0975	0,0975	0,0975	0,1	0,08	0,04	0,1	0,29	1
8	0,0925	0,0925	0,0925	0,0925	0,09	0,07	0,03	0,1	0,34	1
7	0,0925	0,0925	0,0925	0,0925	0,09	0,07	0,03	0,1	0,34	1
6	0,0875	0,0875	0,0875	0,0875	0,09	0,07	0,03	0,1	0,36	1
5	0,0875	0,0875	0,0875	0,0875	0,08	0,06	0,03	0,1	0,38	1
4	0,0825	0,0825	0,0825	0,0825	0,08	0,06	0,03	0,1	0,4	1
3	0,075	0,075	0,075	0,075	0,08	0,06	0,03	0,1	0,43	1
2	0,0675	0,0675	0,0675	0,0675	0,07	0,05	0,02	0,1	0,49	1

1	0,0575	0,0575	0,0575	0,0575	0,07	0,05	0,02	0,1	0,53	1
0	0,045	0,045	0,045	0,045	0,07	0,05	0,02	0,1	0,58	1

Appendix 3 – Interest rate

The table below gives an overview of the nominal interest rate term structure of July, 31st 2016 that is used in the thesis. Additional the rates used in the sensitivity analysis are presented.

	31-07-16			
Duration in years		+1%	+2%	-1%
1	-0,204	0,796	1,796	-1,204
2	-0,220	0,78	1,78	-1,22
3	-0,237	0,763	1,763	-1,237
4	-0,200	0,8	1,8	-1,2
5	-0,153	0,847	1,847	-1,153
6	-0,086	0,914	1,914	-1,086
7	-0,002	0,998	1,998	-1,002
8	0,091	1,091	2,091	-0,909
9	0,185	1,185	2,185	-0,815
10	0,274	1,274	2,274	-0,726
11	0,359	1,359	2,359	-0,641
12	0,430	1,43	2,43	-0,57
13	0,494	1,494	2,494	-0,506
14	0,550	1,55	2,55	-0,45
15	0,597	1,597	2,597	-0,403
16	0,627	1,627	2,627	-0,373
17	0,654	1,654	2,654	-0,346
18	0,677	1,677	2,677	-0,323
19	0,699	1,699	2,699	-0,301
20	0,718	1,718	2,718	-0,282
21	0,724	1,724	2,724	-0,276
22	0,740	1,74	2,74	-0,26
23	0,762	1,762	2,762	-0,238
24	0,789	1,789	2,789	-0,211
25	0,820	1,82	2,82	-0,18
26	0,854	1,854	2,854	-0,146
27	0,889	1,889	2,889	-0,111
28	0,926	1,926	2,926	-0,074
29	0,964	1,964	2,964	-0,036
30	1,003	2,003	3,003	0,003
31	1,042	2,042	3,042	0,042
32	1,080	2,08	3,08	0,08
33	1,119	2,119	3,119	0,119
34	1,156	2,156	3,156	0,156
35	1,194	2,194	3,194	0,194
36	1,230	2,23	3,23	0,23
37	1,266	2,266	3,266	0,266
38	1,301	2,301	3,301	0,301
39	1,335	2,335	3,335	0,335
40	1,369	2,369	3,369	0,369
41	1,401	2,401	3,401	0,401
42	1,433	2,433	3,433	0,433
43	1,463	2,463	3,463	0,463
44	1,493	2,493	3,493	0,493
45	1,522	2,522	3,522	0,522
46	1,550	2,55	3,55	0,55
47	1,578	2,578	3,578	0,578
48	1,604	2,604	3,604	0,604

49	1,630	2,63	3,63	0,63
50	1,655	2,655	3,655	0,655
51	1,679	2,679	3,679	0,679
52	1,702	2,702	3,702	0,702
53	1,725	2,725	3,725	0,725
54	1,747	2,747	3,747	0,747
55	1,768	2,768	3,768	0,768
56	1,789	2,789	3,789	0,789
57	1,809	2,809	3,809	0,809
58	1,828	2,828	3,828	0,828
59	1,847	2,847	3,847	0,847
60	1,866	2,866	3,866	0,866
61	1,884	2,884	3,884	0,884
62	1,901	2,901	3,901	0,901
63	1,918	2,918	3,918	0,918
64	1,934	2,934	3,934	0,934
65	1,950	2,95	3,95	0,95
66	1,966	2,966	3,966	0,966
67	1,981	2,981	3,981	0,981
68	1,995	2,995	3,995	0,995
69	2,009	3,009	4,009	1,009
70	2,023	3,023	4,023	1,023
71	2,037	3,037	4,037	1,037
72	2,050	3,05	4,05	1,05
73	2,063	3,063	4,063	1,063
74	2,075	3,075	4,075	1,075
75	2,087	3,087	4,087	1,087
76	2,099	3,099	4,099	1,099
77	2,111	3,111	4,111	1,111
78	2,122	3,122	4,122	1,122
79	2,133	3,133	4,133	1,133
80	2,144	3,144	4,144	1,144
81	2,154	3,154	4,154	1,154
82	2,164	3,164	4,164	1,164
83	2,174	3,174	4,174	1,174
84	2,184	3,184	4,184	1,184
85	2,194	3,194	4,194	1,194
86	2,203	3,203	4,203	1,203
87	2,212	3,212	4,212	1,212
88	2,221	3,221	4,221	1,221
89	2,230	3,23	4,23	1,23
90	2,238	3,238	4,238	1,238
91	2,246	3,246	4,246	1,246
92	2,255	3,255	4,255	1,255
93	2,263	3,263	4,263	1,263
94	2,270	3,27	4,27	1,27
95	2,278	3,278	4,278	1,278
96	2,285	3,285	4,285	1,285
97	2,293	3,293	4,293	1,293
98	2,300	3,3	4,3	1,3
99	2,307	3,307	4,307	1,307
100	2,314	3,314	4,314	1,314