THE AUSTRALIAN NATIONAL UNIVERSITY

DEMOGRAPHY AND SOCIOLOGY PROGRAM

RESEARCH SCHOOL OF SOCIAL SCIENCES

AGE-TIME INTERACTIONS IN MORTALITY PROJECTION: APPLYING LEE-CARTER TO AUSTRALIA

Heather Booth

John Maindonald

Len Smith

Working Papers in Demography

Revised 2 August 2002

AGE-TIME INTERACTIONS IN MORTALITY PROJECTION: APPLYING LEE-CARTER TO AUSTRALIA

Heather Booth[1], John Maindonald[2] and Len Smith[3]

Abstract

Application of the Lee-Carter method to Australian data shows that model assumptions are not always met because of age-time interactions. The Lee-Carter method is adapted to take account of departures from linearity in the dominant time component and the failure to satisfy the assumption of an invariant age component. The most significant adaptation is a methodology for determining the 'optimum' fitting period in order to address non-linearity in the time component. In the Australian case, this has the additional effect that the assumption of an invariant age component is better met. Additional technical adaptations are also made. The model is expanded to take account of age-time interactions by incorporating the second and higher terms, but these are not easily incorporated into forecasts. The adapted methodology produces forecasts of life expectancy that are higher than official projections.

Introduction

The populations of the industrialised world underwent a major mortality transition over the course of the twentieth century. Mortality rates declined dramatically. In Australia, life expectancy increased from 57 years in 1901-10 to 78 years in 1995-97. Two trends dominated the mortality decline. The first was a reduction in mortality due to infectious diseases affecting mainly young ages, particularly in the first half of the century. The second was increasing and then from about 1970 decreasing deaths due to chronic diseases affecting older ages. In Australia, the balance of these two trends resulted in a period of almost constant mortality in the 1960s. Other trends also emerged: the post-World War 2 accident 'hump' and, most recently, decreasing mortality at very old ages. The current pace of mortality decline in low-mortality populations, including Australia, shows no sign of deceleration, leading to international debate about whether there are limits to human longevity and how soon these might be reached (Wachter and Finch 1997).

---

[1] Demography and Sociology Program, Research School of Social Sciences, ANU. Correspondence: heather.booth@anu.edu.au.
[2] Statistical Consulting Unit, Graduate School, ANU.
[3] Australian Centre for Population Research and National Centre for Epidemiology and Population Health, ANU.

In order to forecast mortality, it must be accurately modelled. Many different models have been employed (Tabeau 2001). Among recent models, the 8-parameter model developed by Heligman and Pollard (1980) has been found to fit a wide range of mortality schedules. However, the large number of parameters limits the model's potential for projection. Further, the extrapolation of model parameters for forecasting generally involves a certain degree of subjective judgment.

Lee and Carter (1992) proposed a method for forecasting future mortality, based on a log-additive model of age-specific death rates with a dominant time component and a fixed relative age component. They used a matrix decomposition method to identify the two components, using US mortality data for 1900-89. They then projected the time component to 2065, using a time series linear forecasting model (autoregressive integrated moving average (ARIMA)). The log-additive model and ARIMA constitute the main features of the Lee-Carter method.

The Lee-Carter method is regarded as the state-of-the-art in mortality forecasting and has been adopted widely in North America (see for instance, *North American Actuarial Journal* 2(4) 1998; Lee 2000). Several advantages are claimed: a parsimonious demographic model is combined with statistical time-series methods, the method involves no subjective judgements, forecasting is based on persistent long-term trends and probabilistic confidence intervals are provided for the forecasts (Lee and Carter 1992). Tuljapurkar, Li and Boe (2000) applied the method to data for 1950-94 from the seven richest countries (G7), identifying a 'universal pattern' (p.789) of mortality decline for this period characterised by roughly linear decline in the dominant time component for each country, and projected the declines forward to 2050.

In an assessment of the method, Lee and Miller (2000) acknowledged that the main methodological problem is the assumption that the age component is invariant over time. Certainly, the mortality experience of the industrialised world over the course of the twentieth century would suggest substantial age-time interaction: the two dominant trends affected different age groups at different times. Lee and Miller (2000:9) suggested that a 'simple and satisfactory solution' to this problem is to base the forecast on data since 1950, as Tuljapurkar *et al.* (2000) had done, so that the assumption that the age component is invariant applies to only half the century. While this may be expected to reduce the extent of the age-time interaction, some interaction is likely to remain.

In the case of Australia, past mortality patterns suggest that whether the entire century or only the latter half is taken into account, the assumption of invariance in the age component will be violated (AIHW 2000:ch.8). This calls into question the applicability of the log-additive model and hence the Lee-Carter method. Further, in an application of the method to Australian data for 1950-94 for comparison with the G7 countries, Booth, Chauhan, Maindonald and Smith (2000) found a marked departure from linearity in the dominant time component, calling into question the 'universal pattern' of mortality decline and hence the applicability of the linear model for forecasting Australian mortality. Thus, there is clearly a need to examine ways to improve the applicability of the Lee-Carter method in the Australian context with a view to increasing model performance and forecasting power.

Aims

The overall aim of this paper is to improve the methodology for forecasting Australian mortality in order to enhance model performance and increase forecasting power. In particular, the aim is to find ways to increase the applicability of the Lee-Carter method given departures from linearity in the dominant time component and the existence of age-time interactions. This includes seeking to develop methodologies for identifying the most appropriate period to which to apply the Lee-Carter method and for incorporating age-time interactions into the Lee-Carter model.

Methods and Data

The point of departure for this research is the Lee-Carter method with linear forecasting model. Several adaptations to this method are proposed in order to increase its applicability in the Australian context. The Lee-Carter method and proposed adaptations are described below. Details of the available data are also given.

The Lee-Carter method

The Lee-Carter method comprises a model of mortality and a methodology for fitting the model, plus a time series model of the dominant parameter which is used for extrapolative purposes in forecasting. The Lee-Carter model is:

$$\ln[m(x,t)] = a(x) + b(x)k(t) + e(x,t)$$

where	m(x,t)	is the central death rate at age x in year t
	k(t)	is an index of the level of mortality
	a(x)	is a set of age-specific constants describing the general pattern of mortality by age
	b(x)	is a set of age-specific constants describing the relative speed of change at each age
	e(x,t)	is the residual at age x and time t

The time component, k(t), or the 'dominant temporal signal' (Tuljapurkar *et al.* 2000:789) captures the main time trend on the logarithmic scale in death rates at all ages. The model makes no assumptions about the nature of the trend in k(t). The age component, b(x), modifies the main time trend according to whether change at a particular age is faster or slower than the main trend (and in the same or opposite direction). The model assumes that b(x) is invariant over time.

In order to obtain a unique solution for the system of equations of the model, a(x) is set equal to the means over time of ln[m(x,t)] and b(x) is constrained to sum to unity. The k(t) values sum to zero. Under this normalisation, b(x) is the proportion of the change in overall mortality on the logarithmic scale attributable to age x. To find a solution, a(x) is subtracted from ln[m(x,t)] and the model becomes log-additive.

Lee and Carter used singular value decomposition (SVD) to find a least squares solution in the estimation of b(x) and k(t). The SVD (Trefethen and Bau 1997) decomposes the matrix of ln(m(x,t)) into the product of three matrices, representing the age component, the singular values and the time component. In the rank 1 approximation used by Lee and Carter, k(t) is derived from the first vector of the time-component matrix and the first singular value, and b(x) is derived from the first vector of the age-component matrix. The second and higher vectors of the SVD comprise the residual. Typically for low-mortality populations, this first rank 1 approximation accounts for about 95 per cent of the variance in the ln[m(x,t)].

Since the model estimates the logarithms of rates, a given ratio of actual to estimated mortality rates has the same weight in the SVD regardless of whether the rates involved are large or small. In order to correct for this distorted weighting, Lee and Carter adjusted each k(t) to reproduce annual total deaths, D(t), while leaving a(x) and b(x) unchanged. This adjustment gives greater weight to ages at which numbers of deaths are large.

An appropriate ARIMA model is used to achieve a time series fit to adjusted k(t) (Diggle 1990). In principle, this could be linear or non-linear. For the data used by Lee and Carter (and for those used by Tuljapurkar *et al.*) a random walk with drift was judged to be most appropriate. The model is

$$k'(t) = k'(t-1) + d + \varepsilon(t)$$

where k′(t) denotes adjusted k(t), d is ‘drift’ (a constant rate of change in k′(t)) and ε(t) are uncorrelated errors. This model, which is linear in t, is then used to extrapolate k′(t) into the future. The standard error in d and ε(t) together estimate the uncertainty associated with a one-year forecast, and this estimate is used to produce probabilistic confidence intervals for the forecast values of k(t), which translate into confidence intervals for projected mortality rates and life-table functions such as life expectancy. When k(t) is linear in t, the log-additive model becomes log-linear.

Application of the Lee-Carter method

Five adaptations are made to the Lee-Carter method with linear ARIMA in order to improve its applicability. The main adaptation addresses the departure from linearity in k(t) by introducing a methodology for identifying the most appropriate period to which to apply the Lee-Carter method. As part of this methodology, a further adaptation is introduced: k(t) is adjusted by fitting to the age distribution of deaths rather than to total deaths. Other departures from the original Lee-Carter method include a different jump-off value for extrapolating k(t), the expansion of the log-additive model to include higher terms, and an alternative presentation of b(x). The examination of model performance is an integral and explicit part of these adaptations. The five adaptations are detailed below.

In the original Lee-Carter method, adjustment of the k(t) values to reproduce D(t) renders the minimisation criterion unclear. Adjustment to reproduce e(0), as suggested by Lee and Miller (2000) in order to avoid the need for population data, renders the minimisation criterion more complex and even less clear. In the present study, k(t) is adjusted by fitting a Poisson regression model to the annual number of deaths at each age, D(x,t), while again leaving a(x) and b(x) from the SVD solution unchanged. The Poisson model is

$$\ln[D(x,t)] = \ln[N(x,t)] + \ln[m'(x,t)] + e'(x,t)$$

where $\ln[m'(x,t)] = a(x) + b(x)k'(t)$, $k'(t)$ refers to adjusted $k(t)$ under the present adjustment and $e'(x,t)$ refers to the residuals after adjustment of $k(t)$. An appropriate minimisation criterion is the deviance (Maindonald 1984):

$$\text{deviance}(t) = 2\, \Sigma_x \{D(x,t) \ln[D(x,t)/D'(x,t)] - (D(x,t) - D'(x,t))\}$$

where $D'(x,t)$ denotes fitted deaths. Fitted deaths are obtained by

$$D'(x,t) = N(x,t)\{\exp[a(x) + b(x)k'(t)]\}$$

where $N(x,t)$ denotes the population aged x at time t. To a close approximation, the deviance(t) is equal to the chi-squared statistic of the lack of fit in $D(x,t)$:

$$\chi^2(t) = \Sigma_x \{ [D(x,t) - D'(x,t)]^2 / D'(x,t) \}$$

The deviance is used for the actual minimisation. However, because the χ^2 statistic is a more familiar measure, it is used to describe the lack of fit.

