Debón Aucejo, AM.; Montes, F.; Puig, F. (2008). Modelling and forecasting mortality in Spain. European Journal of Operational Research. 189(3):624-637. doi:10.1016/j.ejor.2006.07.050.

Modelling and Forecasting Mortality in Spain

A. Debón[1*], F. Montes[2] & F. Puig[2]

[1] Universidad Politécnica de Valencia, [2] Universitat de València

Abstract

Experience shows that static life tables overestimate death probabilities. As a consequence of this overestimation the premiums for annuities, pensions and life insurance are not what they actually should be, with negative effects for insurance companies or policy-holders. The reason for this overestimation is that static life tables, through being computed for a specific period of time, cannot take into account the decreasing mortality trend over time. Dynamic life tables overcome this problem by incorporating the influence of the calendar when graduating mortality. Recent papers on the topic look for the development of new methods to deal with this dynamism.

Most methods used in dynamic tables are parametric, apply traditional mortality laws and then analyze the evolution of estimated parameters with time series techniques. Our contribution consists in extending and applying Lee-Carter methods to Spanish mortality data, exploring residuals and future trends.

Keywords: Forecasting, Dynamic Life Tables, Lee-Carter, Bootstrap confidence intervals.

*Corresponding author: A. Debón. Dpt. Estadística e Investigación Operativa Aplicadas y Calidad. Universidad Politécnica de Valencia. E-46022. Valencia. Spain. Tel. nº: +34 963877007 (Ext. 74961). Fax nº: +34 963877499. E-mail: andeau@eio.upv.es

1 Introduction

The most recent actuarial literature recognises the fact that mortality evolves over time. Mortality experiences that correspond to different periods display different death probabilities for the same age. This is supported by the fact that mortality has been seen to gradually decline over time, although the decrease is not necessarily uniform across age groups. It is therefore important to be able to measure mortality changes over time accurately, as life insurance policy gains depend on survival. If the static tables used to calculate annuities and reserves predict higher mortality probabilities than is actually the case among policy-holders, the latter will have been undercharged and the insurance company will make a loss.

The concept of a dynamic table seeks to solve this problem by jointly analysing mortality data corresponding to a series of consecutive years. This approach allows the calendar effect's influence on mortality to be studied. A review of dynamic models and their application can be found in Benjamin and Pollard (1992), Pitacco (2004) and Tabeau et al. (2001).

Since Lee and Carter proposed dynamic life tables in 1992 and use them to study projections of the mortality rate in US (Lee and Carter, 1992), there have been several papers forecasting population mortality in other developed countries such as Canada (Nault, 1993), Chile (Lee and Rofman, 1994), Japan (Wilmoth, 1996), Belgium (Brouhns et al., 2002), Austria (Carter and Prkawetz, 2001), England and Wales (Renshaw and Haberman, 2003a), Australia (Booth and Tickle, 2003) and Spain (Guillen and Vidiella-i-Anguera, 2005).

This paper applies the Lee-Carter model with different methods for estimating parameters, compares all these approaches using mortality data for Spain during the 1980-1999 period, paying attention to the analysis and study of residuals, and predicts death probabilities and expected remaining lifetime for future years. In the prediction of expected remaining lifetime, parametric and non parametric bootstrapping procedures are used as a way to take into account different sources of uncertainty.

The paper is structured as follows. Section 2 briefly presents the Lee-Carter methodology to be used and the bootstrapping techniques. Section 3 is devoted to the application of the Lee-Carter method to the analysis of Spanish mortality data, including the prediction of specific mortality rates in the year 2000 ($q_{x,2000}$) obtained using the adjusted models and the computation of bootstrap confidence intervals for expected remaining lifetimes. Section 4 establishes the conclusions to be drawn from the results in the previous section.

2 Adjustment and prediction of $\mathbf{q_{xt}}$

The Lee-Carter Model, developed in Lee and Carter (1992), consists in adjusting the following function to the central mortality rates,

$$m_{xt} = \exp(a_x + b_x k_t + \epsilon_{xt})$$

or, its equivalent

$$\ln(m_{xt}) = a_x + b_x k_t + \epsilon_{xt}. \tag{1}$$

In the previous two expressions, the double subscript refers to the age, x, and to the year or unit of time, t. a_x and b_x are age-dependent parameters and k_t is a specific mortality index for each year or unit of time. The errors ϵ_{xt}, with 0 mean and variance σ_ϵ^2, reflect the historical influences of each specific age that are not captured by the model.

In Lee (2000) the author remarks that nothing ensures that the m_{xt} estimations obtained from (1) will not exceed 1, although this problem can be avoided by modelling the logit death rates. It is for that reason that we apply this model to logit death probability q_{xt},

$$\ln\left(\frac{q_{xt}}{1-q_{xt}}\right) = a_x + b_x k_t + \epsilon_{xt}. \tag{2}$$

Booth et al. (2002) and Renshaw and Haberman (2003b) indicate that the interaction between age and time can be captured better by adding terms to (2), which would become

$$\ln\left(\frac{q_{xt}}{1-q_{xt}}\right) = a_x + \sum_{i=1}^{r} b_x^i k_t^i + \epsilon_{xt}. \tag{3}$$

In our application to the Spanish data of mortality we have used (3) with $r = 1$ and $r = 2$, consequently the corresponding models will be named LC and $LC2$, respectively.

2.1 Estimation

Given a solution of (2), (a_x, b_x, k_t), any transformation of the type $(a_x, b_x/c, ck_t)$ or (a_x+cb_x, b_x, k_t-c), $\forall c$, is also a solution. In order to avoid this trouble and to get a single solution, Lee and Carter (1992) propose constraining the b_x to sum to 1 and the k_t to sum to 0. The estimations of the parameters of the model can be carried out following the Lee-Carter's approach, which uses the first term of the singular value decomposition, SVD, the method proposed by Currie et al. (2004), which use conditional generalized linear models, GLM, or by means of maximum-likelihood estimations, ML, as Brouhns et al. (2002) propose and that we have adapted for the equation (2). Next we briefly present how to obtain the parameter estimation by means of the methods SVD and GLM.

2.1.1 Lee-Carter-SVD

a) The estimation of a_x is,

$$\breve{a}_x = \frac{\sum_t \ln\left(\frac{q_{xt}}{1-q_{xt}}\right)}{T},$$

with T number of years.

b) The values b_x and k_t are estimated from the first term of SVD applied to the matrix,

$$\ln\left(\frac{q_{xt}}{1-q_{xt}}\right) - \check{a}_x = \sum_{i=1}^{min(n,T)} s^i u_x^i v_t^i, \tag{4}$$

with n number of ages, and s^i, u_x^i and v_t^i the (ordered) singular values and respective left and right singular vectors (Renshaw and Haberman, 2003b).

