Network for Studies on Pensions, Aging and Retirement

NETSPAR ACADEMIC SERIES

Forecasting pension release using a fund specific experience mortality model

Stijn Johannes Verhagen

MSc 12/2018-017

Forecasting pension release using a fund specific experience mortality model

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

Master of Science in Quantitative Finance and Actuarial Science

5 December 2018

Tilburg University

Tilburg School of Economics and Management

Author:

Stijn Johannes Verhagen

338055

Supervised by

dr. A.G. Balter (Tilburg University)

M. van der Werf MSc (PGGM N.V.)

Abstract

Mortality models that forecast deaths well can perfrom bad when predicting pension release. In this thesis we create two models that forecast pension release and one model that forecasts deaths. The three models are based on differences in mortality between income classes. All our data is from Dutch origin. We then develop two new backtesting methods to evaluate our models on forecasting pension release. We apply our model and on real fund data. To backtests quantify the difference between the models. Our results imply that there is no significant difference between our three models. After doing a robustness check we conclude that our results are robust.

Acknowledgements

Before going to main content I would like to desginate some space for written gratitude towards a number of people.

First of all I would like to thank my supervisor A.G. Balter for her time and extensive feedback. I felt very lucky having such a committed supervisor. I want to thank my internship supervisor M. van der Werf for his help finding a suitable topic for my thesis, his guidance and for introducing me to PGGM N.V.. I also want to thank my other colleagues at PGGM N.V..

This thesis also concludes my time as an econometrics student at Tilburg University. I am greatfull towards all my family and friends for their support during this journey. I am thankfull towards the professors, teachers, assistants and the rest of the staff at the econometrics department. Special thanks to my brother Sjoerd for mentoring through my first years. And last but not least special thanks to my supportive girlfriend Kim.

Table of contents

1. Introduction

It is important for a pension fund to get a good forecast of the pension release. Pension release is money that a pension fund does not need to pay out due to the death of the recipient(s). In this thesis we will investigate how to forecast pension release. More specific, we investigate how to adjust fund specific death probabilities such that they forecast pension release properly. The main contributions of this thesis are: two new models to predict pension release and two new backtesting methods to evaluate a models performance on forecasting pension release.

The population of a pension fund is a subpopulation of the total population. It is reasonable to assume that the death probabilities of the funds participants differ from those of the general population. The mortality models from the literature are not directly useable for a pension fund. This is due to limited historical data about the funds participants mortality. Another reason is the relatively small size of most funds compared to the population size for which the models are designed. To overcome these issues pension funds use a fund specific experience mortality factor. This factor serves as a correction on the death probabilities from a reference population. These corrected death probabilities are then suitable for the fund.

An estimate for the future amount of people passing away in a funds population is essential, but even more important is an estimate of the total amount of pension release. Mortality models are designed to estimate the amount of deaths, however these models are not directly applicable when predicting the pension release. This can be explained best by the following the example.

Take a fund with two participants: participant A with an 80% chance of passing away within a year and participant B with a 20% chance. If participant A passes away, 10 euros will be released, if participant B passes away 90 euros will be released. We now want the general probability that a participant passes away in the upcoming year. We determine this by averaging the death probabilities: $\frac{80\%+20\%}{2} = 50\%$. So both participants have on average a 50% chance of passing away. The expected amount of participants passing away in the upcoming year is $(0.5 * 2) = 1$. This is correct since $(0.8 * 1 + 0.2 * 1) = 1$ as well. Hence we can use the average chance of passing away to predict the amount of participants passing away. However, if we want to predict the pension release we run into difficulties. Because by using the average probability we estimate a pension release of $(0.5 \text{ x } 10 + 0.5 \text{ x } 90) = 50$ euros. Using the individual chance of passing away we find $(0.8 \text{ x } 10 + 0.2 \text{ x } 90) = 26$ euros. Here, the latter is the correct expected amount of pension release. This shows that estimates which are suitable for estimating the amount of deaths can be unsuitable for predicting pension release.

The structure of this thesis is as follows: section two discusses current experience mortality models in the literature, section three gives an overview of the data we use, section four and five discuss the methodology behind our model and tests respectively, section six displays our results, section seven discusses a robustness check on our results and section eight concludes our research.

2. Literature review

Before we review the literature we will first define experience mortality using equation (1):

$$q_{x,t}^{subpopulation} = e_{x,t} * q_{x,t}^{population} \quad (1)$$

Here $e_{x,t}$ is the experience mortality for age x in year t, $q_{x,t}^{subpopulation}$ is the death probability for age x in year t of the subpopulation, and $q_{x,t}^{population}$ is the death probability[1] of age x in year t of the population.

The first model we will discuss is the Cox Proportional Hazard model from Brouhns and colleagues (2002). They determine the death probabilities of annuitants in proportion to the whole population. His model is specified as follows:

$$ln(q_{x,t}^{PORT}) = \theta_1 + \theta_2\, ln\left(q_{x,t}^{POP}\right) + \epsilon_{x,t} \quad (2)$$

Where $q_{x,t}^{PORT}$ are the death probabilities of the portfolio, $q_{x,t}^{POP}$ are the death probabilities for the whole (Belgian) population, and $\epsilon_{x,t}$ is an independent identical distributed error term with zero mean and σ_ϵ^2 variance. Downside of this model is that it is based on a linear model, but there is no evidence that the relationsship is linear.

Pitacco and colleagues. (2009) determine the portfolio specific death probabilities using the following relation:

$$\ln\left(q_{x,t}^{PORT}\right) = f(x) + \ln\left(q_{x,t}^{POP}\right) + \psi_{x,t} \quad (3)$$

Where $\psi_{x,t}$ is an independent identical distributed error term with zero mean and σ_ψ^2 variance, $q_{x,t}^{POP}$ is the total population from the country where the fund is located. The age dependent constant $f(x)$ determines the relationship between the portfolio and the population in this model. Since $f(x)$ is time invariant they assume this relationship is constant over time. The validity of this assumption is doubtful since Chetty and colleagues (2016) find that the difference in life expectancy between poor and rich in the United States is increasing.

[1] Instead of saying “chance of passing away” we use death probability which is dependent on age. In other words, the chance someone passes away between two birthdays.

Villegas and Haberman (2014) also use a reference population and a subpopulation. They first determine the death probabilities of groups in the reference population. These groups are divided by characteristics. As characteristics they pick five socio-economic classes. For each class they determine the death probabilities using a Lee-Carter + cohort model. This model is defined as:

$$ln(q_{x,t}) = \alpha_x + \beta_x \kappa_t + \gamma_{t-x} \quad (4)$$

Where α_x is a constant for age x, β_x is the time coëfficiënt for age x, κ_t is the time trend and γ_{t-x} the cohort effect. For each class $v \in 1, \dots, 5$ they determine the death probabilities as:

$$ln\,(q_{x,t,v}) = \alpha_{x,v} + \beta_x \kappa_{t,v} + ln\,(q_{x,t}^{ref}) \quad (5)$$

Where $q_{x,t,v}$ is the death probability of someone age x, in year t from class v and $q_{x,t}^{ref}$ the death probability of someone age x in year t of the reference population. To find the death probabilities of the subpopulation they project the death probabilities of the classes on the subpopulation. In order for this method to work, one has to pick characteristics which fully cover the whole subpopulation. Besides this application they also provide a general framework for applying this method using other characteristics. Our model is a variant on this approach. We will discuss it further in section four.

At last I want to briefly discuss Plat (2009). He focusses on the portfolio specific factors ($P_{x,t}$ which determine the difference between population and portfolio mortality.

$$P_{x,t} = \frac{q_{x,t}^{PORT}}{q_{x,t}^{POP}} \quad (6)$$

Plat (2009) estimates $P_{x,t}$ as a linear trend on age:

$$P_{x,t} = a_t + b_t x + \epsilon_{x,t} \quad (7)$$

Where $\epsilon_{x,t}$ is an independent identical normal distributed error term with zero mean and σ_ϵ^2 variance. Using fourteen year of data he calibrates his parameters with the linear regression from equation (7). This model is not useable for fund without a decent amount of historical data. This finishes our literature review, we will now discuss our data in the next section.

3. Data

This section discusses the data that we use for our investigation. We make use of data from the following four sources: 'CBS', this is the Dutch national statistics office, 'DNB' this is the Dutch central bank, 'A.G.' this the actuarial society and at last we use fund data. We discuss each source separately.

3.1 CBS data

Our first source is the CBS, they provide information about the national population and the national mortality. The dataset is called "160614 Maatwerk TW Sterftetabellen 2008-2013" and is publicly available. It contains national population size for ages between zero and one-hundred year and is grouped by gender. The data is further broken down to people with and without income. Income is money obtained from labor or pension rights. Social welfare is not seen as income. In 2013 the Netherlands had a total population of 16.8 million. The figure 1 shows the total amount of people with income and the total amount of people without income. People without income are those who did not receive money or a pension fund. Income from capital is unknown in this dataset. The difference of the two lines consists of people without income or unknown income.

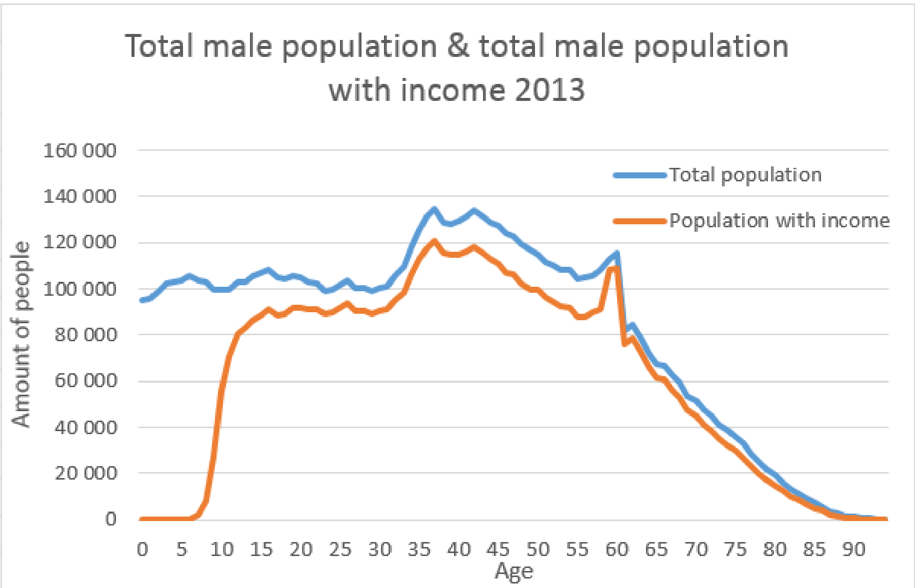

Figure 1: Total population and total working population

The data is divided into six income classes. These classes are determined such that the first six-tile of the population with income is in the first income class and the last six-tile of the population with income is in the sixth income class. People without income are excluded when deterring the income classes. The population without income can be seen as income class 0. This class is not restricted to having a certain percentage of population in it. Based on the way the income groups are divided, each income group for a certain age has a minimum and maximum amount of income assigned to it. These amounts represent the border between two income classes.

The dataset also contains information about the mortality of the Dutch population between 2008- 2013. It shows the amount of deaths for each age split by income class. Unlike the data about population size and income, our mortality data is not available for a specific year. It is accumulated over 2008-2013. So the amount of deaths for each age and income class is the accumulated amount of deaths over these six years. Therefore we could for example see that in total 1,000 males aged 30 in income class-4 passed away between 2008 and 2013. But we do not know how many 30 year old males from income class-4 died in 2010. We only know it is less or equal to 1,000. Figure 2 below shows 10% of the accumulated male population and the accumulated amount of male deaths of those years for each age. The accumulated death curve is too flat to be informative if we use 100% of the actual accumulated population.

