Society of Actuaries in Ireland

Review of best estimate mortality projection methods

September 2020

1. Introduction

The Demography Committee of the Society of Actuaries in Ireland ("Society") is pleased to present the results of the analytical review of methodologies in projecting best estimate mortality improvements for the Irish population. Users of this paper are reminded that when interpreting or applying information from the paper, judgement must be exercised in relation to the appropriateness, sufficiency and suitability of this information for the particular purpose.

Overview of paper

This paper summarises the analytical review performed by the Demography Committee in respect of a range of potential mortality improvement models in the context of best estimate, mortality improvement projections. Section 2 is a summary of the conclusions; section 3 sets out the models selected and the assessment criteria; section 4 details how the models were calibrated; section 5 sets out the projected results (including life expectancies); section 6 compares the models relative to the assessment criteria and section 7 concludes on the findings of the analysis.

The summary contains an overview of the main findings of the investigation, although readers are encouraged to read the full paper.

Governance and Authorship

This paper has been prepared in accordance with ASP PA-2 General Actuarial Practice[1] and the Code of Professional Conduct[2]. Authors of the paper consist of members of the Demography Committee who are Fellows of the Society. The paper was approved by Council as a Society paper. Details of the authors are recorded in a Governance Document. The Governance Document outlines or references, as appropriate, the governance and associated process controls of the mortality improvements project.

Peer Review

In line with best practice, this paper, together with the Governance Document, has been independently peer reviewed. The peer review encompassed a review of the analysis and commentary in the final paper only.

Potential Impact of Covid-19

The current Covid-19 pandemic and the recession which is expected to follow it are likely to have a significant impact on short-term base mortality rates but it is too early to comment on the potential impact on the longer-term term view which drives mortality assumptions. Improvements are generally accepted to be a longer term assumption and therefore the impact of any fluctuations will be dampened. Furthermore, it is too early to undertake any credible statistical analysis as there is little or no data available. Therefore, no adjustments have been made to this paper to reflect (other than this disclosure) to reflect the potential impact of covid-19.

---

[1] https://web.actuaries.ie/standards/asp/asp-pa-2
[2] https://web.actuaries.ie/standards-regulation/code-professional-conduct

Disclaimer

Whilst care has been taken to ensure the accuracy of the information in this document, the Society does not accept any responsibility or liability for any errors and/or omissions, including any errors and/or omissions in the data on which this document is based. This document does not constitute advice and should not be relied upon as such. The Society does not accept any responsibility or liability for any loss to any person or body as a result of any action taken, or any decision taken not to act, on the basis of any statement, fact, figure or expression of opinion or belief contained in this document.

Further Information

The Demography Committee welcomes feedback on the paper and suggestions for improvements. Comments or questions can be submitted to info@actuaries.ie.

2. Summary

The Demography Committee of the Society of Actuaries in Ireland has published a number of studies of base mortality in respect of annuity product, assurance products and self-administered pension schemes. The Committee has also previously published its views on mortality improvement assumptions for Irish lives in respect of illustrations and transfer values. The Committee felt that a more detailed paper exploring a range of options for mortality improvements might be of interest to members.

The investigation into different mortality improvements models was carried out using Irish population mortality data covering the period 1986-2017. A range of different models were assessed against a list of generally accepted assessment criteria. As a measure of the relative significance of the model chosen in the context of some of the common uses of improvement assumptions, the life expectancies produced by the different sets of improvement rates were compared at key ages.

It is expected that this paper will be of use to actuaries working in Ireland as part of their considerations of mortality improvement assumptions for different purposes. No single approach is best in all situations and it is the responsibility of users of mortality improvement projections to ensure that they are appropriate for the particular purpose for which they are being applied. However, for purpose such as point-in-time valuation at retirement, the CSO model generally provides a good balance between the different criteria.

Throughout this paper, analysis has been sub-divided into genders of male and female to facilitate comparative analysis. In carrying out the analysis for these two groups only, we do not imply that these are the only genders possible but we were unable to identify other genders in the data used.

Performance against assessment criteria

The following table shows of the suitability of each of the models assessed against the criteria discussed throughout this paper. The choice of model for any given purpose depends on the relative importance the user places on each measure for that purpose.

Criteria	CSO	Lee-Carter	Cairns-Blake-Dowd	CMI
Trends in Mortality	4	2	3	1
Transparency	1	2	2	4
Ease of Use	1	3	3	2
Flexibility	2	3	3	1
Model Updates	1	2	2	3
Robustness	N/A	1	1	2
Parsimony	1	3	2	4

Table 1: Subjective Ranking of Models by Selected Criteria

The table above clearly shows that no single model outperforms the rest in all categories. This is largely unavoidable given the natural trade-off between certain characteristics – for example, adding more parameters will generally improve the fit to trends but reduce parsimony. Similarly, added flexibility can reduce the ease of use.

If all criteria are equally weighted by the user then the CSO model appears to be the most appropriate choice as it perform well against the majority of criteria. The one criteria against which it performs poorly is the ability to reflect mortality trends. For many purposes, users will apply higher weighting to this criteria but when considered in conjunction with the life expectancies set out in Table 2, this criteria is less significant for some purposes (e.g. point in time valuation at retirement).

Comparison of life expectancies

The table below shows the calculated cohort life expectancies at key ages. All life expectancies are as at 1 January 2020 and were derived using a base mortality assumption of 50% ILT15.

Model	Males			Females		
	60	65	68	60	65	68
CSO_2018	29.4	24.5	21.7	32.3	27.3	24.3
Lee-Carter	29.8	24.6	21.6	32.4	27.3	24.3
Cairns-Blake-Dowd	30.0	24.8	21.8	32.7	27.5	24.5
CMI_2018 – Consistent	29.1	24.3	21.5	31.9	26.9	24.0
CMI_2018 – Core	29.1	24.3	21.5	32.0	27.0	24.1

Table 2: Cohort Life Expectancies at Ages 60, 65 and 68, Calculated for Each Model, (50% ILT15, Base Year 2020)

For both genders, the range of cohort life expectancies produced by the various models is relatively narrow. The age 65 projections fall within a 0.5 year range for males and a 0.6 year range for females.

3. Purpose and approach of the paper

3.1. Purpose

The purpose of this paper is to summarise the analytical review performed by the Demography Committee in respect of a range of potential mortality improvement models in the context of best estimate, mortality/longevity improvement assumptions for the Irish population. We expect this paper to be of interest to actuaries working in the Irish market. It also expected that the analysis set out in this paper will form part of the Society's considerations when reviewing the mortality assumptions in the various ASPs.

This paper, and the work underlying it, is focussed on best estimate mortality/longevity improvement modelling. Although the criteria we have used in this paper are common market/academic criteria, other criteria may be important when using stochastic projection methodologies for other purposes.

This paper sets out the relative merits of a number of mortality improvement methodologies against our specified criteria. Where data permitted, the full age-range was considered in the analysis underlying this paper. However, data is more credible at longevity ages (i.e. age 60+) and the market as a whole, has a greater interest in best estimate assumptions at these ages, including the Society's own use of these in ASPs. Therefore, the analysis presented in this paper is primarily focused on longevity ages.

The analytical approach set out in this paper may also aid actuaries in developing their own analysis of mortality improvement projections for other data sets of interest.

The paper is technical and assumes users have some prior knowledge of this area.