Then $\hat{k}_t^{SVD} = s^1 v_t^1$ and $\hat{b}_x^{SVD} = u_x^1$. As $\sum_x \hat{b}_x^{SVD} = c \neq 1$, we normalize with the transformation $\hat{b}_x = \dfrac{\hat{b}_x^{SVD}}{c}$ and $\check{k}_t = c\hat{k}_t^{SVD}$.

c) The estimated values in b) can show discrepancies between observed and predicted deaths due to the fact that they are based on logit of deaths rates. To avoid this problem, k_t is reestimated, using $\check{k}_t$ as initial values, by means of the equation

$$D_t = \sum_x \left(E_{xt} \frac{\exp(\check{a}_x + \sum_i k_t \hat{b}_x)}{1 + \exp(\check{a}_x + \sum k_t \hat{b}_x)} \right), \tag{5}$$

where E_{xt} is the number of initial exposed to risk at age x in year t and $D_t = \sum_x D_{xt}$ is total deaths in year t, with D_{xt} the number of death at age x in year t. The solution of (5) is $\tilde{k}_t$. The final solution requires a translation,

$$\hat{k}_t = \tilde{k}_t - \text{mean}(\tilde{k}_t) \qquad \text{and} \qquad \hat{a}_x = \check{a}_x + \hat{b}_x \text{mean}(\tilde{k}_t).$$

2.1.2 Lee-Carter-GLM

The former solution $(\hat{a}_x, \hat{b}_x, \hat{k}_t)$ can be improved by conditional GLM models, as Currie et al. (2004) did when applying GLM to force of mortality. The GLM fitting process allows random components to have a distribution other than the normal. We have considered a Binomial distribution for $D_{xt} \sim Bi(E_{xt}, q_{xt})$ with a logit link as more appropriate for the death rate with values between 0 and 1. In fact, the law of mortality proposed by Heligman and Pollard (1980) already suggested the

use of logit death rate as a response variable. Other advantages of conditional GLM are described in Booth et al. (2002).

For the GLM fitting we proceed as follows,

a) for each x, using

$$\ln\left(\frac{q_{xt}}{1-q_{xt}}\right) = \text{offset}(\hat{a}_x) + b_x\hat{k}_t,$$

where $\hat{a}_x$ and $\hat{k}_t$ are the estimations obtained with the previous method, we obtain $\check{b}_x^{GLM}$ with $\sum_x \check{b}_x^{GLM} = s \neq 1$,

b) then for each t, using

$$\ln\left(\frac{q_{xt}}{1-q_{xt}}\right) = \text{offset}(\hat{a}_x) + \check{b}_x^{GLM}k_t,$$

we obtain $\check{k}_t^{GLM}$ with $\sum_x \check{k}_t^{GLM} \neq 0$,

c) finally,

$$\begin{aligned}
\hat{b}_x^{GLM} &= \frac{\check{b}_x^{GLM}}{s} \\
\hat{k}_t^{GLM} &= s\check{k}_t^{GLM} - \text{mean}\left(s\check{k}_t^{GLM}\right) \\
\hat{a}_x^{GLM} &= \hat{a}_x + \sum_i \hat{b}_x^{GLM(i)}\text{mean}\left(s\check{k}_t^{GLM}\right).
\end{aligned}$$

Now the solution satisfies the constraints. The correction (5) has not been done now because, although the total number of annual deaths may not be reproduced exactly, this loss is outweighed by the greater accuracy in reproducing the age distribution of deaths (Booth et al., 2002).

If we want to estimate (3) with $r = 2$ we will use two terms of the SVD decomposition, obtaining two estimations for b_x and k_t from the initial values, $\hat{k}_t^{SVDi} = s^i v_t^i$ and $\hat{b}_x^{SVDi} = u_x^i, i = 1, 2$.

2.2 Residuals

Little attention appears to have been paid to the definition and analysis of residuals in the literature. Renshaw and Haberman (2003a) and Booth et al. (2002) are exceptions. Our work focuses on SVD residuals,

$$\hat{\epsilon}_{xt} = \ln\left(\frac{q_{xt}}{1-q_{xt}}\right) - \left(\hat{a}_x + \sum_{i=1}^{r} \hat{k}_t^i \hat{b}_x^i\right), \tag{6}$$

where r is the number of singular values used.

The graphs of residuals against age and year are used to detect if any pattern is present. The model performance is evaluated with two measures, the Mean Absolute Percentage Error ($MAPE$) and Mean Square Error (MSE). The first, used by Felipe et al. (2002), is defined by,

$$MAPE(\hat{q}_{xt}) = \frac{\sum_x \frac{|\dot{q}_{xt} - \hat{q}_{xt}|}{\dot{q}_{xt}}}{n},\ t = 1, \ldots, T$$

and it measures the mean absolute error weighted with the inverse of the crude estimates $\dot{q}_{xt}$. These weights allow us to reduce the effect of the errors associated with high values of $\dot{q}_{xt}$, usually associated with intermediate and old ages. The second, defined by,

$$MSE(\hat{q}_{xt}) = \sqrt{\sum_x \frac{(\dot{q}_{xt} - \hat{q}_{xt})^2}{n}},\ t = 1, \ldots, T$$

measures the error of estimations without any correction.

2.3 Model forecasts

Forecast of q_{xt} are generated by first modelling $\hat{k}_t$ as a time series by using Box-Jenkins methodology. Usually, in many of these applications, a good model for the k_t is an $ARIMA(0,1,0)$,

$$\hat{k}_t = c + \hat{k}_{t-1} + u_t,$$

where c a is constant and u_t is a white noise. With this model, the prediction of k_t varies in a linear way and each death rate predicted varies at a constant exponential rate. Booth et al. (2002) have introduced some modifications to the Lee-Carter method, particularly an improved fit of the base model, the identification of the optimal fitting period for the lineal ARIMA, and the incorporation of age-time interaction in the base model.

2.4 Bootstrap confidence intervals for expected remaining lifetimes

The Lee-Carter model is designed to produce values for the dynamic expected remaining lifetimes at different ages. For a year t, the hypothetical number of people alive at the start of each age interval $[x, x+1)$ is given by iterative formula $l_{(x+1)t} = l_{xt}(1 - q_{xt})$, with arbitrary value l_0. This leads to a number of deaths $d_{xt} = l_{xt} - l_{(x+1)t}$, and the corresponding number of person-years $L_{xt} = l_{(x+1)t} + 1/2 d_{xt}$. Therefore, the total future remaining lifetime of L_{xt} people who attain age x is $T_{xt} = \sum_{i \geq x} L_{it}$. The future remaining lifetime for individuals from age-grouping x is given by

$$e_{xt} = \frac{T_{xt}}{l_{xt}}.$$

In their original paper, Lee and Carter predict confidence intervals for the expected remaining life time taking into account only forecast errors in the projected $ARIMA$ k_t parameters. Nevertheless, another source of error is due to sampling errors in the parameters of the Binomial model. A way to combine these two sources of uncertainty is to use bootstrapping procedures as Brouhns et al. (2005) and Koissi et al. (2006) do. In the first paper the authors use a parametric bootstrapping, while the second paper is devoted to a nonparametric one.

We have carried out both bootstrapping procedures. For the parametric case we use the Binomial distribution of D_{xt}. Starting from the observations (E_{xt}, d_{xt}), we simulate N bootstrap samples (E_{xt}, d^n_{xt}), $n = 1, 2, \ldots, N$, where d^n_{xt} are realizations from Binomial distribution with

parameters $(E_{xt}, \dot{q}_{xt})$. For each bootstrap sample, the a_x's, b_x's and k_t's are estimated, and the k_t's are then projected on the basis of the $ARIMA$ model selected from the original data. This yields N realizations of a_x^n, b_x^n, k_t^n and projected k_t^n which are used to compute expected remaining lifetimes. The confidence intervals are the percentile intervals, $IC_{95} = [p_{0.025}, p_{0.975}]$.

For the non parametric case, the N bootstrap samples are simulated from the SVD residuals (6) obtained with the original data. Each sample furnishes an estimated $\widehat{logit(q_{xt})}^n$ by means of the inverse of formula (6)

$$\widehat{logit(q_{xt})}^n = logit(\dot{q}_{xt}) - \hat{\epsilon}_{xt}^n,$$

where $logit(\dot{q}_{xt})$ are computed from the observations (E_{xt}, d_{xt}). From this point we proceed as in the parametric case. The extension of bootstrap methods to the $LC2$ models is immediate.

