NOVA
IMS
Information Management School

MGI

Mestrado em Gestão de Informação
Master Program in Information Management

GENETIC ALGORITHMS APPLIED TO ASSET & LIABILITY MANAGEMENT

Life Insurers' Perspective

João Filipe Clérigo Ribeiro de Almeida

Dissertation presented as a partial requirement for the degree of Master of Information Management

NOVA Information Management School
Instituto Superior de Estatística e Gestão de Informação
Universidade Nova de Lisboa

2016

Title: GENETIC ALGORITHMS APPLIED TO ASSET & LIABILITY MANAGEMENT

Subtitle: Life Insurers' Perspective

João Filipe Clérigo Ribeiro de Almeida

MGI

NOVA Information Management School
Instituto Superior de Estatística e Gestão de Informação
Universidade Nova de Lisboa

GENETIC ALGORITHMS APPLIED TO ASSET & LIABILITY MANAGEMENT

by

João Filipe Clérigo Ribeiro de Almeida

Dissertation presented as a partial requirement for the degree of Master of Information Management, Specialization in Information Systems and Technologies Management

Advisor: Mauro Castelli

Co-advisor: Jorge Bravo

August 2016

ACKNOWLEDGEMENTS

I would like to thank Professors Mauro Castelli and Jorge Bravo for the excellent guidance they provided me throughout the process of writing this thesis. Their recommendations on what to read in order to deepen my knowledge in the topics of genetic algorithms and asset liability management were invaluable in order to carry out this project properly and would have been invariably poorer result without their steering.

I would also like to thank my family and friends for encouraging me along the sometimes frustrating and difficult process of researching and writing for this project. A special thank you goes to Joana Vaz, my fiancée, for her incredible support and understanding all the way through this process.

ABSTRACT

Effective asset liability management is at the core of what a life insurance company must do, particularly in what concerns defined benefits pension fund products. The life insurer faces a complex problem whereby multiple and sometimes conflicting objectives must be addressed at the same time, such as achieving higher returns while reducing the portfolio's exposure to a plethora of risks. To achieve these goals, pension fund managers must then carefully choose asset allocation strategies for their portfolios from an infinite pool of asset combinations and weights. Given the nature of this problem, the use of genetic algorithms seems to be adequate, as this method is particularly well suited to deal with very large and multi-modal solution spaces. The main purpose of this dissertation is to assess how well the genetic algorithm method performs in solving this specific problem, and compare the results with other simpler methods. The results of Genetic Algorithms application were satisfactory and the results of this study suggests that Genetic Algorithms are a useful tool to solve ALM problems.

KEYWORDS

Genetic Algorithms; Evolutionary Algorithms; Pension Funds; Life Insurance; Asset Liability Management

INDEX

List of Figures

Acronyms and Abbreviations

ALM Asset Liability Management

DC Defined Contribution

DB Defined Benefits

MV Mean-variance

AI Artificial Intelligence

GA Genetic Algorithms

1. Introduction

Characterization of the Problem

Pension funds are amongst the world's biggest capital market participants and serve as an important long-term saving mechanism that enables any kind of employee, be it private or public sector, to save money throughout their careers and then enjoy a cash flow stream when they retire. There are two main types of pension funds in what regards relation between contributions and benefits: defined contribution (DC) and defined benefits (DB). In a DC scheme, also called funded individual accounts, each member (and sometimes his employer) pays into an account a fixed fraction of his or her earnings. These contributions are used to purchase assets, which are accumulated in the account, as are the returns earned by those assets. At retirement, the assets in the account finance post-retirement consumption through an annuity or some other payout option. In a DB scheme, employers use formulas that usually depend on the number of years worked and salary of the employee to calculate *a priori* the fixed amount the employee will be entitled to receive when he retires (Bodie, Marcus, & Merton, 1988). The main difference between these two types of pension funds is where the risk lies. In DC scheme annuities protect the individual against the non-systematic risks associated with longevity, but leaves him exposed to a wide range of risks, associated with stochastic real rates of return to pension assets, the risks of future earnings trajectories or the future pricing of annuities. In an employer DB scheme, the risk of varying rates of return to pension assets falls on the employer (sponsor), which is to say on some combination of the industry's current workers (through effects on wages), its shareholders and the taxpayer (through effects on current and future profits), its customers (through effects on prices), and/or its past or future workers (if, e.g., the company modifies the benefit formula relative to expectations).

It is in the particular case of a DB scheme that ALM becomes of paramount importance for plan sponsors and the life insurance companies that many times manage these funds for them. It is by using ALM techniques that these companies need to make sure they will have a sufficient asset inflow to be able to make the future defined payments, in many cases with decades of uncertainty between the initial agreement and the actual payment of employee benefits. Fund managers need then to choose from a number of existing techniques, such as gap analysis, cash flow matching, standard immunization (Redington, 1952; Boyle, 1978) or portfolio insurance (Leland, 1980; Leland and Rubinstein, 1981) to help them make proper investment decisions (van der Meer & Smink, 1993).

It is in addressing this very complex problem that the use of Genetic Algorithms can produce good results. If one considers a DB scheme as a fixed stream of future cash outflows, then the goal of the fund manager is to find a portfolio that satisfies those payments, while minimizing the risk the portfolio is exposed to (i.e. downside risk). One of the complications of doing so is that the solution space (e.g. the combinations of assets and asset weights a portfolio can have) is virtually infinite. This leads fund managers to have to apply some constraints (e.g. reduce the number of possible assets in the portfolio) in their allocation search, which may not lead to optimal results. This asset allocation problem has some properties that make genetic algorithms a good fit for solving them, namely having a huge solution space, the optimization target is not unimodal and good answers (i.e. portfolios) are composed of good individual components, which in this case are individual assets (Wadsley, 2011).

This dissertation will focus on exploring the possibilities of applying genetic algorithms to specific ALM problems such as optimal portfolio selection and comparing those outcomes with what more traditional ALM methods would achieve.

Relevance of the Problem

Pension funds are amongst the world's largest institutional investors. It is estimated that the top 300 world's biggest pension funds had assets under management of USD$ 14.9 trillion by year end 2013 (Watson, 2014), 67% of which were defined benefits funds. To put this number into perspective, the GDP of the United States of America was approximately USD$ 14.9 trillion in the same year (World Bank, 2014).

In what concerns DB pension funds, the consequences of asset liability mismatches can vary in significance, potentially leading the plan sponsor to make higher contributions to the fund or, in more extreme cases, can indirectly cause the bankruptcy via impairment of the credit rating (Ryan, 2002).

Considering the sheer size and potential adverse effects that improper ALM can bring to pension funds, insurance companies and society in general, it becomes imperative to use every means available to ensure that these negative outcomes are avoided.

Prior research suggests that genetic algorithms can indeed deliver interesting results when applied to the optimization of asset allocation to support ALM. Some research suggests that using these techniques can help fund managers achieve higher returns or lower volatility by using evolutionary algorithm based simulations (T., Huang, Chen, & Lin, 2012). Further research suggests that evolutionary algorithms can be successfully applied to portfolio optimization (Branke *et al*., 2009).

Aims and Objectives

This dissertation aims to assess the adequacy of the genetic algorithms in optimizing asset liability management problems. This main goal spreads out into the following objectives:

- Evaluate the performance of the genetic algorithm method in the optimization of asset liability management issues;
- Compare the performance of genetic algorithm based optimization to that of more traditional ALM methods.

2. Literature Review

Asset Liability Management

Numerous techniques and strategies have been put forward by scholars and industry practitioners to address the main problem of ALM, to ensure that asset and liability cash flows are coordinated (van der Meer & Smink, 1993). Among the simplest and easiest techniques to use, we can find some **static** techniques such as the cashflow payment calendars, where all major cash inflows and outflows are mapped through a given time period in order to identify imbalances, the gap analysis technique (Clifford, 1981), which looks at balance sheet difference in value between assets and liabilities to identify interest rate exposure, the segmentation technique (Attwood & Ohmna, 1984), where liabilities are segregated in homogeneous groups and asset portfolios are built to match the characteristics of the liability groups, the cashflow matching technique (Fabozzi, Tong, & Zhu, 1995), which uses linear programming to minimize the difference between asset and liability cash flows or classical immunisation that seeks to generate an assured return on the pension assets over the investment horizon. A major concern when using only the aforementioned static techniques would be that they ignore risk-return dimension of assets.

