# PLSA Longevity Model

## Supporting technical appendices

March 2018

# Introduction

Welcome to the technical appendices supporting the PLSA's longevity trends model report.

This document describes the key data, assumptions and analysis behind Club Vita and the PLSA's collaborative research into longevity trends. Note that this document is intended to be read alongside the technical appendices produced for the previous analysis[1] (published in 2014). As such, in a number of places we refer readers back to this previous document, rather than replicating the details here (although relevant updates are clearly identified).

This document, like the 2014 version, is deliberately technical in nature, as it is designed to provide confidence in the rigour of the research and the necessary supporting documentation to enable actuarial users to be comfortable in referring to this work as part of forming their advice.

We start (Section 1) by providing an overview of the data we have used in our updated research – including its origin, how we verified it, the types of data available to us, and, crucially, the data volumes used in our analyses.

Key rating factors are identified, both a postcode based measure and an affluence measure. In doing so we need to create a measure of deprivation that is comparable across all of our data (see Section 1.4.1).

In order to maximise the insights we can gain from the data we have adopted a practical approach to handling missing data (Section 1.5).

We then describe how we grouped the data by socio-economic group, using deprivation and (for men) affluence to create what we call 'VitaSegments' (Section 2). This process is identical to that followed in our 2014 analysis – we have carried out checks to ensure that our groupings remain valid.

In order to bring more 'colour' to our VitaSegments, we have, in collaboration with ELSA, looked at the wider characteristics of 'typical' members of each segment (Section 3).

We discuss the technical details of our calculations of both historical life expectancy (Section 4) and improvements in historical mortality (Section 5).

The historical data is then embedded into the widely used model for projecting mortality improvements (Section 6). This provides a starting point for trustees and sponsors seeking to reflect DB pension scheme data in their longevity improvement assumptions.

We conclude by looking at the four example schemes used to illustrate the financial impact of our results (Section 7), and the construction of our eight different scenarios for longevity improvements (Section 8).

On behalf of all the team we thank you for your interest in this research and we would be delighted to respond to any questions you may have.


Steven Baxter
steven.baxter@clubvita.co.uk

Conor O'Reilly
conor.oreilly@clubvita.co.uk

Steve Hood
steven.hood@clubvita.co.uk


[1] See https://www.clubvita.co.uk/collaborative-research/supporting-technical-appendices

## Reliances and Limitations

The Pensions and Lifetime Savings Association ("PLSA") and Club Vita LLP ("CV LLP") have provided, to the UK pensions industry as a whole, both: an understanding of how differently longevity has been improving for different groups of defined benefit ("DB") pensioners (such as those at different ends of the deprivation spectrum); and materials that pension schemes, and their advisors, can use in practice to better inform the assumptions that are adopted for longevity trends (together, the "Research").

The Research is based upon the PLSA and CV LLP's actuarial understanding of legislation and events as at May 2017 and therefore may be subject to change. The Research is the PLSA and CV LLP's understanding of how longevity has been improving for different groups of DB pensioners and is not, nor is it intended to be, specific to the circumstances of any particular pension scheme.

The information contained herein is therefore not to be construed as advice and should not be considered a substitute for specific advice in relation to individual circumstances. Where the subject of the Research refers to legal matters please note that neither the PLSA nor CV LLP are qualified to give legal advice therefore we recommend that you seek legal advice. Neither the PLSA nor CV LLP (nor their respective licensors) accept liability for errors or omissions in the Research and neither the PLSA nor CV LLP (nor their respective licensors) owe nor shall accept any duty, liability or responsibility in regards the use of the Research except where we have agreed to do so in writing.



If you are seeking to use the information contained in the Research some time after the date it was produced then please be aware that the information may be out of date and therefore inaccurate. Please consult the PLSA and Club Vita websites for updates or contact enquiries@clubvita.co.uk

We recommend that you speak with your appointed longevity consultant and/or other professional advisers should you have any queries in relation to applying the Research findings within your scheme. Alternatively please contact Joe Dabrowski, Head of Governance & Investment of the PLSA at joe.dabrowski@plsa.co.uk or Steven Baxter of Club Vita LLP at steven.baxter@clubvita.co.uk, who will be pleased to discuss any of the issues highlighted by this research in greater detail.

# Contents

# 1 Data underpinning our analysis

## 1.1 Club Vita dataset

The Club Vita database (VitaBank) is a pool of data of individual pension scheme member records, submitted by over 200 participating occupational DB pension schemes. This database (as at February 2017) consists of nearly 6 million member records; including:

- Over 2.5 million pensioners and widow(er)s; and
- 1 million deaths.

The records collected include personal, but non-sensitive, information recorded by pension scheme administrators. This includes information relevant to predicting longevity, such as date of birth, sex, postcode, pension, final salary and retirement health.

## 1.2 Data pre-processing

Only data which has been through our initial quality control process enters the statistical analysis. The data quality control process is designed to ensure the data for each pension scheme is as reliable as possible. However it also recognises that the quality of the data is often dependent on historical record keeping processes and so may have some inherent shortcomings.

A suite of checks are carried out on the data received to ensure it is correct and reliable, and, where necessary, corrections are made if possible. Where a member record has a predictor which our checks suggest is unreliable it is excluded from analysis for that predictor. We also check for concentrations of unreliable records within schemes and biases in exclusions between living and deceased records, and limit a scheme's inclusion in our analysis where there is a risk of bias.

### 1.2.1 Ensuring a complete history of deaths

We recognise that some schemes may not have a complete record of deceased pensioners prior to some point in time. For each scheme we have determined an "earliest useable date" (EUD) – the date from which we believe we have a complete history of deaths.

The mortality data we receive includes experience data up to a date shortly before it was extracted from the pension scheme's administration system. As such it is liable to 'incurred but not reported deaths' i.e. an understatement of deaths in the most recent weeks of the extract as a result in the delay in reporting deaths.

In order to ensure that mortality rates are not underestimated we carry out similar analysis to that described above to verify the point up to which we believe we have full and complete death data. This leads to a "latest useable date" (LUD) for each scheme, which is used to right censor the data (i.e. no observations of survival beyond this date are included in our analysis). Typically the latest useable date excludes between 1 and 2 months' worth of data.

Since we are analysing mortality by calendar years, we need to take care to avoid seasonal biases resulting from including part years, therefore we have for these purposes restricted our analysis for each scheme to the period from the first 1 January on or after the EUD to the last 31 December on or before the LUD for each scheme.

When analysing the patterns in longevity by specific factors, for example pension amount, we also check whether we have complete information on that factor from the EUD onwards. Where this is not the case we use a factor-specific EUD for that scheme.