3 Analysis of mortality data in Spain

3.1 Data

The models described in Section 2 have been used to adjust Spanish mortality data corresponding to the period 1980-1999, for a range of ages from 0 to 96. Data for men and women are adjusted separately. The population of Spain in the last census carried out in 2001 was 40,847,371.

The crude estimates of q_{xt} have been obtained by means of the procedure used by the Spanish National Institute of Statistics (INE, Instituto Nacional de Estadística),

$$\dot{q}_{xt} = \frac{1/2(d_{xt} + d_{x(t+1)})}{P_{xt} + 1/2 d_{xt}},$$

where d_{xt} are the deaths in year t at age x, $d_{x(t+1)}$ are the deaths in year $t+1$ at age x, and P_{xt} is the population that on December 31st of year t was aged x. The formula can be applied to all ages, except for 0, due to the concentration of deaths in the first few months of life. The expression

used for age 0 is,

$$\dot{q}_{0t} = \frac{0.85d_{0t} + 0.15d_{0(t+1)}}{P_{0t} + 0.85d_{0t}}.$$

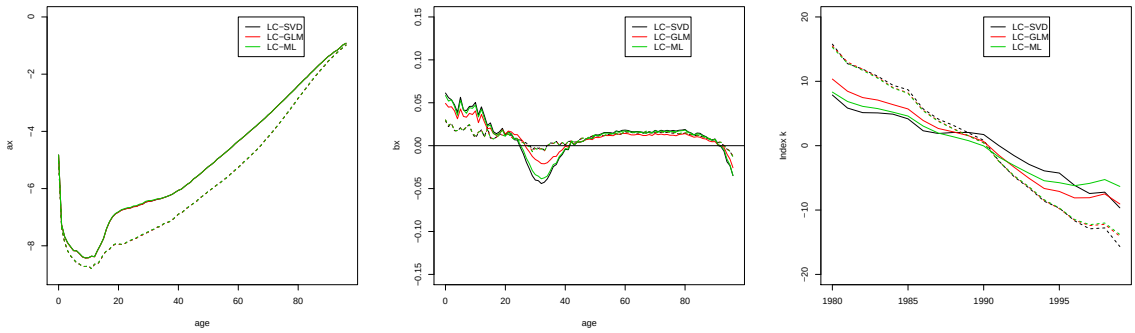
3.2 Model adjustment

The high number of parameters estimated in the LC model, $97 \times 2 + 20 = 214$ for men and women, cannot be fully presented in a paper of this extent. We prefer to present them instead in the form of a graph in Figure 1. As a general comment, we must point out that the differences among the estimations obtained with SVD method (LC-SVD), GLM method (LC-GLM) and ML method (LC-ML) are very small in a_x values for both sexes. For b_x and k_t, they are not appreciable for women, but this is not the case for men.

More specifically, the comparison of parameter a_x for both sexes shows that mortality for women is lower than for men. The hump in Figure 1(a) reveals an increase of mortality in the range of ages from 11 to 40 for men that some authors (Guillen and Vidiella-i-Anguera, 2005) attribute to an accidental mortality.

The negative values of parameter b_x for intermediate ages (24 to 40 for men and 28 t0 32 for women) and for advanced ages (more than 92 for both sexes) indicate that mortality in these age groups does increase over time. For the range of ages from 24 to 40 this increasing over time could be explained by the effect of the AIDS epidemic as Felipe et al. (2002) and Guillen and Vidiella-i-Anguera (2005) point out.

Figure 2 shows the estimations obtained with the $LC2$ model. The behaviour of a_x, b_x^1 and k_t^1 is similar to the one observed for the model LC. b_x^2 (Figure 2(d)), corresponding to the second term, shows larger values for ages in the range from 20 to 40, which implies that the effect of adding a second term acts more specifically on this age group, particularly for *LC2-SVD* and *LC2-GLM* methods. With regards to k_t^2 (Figure 2(e)), *LC2-SVD* and *LC2-GLM* methods do not display any clear trend, it decreases for the first two years, increases until 1990 and decreases for the later



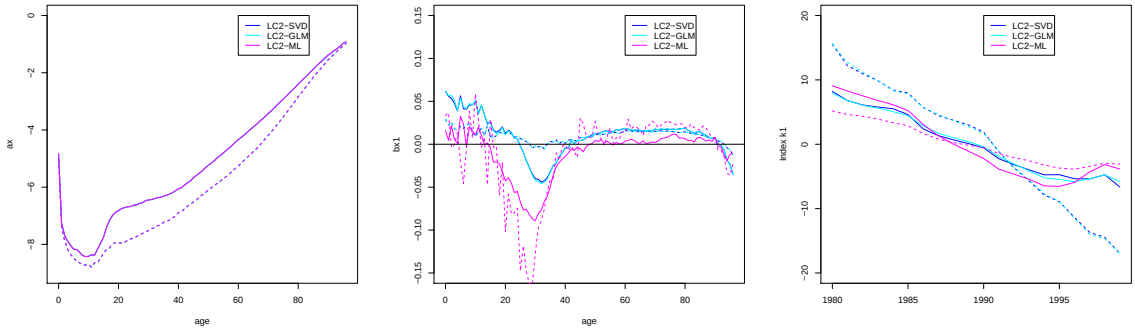
(a) a_x (b) b_x (c) k_t

Figure 1: Estimated values for men (solid line) and for women (dotted line) for the *LC* models.

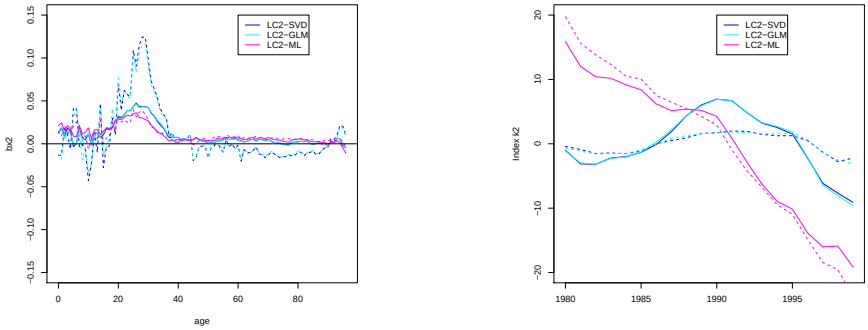
years, with fewer variations for the women than for men. It can be assumed, for both methods, that a slightly increasing trend exists for the first years, which changes to a decreasing trend later. In the *LC2-ML* method the estimation of k_t^2 shows a clearly decreasing trend.

As far as we know, only Felipe et al. (2002) and Guillen and Vidiella-i-Anguera (2005) have studied the evolution of Spanish mortality with dynamic models. The first use the Heligman-Pollards laws to research the way in which the calendar time (1975-1993) affects mortality patterns in the Spanish population for a range of ages from 0 to 90. The second use a Poisson log-bilinear version of the Lee-Carter model, proposed by Wilmoth (1993) and Brouhns et al. (2002), for Spanish mortality data during the 1975-1998 period and a range of ages from 0 to 105. Our results about the evolution of mortality rates are similar to those obtained by these authors.

In recent times, statistical techniques have improved the study of mortality behaviour at very old ages. Pitacco (2004) remarks that the mortality at very old ages is slowly increasing in many countries. Spanish mortality evolution holds with this fact as Figure 3 shows.