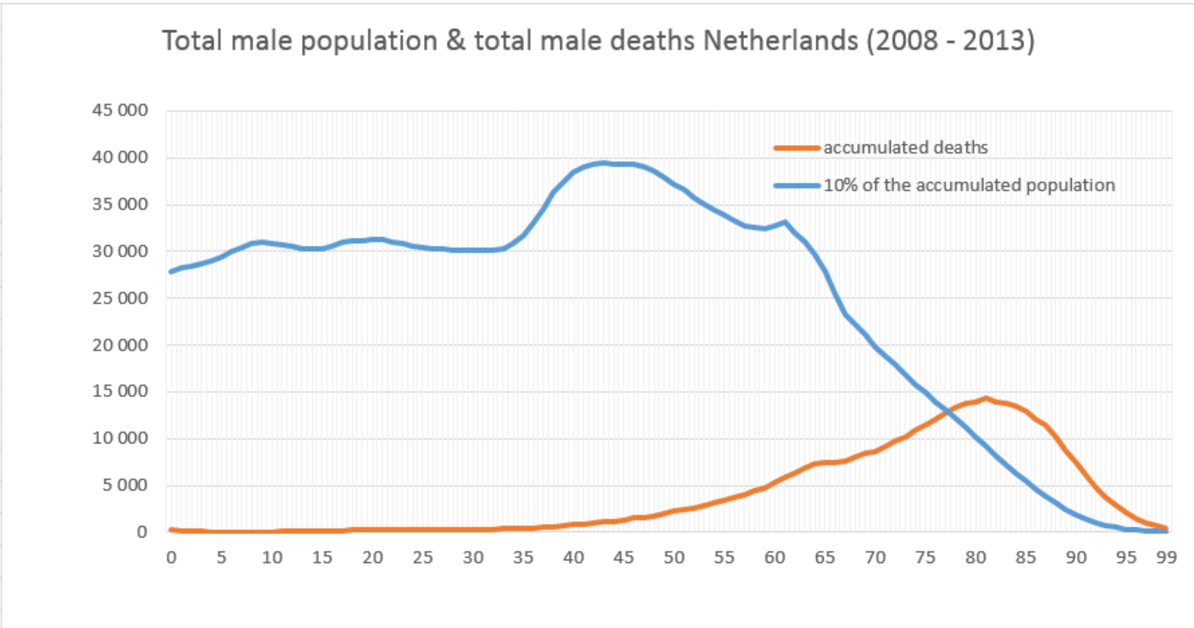

Figure 2: Accumulated amount of males who passed away between 2008-2013 And 10% of accumulated males over 2008-2013.

Figure 2 shows a steady decline in population starting from 60 years. The deaths are relatively high for 0 and 1 year olds. Starting from the age of 30 we see a clear increase in amount of deaths. It is generally known that mortality rates increase with age. In our data this increase in mortality rates outweighs the decrease in total population up to the age of 88. From then on the total amount of deaths decreases. By dividing the accumulated amount of deaths by the accumulated population per age we find the average death probability per age. Figure 3 below shows these probabilities.

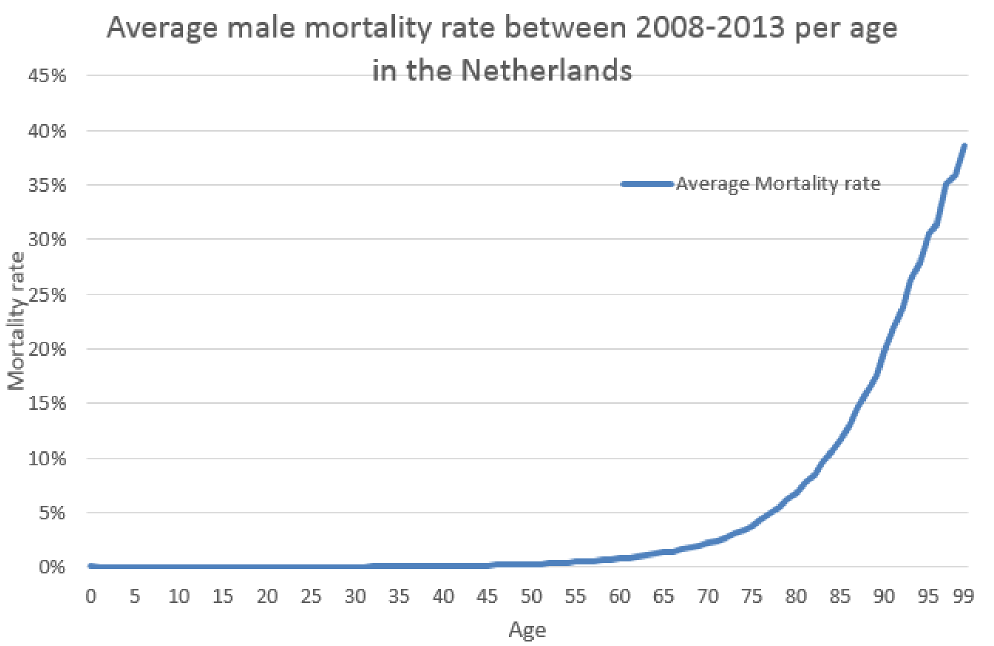

Figure 3: Average percentage of people passing away yearly for all ages between 2008-2013

Up to the age 53 the probability on passing away is less than 1%. From then on we see an exponential increase in death probabilities. This finishes our first data source, we will now discuss the DNB data.

3.2 DNB data

The second source is the DNB, we use their interest rate term structure. We chose them because pension funds are obliged to use it as well. We need this data in order to determine the present values of liabilities. We will discuss the use of the data further in section five. The data from 2016 contains the zero coupon interest rates up to the year 2216. These interest rates are used to determine the discount factors Figure 4 below shows the interest term structure:

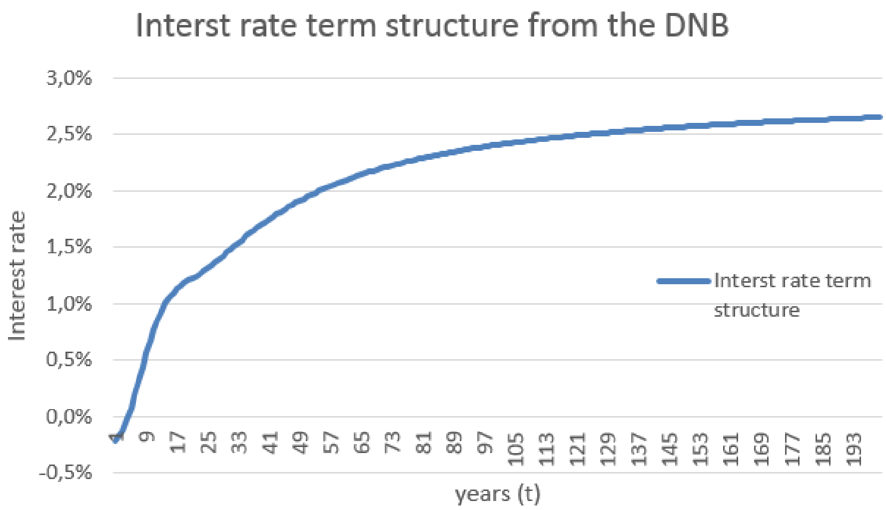

Figure 4: Interest rate term structure provided by the DNB

The figure shows a term structure which converges towards 2.7%. We now go to our second data source.

3.3 A.G. data

We use the forecasted death probabilities produced in 2016 by the A.G.. These probabilities are based on the multi-population mortality model as proposed by Li and Lee (2005). The A.G. updates their forecast every two years. During the construction of this thesis the A.G. did such an update. But it was too late for us to use the updated data. The data contains a forecast for death probabilities per age and gender for the Dutch population. This forecast is up to the year 2216, this means the data contains the forecasted death probabilities per age and gender for each year between 2016 and 2216. Figure 5 shows the forecasted one year death probabilities for males in 2018:

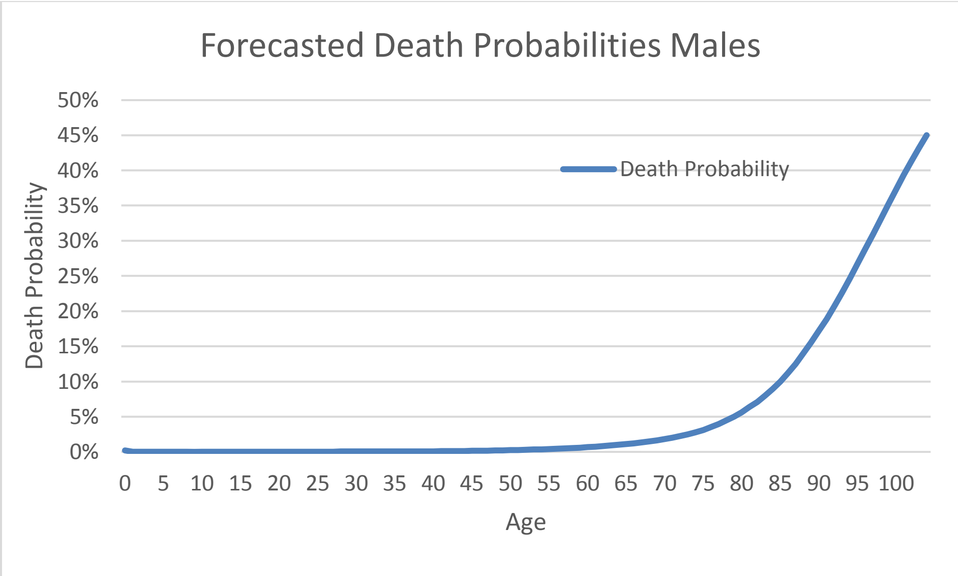

Figure 5: Estimated death probabilities in 2018 for males by the A.G.

Even though figure 3 represents male mortality between 2008-2013 it looks similar to figure 5. We now go to our last data source.

3.4 Fund specific data

At last this thesis makes use of fund specific data which is provided by PGGM N.V.. This dataset contains approximately 110.000 participants over 5 years between 2013 and 2017. Only 8% of participants is female, this means we only have around 10.000 female observations per year. It is generally known that the death probabilities of females and males differ substantially. So we have to split our sample into a female and male sample. Because the female sample size is too small we will only look at the male population in the fund. The fund specific data contains the following variables: “Gender”, “Day of Birth”, “Income”, “Elderly Pension”, “Partner Pension”, “Year” and “Passed away (yes or no)”. Based on the income classes of the CBS data we can attach the “income class” dimension to this data. One side note is that the salaries of those income classes are from 2008-2013 and our pension fund data is between 2014-2017. So we do not know the amounts of incomes that would determine the border between the income classes. Moreover the CBS data covers the years following the 2008 financial crisis. It is reasonable to believe that the income borders in 2013-2017 differ from those between 2008-2013. Figure 6 shows how the average borders in the CBS data evolved between 2008-2013.

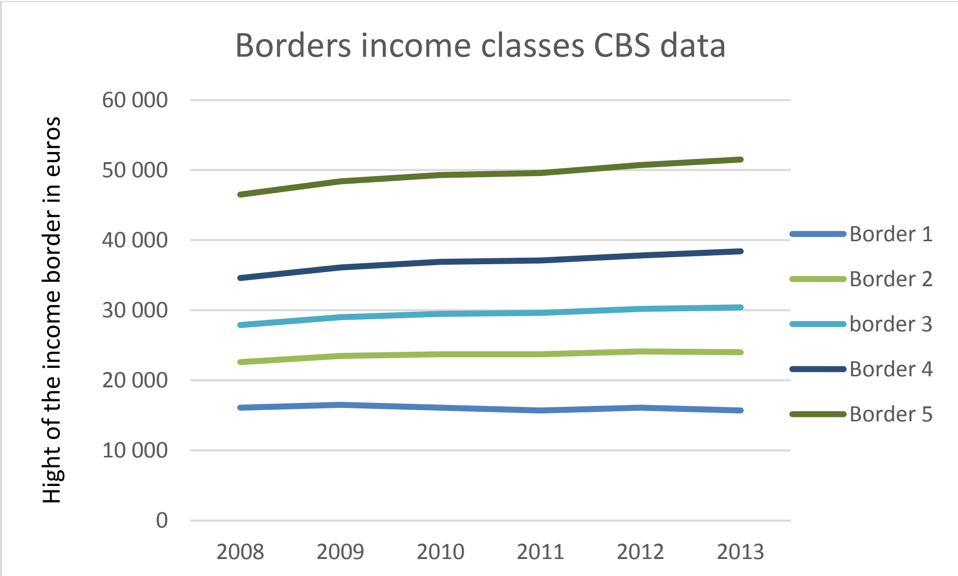

Figure 6: Evolution of the six-tile borders between 2008-2013

In order to determine the income class borders between 2013-2017 we index the income class borders obtained in 2013. Our indexation equals the consumer price index (CPI) provided by CBS. Table 1 shows the CPI for the Netherlands between 2013 and 2017:

2013	*2014*	*2015*	*2016*	*2017*
2.5%	1.0%	0.6%	0.3%	1.4%

Table 1: The CPI between 2013 and 2017

To get better insight into the data we will provide some descriptive statistics. The figure 7 shows the number of male participants for each age:

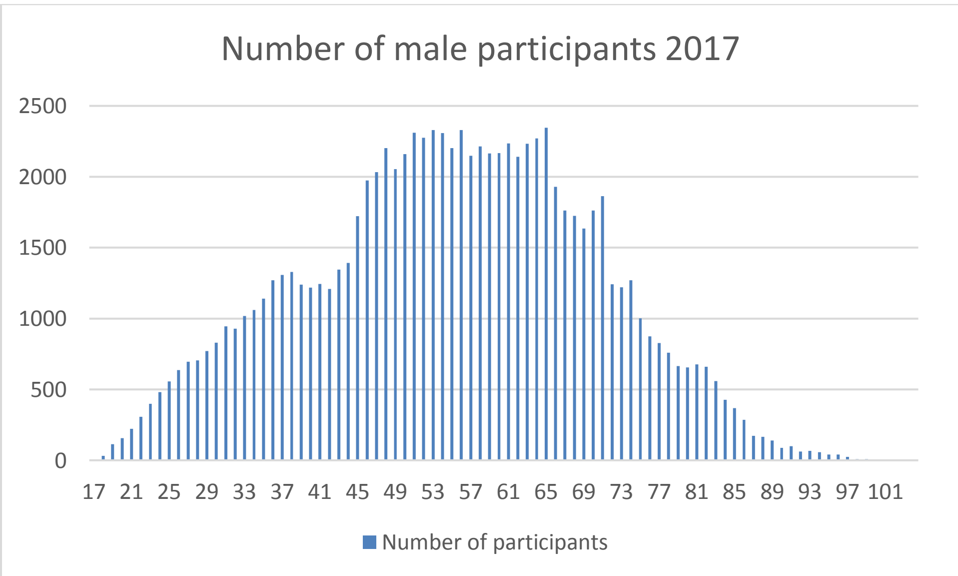

Figure 7: Number of male participants per age in the fund

The figure shows that the bulk of participants is between 45 and 65 years old. So there are relatively many old people in this fund, this means the fund is greying.

Elderly Pension is the yearly amount of money a participant receives starting from retirement, partner pension is the yearly amount of money a participant's partner obtains starting from retirement if the participant passes away. The average elderly pension and partner pension by age for year 2017 is depicted in the figure below.

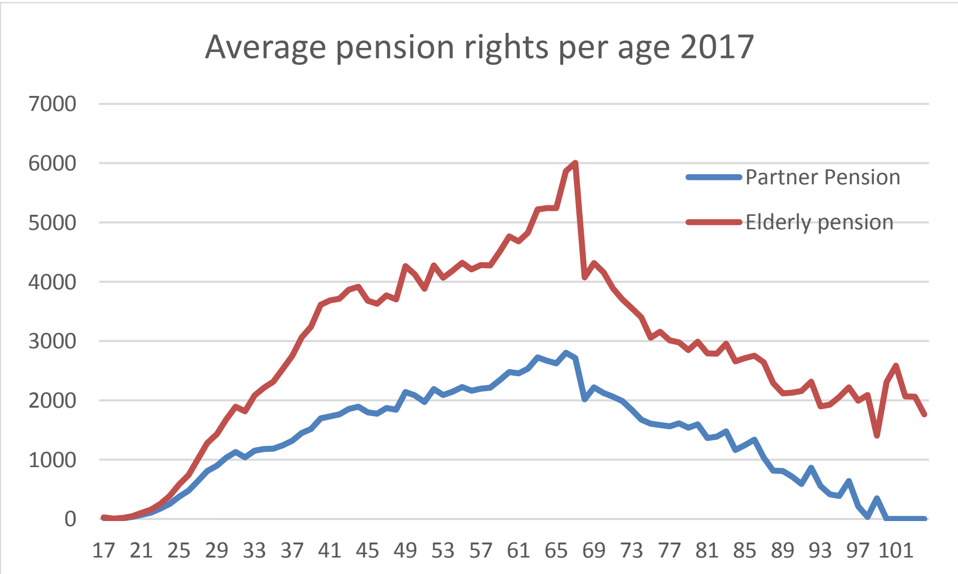

Figure 8: Average amount of Elderly Pension and Partner Pension per age in the fund anno 2017.

As can be seen in the figure, 67 years olds receive on average 4000 euros a year. The average yearly amounts seem low, but take into account that this is also averaged over people who recently entered the pension fund and have little rights. The figure below shows the average salary for men below the age of 65:

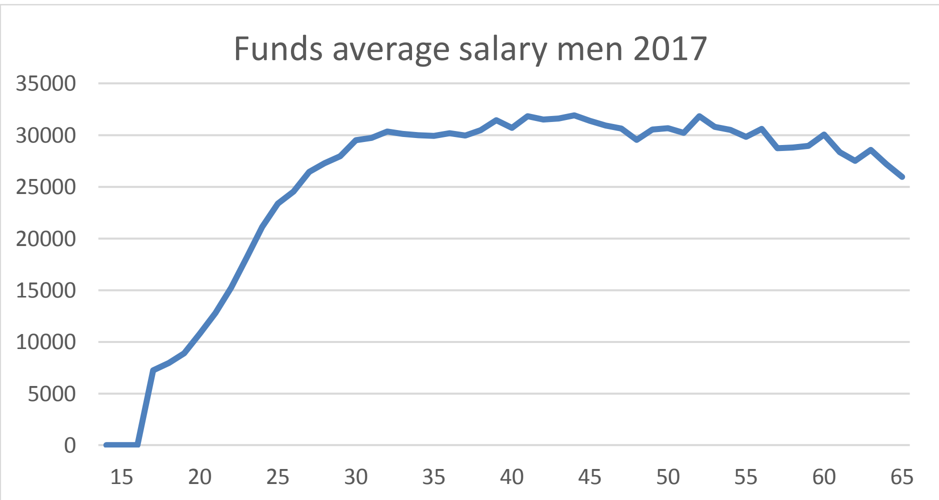

Figure 9: Funds average salary men in 2017

The figure shows a steady increase until the age of 35. After that the average salary is stable until 55 when it starts declining. Most of the participants do hard labor. So a possible explanation is that older participants start working part time because they need more rest. This finishes our data description, we will now continue to the next section in which our model will be discussed.

4. Experience Mortality Model

Our model is based on the 'characterization approach' described by Villegas and Haberman (2014). In line with this approach we determine the experienced mortality of our pension fund based on the characteristics of a reference population. As a reference population we take the total Dutch population over 2008-2013. This population is a lot larger and therefore more reliable than the pension fund population. However we cannot directly project the mortality rates of this population on our fund population. We expect to get better results if we project so called characteristics of the Dutch population on our fund population. It is important to choose characteristics which are closely correlated with mortality and which are applicable on the fund population. For example, you want to determine the mortality rates of a small population of Cuban cigar factory employees. All of these employees happen to be moderate smokers, then it would be reasonable to project the mortality rates of Cuban smokers on the cigar factory employees. The characteristic is then smoking. Since smoking leads to higher mortality rates, the subpopulation (the cigar employees) should have higher mortality rates as well. If we would not use this characteristic we expect to underestimate the mortality rates. Take the same example but now 80% of the employees smoke. This time it might be better to project the mortality rates of Cuban smokers on the 80% smokers and assign the mortality rates of Cuban non-smokers to the remaining 20%. We now use two characteristics, smokers and non-smokers.

We use the income classes as characteristics. Our CBS data provides us with the mortality rates of different income classes. According to Van Berkum (2018) income is one of the most relevant factors that determine death probabilities in pension portfolios. Kulhánová and colleagues (2014) and Von Gaudecker (2007) also identify income as an important factor that explains death probabilities. So we consider it reasonable to use income classes. As mentioned, the CBS data is segmented in seven income groups. One of those income groups is for people without any income. We will not use the latter segment because each participant has income when he is in the pension fund. When a participant leaves the pension fund we lose track of him or her. Therefore we assume that participants who leave the pension fund will stay in the same income class as they had when they left.

We determine the experienced mortality (e) for each separate income class from our CBS dataset. We then multiply the experienced mortality with the death probabilities from the A.G. ($q_{x,t}^{A.G.}$) in order to obtain $q_{x,t,v}^{Class\ (v)}$, this looks as follows:

$$q_{x,t,v}^{Class\ (v)} = e_v * q_{x,t}^{A.G.} \tag{8}$$

Where e_v is defined as:

$$e_v = \frac{q_{x,v}^{ref}}{q_x^{ref}} \tag{9}$$

Here $q_{x,v}^{ref}$ is the mortality rate for age x and class v in the reference population. And q_x^{ref} is the death probability of the whole population of age x. We calculate $q_{x,v}^{ref}$ as follows:

$$q_{x,v}^{ref} = \frac{\sum_{j=2008}^{2013} D_{j,x,v}^{ref}}{\sum_{j=2008}^{2013} T_{j,x,v}^{ref}} \tag{10}$$

Where $D_{j,x,v}^{ref}$ is the amount of deaths in the reference population for year j, age x and class v. And $T_{j,x,v}^{ref}$ the total amount of population in the reference population in year j for age x. We calculate q_x^{ref} in a similar way only without a distinction in class.

$$q_x^{ref} = \frac{\sum_{j=2008}^{2013} D_{j,x}^{ref}}{\sum_{j=2008}^{2013} T_{j,x}^{ref}} \tag{11}$$

For simplicity we stick with experience mortality instead of the death probabilities. The figure below summarizes the experience mortality e_v for the seven different income classes.

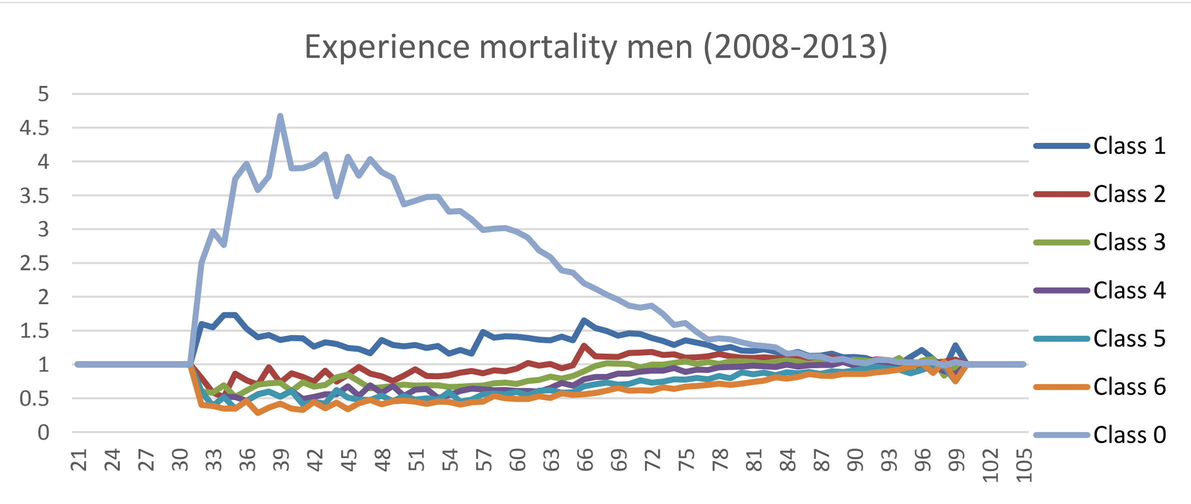

Figure 10: National experience mortality men per age and income class.

We set the experienced mortality of people below the age of 30 and above 99 equal to 1. Since there are only few deaths among young people in the CBS data, these ages are too volatile. Thereby young people do not have much impact on the release of pension rights since their small pension claims are discounted over $37+$ years. The figure shows a high experience mortality for people without income. The high magnitude of this class makes the rest of the figure hard to interpret. In order to make the figure easy to read we smoothed the lines and removed income class 0. We smoothed the curve by fitting a polynomial on it. The degree of the polynomial minimizes the Aike Information Criterion[2] (AIC).