## 1.3 Data extract used in this analysis

### 1.3.1 Exposed to risk & deaths

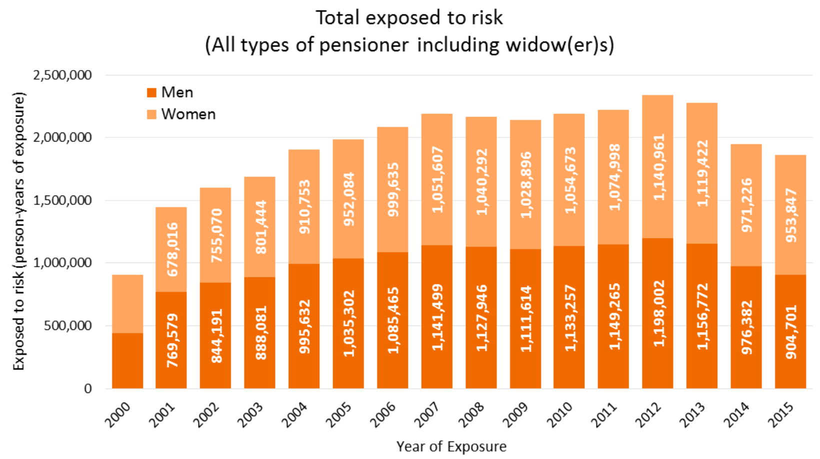
Club Vita collects data annually from each of its subscribers, with these data feeds spread over the calendar year. As such it is regularly refreshed with the latest longevity data. For the purposes of our analysis we have focussed on an extract of the database as at February 2017 throughout.

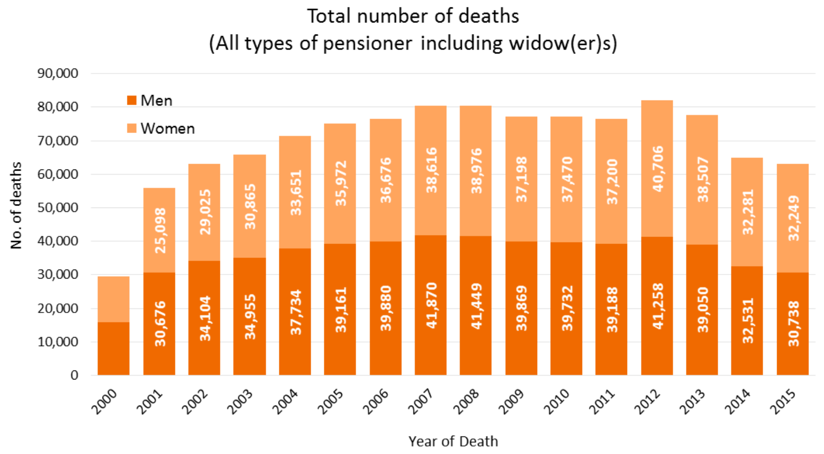
The charts (right) show the pattern of (pensioner and dependant) ‘exposed to risk’[2] and deaths over time for men (dark orange bars) and women (light orange bars) within the data analysed in this report.

We can see how:

- The exposures increase over time reflecting
  - schemes within the Club having reliable data starting at different points in time due to historical administration practices; and
  - the maturation of pension schemes leading to larger numbers of pensioners
- There is a step-up in 2001 – the point at which a number of the larger schemes first have reliable data.
- The deaths follow a similar pattern to the exposed to risk.

We have seen more than a 10% increase in overall data volume since our first Longevity Model report was published in 2014.




However, as a result of the quality checks set out in Section 1.2, not all of the data shown here was used in the analysis presented in this paper. In practice we use around 65% of the data shown here in the analysis.

[2] Broadly speaking a measure of the number of lives in each year, but adjusted to allow for the fact that some individuals were only in the analysis for part of that year. As such, exposed to risk is typically slightly lower than a lives count.

### 1.4 Key rating factors

By collecting information at the individual level, VitaBank contains a wide range of rating factors potentially relevant to both baseline mortality and improvements coming through over time. These rating factors include gender, retirement health, pensioner type (pensioner or dependant), postcode based socio-economic measures (such as Index of Multiple Deprivation), affluence (pension and salary), age and occupation (manual and non-manual)[3].

The 2014 study looked at a number of potential rating factors. We identified two factors, the Index of Multiple Deprivation ("IMD") and pension amount, as the most appropriate to use, given the constraints of requiring rating factors which identified different patterns over time **and** were readily available to all pension schemes. See the previous technical appendices[4] for further discussion of these and alternative rating factors.

For the purposes of this analysis we have retained these two rating factors, as discussed below.

#### 1.4.1 Postcode based measure: Index of Multiple Deprivation

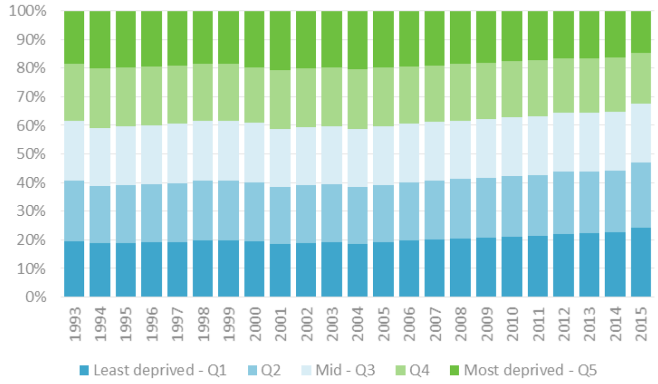
The statistics agencies of each of the nations within the UK measure the deprivation of local areas via an index which captures multiple indicators, typically including such factors as income, employment and crime.

The scores are publicly available at a local level. For example, within England they are available for regions known as 'Lower-layer Super Output Areas' (LSOAs), which cover between 400 and 1,200 households each. However, they are not directly comparable across countries within the UK, with the weighting to the different factors varying from country to country (and indeed in some countries factors are included which are not included in other countries). Further, many of the factors are measured relative to the country-specific average value.

Therefore we needed to generate an index which spans all of the UK. Our method for doing this is detailed in our 2014 technical appendices[5], but, in summary, involves:

- choosing a small number of factors which are used, and significantly weighted, in each country, based on published underlying data (specifically Income and Employment factors);
- carrying out linear regression against the chosen factors; and
- rebasing the values against a base country (England).

The chart below shows the split of our data (men and women combined) between the resultant IMD quintiles. We can see that the split has been relatively stable over time.




[3] See Madrigal et al (2012) for more detail on how Club Vita have determined the key ratings factors for mortality levels.

[4] See https://www.clubvita.co.uk/collaborative-research/supporting-technical-appendices

[5] See https://www.clubvita.co.uk/collaborative-research/supporting-technical-appendices

### 1.4.2 Affluence measure: Pension amount

The Club Vita data contains two measures of affluence: pension and last known salary.

Pension size can be a poor proxy for overall affluence as it depends not only on earnings but length of service in the pension scheme – a modest pension could arise from long service on low pay, or very short service on high pay. However, whilst salary is a better measure of affluence, pension will almost always be available, whereas salary (whilst generally available) may be harder to extract from some pension scheme records. We therefore, as before, use pension amounts in our analysis.