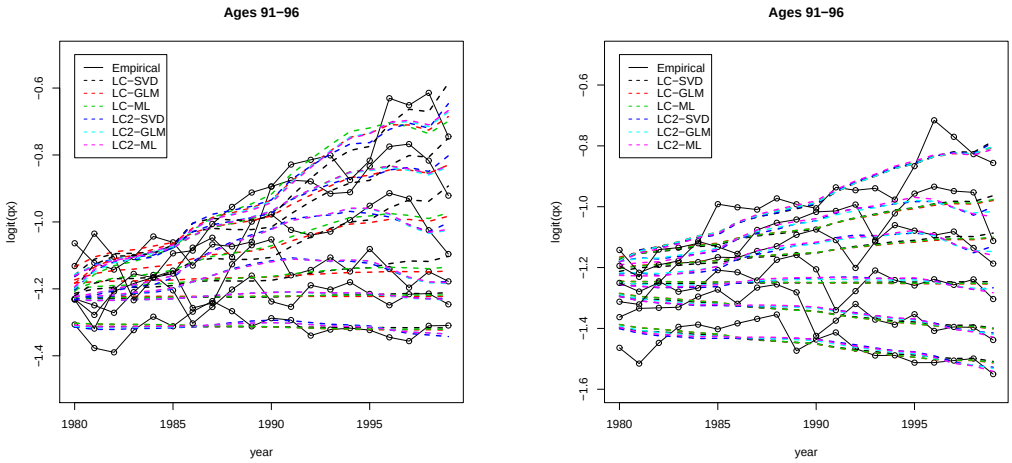


(a) a_x (b) b_x^1 (c) k_t^1



(d) b_x^2 (e) k_t^2

Figure 2: Estimated values men(solid line) and women(dotted line) for $LC2$ models.

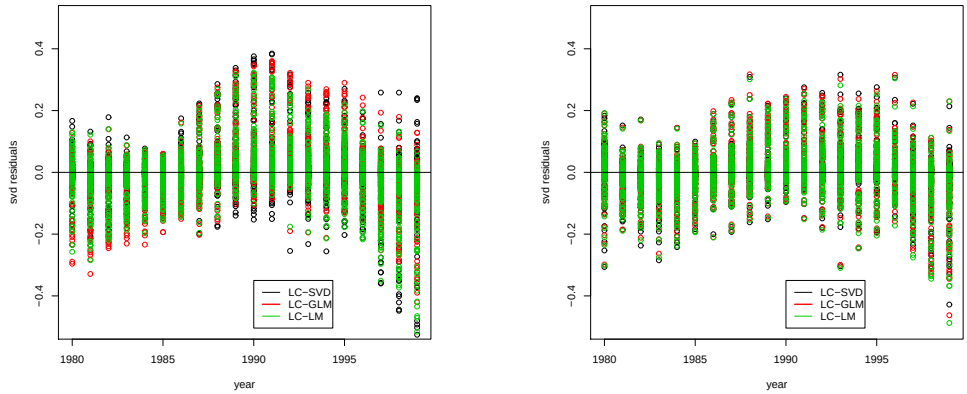


(a) Men (b) Women

Figure 3: Logit rates versus year (1980-1999) for old ages.

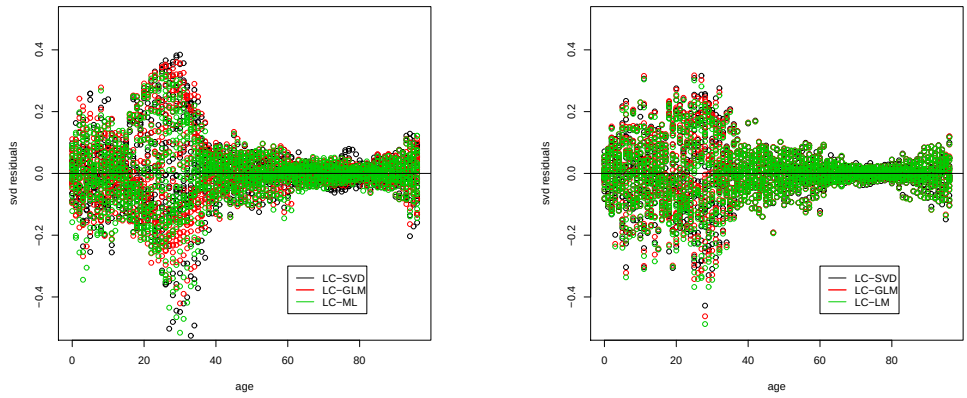
3.3 Goodness-of-fit

Residuals corresponding to the adjustment of the LC model with the three methods are plotted in Figures 4 and 5, displaying a cyclic pattern indicating a poor fit. A better fit is obtained when a second term is introduced in the Lee-Carter model, $LC2$ model, as the plot of the corresponding residuals shows in Figures 6 and 7.



(a) Men (b) Women

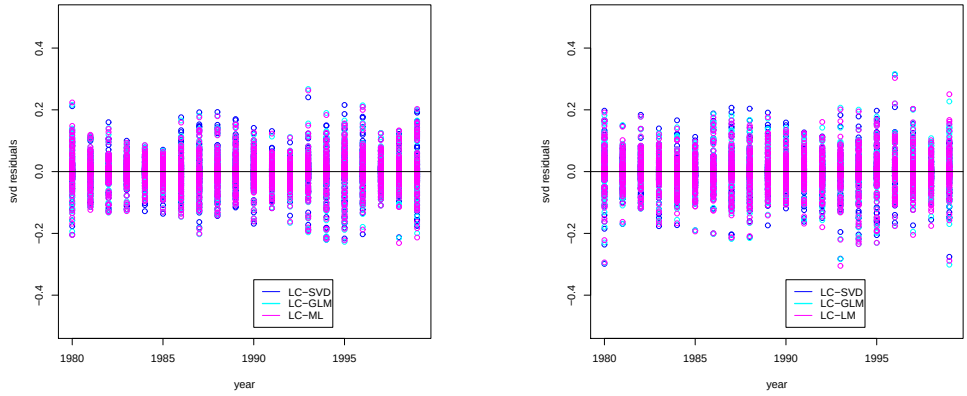
Figure 4: SVD residuals for the LC model versus year



(a) Men (b) Women

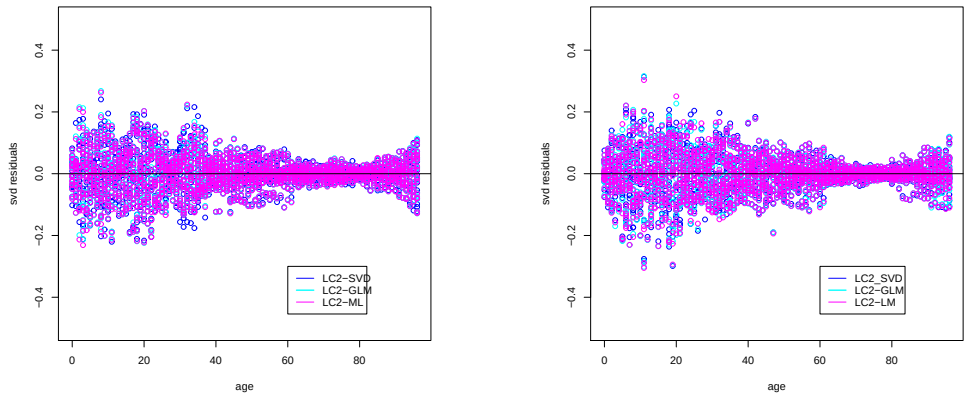
Figure 5: SVD residuals for the LC model versus age