As a **hybrid** strategy between static and dynamic techniques we have the multiscenario analysis technique (Winklevoss, 1982), which forecasts the cashflows from assets and liabilities under different macroeconomic scenarios and shows what the effects of those scenarios would be in terms of asset-liability mismatches.

As for **dynamic** strategies we can find the standard immunization technique (Redington, 1952; Boyle, 1978; Granito, 1984), which proposes to match asset and liabilities' durations such that moves in interest rates will have opposite effects of the same magnitude in both, making their value stay the same. This is only applicable to very small changes in interest rates and for very short periods of time, which creates the need for continuously rebalance the asset portfolio. Another dynamic strategy, this one return driven, is the required rate of return strategy (Miller, Rajan, & Shimpi, 1989), which focuses in calculating the required returns needed to address the future cashflows of liabilities and then uses these returns to select an asset portfolio. Although this strategy is well suited for identifying trading possibilities it puts a lot of stress on returns and may lead to additional risk being carried by the portfolio.

Another problem that fund managers must face following the matching of assets and liabilities is how to engineer a portfolio that can achieve its matching objectives, while at the same time ensuring the highest possible return at the minimum possible risk. The portfolio selection process that dominated investment management for decades was the single-period mean-variance approach (Markowitz, 1952), in which an investor attempts to build an efficient portfolio by combining assets such that for a given level of returns the inherent risk is minimized or, in another perspective, for a given risk level the returns are maximized. Other authors proposed more sophisticated multi-period contexts for asset allocation decision, using continuous time models and dynamic programming techniques (Merton, 1971).

Other scholars use alternative dimensions to classify ALM models (Rosen & Zenios, 2006) and these can also be useful in order to understand how this problem has been approached and evolved through time. Figure 1 shows how ALM models can be classified according to time and type of factor.

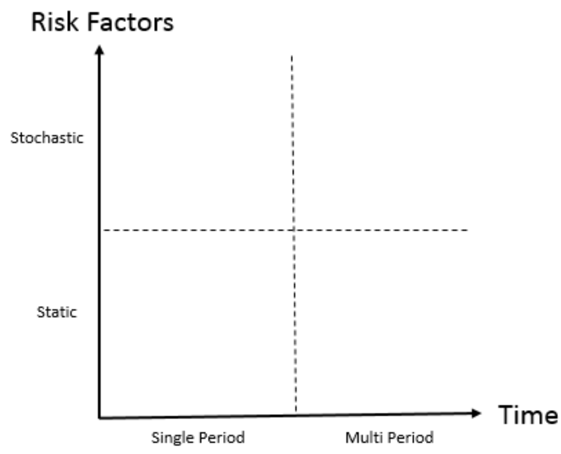

Figure 1. Grouping of ALM optimization models by time and type of risk factor dimensions (adapted from Rosen & Zenios, 2006)

The time dimension is split into two possible options: single period or multi period. In single period models, there is only one decision to be made regarding the portfolio composition and it is made in the present, without taking into consideration future time periods. Multi period models allow for portfolio decisions to be made not only in the present but also in future time periods. Naturally, allowing for portfolio decisions to be made across time leads to more realistic and accurate ALM models.

The risk factors dimension presents two available options: static or stochastic. The static view of risk factors consists of assuming that the risk factors present in the model (i.e. the parameters which bring uncertainty to the model) will remain as they are in the present through time. On the contrary, stochastic models allow for the evolution of the risk factors over time according to some assumed or estimated probability distribution. Unsurprisingly, by allowing risk factors such as economic growth or asset returns to fluctuate, stochastic models are considered more advanced and permit more realistic models.

The combination of these two dimensions allows us to characterize the most important ALM models available.

Single period static models. These models include the ones that allow for the concept of portfolio immunization, meaning following a strategy that allows the value of portfolio to remain unchanged in relation to small changes of a given risk factor (i.e. interest rates). The concept was introduced by Redington (Redington, 1952) and revisited recently by La Grandville (de La Grandville, 2001).

Single period stochastic models. The central idea behind these models is that they take into consideration the distribution of the risk factors instead of assuming these are constant (i.e. they allow for volatility). This is the central idea behind the popular and already discussed mean-variance model (Markowitz, 1991).

Multi period static models. This type of model allows for future decisions to take place, such as portfolio rebalancing, while assuming that risk factors remain constant. Some well-known models of this category include the models by Fama (Fama, 1970) and Hakansson (Hakansson, 1971).

Multi period stochastic models. This type of model allows not only for the fluctuation along time of the risk factors values but also for portfolio related decisions to be made in future periods. These include modern general stochastic programming models proposed by Birge and Louveaux (Birge & Louveaux, 2011) and the ones developed specifically for ALM issues by Ziemba and Mulvey (Ziemba & Mulvey, 1998).

Artificial Intelligence, Evolutionary Algorithms & Genetic Algorithms

Since the advent of the computer age, scientists have been very interested in the possibility of these artifical machines showing signs of human intelligence. Alan Turing called this possibility “machine inteligence” (Turing, 1948), while others such as Arthur Samuel put forward the concept of “machine learning” (Samuel, 1983). These ideas about Artificial Intelligence (AI) were at the foundation of what is called today “evolutionary computing”.

Evolutionary computing is a subfield of AI, which uses Darwin’s Evolution Theory principles to simulate organic evolution in order to solve problems using computers, most notably through the use of Genetic Algorithms (GAs). Figure 2 shows how these scientific areas relate to each other.

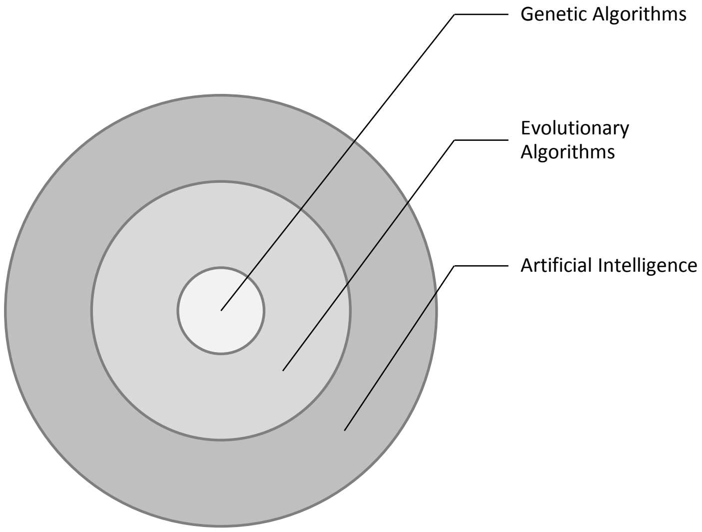

Figure 2. Genetic Algorithm’s position within the Artificial Intelligence field. This is an original figure based in Mitchell, 1998

The invention of GAs is attributed to John Holland (Holland, 1975), where he sought to “*study the phenomenon of adaptation as it occurs in nature and to develop ways in which the mechanisms of natural adaptation might be imported into computer systems*” (Mitchell, 1998). It’s in Holland’s research where the concepts of genes and chromosomes were combined with the bio-inspired mutation and crossover operators to solve problems by means of computers. Figure 3 shows the essential idea of how genetic algorithms work as put forward by Holland.

