Continuous Mortality Investigation

Life Office Mortality Committee

Revised Working Paper 25

Stochastic projection methodologies: Lee-Carter model features, example results and implications

IMPORTANT NOTE:
Working Paper 25 was originally published in April 2007. This revised version was published in November 2007 (together with a revised version of Working Paper 20).

A description of the issues relating to the exposure data and the deviance residuals that gave rise to the amendments in this Working Paper is contained in a separate document entitled "Errata to CMI Working Papers 20, 25 and 27 on Mortality Projections".

This version of Working Paper 25 has NOT been updated for any developments subsequent to its original publication, other than the noted errata.

November 2007

Continuous Mortality Investigation

Life Office Mortality Committee

Working Paper 25

Stochastic projection methodologies: Lee-Carter model features, example results and implications

1. Introduction

1.1. The CMI's Mortality Projections Working Party was established to explore possible projection methodologies for use with the "00" Series Tables. Working Paper 15[1] set out its initial consideration and discussion of projection methodologies and stated that it would consider the P-Spline and Lee-Carter models in detail subsequently.

1.2. The Working Party completed its assessment of the P-Spline model in early 2006 but various computer and technical issues meant that its work on the Lee-Carter model had not been completed. The Working Party recognised that practitioners would need time to become familiar with the model and its features and therefore felt that it could not delay publication of its assessment of the P-Spline model until its work on the Lee-Carter model was completed. It therefore decided to publish separate papers on the P-Spline and Lee-Carter models.

1.3. Working Paper 20[2] was published in April 2006. It provided a summary of the work carried out to date by the Working Party and provided practical advice on using the P-Spline model. It also set out the Working Party's observations on the features of this model, provided example results and discussed their implications. Illustrative software made available at the same time included both P-Spline and Lee-Carter functionality but only the P-Spline functionality was enabled.

1.4. The Working Party has now completed its assessment of the Lee-Carter model and its conclusions are presented in this paper. Simultaneously with the publication of this paper, a revised version of the illustrative software is being made available that includes updated Lee-Carter functionality. All the functionality in the software is now enabled.

1.5. Subsequent sections of this paper provide the background to examples based on the Lee-Carter methodology. Section 2 briefly describes the model and

---

[1] CMI Working Paper 15 (July 2005)
[2] CMI Working Paper 20 (April 2006)

datasets used. Section 3 then describes how to use the output from Lee-Carter functionality in the illustrative software made available by the CMI. Section 4 discusses various features of the Lee-Carter model. Section 5 provides sample annuity values for use in 2004 using data to 2003 and the Lee-Carter model. This section also provides results from back-testing the Lee-Carter model between 1993 and 2004 as more data became available. Section 6 sets out conclusions regarding the Lee-Carter model.

1.6. Whilst engaged in assessing the Lee-Carter model, the Working Party became aware of a paper, published during 2006, by Renshaw and Haberman[3] on extending the Lee-Carter model from an age-period model to an age-period-cohort model. This paper also presented results from applying the extended model to the ONS dataset. The Working Party therefore decided to carry out some initial investigations into the features of this model and these are presented in Section 7.

1.7. Early in its work, the Working Party agreed some high-level objectives desirable of projection models. These formed the criteria against which the Working Party judged the P-Spline and Lee-Carter models. Section 8 discusses these objectives and assesses the P-Spline, Lee-Carter and the Lee-Carter age-period-cohort models against them. It then concludes the paper with a summary of the Working Party's views on these models.

1.8. This Working Paper has been prepared for the Life Office Mortality Committee of the CMI by a Working Party consisting of Angus Macdonald, Adrian Gallop, Keith Miller, Stephen Richards, Rajeev Shah and Richard Willets. It has been approved by the Committee.

2. Models and datasets considered

2.1. In this paper we show results using the Lee-Carter model based on the same datasets used in Working Paper 20: CMI assured lives data for males from 1947 to 2003 and the ONS England and Wales population data from 1961 to 2003. The age ranges of the datasets used for modelling are consistent with those used for Working Paper 20 on the P-Spline models. The choice of the age ranges used is discussed in Working Paper 20 and so will not be repeated here.

2.2. The ONS datasets used by the CMI are not yet publicly available and were only made available to the CMI for research purposes. Therefore, the CMI can only show the results from modelling these datasets and cannot provide the datasets to others.

2.3. Working Paper 15 described the Lee-Carter model and the process of fitting the model to data. As described in Working Paper 15, the deviance residuals used

[3] Renshaw A. E. and Haberman S. (2006) A cohort-based extension to the Lee-Carter model for mortality reduction factors. **Insurance: Mathematics and Economics** 38 (3), 556-570.

in the boot-strapping process were generated by fitting various models to the actual data.

3. How to use the model output

3.1. The deviance residuals from the fits of three models (Lee-Carter, P-Spline age-cohort and P-Spline age-period) and the original dataset are used in the bootstrapping process to produce a synthetic dataset. This process is the basis for each simulation. For the purpose of this paper 1,000 simulations were generated. It did not appear that increasing the number of simulations would materially affect the results.

3.2. The Lee-Carter model is then fitted to each synthetic dataset. As outlined in Working Paper 15, each fit produces values for the alpha parameters, which describe the average level of the log $\mu_{x,t}$ surface over time, the kappa parameters, which reflect the change in overall mortality over time, and the beta parameters, which describe the pattern of deviations from the age profile as the kappa parameter varies. If the beta parameters are particularly high for some ages, it means that mortality rates change faster at these ages than in general. Negative values for beta parameters indicate worsening mortality at those ages. It is worth noting that although we refer to $\mu_{x,t}$ this is only relevant to the CMI datasets, for the ONS datasets $\mu_{x+½,t}$ are produced.

3.3. The kappa parameters are treated as a time series and can be projected. Consistent with our work on the P-Spline models, we projected the kappa parameters for 100 years for all our calculations. The fitted and projected kappa parameters, together with the fitted alpha and beta parameters were used to estimate values of $\mu_{x,t}$ for both the region of the original dataset and the region of the projection.

3.4. Each set of $\mu_{x,t}$ related to a particular sample path. As explained in Working Paper 15, the Lee-Carter model produces sample paths rather than percentiles. Where the Lee-Carter model is being compared to the results from other models that produce percentiles, care is required to ensure that sample paths and percentiles are considered consistently.

3.5. The estimated values of $\mu_{x,t}$ were used to estimate mortality improvements which were then used with a base table of mortality rates to calculate annuities. This approach is described in greater detail in paragraphs 6.7 and 6.8 of Working Paper 20.

3.6. This approach to calculating annuity values allows only for uncertainty relating to projected mortality improvements. We have made no allowance for stochastic variability arising from the size of a portfolio to be valued or for any heterogeneity within a portfolio. Working Paper 15 discussed the additional uncertainty from these sources and how it could be reflected in the annuity calculations.

3.7. The average annuities were estimated for each age and year. The confidence intervals for each age and year were estimated by ordering the annuity values for each sample path.

4. Lee-Carter model features

Minimum data requirements

4.1. For any projection model, the quality and size of the dataset fitted will be crucial factors affecting the reliability of the fit. In our view, the minimum data requirements for the Lee-Carter model are similar to those for the P-Spline model described in Working Paper 20.

4.2. However, data covering a shorter period than 20 calendar years could be used with the Lee-Carter model if a narrower age range is being fitted. In general, the larger the age range covered, the higher the number of calendar years' data required.

4.3. As the Lee-Carter model does not specify unique choices for the parameters, widening the age range without increasing the number of calendar years covered by the data will result in a greater proportion of the volatility being explained by the model as the age-specific variations in the period effects. Therefore, the extent of smoothing carried out by the model would be reduced, making it more difficult to capture and project the true underlying period effects.

Choice of model for producing deviance residuals

4.4. As explained in Working Paper 15, the Lee-Carter fitting process does not take account of the uncertainty relating to the parameter estimates. Therefore, deviance residuals are used in a bootstrapping process to make some allowance for this uncertainty. The best estimate is the 50^{th} percentile of the simulations produced using the bootstrapping process.

4.5. We investigated the use of deviance residuals produced using 3 different models – Lee-Carter, P-Spline age-cohort and P-Spline age-period. As shown in Table M1 below, the choice of model used to produce deviance residuals does not appear to significantly affect the best estimates in results from the simulations.

Time series parameters

4.6. Projecting mortality rates using the Lee-Carter model requires fitting a time-series to the kappa parameters and then projecting this time-series forwards using ARIMA processes. The ARIMA process used depends on the shape of the kappa parameters in the region of the data.

4.7. For all the Lee-Carter model results presented in this paper, we have fitted an ARIMA(1,1,0)[4] model to the kappa parameters in the region of the data and projected these forwards.

Projecting cohort features

4.8. Using results from simulations produced by the Lee-Carter model, it is difficult to analyse features present in the projected improvements. Instead, the process we have used to analyse some of these features was to:

- Fit the Lee-Carter model to the base data and estimate the alpha, kappa and beta parameters in the region of the data;
- Fit a time-series to the kappa parameters in the region of the data;
- Project the kappa parameters using the time-series model but without the stochastic error (i.e. use the expected values from the time series projection);
- Estimate the mortality rates in the region of the data and projection using the fitted and projected kappa parameters, together with the fitted alpha and beta parameters;
- Analyse the improvement rates indicated by these estimated mortality rates; and
- Analyse the deviance residuals of the fitted mortality rates.

4.9. The results of these projections are illustrated in the cohort maps shown in Appendix B and the deviance residuals are shown in Appendix D.

4.10. Based on the cohort maps in Appendix B, the Lee-Carter model does not seem to smooth out volatility in mortality rates between calendar years to the same extent as the P-Spline model. The volatility remaining in the mortality rates fitted makes it difficult to identify trends, such as cohort features, in the region of the data.

4.11. In the region of the projection, no cohort features can be seen even though such features have been widely recognised in the CMI and ONS datasets. Given the extent of volatility left unsmoothed in the region of the data, we consider that it would be very difficult for the Lee-Carter model to recognise any cohort features in the region of the data. Therefore, it may be of little surprise that cohort features are not present in the region of the projection.

[4] An Auto-Regressive Integrated Moving Average model. An ARIMA (p, d, q) model is specified by the order of the autoregressive (p), integrated (d) and moving average (q) components.

4.12. The deviance residuals for the ONS datasets in Appendix D show that there are clear patterns when analysed by cohort, with high positive deviances for the generation born around 1926. The Assured Lives dataset shows a similar pattern though not as clear. This indicates that the Lee-Carter model is not a good fit for these datasets.

4.13. The deviance residuals analysed by period show a saw tooth effect. This indicates that there may be some auto-regressive component in the period effects (i.e. the kappa parameters) that is not captured by the Lee-Carter model. A possible explanation for the auto-regressive features could be the impact of severe winters in some years resulting in high mortality, followed by milder winters where the mortality will be light.

5. What if these models had been used previously

Lee-Carter Results

5.1. Table M1 gives some annuity values based on the Lee-Carter model using the CMI dataset for 1947 to 2003 and the ONS dataset for 1961 to 2003. Results are provided for all three of the methods for calculating deviance residuals for each of the datasets. The age ranges and the time-series models fitted are described in Appendix A.

5.2. For each simulation, mortality improvement factors were estimated using the process described in section 3 above. A two-way table of q_x was then produced by applying these improvement factors to adjusted “92” Series base tables of q_x. The “92” Series tables provide a set of graduated q_x, derived from data for 1991-94, which are assumed to be applicable to 1992. The table of q_x was adjusted by age-banded 100A/Es[5], for the Life Office Pensioner experience, for 2003, to allow for actual mortality improvements up to 2003. (Note: this also has the effect of removing some of the smoothing implicit in the graduated base table.) Values for annuity-due ($\ddot{a}_x$) at the ages shown were then calculated as at 2004 using a 4.5% p.a. interest rate.

5.3. For each set of 1,000 simulations, the annuity values were ordered by age and year of use and the best estimates and confidence intervals for annuity values derived.

5.4. For comparison, the first three rows in the table show annuity values for year of use, u=2004 using the “92” Series base table updated to 2003 by applying 100A/Es for 2003 and applying projected improvements using the long, medium and short cohort projections from 2004. As these projections are based on data only to 2000, they are not strictly comparable with the annuity values described in the previous paragraph as they were calculated using projections

[5] See the CMI note on *Annuity values presented at CILA on 28 September 2005*

based on data to 2003. However, for the purpose of this paper this difference is not material.

5.5. Table M2 shows the annuity values described as a percentage of the annuity values calculated on the "92" Series basis with medium cohort adjustment for experience to 2003.