This adaptation has several advantages. First the use of the deviance as the minimisation criterion is a standard statistical practice, which is useful in comparisons, in particular in determining the fitting period (see below)[4]. Second, the fit to $D(x,t)$ gives greater weight to ages at which numbers of deaths are large and these weights are known (whereas in the Lee-Carter adjustment, although the weights are generally greater at these ages, the weighting is complex and does not conform to standard statistical criteria). While the total number of annual deaths may not be reproduced exactly, this loss is outweighed by the greater accuracy in fitting to the age distribution of deaths. Third, this is a (conditional) maximum likelihood procedure assuming quasi-Poisson errors, the statistical properties of which are well-understood.[5] Fourth, it is easier to calculate than the Lee-Carter minimisation because it does not require finding the zero of a non-linear function.

As noted above, the main adaptation introduces a methodology for identifying the most appropriate period to which to apply the Lee-Carter method. In doing so, it is assumed *a priori* that $k'(t)$ is linear. When $k'(t)$ departs from the linear trend (as for example in Figure 1), the linear fit to $k'(t)$ can be improved by appropriately restricting the fitting period. The ending year of the fitting period is determined by the latest available data. The optimal choice of period is thus determined by the starting year, denoted S. In this analysis, the choice of S is based on two criteria involving statistical measures of relative lack of fit.

The total lack of fit to the log-linear model is composed of two parts: the base lack of fit from the log-additive model and the adjustment of $k(t)$, and the additional lack of fit from the imposition of the ARIMA linear model on $k'(t)$. The base lack of fit from the log-additive model for the period S to 1999 is measured by $\chi^2_{logadd}(S) = \Sigma_t \chi^2_{logadd}(t)$ where the $D'(x,t)$ are derived from the $k'(t)$. The total lack of fit to the log-linear model is measured by $\chi^2_{loglin}(S) = \Sigma_t \chi^2_{loglin}(t)$ where the $D'(x,t)$ are derived

[4] Clearly, the deviance or χ^2 could also be calculated after adjusting $k(t)$ to reproduce $D(t)$. This would allow comparisons to be made but would not involve minimisation of these statistics.

[5] Wilmoth (1993) also examined alternative approaches to the estimation of $k(t)$ and $b(x)$ using weighted least squares and maximum likelihood estimation.

from the linear fit to $k'(t)$. This total lack of fit will be greater than or equal to base lack of fit. In order to compare $\chi^2_{loglin}(S)$ with $\chi^2_{logadd}(S)$ they are each first divided by the relevant degrees of freedom (df) to produce *mean-χ^2 statistics*. For n age categories and m years in the fitting period (m = 1999-S+1), the df for $\chi^2_{logadd}(S)$ is (n-1)(m-2) and the df for $\chi^2_{loglin}(S)$ is n(m-2).[6]

The choice of S is determined by the extent of the additional lack of fit relative to the total. If the fit of the ARIMA linear model is good, the additional lack of fit will be small. The first statistical measure of additional lack of fit adopted here is the ratio of total to base lack of fit,

$$R(S) = \{\chi^2_{loglin}(S)/[n(m-2)]\} \ / \ \{\chi^2_{logadd}(S)/\{(n-1)(m-2)]\}.$$

The second measure is of the *marginal* effect of including year S in the fitting period and is defined as the ratio of the differences in total and base mean-χ^2 statistics for S and S+1:

$$RD(S) = \{[\chi^2_{loglin}(S) - \chi^2_{loglin}(S+1)] \ / \ n \ \} \ / \ \{[\chi^2_{logadd}(S) - \chi^2_{logadd}(S+1)] \ / \ (n-1)\}.$$

Small values of R(S) and RD(S) indicate that the additional lack of fit is relatively small. The two criteria for choice of S are that R(S) and RD(S) be substantially smaller than corresponding statistics for preceding values of S, indicating that the inclusion of year S-1 (and preceding years) in the fitting period results in a relatively large reduction in the goodness of fit of the ARIMA linear model. A change-point test could be used to formalise these criteria (Kuan and Hornik 1995).

A further adaptation of the original Lee-Carter method is in the choice of jump-off value for extrapolating k(t). This is taken as $k'(1999)$, which is based on fitting to D(x,t). The original Lee-Carter model uses their adjusted k(1999) (based on reproducing D(t)), which has been acknowledged by Lee and Miller (2000:6) to be biased.

Having established the optimal fitting period based on lack-of-fit criteria to the linear model, the issues of applicability and performance of the log-additive model (after adjustment of k(t)) need to be examined for this period. The applicability of the model in terms of the assumption of invariant b(x) is assessed by comparing b(x) schedules for a series of subperiods within the fitting period. Model performance is measured in terms of the randomness of the residuals. In the Lee-Carter model these reflect 'particular age-specific historical influences not captured by the model' (Lee and Carter 1992:660) and are comprised of the second and higher terms of the SVD with mean 0. It is possible that systematic variation such as age-time interactions could be captured by the second and higher terms and a lack of randomness would indicate the presence of age-time interactions not modelled by b(x)k(t). Further, the adjustment of k(t) may have introduced systematic (non-random) changes to the residuals. The examination of model performance is in fact based on $e'(x,t)$ calculated using $k'(t)$.

[6] The number of age-specific death rates or entries is mn. For two years, each model requires 2n entries. The log-additive model requires one additional parameter for each additional year, giving m-2 additional parameters for m years, and the number of independent entries is mn-2n-(m-2) = (n-1)(m-2). For the log-linear model, there are no additional parameters, so that there are mn-2n = n(m-2) independent entries.

In order to take account of age-time interactions, a further adaptation of the Lee-Carter method is introduced. The second and possibly higher terms of the SVD are incorporated into the model. The complete expanded model is:

$$\ln[m(x,t)] = a(x) + b_1(x)k_1(t) + b_2(x)k_2(t) + \ldots\ldots + b_n(x)k_n(t)$$

where $b_i(x)k_i(t)$ is referred to as the *i*th rank 1 term in the rank n approximation, where i = 1 to n and n is the number of age groups (Golub and Van Loan 1983**).** In this paper, only the second and third terms are actually incorporated. An iterative procedure is used to progressively refine the fit to $\ln[m(x,t)]$ by adding successive $b_i(x)k_i(t)$ after adjusting $k_i(t)$ to minimise the lack of fit to D(x,t). In progressing from the rank 1 approximation to a rank 2 approximation, the fact that $k_1(t)$ has been adjusted must be taken into account because this adjustment changes the residuals. Similarly, in progressing to a rank 3 approximation, the adjustment of $k_2(t)$ must be accounted for. Each additional term is taken as the rank 1 approximation to the residuals from the previous approximation.

In order to incorporate the second term, an SVD is performed on the residuals of the logarithms of death rates, $\ln[m(x,t)]$, after fitting $b_1(x)k'_1(t)$, that is on $e'_1(x,t) = \ln[m(x,t)] - \ln[m'_1(x,t)]$. The first vectors from this SVD would correspond to the second vectors in the original SVD, apart from the fact that the adjustment has been made to $k_1(t)$. These are denoted by $b^*_2(x)$ and $k^*_2(t)$. Then $k^*_2(t)$ is adjusted to optimise the fit to D(x,t) for each year. The following Poisson regression model is used to find $k^{*\prime}_2(t)$:

$$\log(D(x)) = \log(N(x)) + \log(m'_1(x,t)) + b^{*\prime}_2(x)k^{*\prime}_2(t) + e'_2(x,t)$$

The resulting $k^{*\prime}_2(t)$ is the adjusted value of $k_2(t)$, henceforth referred to simply as $k'_2(t)$. Using $k'_2(t)$, fitted death rates, $m'_2(x,t)$, are calculated. In order to incorporate the third term, this procedure is iterated.

The final adaptation is the rescaling of b(x) for presentation purposes such that they are more readily interpretable and changes in mortality are better understood. This rescaling is based on $k'(t)$, which is also rescaled. The rescaling of $b_1(x)$ and $k'_1(t)$ is done slightly differently to the rescaling of second and higher terms. For the first terms, the average annual change in rescaled $k'_1(t)$ over the fitting period is set equal to 1. Values of $k'_1(t)$ are thus divided by C_1, where $C_1 = \text{mean}_t[k'_1(t+1) - k'_1(t)]$. This implies that when $k'_1(t)$ is roughly linear, the range of rescaled $k'_1(t)$ is approximately equal to the length of the fitting period.[7] These rescaled values are denoted $K'_1(t)$. The $b_1(x)$ are correspondingly rescaled by multiplying by C_1. The rescaled values, denoted $B_1(x)$, describe the annual rate of change at each age corresponding to an average annual change of 1 in $K'_1(t)$. Algebraically, $B_1(x)$ estimates the contribution of the first term[8] to the annual average of $\ln[m(x,t+1)/m(x,t)]$.[9] Since the $B_1(x)$ relate to an average annual change of 1 in $K'_1(t)$, they are comparable over both age and time.

---

[7] Rescaled $k'_1(t)$ are not comparable between different time periods or countries.

[8] This assumes the expanded Lee-Carter model. If the original Lee-Carter model is used, B(x) estimates the annual average of $\ln[m(x,t+1)/m(x,t)]$.

For second and higher terms (i = 2, n), the trend in $k'_i(t)$ is not linear and $b_i(x)$ may range from negative to positive values. Thus, the average annual *absolute* change in rescaled $k'_i(t)$ over the fitting period is set equal to 1. Thus C_i becomes $mean_t[|k'_i(t+1) – k'_i(t)|]$. The $B_i(x)$ are equal to relative change at each age referring to an average annual *absolute* change of 1 in $K'_i(t)$. Algebraically, $B_i(x)$ estimates the contribution of the *i*th term to the annual average of ln[m(x,t+1)/m(x,t)]. The $B_i(x)$ are comparable over time for simple patterns, such as a fairly consistent decrease followed by a fairly consistent increase.