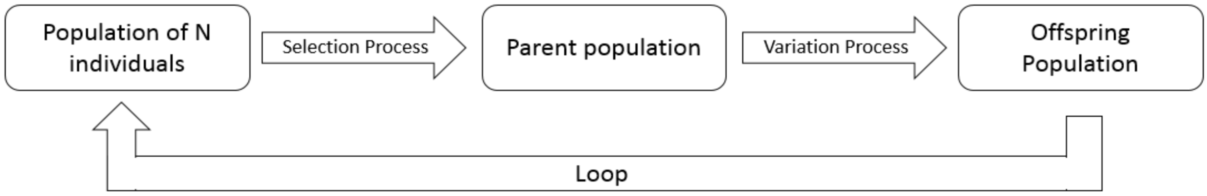

Figure 3. General Flow of a Genetic Algorithm. This is an original flow based in Holland, 1975

The algorithm is initialized with a population of individuals, which in this context means potential solutions to a given problem. The set of initial individuals are called the first generation, since as the GA loops new generations will be created. Then, the individuals are evaluated (i.e. their fitness to solve the problem is calculated) and the best individuals of the generation are selected to survive. The are many selection algorithms that can be used and the most common are fitness proportional, ranking and tournament (Goldberg & Deb, 1991) but the underlying idea is that individuals with good fitness (i.e. good solutions to the problem at hand) should be allowed survive. After the selection process the solutions go through what is referred to as the variation process, which consists of applying genetic operators such as mutation or crossover (Fogel & Atmar, 1990) in order to produce new solutions, which indirectly leads to explore the solution space. Once the variation process is finished, the new individuals (i.e. new generation) go through the same process over and over again until a termination condition is achieved, namely a certain number of generations is reached or a godd enough solution is found.

Genetic Algorithms Applications

Genetic algorithms have been used to solve and optimize problems in a vast number of different fields. This is possible due to the GA's intrinsic nature, namely that they are particularly well suited to search vast solution spaces, able to solve multi-modal problems, provide good near-optimal for problems where finding the optimal solution is too computationally costly or when the shape of good solutions is not known beforehand. Figure 4 shows some of the fields and problems that have made use of genetic algorithms.

Genetic Algorithms Applications
Analysis of Insolvency Risk (Varetto, 1998)
Neural Networks (Leung, Lam, Ling, & Tam, 2003)
Project Scheduling (Hartmann, 1998)
Function Optimization (Mühlenbein, Schomisch, & Born, 1991)
3D Medical Imaging (Cagnoni, Dobrzeniecki, Poli, & Yanch, 1999)
Flowshop Scheduling (Murata, Ishibuchi, & Tanaka, 1996)
Electromagnetics Engineering (Johnson & Rahmat-Samii, 1997)
Physics (Deaven & Ho, 1995)
Optimal Control Problems (Michalewicz, Janikow, & Krawczyk, 1992)
Finance (Oh, Kim, & Min, 2005)

Figure 4 List of GA Applications. This is an original table based in Mitchell, 1998

Genetic Algorithms and Portfolio Optimization

One of the interesting uses of genetic algorithms can be found in Finance, particularly in the sub-field of portfolio optimization. The portfolio optimization problem consists generally in selecting assets for a portfolio in such a way that the portfolio is optimized according to some criterion, namely maximizing the expected value of or minimizing its variance. Genetic algorithms are very well suited to perform portfolio optimization because the problem's features match closely GA's strengths, namely that the solution space is very large (i.e. combinations of all weighted assets in a portfolio), it is a multi-objective problem (i.e. maximize returns while minimizing risk), usually a local optimum is acceptable and the composition of a good portfolio is not known beforehand. This can be shown by the considerable number of authors that have used GAs in different ways to tackle this problem with good results. See (Branke et al., 2009), (Yang, 2006) and (Lai et al., 2006) for some interesting examples.

Genetic Algorithms and Asset Liability Management

It is straightforward to see that the ALM problem is very similar to the portfolio optimization problem. In other words, we can consider that portfolio optimization is part, but not the totality, of the ALM problem, meaning that an asset manager still has to construct an efficient portfolio but in ALM there is added complexity in the form of restrictions on asset allocation, liquidity constraints and individual's requirements. This assertion hints that GAs may also be of use when applied to the ALM problem because it of its similarities with portfolio optimization. Indeed, some authors have used GAs to solve successfully ALM problems while coping with many of the constraints classical model struggle or can't handle, and achieved good results (Baglioni et al., 2000), while other scholars have combined classical portfolio selection techniques such as mean-variance and Bayesian approaches with genetic algorithms to achieve enhanced results (Yang, 2006).

The focus of this thesis will be to evaluate the validity of using GAs to solve the specific problem of ALM for defined benefits pension plans by using actuarial methods to generate the liabilities of a simulated pension fund and then using GAs to build efficient portfolio that match the liabilities and respect the fund's restrictions.

3. Methodology

Experimental Design

In order to assess the usefulness of using Genetic Algorithms to solve an ALM problem this paper resorts to simulating a Defined Benefits Pension Fund and using GAs to optimize the portfolio choices. The simulation was designed using SAS BASE software and consists of the following eight sequential steps which will now be detailed.

1) Pension Fund Features. The very first step of the simulation was to define the specific features of the DB Pension Fund. Several assumption were made in this step in order to not overcomplicate the simulation with non-essential features. The fund consists initially of the hypothetic situation where a non-profit organisation is managing the funds on the behalf of the employees and sponsor. This means that any surplus generated is used to pay the pensioner in the future and is not used to pay dividends to the shareholders of the managing organization nor to lower the employees and/or sponsor contributions. In the simulations, we assume the fund consists of 1.000 members, all starting to contribute at age 25 and retiring at 65. The members and the sponsor of the plan both contribute equally with 5% of the yearly salary to the fund during the member's working life period. The starting salary is 50.000 units of income for all members at the beginning of their careers and it grows based on inflation and performance. It is not possible to exit the fund (i.e. leave this fund when changing jobs). The amount of cash available to invest in assets is then the yearly contribution and the returns/losses from the previous year. When the fund's members reach age 65 they retire, stop contributing to the fund and start receiving a yearly fixed income stream corresponding to 70% of their last salary. This yearly income is fixed and is received by the pensioner until they die. Maximum life expectancy was capped at 110 years of age. Figure 5 illustrates visually the CFs during the life of the fund under these assumptions, the Y-axis captures the cash flows while the X-axis shows the passage of time.

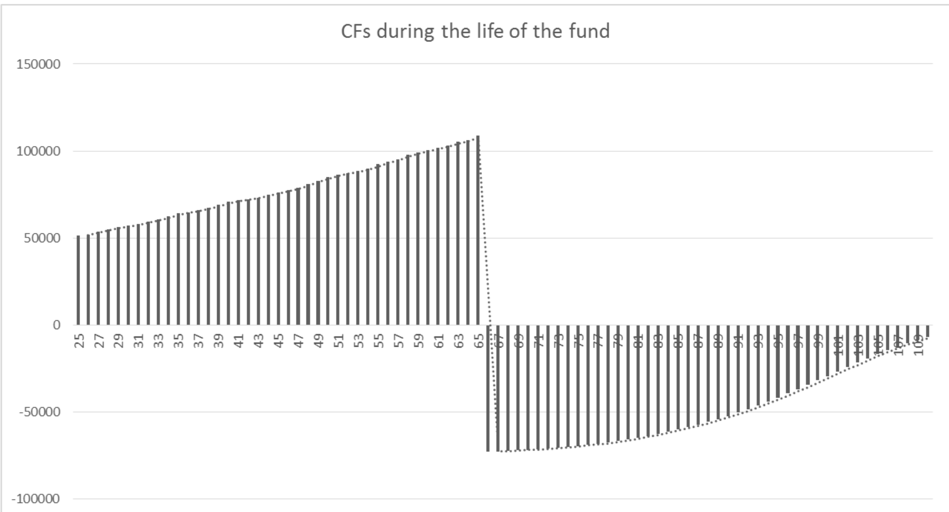

Figure 5 CFs during the life of the Fund

2) Longevity assumptions. Subsequent to defining the fund's features was to model the life expectancy of the future pensioners. To achieve this end an actuarial approach was utilised. A UK life table sourced from www.mortality.org (Appendix I) was used as input to this section, which contained a variable q(x) that represented the probability of death between ages *x* and *x+1*. The simulation used data for people born in 1996 and over 24 years old to improve the accuracy of a fund starting in 2016. The simulation looped through all years from age 25 to age 110 for all 1.000 members and calculated an indicator, named *life_death_flag,* that for each year would randomly generate a number between 0 and 1 and compare it to the probability of survival for a given year. If the random number generated is less than the probability of survival the indicator is set to 0 and the individual survives the period, otherwise it is considered deceased. Figure 6 shows a sample of the death probabilities corresponding to the life table used. It is straightforward to see that the chances of dying in a given year rise dramatically in old age.