Table M1

	Projection based on male assured lives, 1947-2003			Projection based on male ONS data, 1961-2003		
	4.5% annuity value at age			4.5% annuity value at age		
Mortality basis	60	65	75	60	65	75
PMA92u04mc	15.218	13.640	9.902	15.218	13.640	9.902
PMA92u04lc	15.620	14.154	10.355	15.620	14.154	10.355
PMA92u04sc	15.044	13.415	9.599	15.044	13.415	9.599
PMA92u04LC50psac	15.056	13.411	9.549	15.001	13.349	9.481
PMA92u04LC97.5psac	15.338	13.685	9.775	15.261	13.598	9.661
PMA92u04LC2.5psac	14.765	13.136	9.321	14.750	13.113	9.296
PMA92u04LC50psap	15.057	13.412	9.550	14.998	13.347	9.480
PMA92u04LC97.5psap	15.334	13.686	9.772	15.257	13.588	9.656
PMA92u04LC2.5psap	14.771	13.140	9.322	14.747	13.115	9.297
PMA92u04LC50LC	15.059	13.414	9.552	14.993	13.343	9.478
PMA92u04LC97.5LC	15.333	13.688	9.782	15.254	13.598	9.654
PMA92u04LC2.5LC	14.763	13.134	9.320	14.752	13.116	9.303

Table M2

Annuity values in shaded cells. Other cells show the values in Table M1 as percentages of values in the shaded cells.						
	Projection based on male assured lives, 1947-2003			Projection based on male ONS data, 1961-2003		
	4.5% annuity value at age			4.5% annuity value at age		
Mortality basis	60	65	75	60	65	75
PMA92u04mc	15.218	13.640	9.902	15.218	13.640	9.902
PMA92u04lc	102.6%	103.8%	104.6%	102.6%	103.8%	104.6%
PMA92u04sc	98.9%	98.3%	96.9%	98.9%	98.3%	96.9%
PMA92u04LC50psac	98.9%	98.3%	96.4%	98.6%	97.9%	95.7%
PMA92u04LC97.5psac	100.8%	100.3%	98.7%	100.3%	99.7%	97.6%
PMA92u04LC2.5psac	97.0%	96.3%	94.1%	96.9%	96.1%	93.9%
PMA92u04LC50psap	98.9%	98.3%	96.4%	98.6%	97.8%	95.7%
PMA92u04LC97.5psap	100.8%	100.3%	98.7%	100.3%	99.6%	97.5%
PMA92u04LC2.5psap	97.1%	96.3%	94.1%	96.9%	96.1%	93.9%
PMA92u04LC50LC	99.0%	98.3%	96.5%	98.5%	97.8%	95.7%
PMA92u04LC97.5LC	100.8%	100.3%	98.8%	100.2%	99.7%	97.5%
PMA92u04LC2.5LC	97.0%	96.3%	94.1%	96.9%	96.2%	94.0%

5.6. The naming convention for the mortality bases used in these tables is similar to the conventions previously used by the CMI. The first part of the name (i.e. PMA92) refers to the base mortality table ("92" Series, pensioners, males, amounts). "u04" specifies that the calculation is done using the set of q_x for lives aged 60, 65 or 70 in 2004 and following them as they age through successive calendar years to the end of the table i.e. following diagonals for particular years of birth. The next part of the basis name ("LC") refers to the Lee-Carter projection method and the number appended to that represents the related confidence interval. Lastly, the letters "psac", "psap" or "LC" have been added to denote the type of model, P-Spline age-cohort, P-Spline age-period or Lee-Carter, used to derive the deviance residuals that were used to generate the synthetic dataset.

5.7. The 50th percentile annuity values for the assured lives dataset calculated using the deviance residuals from all three models are very similar. The widths of the confidence intervals are similar for the annuities based on the P-Spline age-cohort and P-Spline age-period deviance residuals. However, the confidence intervals for the annuities based on the Lee-Carter deviance residuals are slightly wider. This may be a spurious feature due to the small number of simulations run.

5.8. The results derived from the ONS data are similar but the confidence intervals are narrower than those for the assured lives dataset when using the Lee-Carter deviance residuals. The narrower confidence intervals may partly be due to the greater number of deaths and exposure in the ONS data.

5.9. For both the CMI and ONS datasets, the 50th percentile annuity values are lower than those produced using the Medium Cohort projections. Even at the 97.5th percentile, half of the annuity values using the Lee-Carter model are lower than annuity values producing using the Medium Cohort projections.

Comparison with Results in Working Paper 20

5.10. Comparing these results to those published in Working Paper 20, the annuity values for both the CMI and ONS datasets using the Lee-Carter model are considerably lower than those based on the P-Spline model. Further, the 50th percentile annuity values produced by the Lee-Carter model are generally lower than those calculated on the "92" Series basis with short cohort adjustment.

5.11. Compared to the P-Spline projections, the confidence intervals for the assured lives data are generally slightly narrower but for the ONS data they are much narrower.

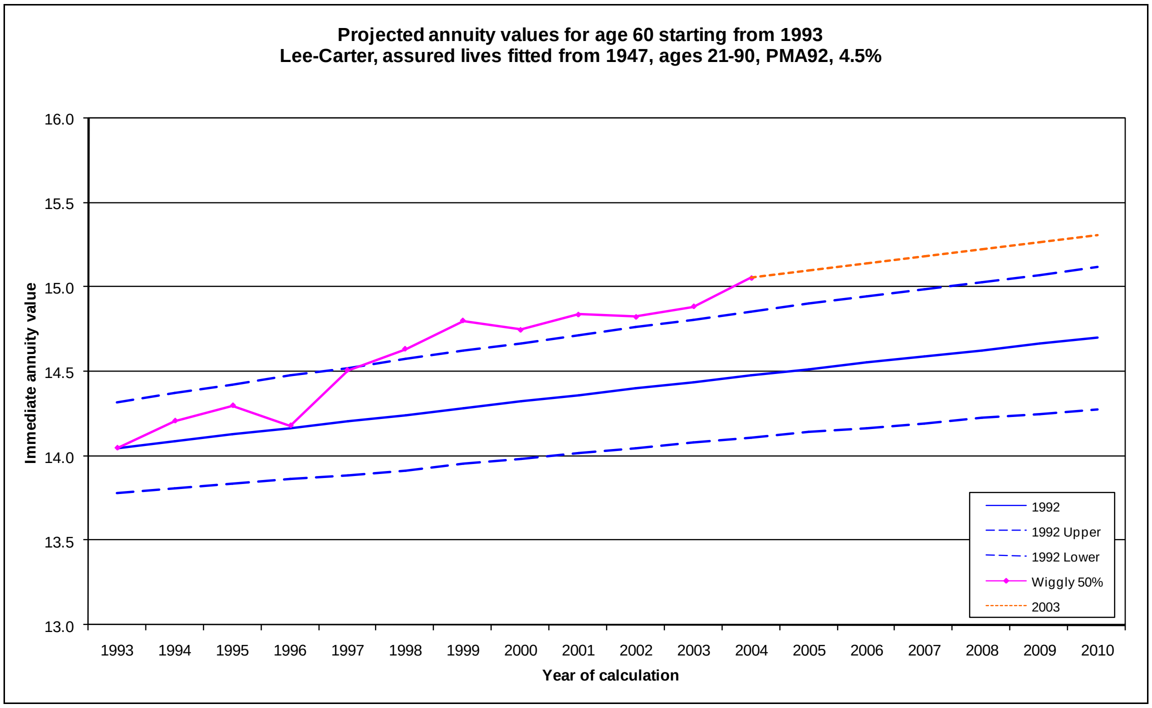
Back-testing

5.12. An obvious way to test a projection model is to carry out back-testing i.e. to consider what results would have been produced if the model had been used in the past. We used this approach to test the P-Spline model and have used it again to test the Lee-Carter model.

5.13. Using this approach, we can rebase the projections and calculate annuity values from 1993 onwards by adding new data for succeeding years. For example, in 1997, assuming data to 1996 is available, the table of base q_x can be adjusted to update it to 1996 by applying appropriate 100A/Es. The Lee-Carter model could also be refitted to that data and new annuity values can then calculated for use in 1997. In this way the driver of the development of the mortality element of the annuity basis over successive years is the availability of the latest mortality data rather than the infrequent production of new base tables or adjustments to older tables.

5.14. Figure 1 shows the 50th percentile annuity values using the Lee-Carter model. The deviance residuals arising from fitting a P-Spline age-cohort model to assured lives data have been used. The "wiggly line" in the graph shows how an annuity value for a 60 year old male in each year would have progressed over the period 1993 to 2004 as additional data became available and the annuity values were recalculated taking account of each year of additional data. The recalculation involves the replacement of the projected mortality rates by the actual rates experienced for the year for which additional data became available and the recalibration of projected future improvements. The values shown on the "wiggly line" for 2004 are those described in the table above. From the end of the "wiggly line", in 2004, the dotted line shows the projection of future annuity due values, based on all data to 2003. The central blue line is the projected annuity values based only on the data as at 1992 (i.e. the base tables and the projections for the annuities in the line are not adjusted for actual experience in later years as it becomes available), starting the calculation in 1993. The two 'dashed' lines show projected annuity values based on the 1992 data but calculated from the 2.5th and 97.5th percentiles respectively. That is, they represent the confidence interval applicable to the 1992 based annuity values.

Figure 1

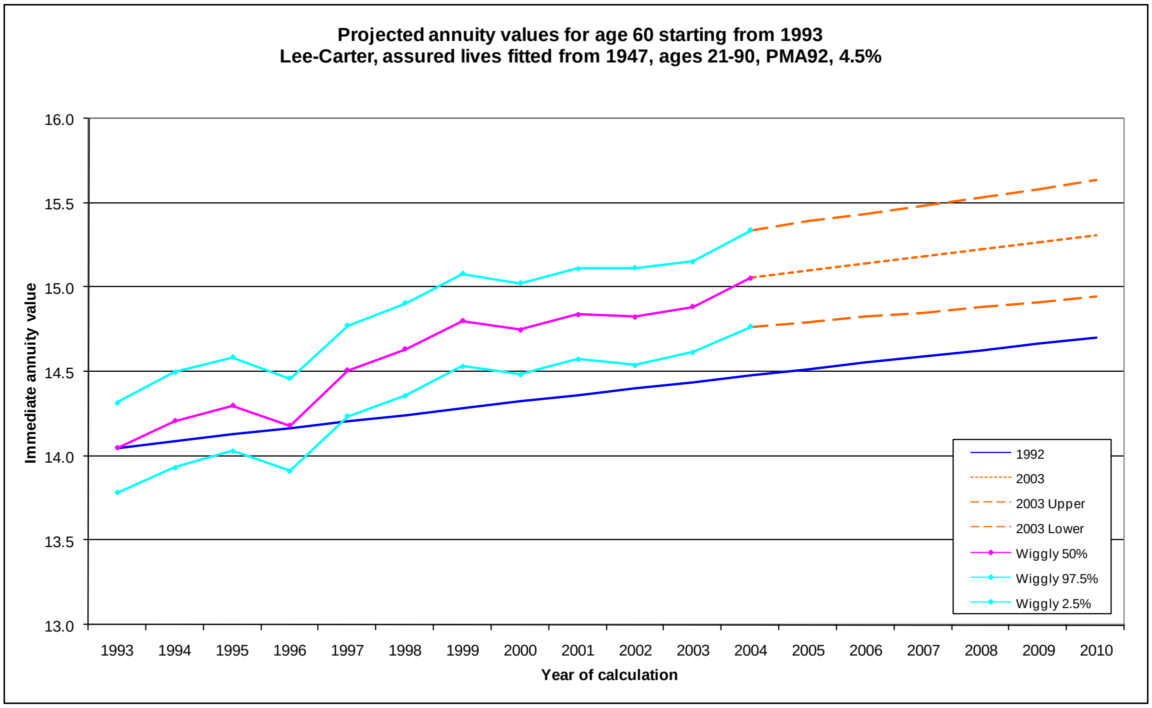


5.15. Figure 1 shows that using data to 1992, the Lee-Carter model would have produced projections that quickly became outdated. Revised projections using a further 5 years data to 1997, would have been outside the 95% confidence interval of the projections using data to 1992. We consider that this poor "predictive power" of the Lee-Carter model is due, at least in part, to its inability to recognise cohort effects at an early stage.

5.16. The results presented in Figure 1 above are directly comparable to Figure 5 in Working Paper 20 which presented the results of similar back-testing of the P-Spline model. This comparison shows that the "predictive power" of the Lee-Carter model does not appear to be as good as the P-Spline model for the period in question and using the Assured Lives dataset.

5.17. Figure 2 plots the changes in the confidence intervals (2.5th and 97.5th) over this period. It is comparable to Figure 6 in Working Paper 20.

Figure 2



5.18. Figures 3 to 5 show a comparison of annuity values based on the Lee-Carter model, using the deviance residuals arising from fitting the P-Spline age-cohort model to assured lives data, against the interim cohort projections (from 2000 onwards), for males at ages 60, 65 and 75. In each case, the projections have been updated using 100 A/Es for each year up to 2004. The “wiggly” lines are as per those in the Figure 2.

5.19. The results presented in Figures 3 to 5 below are directly comparable to Figures 9 to 11 shown in Working Paper 20 relating to the P-Spline model.