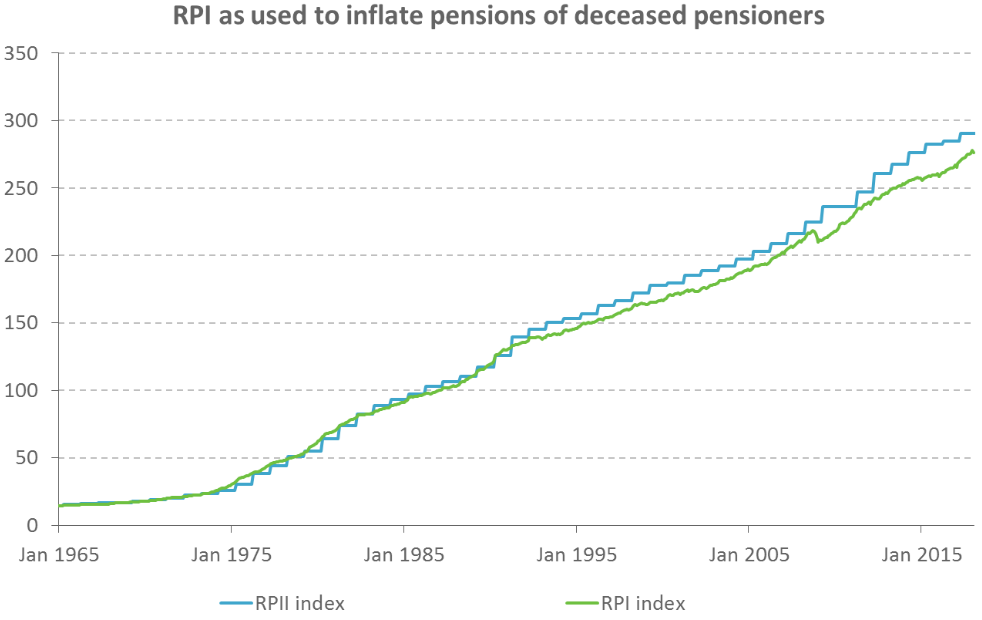
To allow for inflation, pension amounts are revalued from their 'as at' date to a common date (1 January 2010). For living pensioners this is simply done in line with the change in RPI. For deceased pensioners the revaluation of pension amounts is performed to enable broad consistency with the pension increases paid historically to surviving pensioners, which will typically be a mix of full RPI, limited price inflation and nil increases, as follows:

- Identify the 'as at' date for current living pensioners in the scheme (using the most common date where multiple dates exist).
- Determine the 'adjusted' RPI value ('RPII') at both the 'as at' date for live pensioners, and the date of death for each deceased pensioner. The RPII index is increased each April based on the RPI over the year to the previous September (subject to a minimum increase of 0%).
- Determine the proportion of full inflationary pension increases likely to have been awarded ('PpnRPI'), a value that depends on year of retirement and differs for public and private sector schemes, and for men and women.
- Revalue pension from date of death to 'as at' date as follows:

$$Pension\ at\ date\ of\ death * \left(\frac{RPII_{current}}{RPII_{death}}\right)^{PpnRPI}$$

- Revalue pension from 'as at' date to 1 January 2010 using 'full' RPI.

The chart below shows the difference between actual RPI index (green) and our adjusted 'RPII' index value (blue).




We can see how the RPII index has typically been above the RPI index (with the main exception of period of high inflation in the late 1970s).

## 1.5 Making maximum use of available data

In Section 1.2 we discussed how the scheme data used in our analysis has undergone a thorough data quality control process to determine what data will be used in the onward analyses and ensure reliability of data. This is done both at the scheme level and at the covariate level (so, for example, a particular scheme may have reliable postcode data but suspect pension amounts in a particular year).

Levels of unknown covariates can be expected to increase as we go further back in time (due to having less stringent administration standards historically, records not being updated, etc.). In particular these issues are more likely to affect deaths (i.e. higher levels of unknowns), so there is the possibility that we could be biasing the results by excluding more deaths relative to living pensioners in a given calendar year.

At a scheme level, the proportions of 'unknowns' is again likely to increase as we go back in time, until, in some cases, reaching the exclusion 'trigger' level – the point in time before which no exposures are included (the EUD discussed in Section 1.2).

There is, therefore, a growing risk of understating rates of mortality historically (if we exclude more deaths than lives, we are reducing the mortality rate). This will have a knock on effect on mortality improvements, which will again be lower than their 'true' level, due to historical mortality rates being lower.

We have sought to overcome this issue, and maximise the available data, without compromising on overall data quality, by reallocating 'unknown' data, using the same process as in our 2014 analysis. We discuss this briefly below – for more details see our original technical appendices[6].

### 1.5.1 Adjusting for missing data

We have sought to maximise the amount of data used by re-allocating lives and deaths with 'unknown' covariates across the covariate groups, as follows.

We initially take the (cleaned) submitted data, and allocate individual members (lives and deaths) to the appropriate IMD quintiles and (for men) pension bands (including 'unknowns' for each covariate as appropriate).

Where one of the covariates (e.g. IMD quintile) is unknown, then the exposures and deaths for the group are assumed to be spread across the unknown covariate in the same proportions to where the covariate is known. This minimises the risk that the mortality rates, as measured over time, are polluted by any imbalances in data coverage between lives and deaths.

The proportions to use for spreading data across the unknown covariate are volatile from one age to the next. To smooth this out, we average across the 5 year age bracket centred on each age when determining the ratios.

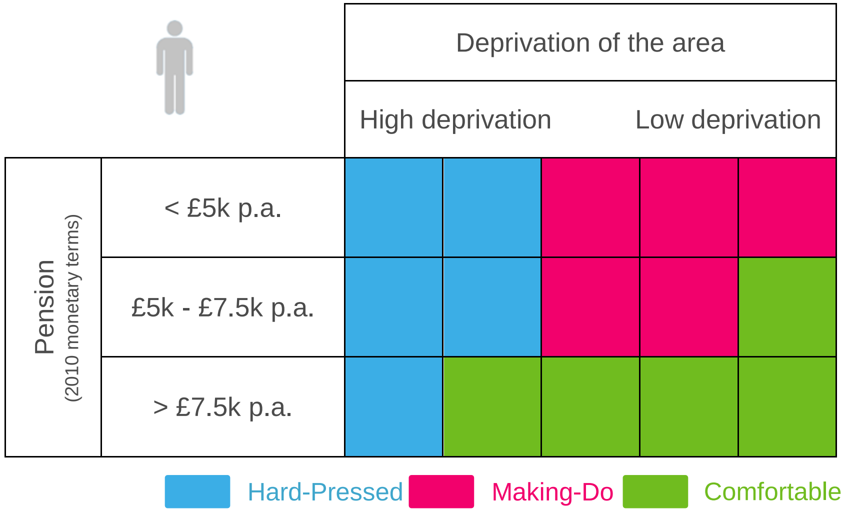
The same approach of reallocating unknowns is applied to men and women. However as we only use one covariate – IMD – for women, the calculations are less complex than for men, although the levels of unknowns are similar.