Age	qx (Probability of death)	1 - qx (Probability of survival)
25	0.07%	99.94%
35	0.10%	99.90%
45	0.23%	99.77%
55	0.59%	99.41%
65	1.68%	98.32%
75	4.46%	95.54%
85	11.12%	88.88%
95	25.74%	74.26%
105	44.92%	55.08%
110	100.00%	0.00%

Figure 6 Probabilities of life and death by Age. Source: www.mortality.org

Figure 7 displays how the life_*death*_lag indicator works: if the randomly generated number is inside the blue section on a given year the individual survives to the next period and is faced with a decreasing survival chance. If the random number is inside the red section the individual is considered deceased in the 1st year that it happens.

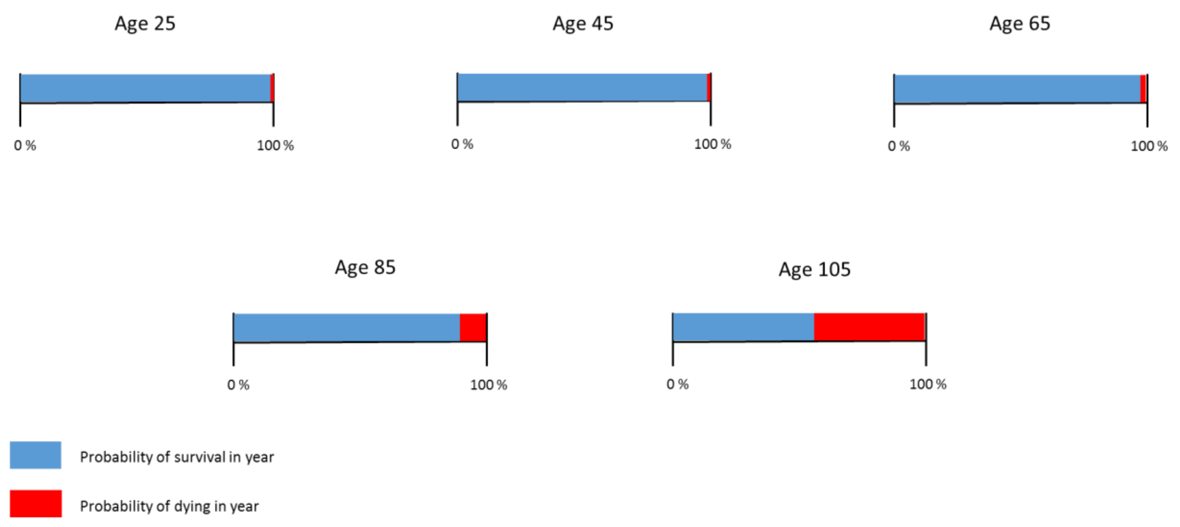

Figure 7 Probability of surviving vs dying in select years. This is an original figure based in www.mortality.org data

3) Pensioner's income. To forecast the future income of the fund members it was assumed that every member starts with a salary of 50.000 units of income and it grows each year until retirement in a fixed and a variable way. The salary has a 2% growth rate to account for an assumed 2% inflation rate and thus to avoid loss of purchasing power. The salary has also a random variable growth component, between 0% and 2% on any given year, to reflect differences in performance of the member. The final salary is of particular importance since it will be the basis for pension benefits calculation.

The second stage of calculating the member's income relates to yearly pension for all pensioners. To calculate this amount we calculate the yearly pension (i.e. 70% of last yearly salary), remove members that were deceased before age 65, retrieved from the simulated age of death from the life expectancy section and then proceed to calculate the undiscounted cash flows for each member according to the following formula:

$$Total\ Pension = Yearly\ Pension\ \times (Age\ of\ Death - Retirement\ Age)$$

The final step of this section is to calculate the present value of pension benefits at the time of retirement. The discount factor used was 5%, which is in agreement with current market practices (see www.otpp.com for an example) for this type of operation, even though we acknowledge that there is an ongoing debate in the industry regarding this rate. The objective of the fund will be to invest in a way such that the pension contributions and investment returns are sufficient to cover the discounted total pension at age 65.

4) Employee and Sponsor contributions. The next step of the simulation is to calculate all the contributions to the pension fund made by the plan's sponsor and the member. As stated previously this will be a fixed percentage of the annual salary (5% from sponsor and 5% from member) during active working years (i.e. 25 to 65 years of age), taking into account simulated life expectancy, meaning that members that are deceased between ages 25 and 65 stop making contributions when they die. The differences between contributions amongst members exists only due to the performance part of the salary which is randomly attributed and can generate some differences in yearly salaries. Notice that these contributions are independent from investment returns and are fairly stable and predictable growth rate in comparison.

5) Target rate of return for Assets. The following step is to compute the rate of return at which the plan's contributions must be invested such that they cover the discounted total pension at age 65. Essentially, it is calculating the rate at which the contributions must be compounded to equal the liabilities at retirement. To this effect, several rates were simulated starting at 1% and growing in 0.001 steps until a rate was found that could cover the liabilities. This rate is the minimum rate of return required to cover the liabilities and was generally around 5% under the followed assumptions.

6) Asset return scenarios. We assume the fund can invest in 3 different assets: the S&P 500 index, U.S. Gov. 10-year Treasury Bonds and U.S. Gov. 3-month Treasury Bills, each of them intended to represent the stocks, bonds and risk-free asset classes, respectively. The already calculated historical yearly returns from 1928 until 2015 were sourced from Professor A. Damodaran's website and the original raw data was obtained from S&P and the FED (Appendix II). The mean and standard deviation was computed for each of the asset classes and were used to generate random scenarios in which the asset prices would fluctuate randomly following a normal distribution with the historical mean and standard deviations observed for the past 85 periods. The equation used to simulate the asset prices is given by:

$$Asset\ Price(t\ +1) = \ Asset\ Price(t) \times (\ 1 + invNorm(\gamma)\ \times\ \sigma(Asset\ Class) +\ \mu(Asset\ Class)$$

Where:

Asset Price (t + 1): price of the asset in the next period

Asset Price (t): price of the asset in the current period

invNorm(γ): inverse of the gaussian cumulative density function for a random number that follows a Uniform Distribution in the interval [0,1]

$\sigma_{\text{(Asset Class)}}$: standard deviation for asset class ($\sigma > 0$)

$\mu_{\text{(Asset Class)}}$: mean for asset class

The asset price of the next period is computed starting from the asset price of the current period and multiplying that value with one plus the inverse of the Gaussian CDF for a random number that follows a uniform distribution. The inverse of the Gaussian CDF has the standard deviation and the mean of the assets class being calculated[1].

Correlation between assets is not explicitly considered. However, since we used historical means and standard deviation to calculate the normal distribution parameters they are implicitly present in these parameters.

Figures 8, 9 and 10 show ten scenario simulations for each of the assets. The Y-axis adapts to the figures in each graphic, by looking closely at each Y-axis it is easy to see the increasing volatility.