3.2. Model selection

The mortality improvements models considered in this paper were as follows:

- CSO_2018 (the CSO's latest mortality improvements model underlying its Irish Population and Labour Force Projections 2017 to 2051)
- The Lee-Carter model
- The Cairns-Blake-Dowd model.
- CMI_2018 (the Continuous Mortality Investigation's ("CMI") mortality improvements model, version 3 of CMI_2018[3], published in 2019).

In selecting models to be included in our analysis, we considered the main model families in use in the market and in academic research. Our selected models include one from each major family except for P-Splines. Some variants of the above models were also considered but we concluded that the chosen

[3] CMI_2019 is available as at the time of writing. However, the only change from the previous iteration is updated data for England Wales. CMI_2018 has been in more regular use and testing by users, with some updates to the model since the initial release.

models provided a broad coverage. Consideration was also given to the Renshaw-Haberman extension of the Lee-Carter model, which incorporates cohort effects. There was difficulty in fitting the model and the resulting improvement rates were very volatile by age. We therefore do not believe that this model is suitable for use with Irish population data. We excluded P-Splines as, despite much academic research and CMI promotion, P-Splines are not widely used in industry and the CMI's own projection model methodology moved away from P-Splines with the release of CMI_2016.

The main projection methodologies of the models considered are targeting and extrapolation, defined as follows:

- Targeting - interpolation in some manner between current rates of mortality improvement and the targeted future long-term rate.
- Extrapolation - "business as usual" continuation of historical trends.
- Hybrid – combine features of both extrapolation and targeting projection type.

The CSO model falls into the targeting category. The Lee-Carter and Cairns-Blake-Dowd model belong to the extrapolation family and the CMI model is a hybrid model which uses an Age-Period-Cohort model to project short-term mortality rates and then converges to a pre-defined long-term rate.

More detail on the form of each model can be found in Appendix A. An important point to note is that he Lee-Carter and Cairns-Blake-Dowd model only incorporate age- and period-effects whereas the initial component of the CMI model incorporates cohort-effects, as well as age- and period-effects.

This paper focusses exclusively on Irish data. Therefore, one form of model not explored is multi-population models. These are sometimes used for smaller populations. Data from other regions which have similar characteristics are used in the projection of mortality trends. This form of model may be an area of consideration in future work on this topic.

3.3. Assessment criteria

Considering the main model criteria adopted both in academic and industry studies, the criteria applied for the assessment outlined in this paper are:

- **Incorporation of trends in mortality** – a model should reflect the experienced trends in mortality. It should provide a good fit to the historical data used to calibrate it. Where they are present, separation between age, period and cohort-effects should be allowed for by the model.
- **Transparency** – users of the model should be able to understand the elements that are driving the projections. Ideally, a user should be able to understand why a change in input data or parameters have the impact that they do. Essentially, "black-box" methodologies should be avoided.
- **Ease of use and implementation** – models should not be unnecessarily complex to set up or run. The more difficult a model is to set up and/or run, the greater the risk of mistakes and the greater the likely dependence on individual specialist users.
- **Flexibility** – a model should allow users the flexibility to modify projections to suit their own views and purpose. Users may have views on how past drivers of mortality improvements will change over time. It is desirable to have the ability to allow these views to be reflected in the projections.
- **Model updates** – a model can be regularly updated over time to reflect emerging experience
- **Robustness of parameter estimates and forecasts** – large changes in projected mortality rates would not be expected from the addition of a small number of years' mortality data. Any methodology that leads to large fluctuations in mortality rates, or expectations of life, is not likely to be seen as appropriate.

- **Parsimony** – models should use as few parameters as possible. Additional parameters are only included when they lead to a significant improvement in the fit of the model.

4. Model Calibration

With the exception of the CSO model, all of the models being investigated require a dataset for calibration, as well as assumptions to underlie the projections. In this section, we will explore the decisions that need to be made with regards to model calibration. The main aim for the calibration is consistency between the different models so that the results are directly comparable. Other users may decide that they wish to use the models differently.

The CMI model has a wide array of parameters that can be selected by the user. This paper will not seek to derive the most appropriate calibration of the model for the Irish population but will consider the model calibrated with its core values. For example, the period-smoothing parameter (S-kappa) will be set to 7.0. An exception to this is the calibration period, which will be adjusted. As the core values for some parameters are not consistent with the assumptions for the other models under investigation, two projections of the CMI model will be produced: a "core" calibration, and a "consistent" calibration.

4.1. The Calibration Data

In order to calibrate the range of projection models we had to choose an appropriate set of historical experience data. The only available dataset of an appropriate duration is Irish population data.

The Central Statistics Office (CSO) publish national statistics for the Republic of Ireland, including annual population size estimates and death volumes split by gender and age. The Human Mortality Database (HMD)[4] is an online initiative which collates population data for countries around the world. The HMD apply refinements in order to produce datasets which allow consistent analysis of mortality and mortality trends. For Ireland, the HMD uses CSO data in creating its output. HMD data is widely used for academic purposes and the methodology which it follows is extensively documented.

Considering the long-term nature of mortality/longevity improvements and the manner in which improvement assumptions are used in practice, we decided that annual exposure and death figures over as long a period as the data was reliable was the most appropriate input for calibrating improvements.

4.2. The Calibration Period

It is necessary to choose the period of experience to which the various models will be calibrated. Generally, a balance will be sought between using a large enough dataset and using experience that is relative to current trends. There is no standard approach but, for reference, the CMI currently uses 41 years of data in the core calibration of their model.

There is no reason not to include the most recent data in the calibration. The latest year for which death figures are available is 2017. Therefore, this will be the end point of the calibration data. The Irish data has specific features which influence the decision on calibration period. There are serious limitations in relation to data available prior to 1950 and the HMD does not provide any data before

[4] https://www.mortality.org/

this point. While data is provided for the 1950-1985 period, there are known data quality issues. One significant issue is age-heaping whereby the recorded age for individuals is rounded.

On the basis of concerns about the potential impact of invalid data, we decided to primarily investigate models calibrated to data for the period 1986-2017. However, we did perform analysis on the impact of including the data covering 1950-1985. This did have a significant impact on the projected improvement rates and life expectancies using the Lee-Carter and Cairns-Blake-Dowd model, with life expectancies up to two years longer when the more recent period was used. The impact on the CMI model results is much smaller, less than half a year, due to the emphasis placed on the most recent years in the input data.

4.3. The Calibration Age-Range

The age-range to which the models are calibrated to is another area of judgement.

The volumes of historic data at ages above 90 is relatively small, which results in significant volatility in the annual mortality rates at these ages. We therefore decided to use a maximum fitted age of 90 for all models except for the "core" calibration of the CMI model as this uses a maximum age of 100 by definition.

The form of some of the models influenced our decision as to what is an appropriate age-range. Including ages below 60 does not have an impact on the projections using the Lee-Carter model, which we confirmed with investigations using both 25-90 and 60-90 fitted age-ranges. The Cairns-Blake-Dowd model assumes a smooth pattern of improvements by age, generally downwards-sloping. For this reason, it is typically used for older age ranges where this pattern often mirrors experience. Including younger ages in the calibration data may cause the model to insufficiently capture the pattern of improvements at the longevity ages we are most interested in (i.e. ages 60+). Where data is sufficiently credible, the models could be calibrated separately for younger age ranges to provide an indication of mortality improvements for this age group rather than considering it on a combined basis.