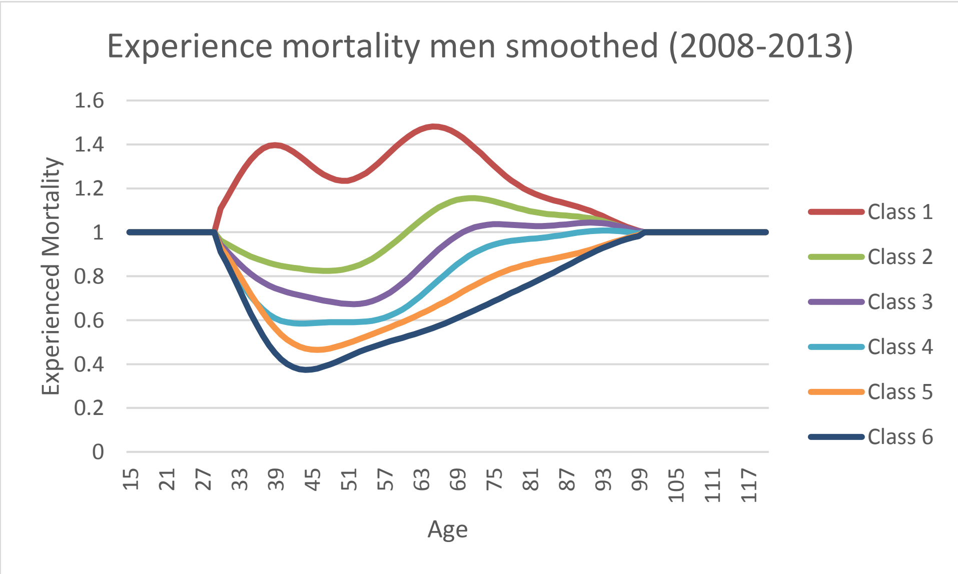

Figure 11: National experience mortality men per age and income class 1 t/m 6 smoothed.

Figure 11 shows that men in high income classes have lower experienced mortality compared to their counterparts in low income classes. For example, the probability of passing away at the age of 45 is $\frac{1.4}{0.4} = 3.5$ times as large for a man in income class-1 compared to a man in income class-6. Also experienced mortality converges to 1 as age increases. Note that these income classes are determined such that they consist a six-tile of the total population for a certain age. Since lower income six-tiles have higher death probabilities these people pass away more often. These six-tiles then have to be supplemented by higher income six-tiles in the next year in order to stay a six-tile. This causes the average salary to increase in those six-tiles which causes the experience mortality to decrease. This is partly what causes the class-1 experience mortality to converge to one as

[2] The AIC is defined as $AIC = 2k - 2\ln(\hat{L})$ where k is the number of parameters and $\hat{L}$ is the maximum value of the maximum value function. So the AIC only accepts more parameters if the likelihood increases sufficiently.

age increases. Higher income classes also converge to 1 with similar reasoning. This shifts the average death probabilities more towards the death probabilities of the higher income classes.

In line with Villegas and Haberman (2014) we now project the mortality rates of the income classes on the subpopulation, which is the pension fund. Projecting the experience mortality on the subpopulation will result in a pension fund specific mortality. We project the classes on the pension fund by taking a weighted average. A feature of our model is that we can apply different weighting methods for calculating the average. The first weighting method is based on number of people in each income class. We introduce the following indicator function $I_{i\in n_v}$, define n_v as the part of the population in income class v, then:

$$I_{i\in v} = \begin{cases} 1, & if\ i \in n_v \\ 0, & if\ i \notin n_v \end{cases} \quad (12)$$

The weight of an income class using people as weights is then determined as follows:

$$W_v^{people} = \frac{\sum_{i=1}^{n} I_{i\in n_v}}{n} \quad (13)$$

Where W_v^{people} is the weight of income class v and n is the total amount of people in the pension fund below the age of 65. We only include people below 65 because we only consider unretired people. Once people retire their income data becomes less reliable. Retired participants might also receive income from other pension funds if they switched fund during their working life. We will refer to the weighting method defined in (14) as "W^{people}". We end up with a weight for each income class. The experience mortality of the pension fund is then determined as:

$$e_x^f = \sum_{v=1}^{6} e_{x,v} * W_v \quad (14)$$

Where e_x^f is the age x experience mortality of the fund and W_v is the weight from one of the weighting methods with class v. We assume that the composition of income classes in our retired and unretired population is the same. So we use the weights also for the retired participants.

Another weighting method is by income. The weight of the income classes is then determined by the total income generated in the income class. The formula below depicts this way of calculating the weights. We will call this weighting method W^{income}.

$$W_v^{income} = \frac{\sum_{i=1}^{n} I_{i \in n_v} * s_i}{\sum_{i=1}^{n} s_i} \qquad (15)$$

Where s_i is the income (salary) of participant i and n is the population in the pension fund below 65. At last we use the pension rights to determine the weights. The weights of the income classes then depend on the present value of the pension rights in the income class. We will call this weighting method $W^{release}$ since it is the potential pension right 'release' of the participant. For now just take this concept for granted. We will provide more information on this and the way we determine the present value of the rights in the next section. The formula below shows this weighting method:

$$W_v^{release} = \frac{\sum_{i=1}^{n} I_{i \in n_v} * pr_i}{\sum_{i=1}^{n} pr_i} \qquad (16)$$

Where pr_i are the pension rights of participant i. In the rest of this thesis we will refer to these weighting methods as "W^{people}", "W^{income}" and "$W^{release}$" respectively. The results of the weighting methods are given in the table below:

	Class 1	***Class 2***	***Class 3***	***Class 4***	***Class 5***	***Class 6***
W^{people}	0.1611	0.5752	0.1737	0.0795	0.0101	0.0004
W^{income}	0.067	0.5941	0.2105	0.1121	0.0149	0.0008
$W^{release}$	0.0957	0.5714	0.2156	0.1021	0.01457	0.0007

Table 2: Calculated weights for each weighting method.

Income weighting gives larger weights to higher income classes compared to people. This is as expected since higher income classes have per definition higher income and therefore earn higher relative weight in this method. We do similar observations for $W^{release}$. People in higher income classes will on average have built up higher pension rights compared to their counterparts in lower classes. Higher pension rights in turn lead to more release. We already discussed that higher income classes have lower experience mortality, so when applying the W^{income} and $W^{release}$ we expect to find lower experience mortality compared to the W^{people}. The table also shows the low representation of high income classes in the fund. If the fund population would match the national population every income class would contain $1/6$ of the total fund population. But as we can see in the W^{people} row the highest income class is virtually empty and income class 2 contains almost $3/6$ of the population. This means that we are looking at a pension fund with most participants in the lower income classes. Using the three different weighting methods to determine experience mortality we obtain the following figure:

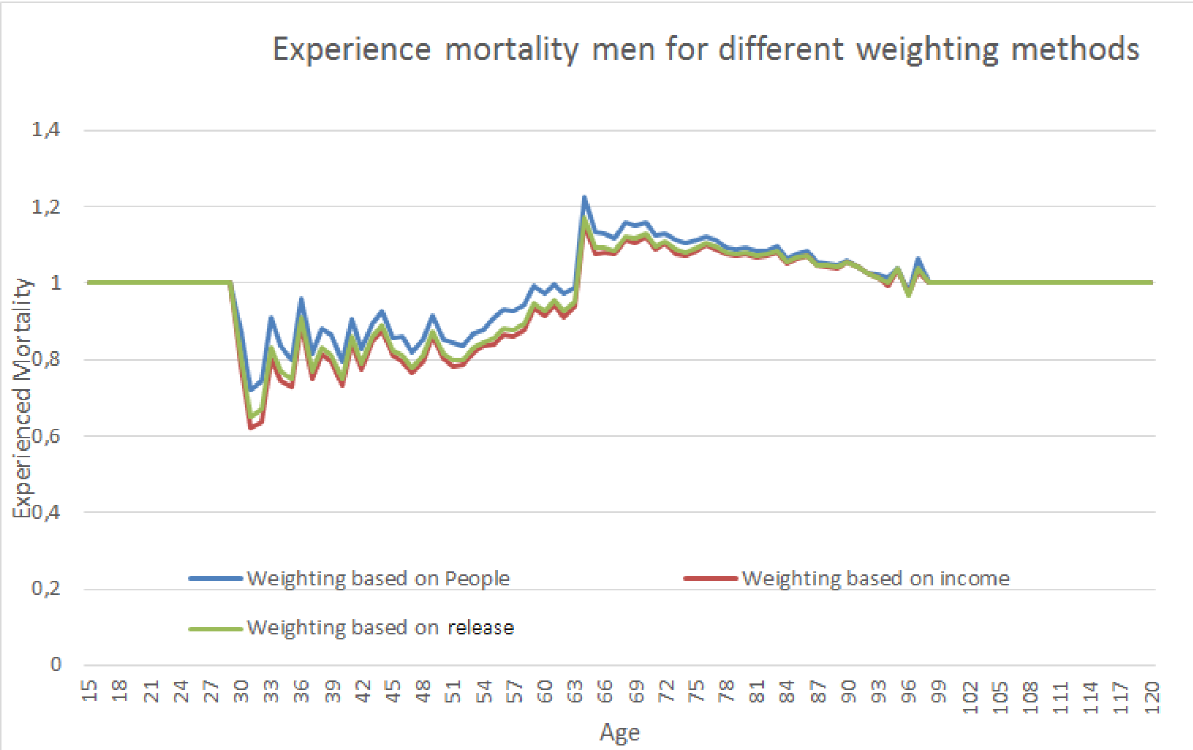

Figure 12: Fund specific experience mortality per weighting method

Figure 12 shows that W^{people} results in the highest experience mortality, $W^{release}$ results in lower experience mortality than W^{people} and W^{income} gives the lowest experience mortality, but is quite similar to W^{people}. By multiplying the experienced mortality of age x (e_x^f) from one of the three models with the death probability of age x in year t from the A.G. tables we find the death probability of participants which are aged x in year t.

From now on we will only use the death probabilities generated by the models. These death probabilities are calculated using the experienced mortality and the A.G. tables (see eq. (9)). In the next sections, we will refer to this death probability with q_x or q_i when we speak about the death probability for age x or participant i. To further investigate the weighting methods we do a number of so called 'backtests'. In the next section we will discuss the technical aspects of our tests.

5. Backtest models

In this section we discuss how to test the performance of our different experience mortalities when forecasting pension release. It is important to evaluate the performance of our models before we draw conclusions. We evaluate our models on past data from the years 2013 until 2017. This means we are performing so called 'backtests'. We propose three main backtests. To get better insight into the dynamics of the variance of these tests we also formulate some variants. Each of these main tests constructs confidence intervals around the expected pension release of our three models. If the real pension release is outside the confidence interval which corresponds to our α we reject the model. In our backtest we will use an alpha of 5%. First we will show how we determine the actual (real) pension release in the past years. Afterwards we will discuss our (back)tests. As discussed in our data section, each of our participants has a certain amount of pension rights (pr). When the participant goes into retirement these rights translate into yearly annuities until the participant dies. Once a participants passes away, his pension rights are 'released'. In order to measure the size of his release we have to calculate the present value of his pension rights. The present value depends on the size of the pension rights, the age of the participant, the interest rate term structure and the death probabilities of the participant. Take participant i age x with pension rights (or annuity size) pr_i and death probability q_x. Furthermore, take the participant to be retired. Then the present value the annuity he receives at age $x + 10$ ($R_{i,x+10}$) is calculated as:

$$R_{i,x+10} = pr_i * \prod_{y=x}^{x+10}(1 - q_y) * e^{-r_{10}*10} \qquad (17)$$

Here $\prod_{y=x}^{x+10}(1 - q_y)$ is the probability participant i is alive 10 years from now, and r_{10} is the interest rate for $t = 10$. We will test three different models which generate death probabilities q_x. So the present value of the pension rights of a participant depends on the model we are using. The present value of all annuities of participant i is the summation of all remaining years a participant might live:

$$R_i = \sum_{t=1}^{\infty}(pr_i * \prod_{y=x}^{x+t}(1 - q_y) * e^{-r_t*t}) \qquad (18)$$

Our summation is until infinity, but starting from age 121 the A.G. tables report a death probability of 100%. Our experienced mortality is also 1 in each model at that age, so after

$(120 - x)$ years participant i will die 100% if he manages to survive until then. This means the present value of those annuities is zero, so in effect the summation is limited but we will keep ∞ since it looks clearer in the formula.