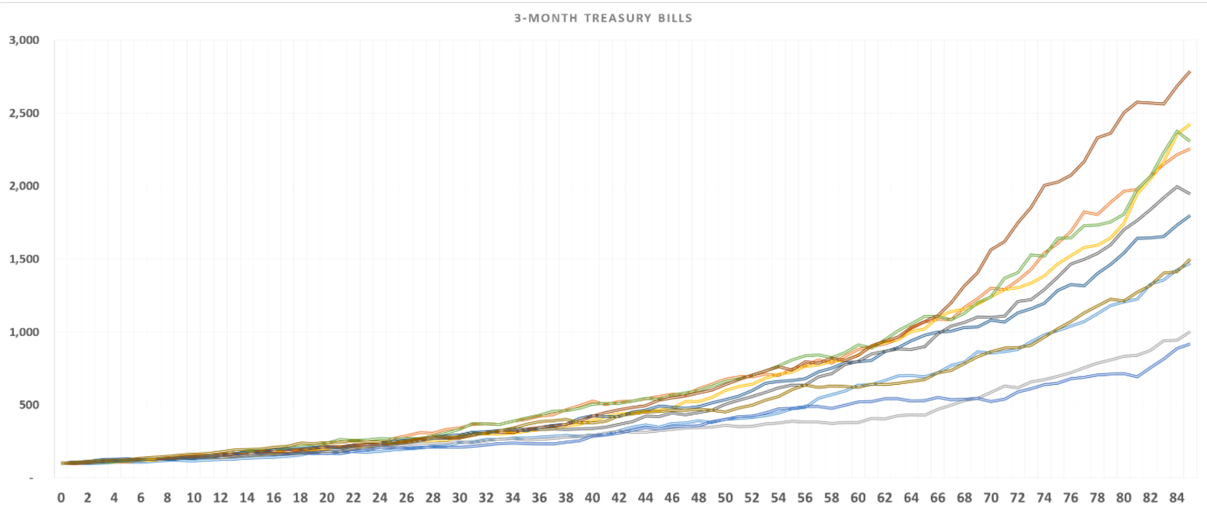

Figure 8 Randomly generated scenarios for 3-month Treasury Bills

[1] For more details on simulating the values of a normal random variable see Winston, W: *Microsoft Excel 2010 Data Analysis and Business Modelling*, Pearson Education, 2011

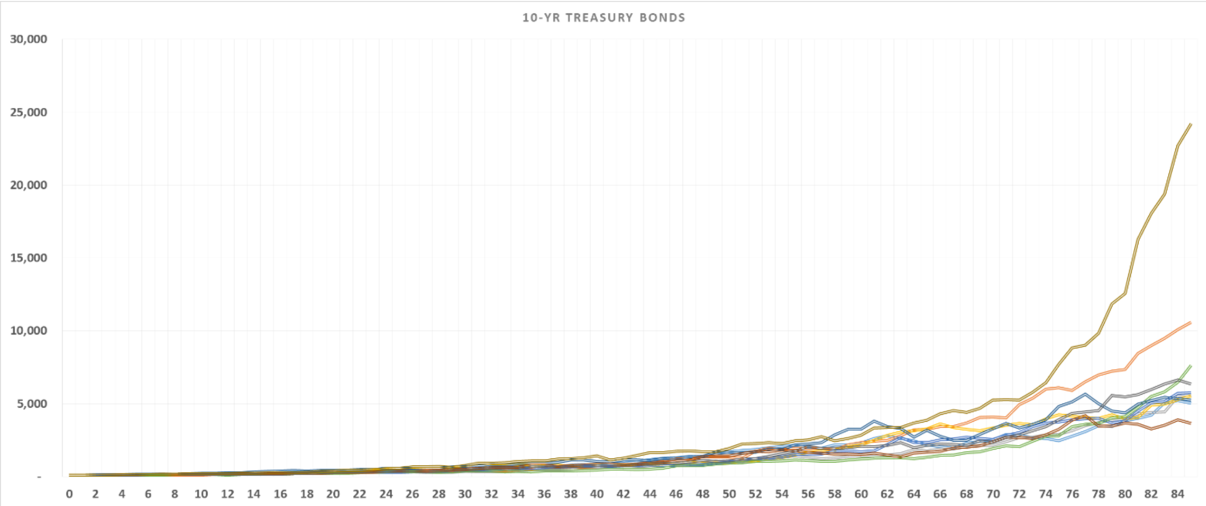

Figure 9 Randomly generated scenarios for 10-Year Treasury Bonds

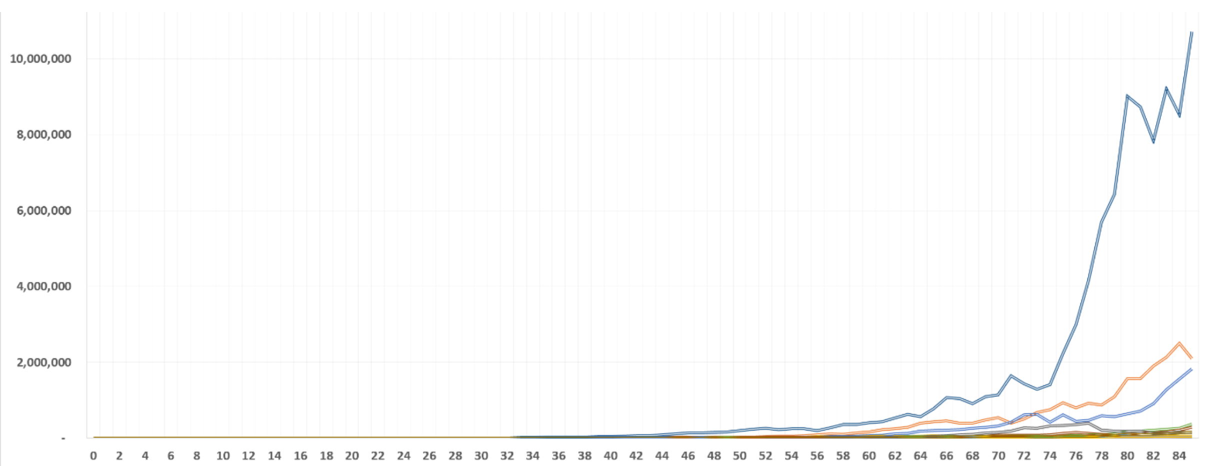

Figure 10 Randomly generated scenarios for the S&P 500 Stock Index

7) Creating efficient portfolios using Genetic Algorithm. The final section of the model is to use Genetic Algorithms to create efficient portfolios of assets. The simulation used 3 loops, one inside of each other to achieve this objective.

The outer loop feeds the model different scenarios, where asset prices randomly change through time. The middle loop presents the model different years for each scenario, being that one 1 scenario consists of 85 years of simulated asset returns. For each new year presented to the model, the previous 30 years are used to compute the asset's means, variance and covariance, which will be used in the inner loop. The inner loop is where the GA is used to generate and evolve solutions to create efficient portfolios. In this context, solutions consist of different combinations of weights attributed to each of the 3 assets in order to create a portfolio, being that the sum of all weights must be equal to 100%. Short positions are not allowed in this model, as they are not usually used in pension fund management.

For each year, 100 solutions are randomly generated in the 1st generation and for each solution the mean return and variance of the portfolio they create is computed. Then, the fitness of each solution is evaluated by minimizing the difference between the achieved mean return of the portfolio with the minimum rate of return required to cover the liabilities plus 1% (to create some slack), which was usually circa 6%. After ranking the solutions by this 1st criteria, the second fitness criteria, minimizing the standard deviation of the portfolio, would do a 2nd ranking of the solutions. The top solution from each generation was stored permanently in a separate table. To evolve the solutions, the 30% best solutions would go through the variation process to generate offspring to be considered in the next generation. Strict vertical crossover was used in the variation process, meaning that parent solutions would contribute their weights to offspring but only inside the same asset class. This was put in place because if a solution has a high fitness it must have been due to generating asset allocations that approximated well the target expected return while minimizing variance of the portfolio. Changing the weights across asset classes would risk generating offspring with less fitness than their parents and it was avoided. The same line of thought explains why mutation was also not considered: changing randomly the solution's weights would likely generate worse solutions. The case might have been different had the simulation used lots of assets, in those circumstances the benefit of searching the solution space better would probably outweigh generating slightly worse offspring. After the top 30% solutions went through crossover another 70 random solution would be generated and the 100 solutions would loop again to create new generations. Five generations were computed per year in the simulation. Figure 11 shows the distribution of a generation of 100 solutions across the portfolio return and standard deviation dimensions.