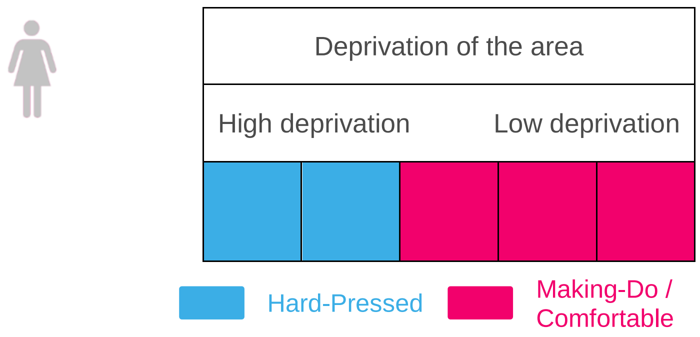
The impact on our results of this reallocation approach is relatively small (increasing the total exposure allocated to our groups by 4-5%). However we can be confident that we have removed a possible area of bias in our analysis of historical improvements.

[6] See https://www.clubvita.co.uk/collaborative-research/supporting-technical-appendices

# 2 Generating our socio-economic groups

For our 2014 analysis we generated 3 (2 for women) distinct socio-economic groups (VitaSegments), based on a combination of IMD quintile (adjusted as discussed in Section 1.4.1 to be UK equivalent) and (for men) pension band.




| Group | Characterisation |
|---|---|
| Hard-Pressed | Living in more deprived areas and generally with lower levels of retirement income. |
| Making-Do | Modest retirement income levels and living in areas of average to low levels of deprivation. |
| Comfortable | Higher levels of retirement income (over £7,500 p.a. unless living in the least deprived 20% parts of the UK when this can be reduced to £5,000 p.a.). This group naturally includes some pensioners with retirement incomes much higher than £7,500 p.a. |

For this updated analysis we have retained the same socio-economic groups. In particular we have retained the same pension band thresholds (in 2010 terms) and used the same underlying IMD mappings as in the 2014 research.

We have however carried out checks to ensure that the approach remains appropriate in light of the additional data now available, applying the same range of statistical tests as applied to the 2014 data. In particular we used clustering techniques to help identify possible groupings (both partitioning about medoids (PAM) and Fuzzy Analysis).

For more for more details on the methods used to create these socio-economic groups, see our original technical appendices[7].

[7] See https://www.clubvita.co.uk/collaborative-research/supporting-technical-appendices

# 3 Understanding the VitaSegments

## 3.1 Background

The Club Vita data used to construct our VitaSegments includes a range of factors which may help predict life expectancy (such as affluence and socio-demographic details). However it does not hold any information on an individual's lifestyle habits or personal circumstances that could help us to build up a picture of the 'typical' characteristics of each VitaSegment.

In order to provide more 'colour' on the VitaSegments we have looked at the information held by the English Longitudinal Study of Ageing[8] (ELSA).

ELSA began in 2002. It is a large scale study of people, initially aged 50 or over on 1 March 2002, and their partners, living in private households in England. The same group of respondents have been interviewed at two-yearly intervals, known as 'waves', with some of these waves also incorporating additional lives. The interviews ask a wide variety of questions which enable the study to measure changes in their health, economic and social circumstances, covering such areas as:

- Household and individual demographics
- Health – physical and psychosocial
- Social care
- Work and pensions
- Income and assets
- Housing
- Cognitive function
- Social participation
- Walking speed

## 3.2 Incorporating ELSA data in our analysis

By combining the ELSA data with our VitaSegments we can:

- deepen our understanding of why the groups have historically had different life expectancy expectations; and
- form a view as to whether this is likely to continue in future, or whether any groups are likely to increase or decrease.

We have based our analysis on the anonymised data for the Wave 7 respondents (the most recent wave). This contains data that was collected over the period 1 June 2014 to 31 May 2015 from a total of 9,670 individuals.

In addition, NatCen (who manage the ELSA dataset) have kindly supplied additional information on the deprivation quintile of the area in which each individual lives. This additional information, combined with the information in the ELSA dataset on an individual's pension income, has enabled us to map the individuals in the ELSA data to our VitaSegments.

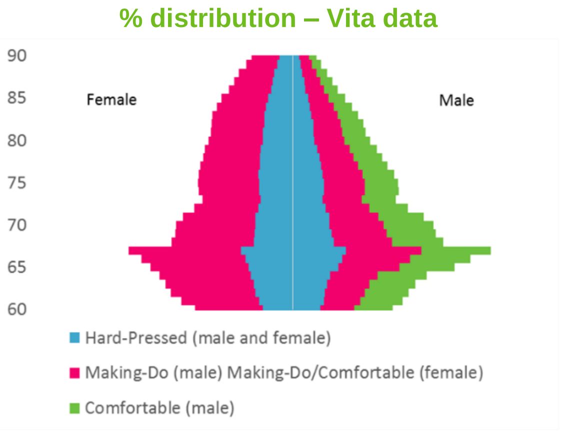
The ELSA data includes a representative sample of individuals aged 50 and over. In order to make direct comparisons to the data underpinning our analysis we have restricted our attention to the 3,694 individuals who met the criteria that:

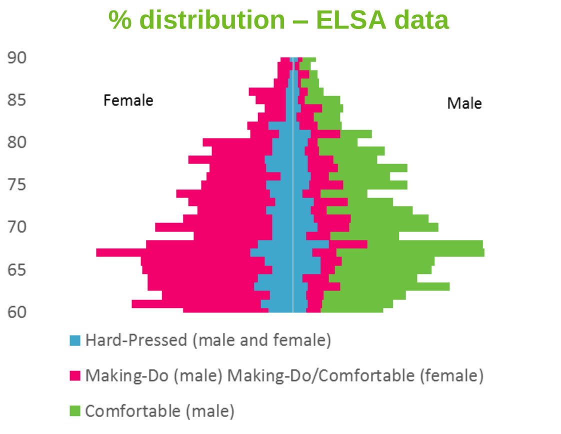
- they were in receipt of a pension; and
- a proportion of their pension related to DB (so not exclusively DC).

[8] See https://www.elsa-project.ac.uk/

## 3.3 Comparing datasets

We have carried out a number of checks to ensure that the resulting individuals are likely to be representative of people in the VitaSegments, including verifying that the datasets have a similar age distribution (both in aggregate and for each VitaSegment). The charts below show the comparative spread of lives across the VitaSegments, for men and women.




It is notable that a higher proportion of men in the ELSA data (c63%) are allocated to the Comfortable group than in the Club Vita data (c34%).

Within the ELSA data the pension amount recorded is the total across all (non-state) pensions, so may include multiple DB and/or some additional DC pensions. However, when filtering the data for men with only one DB pension we continue to see a bias towards the Comfortable group.

Looking at the spread of ELSA data over IMD quintiles, we see that there is relatively high coverage in the lowest deprivation quintiles (around 30% in both quintile 1 and 2), while the coverage in the most deprived quintiles is lower (14% in quintile 4 and just 7% in quintile 5, the most deprived).

This suggests any bias may simply be a feature of the ELSA sample.