Testing on the CBD model using a wider or narrower age-range was performed to show that including younger ages for Irish data produced results which did not fully reflect the experience. A fitted age-range of 20-90 produced life expectancies for males that were roughly half a year longer than 60-90. Analysis of a scatterplot of standardised deviance residuals by age provided support for the 60-90 age range, which was used for the results presented throughout this paper. It is desirable for the fitted residuals to be normally distributed, with no obvious patterns by age. Such patterns could indicate age-effects which are not being accurately modelled. The CBD residuals for the 20-90 age-range are clearly not normally distributed and it is evident that the patterns of mortality by age are not being captured appropriately. We opted to use the 60-90 age-range for the results presented throughout this paper.

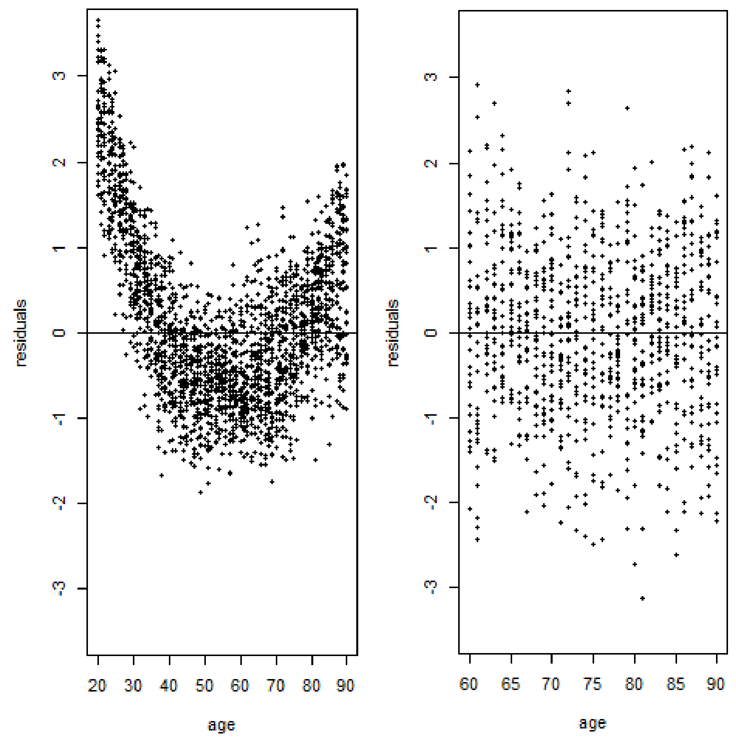

Figure 1: Plot of Male Residuals by Age for the Cairns-Blake-Dowd Model, Fitted to Ages 20-90 (Left) and 60-90 (Right)

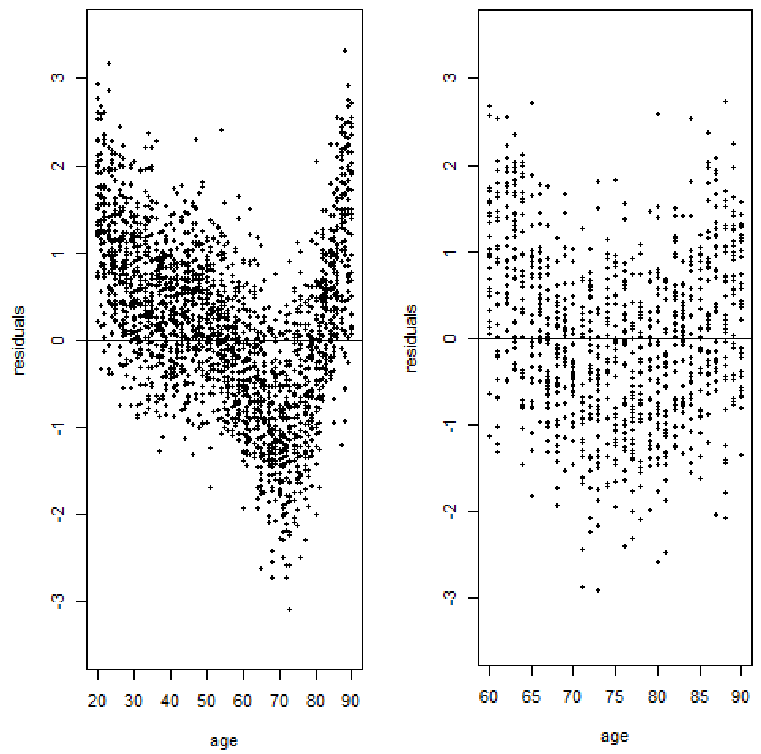

Figure 2: Plot of Female Residuals by Age for the Cairns-Blake-Dowd Model, Fitted to Ages 20-90 (Left) and 60-90 (Right)

In contrast, models which incorporate cohort effects will require historic data at ages below 60 to project forward trends at ages above 60. Therefore, for these models data from age 20 was used.

Considering the factors discussed above, although the core setup of the CMI model uses a calibration age-range of 20-100, we adopted a calibration age-range of 20-90 for the CMI model for the "consistent" calibration.

4.4. Improvements at Older Ages

As discussed in section 4.3, the models have been fitted to data with a maximum age of 90 or 100. We do also require estimates of improvements at older ages. Typically, it is assumed that improvements converge to 0% at some age. This reflects the observed experience that improvements rates have tended to decline at the oldest ages.

Age Group	Average Annual Improvement Rate
40-59	2.66%
60-79	2.96%
80-100	1.44%

Table 3: Average Annual Improvement Rate by Age-Band, Both Genders Combined, 1986-2917, Measured by SMR

The CSO model assumes that improvement rates converge to 0% between ages 90 and 100. The core setup of the CMI model assumes that initial improvement rates converge to 0% between ages 100 and 110, and that long-term rates converge to 0% between ages 85 and 110. We have decided that converging over ages 90-100 is a sensible approach and have also used this as our assumption for the models in this paper, with the exception of the core parametrisation of the CMI model.

4.5. Long-Term Improvement Rate

The CSO model has a default long-term improvement rate of 1.5%. We have decided to use this value for consistency between the models as the CMI model requires an assumption to be made for the long-term improvement rate. There is no core value for this parameter and the CMI advises users to choose a value based on their own views on long-term trends.

4.6. Summary of Model Calibration

Model	Projection Type	Calibration			Initial Improvements	Long-Term Improvements		Old-Age Convergence to 0%
		Data	Period	Age-Range		Rate	Convergence	
CSO_2018	Targeting	N/A	N/A	N/A	Males - 2.5% Females - 2.0%	1.5%	25 Years	90-100
Lee-Carter	Extrapolation	HMD-Irish Population	1986-2017	60-90	From Model	From Model	N/A	90-100
Cairns-Blake-Dowd	Extrapolation	HMD-Irish Population	1986-2017	60-90	From Model	From Model	N/A	90-100
CMI_2018 – Consistent	Hybrid (Age-Period-Cohort)	HMD-Irish Population	1986-2017	20-90	From Model	1.5%	5-40 Years Varies by Age	Initial – 90-100 Long-Term – 90-100
CMI_2018 – Core	Hybrid (Age-Period-Cohort)	HMD-Irish Population	1986-2017	20-100	From Model	1.5%	5-40 Years Varies by Age	Initial – 100-110 Long-Term – 85-110

Table 4: Summary of Model Calibration

[5]

[5] There are different convergence periods for the age-period and cohort effects in the CMI model. They also vary by age. More details are available in Appendix A

4.7. Fitting the Models

The CSO model is pre-determined and does not require fitting.

The Lee-Carter, Renshaw-Haberman, and CBD models were fitted using the “StMoMo” ‘R’ library for stochastic mortality modelling. The period effects in both the Lee-Carter and CBD models are projected using time-series methods.

The CMI model is Excel-based. All parameters can be changed via the user interface and the calibration dataset can be easily replaced with the desired population.