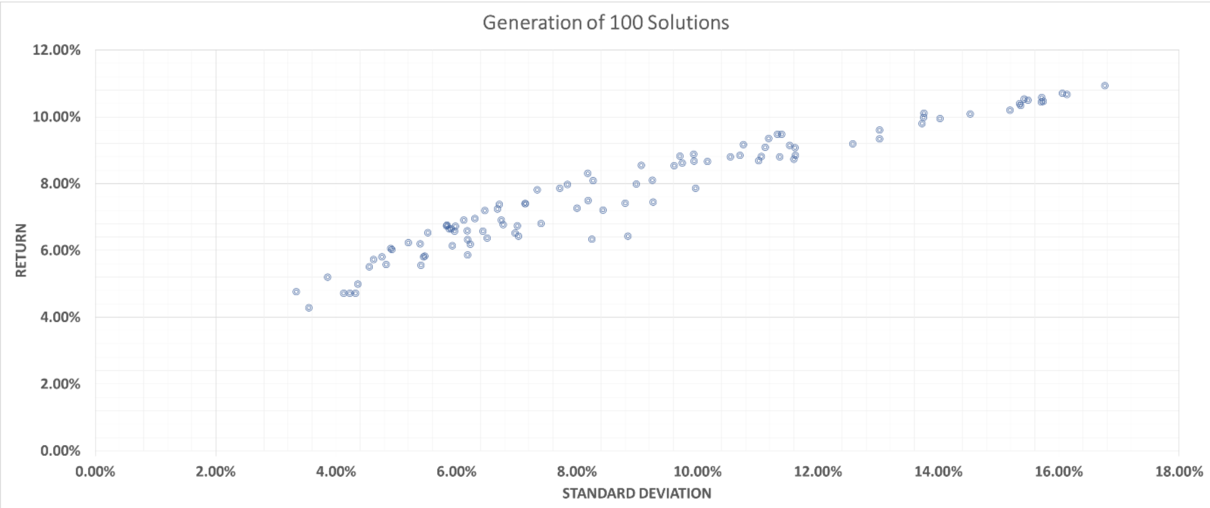

Figure 11 Generation of 100 Solutions

8) Outcomes Evaluation. The final step of the model is to combine the best portfolios generated in the previous step with the contributions to the fund and the simulated returns for the scenarios considered in the simulation. This created a table which can the used to calculate surplus / deficits in the retirement year and to compare with other methodologies by changing the weights of each asset.

Assumptions

The simulation model incorporates the following assumptions:

- The Pension Fund is not for profit (surplus is kept by pension fund)

- 110 yrs. old is the maximum age allowed in the model
- Contributions start at age 25
- Inflation Rate is fixed at 2%
- The salary growth rate is inflation plus a random increase between [0%-2%] (performance)
- It is not possible to exit the fund
- Disabilities are not considered
- Investment in any given available asset class is not limited
- Widowers are not considered
- The pension is 70% of last salary and is not indexed over time
- Starting Salary is 50.000 units of income for all pensioners
- The employer and sponsor contributions already consider the simulated life expectancy of each pensioner
- The employer and sponsor contributions are 5% each and are fixed
- Target rate of return depends on liabilities
- If a pensioner dies between age 25-64 the contribution from that year does not occur
- The available assets are the S&P 500 index (stock), U.S. 10-year T-Bond (bond) and U.S 3-month T-bills (risk-free)
- Scenarios for the movement of future asset prices are created using normal distribution simulation
- Input Data for the simulation of future asset prices are the historical returns for the assets being used
- Asset Prices are assumed to have a value of 100 for each asset unit at the beginning of the simulation (what matters is distribution of the returns)
- The Genetic Algorithm will search for weight allocations across the 3 available assets to create efficient portfolios
- No liquidity constraints are considered
- Member’s preferences are not taken into account
- Regulatory requirements, such as capital buffers, were not considered
- Transactions costs are not considered

4. Results and Discussion

Experimental Results and Discussion

For the purpose of evaluating if the model produced adequate results, 10 scenarios were executed and the surplus or deficit of the pension fund at retirement age was calculated. The choice of executing 10 scenarios was taken due to technical reasons (i.e. time and computer memory constraints), namely the fact that with 10 scenarios the optimization would take 20 minutes to execute. This happens because for each scenario 85 years of asset returns are simulated, with 100 randomly generated solutions a year and that are optimized in 5 generations. Figure 12 displays the value of the pension fund against the NPV of the liabilities at retirement age for the executed scenarios. Figure 13 shows some statistics of the fund value.

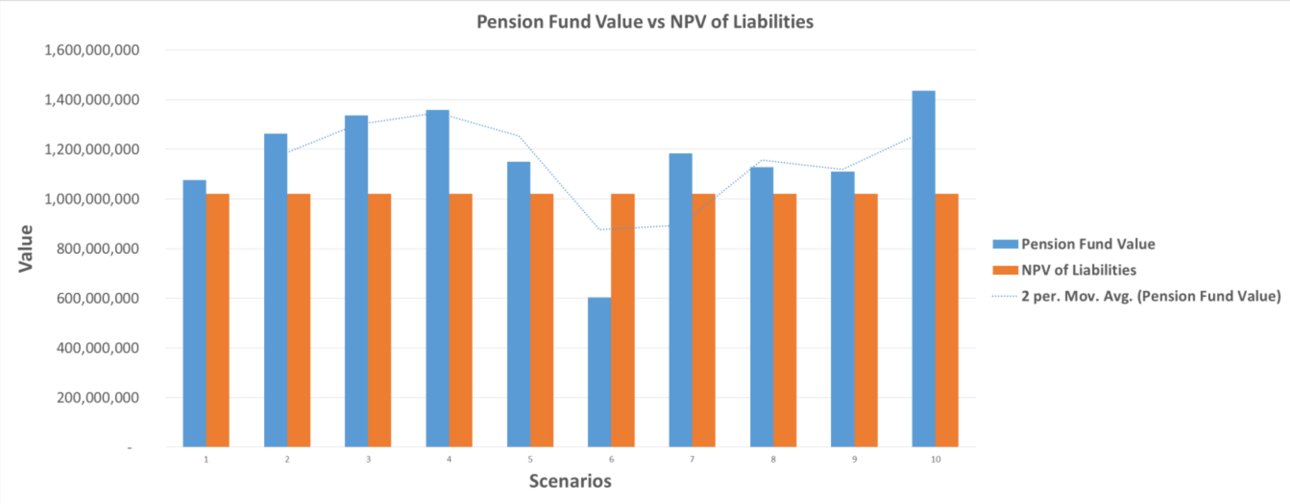

Figure 12 Pension Fund Value vs NPV of Liabilities at Retirement

Statistics	Values
Expected Value	1,164,719,931
St. Dev.	230,417,764
Max	1,436,097,783
Min	603,493,047

Figure 13 Pension Fund Descriptive Statistics

The results of the simulation are positive and suggest that the GA portfolio optimization produced the desired results, meaning achieve a rate of return such that the liabilities can paid by the fund, while minimizing variance of returns. In 9/10 scenarios the pension fund achieved sufficient value to cover all the liabilities. We can also note that the results are fairly consistent across iterations. This also suggests that the sponsor would be unlikely asked to make extra contributions.

The weight distribution of assets in the portfolio across time was also analysed for one scenario, to assess how it varied over time. The mean allocation for stocks, T-bonds and T-bills was 25%, 44% and

32%, respectively. This is in accordance expectations, in the sense that model was targeting a relatively low return rate, which forced the composition of the portfolios to be made of less risky assets. Figure 14 shows the distribution of weights across the assets through the investing horizon. We can see that there is considerable movement of the allocation of the assets between periods, which could cause difficulties if the model considered securities transaction costs.