Figure 3

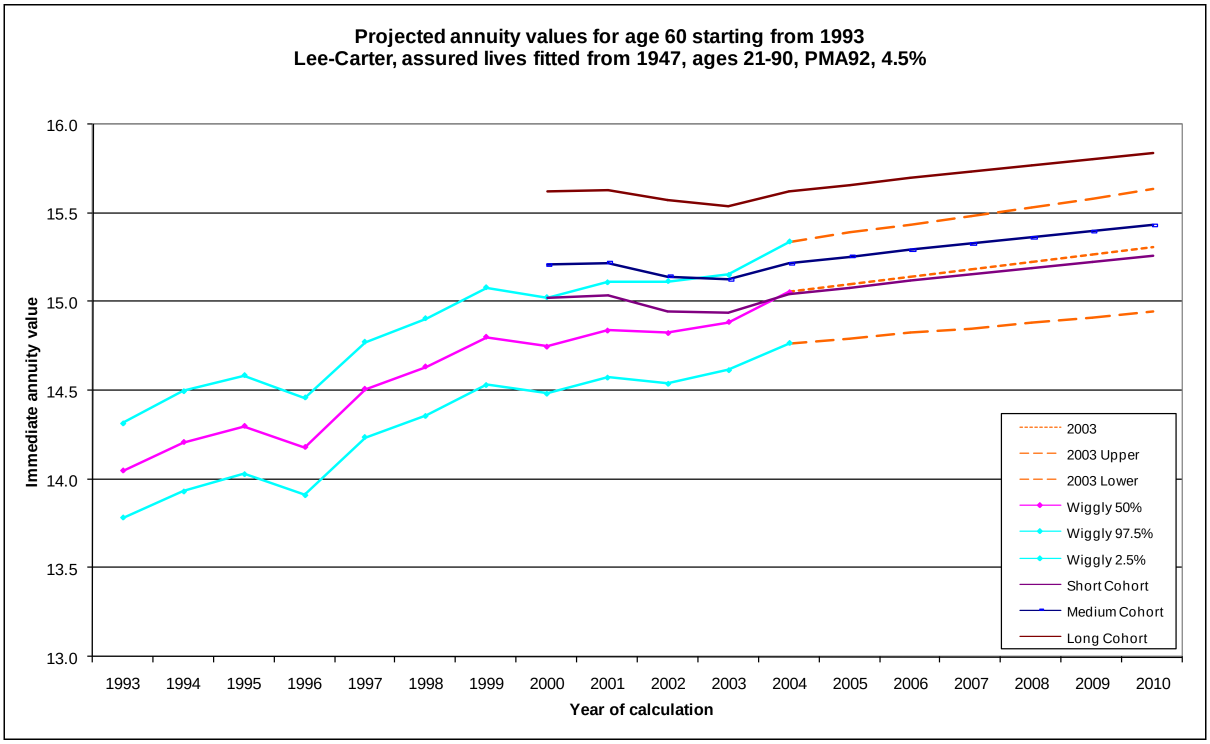


Figure 4

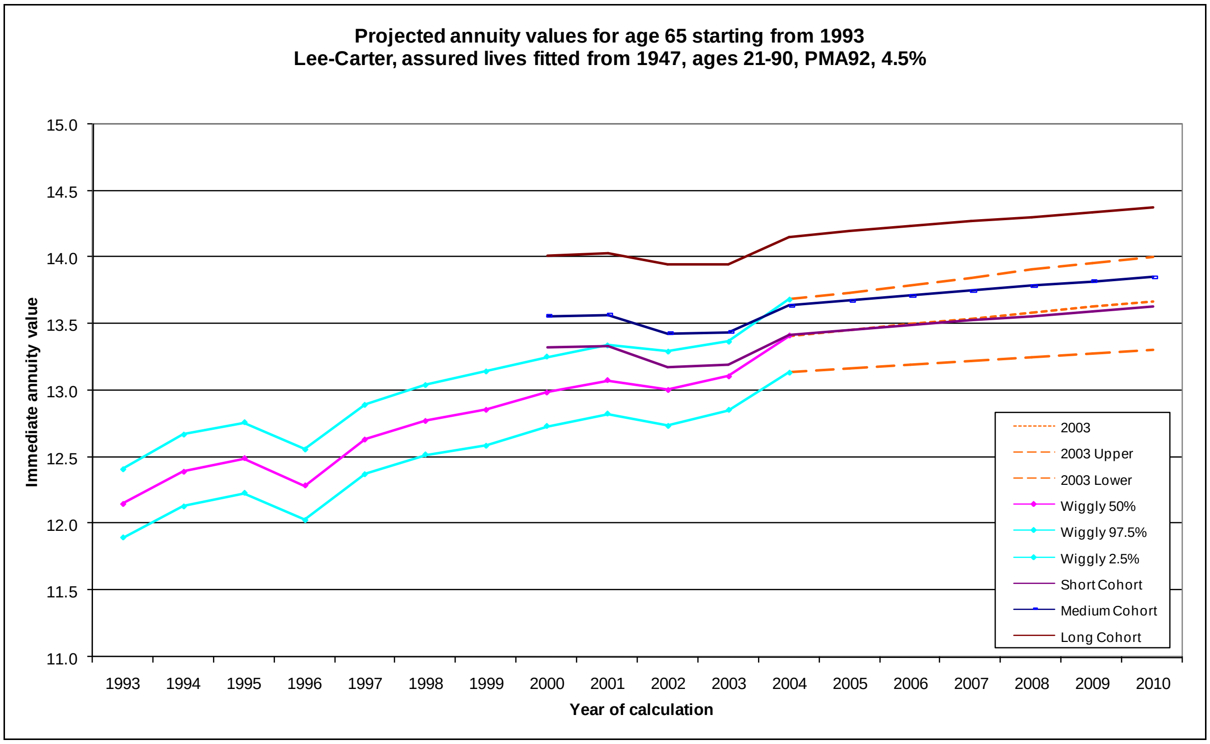
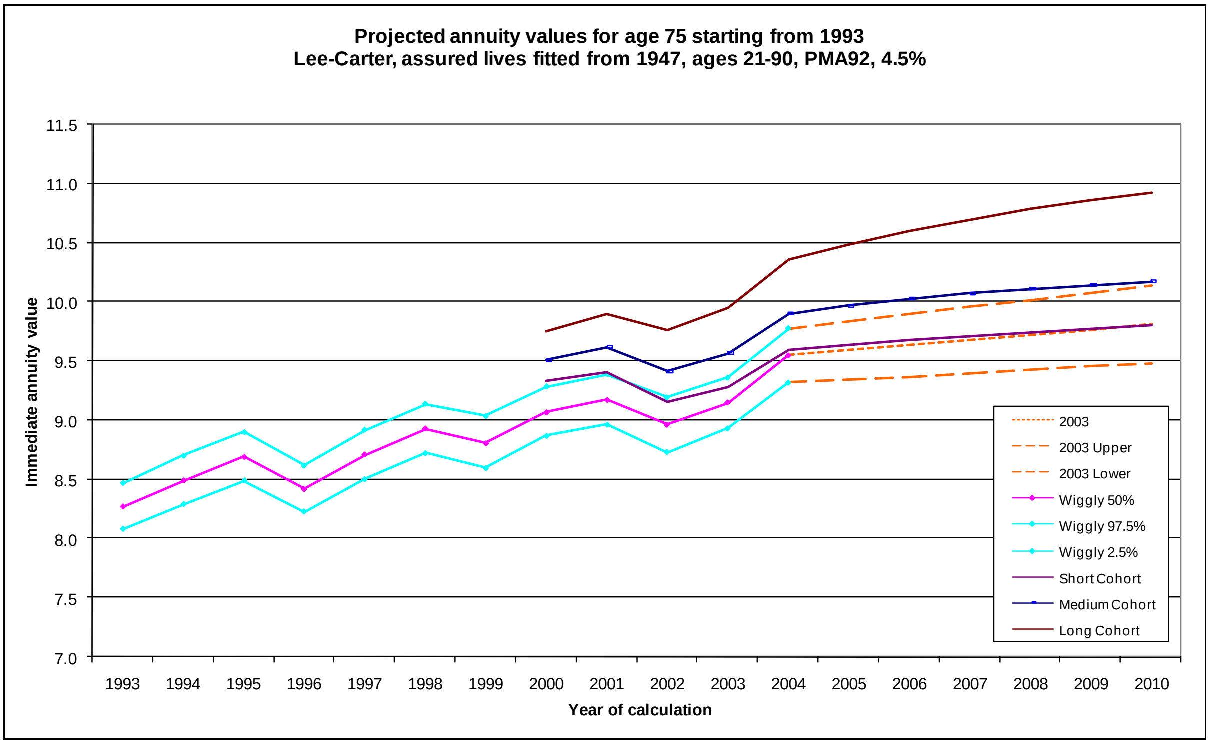


Figure 5



5.20. Figures 3 to 5 show that the annuity values for the long and medium cohort projections are greater than the 50th percentile Lee-Carter projection for all years. Annuity values for the short cohort projections are greater than the 50th percentile Lee-Carter projection prior to 2004 but from 2004 the annuity values become very similar.

6. Conclusions regarding the Lee-Carter model

6.1. The back-testing described in Section 5 demonstrates that projections using the Lee-Carter model, based on data to 1992, would not have worked well in recent years. Whilst this does not mean that the model would not work well in the future either, there are particular features that make this model unsuitable for projecting the CMI and ONS datasets.

6.2. The analyses in Section 4 indicated that the Lee-Carter model does not seem to be able to project forward any cohort effects. Both the CMI and ONS datasets have clear cohort effects that have been widely recognised for some time. Clearly, if users view cohort features as important and consider that they will continue for some time, the Lee-Carter model will not be satisfactory.

6.3. The analyses also show that the Lee-Carter model does not sufficiently smooth out the volatility in mortality rates between calendar years. This makes it very difficult to use results from the Lee-Carter model to identify features in the region of the data.

7. The Lee-Carter age-period-cohort model

7.1. Around the time the Working Party was reaching its conclusions regarding the Lee-Carter model, it became aware of the paper by Renshaw and Haberman that extended the Lee-Carter model from an age-period model to an age-period-cohort model (the Lee-Carter APC model). Given the conclusions the Working Party were reaching regarding the Lee-Carter model, we considered it would be worthwhile exploring the Lee-Carter APC model. However, given the time constraints, we have only been able to carry out an initial exploration of this model.

Model structure

7.2. The Lee-Carter APC model is a bilinear model in the variables x (age), t (period) and c (cohort) of the following form:

$$\log \mu(x, t, c) = a(x) + b_1(x)k(t) + b_2(x)I(c) + z(x, t, c)$$

where $\mu(x, t, c)$ is the force of mortality at age x in year t for generation c and $z(x, t, c)$ is a random error term. The $a(x)$ coefficients describe the average level of the $\log \mu(x, t, c)$ surface over time. The $b_1(x)$ and $b_2(x)$ coefficients describe the pattern of deviations from the age profile as $k(t)$ and $I(c)$ respectively vary.

The $k(t)$ parameter describes the change in overall mortality over time while $I(c)$ describes the change in mortality between generations.

7.3. The model does not specify unique choices for the parameters because $b_1(x)$ and $k(t)$ along with $b_2(x)$ and $I(c)$ only appear through their products $b_1(x)k(t)$ and $b_2(x)I(c)$. Further, there is a linear relationship between x, t and c ($c = t - x$). Therefore, constraints have to be applied to the fitted parameters in order to produce unique solutions. Renshaw and Haberman have suggested the following constraints:

$$\sum_x b_1(x) = 1, \ \sum_x b_2(x) = 1 \text{ and either } I(t_1 - x_k) = 0 \text{ (or } k(t_1) = 0)$$

7.4. The model can be fitted using standard likelihood methods, assuming a Poisson model for the numbers of deaths at each age and calendar year for each generation. Projected mortality rates are produced by fitting time-series models to the fitted $k(t)$ and $I(c)$ parameters and then projecting this time-series forward using ARIMA processes.

7.5. We have used univariate time-series models, effectively assuming independence between the fitted $k(t)$ and $I(c)$ parameters, so that we could reproduce the results presented in the Renshaw and Haberman paper. While this assumption seemed to work well for the ONS data, multivariate time-series models may be necessary for other datasets.

Sample results

7.6. Table M3 gives some annuity values based on the Lee-Carter APC model using the CMI dataset for 1947 to 2003 and the ONS dataset for 1961 to 2003. The approach used is similar to that used to calculate annuity values using the Lee-Carter model as shown in Table M1. However, due to time constraints, only deviance residuals from the Lee-Carter APC model were used to generate the simulations. The Lee-Carter APC model was fitted using the iterative process described in the Renshaw and Haberman paper.

7.7. The age ranges and the time-series models fitted are described in Appendix A. The age ranges for the CMI and ONS datasets differ from those used for the Lee-Carter model as we had difficulty in getting the model to converge, particularly when the younger ages were included. The starting parameter values used in the iterative fitting process also differ between the datasets as we were unable to find a single set of starting parameters that enabled the model to converge for both datasets and for different age ranges.

7.8. Table M4 then shows the annuity values given in Table M3 described as a percentage of the annuity values calculated on the “92” Series basis with medium cohort adjustment and adjusted for experience to 2003.

Table M3

	Projection based on male assured lives, 1947-2003			Projection based on male ONS data, 1961-2003		
	4.5% annuity value at age			4.5% annuity value at age		
Mortality basis	60	65	75	60	65	75
PMA92u04mc	15.218	13.640	9.902	15.218	13.640	9.902
PMA92u04lc	15.620	14.154	10.355	15.620	14.154	10.355
PMA92u04sc	15.044	13.415	9.599	15.044	13.415	9.599
PMA92u04LCapc50	15.126	13.497	9.598	16.348	14.575	10.217
PMA92u04LCapc97.5	15.557	13.899	9.908	17.151	15.386	10.759
PMA92u04LCapc2.5	14.675	13.081	9.296	15.640	13.907	9.789

Table M4

Annuity values in shaded cells. Other cells show the values in Table M3 as percentages of values in the shaded cells.						
	Projection based on male assured lives, 1947-2003			Projection based on male ONS data, 1961-2003		
	4.5% annuity value at age			4.5% annuity value at age		
Mortality basis	60	65	75	60	65	75
PMA92u04mc	15.218	13.640	9.902	15.218	13.640	9.902
PMA92u04lc	102.6%	103.8%	104.6%	102.6%	103.8%	104.6%
PMA92u04sc	98.9%	98.3%	96.9%	98.9%	98.3%	96.9%
PMA92u04LCapc50	99.4%	99.0%	96.9%	107.4%	106.8%	103.2%
PMA92u04LPapc97.5	102.2%	101.9%	100.1%	112.7%	112.8%	108.7%
PMA92u04LCapc2.5	96.4%	95.9%	93.9%	102.8%	102.0%	98.9%

7.9. Using the CMI dataset, the 50th percentile annuity values using the Lee-Carter APC model are lower than those produced using the Medium Cohort projections. Compared to the Lee-Carter model, the 50th percentile values using the Lee-Carter APC model are higher and the confidence intervals wider.

7.10. Using the ONS dataset, the 50th percentile annuity values using the Lee-Carter APC model are markedly higher than those produced using the Medium Cohort projections. They are also higher than the annuity values produced using the P-Spline age-cohort projections shown in Table M4 of Working Paper 20 though the confidence intervals are comparable.