5. Projected Results

5.1. Comparison of Life Expectancies

Cohort life expectancies as at 1 January 2020 were derived using a base mortality assumption of 50% ILT15. This base mortality assumption is expected to be reasonably representative of Irish assured lives, annuitants or pensioners. In this context, we are concerned with the relative performance of each improvement model and therefore the base mortality assumed is not a significant factor.

The table below shows the calculated life expectancies at key ages.

Model	Males			Females		
	60	65	68	60	65	68
CSO_2018	29.4	24.5	21.7	32.3	27.3	24.3
Lee-Carter	29.8	24.6	21.6	32.4	27.3	24.3
Cairns-Blake-Dowd	30.0	24.8	21.8	32.7	27.5	24.5
CMI_2018 – Consistent	29.1	24.3	21.5	31.9	26.9	24.0
CMI_2018 – Core	29.1	24.3	21.5	32.0	27.0	24.1

Table 5: Cohort Life Expectancies at Ages 60, 65 and 68, Calculated for Each Model, (50% ILT15, Base Year 2020)

For both genders, the range of cohort life expectancies produced by the various models is relatively narrow. The age 65 projections fall within a 0.5 year range for males and a 0.6 year range for females. The CSO model produces life expectances that are significantly lower than both the LC and CBD models at age 60 for males, but the life expectancies at ages 65 and 68 are more closely aligned. For females, the CSO model produces results which are very similar to the LC model. The CBD model consistently produces the highest life expectancies. The differences in life expectancy between the "core" and "consistent" CMI projections are minimal. The CMI model produces the lowest life expectancy for both genders.

5.2. Comparison of Initial Improvements by Age

Figure 3 and Figure 4 below show the projected improvements in 2020 for ages 60-90. 2020 does not represent the "initial" year for the CSO and CMI model so a small level of convergence towards long-term rates has already occurred. The initial projection year for these models are 2016 and 2018 respectively.

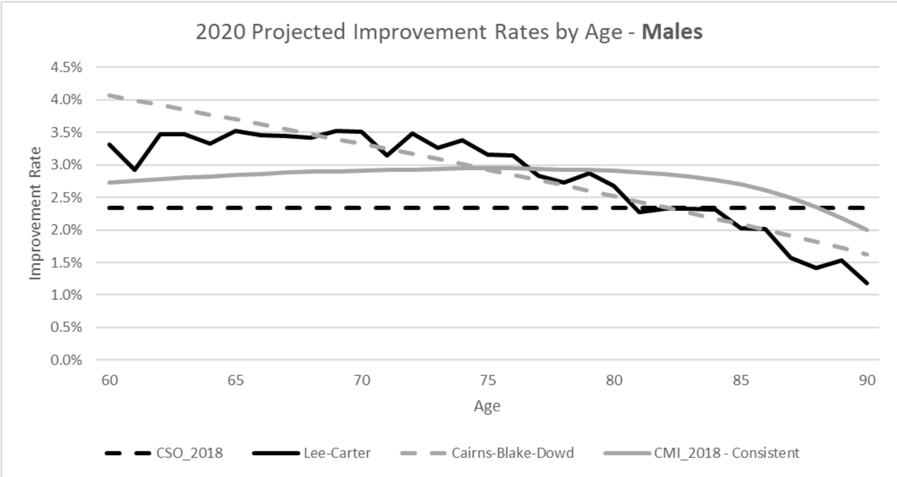

Figure 3: Projected Male Improvements in 2020 for Ages 60-90

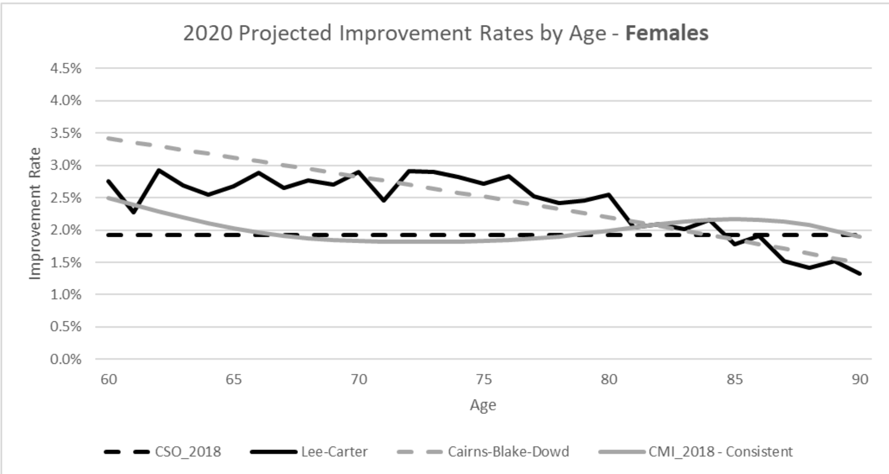

Figure 4: Projected Female Improvements in 2020 for Ages 60-90

The shapes produced follow the patterns we would expect for each model – uniform for the CSO model, volatile for the LC model, smooth and straight for the CBD model, and smooth but curved for the CMI model.

The CBD model rates generally appear like smoothed versions of the LC rates, although the projected rates for the LC model are flatter in the 60-70 age-range.

There is reasonably close alignment between the initial rates projected by both the CSO and CMI models, particularly for females. Both models are intended to pick up the most recent trends in their initial projections.

It is important to show the composition of the initial projected rates for the CMI model, as the split between cohort- and age-effects will have a large impact on the way the improvement rates are projected forward into the future.

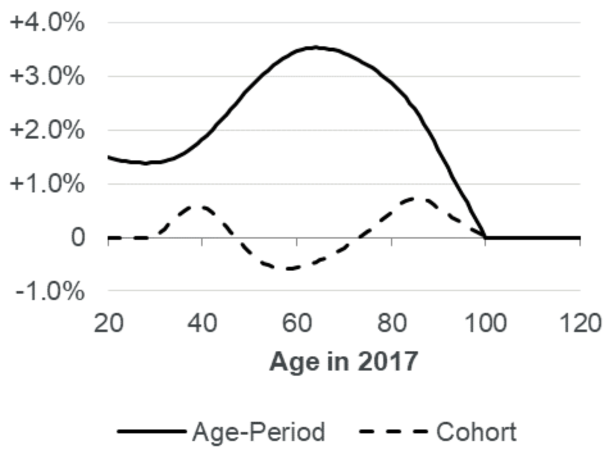

Figure 5: Split in Projected Male Initial Improvements for CMI Model

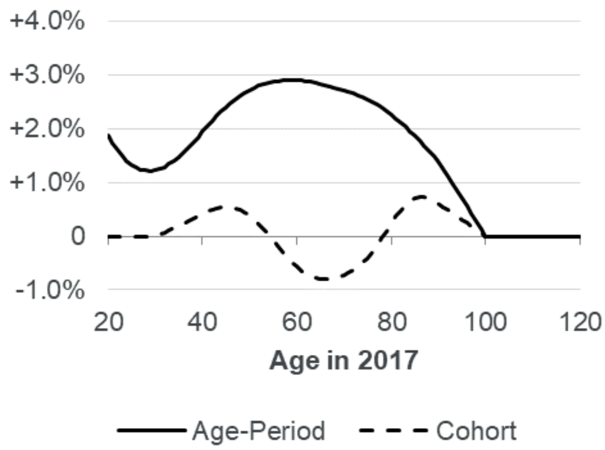

Figure 6: Split in Projected Female Initial Improvements for CMI Model

The patterns are quite similar for both genders, with negative cohort effects being projected for individuals currently aged in the typical range for retirements. However, the age-effects being projected are much larger in magnitude, particularly for males. The negative cohort effects are the main reason that the CMI model produced the lowest life expectancies in Table 5.