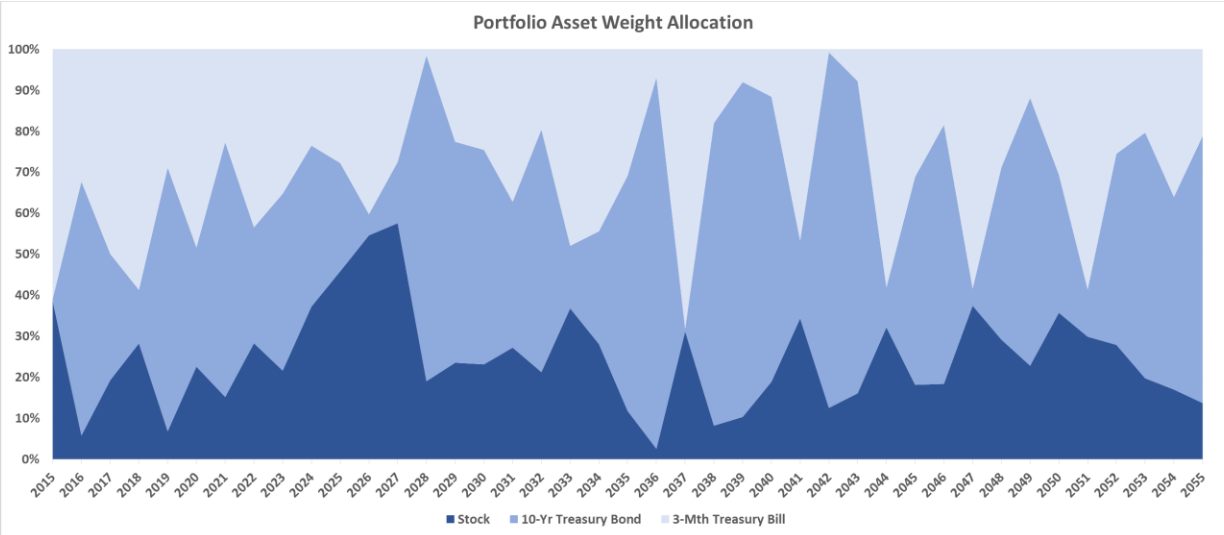

Figure 14 Distribution of weights across by asset through the investing horizon

Comparison to Other Methodologies

To assess the effectiveness of the model we compared the generated solutions with the solutions generated via two other different methods: (1) a portfolio split equally across all assets and (2) a portfolio resulting from a 1 period static Minimum-Variance optimization only.

Figure 15 shows statistics for the generated fund value of each method for the 10 scenarios in analysis.

By comparing the model portfolio with the equally weighted portfolio we can see that the equally weighted portfolio yields a higher expected value at the expense of a bigger standard deviation. This happens because the return objective of the model portfolio is relatively low and shifts asset allocation towards less risky assets whereas in the equally weighted portfolio there are no such considerations. We also noted that the equally weighted portfolio generated a deficit in the same scenario as the model portfolio, even though it was smaller than the model portfolio.

When assessing the model portfolio against the 1 period static optimization portfolio we see that the statistics are fairly comparable. We note however that the latter generated deficits in 3 scenarios while the model portfolio only had 1. This may be due to the static nature of the static optimization, in the sense that it fails to take into account relevant changes in the underlying market circumstances, which can be particularly exacerbated when considering long term investment horizons.

We also computed the Sharpe Ratio and the Calmar Ratio for each portfolio. The model portfolio exhibited the best Sharpe Ratio of the three, mainly due to a good balance between the realised return (i.e. ratio was calculated ex post) and its standard deviation. There were no significant differences between the portfolio's Sharpe Ratios. The Calmar Ratio produced interesting results, putting the 1 Period in a much better standing than the other two portfolios. This happened because the CAGRs for all portfolios were very similar (i.e. between 6-7%) but the Maximum Drawdown was significantly higher for both the model and equal weights portfolios. This happened because these portfolios had significant exposures to equities in years that these asset classes had very poor performances. It is also worthwhile to notice that using the 40 year period to assess the Maximum Drawdown would tend to penalize most dynamic portfolios.

Statistics	Model Portfolio	Equal Weights Portfolio	1 Period Optimization Portfolio
Expected Value	1,164,719,931	1,399,693,198	1,144,947,959
St. Dev.	230,417,764	378,373,568	219,800,088
Max	1,436,097,783	2,219,577,629	1,478,081,799
Min	603,493,047	844,414,786	837,234,849
Sharpe Ratio	1.78	1.70	1.77
Calmar Ratio	1.26	1.87	3.50

Figure 15 Comparison of Outcomes between diferent asset allocation methods

5. Conclusions

In this study we produced a model that simulates and solves the ALM problem faced by Defined Benefits Pension Funds. We defined the features of the fund, used actuarial methods to calculate the life expectancy of the fund's members and used that information to calculate the fund's liabilities and contributions to the fund. Based on the liabilities we calculated the minimum target return rate that the portfolio of assets needs to obtain to cover the liabilities. We then used historical asset prices to generate scenarios of future asset returns and used Genetic Algorithms to create efficient portfolios of assets that would achieve the desired rate of return and minimizing volatility. The results were positive in the sense that the model reaches its objectives (i.e. generate a surplus) in the majority of the scenarios in the study, while minimising risk. The model also compares well with other simpler asset allocation methods, such as single period optimization and equal weight asset allocation. Using well known statistics to assess the performance of the various models, we could verify that the dynamic multi-period optimization tends to produce the best risk-adjusted returns but it is still exposed to over-allocation to some asset classes, as the results of the CALMAR ratio indicated. This study suggests that Genetic Algorithms can be a useful tool to solve ALM problems using a dynamic, multi-period approach.

6. Limitations and Recommendations for Future Research

Limitations of the Model

This study has some limitations caused by some of the assumptions used. This section analyses them and proposes ways to solve or mitigate them.

Number of assets used. The model uses a limited number of assets to create portfolios and while this feature was beneficial to speed up executions of the model it would be interesting to see how the mode would behave given more assets. Achieving this would not be very complex (i.e. it would require adjusting some of the formulas).

Exiting the fund, disabilities and widowers. The model does not allow for members to exit the fund, become disabled or give back their contributions to their widow. While this would be interesting it was not deemed critical to achieve the objectives of this study. For examples of how this could be achieved see (Dondi, Herzog, Schumann, & Geering, 2006).

Initialization of the fund. The model initializes the fund with very simplistic terms. It would be interesting to add some dynamism to this, namely starting from a point with pensioners and active members mixed together.

Asset allocation reaction. The asset allocation is driven fully by the required rate of return to meet the future liabilities. It does not take into consideration the actual results obtained to guide future investing behaviour. To see an example of how to achieve these outcome see (Infanger, 2006).

Pension Fund Features. The fund's features do not take into consideration the full spectrum of what can occur. For instance, it was assumed that any generated surplus always stays with the fund but that needs not be the case. See (Boender, Dert, Heemskerk, & Hoek, 2006) for a more comprehensive definition of fund's features.

Recommendations for Future Research

Some recommendations for future research are as follows:

- It would be interesting to see such a simulation applied to real pension fund with DB
- The model could be extended to mitigate many of its current limitations. It would be interesting to see the impact of those extensions
- It could be interesting to use GAs to estimate the future asset returns more precisely, thus extending their role in the ALM problem
- Compare the results of this model with more advanced multi-period stochastic models

7. Bibliography

Attwood, J., & Ohman, C. (1984). *Segmentation of Insurance Company General Accounts.* Georgia Life Office Management Association, Inc.

Baglioni, S., Pereira, C. D., Sorbello, D., & Tettamanzi, A. G. (2000). An evolutionary approach to multi-period asset allocation. *Lecture Notes in Computer Science*, 225–236.

Bank, W. (27 de June de 2014). *World Bank*. Obtido de World Bank: http://data.worldbank.org/indicator/NY.GDP.MKTP.CD/countries/US?display=graph

Birge, J. R., & Louveaux, F. (2011). *Introduction to stochastic programming.* Springer Science & Business Media.

Bodie, Z., Marcus, A., & Merton, R. (1988). Defined Benefit versus Defined Contribution Pension Plans: What are the Real Trade-offs? Em N. B. Research, *Pensions in the U.S. Economy* (pp. 139-162). Chicago: University of Chicago Press.

Boender, C., Dert, C., Heemskerk, F., & Hoek, H. (2006). A scenario approach of ALM. Em *Handbook of Asset Liability Management.*

Bojarczuk, C. C., Lopes, H. S., & Freitas, A. A. (2000). Genetic programming for knowledge discovery in chest-pain diagnosis. *IEEE Engineering in Medicine and Biology Magazine*, 38–44.

Boyle, P. (1978). Immunisation under stochastic models of the term structure. *Journal of the Institute of Actuaries 105*, 177-187.

Branke, J., Scheckenbach, B., Stein, M., Deb, K., & Schmeck, H. (2009). Portfolio optimization with an envelope-based multi-objective evolutionary algorithm. *European Journal of Operational Research*, 684–693.