Projecting cohort features

7.11. Using a similar process to that described in paragraph 4.8, the features present in the projected improvements produced by the Lee-Carter APC model were analysed. The results of these projections are illustrated in the cohort maps shown in Appendix C and the deviance residuals are shown in Appendix E. The fitted parameter values are shown in Appendix F.

7.12. Though some volatility remains unsmoothed, the cohort maps show that compared to the Lee-Carter model, the Lee-Carter APC model is better at smoothing the volatility in mortality rates between calendar years. Features, including cohort effects, are also more easily identified in the region of the data. Cohort features are also projected and appear to be consistent with the cohort features identified in the region of the data.

7.13. A feature of the model is that there seems to be significant volatility in the fitted mortality rates between succeeding generations. It is difficult to say whether this volatility is a genuine feature of the data or whether the model applies insufficient smoothing between generations.

7.14. In the region of the projection, cohort effects are only projected for generations that exist in the dataset. The model seems to assume that mortality rates for younger generations not included in the dataset will not have any cohort features. Therefore, there is not a smooth transition in the projected mortality rates between generations included in the dataset and younger generations. However, this would only be an issue if the user was interested in the projected mortality rates for these younger generations.

7.15. Based on the deviance residuals in Appendix E, the Lee-Carter APC model seems to capture the cohort effects much better than the Lee-Carter model. The deviance residuals for the Assured Lives dataset seem randomly dispersed when analysed by cohort.

7.16. When analysed by period, the deviance residuals continue to exhibit a saw tooth effect, similar to that seen for the Lee-Carter model. This indicates that there may be some auto-regressive component in the period effects that is not captured by the Lee-Carter APC model.

7.17. The fitted parameter values for the ONS and Assured Lives dataset in Appendix F show that the cohort parameters ($I(c)$) seem to capture well the cohort effects for the generation born between 1922 and 1946. Due to significant volatility being retained in the fitted rates, it is not possible to identify this feature in the cohort map shown in Appendix C.

7.18. Using the same constraints used to create the graphs in Appendix F, the fitted cohort parameters appear to be stable as the age range fitted is changed. Figure G1 in Appendix G shows the fitted parameters for the Assured Lives dataset with the age range extended to cover ages 25-90. However this set of constraints did not work for all age and year ranges. Using alternative constraints that fit some of these subsets of data Figure G2 shows that this stability is affected. Using this alternative set of constraints illustrates that the cohort parameters do not appear to capture the cohort effects for the generation born between 1922 and 1946. For this fit, some of these cohort effects may have been captured in the coefficients describing the pattern of deviations from the age profile ($b_2(x)$) as the cohort parameters vary.

Further assessment of the Lee-Carter APC model

7.19. The Working Party considered that some of the theoretical features of the model needed further examination. However, this work was not undertaken due to the time constraints and the resources necessary for such research work. The time constraints also meant that we could only carry out an initial assessment of the Lee-Carter APC model. This meant that we did not manage to resolve all the computing issues that arose during our testing of this model. The computing issues and theoretical features are discussed below.

7.20. While we managed to fit the full ONS dataset to 2003 and reproduce the results presented in the Renshaw and Haberman paper, we ran into computing issues when carrying out back-testing. We were unable to get the model to converge for certain subsets of the dataset (e.g. using data from 1961 to 1999). We also had more difficulty in fitting the model to subsets of the CMI dataset.

7.21. On further analysis, we found that we could generally get the model to converge by changing the starting parameter values used in the iterative fitting process. However, when back-testing a dataset or fitting a different age range, we were unable to find a set of starting parameter values that consistently worked for the different subsets of the data. Where a number of sets of starting parameter values worked for a particular dataset, we also found that the fitted values could differ materially.

7.22. Due to the linear dependence between the age, period and cohort parameters, constraints have to be placed on the fitted parameters in order to get unique solutions. Unique solutions are necessary to produce fits that remain relatively stable as additional calendar years' data become available. The convergence problems highlighted by our testing indicate that the suggested constraints may not fully resolve this linear dependence issue and we consider further work is necessary in this regard.

7.23. A key assumption underlying the model is that the time series fitted to the period and cohort parameters can be projected independently. For the ONS dataset over the period 1961 to 2003, this assumption seemed to work well. However, we are not convinced that this assumption would work well for all datasets.

7.24. Though the Lee-Carter APC model extends the structure of the original Lee-Carter model, it is effectively a new model as it seeks to resolve the linear dependencies between the age, period and cohort parameters. The Working Party's overall view was that the Lee-Carter APC model was interesting and may have resolved some of the disadvantages of the Lee-Carter model. However, we consider that it would benefit from greater testing.

8. High-level objectives for mortality projection models

8.1. We believe that both the P-Spline and Lee-Carter models have particular features that make them suitable for certain purposes but consider that neither of these models meets all the desirable objectives for projection models. Therefore, we are not able to recommend any particular projection model for use with the CMI "00" Series tables.

8.2. However, we are also aware that there are a number of other projection models available, both public and proprietary, but we did not have the resources to investigate them all. Therefore, we consider that actuaries may find it helpful if we set out the objectives we consider are desirable for projection models:

- **Ease of use** – An obviously desirable feature for any model is the ease of using the model. An easy to use model should also generally be easier to understand and explain to others. This would also allow the reasons for the main features of any fit or projection to be explained to and better understood by non-experts.

- **Ability to interpret the parameters** – In addition to being easy to understand and explain, another desirable feature for a model is whether the parameters can also be interpreted in a way that can be easily understood and assessed by users. This would allow the fitted parameters to be considered for reasonability and features explained to non-experts.

- **Model structure and fit** – As noted in Working Paper 3[6], a key requirement of a projection model is that it should be sensible in the region of the data, producing fits that represent the data well given the usual requirements of parsimony and adherence to the data. This usually requires a trade-off between smoothness and goodness-of-fit. The model should also produce a smooth transition between the region of the data and the region of the projection.

  As well as providing good fits for a range of datasets, the model should produce fitted values that are stable so that the availability of another year's data does not result in an entirely different fit to the data. This means that the projection should not give undue weight to the final year or the extreme ages in the dataset.

- **Cohort effects** – Where cohort effects are known to exist in the dataset, the model should be able to reflect these effects in the projection. However, if users are sure that cohort effects do not exist or are not significant, this objective is not as relevant.

- **Best estimate** – A key requirement of a stochastic projection model is that it should produce best estimate projections that are reasonable and plausible. The projections should be consistent with the recent past and take account of as many relevant trends as possible.

- **Confidence intervals** – Uncertainty regarding future mortality rates can be better illustrated, helping understanding of the key risk, if probabilistic statements can be associated with the projections. Such information can also inform businesses' risk appetites and financial controls.

- **Ability to generate sample paths** – Most asset models generate sample paths and so it is desirable that mortality projection models

[6] CMI Working Paper 3 (March 2004), Section 4.2

used in conjunction with such asset models also produce sample paths. For longevity business, if the model can produce percentiles, this may be a reasonable substitute if suitably calibrated. However, percentiles will not be reasonable substitutes for modelling business, such as term assurance, where the volatility in mortality rates between calendar years can be much more important than long-term trends.

8.3. The Working Party's assessment of the P-Spline age-cohort, P-Spline age-period, the Lee-Carter and the Lee-Carter APC models against the above objectives is summarised in the table below followed by a discussion.

Objective	P-Spline age-cohort	P-Spline age-period	Lee-Carter	Lee-Carter APC
Ease of use	Y	Y	Y	Y
Ability to interpret the parameters	N	N	Y	Y
Model structure and fit	Y	Y	N	?
Cohort effects	Y	N	N	Y
Best estimate	Y	Y	Y	Y
Confidence intervals	Y	Y	Y	Y
Ability to generate sample paths	N	N	Y	Y

Ease of use

8.4. The Lee-Carter model is easy to use and has been used widely for a number of years and so is currently well understood. However, the requirement for boot-strapping, to allow for the uncertainty relating to parameter estimation, reduces the ease of use. Fitting the model and running a single simulation takes little computer time. However, running a large number of simulations requires significant computer resources and time.

8.5. The P-Spline models are also easy to use though the mathematics underlying the model may be considered more complex to program. As relatively new projection models, they are likely to be better understood once a user community develops. The models do take longer to fit than the Lee-Carter model but at the same time there is no need to run a large number of simulations.

8.6. The Lee-Carter APC model is similar in concept to the Lee-Carter model. However, the linear relationship between the age, period and cohort parameters makes it difficult to set constraints that will allow the model to converge to unique solutions for all datasets. The assumption that the cohort and period parameters can be modelled using univariate time-series models also needs further investigation. This model can take longer to fit than the P-Spline models and the need to run a large number of simulations means that the total computer resources and time required can be considerable.

Ability to interpret the parameters

8.7. The parameters in the Lee-Carter and Lee-Carter APC models can be readily interpreted. The shape of and trend in the age, period and cohort parameters can also be assessed for reasonability and can be analysed separately (e.g. the cohort parameters demonstrate which cohort effects are being captured). These models also allow the relative importance of the different sets of parameters to be easily measured.

8.8. Whilst the fitted and projected mortality rates from P-Spline models can be interpreted using various techniques, the underlying parameters are much more difficult to interpret or assess for reasonability.

Model structure and fit

8.9. The P-Spline models seem to produce sensible fits in the region of the data with a reasonable trade-off between smoothness and goodness-of-fit. They also produce smooth transitions in fitted mortality rates between the regions of the data and the projection. Therefore, trends seen in the last few years in the region of the data continue for at least a short period in the region of the projection. The results of back-testing the P-Spline age-cohort model, presented in Working Paper 20, showed that this model appeared to have been stable for the Assured Lives and ONS datasets.

8.10. As discussed in Section 4, the deviance residuals indicated that the Lee-Carter model did not produce sensible fits for the Assured Lives and ONS datasets. Also, the model did not seem to sufficiently smooth out the volatility in mortality rates between calendar years which also meant that the transition in fitted mortality rates between the regions of the data and the projection was not smooth. Back testing of this model indicated that it is less stable than the P-Spline age-cohort model.

8.11. The Lee-Carter APC model produced more sensible fits for the Assured Lives and ONS datasets. However, the fitted mortality rates were still not as smooth in the region of the data as the P-Spline models. Nor was the transition to the region of the projection smooth. We were unable to complete back-testing of this model but the convergence issues we faced indicated that the model may not be as stable as desired.

<u>Cohort effects</u>

8.12. The P-Spline age-period and age-cohort models are both able to identify cohort effects in the region of the data. However, the P-Spline age-cohort model is better able to reflect the cohort effects into the region of the projection.

8.13. Unlike the Lee-Carter model, the Lee-Carter APC model is able to reflect cohort effects into the region of the projection. However, the cohort effects captured by the model, as shown in the fitted parameter values, do not appear stable as the age range in the dataset is changed. As we were unable to carry out back-testing, we could not check whether the cohort effects were stable as additional years data was added.

<u>Best estimate</u>

8.14. All four of these models allow best estimates to be produced. However, whether these best-estimates could be considered reasonable and/or plausible will depend on the model fit and structure as well as the dataset being used. Whether they produce projections consistent with past trends and data will additionally depend on how well the model captures these trends.

<u>Confidence interval</u>

8.15. All four of these models allow uncertainty regarding future trends in mortality rates to be illustrated. Some users may consider that the percentiles produced by the P-Spline models are easier to use in illustrating this uncertainty to non-experts.

8.16. Percentiles could also be produced using the Lee-Carter and Lee-Carter APC models but these have to be derived from simulations and so take greater time and effort.

<u>Ability to generate sample paths</u>

8.17. The Lee-Carter and Lee-Carter APC models produce sample paths that can be used directly or via derived percentiles. The P-Spline models currently can only generate percentiles and so would not be suitable for modelling business, such as term assurance, where the volatility in mortality rates between calendar years, is important.

Summary of the Working Parties views on the models assessed

8.18. The above assessment shows that none of these four models meet all the desired objectives. Therefore, we do not consider that any one of these four models will always be more suitable.

8.19. The ability to generate sample paths is not as important an objective for longevity business. The P-Spline age-cohort model meets all the other objectives and so, in our view, is suitable for modelling the Assured Lives and ONS datasets. We do not consider the P-Spline age-period model to be as suitable for these datasets, which have clear cohort effects in the region of the data, as the preservation of the cohort effects into the region of the projection is not as good.

8.20. The Lee-Carter model is clearly unsuitable for the Assured Lives and ONS datasets as it produces poor fits and is unable to preserve cohort effects into the region of the projection. The Lee-Carter APC model seems to produce better fits and preserves cohort effects. However, we consider that further testing for this model is necessary, particularly regarding model fit and structure, before a judgement could be made regarding its suitability.

8.21. With this paper, the Working Party has completed its work on projection methodologies. There remain many areas regarding projection models that are worthy of further investigation and research and the CMI will retain an interest (and may commission further work) in these areas. The development of projection methodologies and publication of projections are of course areas for the wider profession, and not merely the CMI. Given the importance of projections, the actuarial profession may wish to consider how this should be taken on, perhaps by a Research Group or Working Party.

9. Appendices

Appendix A - Parameters used to generate projections

For all the Lee-Carter projections we have used an ARIMA(1,1,0) model to project the kappa parameters.

The following age ranges were used:

	Lee-Carter Deviance Residuals	**Age-Cohort Deviance Residuals**	**Age-Period Deviance Residuals**
Age range			
- Assured Lives	20-90	21-90	22-90
- Male ONS	20-89	21-89	23-89
- Female ONS	20-89	24-89	23-89

For all the Lee-Carter APC projections we have used an ARIMA(2,1,0) model to project the kappa parameters and an ARIMA(1,1,0) model to project the iota parameters.