## 3.4 Variations between VitaSegments

Our analysis found clear differences in health, lifestyle and care characteristics between the distinct groups – with the Comfortable group of men consistently scoring higher (in factors which were likely to lead to longer life) than the Hard-Pressed group, while the Making-Do group were in between (with the Making-Do/Comfortable women scoring higher than the Hard-Pressed women).

These differences will impact both current longevity and the prospects for future improvements. This helps explain the higher current life expectancy for Comfortable men seen in our analysis.

The data were made available through the UK Data Archive. ELSA was developed by a team of researchers based at the NatCen Social Research, University College London and the Institute for Fiscal Studies. The data were collected by NatCen Social Research. The funding is provided by the National Institute of Aging in the United States, and a consortium of UK government departments co-ordinated by the Office for National Statistics. The developers and funders of ELSA and the Archive do not bear any responsibility for the analyses or interpretations presented here.

# 4 Calculating life expectancy

## 4.1 Smoothing and extending mortality rates

In order to explore variation of experience by socio-economic group we need a method for calculating life expectancies over time. In particular we need a method that:

- applies some smoothing to the underlying data (which can be volatile when looking at individual ages); and
- allows extensions of mortality rates to older ages, above the upper limit of the data set.

The method we use is to assume the Gompertz law of mortality (that $\log \mu_x$ is linear with age $x$) applies at all ages.

In our analysis we adopt an 'individual year' basis (so based on the experience in a given year, without smoothing) when considering the life expectancy in each year. This 'individual year' approach is also used when looking at the increase in life expectancy over successive five year periods.

However we also consider the general trends in historical experience, and for this purpose we use a 3 year smoothing period, which provides some level of year on year smoothing without 'over-smoothing' and so running the risk of smoothing out emerging trends (although we are unable to calculate the equivalent value for 2015).

## 4.2 Detailed calculations

The details of the life expectancy calculations for each year are as follows (where the only differences between the one and three year smoothing is in the period used to find the crude mortality rate).

### 4.2.1 Estimate crude rates

- Calculate the crude mortality rate $q_x$ (the probability of a life aged $x$ dying before age $x+1$) as:

$$q_x = \frac{D_x}{E_x}$$

  Where $D_x$ is the number of observed deaths aged $x$, and $E_x$ is the (initial) exposure aged $x$ in the year.

  In each year we calculate crude rates for ages 60 to 95 (inclusive).

- Estimate crude $m_x$ (the central death rate for age $x$) as:

$$m_x \approx \frac{q_x}{1 - \frac{q_x}{2}}$$

- Approximate crude $\mu_{x+\frac{1}{2}}$ (the force of mortality for age $x + \frac{1}{2}$), using the assumption that deaths are uniformly distributed, as:

$$\mu_{x+\frac{1}{2}} \approx m_x$$

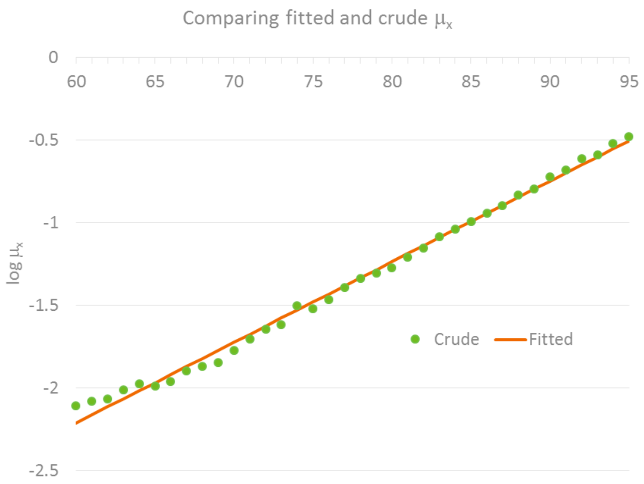
- Calculate crude $\log(\mu_{x+\frac{1}{2}})$

### 4.2.2 Find fitted rates

These calculations provide crude estimates for the (log) force of mortality by age, for each calendar year. However in order to smooth out volatility, and allow us to extend rates to older ages, we need to fit a linear regression line to the calculated crude $\log(\mu_{x+\frac{1}{2}})$ values. This will therefore have the form:

$$log(\mu_{x+\frac{1}{2}}) = \text{intercept} + \text{gradient} * (x + \tfrac{1}{2})$$

The chart below illustrates crude rates (green) with a corresponding fitted (orange) line.




Using this fitted line, we can calculate the fitted $p_x$ (the probability of surviving from age $x$ to $x+1$) as:

$$p_x = \mathrm{G}^{(C^x*(C-1))}$$

where:

$$\mathrm{B} = 10^{Intercept}$$

$$C = 10^{Gradient}$$

$$\mathrm{G} = exp\left(\frac{-B}{\ln C}\right)$$

Note that using this approach we can also calculate $p_x$ for older ages, so enabling extension beyond the oldest age of the data set (and indeed up to 125).

Having found $p_x$ it then remains to calculate $e_x$ (the life expectancy age x) recursively as:

$$e_x = p_x * (e_{x+1} + 1) + \frac{(1 - p_x)}{2}$$

By following this approach we are able to calculate life expectancy (from age 65) for each year from 2000 to 2015, for each individual VitaSegment and the aggregated DB data set.

# 5 Improvements in historical mortality

## 5.1 Calculating age standardised improvements

While we can calculate life expectancy at each age, for each VitaSegment, as set out in Section 4, it is useful when considering improvements in life expectancy over time to summarise improvements over a given period in a single figure. When doing this it is important to isolate differences in the age structure of different populations when making such comparisons, through the process of age standardisation.

The process that we have adopted is as follows:

### 5.1.1 Calculate an age standardised mortality rate

- Calculate the crude ${q^y_x}_x$ rate for given population for each age *x* and year *y*, using the observed initial exposures ($E^y_x$) and numbers of deaths ($D^y_x$).

$$q^y_x = \frac{D^y_x}{E^y_x}$$

- Determine an appropriate 'reference' population to use – in this case we used the exposure data for the England & Wales population in 2010 (separately for men and women), as provided in the illustrative CMI_2016 software published by the CMI[9].

$$E^{Ref}_x = Exposure\ age\ x\ in\ reference\ population$$

- Calculate the 'age standardised' deaths (the number of deaths that would have occurred at a given age *x* and year *y*, if the crude mortality rate had applied to the 'reference' population).

$$D^{Ref\,y}_{\ \ x} = q^y_x * E^{Ref}_x$$

- Calculate the age standardised mortality rate for year *y* by summing the age standardised deaths and exposures (in the ) for the desired age range (in this case 65 to 95)

$$q^y_{Stand} = \frac{\sum_{x=65}^{95} D^{Ref\,y}_{\ \ x}}{\sum_{x=65}^{95} E^{Ref}_x}$$

### 5.1.2 Calculate the implied average annual rate of improvement over specified calendar year periods

- Calculate the annualised improvement in age standardised mortality rate between years *y* and *z* as:

$$MI^{y,z}_{Stand} = 1 - \left(\frac{q^z_{Stand}}{q^y_{Stand}}\right)^{\frac{1}{z-y}}$$

## 5.2 Confidence intervals in age standardised improvements

In calculating the standard errors on annualised age-standardised improvements we have adopted a methodology that is intended to be broadly in line with section A3.3 of CMI Working Paper 97 ('WP97')[10].