5.3. Comparison of improvements by projection year

The following graphs show the projected improvements of the cohort aged 65 beginning in 2020. Table 5 showed that life expectancies are broadly consistent across all models. Nonetheless, the different features of the models mean we expect to see differences in the individual projected improvements by year and this is clearly evident in the graphs below.

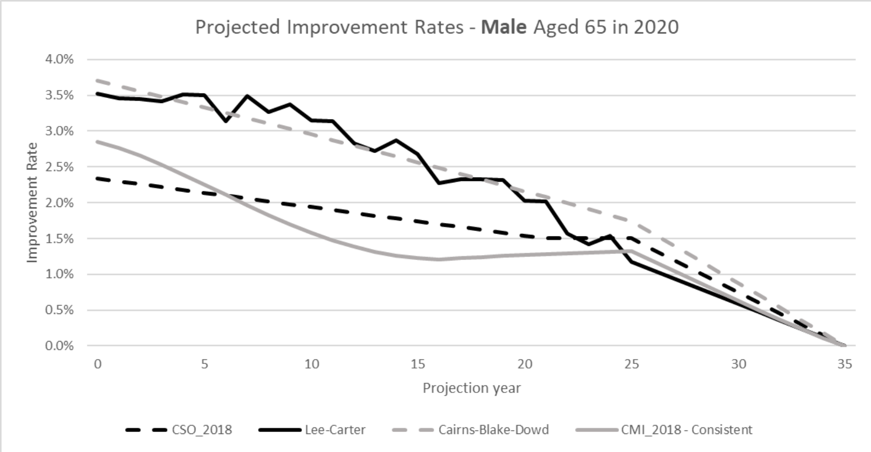

Figure 7: Projected Improvements for Males age 65 in 2020

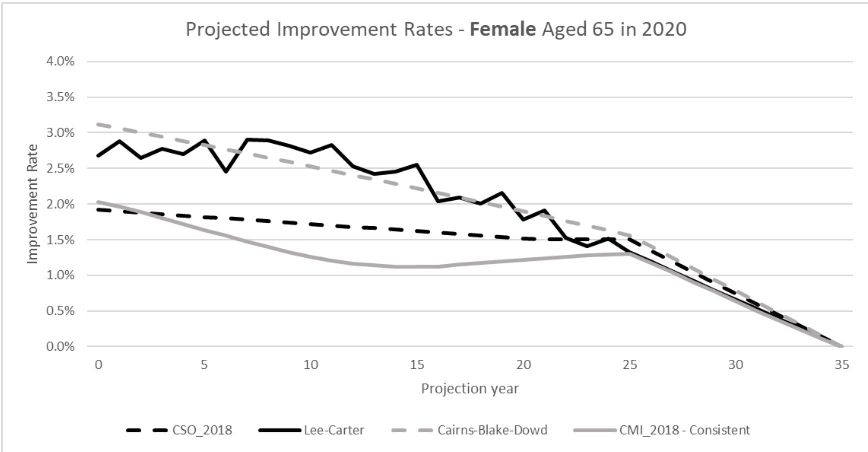

Figure 8: Projected Improvements for Females aged 65 in 2020

The rates projected by the CSO model will linearly reduce from the initial rates to the long-term rates by 2040. The LC and CBD models both have invariant shapes by age in all future projected years. The CMI model includes cohort effects as well as age-effects. Therefore, the shape by age will be different for different projection years, until the long-term rate has been fully achieved.

Both the CSO and CMI models converge to the same long-term rate. However, in the medium-term, there is divergence, even where the initial rates are very similar for females. The CSO model linearly converges to the long-term rate, whereas the CMI model factors in cohort effects and uses non-linear convergence.

6. Consideration of model criteria

6.1. Incorporation of Trends in Mortality

There are some complexities in comparing all models in a consistent and fair way:

- The CSO model does not fit historic improvements and the goodness-of-fit cannot be assessed.
- The CMI model has been fitted using a larger calibration dataset than the LC and CBD models, which means that goodness-of-fit measures cannot be compared.
- The initial rates of the CSO and CMI model are intended to reflect the most recent trends, rather than patterns over the full calibration period, as is the case for the LC and CBD models.

The following table shows the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) for the LC and CBD models. Both tests measure goodness-of-fit, taking into consideration the number of parameters. The BIC penalises overfitting to a greater extent. A lower number is indicative of a better fit.

Measure	Males		Females	
	Lee-Carter	Cairns-Blake-Dowd	Lee-Carter	Cairns-Blake-Dowd
AIC	8973	9024	8689	9083
BIC	9423	9336	9138	9396

Table 6: AIC and BIC Calculated for Lee-Carter and Cairns-Blake-Dowd Models

For males, both models appear to provide a fit of similar quality. The LC has a slight advantage using the AIC but has a slight disadvantage when measured with BIC, owing to the greater number of parameters (172 vs. 64 in this case). On the other hand, the LC model appears to have a clear advantage for females, with significantly lower values for both the AIC and BIC.

Another potential way to compare the goodness-of-fit is through the deviance residuals. These are the standardised deviation between the fitted and actual mortality rates in each cell. We can look at heatmaps of the residuals by age and year to identify clusters of similar residuals which may represent a poor fit. In particular, diagonal lines of similar deviances may indicate that cohort-effects are present in the data but are not being captured.

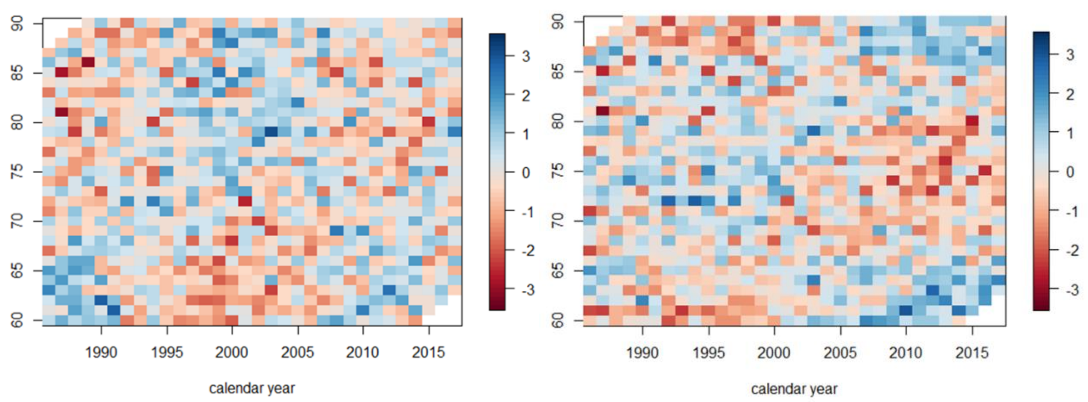

Figure 9: Heat Map of Male Residuals by Age and Calendar Year, for Lee-Carter (Left) and Cairns-Blake-Dowd (Right) Models