Figure 8 shows the MAPE values for different adjustments. These values confirm the best



(a) Men (b) Women

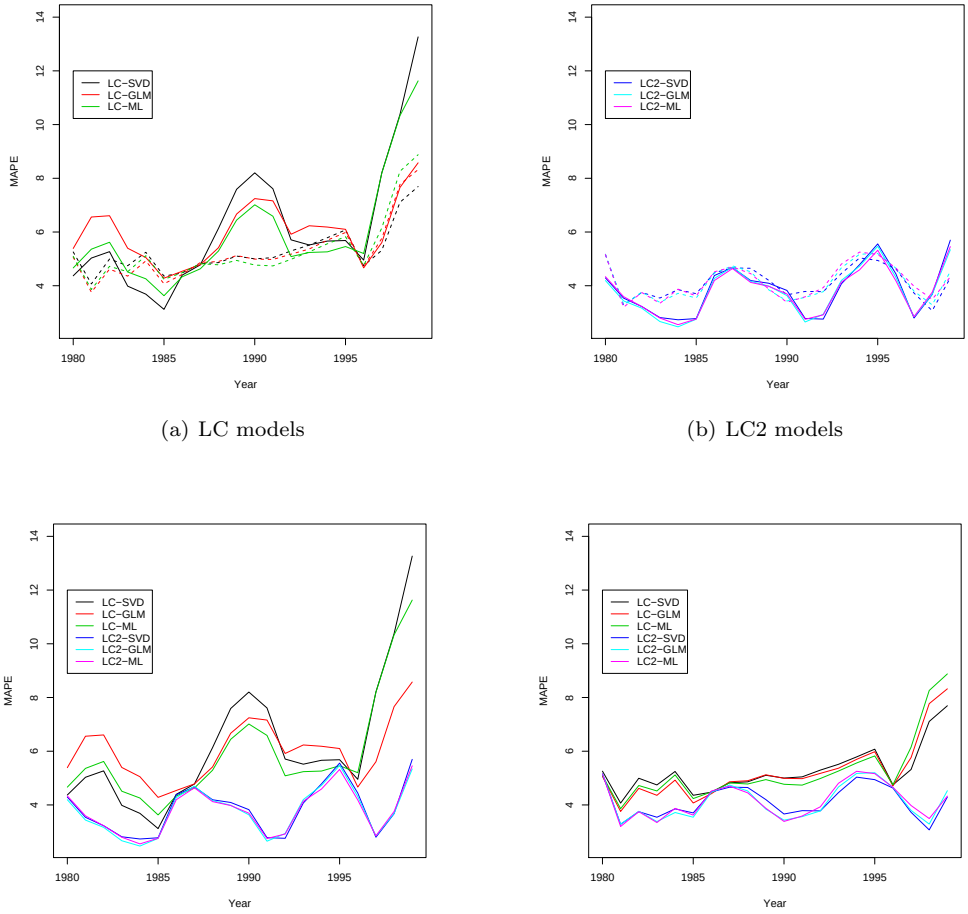
Figure 6: SVD residuals for the $LC2$ model year



(a) Men (b) Women

Figure 7: SVD residuals for the $LC2$ model versus age

behaviour of the $LC2$ model. The LC model fits worse for the most recent years. The explanation of this fact can be found in the diminution of the accident hump in these years. This poorer adjustment is not observed in the case of the $LC2$ model because, as we have already pointed out, the introduction of the second term better adapts the model for the ages involved in the accident hump. The GLM and ML methods produce better adjustments than the SVD method in both models. The adjustment for women is in general better than for men, particularly with the GLM method.



(a) LC models

(b) LC2 models

(c) Men

(d) Women

Figure 8: MAPE values over the years (solid line for men and dotted line for women).

3.4 Forecasting

The prediction of mortality rates needs the previous adjustment of a time series to mortality indexes, k_t, which is carried out by means of the Box-Jenkins methodology using an R code developed by Shumway and Stoffer (2006). As we have already said, the most appropriate model turns out to be an $ARIMA$ model because the first or second differences are seen to be a stationary series for both sexes. The standard Autoregressive Integrated Moving Average (ARIMA) models are denoted $ARIMA(p, d, q)$ where p, d and q are autoregressive, difference and moving average orders, respectively (Diggle, 1990).

The $ARIMA$ models adjusted for each model and method are shown in Table 1.

		SVD		GLM		ML	
		Men	Women	Men	Women	Men	Women
LC	k_t	(0,1,0)	(0,1,0)	(0,1,0)	(0,1,0)	(0,1,0)	(0,1,0)
	constant	-0.9226	-1.6550	-1.0217	-1.5556	-0.7715	-1.5308
LC2	k_t^1	(0,1,0)	(0,1,0)	(0,1,0)	(0,1,0)	(0,1,0)	(0,1,0)
	constant	-0.7815	-1.7115	-0.7222	-1.7170	-0.6804	-0.4341
	k_t^2	(0,2,0)	(0,2,0)	(0,2,0)	(0,2,0)	(0,1,0)	(0,1,0)
	constant	0	0	0	0	-1.8403	-2.2799

Table 1: ARIMA models for mortality index

The underlying assumption is that the two time series corresponding to k_t^1 and k_t^2 are independent. This assumption made also by Booth et al. (2002) and by Renshaw and Haberman (2003b) represents a potencial weakness as Renshaw and Haberman (2003b) recognize. An example of how the model can be expanded to include dependence and co-integration effects is given in Renshaw and Haberman (2003c).

3.4.1 Predictions for year 2000

As Sithole et al. (2000) point out, "... for each experience analysed, determining a model that provides the best fit to the data did not present many problems. The difficulty was in identifying

a model that not only provided a good fit for the data, but also had a good shape for the purpose of making projections."

In order to check how the adjusted models perform when predicting, the predictions of $q_{x,2000}$ have been obtained. The crude estimates, $\dot{q}_{xt}$, are known and we also have mortality data for the year 2001. A summary of these predictions in terms of $MAPE$ and MSE is shown in Figure 9. It can be seen that all adjustments perform better for women than for men, as women display greater linearity and a lesser variability. With respect to models and methods, these results confirm those obtained in Section 3.2 in the sense of a better behaviour of the $LC2$ model. When comparing crude estimates with their predictions, all methods show a lack of fit for high values of $\dot{q}_{xt}$. This can be seen in Figure 10 that shows predictions versus crude estimates for some of the adjustments.

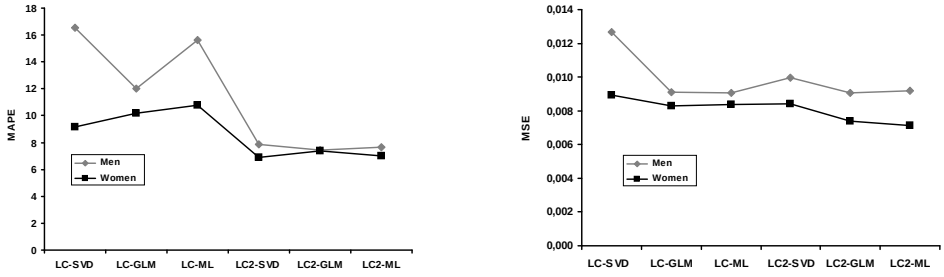


Figure 9: MAPE and MSE of predicted mortality rates for the year 2000.