### 5.2.1 Confidence intervals for DB pension scheme data

When considering the pension scheme data as used in our analysis there are three key differences from the approach taken in WP97:

1 As the data is lives-weighted, the calculations set out in WP97 can be slightly simplified;

[9] See https://www.actuaries.org.uk/learn-and-develop/continuous-mortality-investigation/cmi-working-papers/mortality-projections/cmi-working-papers-97-98-and-99

[10] See https://www.actuaries.org.uk/learn-and-develop/continuous-mortality-investigation/cmi-working-papers/mortality-projections/cmi-working-papers-97-98-and-99

2 The calculations use $q_x$ rather than $\mu_x$, so deaths are assumed to have a Binomial rather than Poisson distribution; and

3 As we are calculating the uncertainty in the average annual improvement over a period, rather than the uncertainty in a specific year's annual improvement rate, we use a different form of geometric annualised improvement (taking the fifth root of the ratio of standardised mortality rates at time t+5 and time t, rather than taking the product of annual improvements in each year and the fourth root)

The approach that we have adopted for pension scheme data is therefore:

- For each age *x* and year *y*, take the crude mortality rate ($\mathrm{q}_x^y$) and standardised mortality rate ($\mathrm{q}_{Stand}^y$), as set out in Section 5.1
- Calculate the variance of mortality rate, $Var(\mathrm{q}_x^y)$, at each age *x* in each year *y*.

$$Var(\mathrm{q}_x^y) = \frac{\frac{\mathrm{D}_x^y}{\mathrm{E}_x^y} * \left(1 - \frac{\mathrm{D}_x^y}{\mathrm{E}_x^y}\right)}{\mathrm{E}_x^y}$$

- Calculate the variance of the standardised mortality rate, $Var(\mathrm{q}_{Stand}^y)$, for each year *y*, where

$$\mathrm{q}_{Stand}^y = \sum_x \mathrm{q}_x^y * \frac{E_x^{Ref}}{\sum_x E_x^{Ref}} = \sum_x \mathrm{q}_x^y * \theta_x$$

Where

$$\theta_x = \frac{E_x^{Ref}}{\sum_x E_x^{Ref}}$$

Here, $Var(\mathrm{q}_{Stand}^y)$ is the weighted mean sum of $Var(\mathrm{q}_x^y)$ over individual ages, where the weights are the squares of the representative proportions in the standard population (see Section 5.1).

$$Var(\mathrm{q}_{Stand}^y) = Var\left(\sum_x \mathrm{q}_x^y * \theta_x\right) = \sum_x \left(Var(\mathrm{q}_x^y) * \theta_x{}^2\right)$$

- Calculate the variance of the improvement in age standardised mortality rate over 5 years, $Var(MI_y)$, where $MI_y$ represents the improvement in the age standardised rate in the "end year" (*y*+5) over the "start year" (*y*):

$$MI_y = \frac{\mathrm{q}_{Stand}^{y+5}}{\mathrm{q}_{Stand}^y}$$

$Var(MI_y)$ is calculated using the formula referenced in WP97, for the variance of X/Y.

$$Var(MI_y) = Var\left(\frac{\mathrm{q}_{Stand}^{y+5}}{\mathrm{q}_{Stand}^y}\right)$$
$$= \left(\frac{\mathrm{q}_{Stand}^{y+5}}{\mathrm{q}_{Stand}^y}\right)^2 * \left\{\frac{Var(\mathrm{q}_{Stand}^{y+5})}{\mathrm{q}_{Stand}^{y+5}{}^2} + \frac{Var(\mathrm{q}_{Stand}^y)}{\mathrm{q}_{Stand}^y{}^2}\right\}$$

Calculate the variance of annualised improvement, $Var(MI_y^{ann})$, using

$$Var(f(X)) = Var(X) * [f'(E[X])]^2$$

Where in this case,

$$X = MI_y$$

$$f(X) = \mathrm{MI}_{Stand}^{y,y+5} = 1 - MI_y{}^{\frac{1}{5}}$$

So

$$Var\left(MI_y^{ann}\right) = Var\left(MI_y\right) * \left(\frac{1}{5} * \left(MI_y\right)^{\frac{1}{5}-1}\right)^2$$

- Take the square root of $Var\left(MI_y^{ann}\right)$ to find the standard error of $\mathrm{MI}_{Stand}^{y,z}$

The above approach can be carried out for each of the VitaSegments, as well as for the aggregated DB pension scheme data.

### 5.2.2 Confidence intervals for England & Wales population data

We repeat the calculations above for England & Wales population data for comparison.

Note that in this case we use $m_x$ rather than $q_x$ in deriving the variance and so confidence intervals. However these confidence intervals are then applied to $q_x$ values for ease of comparison, but this is assumed to be a reasonable approximation.

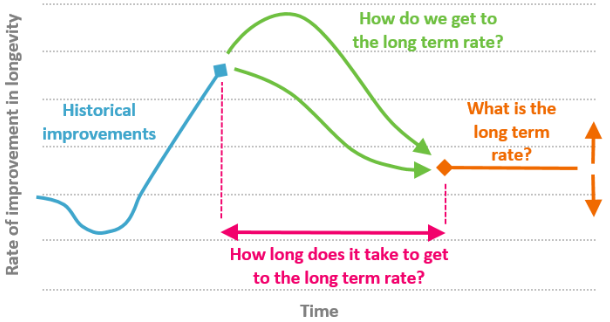
# 6 Projecting future trends using the CMI model

## 6.1 Introduction

The CMI mortality projections model (the 'CMI Model') is currently the most widely used model for mortality improvements in the actuarial industry in the UK. The model is published by the Continuous Mortality Investigation (CMI), part of the Institute and Faculty of Actuaries (IFoA).

## 6.2 The CMI model

The CMI Model is a deterministic model driven by user inputs, based on the assumption that current rates of mortality improvements converge over time to a single[11] long-term rate.




There are broadly three parts to the longevity improvement model[12]:

- Initial rates of improvement
- Long-term rate of improvement
- The "pathway" connecting the short term and long term

The model has been updated roughly annually to reflect emerging experience of the underlying England & Wales population data (with the occasional minor tweak to methodology).

However the latest (at the time of the analysis) version of the model, CMI_2016, introduced a more material revision to the structure of the model, particularly around how the model fits historical data when determining the initial rates of improvement.

In particular new parameters were added which enable users to control the level of smoothing applied in each dimension, so enabling users to, for example, make more or less allowance for recent experience when setting initial rates of improvement.

Prior versions of the model required data to be separately smoothed and disaggregated before being used for calibration, and also featured a number of decisions around the disaggregation itself which had a material impact on projections. The CMI_2016 model removed the need for such extensive 'pre-processing' of data prior to calibrating the model, becoming more of a 'one stop shop' where raw data is input to the model.