The following age ranges were used:

	APC Lee-Carter
Ag range	
- Assured Lives	30-90
- Male ONS	30-89

Appendix B – Contour maps of projected improvements using the Lee-Carter model

Figure B1

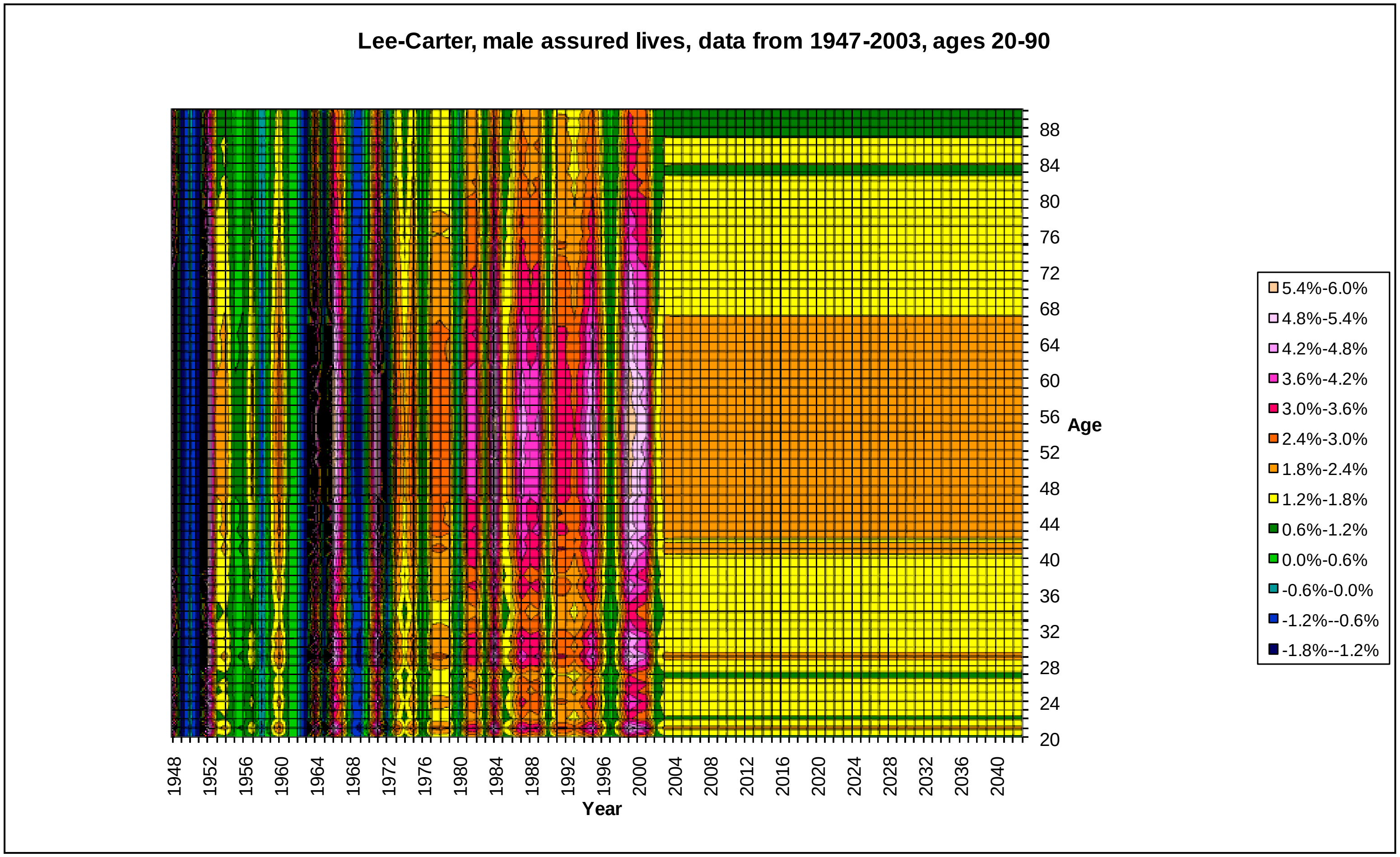
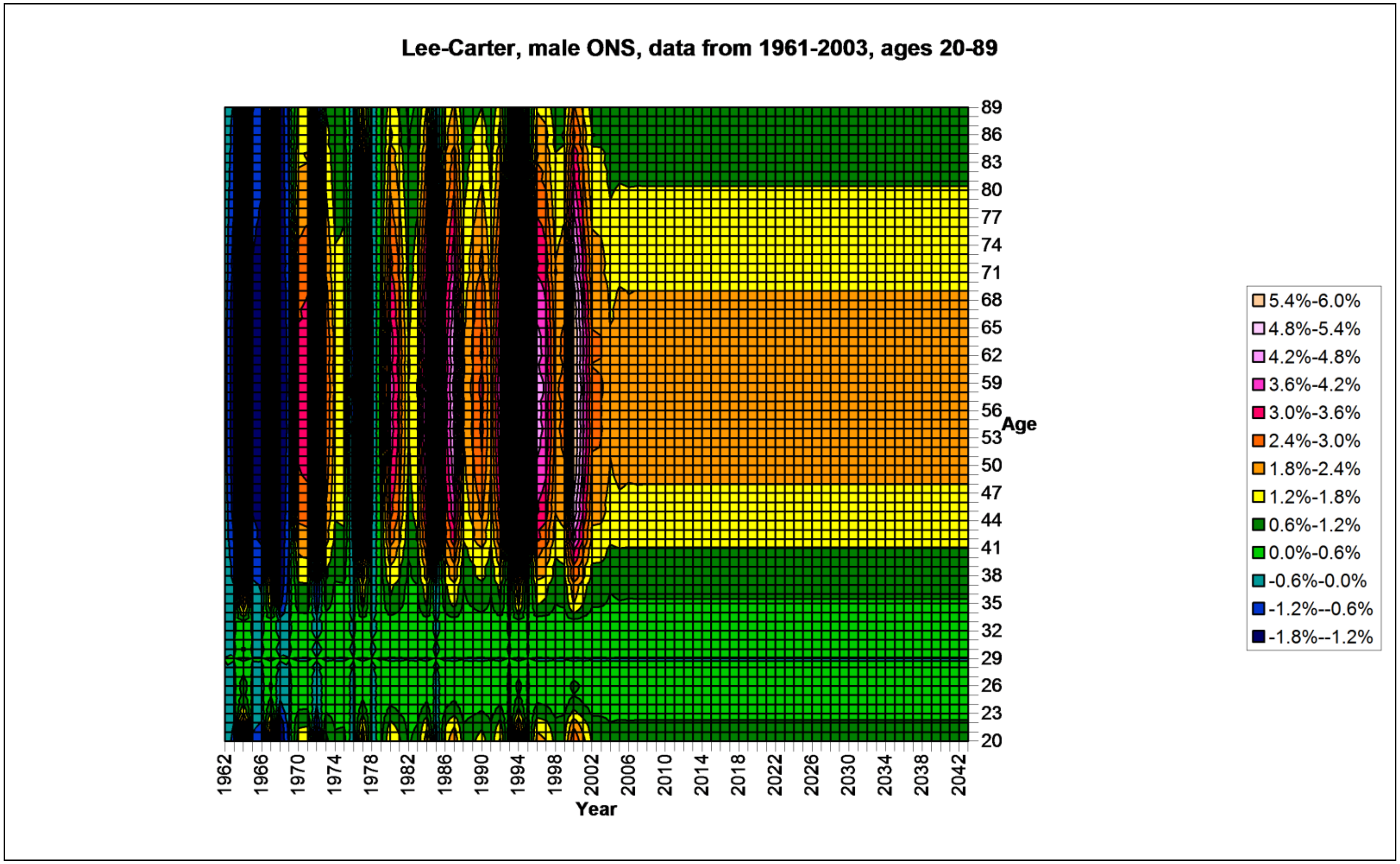
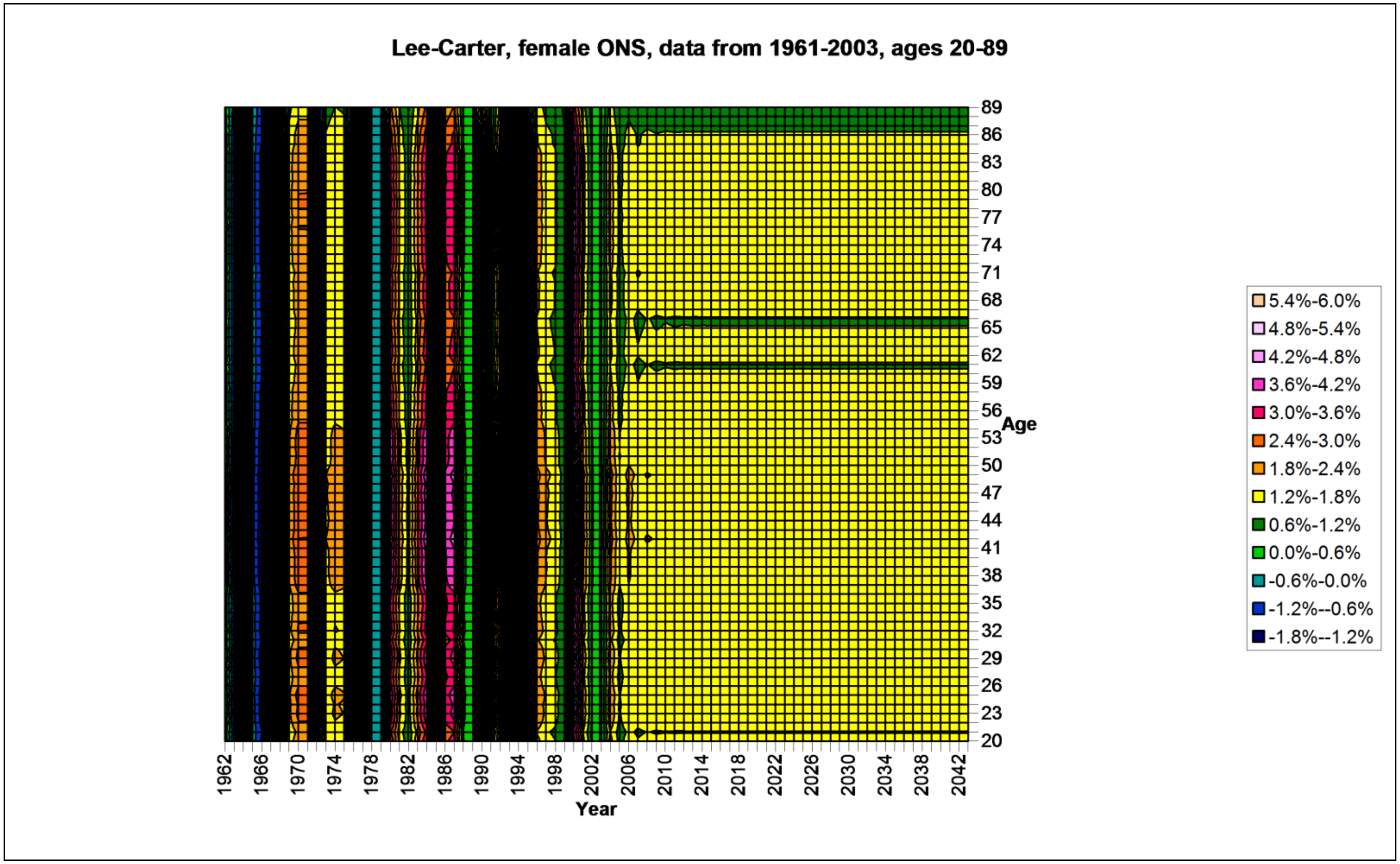


Figure B2



Lee-Carter, male ONS, data from 1961-2003, ages 20-89
89
86
83
80
77
74
71
68
65
62
59
56
53
50
47
44
41
38
35
32
29
26
23
20
Age
1962
1966
1970
1974
1978
1982
1986
1990
1994
1998
2002
2006
2010
2014
2018
2022
2026
2030
2034
2038
2042
Year
5.4%-6.0%
4.8%-5.4%
4.2%-4.8%
3.6%-4.2%
3.0%-3.6%
2.4%-3.0%
1.8%-2.4%
1.2%-1.8%
0.6%-1.2%
0.0%-0.6%
-0.6%-0.0%
-1.2%--0.6%
-1.8%--1.2%

Figure B3



Lee-Carter, female ONS, data from 1961-2003, ages 20-89
89
86
83
80
77
74
71
68
65
62
59
56
53
50
47
44
41
38
35
32
29
26
23
20
Age
1962
1966
1970
1974
1978
1982
1986
1990
1994
1998
2002
2006
2010
2014
2018
2022
2026
2030
2034
2038
2042
Year
5.4%-6.0%
4.8%-5.4%
4.2%-4.8%
3.6%-4.2%
3.0%-3.6%
2.4%-3.0%
1.8%-2.4%
1.2%-1.8%
0.6%-1.2%
0.0%-0.6%
-0.6%-0.0%
-1.2%--0.6%
-1.8%--1.2%

Appendix C – Contour maps of projected improvements using the Lee-Carter APC model