For participants who are already retired we can use the above formula to calculate the present value of their pension rights. For participants below the retirement age we use the following formula:

$$R_i = e^{-r_{(67-x)}*(67-x)} \prod_{z=x}^{67}(1 - q_z) \sum_{t=67}^{\infty}(pr_i * \prod_{y=67}^{t}(1 - q_y) * e^{-r_t*t}) \qquad (19)$$

This formula calculates the present value of the annuities starting from retirement. Then it discounts the present value over the gap between the participants current age and retirement age. Note that we use 67 as retirement age. This due to the fact that the retirement is shifting smoothly towards this age. There is an amount of participants around 60 years old who will retire between 65 and 67. So our formula assigns a retirement age to these people which is too high. Also, chances are there the retirement will change again before all our participants have surpassed 67. In that case these participants will have been assigned a wrong retirement age as well. But these shortcomings affect all three weighting methods equally and will therefore not affect the comparison of our three methods.

A death will be denoted as $(D = 1)$ and surviving as $(D = 0)$. If participant i passes away $(D_i = 1)$ the present value of his pension rights R_i is released. We will use R_i' to indicate the released amount of the present value of the pension rights from participant i. The released amount R_i' depends on D_i in the following way:

$$R_i' = \begin{cases} R_i, & D_i = 1 \\ 0, & D_i = 0 \end{cases} \qquad (20)$$

The real total pension release is then calculated as:

$$R_n' = \sum_{i=1}^{n} R_i * D_i \qquad (21)$$

Where R_n' the total real pension release is of the fund with n participants. We will now discuss our backtests.

5.1 First backtest normality

The mortality process of an individual can be seen as a Bernoulli process. The probability of individual i passing away in a certain year t has probability q_i. The probability of surviving that year is then $(1 - q_i)$. This is written as:

$$P(D_i = 1) = q_i \qquad (22)$$

$$P(D_i = 0) = 1 - q_i \qquad (23)$$

The expected value of a Bernoulli variable with probability q of success is equal to q and its variance is $q(1 - q)$.

$$\mu = q \qquad (24)$$

$$\sigma^2 = q(1 - q) \qquad (25)$$

Our first backtest method is based on Van Berkum (2018) his approach. Van Berkum sees the individual release also as a binomial variable, the expected value of this variable is:

$$E[R'] = q * R + (1 - q) * 0 \qquad (26)$$

This is the present value of the pension rights multiplied by the probability those rights are released. Van Berkum sees the variance of this variable as:

$$\sigma^2 = R^2 * q(1 - q) \qquad (27)$$

So this is the variance of a normal Bernoulli variable multiplied with the squared present value of the pension rights. The expected value and variance of the pension release from the whole fund is then:

$$E[R'_n] = \sum_{i=1}^{n} q_i * R_i + (1 - q_i) * 0 \qquad (28)$$

$$var[R'_n] = \sum_{i=1}^{n} R_i^2 * q_i(1 - q_i) \qquad (29)$$

Van Berkum then assumes the total pension release is normally distributed. The confidence intervals can be constructed using the mean and variance from the expressions above and the formula:

$$100(1 - \alpha)\% \ C.I. = [\mu - z * \frac{\sigma}{\sqrt{n}}, \mu + z * \frac{\sigma}{\sqrt{n}}] \qquad (30)$$

Where z is the μ is the expected value and z is such that $\Phi(z) = (1 - \alpha)$, and Φ is the CDF of the $N(E[R'_n], \sqrt{var[R'_n]})$ distribution.

The distribution of total deaths in a year often has a bell-shaped curve. This is reasonable since a bell shaped curve gives a low probability to everyone or nobody passing away. With the same logic one could model the distribution of total pension release as a bell shaped curve. In section 5.2 we explain how we can simulate total amount of deaths in our fund. Figure 13 shows the shape of our funds total amount deaths distribution based on 15000 simulations.

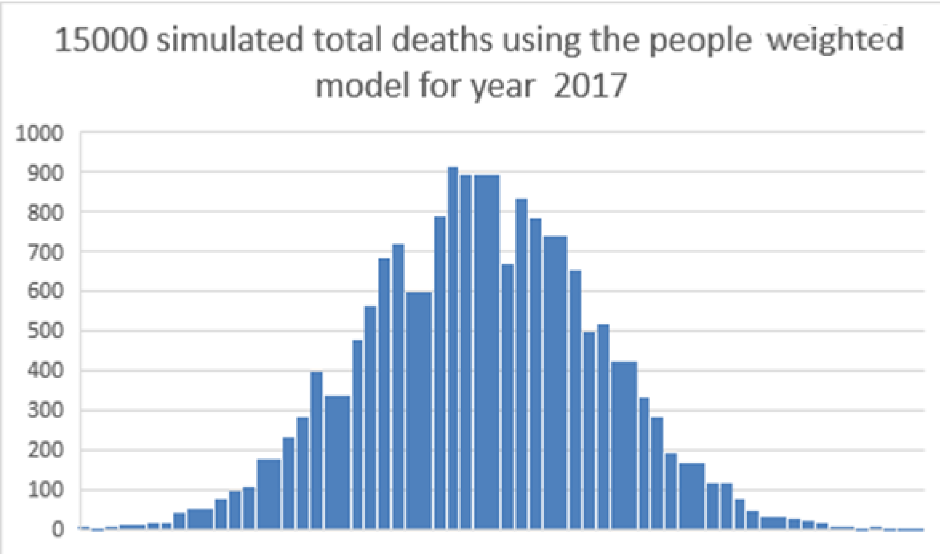

Figure 13: shape of total death distribution

We observe a bell-shape with some seemingly random gabs. The assumption that pension release is normally distributed should be tested using the Jarque-Bera test. Assman (2014) does such a test for his total pension release. The null hypothesis of this test is a joint hypothesis of the skewness being zero and the kurtosis being three. Samples from a normal distribution have an expected skewness of zero and an expected kurtosis of three. By deriving the third moment of pension release Assman (2014) finds that the pension release tends to be positively skewed. Using the kurtosis and squared skewness from the pension release distribution by Assman (2014) we also do a Jarque-Bera test on our fund data for each weighting method. The squared skewness (β_1) and kurtosis (β_2) are:

$$\beta_1[R'_n] = \frac{\left(\sum_{i=1}^{n} (r_i)^3(1-2q_i)(1-q_i)q_i\right)^2}{\sum_{i=1}^{n} (r_i)^2 q_i(1-q_i))^3} \quad (31)$$

$$\beta_2[R_n'] = \frac{\left(\sum_{i=1}^{n} (r_i)^4 (1-6(1-q_i)q_i(1-q_i)\right) q_i}{\sum_{i=1}^{n} (r_i)^2 q_i(1-q_i))^2} + 3 \qquad (32)$$

For the derivation of these I refer to the master thesis of Assman (2014). The Jarque-Bera test is given as follows:

$$JB = \frac{n}{6}\left(\beta_1 + \frac{1}{4}(\beta_2 - 3)^2\right) \qquad (33)$$

If the data comes from a normal distribution, then the JB statistic asymptotically has a χ_2^2-distribution. The statistic increases if the skewness deviates from 0 and if the kurtosis deviates from 3. The critical value for a 1% significance is 9.21, so the null hypothesis for normality will be rejected if $JB \geq 9.21$

We find the following values for $\beta_1[R]$ and $\beta_2[R]$ based on our W^{people} from year 2017. In this year we have a total population of 91,548 males:

$$\beta_1[R] = 0.0059077$$

$$\beta_2[R] = 3.0073839$$

So our JB statistic is:

$$JB = \frac{91548}{6}\left(0.0059077 + \frac{1}{4}(3.0073839 - 3)^2\right) = 90.35$$

This means we reject our null-hypothesis of normality. We repeat the above test for our other two weighing methods, we find the following values for W^{income}:

$$\beta_1[R] = 0.0061719$$

$$\beta_2[R] = 3.0077389$$

$$JB = 94.40$$

And for $W^{release}$:

$$\beta_1[R] = 0.0061073$$

$$\beta_2[R] = 3.0076517$$

$$JB = 93.41$$

Our findings are in line with Assman (2014), he also finds a positive skewness, a kurtosis larger than three and rejects the null-hypothesis with a JB value of 62.99.

The Jarque-Bera test fails, so we will not rely on the normal assumption for our model evaluation. Because it is used in the literature we will evaluate this method further by comparing its output with the other backtests. We will refer to this backtest as the "normality backtest" since it is based on the normality assumption.

5.2.1 Second backtest simulation

We now introduce a second backtesting method. Our models assign death probabilities $q_i \in (0,1)$ to participants $i \in (1, \dots, n)$ in our funds populations. Assuming our model is true the probability participant i dies can be denoted as:

$$P(D_i = 1) = q_i$$

We now draw a variable from an $unif(0,1)$ distribution. The probability this variable is smaller or equal to q_i is also q_i, so take $u_i \sim unif(0,1)$ then:

$$P(D_i = 1) = P(u_i \leq q_i) = q_i \tag{34}$$

By drawing one u_i we can simulate the death of participant i by:

$$D_i = \begin{cases} 1, & u_i \leq q_i \\ 0, & u_i > q_i \end{cases} \tag{35}$$

The probability the present value of the pension rights is released is also q_i. We can write the expected value $E[R_i']$ as:

$$E[R_i'] = q_i * R_i + (1 - q_i) * 0 = P(u_i \leq q_i) * R_i + P(u_i > q_i) * 0 \tag{36}$$

By drawing one u_i we can simulate the realised pension release of participant i directly in the following way:

$$R_i' = \begin{cases} R_i, & u_i \leq q_i \\ 0, & u_i > q_i \end{cases} \tag{37}$$

Or we can simulate it indirectly by first simulating the D_i of participant i and then apply:

$$R_i' = R_i * D_i \tag{38}$$

For each participant in the fund we draw one u_i, we then define indicator function $I_{u_i \leq q_i}$ as:

$$I_{u_i \leq q_i} = \begin{cases} 1, & u_i \leq q_i \\ 0, & u_i > q_i \end{cases} \tag{39}$$

Then total simulated pension release of the fund is determined as:

$$R'_n = \sum_{i=1}^{n} I_{u_i \leq q_i} * R_i \tag{40}$$

Also, we can simply simulate the total amount of deaths D'_n of the pension with equation (42):

$$D'_n = \sum_{i=1}^{n} I_{u_i \leq q_i} \tag{41}$$

Our backtest simulates the total pension release m times for a given year. Under the assumption that model which assigned the q_i's is true, our real pension release should be a possible outcome of this simulation. By taking an interval around the mean which consists 95% of our m simulations we produce a 95% confidence interval. If real pension release is within the 95% confidence interval our backtest passes and the model is validated. Using equation 42 this method can also be used for backtesting the total amount of deaths. We have shown how we can simulate total pension release distribution and the amount of deaths distribution. We refer to these simulations as "simulation of pension release" and "simulation of deaths". In section 5.2.2 we will discuss another simulation.