3.4.2 Prediction for future years

Figure 11 shows the prediction of mortality ratios for years 2000 to 2011 obtained with the different models and methods. In order to facilitate the interpretation of these results we have represented the ages from 10 to 50 years, for every 10 years. The greatest differences between the different adjustments are observed in the ages 20 and 30. The best behaviour is observed for the *LC2-GLM* and *LC2-ML* methods because, as we have already said, the inclusion of the second term adapts the model to the changes of trend better. However, this reasoning cannot be applied to the *LC2-SVD* method, that also incorporates a second term, because, contrary to the other two methods,

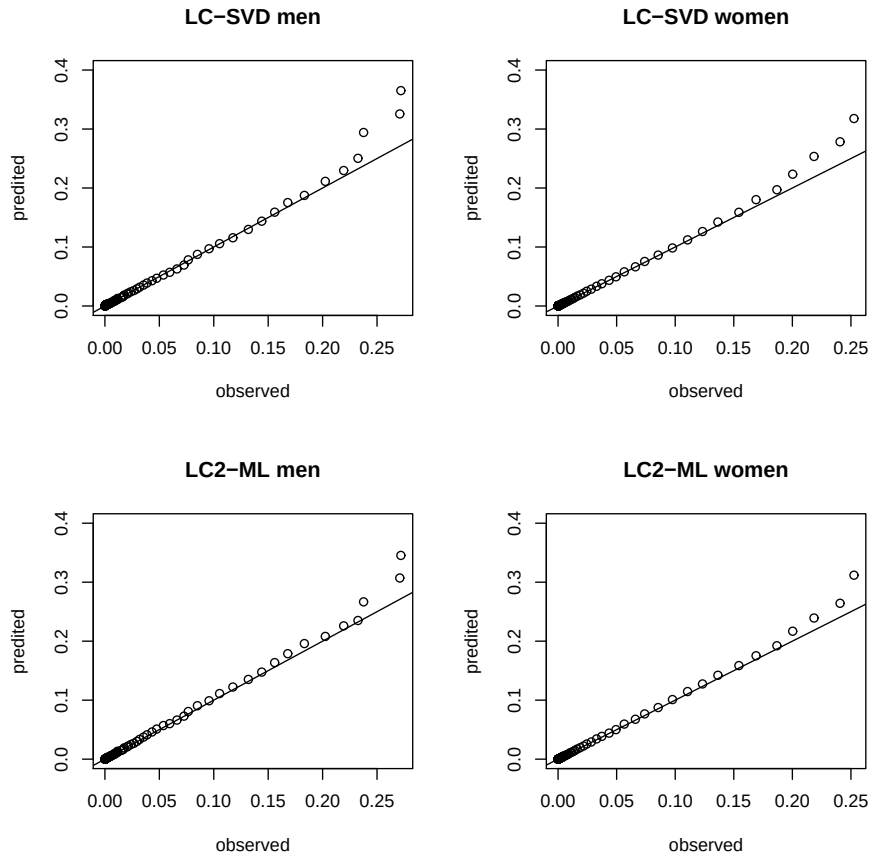


Figure 10: Predicted vs observed values for some methods for the year 2000.

the evolution of its estimations of the mortality index k_t^2 does not allow a good adjustment of the corresponding time series. This fact is well observed in the predictions for women with ages of 20 and 30 years, whose evolution with *LC2-SVD* method increases in a unreasonably. These comments can be extended to the intermediate ages whose graphs do not appear in the figure.

For old ages there are no great differences, although the *LC2-GLM* and *LC2-ML* methods are those showing the smoothest increase.

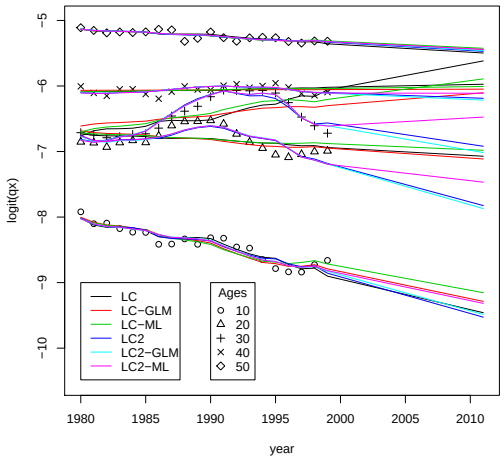
3.5 Bootstrap confidence intervals for e_{65t}

Tables 2 and 3 show the confidence intervals for the prediction of expected remaining lifetimes at the age of 65 years, e_{65t}, obtained by means of bootstrap techniques described in Section 2.4. The prediction has been carried out for the period 2000-2011. The first table contains the results for the Lee-Carter model with only one term, LC. The results for the model with two terms are shown in the second table.

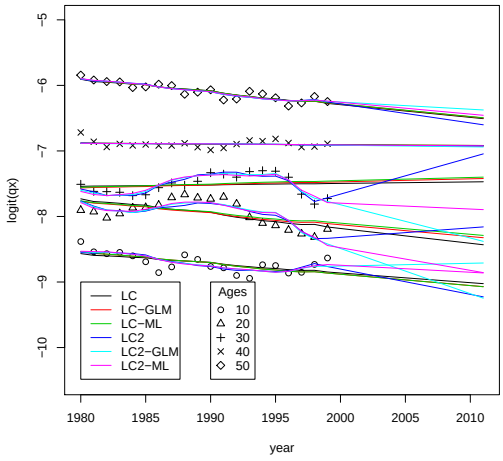
We can see from the two Tables that the expected remaining lifetime is higher for women than for men, and there is a clear trend indicating an increase in e_{65t}. Specifically, this expected increase varies for men from 0.84 to 1.06 years depending on the model, adjustment method and bootstrap method. For women, the increase varies from 1.05 to 1.64 years. These results are slightly higher than those obtained by Guillen and Vidiella-i-Anguera (2005).

As far as the bootstrap method is concerned, the amplitude of the intervals obtained by means of the parametric method is, in general, smaller than the one obtained with the non-parametric one. With regard to the models, the $LC2$ model provides intervals of smaller amplitude for men, although this good behaviour does not hold for women, because only the *LC2-ML* method obtains better intervals than its homologous *LC-ML*.

It must be pointed out that our prediction intervals for forecasted life expectancies in tables 2 and 3 are narrow. This fact has attracted attention of other researchers in this field, Lee and



(a) Men



(b) Women

Figure 11: Predicted rates versus year (1980-2011) for some ages.