Figure C1

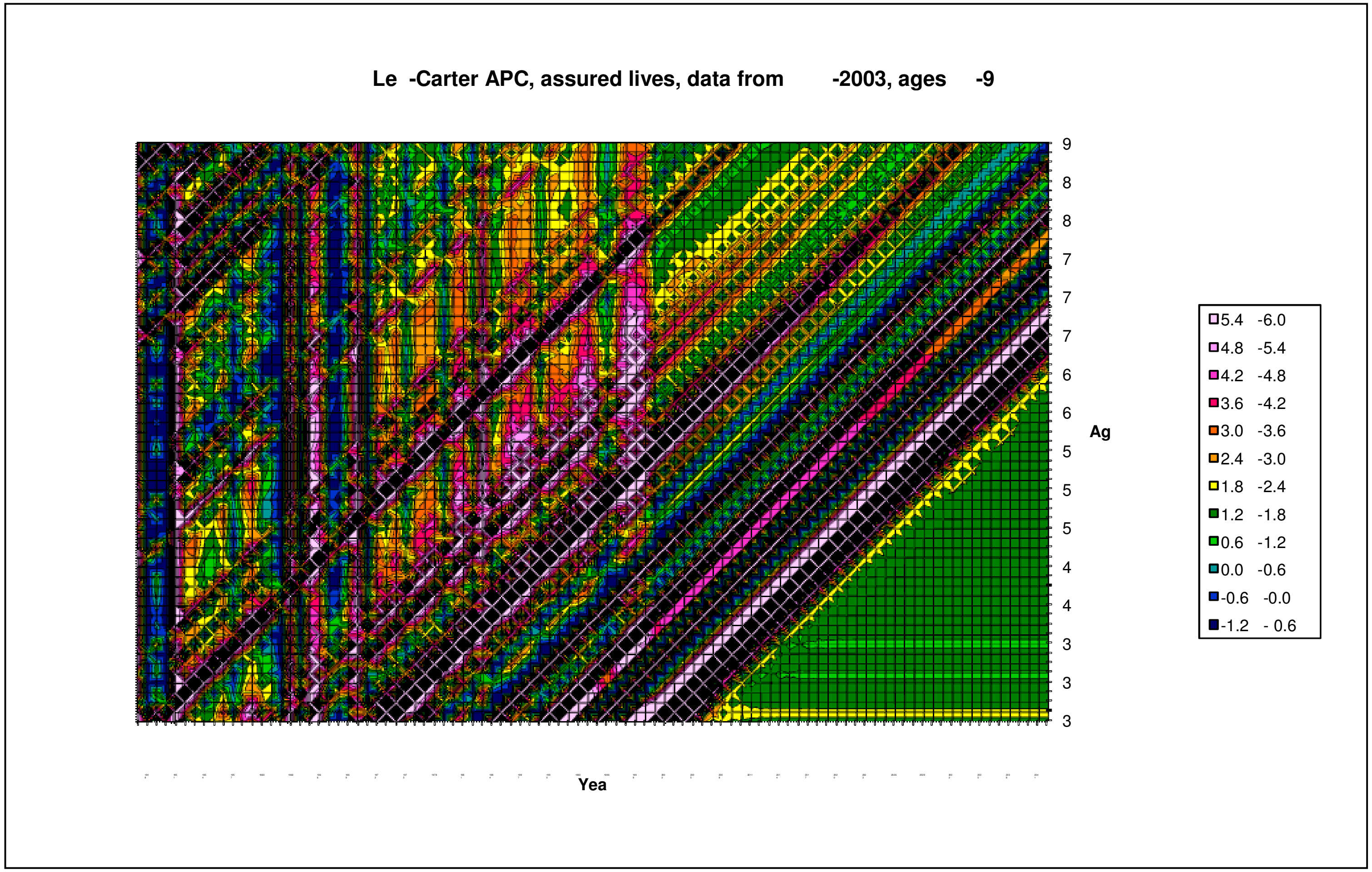
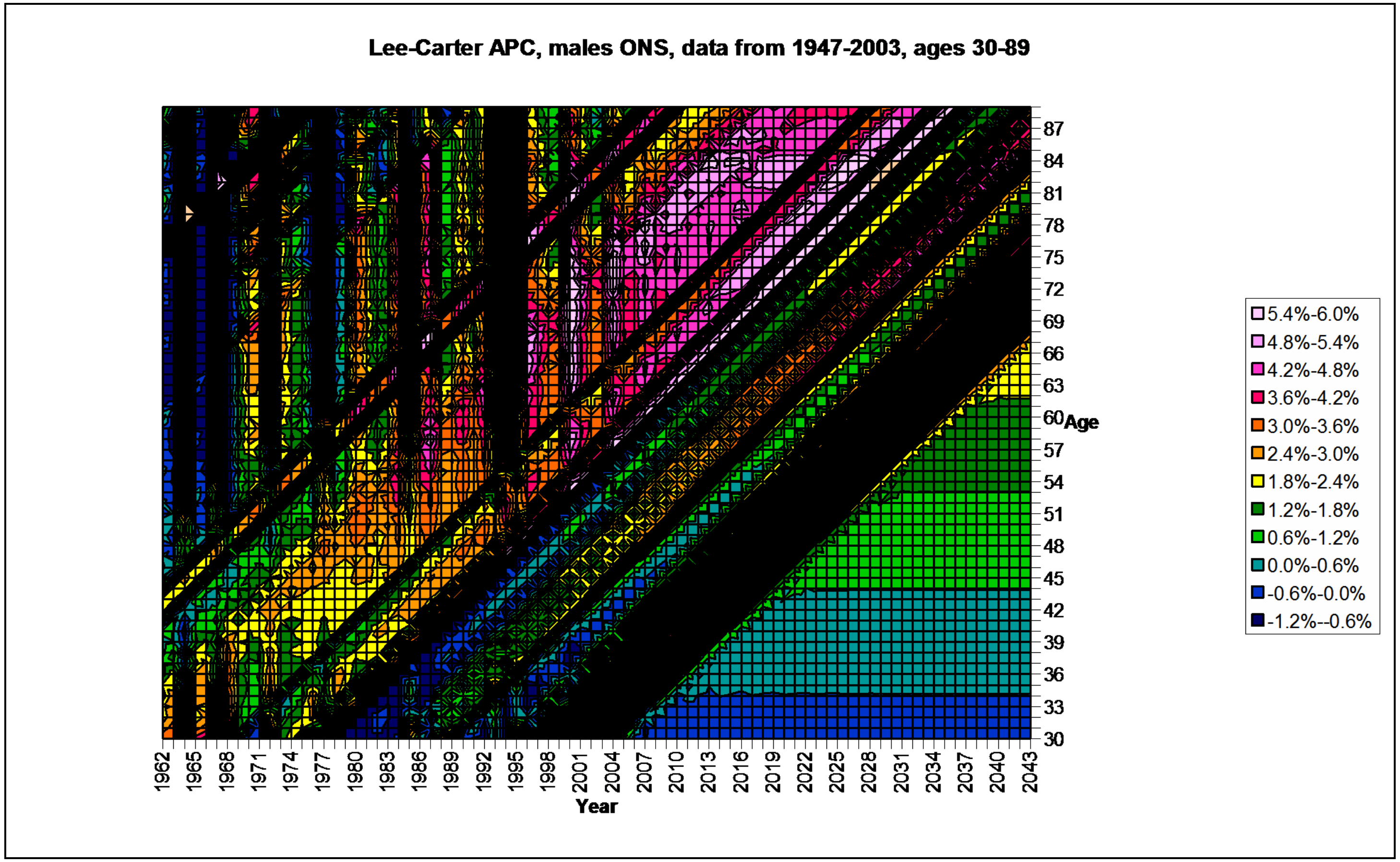


Figure C2

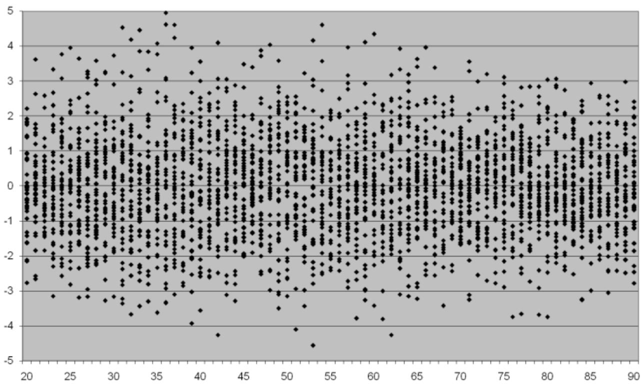


Lee-Carter APC, males ONS, data from 1947-2003, ages 30-89
87
84
81
78
75
72
69
66
63
60
Age
57
54
51
48
45
42
39
36
33
30
1962
1965
1968
1971
1974
1977
1980
1983
1986
1989
1992
1995
1998
2001
2004
2007
2010
2013
2016
2019
2022
2025
2028
2031
2034
2037
2040
2043
Year
5.4%-6.0%
4.8%-5.4%
4.2%-4.8%
3.6%-4.2%
3.0%-3.6%
2.4%-3.0%
1.8%-2.4%
1.2%-1.8%
0.6%-1.2%
0.0%-0.6%
-0.6%-0.0%
-1.2%--0.6%

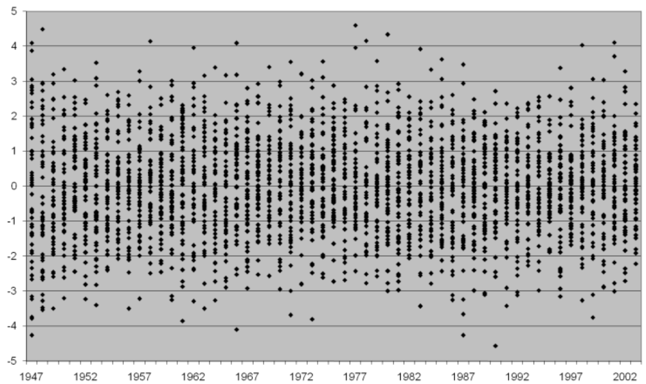
Deviance Residuals of Lee-Carter Fits (see Appendix D of Working Paper 25)

Figure D1 – Assured lives, data from 1947-2003, ages 20-90

(a) Age



(b) Year



(c) Cohort

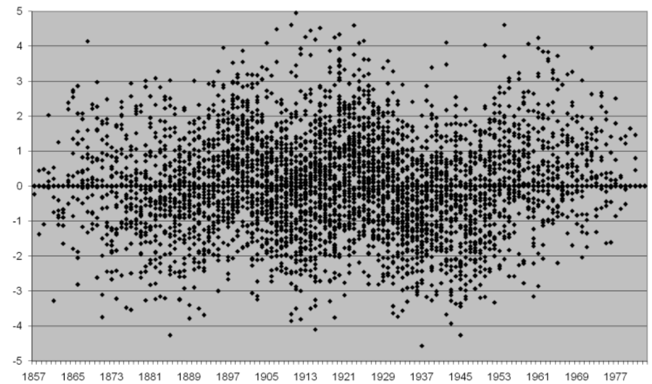
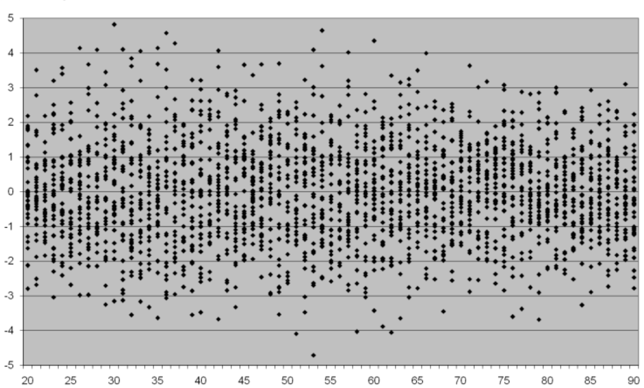
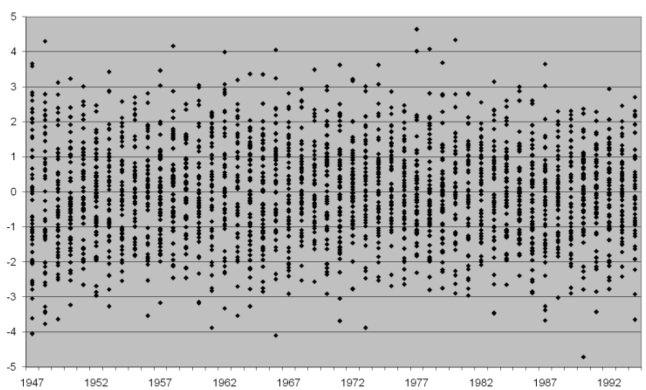


Figure D2 – Assured lives, data from 1947-1994, ages 20-90

(a) Age



(b) Year



(c) Cohort

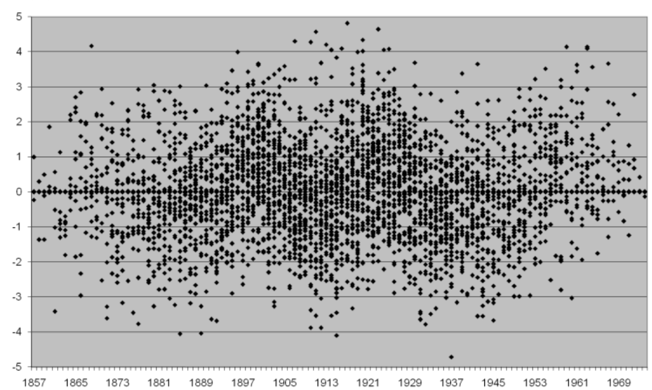
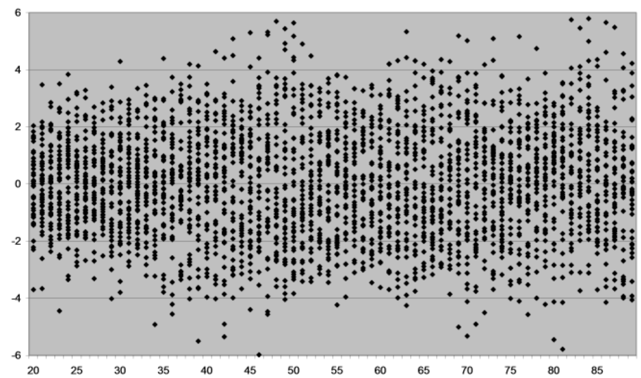
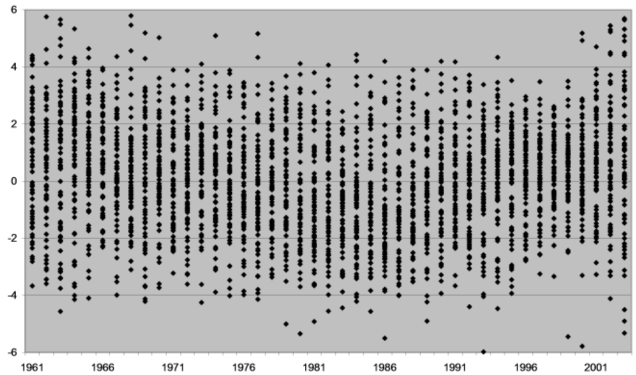