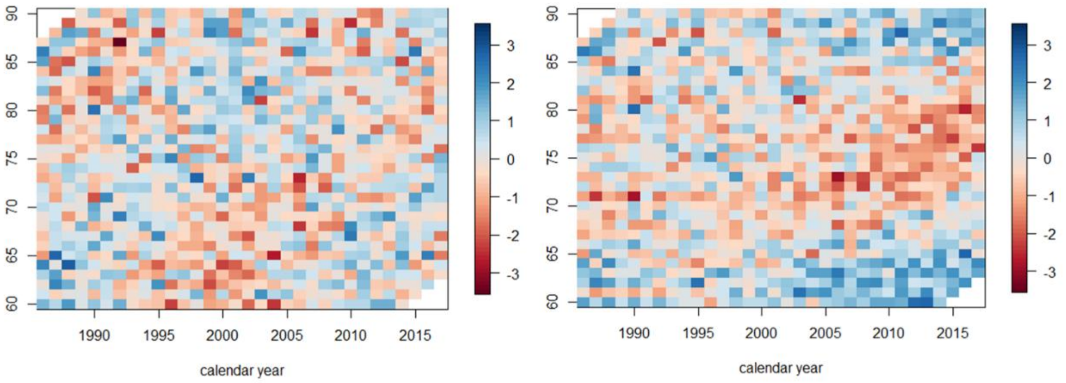

Figure 10: Heat Map of Female Residuals by Age and Calendar Year, for Lee-Carter (Left) and Cairns-Blake-Dowd (Right) Models

There is not significant evidence of clustering or cohort effects in the residuals of the LC model. Some clustering can be seen in the recent years of the CBD model, particularly for females.

Similarly, scatter plots of residuals by age, year, and year of birth, can be used to demonstrate whether any patterns in the data are not effectively captured by the models.

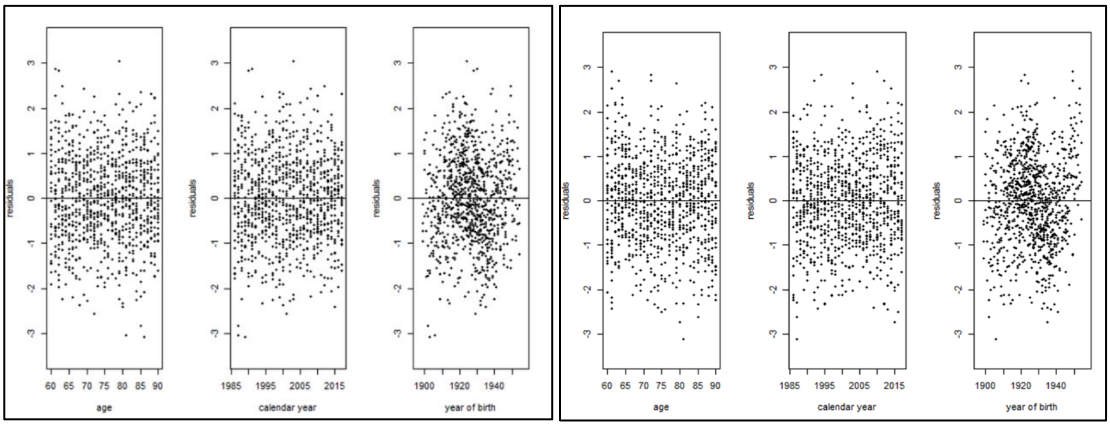

Figure 11: Plot of Male Residuals by Age, Calendar Year, and Year of Birth, for Lee-Carter (Left) and Cairns-Blake-Dowd (Right) Models

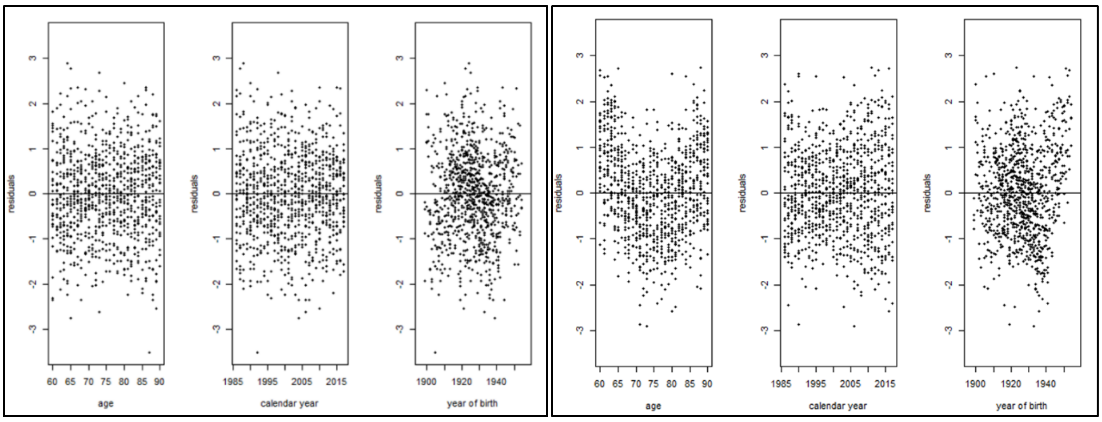

Figure 12: Plot of Female Residuals by Age, Calendar Year, and Year of Birth, for Lee-Carter (Left) and Cairns-Blake-Dowd (Right) Models

There are no clear patterns by age or calendar year for most of the fitted models. The one exception to this is the fit by age for females, where the residuals show a "v-shaped" pattern. A potential method for improving this may be an extension of the CBD model which incorporates a quadratic age-effect. There is some evidence of cohort effects in these graphs, with some downward trend in the residuals by year of birth.

For the CSO and CMI models, it is worthwhile to compare the initial rates of improvement to recent experience. As improvement rates by individual age will be very volatile, we have calculated the average annual improvement in Standardised Mortality Rates for 5-year age-bands. We compared these to the average initial projected rate in these bands.

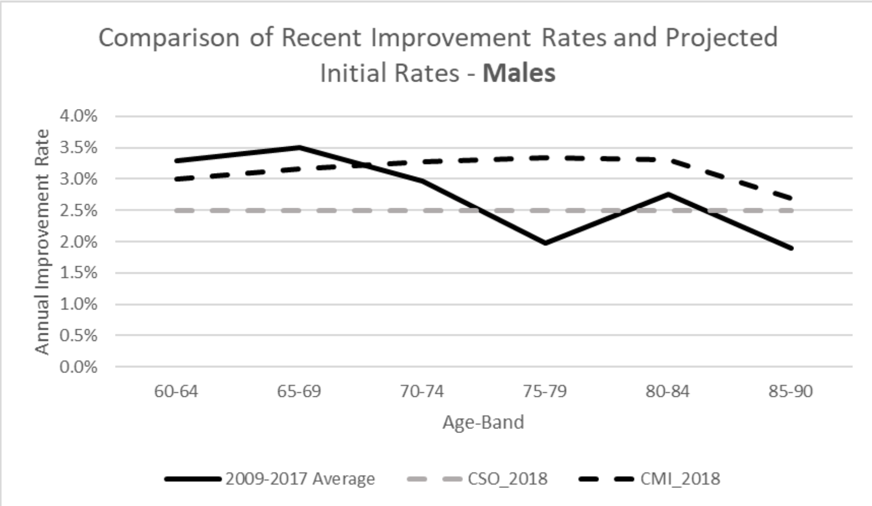

Figure 13: Graph of Average Male Annual Improvements in 2009-2017, Split by Age-Band and Compared to Initial Projected Improvements of CSO and CMI Models