Data

The database for this study is a set of annual age-specific death rates for Australia for 1907 to 1999. Following the approach used by Lee and Carter (1992) and Tuljapurkar *et al.* (2000), data for both sexes combined are used. The age categories are 0, 1-4, 5-9, …., 80-84, 85+. The data were not extended to older ages.

Results

Model fit

The Lee-Carter method with the maximum likelihood adjustment of k(t) was initially applied to the entire data series for 1907-1999. Figure 1 shows k(t) before and after adjustment and the linear fit to k′(t) for the entire period.[10] It is clear that k(t) departs significantly from the linear trend, especially after adjustment. This is also seen in the increase in the additional lack of fit (the difference between total and base lack of fit) in Figure 2. In order to determine the optimal fitting period, R(S) and RD(S) were calculated for all possible fitting periods. Examination of R(S) and RD(S) in Figures 3 and 4 shows that 1968 is the pivotal year. The addition of the preceding year marks the beginning of a period of substantial increase in both indicators of lack of fit. Figure 3 shows that including preceding years increases R(S) consistently, while Figure 4 demonstrates that the marginal effect of including 1967 (and preceding years) is large relative to more recent years.

Figures 3 and 4 also show that R(1950) is close to the maximum and RD(1950) is high with even higher values to the right, indicating that the ARIMA fit for the period 1950-99 is among the least satisfactory. Extension of the fitting period towards a starting year of 1907 results in lower R(S) and RD(S) values, but this is almost entirely due to the increase in the base lack of fit while the additional lack of fit remains large. Compared to 1968-99, 1907-99 results in a relatively poor fit to *both* the log-additive Lee-Carter model and the linear ARIMA model. Violation of the assumption of invariant b(x) may contribute to the lack of fit of the log-additive model for these longer periods. As seen in Figure 5, b(x) differs substantially by fitting period.

The b(x) profiles presented in Figure 5 are shown as rescaled values, B(x), in Figure 6. These rescaled values show the annual average rate of change in mortality at each

---

[9] Since m(x,t+1)/m(x,t) is close to 1 and $B_1(x)$ close to 0, [1 + $B_1(x)$] approximates m(x,t+1)/m(x,t) (for example, $B_1(x)$ = -0.01 is equivalent to m(x,t+1)/m(x,t) = 0.99). Thus in a graph of $B_1(x)$, such as Figure 5, the larger negative values of $B_1(x)$ correspond to greater reductions in mortality.

[10] The linear fit, which is roughly equivalent to the regression of k′(t) on time, is a line of slope *d* through the mean of k′(t).

Figure 1 k(t), adjusted k(t) and linear fit to adjusted k(t), 1907-99

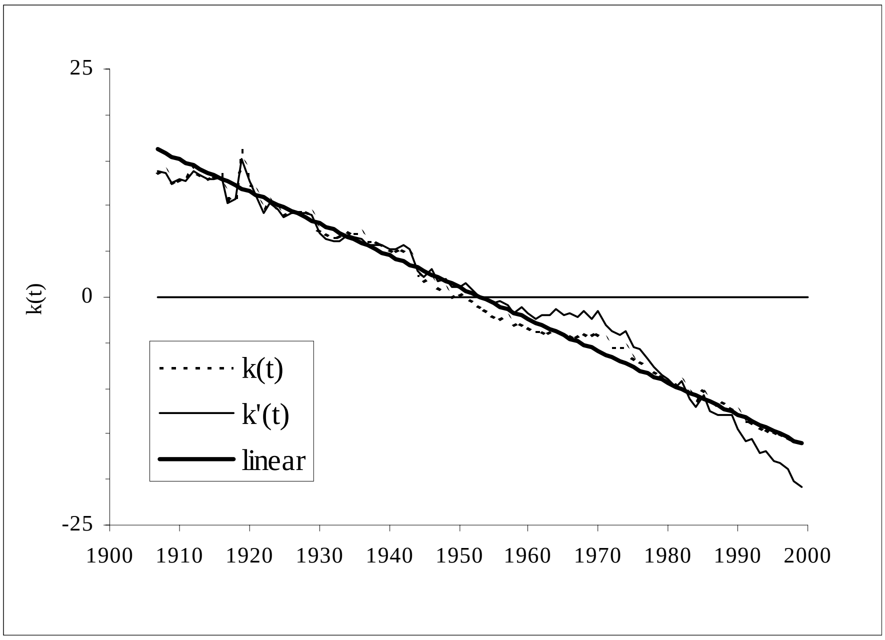


Figure 2 Mean-χ^2 statistics of total and base lack of fit for periods 1907-99 to 1998-99

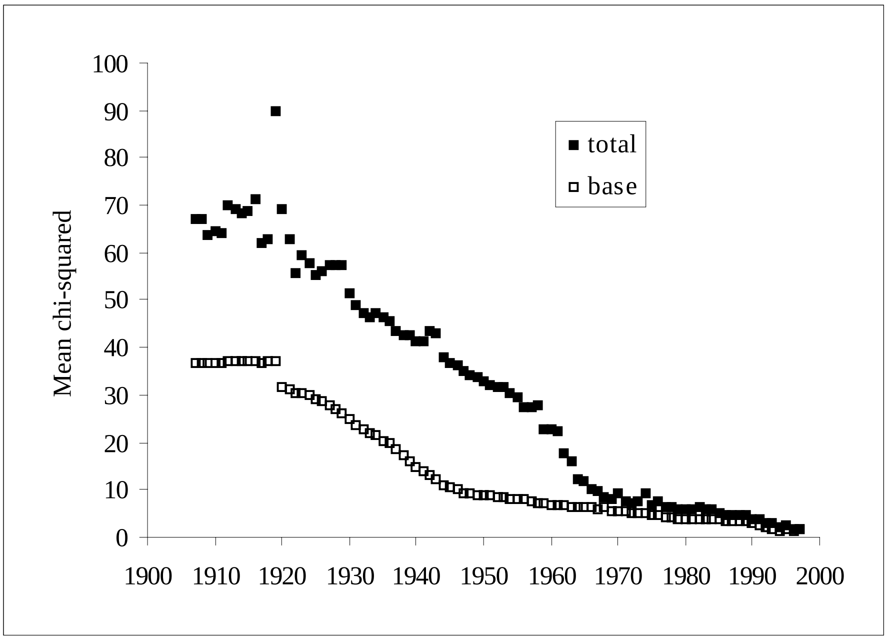


Figure 3 Ratio of total to base lack of fit, R(S), for S = 1907 to 1997

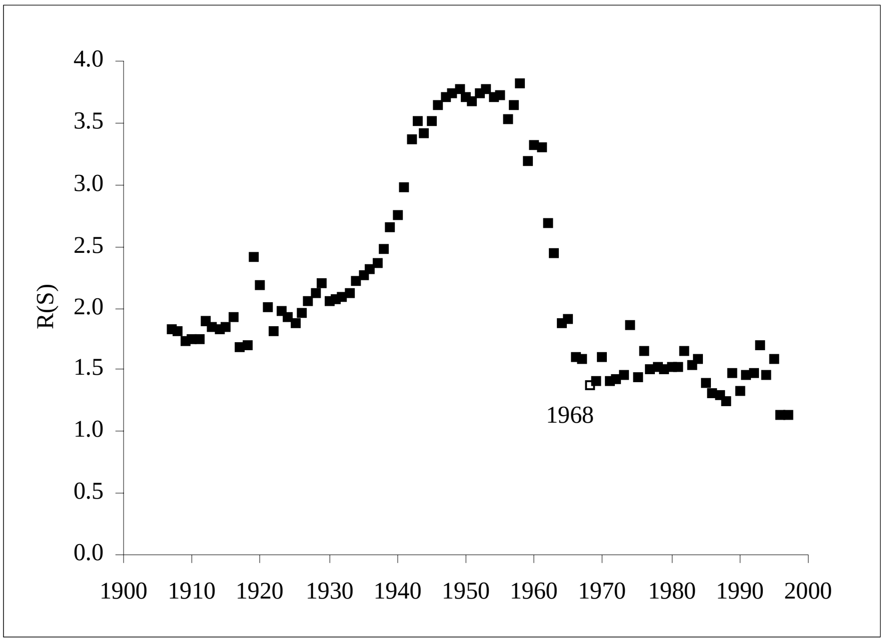
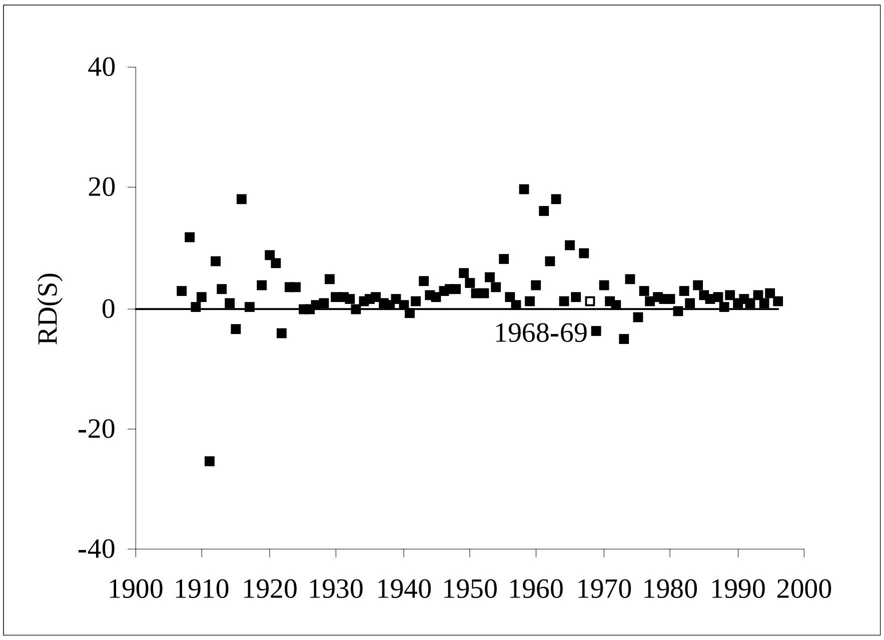
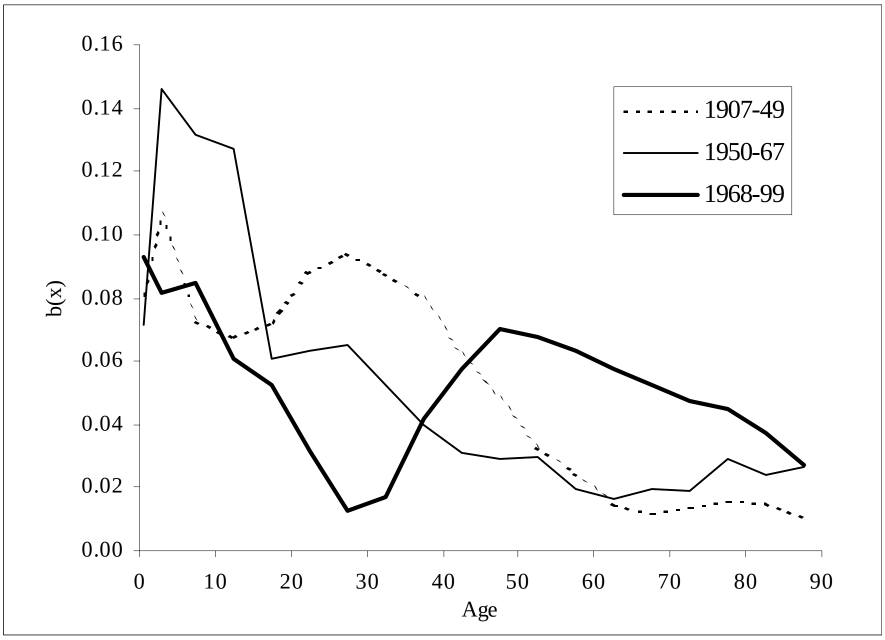