5.2.2 Backtest 2 Simulation with constant deaths.

The total pension release depends on two things, firstly the amount of deaths in the pension fund and secondly the present value of the pension rights each individual who passed away has. Take for example a fund with a five participants: A, B, C, D and E. The table below summarizes the present value of the pension rights each individual has:

Participant	q_i	R_i
A	20%	10
B	20%	25
C	20%	30
D	20%	50
E	20%	90

Table 3: Example: q_i and R_i of the participants

Now three of our participants pass away. Given this the minimal release is obtained when A, B and C pass are the ones who passed away. This leads to 65 euros release. The maximum pension release is obtained when C, D and E pass away. This leads to 205 euros release. In total there are $\binom{5}{3}$ = 10 combinations possible. Because all q_i are equal each combination of three participants has the same probability of happening, this is precisely $1/10$. The distribution of total pension release given three deaths is as summarized in figure 14:

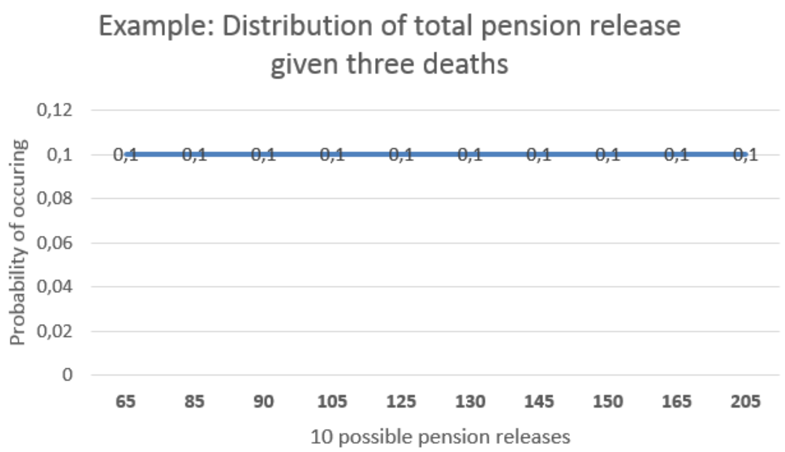

Figure 14: Example distribution of total pension release given 3 deaths

If every participant had the same present value of pension rights then there would only be one outcome given three deaths. Then the variance in total release would only depend on the variance in amount of deaths. But due to differences in present value of pension rights, the total release also depends on who passes away. So we identify two factors which cause total pension release to variate. On one hand we have the variance caused by amount of deaths, on the other hand we have variance caused by the differences in present value of the pension rights. To simulate the latter we have developed the following method:

We are again going to simulate pension releases, but this time we keep the amount of deaths constant. Our models assign death probabilities to ages. So the population participants age x (n_x) have the same death probability q_x, so $\forall i \in n_x;\ q_i = q_x$. Then the expected amount of deaths in n_x is:

$$E[D_x] = q_x * n_x \qquad (42)$$

For each age x in our population we take a random sample size $E[D_x]$. We then add these random samples into one big random sample size $E[D_n]$. Every participant i in the sample gets ($D_i = 1$) assigned, every participant outside the sample gets ($D_i = 0$). Because the samples taken from D_x are random this is a simulation as well. In most cases $E[D_x]$ is not integer. We solve this by giving one participant in the population n_x a value of D_i equal to the amount $E[D_x]$ exceeds the first integer below $E[D_x]$. Our simulated total pension release is:

$$R'_n = \sum_{i=1}^{n} R_i * D_i \quad (43)$$

We repeat this process k times giving us k possible pension releases given $D_n = E[D_n]$. Since the amount of deaths is constant the only source of variance are the differences in R_i. We will call this backtest "simulation of pension rights".

This finishes our second backtest. We will now continue to the third and last backtest.

5.3 Third backtest Poisson-Binomial

As discussed in the first backtest, the mortality process of an individual can be seen as a Bernoulli process. The mortality process of the whole fund can then be seen as n Bernoulli trails. We assume the trails to be independent meaning no correlation between deaths. The mean and variance of the funds mortality process is then:

$$\mu_n = \sum_{i=1}^{n} q_i \quad (44)$$

$$\sigma_n^2 = \sum_{i=1}^{n} q_i(1-q_i) \quad (45)$$

Our population consists of males between 17 and 103 years old. In our model each age group has a different mortality. This means we have 86 different q's split among approximately 100.000 participants. Because the q's differ, we are dealing with a Poisson-Binomial distribution instead of an ordinary Binomial distribution. The Binomial distribution is a special case of the Poisson Binomial distribution where all q's are equal. We used the work of Hong (2013) to find the probability mass function (PMF) and the cumulative distribution function (CDF) of this distribution. The PMF of this distribution is:

$$P(D_n = M) = \sum_{A \in F_M} \prod_{i \in A} q_i \prod_{j \in A^c} (1-q_j) \quad (46)$$

Here F_M is the set of all possible events which result in M deaths. The probability that M occurs is the summation of the probability of each event in F_M. The number of events which result in M deaths in a populations size n is $\binom{n}{M}$. For example, if we have a population size 3 consisting of three people, person A, person B and person C. Then there are three events which result in two deaths, namely: {(A,B}, {A,C} and {B,C}. We assign the following death probabilities: $q_A = 0.1$, $q_B = 0.2$, $q_C = 0.5$. Then using the PMF we can calculate the probability two deaths occur as follows:

$$P(D_n = 2) =$$

$$\big(0.1 * 0.2 * (1 - 0.5)\big) + (0.1 * 0.5 * (1 - 0.2) + \big(0.2 * 0.5 * (1 - 0.1)\big)$$

$$= 0.14$$

We now go to the CDF of this distribution. The CDF is the summation of PMF's up to the point where the CDF is evaluated, this looks as follows:

$$P(R_n' \leq M) = \sum_{g=1}^{M} \sum_{A \in F_g} \prod_{i \in A} q_i \prod_{j \in A^c} (1 - q_j) \qquad (47)$$

If we want to find $P(D_n' \leq 2)$ we have to add the probability of $(D_n = 1)$ and $(D_n = 0)$ to the previous example. The set F_g then becomes ({A,B}, {A,C} ,{B,C},{A},{B},{C},{}), here "{}" refers to the null set (i.e. the empty set). Using this CDF we can assign quantiles to amounts of deaths. Depending on α we can construct a confidence interval in which the realized amount of deaths should occur with $(1 - \alpha)\%$. This method is computational very costly. We have 100.000 participants in our sample, so if want to know the probability there are 1000 deaths we have to evaluate $\binom{100.000}{1000}$ events.

Besides deaths we are also interested in testing our models pension release prediction. The probability that the pension rights of a participant are released in year t is equal to the death probability of that participant. So the expected pension release $(\mathrm{E}[\mathrm{R}_\mathrm{i}'])$ of the participant is then:

$$E[R_i'] = q_i * r_i \qquad (48)$$

Where r_i is the amount of pension rights owned by participant i. The expected pension release of the total fund is then:

$$E[R_n'] = \sum_{i=1}^{n} q_i * r_i \qquad (49)$$

The probability mass function of the pension release is:

$$P(R_n' = X) = \sum_{A \in F_X} \prod_{i \in A} q_i \prod_{j \in A^c} (1 - q_j) \qquad (50)$$

Where $P(R_n' = X)$ is the probability to total release is equal to X. Now F_X is the set which contains all possible events which result in exactly X pension release. So whether an event is part of F_X depends on the pension release of the participants of that set. The probability of the event depends on the death probabilities of all participants. Returning to our example with person A, B and C. If person A has 5000 pension rights, person B has 6000 pension rights and person C has 11.000 pension rights. If $P(R_n' = 11.000)$ then $F_{11.000}$ = ({A,B}, {C}). The probability of this occurring is:

$$\begin{aligned} \mathrm{P}(\mathrm{R} = 11000) &= 0.1 * 0.2 * (1 - 0.5) + 0.2 * (1 - 0.1) * (1 - 0.5) * (1 - 0.5 * 0.1) \\ &= 0.0955 \end{aligned}$$

The CDF is again the summation of the Probability mass functions:

$$P(R_n' \leq M) = \sum_{g=1}^{M} \sum_{A \in F_g} \prod_{i \in A} q_i \prod_{j \in A^c} (1 - q_j) \qquad (51)$$

Here F_g is the set which contains all events which result in a pension release equal to or below M. Returning to the example, for $(\mathrm{R}_\mathrm{n}' \leq \mathrm{M})$ the events ({A,B}, {C}, {A}, {B}, {}). Using this CDF we can assign quantiles to amounts of pension release. Depending on α we can construct a confidence interval in which the realized amount of pension release should occur with $(1 - \alpha)\%$. We determine the realized amount of pension release by calculating the present value of the pension rights.

We will call this backtest the “Poisson-Binomial backtest”. We will refer to the pension release part as “Poisson-Binomial pension release”. The backtest on deaths will be called “Poisson-Binomial deaths”. This finishes section four. We will now discuss our results in section five.

6. Results

In this section we present the results of our backtest. We implemented our model and backtests in the statistical program "R". For the calculation of our third backtest we used the R-package "Poisbin" based on Hong (2013) which we also used for designing the backtest.

6.1 Normality backtest

The boxplots below summarize the results of the backtest based on the normality assumption:

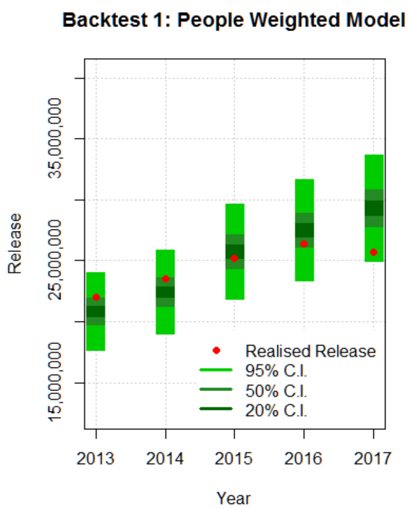

Figure 15

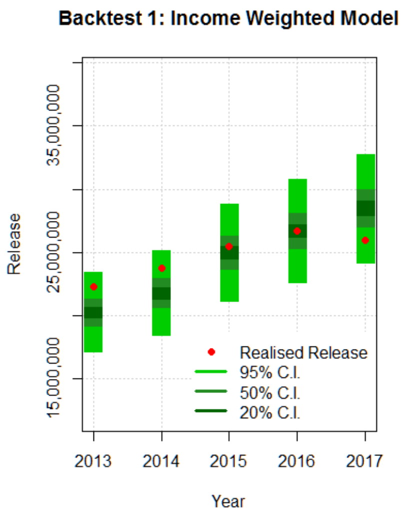

Figure 16

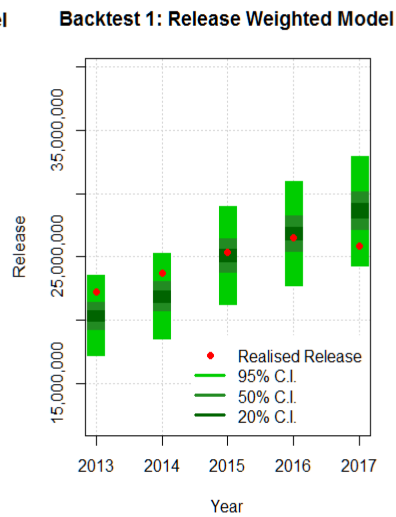

Figure 17

The real pension release is within the 95% confidence interval every year for all methods. Looking at it closely one can notice the real pension release differs for each method. This is the case since the real pension release depends on the death probabilities, and these differ for each method. The 'People Weighted Model' returns the highest death probabilities, because of the way the present value is calculated the real releases are lowest here. With the same reasoning the expected releases of this model should be lower than the other models. But these higher death probabilities increase the probability a participant his pension rights are released. The results show that the expected pension release of the 'People Weighted model' exceeds the other models in every single year. So the higher death probabilities increase the expected amount of pension release. The models produce similar results. The lower and upper bounds of the confidence intervals from the People weighted Model are slightly higher than those of the other two models. This means that W^{people} forecasted more pension release than the other models. Also, the size of the confidence intervals are larger for the 'People Weighted Model'. This can be explained by the way the variance is calculated in equation 8:

$$var[R] = \sum_{i=1}^{n} R_i^2 * q_i(1 - q_i)$$

Because the People Weighted Model assigns higher death probabilities (q_i) the $q_i(1 - q_i)$ term is larger compared to the other models.