Figure D3 – ONS males, data from 1961-2003, ages 20-89

(a) Age



(b) Year



(c) Cohort

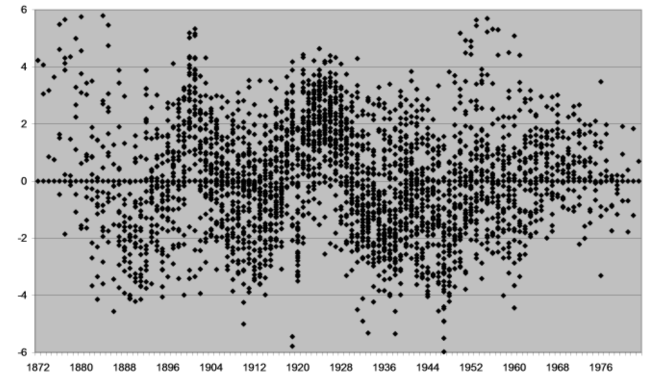
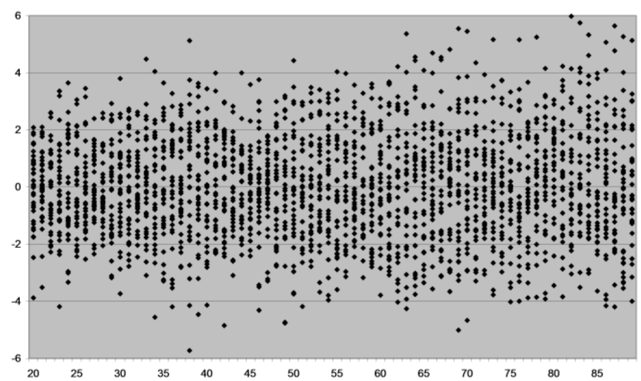
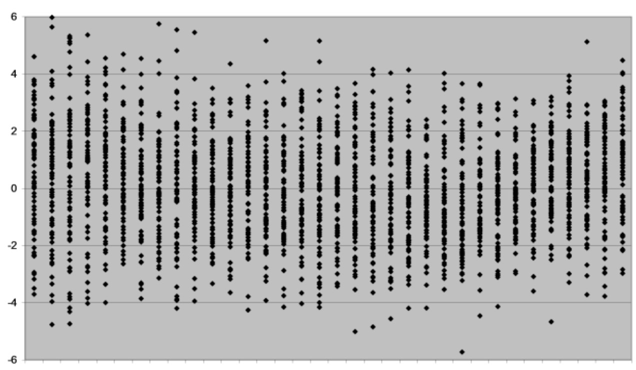


Figure D4 – ONS males, data from 1961-1994, ages 20-89

(a) Age



(b) Year



(c) Cohort

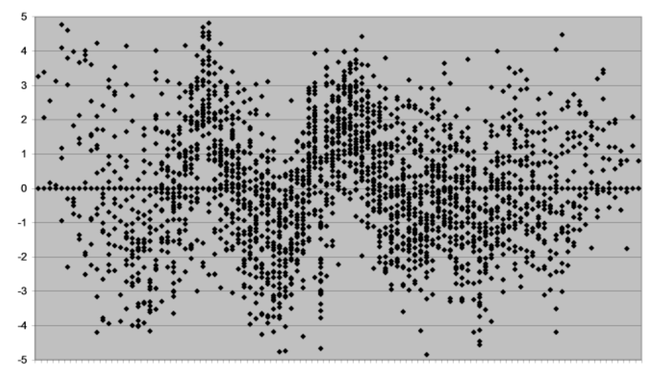


Figure D5 – ONS females, data from 1961-2003, ages 20-89

(a) Age *(b) Year* *(c) Cohort*

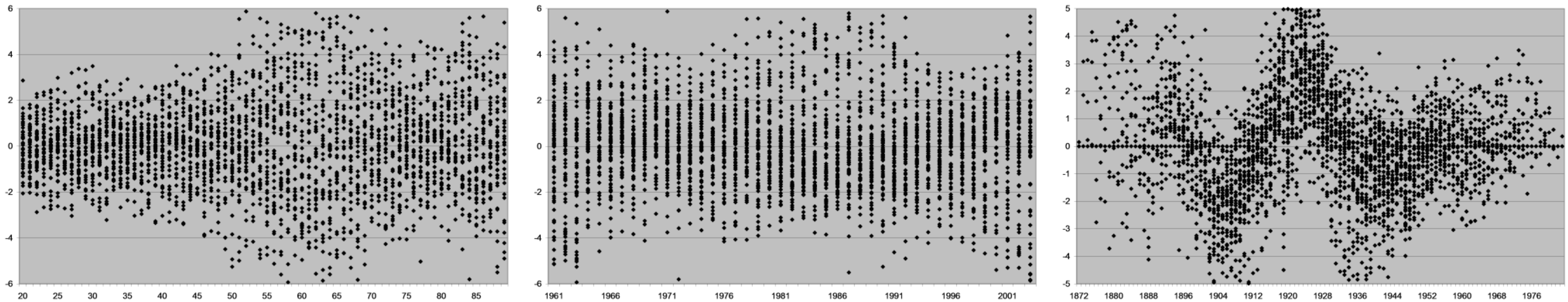
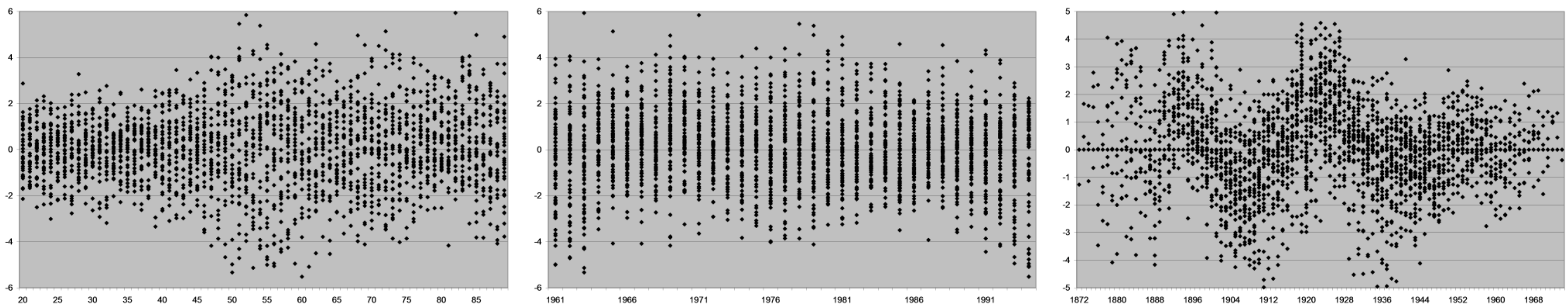


Figure D6 – ONS females, data from 1961-1994, ages 20-89

(a) Age *(b) Year* *(c) Cohort*



Appendix E – Deviance Residuals of Lee-Carter APC Fits

Figure E1 - Assured lives, data from 1947-2003, ages 30-90

(a) Age *(b) Year* *(c) Cohort*

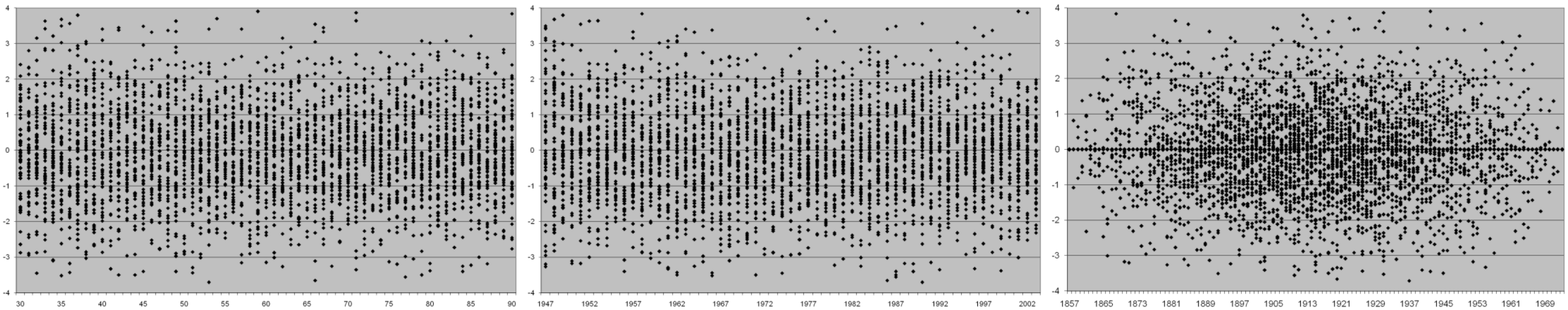
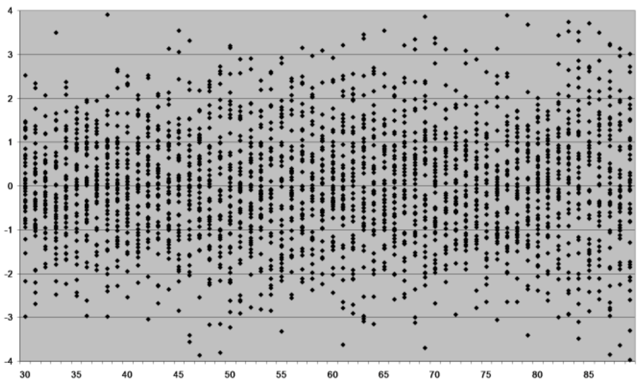
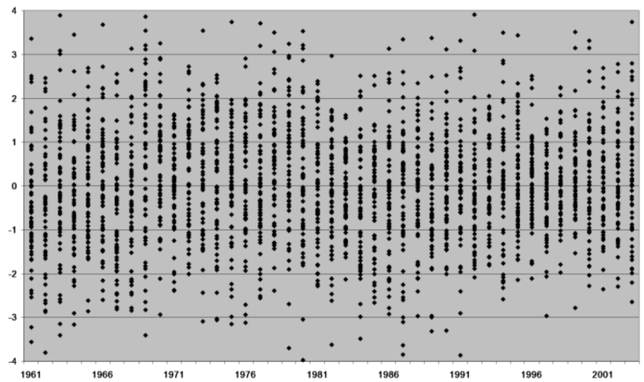


Figure E2 – ONS males, data from 1961-2003, ages 30-89

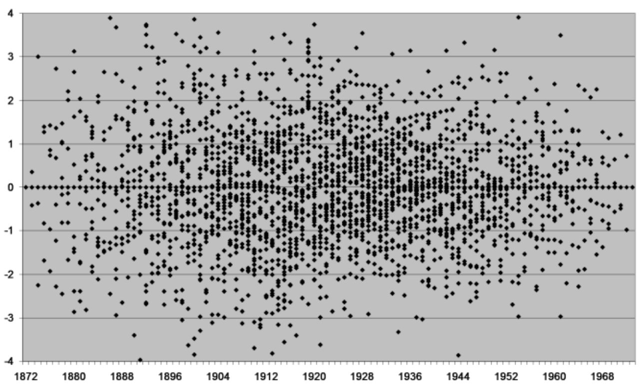
(a) Age *(b) Year* *(c) Cohort*



4
3
2
1
0
-1
-2
-3
-4
30 35 40 45 50 55 60 65 70 75 80 85



4
3
2
1
0
-1
-2
-3
-4
1961 1966 1971 1976 1981 1986 1991 1996 2001

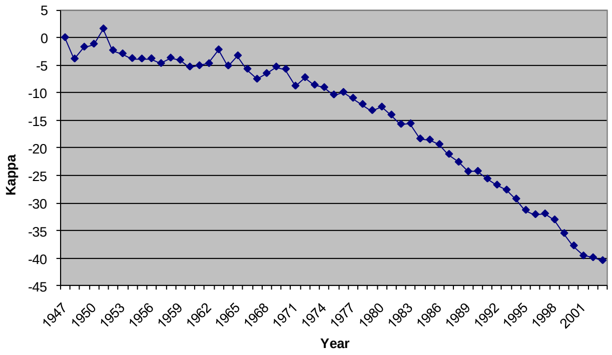


4
3
2
1
0
-1
-2
-3
-4
1872 1880 1888 1896 1904 1912 1920 1928 1936 1944 1952 1960 1968

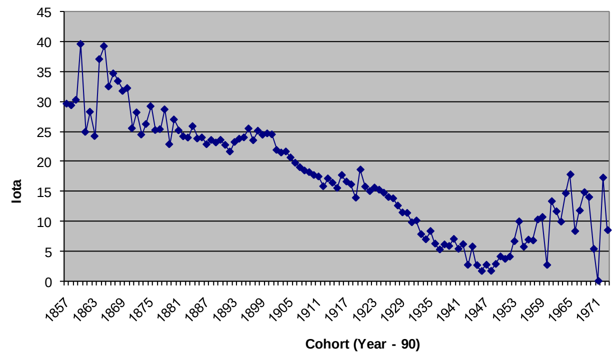
Appendix F – Parameter Graphs of Lee-Carter APC Fits