Figure 4 Ratio of differences in total and base lack of fit, RD(S), for S =1907 to 1996



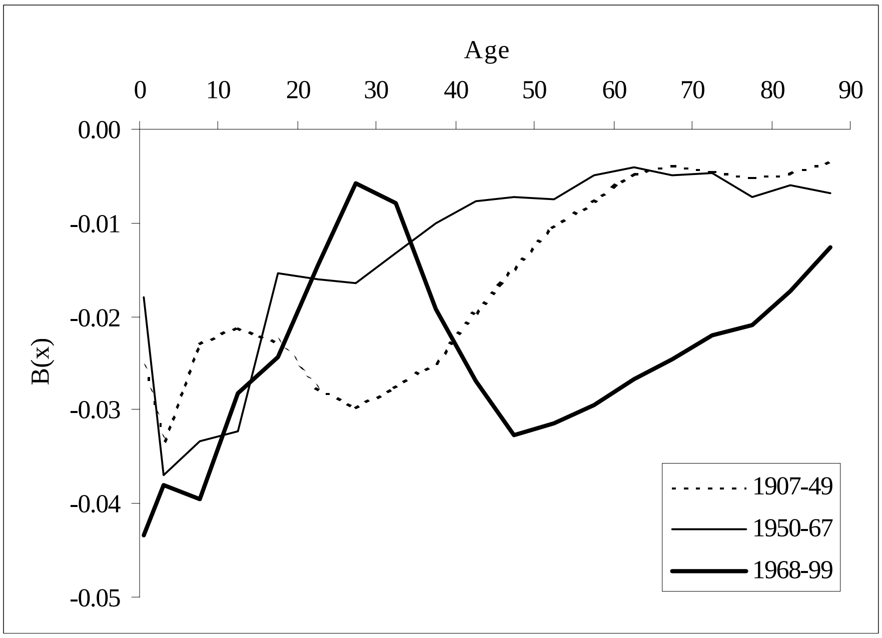
Note: RD(1918) = -85 not shown. The post-World War 1 influenza epidemic occurred in 1919 in Australia.

Figure 5 b(x) values for selected fitting periods, 1907-99



Note: b(x) is not comparable over time.

Figure 6 Rescaled b(x) values, B(x), for selected fitting periods, 1907-99



age corresponding to an annual average change in K′(t) for the consecutive periods, 1907-1949, 1950-67 and 1968-99. Since these are comparable over age and time, they demonstrate the way in which age-specific mortality changed over the course of the twentieth century. It is clearly seen that the greatest relative reductions in death rates occurred in childhood, particularly in the most recent period. Relative reductions in early adulthood declined considerably over time and in the most recent period are close to zero. In contrast, relative reductions at older ages increased dramatically.

For the optimal fitting period, 1968-99, the assumption of invariant b(x) holds fairly well. This is seen in Figure 7 which shows b(x) values for subperiods of the optimal fitting period. Values of k(t), k′(t) and the linear fit to k′(t) are shown in Figure 8, and corresponding values of a(x) and b(x) are shown in Figures 9 and 10. As is to be expected, the linear fit to k′(t) is clearly better than for 1907-99. The first term, b(x)k′(t), is shown in Figure 11 as a surface of age by time effects. Since k′(t) is roughly balanced about zero, positive effects in the early part of the period are balanced by negative effects in the later part. A large b(x) produces greater mortality reductions at age x (for example at age 0), and a small b(x) produces smaller reductions (for example at age 25-29).

Forecasts

Based on the optimal fitting period, k′(t) was extrapolated to 2050. This year was chosen for illustrative purposes. The extrapolated values are simply referred to as k(t). These k(t) and their 95 per cent confidence intervals are shown in Figure 12. They translate into the e(0) values and confidence intervals shown in Figure 13. Life expectancy in 1999 was 79.6 years. By 2050, the forecast value is 92.0 years with a 95 per cent confidence interval of 85.1 to 99.8 years. The asymmetry in this confidence interval is due to the fact that the transformation from k(t) to e(0) is highly non-linear.

Since incorporating choice of fitting period into the Lee-Carter method is designed to improve model fit with a view to increasing the validity of forecasts, it is of interest to compare the above forecast with forecasts based on longer fitting periods. Since that scale of k(t) is arbitrary, the drift and standard errors cannot be compared between fitting periods. Forecasts of e(0) and confidence intervals are shown in Figure 14.

The 1968-99-based forecast of e(0) is compared with the most recent official projection[11] for Australia (ABS 2000) in Figure 15. It is seen that this Lee-Carter forecast is more optimistic than the official projection. From 2027, the official projection falls within the Lee-Carter 95 per cent confidence interval for almost the entire forecast period. The official projection is also lower than the forecast based on 1950-99. When compared with the forecast based on 1907-99, the official projection is higher to 2027 and thereafter lower.

[11] The official projection (ABS 2000) is for 1997-2051 and is based on the following assumptions. Life expectancy was assumed to increase during 1997-2002 at the same rate as in 1986-96. From 2003, the rate of increase was gradually reduced to zero in 2051. This yields a life expectancy in 2051 of 84.9 years. The observed age pattern of change 1971-1996 was assumed to continue for 1997-2027, except that any increases in mortality were assumed not to re-occur. From 2028, the rate of change was held constant over age. Further adjustment to the pattern of change was made to achieve the predetermined mortality levels.