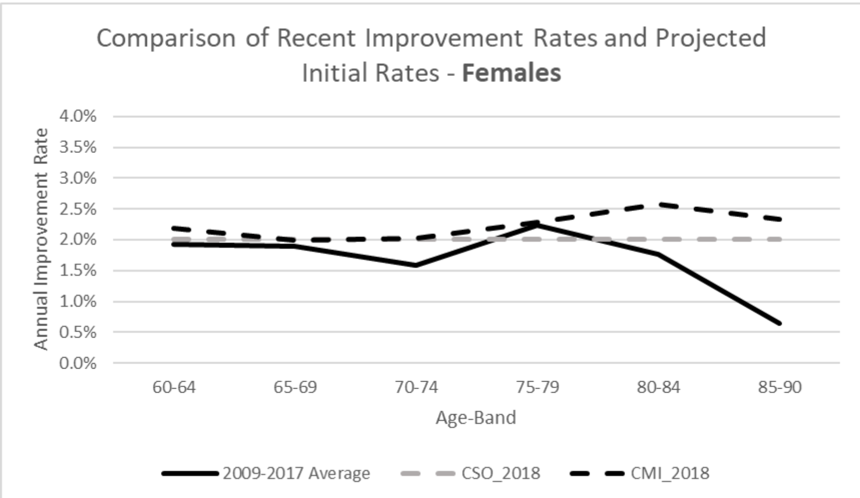

Figure 14: Graph of Average Female Annual Improvements in 2009-2017, Split by Age-Band and Compared to Initial Projected Improvements of CSO and CMI Models

Of course, a flat improvement rate, as with the CSO model, is not intended to capture any differentials by age. However, the initial rates in the CSO model provide a reasonably close match to recent experience, particularly for females. The CSO model projects lower improvements at younger ages for males but provides a good fit match at older ages. The CMI model projects higher improvements at older ages for males, in comparison to recent improvement rates. Both models project higher improvements for females at ages 80+.

6.2. Transparency

The CSO model is very easy to understand, with the improvement rate projections following a pre-determined formula. For both the LC and CBD models, the drivers of mortality projections are relatively straightforward. Broadly, higher improvements at a given age in the calibration data lead to higher projected rates for that age. The CMI model is more complicated – there are additional parameters such as smoothing parameters, and the cohort effects are more difficult to analyse than the simpler age-effects.

6.3. Ease of Use

The CSO model is extremely easy to implement, with very little time or expertise required. The LC and CBD models both have similar levels of complexity and risk in implementation. There will be a higher requirement for expertise in fitting and forecasting improvement rates under both models. The CMI model is user-friendly. Aside from replacing the calibration data, the user can choose to make no other parameter or option changes to the core set-up. The user interface allows options to be changed in a relatively easy manner.

Aside from the practical implementation considerations, judgement may be required in using any of these models. Even for the CSO model, the initial rate, convergence period, or long-term rate could be altered to reflect a differing view on improvements for the general population or the particular group of interest. For the other models, the calibration data, including the period and age-range, must be selected. The CMI models core calibration has been selected in reference to the standard calibration dataset. Users are urged to take care in implementing the model for different populations, with the smoothing parameters being highlighted in particular for consideration.

6.4. Flexibility

For the CSO model, the initial rate, convergence period, or long-term rate could be altered to reflect a differing view on improvements for the general population or the particular group of interest. The only area of flexibility with the LC and CBD models is the calibration data, including the period and age-range. As well as the calibration data, there are numerous areas of flexibility in the CMI model. These include options to change smoothing parameters, convergence features, and the long-term rate.

6.5. Model Updates

The CSO model can be easily updated, either because the CSO has released a new version, or because the user wishes to change the input values. This is on the assumption that the CSO does not make any fundamental changes to their projection model. The LC and CBD models can be very easily updated by adding in new years to the calibration dataset. The CMI model can similarly be updated by adding in new data. However, the user may decide to implement new versions of the CMI model, which are released on an annual basis. These updates sometimes include new core parameter values or changes to the underlying assumptions. If the user is incorporating a different parameterisation of the model, analysis would be required to ensure it remains appropriate for new data.

6.6. Robustness of Parameter Estimates

The CSO model does not have input data and so this item is irrelevant. Given the absence of reliable data prior to 1986, there is limited scope to test the robustness of the LC other models to the calibration period. Cutting the calibration period to less than 30 years would leave a low volume of data for fitting mortality improvements. Section 4.3 demonstrated the impact of widening the calibration age-range of the CBD model, which is not recommended. Broadly speaking, the LC and CBD projections should not be very sensitive to the addition of new data, provided the calibration period is already of sufficient length. The CMI model places a particular emphasis on recent trends through the period-effect. This causes improvements which are quite sensitive to emerging experience that is added. This can be viewed positively or negatively. The level of sensitivity can be altered through the period-smoothing (S-Kappa) parameter.

6.7. Parsimony

The CSO model has a clear advantage here, with an initial rate, convergence period, and long-term rate for each gender. The number of parameters used for the LC and CBD model is dependent on the calibration dataset. In the case of the projections in this paper, the LC and CBD models had 172 and 64 parameters respectively, for each gender. The CMI model uses many more parameters, with cohort effects estimates for each year of birth, and additional parameters such as the smoothing parameters. The AIC, and in particular the BIC, as shown in **Error! Reference source not found.**, incorporate penalties for over-parameterization.

6.8. Summary

The following table shows of the suitability of each of the models assessed against the criteria discussed throughout this paper. . We have used a red-amber-green indicator to allow users to quickly understand the strengths and weaknesses of each model. However, these are primarily relative assessments within the range of models we have assessed in this paper. A red assessment indicates that a model does not performs well as the other models against a given criteria but this does not mean that it is not suitable or even the preferred choice for some purposes. The choice of model for any given purpose depends on the relative importance the user places on each measure for that purpose.

Criteria	CSO	Lee-Carter	Cairns-Blake-Dowd	CMI
Trends in Mortality	4	2	3	1
Transparency	1	2	2	4
Ease of Use	1	3	3	2
Flexibility	2	3	3	1
Model Updates	1	2	2	3
Robustness	N/A	1	1	2
Parsimony	1	3	2	4

Table 7: Subjective Ranking of Models by Selected Criteria

The table above clearly shows that no single model outperforms the rest in all categories. This is largely unavoidable given the natural trade-off between certain characteristics – for example, adding more parameters will generally improve the fit to trends but reduce parsimony. Similarly, added flexibility can reduce the ease of use.

If all criteria are equally weighted by the user then the CSO model appears to be the most appropriate choice as it perform well against the majority of criteria. The one criteria against which it performs poorly is the ability to reflect mortality trends. For many purposes, users will apply higher weighting to this criteria but when considered in conjunction with the life expectancies set out in Table 5, this criteria is less significant for some purposes (e.g. point in time valuation at retirement).

7. Conclusion

7.1. Summary of findings

Across the models considered in this paper, no single model dominates in terms of performance and all calculated life expectancies lie within a relatively narrow range. The decision on which is most suitable is likely to be driven by the needs and constraints of the individual user.

The CSO model is attractive for users for whom ease of use and implementation are important factors. Although it is not explicitly calibrated to Irish data, the assumptions that define this model are set by the CSO with consideration of overall experience of the Irish population.

The two purely statistical models that were analysed provided similar results, but the Lee-Carter model appears to provide a better fit for females. It is also simpler to use the Lee-Carter model when projecting improvements across a wider age-range. One feature of these models which may be seen as an advantage is the fact they do not rely on subjective assumptions about the long-term rate of improvement. They are also less sensitive to the addition of new data, in comparison to the CMI model.

The CMI model provides the greatest flexibility but may require greater resources and expertise in calibration and updating. Unlike the other models investigated, the CMI model is not publicly available and is restricted to licence holders.