		LC-SVD						LC-GLM						LC-ML					
		men			women			men			women			men			women		
	year	$p_{0.025}$	mean	$p_{0.975}$	$p_{0.025}$	mean	$p_{0.975}$	$p_{0.025}$	mean	$p_{0.975}$	$p_{0.025}$	mean	$p_{0.975}$	$p_{0.025}$	mean	$p_{0.975}$	$p_{0.025}$	mean	$p_{0.975}$
non parametric	2000	16.78	17.08	17.35	20.68	20.88	21.06	16.63	16.81	17.00	20.63	20.79	20.95	16.48	16.69	16.90	20.58	20.77	20.94
	2001	16.85	17.17	17.46	20.79	20.99	21.19	16.70	16.89	17.09	20.74	20.91	21.07	16.55	16.77	16.99	20.69	20.89	21.07
	2002	16.93	17.27	17.56	20.90	21.11	21.32	16.77	16.97	17.19	20.84	21.02	21.19	16.62	16.84	17.08	20.80	21.00	21.19
	2003	17.01	17.36	17.66	21.00	21.23	21.44	16.84	17.04	17.27	20.95	21.14	21.32	16.68	16.92	17.17	20.91	21.12	21.32
	2004	17.08	17.46	17.77	21.11	21.35	21.57	16.91	17.12	17.36	21.05	21.25	21.44	16.75	17.00	17.26	21.02	21.23	21.44
	2005	17.16	17.55	17.87	21.21	21.47	21.70	16.97	17.20	17.46	21.16	21.37	21.56	16.82	17.07	17.36	21.12	21.35	21.56
	2006	17.24	17.64	17.97	21.32	21.58	21.82	17.04	17.27	17.54	21.26	21.48	21.68	16.88	17.15	17.44	21.23	21.46	21.68
	2007	17.31	17.73	18.07	21.42	21.70	21.94	17.11	17.35	17.63	21.37	21.59	21.80	16.94	17.23	17.53	21.33	21.57	21.80
	2008	17.39	17.82	18.18	21.52	21.81	22.06	17.18	17.43	17.72	21.47	21.70	21.91	17.00	17.30	17.62	21.43	21.68	21.92
	2009	17.46	17.91	18.28	21.63	21.92	22.17	17.24	17.51	17.81	21.57	21.81	22.03	17.06	17.38	17.71	21.53	21.79	22.04
	2010	17.54	18.00	18.38	21.73	22.03	22.29	17.31	17.58	17.90	21.67	21.92	22.14	17.12	17.45	17.79	21.63	21.90	22.16
	2011	17.61	18.09	18.48	21.83	22.14	22.41	17.37	17.66	17.99	21.77	22.02	22.26	17.18	17.53	17.88	21.73	22.01	22.28
parametric	2000	17.06	17.11	17.16	20.85	20.89	20.92	16.78	16.82	16.85	20.77	20.81	20.84	16.70	16.73	16.77	20.77	20.80	20.84
	2001	17.16	17.21	17.26	20.97	21.01	21.05	16.86	16.90	16.93	20.89	20.92	20.96	16.77	16.81	16.85	20.88	20.92	20.96
	2002	17.25	17.30	17.36	21.09	21.13	21.17	16.94	16.98	17.01	21.00	21.04	21.08	16.85	16.89	16.93	21.00	21.03	21.07
	2003	17.34	17.40	17.46	21.20	21.25	21.29	17.02	17.05	17.09	21.12	21.16	21.19	16.92	16.97	17.01	21.11	21.15	21.19
	2004	17.43	17.49	17.55	21.32	21.36	21.41	17.09	17.13	17.17	21.23	21.27	21.31	17.00	17.04	17.09	21.23	21.27	21.31
	2005	17.52	17.59	17.65	21.43	21.48	21.52	17.17	17.21	17.25	21.34	21.39	21.43	17.08	17.12	17.17	21.34	21.38	21.42
	2006	17.61	17.68	17.75	21.54	21.59	21.64	17.25	17.29	17.34	21.46	21.50	21.54	17.15	17.20	17.25	21.45	21.49	21.54
	2007	17.71	17.77	17.84	21.66	21.71	21.76	17.32	17.37	17.42	21.57	21.61	21.66	17.23	17.27	17.32	21.56	21.60	21.65
	2008	17.80	17.87	17.94	21.77	21.82	21.87	17.40	17.45	17.50	21.67	21.72	21.77	17.30	17.35	17.40	21.67	21.71	21.76
	2009	17.89	17.96	18.03	21.88	21.93	21.98	17.48	17.53	17.58	21.78	21.83	21.88	17.37	17.43	17.48	21.78	21.82	21.87
	2010	17.98	18.05	18.13	21.99	22.04	22.10	17.55	17.60	17.66	21.89	21.94	21.99	17.45	17.50	17.56	21.88	21.93	21.98
	2011	18.06	18.14	18.22	22.10	22.15	22.21	17.63	17.68	17.74	22.00	22.05	22.10	17.52	17.58	17.64	21.99	22.04	22.09

Table 2: Bootstrap confidence intervals for the LC model

Table 3: Bootstrap confidence intervals for the $LC2$ model

		LC2-SVD						LC2-GLM						LC2-ML					
		men			women			men			women			men			women		
	year	$p_{0.025}$	mean	$p_{0.975}$	$p_{0.025}$	mean	$p_{0.975}$	$p_{0.025}$	mean	$p_{0.975}$	$p_{0.025}$	mean	$p_{0.975}$	$p_{0.025}$	mean	$p_{0.975}$	$p_{0.025}$	mean	$p_{0.975}$
non parametric	2000	16.77	16.91	17.07	20.65	20.88	21.07	16.72	16.84	16.97	20.60	20.78	20.95	16.69	16.81	16.94	20.65	20.79	20.92
	2001	16.86	17.01	17.18	20.74	21.04	21.29	16.80	16.93	17.08	20.61	20.91	21.15	16.76	16.89	17.03	20.76	20.90	21.05
	2002	16.94	17.11	17.30	20.83	21.19	21.51	16.88	17.03	17.20	20.63	21.03	21.36	16.83	16.97	17.12	20.87	21.02	21.17
	2003	17.02	17.20	17.42	20.92	21.34	21.72	16.96	17.12	17.31	20.63	21.15	21.58	16.91	17.05	17.20	20.98	21.14	21.29
	2004	17.11	17.30	17.53	21.00	21.49	21.93	17.04	17.21	17.42	20.65	21.27	21.81	16.98	17.13	17.29	21.09	21.25	21.41
	2005	17.19	17.40	17.65	21.08	21.64	22.14	17.11	17.30	17.54	20.65	21.38	22.02	17.05	17.21	17.38	21.19	21.36	21.53
	2006	17.27	17.49	17.77	21.17	21.78	22.34	17.19	17.39	17.66	20.64	21.50	22.22	17.12	17.29	17.46	21.30	21.48	21.64
	2007	17.35	17.59	17.88	21.24	21.92	22.53	17.26	17.48	17.77	20.64	21.62	22.42	17.20	17.37	17.55	21.40	21.59	21.76
	2008	17.44	17.68	18.00	21.32	22.06	22.73	17.33	17.57	17.88	20.63	21.73	22.61	17.27	17.45	17.63	21.50	21.70	21.88
	2009	17.51	17.78	18.11	21.39	22.20	22.92	17.41	17.66	17.99	20.62	21.83	22.79	17.34	17.53	17.72	21.61	21.81	22.00
	2010	17.59	17.87	18.22	21.47	22.34	23.09	17.48	17.74	18.10	20.61	21.94	22.96	17.42	17.61	17.81	21.71	21.92	22.11
	2011	17.67	17.96	18.34	21.55	22.47	23.26	17.55	17.83	18.21	20.59	22.05	23.13	17.49	17.68	17.89	21.81	22.02	22.22
parametric	2000	16.88	16.93	16.98	20.76	20.89	21.06	16.80	16.84	16.88	20.73	20.78	20.83	16.79	16.82	16.86	20.77	20.80	20.84
	2001	16.97	17.03	17.09	20.82	21.05	21.36	16.89	16.94	16.99	20.78	20.88	20.97	16.87	16.91	16.94	20.88	20.92	20.95
	2002	17.07	17.13	17.20	20.88	21.21	21.65	16.98	17.03	17.09	20.84	20.97	21.10	16.95	16.99	17.02	21.00	21.03	21.07
	2003	17.15	17.23	17.31	20.94	21.37	21.93	17.06	17.12	17.19	20.90	21.07	21.24	17.03	17.07	17.11	21.11	21.15	21.19
	2004	17.24	17.32	17.43	20.99	21.52	22.20	17.15	17.22	17.30	20.95	21.17	21.37	17.10	17.15	17.19	21.22	21.26	21.31
	2005	17.32	17.42	17.54	21.04	21.67	22.46	17.23	17.31	17.40	21.01	21.26	21.50	17.18	17.23	17.27	21.34	21.38	21.42
	2006	17.41	17.52	17.66	21.10	21.82	22.72	17.31	17.40	17.50	21.07	21.36	21.63	17.26	17.31	17.35	21.45	21.49	21.54
	2007	17.49	17.61	17.77	21.15	21.96	22.97	17.39	17.49	17.61	21.12	21.45	21.76	17.33	17.39	17.43	21.56	21.60	21.65
	2008	17.58	17.71	17.88	21.21	22.11	23.21	17.47	17.58	17.71	21.18	21.55	21.89	17.41	17.46	17.52	21.67	21.72	21.76
	2009	17.66	17.80	17.99	21.26	22.25	23.45	17.56	17.68	17.81	21.24	21.64	22.02	17.49	17.54	17.60	21.78	21.82	21.88
	2010	17.75	17.90	18.10	21.32	22.39	23.67	17.64	17.77	17.91	21.29	21.74	22.15	17.57	17.62	17.68	21.88	21.93	21.99
	2011	17.83	17.99	18.21	21.37	22.53	23.88	17.72	17.86	18.01	21.35	21.83	22.27	17.64	17.70	17.76	21.99	22.04	22.10