Figure 7 b(x) values for selected fitting periods, 1968-99

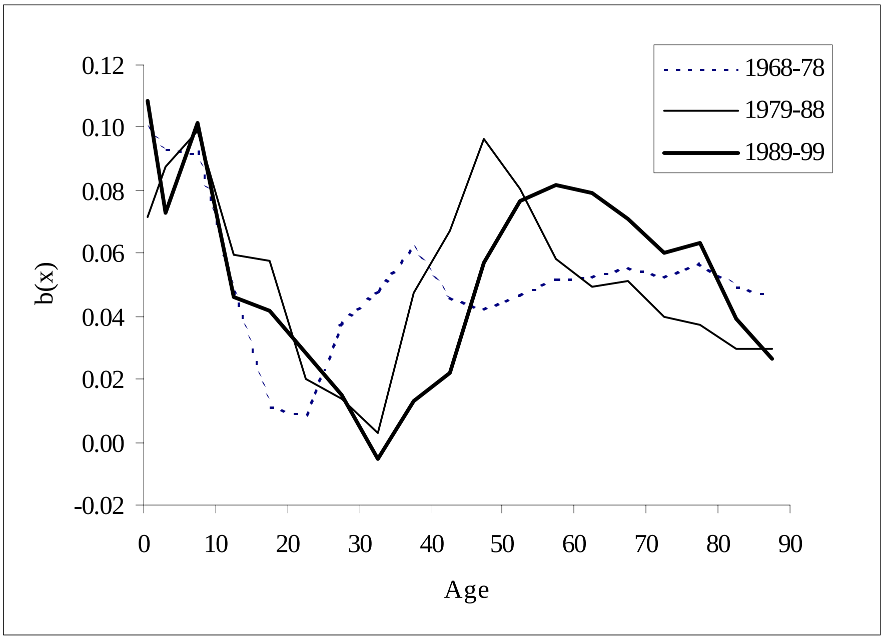


Figure 8 k(t), adjusted k(t) and linear fit to adjusted k(t), 1968-99

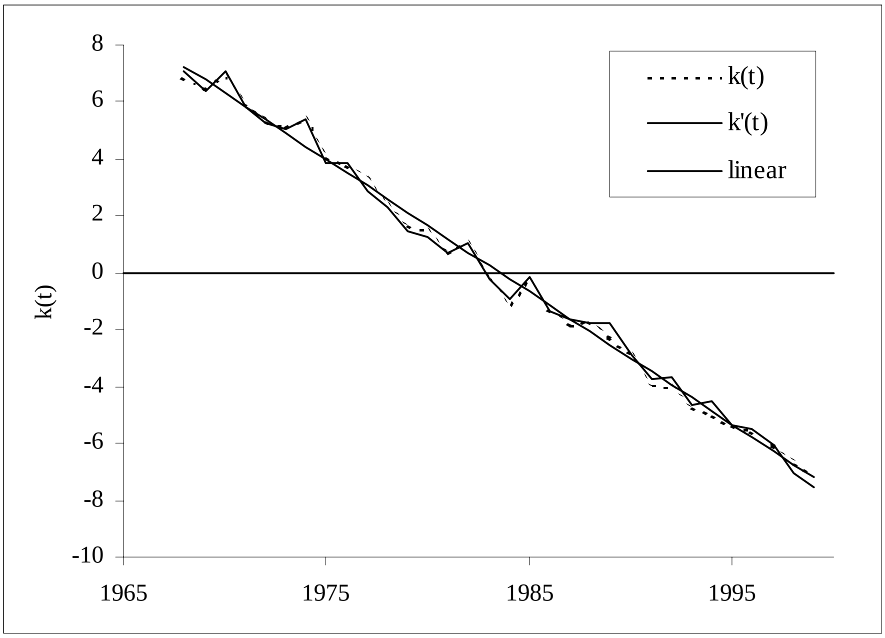


Figure 9 a(x) values, 1968-99

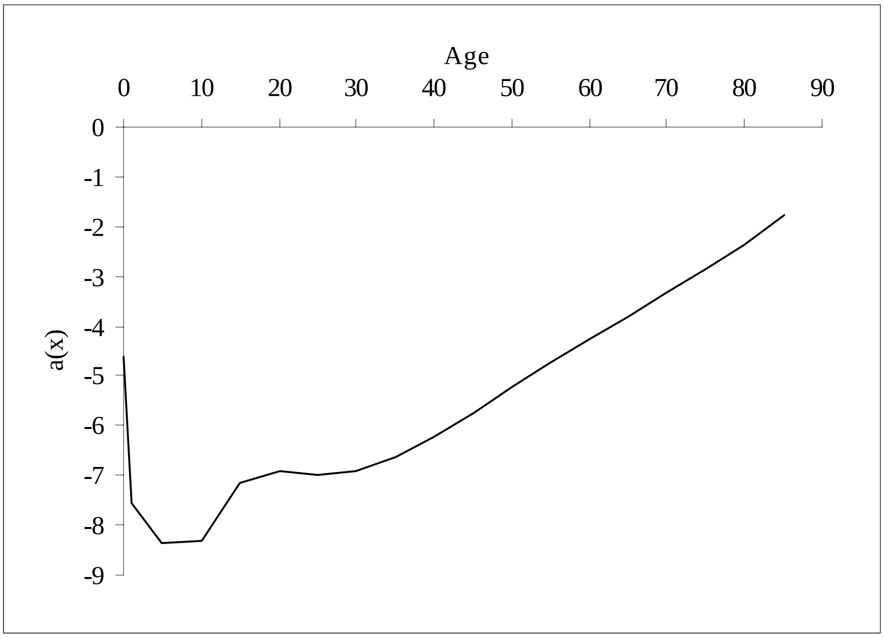
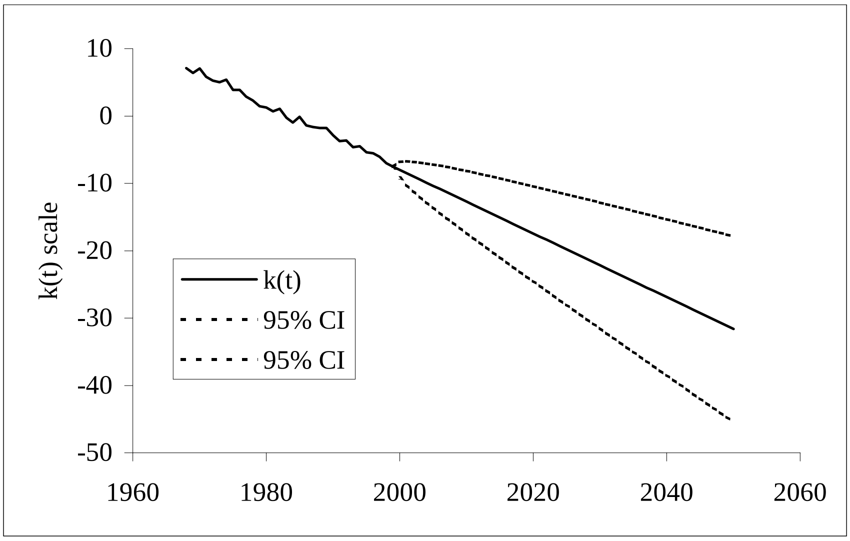


Figure 10 b(x) values, 1968-99

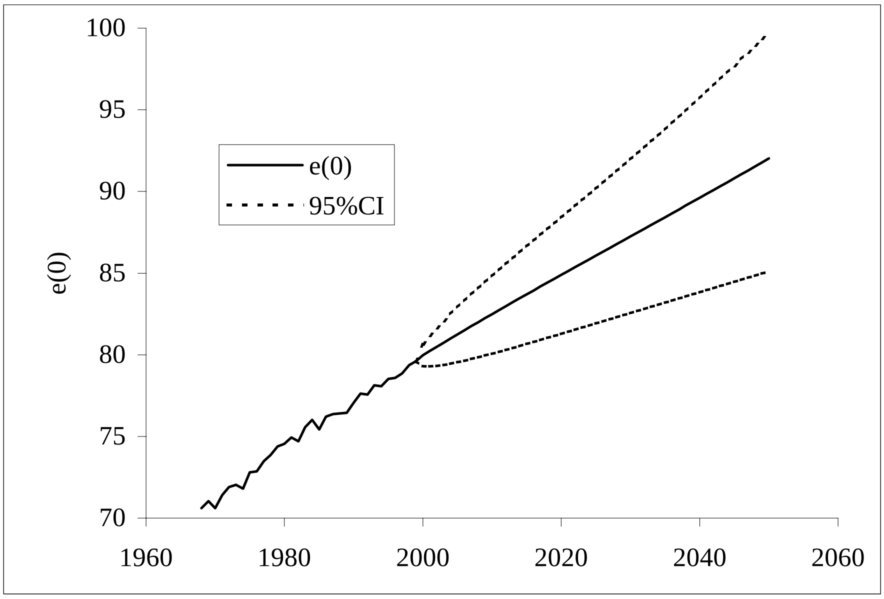
Figure 11 Surface of age by time effects, b(x)k′(t), 1968-99

Figure 12 Forecast values of k(t) and 95 per cent confidence interval for 2000-50 based on 1968-99



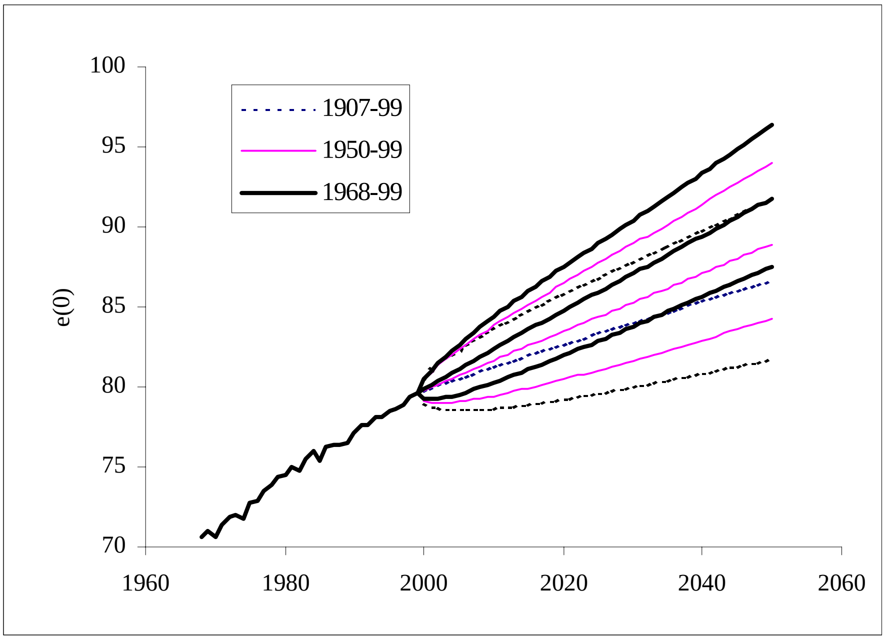
Note: For 1968-99, k′(t) is shown.

Figure 13 Forecast values of e(0) and 95 per cent confidence interval for 2000-50 based on 1968-99



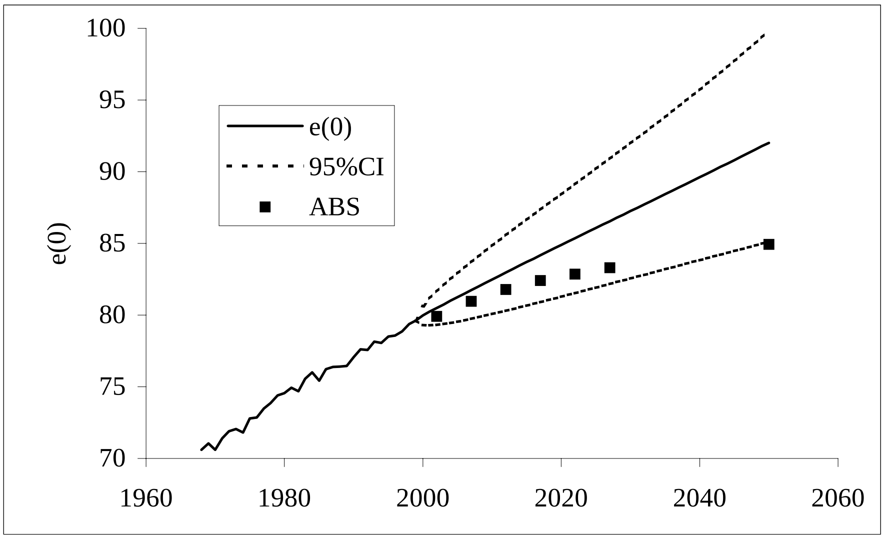
Note: For 1968-99, actual e(0) is shown.

Figure 14 Comparison of forecast values of e(0) and 95 per cent confidence intervals based on selected fitting periods



Note: For 1968-99, actual e(0) is shown.

Figure 15 Comparison of recent official projection with forecast values of e(0) based on 1968-99



Note: For 1968-99, actual e(0) is shown.
Source: Official mortality projection from ABS (2000) *Population Projections, Australia, 1999-2101*. Canberra. Catalogue No. 3222.0.

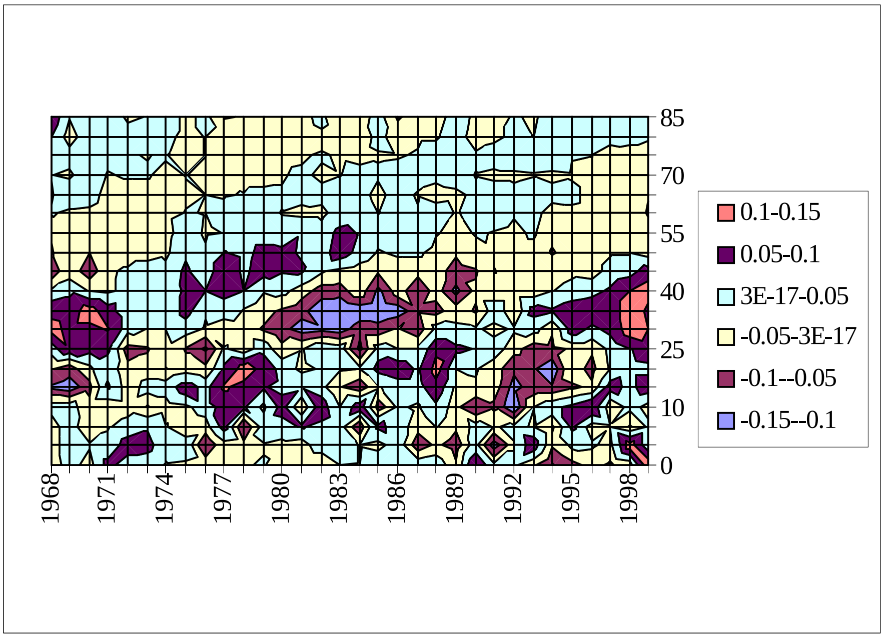
Model performance

The above forecasts are based on the rank 1 approximation of the Lee-Carter model which includes only the first term, $b_1(x)k_1(t)$, of the rank n approximation or complete model. Figure 16 shows the residuals from this rank 1 approximation. It is seen that they exhibit a systematic pattern which appears to be a cohort effect. When expressed as residual death rates, the effect remains (see Figure 17). This effect is in fact rather too steep to be purely a cohort effect and might better be described as a cohort-period effect. As such, it clearly demonstrates the presence of age-time interaction.