The 2014 analysis used the CMI_2013 version of the model for projecting future trends. For our updated analysis we have made use of the increased flexibility of the CMI_2016 model when calibrating future trends.

[11] The model reduces the user input long term rate to 0% p.a. at the oldest ages

[12] See https://www.clubvita.co.uk/collaborative-research/supporting-technical-appendices

### 6.3 Calibrating CMI_2016 to VitaSegments

For our analysis we have calibrated the model to pension scheme data, subdivided into the socio-economic groups set out in Section 2.

In doing so we make a number of adjustments to the ‘core’ parameters of the CMI_2016 model. We explore each of these below (where not mentioned, parameters are in line with core settings).

#### 6.3.1 Data range used

The core setup of the CMI_2016 model is calibrated to (England & Wales) population data, using ages 20 to 100 and calendar years 1976 to 2016. When fitting to pension scheme data we use an age range of 60 to 95 (inclusive) and calendar years 1993 to 2015 (inclusive).

Above age 95 improvements are assumed to taper to 0% p.a. at age 110 (the same age as core).

#### 6.3.2 Constraint on cohort component

The cohort component is constrained to be 0% at ages 60 and below (core age 30). The CMI model requires that this value must be no lower than the lower bound of the data set used to calibrate the model.

#### 6.3.3 Smoothing parameter

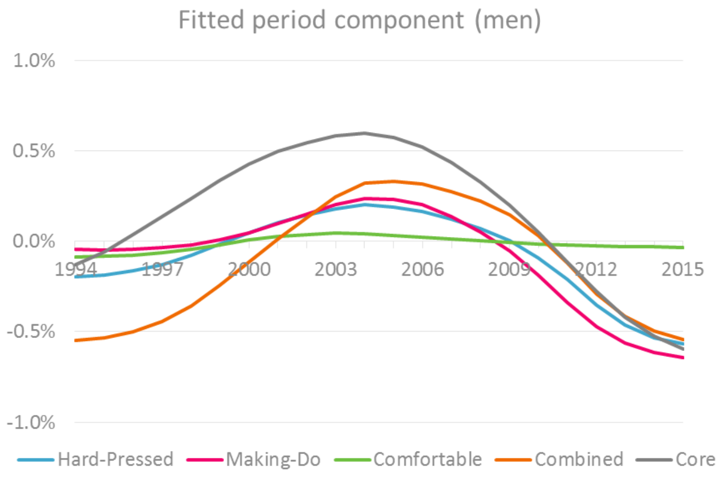
A key parameter of the CMI_2016 model is the level of smoothing that is applied in the period dimension. This reflects a general belief that period effects (the component of improvement due to the individual year, applying to all ages) are a key contributor to year on year improvements, rather than variations from year to year just due to seasonal volatility.

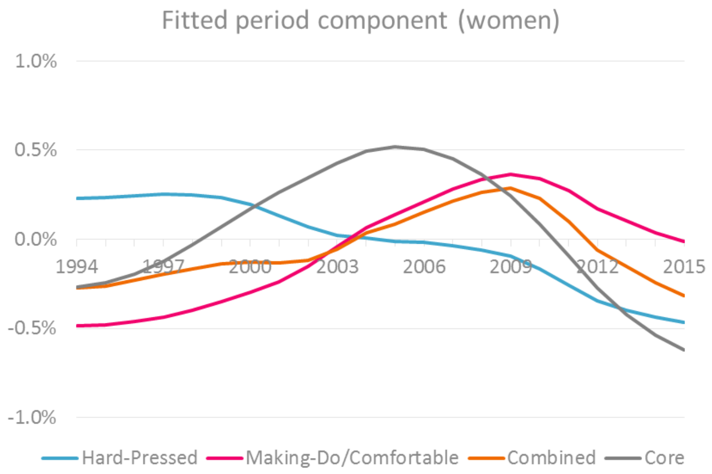
For our calibrations to pension scheme data we have chosen to adjust the period smoothing parameter (referred to in the CMI_2016 model as $S_\kappa$) to a value of 6 (from a core value of 7.5). This has the effect of applying ‘less’ smoothing in the period dimension.

The use of a lower smoothing parameter is to capture some element of a period effect when calibrating to VitaSegments. Note however that the underlying pension scheme data does not readily suggest that there are strong period effects, and so this choice is purely to reflect some element of period effect. We continue to explore this dynamic and hope to provide further details of our analysis in the future.

The resultant fitted period components generated from the CMI_2016 model are shown in the charts below. We have shown the corresponding period component from the core calibration (i.e. using England & Wales population data) for reference.

Note in particular how, even with a reduced level of smoothing, the fitting process struggles to identify a significant period component for the Comfortable male group.





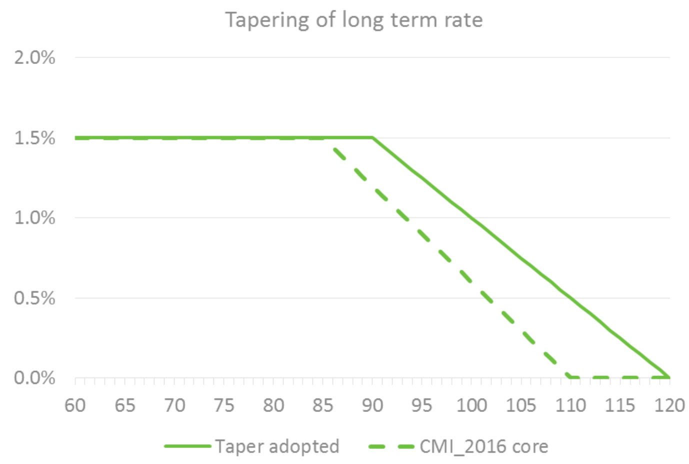

### 6.3.4 Long term rate

For the purposes of our analysis, the long term rate (in the age/period dimension) is assumed to be 1.5% p.a., while we have retained a cohort long term rate of 0% p.a..

This choice of long term rate is purely illustrative, in order to provide a ‘baseline’ projection for comparative purposes, rather than necessarily representing our view of an appropriate long term rate.

In addition, we have assumed that the long term rate declines linearly above age 90 to 0% p.a. at age 120. This reflects the likely ‘aging’ of improvements, and brings it into line with the core assumption of previous versions of the CMI model (up to and including CMI_2015), as well as the previous edition of the PLSA model.

Note that the core setting for CMI_2016 assumes the decline occurs between ages 85 and 110.




## 6.4 Projecting life expectancies

We calculate historical life expectancies (from age 65) based on 3 year smoothed mortality rates for each year before 2015 (while the life expectancy in 2015 is based on smoothed mortality rates in 2015 alone).

For future life expectancy calculations, we can apply the mortality improvements generated by the CMI_2016 model, calibrated as discussed in Section 6.3, to the smoothed historical mortality rates to generate mortality rates for each future year.