6.2 Simulation backtest

The second backtest gives the following results. We did 10,000 iterations per year for each model. The distribution of those 10,000 releases for $W^{release}$ is shown in figure 18:

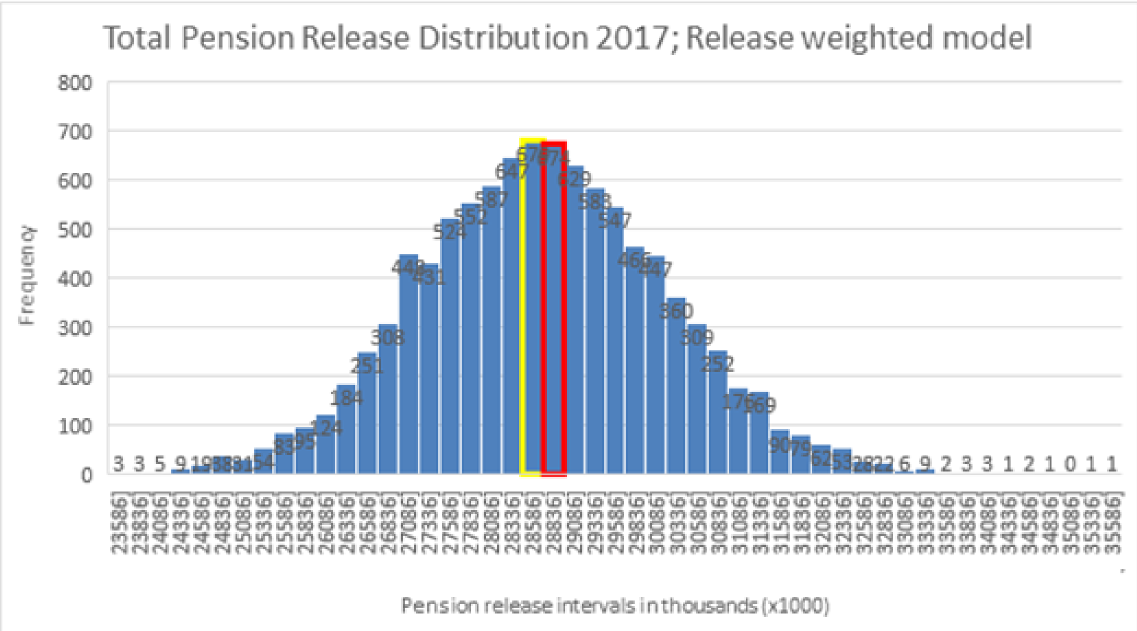

Figure 18: Simulated Pension release distribution for $W^{release}$

When performing the Jarque-Bera test we found a positive skewness and a kurtosis larger than three. This means the distribution in figure 18 should be leptokurtic with a positive tail. A distribution with a positive tail has a mean which is larger than the median. The mean of the distribution in figure 18 is 28,586,493 and the median is 28,563,702. The red block consists the mean and the yellow block the median. The frequency of those blocks is 674 and 679 respectively. So we do find a mean which is slightly larger than the median, meaning we have a positive tail. Also there are more intervals on the right on the mean. Our findings about the tail are in line with our Jarque-Bera test results and Assman (2014). Using the distributions we construct the confidence intervals. The boxplots below show these confidence intervals:

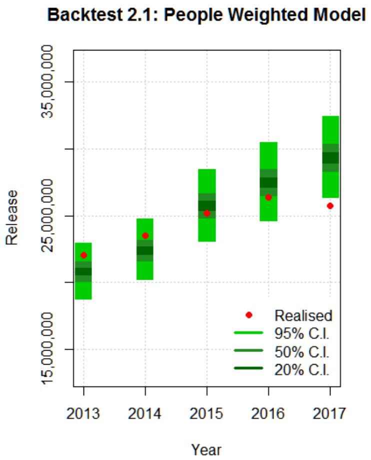

Figure 19

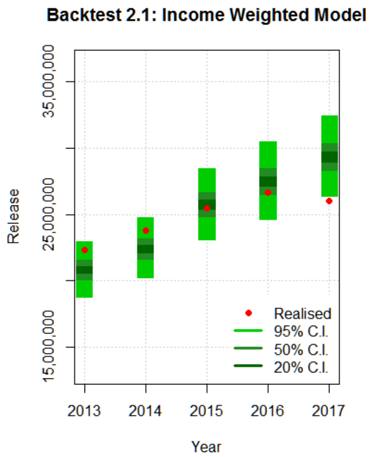

Figure 20

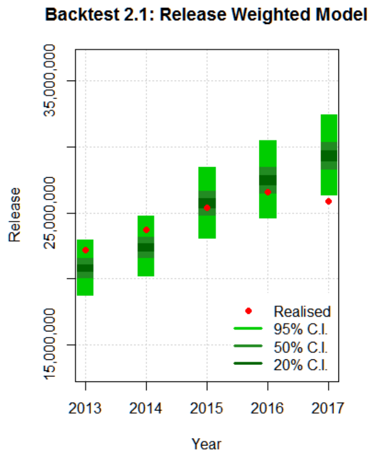

Figure 21

This test generates smaller confidence intervals compared to the normality backtest. If this backtest depicts the real confidence intervals then the normality backtest is too sensitive to type-II errors. Type-II error is the failure of rejecting a false model. The expected amount of pension releases in both tests do coincide. In contrast to the normality backtest, this test does not use the expected release to center the confidence intervals. The center of the confidence intervals is the median of the simulations after ranking the simulations. The relealised release (red dots) are the same as in the normality backtest. The test now fails in 2017 for all methods. So according to this test we have to question the validity of our model. Our results do not give reason to believe one of the weighting methods outperforms the others since they all fail in one year.

We now go to the simulation of deaths. Our W^{people} is designed to forecast the amount of deaths in the fund. The other two models are modifications on this model in order to better predict pension release. So we will only simulate deaths with this model. The boxplot below shows the confidence intervals constructed on 15000 simulations. We executed 15000 simulations because our algorithm for them requires only little computation time.

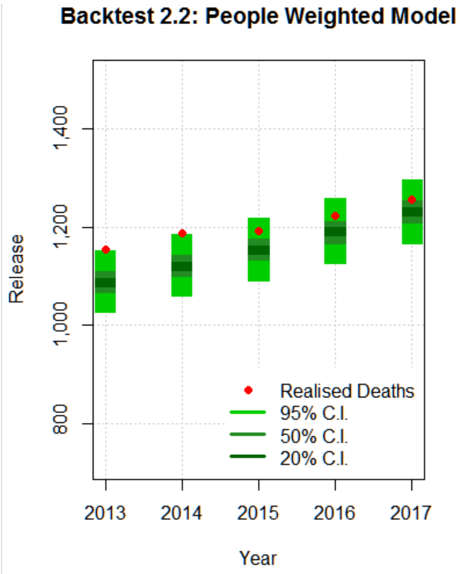

Figure 22

The realised deaths increase every year. It is notable that most deaths occurred in 2017 but we saw before that 2016 brought most pension release. Furthermore, in 2013 and 2014 the backtest fails. Our model underestimated the amount of deaths in those years. When forecasting the total pension release in 2013 and 2014 we saw all tests underestimating the total pension release as well. In 2015, 2016 and 2017 the realised amounts are within the 95% confidence interval. When forecasting total pension release in those years we saw that W^{income} and $W^{release}$ were closer to the realised release. So W^{people} seems to perform worse in years where de total deaths are in the 95% confidence interval. In order to put figure qq in perspective we scale the amount of deaths by the average pension release per death (λ). We calculate this as follows:

$$\lambda_t = \left(\frac{\sum_{i=1}^{n_t} q_{i,t} * R_{i,t})}{n_t}\right) \quad (52)$$

Where λ_t is the average pension release in year t, $R_{i,t}$ is the present value of the pension rights of participant i in year t, and n_t is the total population in year t. Since these λ_t's are constant they do not cause any extra variance. We can see this as simulating total pension release while keeping R_i' constant. This gives us the following figure:

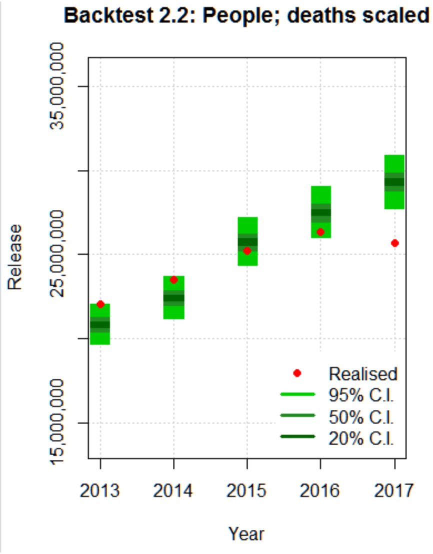

Figure 23

We also added the realised pension release as an anchor point. One can quickly notice that the size of the confidence intervals are approximately half the size of the total pension release simulation. We will present the exact size of the confidence intervals after showing the results of the simulation with constant deaths. The next figures below show 500 simulations of that test. Since that test was computationally costly to simulate we only managed to do 500 simulations. We expect that this because random sampling which is executed often in our simulation is a relatively long operation.

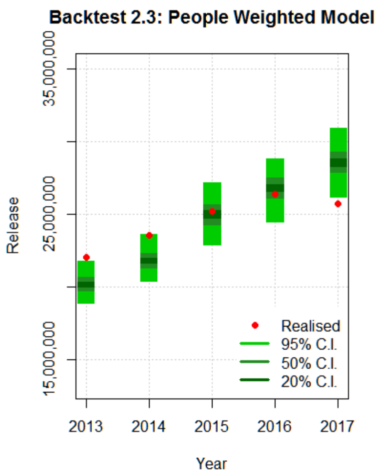

Figure 24

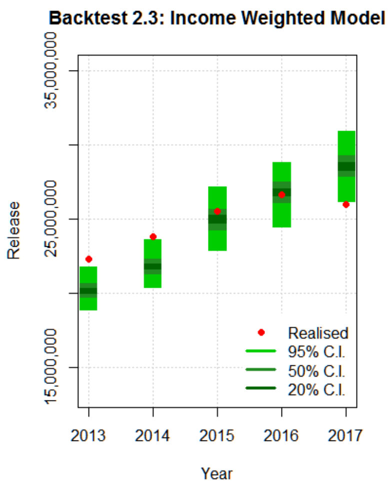

Figure 25

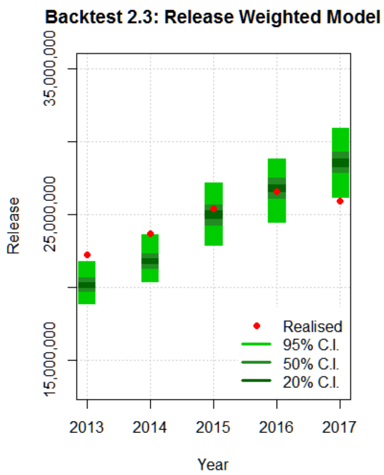

Figure 26

We again added the realised amounts as anchor points. First thing to note is that the middle of the confidence intervals do not precisely coincide with the previous test. This is due to the relative small amount of simulations. The confidence intervals are smaller compared to the simulation of pension rights, and a bit larger than the simulation of deaths (with scaled deaths). The table below gives better insight into the size of the confidence intervals for W^{people}:

Size of 95% confidence interval for the 'People Weighted Method'.				
size 95% interval ->	Pension release backtest	Deaths backtest	Pension rights backtest	Normality backtest
2013	4,166189	2,394,178	2,908,337	6,415,313
2014	4,549,557	2,499,388	3,201,680	6,857,290
2015	5,370,445	2,833,284	4,259,189	7,806,439
2016	5,815,082	3,050,395	4,347,022	8,265,460
2017	6,051,712	3,145,228	4,712,837	8,711,152

Table 4: Difference between the 97.5% and 2.5% quantile for backtests 1 and 2.

The size of the confidence intervals is increasing with the years. This can be explained by the fact that the fund is greying, leading to higher death probabilities. As explained in the first backtest, higher q_i's lead to higher variance. The variance caused from variance in deaths is smaller than the variance caused by differences in R_i's. These variances are not independent, because they do not sum up to the total variance of the pension release, the sum is larger. The size of the confidence intervals from the normality backtest are consistently around 150% of the size of the pension release simulation.

6.3 Poisson-Binomial backtest

The results of the Poisson-Binomial pension release backtest are shown in the boxplots below. We only constructed the 95% confidence interval in order to reduce the computation time for these results.