In order to try to take account of this age-time interaction, the expanded Lee-Carter model was used with successive incorporation of the second and higher terms. After the second term was incorporated, the residuals appeared much less systematic, as seen in Figure 18, suggesting that the second term largely models the cohort-period effect. The corresponding residual death rates (Figure 19) show a similar lack of systematic pattern. However, the residuals cannot be described as random since marked clustering occurs. The third term was thus incorporated, resulting in the residuals seen in Figures 20 and 21. While these show no systematic pattern, some clustering still occurs. It should be noted that, as seen in Figures 16, 18 and 20, the largest residuals occur at younger ages, reflecting both the greater irregularity in death rates at these ages and the smaller weights at ages with smaller numbers of deaths. It is possible that this effect is at least in part due to the adjustment of $k_i(t)$.

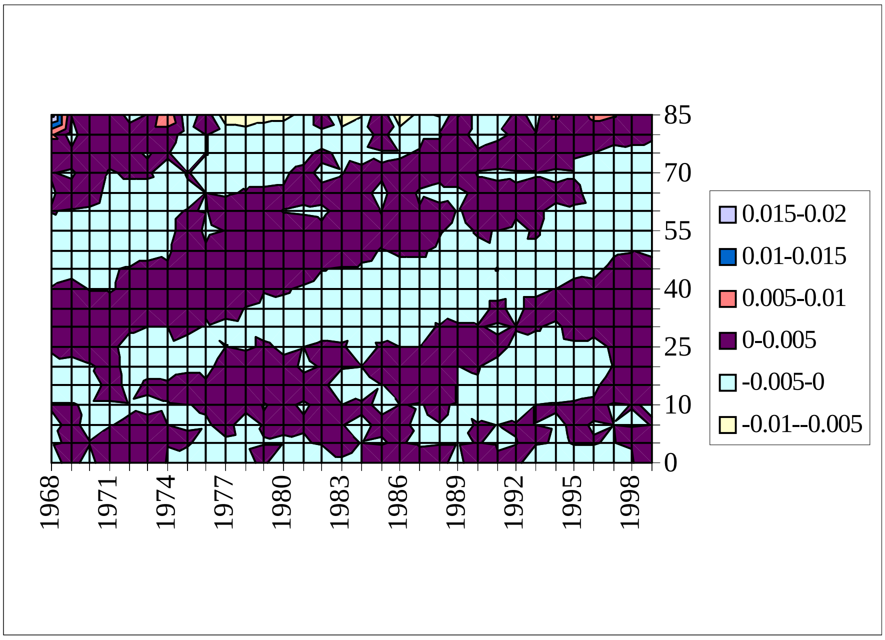
Having identified an age-time interaction modelled by the second term, this term should ideally be taken into account in the forecasting methodology. This would involve the extrapolation of $k'_2(t)$, shown in Figure 22, but this is not straightforward because its orthogonal functional form is constrained within the fitting period. The trend in $k'_3(t)$ is seen in Figure 23.

Figure 16 Residuals after fitting the first term



Note: These residuals are on a logarithmic scale.

Figure 17 Residual death rates after fitting the first term



Discussion

This research has addressed two issues in the application of the Lee-Carter method to Australian mortality, namely the departure from the 'universal pattern' of constant exponential decline over the period 1950-94 identified by Tuljapurkar *et al.* (2000:789) for the G7 countries and the failure to satisfy the assumption of a constant age pattern of mortality decline, b(x). These issues are central to the accuracy of the forecasts. Only when the assumption of invariance in the age component is satisfied will the time component be an accurate indicator of the level of, and hence change in, overall mortality. Further, only when a good fit of the ARIMA (in this case, linear) model is obtained will the past fitting period be a reliable guide for the future.

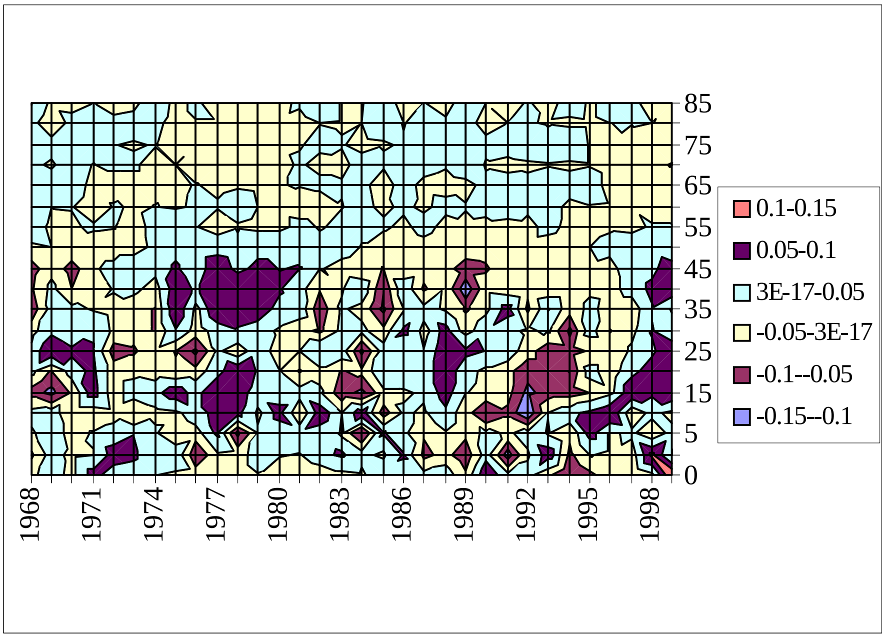
While the focus has been on Australian mortality, these issues are also relevant to other developed countries. The evidence provided by Tuljapurkar *et al.* for the G7 countries and by Lee and Carter for the US in fact demonstrates similar but less pronounced patterns of deviation from linearity in k(t) to that found for Australia. Deviations from linearity are also evident in the 1947-99 data for Austria analysed by Carter and Prskawetz (2001). The fact that the age pattern of mortality decline varies over time has been demonstrated in a range of countries (Kannisto, Lauritsen and Thatcher 1994, Horiuchi and Wilmoth 1998, Wilmoth 1998, Lee and Miller 2000, Carter and Prskawetz 2001). The methodology developed in this paper to take account of these issues could in principle be applied to any population.

Since these two issues relate separately to the log-additive and ARIMA models of the Lee-Carter method, they are addressed separately rather than simultaneously. In this research, the issue of linearity in k(t) was addressed first. The approach was to choose a fitting period for which the assumption of linear k(t) holds good and then to examine variability in b(x) and the performance of the log-additive model over this period. This is much more straightforward than the reverse-order procedure of choosing a fitting period to satisfy the assumption of invariant b(x) and then examining the performance of the linear model for this period. Choosing the fitting period based on invariance in b(x) is problematic due to the lack of a theoretical model for b(x).[12]

The issue of linearity in k(t) is the focus of the main methodological adaptation, that of choosing an 'optimal' fitting period to give a good fit to the linear time series model. In doing so, it is assumed *a priori* that adjusted k(t) is linear. This is equivalent to assuming that age-specific death rates decline exponentially at a constant rate (given that b(x) is assumed to be invariant). Exponential decline is the paradigm within which the Lee-Carter method has been previously applied (Carter and Lee 1992, Lee 2000, Tuljapurkar *et al.* 2000, Carter and Prskawetz 2001). It is the simplest model to describe mortality change and generally conforms to observation. A linear trend in k(t) is also convenient for extrapolation.

[12] Carter and Prskawetz (2001) attempted to choose an optimum fitting period based on a heuristic examination of variation in b(x), but without success. They did not consider the issue of linearity in k(t).

Figure 18 Residuals after fitting the first and second terms



Note: These residuals are on a logarithmic scale.