Branke, J., Scheckenbach, B., Stein, M., Deb, K., & Schmeck, H. (2009). Portfolio optimization with an envelope-based multi-objective evolutionary algorithm. *European Journal of Operational Research*, 684-693.

Cagnoni, S., Dobrzeniecki, A. B., Poli, R., & Yanch, J. C. (1999). Genetic algorithm-based interactive segmentation of 3D medical images. *Image and Vision Computing*, 881-895.

Castillo, F., Kordon, A., & Smits, G. (2006). Robust pareto front genetic programming parameter selection based on design of experiments and industrial data. Em *Genetic Programming Theory and Practice IV.* Springer.

Clifford, J. T. (1981). A Perspective on Asset-Liability Management: Part I. *Bank Management*.

de La Grandville, O. (2001). *Bond Pricing and Portfolio Analysis.* Cambridge, MA: The MIT Press.

Deaven, D. M., & Ho, K. M. (1995). Molecular geometry optimization with a genetic algorithm. *Physical review letters*, 288.

Dondi, G., Herzog, F., Schumann, L. M., & Geering. (2006). Dynamic asset and liability management for Swiss pension funds. Em *Handbook of Asset and Liability Management.*

Fabozzi, T., Tong, T., & Zhu, Y. (1995). Beyond Cash Matching. Em T. Fabozzi, *The Handbook of Fixed Income Securities 4th ed.* (p. Chapter 44). Irwin.

Fama, E. F. (1970). Multiperiod consumption-investment decisions. *The American Economic Review*, 163-174.

Fogel, D. B., & Atmar, J. W. (1990). Comparing genetic operators with Gaussian mutations in simulated evolutionary processes using linear systems. *Biological Cybernetics*, 111-114.

Goldberg, D. E., & Deb, K. (1991). A comparative analysis of selection schemes used in genetic algorithms. *Foundations of genetic algorithms*, 69-93.

Granito, M. (1984). *Bond Portfolio Immunization.* Toronto: D. C. Health and Company.

Hakansson, N. H. (1971). On optimal myopic portfolio policies, with and without serial correlation of yields. *Journal of Business*, 324-334.

Hampo, R. J., & Marko, K. A. (1992). Application of genetic programming to control of vehicle systems. *Proceedings of the Intelligent Vehicles '92 Symposium.* Detroit.

Hartmann, S. (1998). A competitive genetic algorithm for resource-constrained project scheduling. *Naval Research Logistics*, 733-750.

Holland, J. H. (1975). *Adaptation in natural and artificial systems: an introductory analysis with applications to biology, control, and artificial intelligence.* U Michigan Press.

Infanger, G. (2006). Dynamic asset allocation strategies using a stochastic dynamic programming approach. Em *Handbook of asset and liability management* (pp. 199-251).

Johnson, J. M., & Rahmat-Samii, V. (1997). Genetic algorithms in engineering electromagnetics. *Antennas and propagation Magazine*, 7-21.

Koza, J. R. (1992). *Genetic programming: on the programming of computers by means of natural selection.* MIT press.

Koza, J. R. (2010). Human-competitive results produced by genetic programming. *Genetic Programming and Evolvable Machines 11*, 251-284.

Lai, K. K., Yu, L., Wang, S., & Zhou, C. (2006). A double-stage genetic optimization algorithm for portfolio selection. *Conference on Neural Information Processing* (pp. 928-937). Springer Berlin Heidelberg.

Leland, H. (1980). Who should buy portfolio insurance? . *Journal of Finance 35*, 581-594.

Leland, H., & Rubinstein, M. (1981). Replicating options with positions in stocks and cash. *Financial Analysts Journal 37*, 63-72.

Leung, F. H., Lam, H. K., Ling, S. H., & Tam, P. K. (2003). Tuning of the structure and parameters of a neural network using an improved genetic algorithm. *Transactions on Neural networks*, 79-88.

Lin, C. M., & Gen, M. (2007). An effective decision-based genetic algorithm approach to multiobjective portfolio optimization problem. *Applied Mathematical Sciences*, 201-210.

Markowitz, H. (1952). Portfolio selection. *Journal of Finance*, 77–91.

Markowitz, H. (1991). *Portfolio Selection: Efficient Diversification of Investments.* Oxford: Basil Blackwell.

Merton, R. C. (1971). Optimal consumption and portfolio rules in a continuous-time. *Journal of Economic Theory*, 373–413.

Michalewicz, Z., Janikow, C. Z., & Krawczyk, J. B. (1992). A modified genetic algorithm for optimal control problems. *Computers & Mathematics with Applications*, 83-94.

Miller, L., Rajan, U., & Shimpi, P. A. (1989). Realized Return Optimization: A Strategy for Targeted Total Return Investing in the Fixed Income Markets. *The Institutional Investor Focus on Investment Management*.

Mitchell, M. (1998). *An introduction to genetic algorithms.* MIT press.

Mühlenbein, H., Schomisch, M., & Born, J. (1991). The parallel genetic algorithm as function optimizer. *Parallel computing*, 619-632.

Murata, T., Ishibuchi, H., & Tanaka, H. (1996). Multi-objective genetic algorithm and its applications to flowshop scheduling. *Computers & Industrial Engineering*, 957-968.

Oh, K. J., Kim, T. Y., & Min, S. (2005). Using genetic algorithm to support portfolio optimization for index fund management. *Expert Systems with Applications*, 371-379.

Poli, R., Langdon, W. B., McPhee, N. F., & Koza, J. R. (2008). *A field guide to genetic programming.*

Redington, F. (1952). Review of the principles of life office valuations. *Journal of the Institute of Actuaries 18*, 286-315.

Rosen, D., & Zenios, S. (2006). Enterprise-wide asset and liability management: issues, institutions, and models. Em *Handbook of Asset and Liability Management: Theory and Methodology* (pp. 6-23).

Ryan, R. J. (2002). Rethinking Pension Liabilities and Asset Allocation. *Journal Of Portfolio Management*, 28(4), 7-15.

Samuel, A. L. (1983). AI, Where It Has Been and Where It Is Going. *International Joint Conference on Artificial Intelligence*, (pp. 1152-1157).

T., Y., Huang, H., Chen, C., & Lin, Q. (2012). Generating effective defined-contribution pension plan using simulation optimization approach. *Expert Systems with Applications*, 2684–2689.

Turing, A. M. (1948). *Intelligent machinery.*

van der Meer, R., & Smink, M. (1993). Strategies and Techniques for Asset-Liability Management: an Overview. *The Geneva Papers on Risk and Insurance*, 144-157.

Varetto, F. (1998). Genetic algorithms applications in the analysis of insolvency risk. *Journal of Banking & Finance*, 1421-1439.

Wadsley, B. (February de 2011). Are Genetic Algorithms Even Applicable to Actuaries? *Risks & Rewards*, pp. 3-7.

Watson, T. (2014). *Pensions & Investments / Towers Watson 300 analysis.* New York: Towers Watson.

Winklevoss, H. E. (1982). Plasm: Pension Liability and Asset Simulation Model. *The Journal of Finance*, 585-595.

Yang, X. (2006). Improving portfolio efficiency: A genetic algorithm approach. *Computational Economics*, 1–14.

Yang, X. (2006). Improving portfolio efficiency: A genetic algorithm approach. *Computational Economics*, 1-14.

Yu, T., & Chen, S. H. (2004). Using genetic programming with lambda abstraction to find technical. *Computing in Economics and Finance*, 8-10.

Ziemba, W. T., & Mulvey, J. M. (1998). *Worldwide asset and liability modeling.* Cambridge: Cambridge University Press.

8. Annexes

APPENDIX I – Life Table

The file in this appendix contains the UK Life Table used to simulate life expectancy.

uk_life_table.csv

Source: http://www.mortality.org/Public/ExplanatoryNotes.php#CompleteDataSeries

APPENDIX II – Historical Asset Returns

The file in this appendix contains the historical asset returns used to simulate future asset returns.

asset_returns.csv

Source: http://pages.stern.nyu.edu/~adamodar/New_Home_Page/datafile/histretSP.html

APPENDIX III – Source Code

The file in this appendix contains the SAS code used to carry out the simulations.

final_code.sas