Figure F1 - Assured lives, data from 1947-2003, ages 30-90

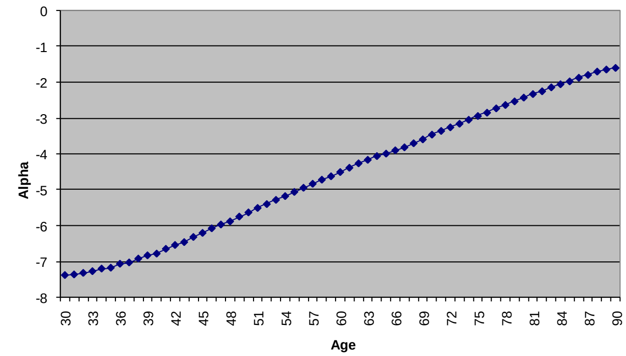
(a) Kappa (k(t))



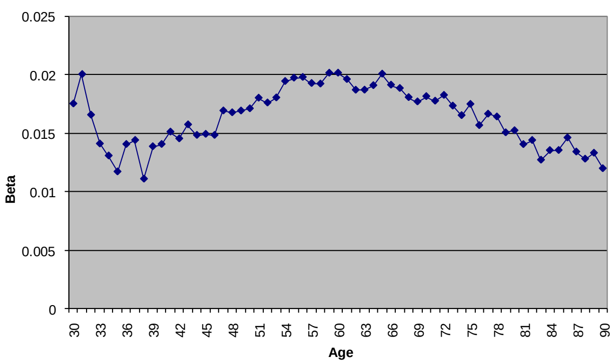
(b) Iota (I(c))



(c) Alpha (a(x))



(d) Beta applying to Kappa ($b_1(x)$)



(e) Beta applying to Iota ($b_2(x)$)

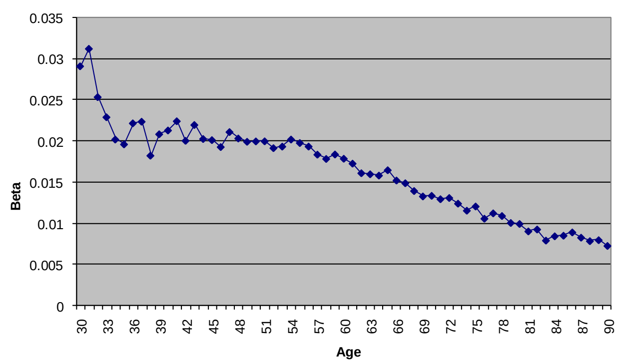
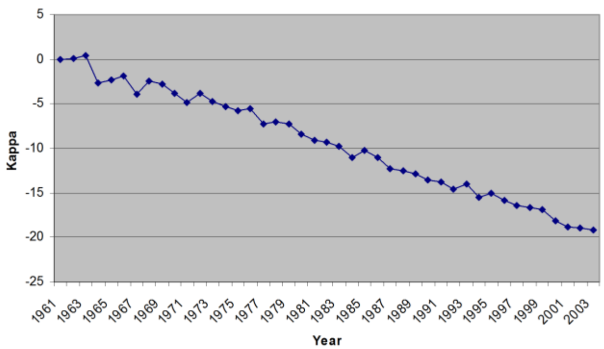
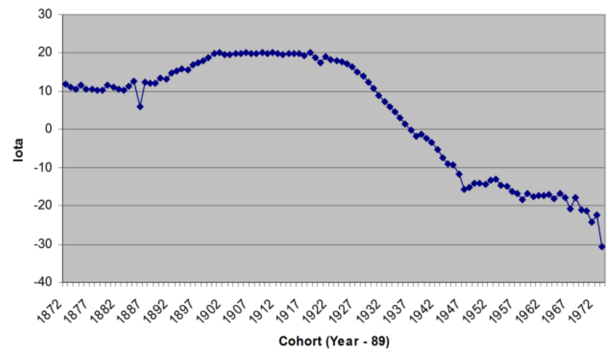


Figure F2 - ONS males, data from 1961-2003, ages 30-89

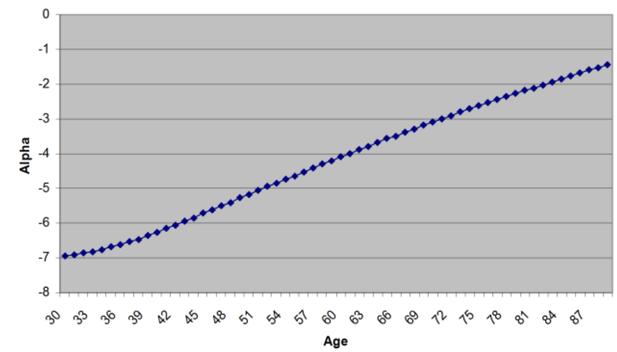
(a) Kappa (k(t))



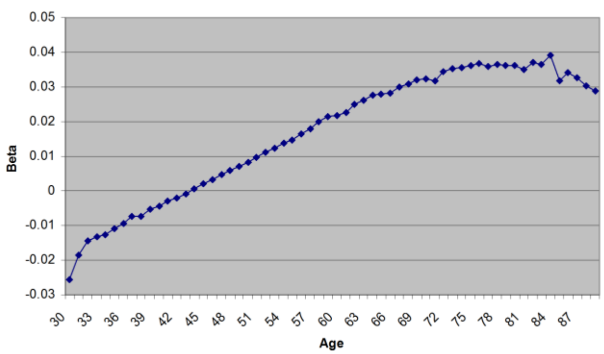
(b) Iota (I(c))



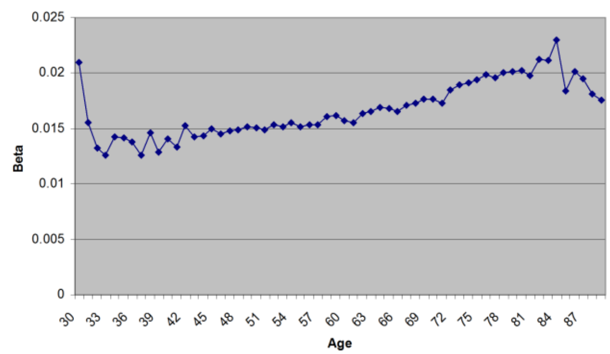
(c) Alpha (a(x))



(d) Beta applying to Kappa ($b_1(x)$)



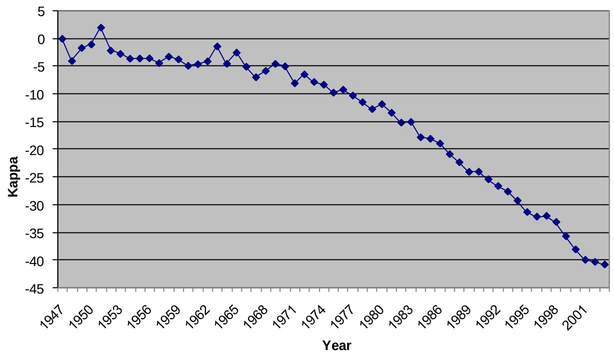
(e) Beta applying to Iota ($b_2(x)$)



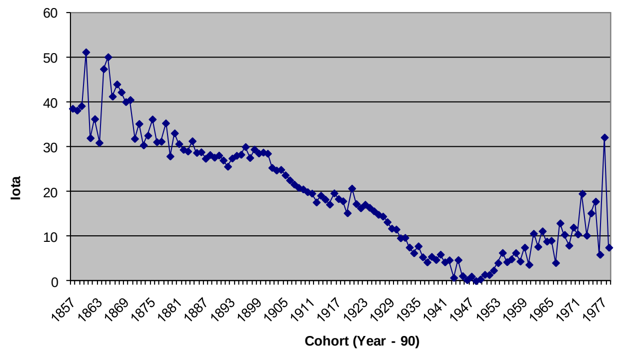
Appendix G – Parameter Graphs of Lee-Carter APC Fits (Extended Age Range)

Figure G1 - Assured lives, data from 1947-2003, ages 25-90

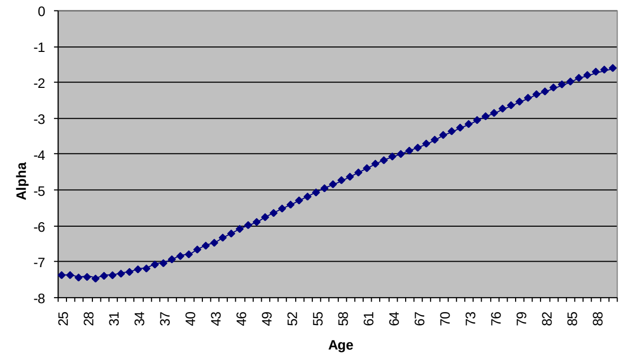
(a) Kappa (k(t))



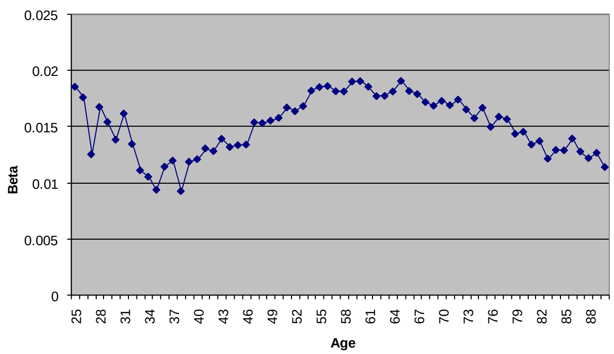
(b) Iota (I(c))



(c) Alpha (a(x))



(d) Beta applying to Kappa ($b_1(x)$)



(e) Beta applying to Iota ($b_2(x)$)

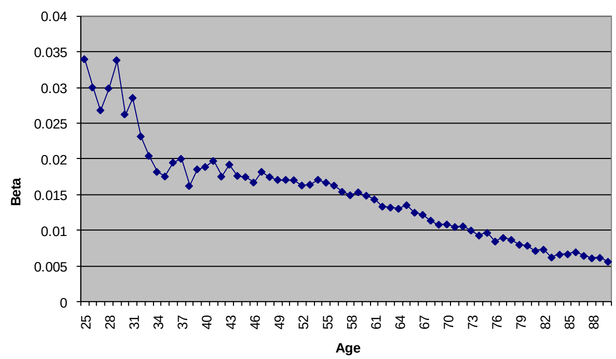
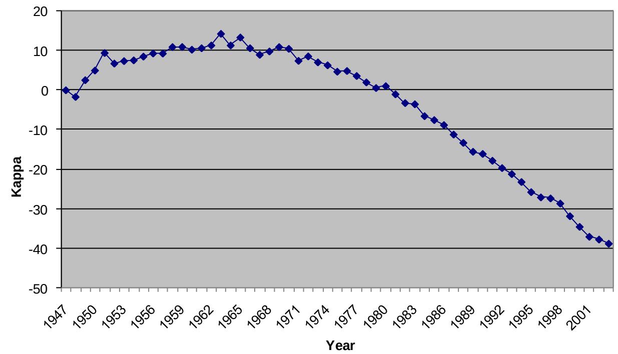
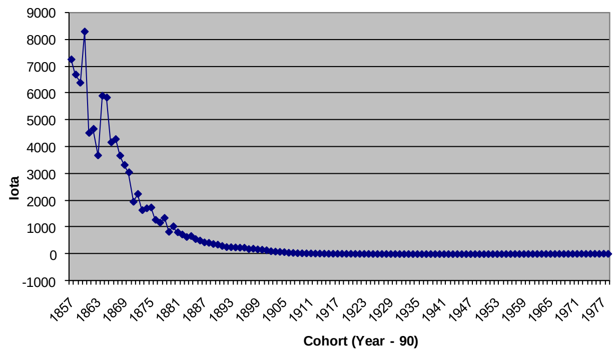


Figure G2 - Assured lives, data from 1947-2003, ages 25-90 (new constraints)

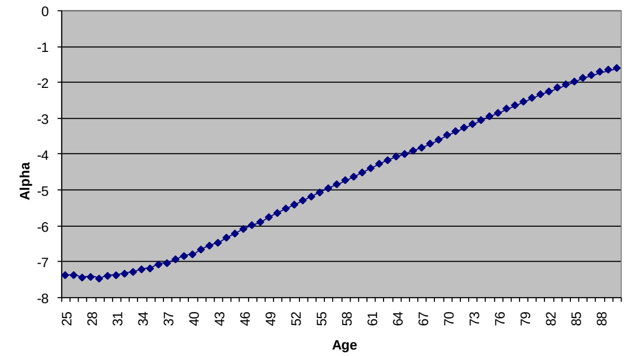
(a) Kappa (k(t))



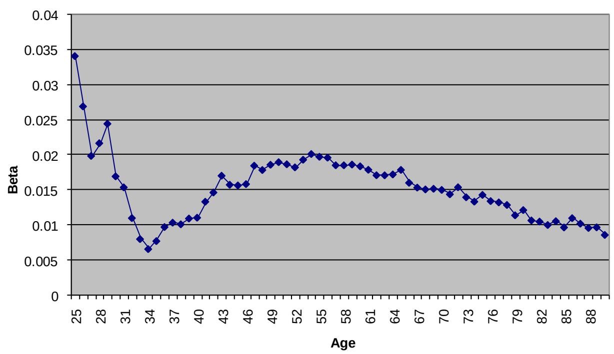
(b) Iota (I(c))



(c) Alpha (a(x))



(d) Beta applying to Kappa ($b_1(x)$)



(e) Beta applying to Iota ($b_2(x)$)