Figure 19 Residual death rates after fitting the first and second terms

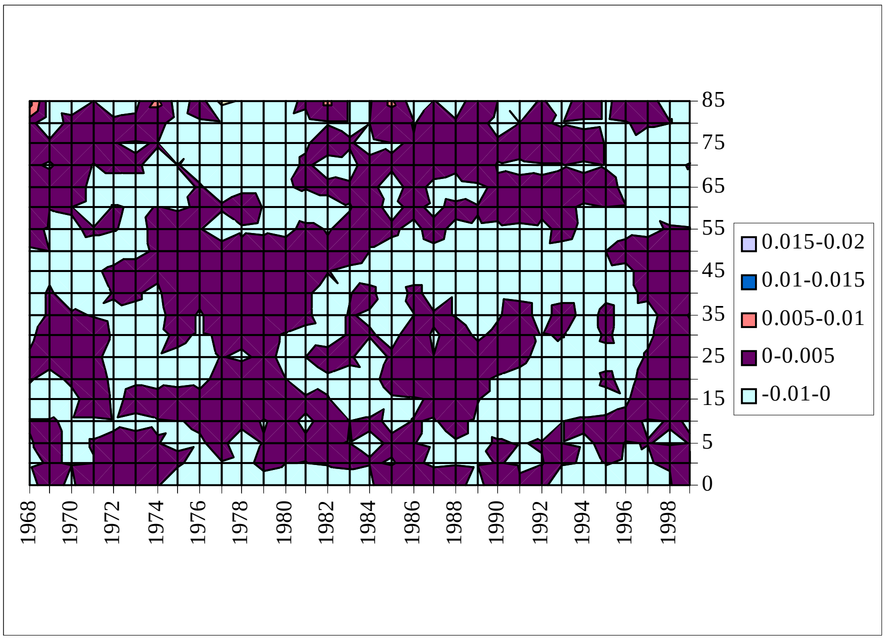
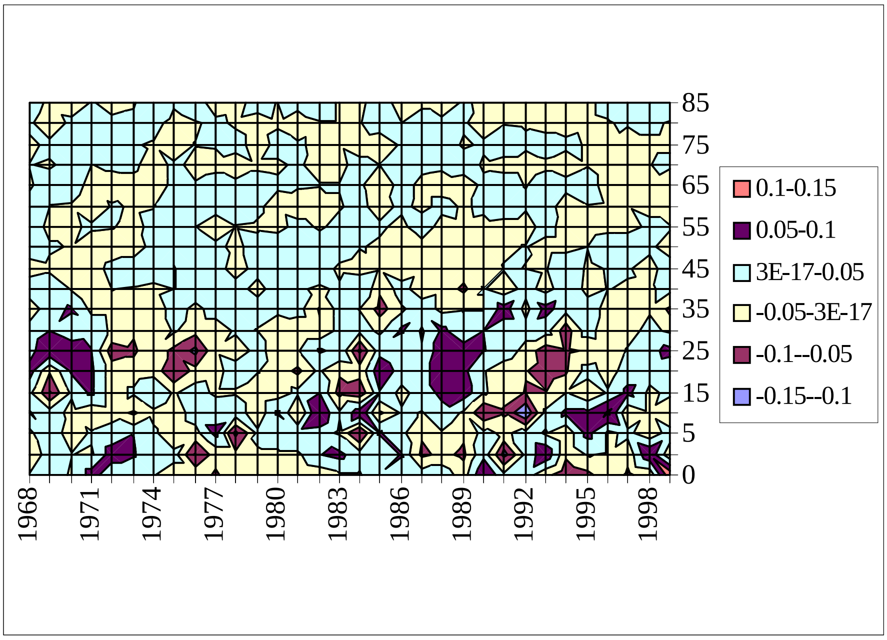


Figure 20 Residuals after fitting the first, second and third terms



Note: These residuals are on a logarithmic scale.

Figure 21 Residual death rates after fitting the first, second and third terms

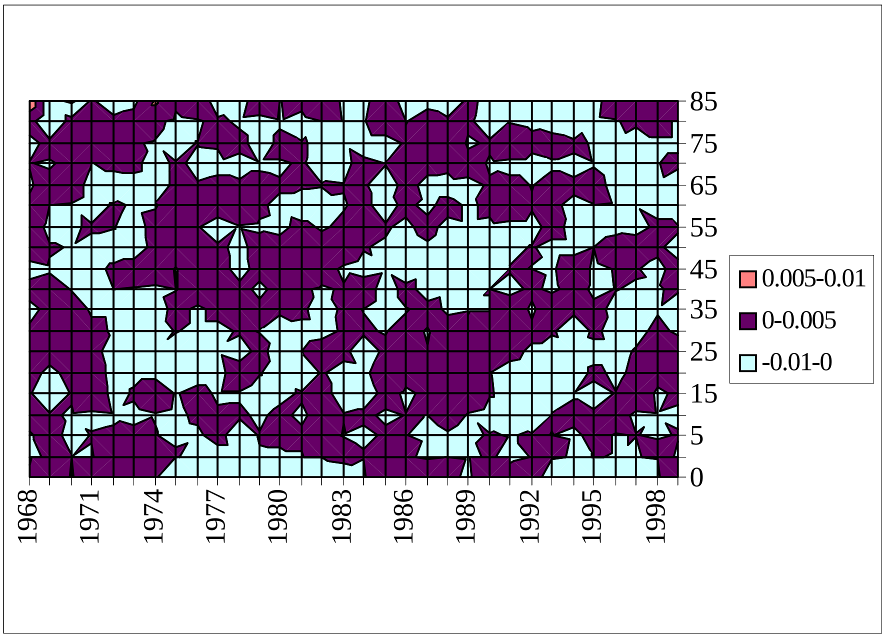


Figure 22 Adjusted $k_2(t)$ based on 1968-99

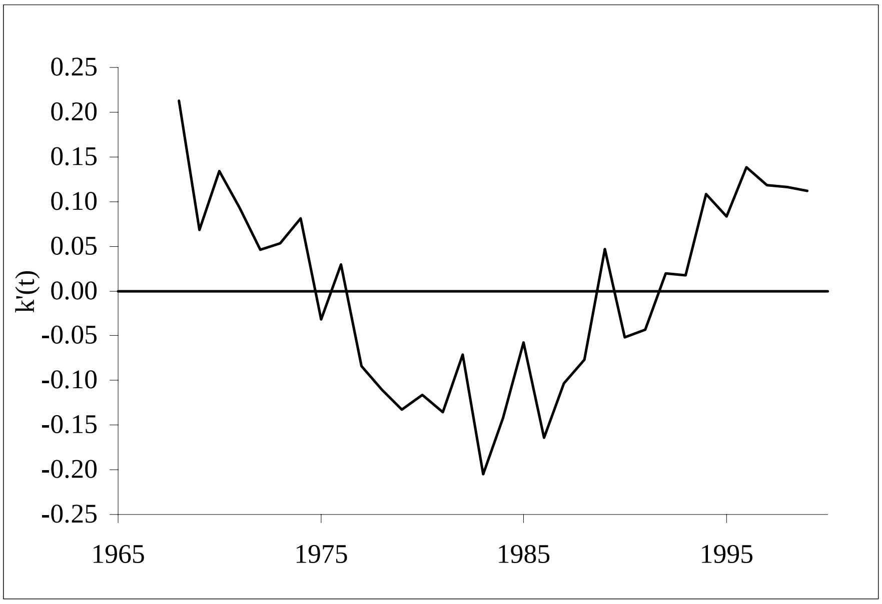
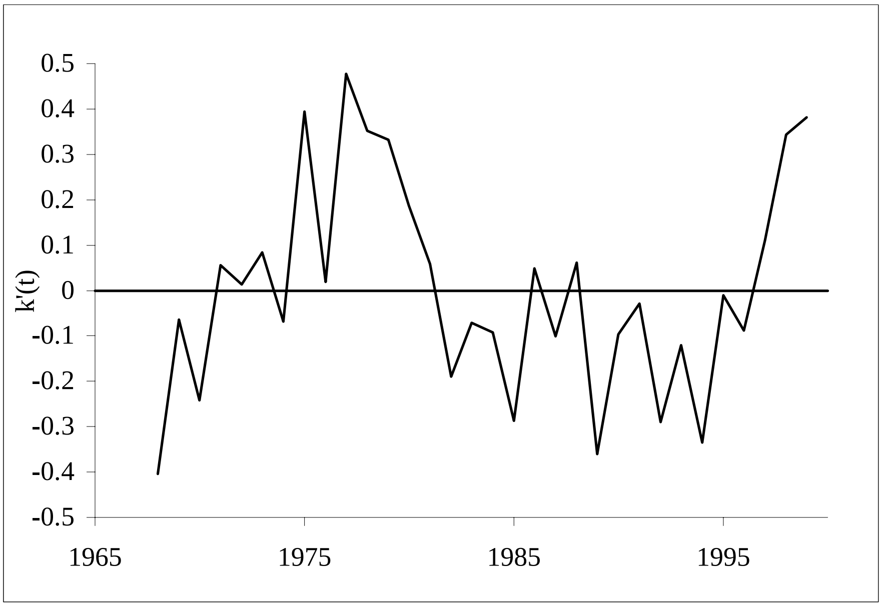


Figure 23 Adjusted $k_3(t)$ based on 1968-99



While provision is made in the Lee-Carter method for applying a non-linear model to k(t) (Lee and Carter 1992:663), it would appear from the literature that this possibility has yet to be pursued.[13] Lee and Miller (2000:4) report that for at least 10 national data sets k(t) is 'highly linear'. This linearity contributes to the method's simplicity and hence its popularity (Alho 2000, Lee and Miller 2000, Carter and Prskawetz 2001). Indeed, the simplest approach is to assume that k(t) is linear no matter what the period of available data, in other words without consideration of either non-linear models of k(t) or of more appropriate fitting periods. The danger is that this approach will be adopted at the expense of forecasting accuracy. It is argued here that this computational simplicity belies the complex issues involved in the choices inherent in applying the method, in particular the choice of fitting period.

The methodological adaptations developed in this paper are designed to strengthen the technical basis of the Lee-Carter method. The main adaptation enhances the technical rationale of the method by establishing objective criteria for the choice of fitting period. The adaptation of adjusting k(t) by fitting to the age distribution of deaths, rather than to total deaths or life expectancy, also strengthens the technical basis of the method, and makes a difference to the fit. The use of k′(1999) based on fitting to D(x,t) as the jump-off value for extrapolation (rather than k′(1999) based on fitting to D(t)) is a further technical enhancement, the effect of which may be substantial. Lee and Miller (2000) used observed death rates for the jump-off year, which made a difference of 0.6 years in the life expectancy value for the jump-off year. The use of observed rates has the disadvantage of extrapolating any idiosyncrasies that may be present.

The effect of limiting the fitting period to 1968-99 in the Australian case was to increase the applicability of the log-additive model: the variation in b(x) was significantly reduced.[14] Thus by choosing the optimal fitting period, the assumption of invariant b(x) was better met and the problem associated with violation of this assumption was largely avoided. The reduced variation in b(x) for 1968-99 suggests that age-time interactions for the optimal period are considerably less than for 1907-99. However, it was shown through examination of the residuals based on only the first term that interaction exists. Modelling age-time interactions through the incorporation of the second term would seem desirable given that it appears to be primarily a cohort effect, but its incorporation in forecasting would be highly problematic. Further research is needed to fully understand the contribution of the second and possibly higher terms in order to incorporate them into the forecasts.

It is recalled that the solution proposed by Lee and Miller to violation of the assumption of invariant b(x) over longer periods was the adoption of 1950 as the starting year for the fitting period. In the case of Australia, 1950-99 encompasses considerable variation in b(x) (as seen in Figure 5) and is among the least satisfactory of possible fitting periods in terms of the additional lack of fit resulting from the imposition of the linear time series model (see Figures 2 to 4). It is likely that this is also true of other developed countries since mortality patterns have been subject to

[13] Lee and Nault (1993) found a superior model was obtained by the addition of a moving average or autoregressive term, but this made almost no difference to the forecast.

[14] It is noted that reduced variation in b(x) does not necessarily follow from 'optimising' the fitting period, but is likely to occur in practice in developed countries given known changes to mortality patterns over the twentieth century.

similar influences. Carter and Prskawetz (2001) demonstrated that in the case of Austria significantly different patterns of b(x) are obtained for different fitting periods.