Note in particular that as we project smooth rates, we start from the values in 2014 (rather than 2015, as the mortality rate in 2015 does not have smoothing applied), so must apply 2 years of increases to get the assumed mortality rates in 2016.

### 6.5 Choosing a typical projection

In our report we chose a 'typical' projection to use as a comparison when projecting life expectancies into the future. Given the near universal use of the CMI model for such purposes, the choice was essentially which version of the CMI model (and so underlying data) to reference.

We elected to use the CMI_2015 model for this purpose, on the grounds that:

- It was, at the time of publishing the report, the immediately preceding version to CMI_2016
- Given CMI_2016 had only just been published, it was felt that few schemes would yet have had the chance to adopt it.
- Therefore schemes which routinely adopted the 'latest' model are likely to have used CMI_2015 most recently.
- It included (population) data in respect of (part of) 2015, the last year included in the pension scheme data.

A case could potentially be made for instead adopting a previous version of the model, given a number of commentators were at the time concerned about adopting CMI_2015, given the lower life expectancies, and so liabilities, that would result. A number of pension schemes therefore elected to retain a previous version of the model. However, given this projection was intended to be purely illustrative, we settled on using CMI_2015.

Having chosen the CMI model version, we also needed to decide on how to calibrate it. For consistency with the pension scheme data based projections, we assumed a long term rate of 1.5% p.a.[13], and used the core settings of the CMI_2015 model.

Note in particular this meant that we retained the core tapering of the long term rate from age 90 to 120, rather than the revised core tapering (85 to 110) which was introduced for the CMI_2016 model. Again this is in line with the approach adopted when calibrating CMI_2016 to VitaSegments.

[13] Note that CMI_2016 introduced a subtle change in 'currency' of long term rate, as it changed to applying to mx rather than qx. As such, a long term rate of 1.5% p.a. in CMI_2016 is actually a slightly weaker assumption than 1.5% p.a. in CMI_2015 (and previous versions).

# 7 Our illustrative schemes

## 7.1 Generating example schemes

In order to illustrate the impact of using socio-economic groups to project future improvements it is helpful to consider some example schemes. We have therefore designed four ‘example’ schemes for this purpose.

Based upon the socio-economic mixes and age profiles seen within Club Vita, and including both pensioner and non-pensioners members, they are designed to be broadly representative of the range of UK DB pension schemes.

| | | |
|---|---|---|
| **A** | Mature, lower socio-economics | Mature scheme skewed to lower socio-economics. Probably closed to future accrual. Similar to schemes from heavy manufacturing industries |
| **B** | Examples of broadly typical mixes | Broad mix of socio-economic groups. Likely to be similar to schemes from consumer services or cyclicals and also local government schemes. |
| **C** | | Mix of socio-economic groups, although biased towards higher groups. Likely to be similar to schemes from technology, pharma and skilled engineering industries. |
| **D** | Higher socio-economics | Long standing scheme. Skewed towards higher socio-economic groups. Likely to be similar profile to schemes from financial services sector. |

Full details of these example schemes (including age and socio-economic profiles) can be found in the Technical Appendix published alongside the original longevity trends model[14].

## 7.2 Impact of improvements on example schemes

In order to assess the expected impact on these example schemes of allowing for socio-economic group when setting future improvements, we have calculated approximate liabilities for each scheme.

In doing so we have made a number of simplifying assumptions, both around benefit structures and demographic and financial assumptions.

Again fuller details are provided in the Technical Appendix published alongside the original longevity trends model.

We have however we have made a number of changes, as set out below.

- The valuation year is now assumed to be 2017.
- Baseline mortality is taken to be as (3 year smoothed) experience for 2014 (the latest year available where 3 year smoothing can be applied to pension scheme data).
- As before, we extrapolate mortality rates at older ages (above 95) using the approach adopted by the CMI SAPS committee in S2 series mortality tables.
- Net discount rates are assumed to be 0% p.a. pre-retirement and -1% p.a. post retirement (to reflect market conditions at publication).

In each case we compare the approximate liabilities against a ‘typical’ basis of CMI_2015 with a 1.5% p.a. long term rate.

[14] See https://www.clubvita.co.uk/collaborative-research/supporting-technical-appendices

# 8 Creating scenarios for future improvements

## 8.1 Introducing our scenarios

In our 2014 report we introduced a series of six health 'scenarios' for future improvements in life expectancy. The scenarios included two which assumed low/negative future increases, two which assumed relatively high future increases, and two more 'central' assumptions. In each case we created a 'real world' narrative around the scenario.

We have updated each of these scenarios for the passage of time for our 2017 report. In most instances this simply means launching the scenarios off from 2015 (rather than 2010 as used previously), and using the CMI_2016 model. However a number of scenarios (notably our 'Health Cascade') have received further refinements to reflect experience from 2010 to 2015.

We have also taken the opportunity to introduce two new 'central' scenarios in light of the uncertainty surrounding recent trends:

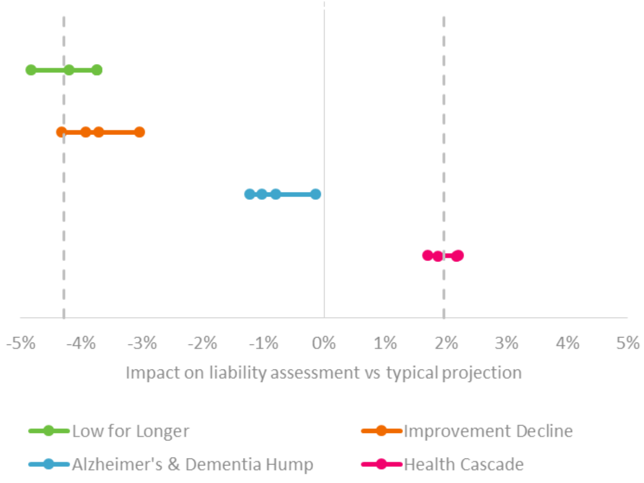
- '**Low for longer**': This scenario considers the impact of sustained low economic growth / austerity on longevity, and how this may impact the socio-economic groups differently.
- '**Alzheimer's & Dementia wave**': This scenario builds on the recent rise in numbers of deaths attributed to Alzheimer's & dementia. It continues this rise for a few years, before a period of rapid decline as a result of successful interventions / cure.

For full details of these eight scenarios, including specifics around the calibrations of the CMI_2016 model, please see our separate guide[15].

## 8.2 Assessing the impact

We carry out calculations of the likely impact of each scenario on our four example schemes, using the same approach discussed in Section 7. As before, the impacts are compared to using a 'typical' projection of CMI_2015 with a 1.5% p.a. long term rate.

The chart below summarises the impacts on the four sample schemes of the four 'central' scenarios. These scenarios are broadly in keeping with what many DB pension schemes use for funding purposes.




The broad spread between scenarios for any given scheme is around 6%. In contrast the variation within any given scenario is around 1½%. This highlights the importance of considering the socio-economic mix of a pension scheme's membership when setting the funding assumption.

[15] See 'A guide to the PLSA longevity trends model Scenarios'