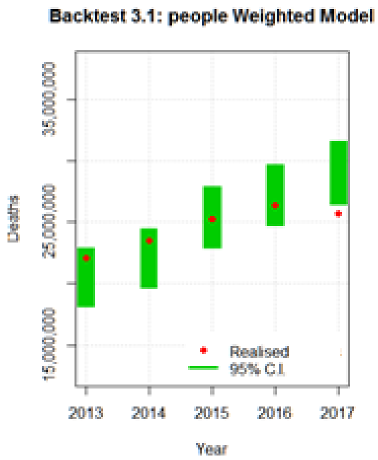

Figure 27

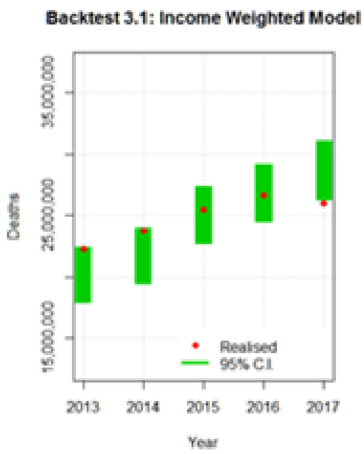

Figure 28

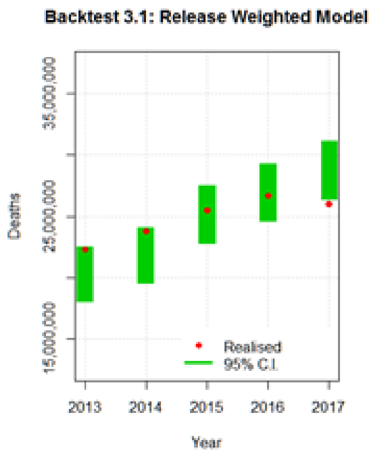

Figure 29

The size of the confidence intervals coincides with those constructed in the simulation of pension rights. The confidence intervals are a bit smaller compared to this test. But this test also rejects all models in 2017. The similarity of the confidence intervals of test simulation of pension rights and the Poisson-Binomial pension release makes us believe these are the correct confidence intervals.

We now go to the second part of this backtest, the Poisson-Binomial deaths. The confidence intervals around the expected amount of deaths based on the poison-binomial CDF are shown in the figure below:

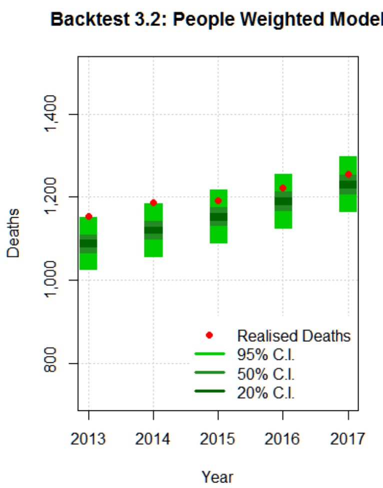

Figure 30

These results coincide with the results from the simulation of deaths. In order to compare the results we show the values of the quantiles produced by both methods in the two tables below.

Simulation backtest (deaths)						
Quantiles ->	97.5%	75.0%	60.0%	40.0%	25.0%	2.5%
2013	1152	1109	1096	1079	1066	1027
2014	1184	1142	1129	1112	1099	1059
2015	1218	1176	1162	1145	1132	1091
2016	1257	1212	1198	1182	1167	1125
2017	1297	1253	1239	1222	1208	1165

Table 5: quantiles backtest deaths simulations

Poisson-Binomial backtest (deaths)						
Quantiles ->	97.5%	75.0%	60.0%	40.0%	25.0%	2.5%
2013	1151	1110	1096	1080	1067	1026
2014	1185	1143	1129	1113	1099	1058
2015	1218	1176	1162	1145	1132	1090
2016	1256	1212	1198	1182	1168	1125
2017	1297	1253	1239	1222	1208	1165

Table 6 quantiles backtest deaths Poisson-Binomial

So most quantiles are exactly the same, the others differ no more than one death. This strengthens our belief that the simulation and Poisson-Binomial backtest depict the real confidence intervals. This finishes our result section, we now go to the robustness check.

7. Robustness Check

In this chapter we will assess the robustness of our results when altering the interest rate term structure. Instead of using the interest rate term structure of the DNB we will use an arbitrarily chosen constant interest rate of 1%. Our interest rate term structure equals approximately 1% after $t = 15$, so annuity payments which are to be paid within 15 years are discounted more heavily. On the other hand are annuity payments paid after 15 years discounted less. We consider the simulation and Poisson-Binomial backtest as our most reliable methods. These backtest give similar results as well. So both give similar output but the simulation has the shortest computational time of the two we will only perform this backtest. Figure 32, 33 and 34 below summarizes our results after 5000 simulations:

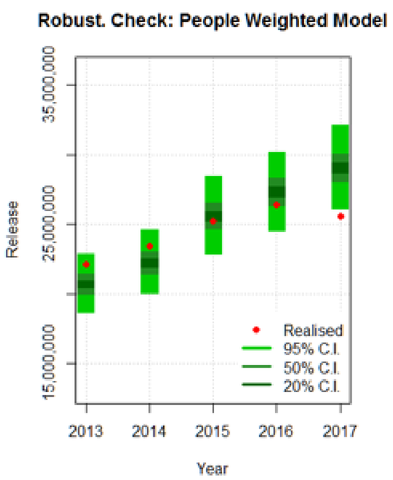

Figure 31

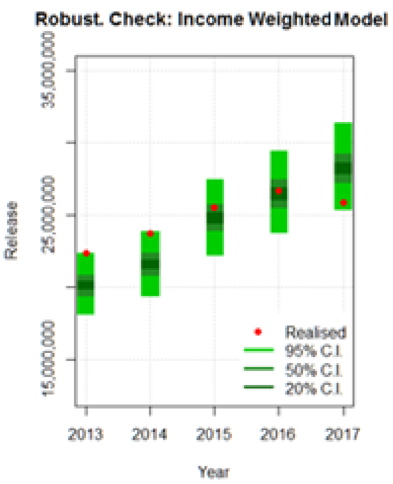

Figure 32

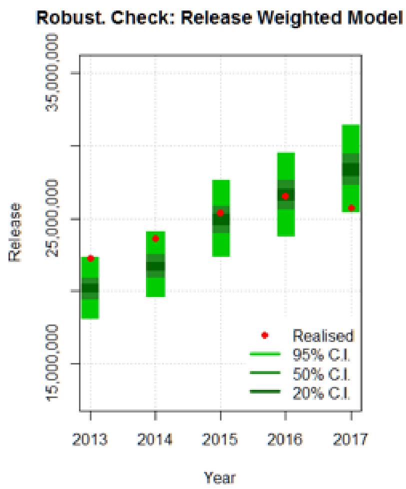

Figure 33

The outcome for W^{people} remains the same. The realised pension release is inside the 95% confidence interval in every year except for 2017. The realised and expected pension releases for this model changed only a little bit. Our W^{income} and $W^{release}$ differ a bit from our earlier results. The W^{income} now (be it barely) fails in 2013 and it succeeds in the other years. At last we find that $W^{release}$ succeeds in every year. Whether 2013 fails or succeeds is hard to read from the figure. To see the exact levels of the realised pension release and the confidence interval we refer to appendix 1.

8. Conclusion

This thesis investigated the effect of different weighting methods on estimating total pension release. The relationship between the weighting methods and the total pension release is as follows. The weightings affect the death probabilities generated by our model, these death probabilities are then used to estimate the total pension release. We use three weighting methods based on People, Income and Release. Our fund data indicates we are dealing with a grey fund with most participants inside the lowest two income classes. Our model is based on the characteristic approach with income classes as characteristics. The people weighted model resulted in the highest experience mortality. This is due to the fact that this method allocates more weight to higher income classes, and higher income classes have lower death probabilities. We backtested our different weighting methods using a normality assumption, simulation and the Poisson-Binomial CDF. We found that simulation and the Poisson-Binomial CDF are suitable for testing the models. We considered the backtest based on the normality assumption not useable. We based this on the skewness and kurtosis of the pension release distribution which is too large for this assumption. The results of the other two backtests show that the three weighting methods give similar output. The similarity of our outcomes can be traced back to the income class composition of our fund data. If the distribution among income classes was representative for the working Dutch people, each income class should contain 1/6 of the population. But we find 4/6 in our lowest two income classes and our highest class is virtually empty. We showed that higher income classes accumulate more weight in W^{income} and $W^{release}$, but due to the low initial weight they still only have little impact on the funds experience mortality. This explains why the different weighting methods give similar output. Our backtests showed no significant difference between our methods. Either all methods were inside or outside the 95% confidence interval, for a given year. So for our fund data we find no significant difference between the methods. Our results not robust for different interest rate term structures. The simulation backtest gave similar results after changing the interest rate term structure to a constant 1%. The same investigation on a fund with a different composition of participants among income classes might result in significantly different outcomes. Also the CBS data used in this investigation did not provide year specific information. So we could not calibrate time variant factors in our model. Further research should use fund-data in which participants are divided more equally among income classes.

References

Assman, C. (2014). Backtesting the fund specific expected experience mortality. *Maastricht University Master Thesis.*

Brouhns, N., Denuit, M., & Vermunt, J. (2002), 'A Poisson log-bilinear regression approach to the construction of projected lifetables', *Insurance: Mathematics and Economics, 31*(3), 373-393.

Chetty, R., Stepner, M., Abraham, S., Lin, S., Scuderi, B., Turner, N., ... & Cutler, D. (2016). The association between income and life expectancy in the United States, 2001-2014. *Jama, 315*(16), 1750-1766.

Hong, Y. (2013). On computing the distribution function for the Poisson binomial distribution. *Computational Statistics & Data Analysis, 59,* 41-51.

Kulhánová, I., Hoffmann, R., Eikemo, T.A., Menvielle, G., & Mackenbach, J.P. (2014). Educational inequalities in mortality by cause of death: first national data for the Netherlands. *International Journal of Public Health, 59*(5), 687-696.

Li, N. & Lee, R. (2005). Coherent mortality forecasts for a group of populations: An extension of the lee-carter method. *Demography, 42*(3), 575-594.

Pitacco, E., Denuit, M., Haberman, S., & Olivieri, A. (2009). *Modelling longevity dynamics for pensions and annuity business.* Oxford University Press.

Plat, R. (2009). Stochastic portfolio specific mortality and the quantification of mortality basis risk. *Insurance: Mathematics and Economics, 45*(1), 123-132.

Van Berkum, F. (2018). Models for population-wide and portfolio-specific mortality. *University of Amsterdam.*

Villegas, A.M. & Haberman, S. (2014). On the modeling and forecasting of socioeconomic mortality differentials: An application to deprivation and mortality in England. *North American Actuarial Journal, 18*(1), 168-193.

Von Gaudecker, H.M., & Scholz, R.D. (2007). Differential mortality by lifetime earnings in Germany. *Demographic Research, 17*, 83-108.

Appendix 1

The tables below show the exact numbers from the backtest displayed in figures 30, 31 and 32. The columns 4 through 9 of table 8, 9 and 10 are quantiles.

Robust. Check: People weighted model								
Year	Expected	Realised	97.5%	75.0%	60.0%	40.0%	25.0%	2.5%
2013	20708854	22138349	22854367	21434464	20980943	20446795	19993509	18688929
2014	22240730	23458561	24605256	23051647	22537625	21933692	21464621	20058452
2015	25613798	25218195	28472668	26568031	25944020	25248527	24680981	22909258
2016	27327453	26389005	30201696	28300377	27682697	26937138	26337594	24532339
2017	29092039	25583015	32090908	30065184	29429121	28658742	28046307	26132881

Table 7: Numerical output robust. test figure 30.

Robust. Check: Income weighted model								
Year	Expected	Realised	97.5%	75.0%	60.0%	40.0%	25.0%	2.5%
2013	20112874	22387016	22352580	20847699	20404344	19873369	19424533	18099960
2014	21601854	23739941	23868431	22351469	21890845	21293710	20836229	19398105
2015	24842898	25512866	27504895	25735506	25162622	24464914	23935663	22238581
2016	26514250	26684683	29405439	27467610	26847115	26090546	25504806	23758659
2017	28244542	25866288	31398697	29250596	28635862	27862028	27243581	25376127

Table 8: Numerical output robust. test figure 31.

Robust. Check: Release weighted model								
Year	Expected	Realised	97.5%	75.0%	60.0%	40.0%	25.0%	2.5%
2013	20218355	22292654	22314582	20916464	20475074	19935471	19500141	18111305
2014	21714162	23631207	24084199	22518594	22016452	21429051	20956390	19592327
2015	24984022	25402171	27639292	25866091	25292428	24607205	24040870	22369914
2016	26662643	26571858	29551044	27661384	27044603	26271426	25671690	23858285
2017	28398577	25755785	31427928	29445536	28786322	27975378	27340749	25435160

Table 9: Numerical output robust. test figure 32.