In order to take account of this variation in b(x), Carter and Prskawetz (2001) proposed ‘an extension of the original Lee-Carter method that takes into account the changing age pattern of mortality….[and] can be applied to locate structural changes in the historical pattern of mortality that could ultimately be used to determine the time horizon of past mortality rates on which forecasts should be based’ (p.2). Their ‘extended’ method involves the application of the Lee-Carter method to the 30 overlapping 24-year subperiods of their 53-year dataset. They conclude that their analysis ‘should be taken as evidence that the choice of the base period may not be negligible’ (p12) in forecasting. However, they ‘refrain from drawing any definite conclusions about the optimal period on which to base forecasts’ (p.12).

The approach taken by Carter and Prskawetz (2001) may be contrasted with the approach of the present research. Carter and Prskawetz chose to address the issue of invariant b(x) and aimed to choose a fitting period by examining the behaviour of b(x). The issue of linearity in k(t) was not addressed and it was implicitly assumed that k(t) was linear for all fitting periods. That they were unable to draw definite conclusions might be attributed to the heuristic nature of their approach, which may derive from the difficulties already noted arising from the lack of a theoretical model for b(x). In fact, it is not immediately clear how their methodology would allow identification of the optimal period. In the present paper, an independently-developed methodology for determining the ‘optimal’ fitting period is proposed and is shown empirically to better satisfy the assumption of invariance in b(x) while ensuring a better fit to the linear model of k(t).

The process of statistically determining the fitting period may be contrasted with the common practice of taking the maximum period for which data are available, based on a belief that data should not be discarded and implying that all experience is useful and relevant. This practice appears to have been employed by Lee and Carter (1992:659) for the US and by Tuljapurkar *et al*. (2000:790) for the G7 countries,[15] though in both cases it is clear that their approach is to base long-term forecasting on long-term trends (Lee and Carter 1992:668; Tuljapurkar *et al.* 2000:790-1). Lee and Carter make forecasts for 76 years (to 2065) based on data for 90 years (1900-89), while Tuljapurkar *et al.* make forecasts for 56 years (to 2050) based on data for 45 years (1950-94). This is in keeping with convention: it is generally the case that long-term trends are used as a basis for long-term forecasts and short-term trends for short-term forecasts. The inappropriateness of basing long-term forecasts on short-term trends stems from the possibility that the short-term trend may be dominated by temporary deviations from a more persistent underlying trend. Similarly, it may be inappropriate to base short-term forecasts on long-term trends.

Whether a long- or short-term approach is adopted, the central issue in relating past experience to future forecasts is theoretical and substantive. How relevant is average past experience to the future? In taking a long-term approach, it is implicit that past

[15] The need to readily obtain comparable data would have curtailed the period.

long-term trends will be of continuing relevance to the long-term future. A short-term approach implies that only recent experience is relevant for the short-term future. In the case of Australia, the optimal choice of the 32-year fitting period implies the adoption of a relatively short-term approach. In basing forecasts on 1968-99, the experience of the early part of the century and the mortality stagnation of the 1960s are not considered relevant for future mortality. Rather, only the relatively recent experience of rapid decline is taken into account.

For how long will this experience continue to be relevant? While in this paper, a forecast has been made for 51 years (to 2050) based on only 32 years of experience, this is for illustrative and comparative purposes only. Convention and caution would dictate that the length of the forecasting period be rather shorter than 51 years, and a maximum of 32 years would seem more appropriate. The need for caution in determining forecasting length is underlined by two important considerations. The first is the fact that a continued rapid decline will lead to forecasts of life expectancy that are beyond what is generally considered feasible (see, for example, Olshansky, Carnes and Desesquelles 2001). The forecast of 92 years in 2050 is indeed high in relation to other forecasts for developed countries (United Nations 2001). The second cause for caution is the width of the confidence interval. The fact that the optimal fitting period is chosen on the basis of goodness-of-fit criteria to the linear model results in relatively small deviations from this short-term linear trend (see Figure 7) but the shorter fitting period results in a more rapid widening of the confidence interval. While it may be reasonable to expect small deviations in the immediate future, such an expectation becomes increasingly tenuous as the forecasting period extends.

In the short term, the forecast based on 1907-99 has the widest confidence interval. This is due to the fact that the relatively large past deviations from the long-term trend (see Figure 1), which comprise both random error and systematic deviations from the linear model, are treated as entirely random. Thus the implication is that future random deviations from the forecast trend could be as great as the systematic deviations due to the stagnation of the 1960s and the accelerated decline in recent decades. In other words, the wide confidence interval reflects in part the fact that the linear model is not entirely appropriate for this longer series of data.

Compared with official projections for Australia, the forecast based on 1968-99 gives more optimistic results. This is to be expected since only the recent trend of rapid decline is taken into account in the forecast. However, this relative optimism is also true of the long-term forecasts based on 1950-99 and in the longer term 1907-99. It is noted that the forecasts of life expectancy for the G7 countries made by Tuljapurkar *et al.* (2000), which are based on 1950-94, are also higher than official projections for all or the later part of the forecasting period. Forecasts for Austria of sex-specific life expectancy to 2050 based on 1947-99 also indicate a higher both-sexes life expectancy than official projections throughout the forecasting period (Carter and Prskawetz 2001). For the US, Lee and Carter (1992) forecast life expectancy based on 1900-89 to be 5.6 years higher in 2065 than the official projection. Combined with historical demonstrations both of the greater accuracy of the Lee-Carter method over official methods and of a systematic tendency of the Lee-Carter method to under-predict gains in life expectancy in the long term (Bell 1997, Lee and Miller 2000), these more optimistic results suggest that current official projections are highly likely to underestimate future mortality decline particularly in the long term. This will have far-reaching implications for a range of mortality applications including projections of future population size and structure.

To conclude, the Lee-Carter methodology has been adapted to take departures from constant exponential decline into account by restricting the fitting period. The optimum fitting period is based on criteria of lack of fit of the log-linear model. In the case of Australia, the assumption of an invariant age pattern of mortality decline is better satisfied for the optimum period. A methodology for detecting and modelling age-time interactions has been proposed. However further research is needed to better understand age-time interactions with a view to their incorporation into forecasts of future mortality. This might usefully include analysis by sex and cause of death.

Acknowledgments

This paper was presented at the Population Association of America 2001 Annual Meeting, Washington DC, 31 March. The analysis was undertaken using purpose-written code in R. Thanks are due to Rajesh Chauhan for research assistance and to Leonie Tickle for valuable comments.

References

Alho, JM (2000) Discussion. *North American Actuarial Journal* 4(1): 91-93.

Australian Bureau of Statistics (ABS) (2000) *Population Projections, Australia, 1999-2101*. Canberra. Catalogue No. 3222.0.

Australian Institute of Health and Welfare (AIHW) (2000) *Australia's Health 2000. The seventh biennial report of the Australian Institute of Health and Welfare.* Canberra. AIHW Catalogue No. 19.

Bell, WR (1997) Comparing and assessing time series methods for forecasting age specific demographic rates. *Journal of Official Statistics* 13(3):279-303.

Booth, H, Chauhan, RK, Maindonald, J and Smith, L (2000) *The future of the Australian mortality transition.* Paper presented at the 2000 Biennial Conference of the Australian Population Association, Melbourne, 29 November – 1 December.

Carter, LR and Lee, RD (1992) Modeling and forecasting US sex differentials in mortality. *International Journal of Forecasting* 8: 393-411.

Carter, LR and Prskawetz, A (2001) *Examining structural shifts in mortality using the Lee-Carter method.* MPIDR Working Paper WP 2001-007. Rostock: Max Planck Institute for Demographic Research.

Diggle, PJ (1990) *Time Series: A Biostatistical Introduction*. Oxford: Clarendon Press.

Golub, GH and Van Loan, CF (1983) *Matrix Computations*. Baltimore: Johns Hopkins.

Heligman, L and Pollard, JH (1980) The age pattern of mortality. *Journal of the Institute of Actuaries* 107:49-80.

Horiuchi, S and Wilmoth JR (1998) Deceleration in the age pattern of mortality at older ages. *Demography* 35(4): 391-412.

Kannisto, V, Lauritsen, J, Thatcher, AR and Vaupel, JW (1994) Reductions in mortality at advanced ages: several decades of evidence from 27 countries. *Population and Development Review* 20(4): 793-810.

Kuan, C-M and Hornik, K (1995) The generalized fluctuation test: a unifying view. *Econometric Reviews* 14: 135 - 161.

Lee, RD (2000) The Lee-Carter method for forecasting mortality, with various extensions and applications. *North American Actuarial Journal* 4(1): 80-91.

Lee, RD and Carter, LR (1992) Modeling and forecasting US mortality. *Journal of the American Statistical Association* 87(419): 659-671.

Lee, RD and Miller, T (2000) *Assessing the performance of the Lee-Carter approach to modeling and forecasting mortality*. Paper presented at the 2000 Meetings of the Population Association of America, Los Angeles, 23-25 March.

Lee, RD and Nault, F (1993) *Modeling and forecasting provincial mortality in Canada*. Paper presented at the World Congress of the International Union for the Scientific Study of Population, Montreal.

Maindonald, JH (1984) *Statistical Computation*. New York: Wiley.

Olshansky, SJ, Carnes, BA and Desesquelles, A (2001) Prospects for human longevity. *Science* 291(5508): 1491-1492.

Tabeau, E (2001) A review of demographic forecasting models for mortality. In E Tabeau, A van den Berg Jeths and C Heathcote (eds) *Forecasting Mortality in Developed Countries: Insights from a Statistical, Demographic and Epidemiological Perspective.* Dordrecht: Kluwer Academic Publishers.

Trefethen, LN and Bau, D (1997) *Numerical Linear Algebra*. Philadelphia: Society for Industrial and Applied Mathematics.

Tuljapurkar, S, Li, N and Boe, C (2000) A universal pattern of mortality decline in the G7 countries. *Nature* 405:789-792.

United Nations (2001) *World Population Prospects: The 2000 Revision. Highlights.* New York: United Nations, ESA/P/WP.165.

Wachter, KW and Finch, CE (eds) (1997) *Between Zeus and the Salmon: The Biodemography of Longevity*. Washington DC: National Academy Press.

Wilmoth JR (1993) *Computational methods for fitting and extrapolating the Lee-Carter model of mortality change*. Technical Report, Department of Demography, University of California at Berkeley.

Wilmoth JR (1998) Is the pace of Japanese mortality decline converging toward international trends? *Population and Development Review* 24(3): 593-600.