If all criteria are equally weighted by the user then the CSO model appears to be the most appropriate choice as it performs well against the majority of criteria. The one criteria against which it performs poorly is the ability to reflect mortality trends. For many purposes, users will apply higher weighting to this criteria but when considered in conjunction with the life expectancies at key retirement ages this criteria is less significant for some purposes (e.g. point in time valuation at retirement).

8. Appendices

Appendix A – Overview of mortality improvement models

CSO model

In the CSO model mortality rates are projected by estimating the current rate of improvement for each sex and assuming that this rate of improvement will decline over a twenty-five year period to a long-term average improvement rate not dissimilar to the rates observed in the long-term past. It is assumed that there will be no mortality improvements at ages from 100 years and upwards.

In its 2018 analysis[6], the CSO Expert Group judged it reasonable to apply a long-term rate of decline of 1.5% per annum for all ages up to age 90 years from 2041 onwards. For each year between 2016 and 2040, the mortality declines for that year are calculated by linear interpolation.

The approach adopted by the CSO is to assume a gradual decline in mortality improvements from the levels recorded between 2011 and 2015 towards a long-term rate of mortality improvement that would apply from 2041 onwards. The CSO's assumptions use as a starting point the average recorded annual mortality rate improvement from 2011 to 2015, of 2.5% for males and 2% for females. By comparison, the previous set of CSO assumptions, prepared for the 2016 – 2046 population and labour force projections, had been based on initial annual mortality improvement rates of 3% for males and 2.5% for females.

Use of a constant rate of improvement by gender which does not vary by age avoids potential inconsistencies in the projections with, say, age x+1 having lower mortality than age x at some point in the future or, in general, the projections producing a very oddly shaped mortality curve. Justification for the strong smoothing adopted is that life expectancies at ages 0 up to age 65 years showed no significant differences in each future year whether the unsmoothed rates of improvement or the strongly smoothed rates of improvement were used.

The CSO assumes that the short-term mortality improvement rates would apply in 2016 for lives aged up to 90, with 0% mortality improvement for ages of 100 years and over. For ages 91 to 99, the rate of improvement is estimated by linear interpolation between the assumed rate of improvement at age 90 years (2.5% for males and 2% for females) and the zero percent rate of improvement assumed to be applicable at age 100 years.

A long-term average annual improvement rate of 1.5%, to apply from 2041 onwards for both males and females, is assumed for all ages up to 90 years. The projections assume that there is zero long-term mortality improvement from age 100 upwards. Improvements for ages 91 to 99 for 2041 onwards are estimated by linear interpolation between the rate of 1.5% assumed to apply at age 90 and the rate of 0% assumed to apply at age 100.

[6] https://www.cso.ie/en/releasesandpublications/ep/p-plfp/populationandlabourforceprojections2017-2051/mortalityassumptions/

Assumed mortality improvements between 2016 and 2041 are set by the CSO using linear interpolation at each age and sex between the rate applying in 2016 at that age and the long-term rate applying to that same age from 2041 onwards.

Lee-Carter model

This model takes the parametric form

$$\log[m(t,x)] = \alpha(x) + \beta(x)k(t) + \varepsilon$$

with vectors α and β along the age dimension and vector k along the time dimension (or a "period effect"). The vector α can be interpreted as an average age profile of mortality, the vector k tracks mortality changes over time, and the vector β determines how much each age group changes when k(t) changes.

It is probably the most common model in use for mortality modelling.

First published in 1992[7], the Lee-Carter is a one-factor model meaning and therefore improvements in mortality are implicitly assumed to be perfectly correlated at all ages. The model, as originally published, does not project cohort effects although various authors have published extensions to the model which allow it to do so.

It can also result in a lack of smoothness in the estimated age effect.

The main input is the set of underlying data.

[7] Modeling and Forecasting the Time Series of U.S. Mortality (Lee & Carter, Journal of the American Statistical Association 87: 659–671)

CBD model

This model takes the parametric form

$$logit[q(x,t)] = k1(t) + k2(t)(x - \quad)$$

where

- logit(x) = log(x/1-x)
- k_1 which can be interpreted as the ‘level’ of mortality has a downwards trend, reflecting generally improving mortality rates over time
- k_2 the ‘slope’ coefficient has a gradual upwards drift, reflecting the fact that, historically, mortality at high ages has improved at a slower rate than at younger ages.
- $\bar{x}$ is the mean age in the sample range

The CBD model works with a logistic transform of death probabilities or death rates.

It was first published in 2006[8] and relies on the fact that logit[q(x,t)] is typically reasonably linear in x particularly for older age ranges in most populations.

The model was designed for annuities and pensions, not for short-term mortality risk. In particular, it was not designed for lower age ranges (i.e. below age 60). Cairns et al commented that the overall fit of the model is typically not as good as the Lee Carter model and the Lee Carter model is better able to pick up small non-linearities in mortality data.

This is a two-factor model which assumes smoothness across ages in the same projection year but makes no assumption about smoothness in different years. It can be extended to include a cohort effect.

It uses two period-effect parameters to capture the trend improvement in mortality rates (the intercept or level term) and the differential higher age dynamics (slope term).

The main input is the set of underlying data.

[8] A two-factor model for stochastic mortality with parameter uncertainty: Theory and calibration (Cairns et al, Journal of Risk and Insurance 73: 687-718)

CMI model

The CMI model is described extensively in various publications[9].

This model is set up to project a transition from current (or initial) rates of mortality improvement towards a long-term rate of improvement to be reached at some point in the future. The model can be parameterised for bespoke purposes by users, but also offers a default set of parameters, which only require the user to select the assumed long-term rate of improvement (and any margin of prudence or underpin, if required). The default set of parameters have been selected based on the actual experience and model output for the England & Wales population dataset. When using the model for different purposes, as well as using different calibration period lengths and age ranges, it is advised to not simply use the default parameterisation.

The CMI approach is based on an underlying Age-Period-Cohort model blended with a significant weighting towards most recent experience and a view on long-term convergence.

The model projects a convergence from the initial rates of mortality improvement to the long-term rate of mortality improvement, with the pace of convergence varying by age. Age-Period and cohort-effects also converge over different lengths of time.

Figure 15: Convergence Period of the CMI model

The long-term rate of mortality improvement (by default a constant rate selected by the user but there are more complex options) is assumed to apply at ages 20 to 85, with a 0% long-term rate of mortality improvement applying at age 110. For ages between 85 and 110, the long-term rate of mortality improvement is found through linear interpolation between those rates applying at age 85 and at age 110.

The CMI model is restricted to licensee use.

[9] CMI Working papers, including WP 105 and WP 119

Appendix B – Human Mortality Database

The Human Mortality Database (HMD – www.mortality.org) was created to provide detailed mortality and population data to researchers, students, journalists, policy analysts, and others interested in the history of human longevity. It is the work of two teams of researchers in the USA and Germany, with the help of financial backers and scientific collaborators from around the world.

The HMD contains original calculations of death rates and life tables for national populations (countries or areas), as well as the input data used in constructing those tables. The input data consist of death counts from vital statistics, plus census counts, birth counts, and population estimates from various sources.

The main goal of the HMD is to document the longevity revolution of the modern era and to facilitate research into its causes and consequences. As much as possible, four guiding principles are followed in creating this database: comparability, flexibility, accessibility, reproducibility.

The database is highly tractable and provides a relatively clean source for completing projections. For the Irish portfolio, data is available from 1950 to 2014 on births and deaths (by age per year).Further details on the background of the Irish dataset, as well as the methodology followed by the HMD is available at the website referenced above.