Carter (1992), Lee (2000), Booth et al. (2002) and Koissi et al. (2006), which provide different explanations. In our case the resultas agree with the evolution of values published by INE.

4 Conclusions

The most important conclusion drawn for the results of our analysis is that the $LC2$ model performs better than the LC model. It fits the Spanish mortality data better, it also predicts the values of $q_{x,2000}$ better and provides bootstrap confidence intervals that display in general smaller amplitude.

When analysing the results by sex, the $LC2$ model also display better behaviour. This is more evident in the case of men, due to the fact that male mortality fluctuations for the ages in the accident hump are difficult to capture over the period of time under consideration, and the inclusion of a second term in (3) improves the versatility of the model. As women display greater linearity and homogeneity, their fits with LC and $LC2$ models are more similar.

When comparing our results with the work of Felipe et al. (2002), who use the Heligman and Pollard second law for studying Spanish mortality data, they are similar in the case of men but the $LC2$ model fits better in the case of women. The reason for a poorer behaviour in women for the Heligman and Pollard model is because, unlike men, mortality in Spanish women does not present the accident hump that this law presupposes for the intermediate ages.

In Guillen and Vidiella-i-Anguera (2005), the authors study the impact on forecast of accidental mortality splitting Spanish mortality data. They obtain confidence intervals for e_{65t} based on projected mortality index k_t for both, natural and overall mortality. Compared with the bootstrap confidence intervals we have computed, the average values of e_{65t} are similar, but with a range slightly higher in our case. As far as precision is concerned, the intervals that we have obtained by means of parametric bootstrap for the $LC2$ model are narrower, even than those they have obtained for data of mortality by natural causes.

Let us indicate finally that, although the results are satisfactory, these models continue to show a poor adjustment for the highest mortality ratios (see Figure 10), most of them corresponding to old ages. All these models assume independence in the observations, a hypothesis difficult to maintain as the graphical representation of the residuals demonstrates. A solution to this problem may consist in modelling the residuals dependence by means of suitable techniques, a future line of research to explore.

Acknowledgments

The authors are indebted to the anonymous referees whose suggestions improved the original manuscript. This work was partially supported by a grant from MEyC (Ministerio de Educación y Ciencia, Spain, project MTM-2004-06231). The research of Francisco Montes has also been partially supported by a grant from DGITT (Direcció General d'Investigació i Transferència Tecnològica de la Generalitat Valenciana, project GRUPOS03/189)

References

Benjamin, B. and Pollard, J. (1992). *The Analysis of Mortality and Other Actuarial Statistics*. Butterworth-Heinemann, London, 6$^{\text{th}}$ edition.

Booth, H., Maindonald, J., and Smith, L. (2002). Applying Lee-Carter under conditions of variable mortality decline. *Population Studies*, 56:325–336.

Booth, H. and Tickle, L. (2003). The future aged: new projections of Australia's ederly population. *Population Studies*, 22(4):38–44.

Brouhns, N., Denuit, M., and Keilegom, I. V. (2005). Bootstrapping Poisson log-bilinear model for mortality forecasting. *Scandinavian Actuarial Journal*, (3):212–224.

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

Carter, L. and Prkawetz, A. (2001). Examining structural shifts in mortality using the Lee-Carter method. Mpidr wp 2001-007, Center for Demography and Ecology Information, University of Wisconsin-Madison.

Currie, I., Kirkby, J., Durban, M., and Eilers, P. (2004). Smooth Lee-Carter models and beyond. In *Workshop on Lee-Carter Methods*, www.ma.hw.ac.uk iain/workshop/workshop.html (4th March 2005).

Diggle, P. J. (1990). *Time Series. A Bioestatistical Introduction*. Oxford University Press, New York.

Felipe, A., Guillén, M., and Pérez-Marín, A. (2002). Recent mortality trends in the Spanish population. *British Actuarial Journal*, 8(4):757–786.

Guillen, M. and Vidiella-i-Anguera, A. (2005). Forecasting Spanish natural life expectancy. *Risk Analysis*, 25(5):1161–1170.

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

Koissi, M., Shapiro, A., and Högnäs, G. (2006). Evaluating and extending the Lee-Carter model for mortality forecasting confidence interval. *Insurance: Mathematics & Economics*, 38(1):1–1.

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

Lee, R. and Carter, L. (1992). Modelling and forecasting U. S. mortality. *Journal of the American Statistical Association*, 87(419):659–671.

Lee, R. and Rofman, R. (1994). Modelación y proyección de la mortalidad en Chile. *Notas Población*, 22(59):182–213.

Nault, R. L. F. (1993). Modeling and forecasting provincial mortality in Canada. Montreal. World Congress of the International Union for Scientific Study of Population.

Pitacco, E. (2004). Survival models in dynamic context: a survey. *Insurance: Mathematics & Economics*, 35(2):279–298.

Renshaw, A. and Haberman, S. (2003a). Lee-Carter mortality forecasting: a parallel generalized linear modelling approach for England and Wales mortality projections. *Journal of the Royal Statistical Society C*, 52(1):119–137.

Renshaw, A. and Haberman, S. (2003b). Lee-Carter mortality forecasting with age specific enhancement. *Insurance: Mathematics & Economics*, 33(2):255–272.

Renshaw, A. and Haberman, S. (2003c) Lee-Carter mortality forecasting incorporating bivariate time series. Actuarial Research Paper 153, City University, London.

Shumway, R. and Stoffer, D. (2006). *Time Series Analysis and its Applications with R Examples*. Springer, New York.

Sithole, T., Haberman, S., and Verrall, R. J. (2000). An investigation into parametric models for mortality projections, with applications to immediate annuitants' and life office pensioners' data. *Insurance: Mathematics & Economics*, 27(3):285–312.

Tabeau, E., van den Berg Jeths, A., and (Eds), C. H. (2001). *A Review of Demographic Forecasting Models for Mortality. Forecasting in Developed Countries: From description to explanation*. Kluwer Academic Publishers.

Wilmoth, J. (1993). Computational methods for fitting and extrapolating the Lee-Carter model

of mortality change. Technical report, Departament of Demography, University of California, Berkeley.

Wilmoth, J. (1996). *Health and Mortality among Elderly Populations*, chapter Mortality projections for Japan: A comparison of four methods, pages 266–287. Oxford University Press.