CONTRIBUTI DI RICERCA CRENOS

MODEL EFFECT ON PROJECTED MORTALITY INDICATORS

Ana Debón
Steven Haberman
Francisco Montes
Edoardo Otranto

WORKING PAPERS

2012/15

CUEC

CENTRO RICERCHE ECONOMICHE NORD SUD
(CRENOS)
UNIVERSITÀ DI CAGLIARI
UNIVERSITÀ DI SASSARI

CRENOS was set up in 1993 with the purpose of organising the joint research effort of economists from the two Sardinian universities (Cagliari and Sassari) investigating dualism at the international and regional level. CRENoS' primary aim is to improve knowledge on the economic gap between areas and to provide useful information for policy intervention. Particular attention is paid to the role of institutions, technological progress and diffusion of innovation in the process of convergence or divergence between economic areas. To carry out its research, CRENoS collaborates with research centres and universities at both national and international level. The centre is also active in the field of scientific dissemination, organizing conferences and workshops along with other activities such as seminars and summer schools.
CRENoS creates and manages several databases of various socio-economic variables on Italy and Sardinia. At the local level, CRENoS promotes and participates to projects impacting on the most relevant issues in the Sardinian economy, such as tourism, environment, transports and macroeconomic forecasts.

www.crenos.it
info@crenos.it

CRENoS – CAGLIARI
VIA SAN GIORGIO 12, I-09100 CAGLIARI, ITALIA
TEL. +39-070-6756406; FAX +39-070- 6756402

CRENoS - SASSARI
VIA TORRE TONDA 34, I-07100 SASSARI, ITALIA
TEL. +39-079-2017301; FAX +39-079-2017312

Title: MODEL EFFECT ON PROJECTED MORTALITY INDICATORS

ISBN: 978 88 84 67 749 5

First Edition: June 2012

Via Is Mirrionis, 1
09123 Cagliari
Tel./Fax 070 291201
www.cuec.it

Model effect on projected mortality indicators

Ana Debón
Universidad Politécnica de Valencia
Steven Haberman
City University, United Kingdom
Francisco Montes
Universitat de València
Edoardo Otranto
University of Messina and CRENoS

Abstract

The parametric model introduced by Lee and Carter in 1992 for projecting mortality rates in the US has been a seminal development and has been widely used since then. Different versions of the model, incorporating constraints on the data, and different adjustment methods have led to improvement. All of these changes have increased the complexity of the model with a corresponding improvement in goodness of fit, however, there is little change in the accuracy of forecasts of life expectancy in comparison with the original Lee-Carter model, according to some authors.

To evaluate to what point the increments in the complexity and computational cost of the models are reflected in the forecast of such indices as life expectancy and modal age at death, among others, we have compared three different models: the original Lee-Carter with one parameter and the Lee-Carter model with two temporal parameters forecasted by means of two independent time series or by means of a bivariate one. The three sets of predictions so obtained are compared using a mixture of block-bootstrap techniques and functional data analysis.

Keywords: mortality indicators, block-bootstrap, functional data analysis.
JEL Classification: C53.

1 Introduction

"Human senescence has been delayed by a decade. This finding, documented in 1994 and bolstered since, is a fundamental discovery about the biology of human ageing, and one with profound implications for individuals, society and the economy. Remarkably, the rate of deterioration with age seems to be constant across individuals and over time: it seems that death is being delayed because people are reaching old age in better health. Research by demographers, epidemiologists and other biomedical researchers suggests that further progress is likely to be made in advancing the frontier of survival — and healthy survival — to even greater ages" (Vaupel, 2010). In this context of recent demographic changes, the development of new models for building life tables and their projection is presented as a point of key research. Life expectancy reflects these changes but its effects are diminished due to its robustness. If, moreover, we bear in mind that life expectancy offers no information as to whether this improvement is the same for different age groups, it is important and necessary to turn to other mortality indicators whose past and future evolution in Spain we are going to study.

An appropriate set of indicators for the study of all these phenomena should include an indicator of infant mortality, life expectancy, the Gini index, the modal age at death. These indicators were applied to Spanish mortality data for the period 1981-2008 for the age range 0 to 99 in Debón et al. (2011). The main conclusions obtained were:

- mortality in Spain improved in both the observed period, 1981-2008, and the forecast period, 2009-2028,
- future improvement is expected to be more sustained than that experienced during the period observed,
- the evolution of the modal age at death, the Lorenz curve and Gini index also confirmed that Spanish mortality displays both expansion and rectangularization,
- the mortality experience of women is better than men, meaning longer life expectancy, higher modal age at death, and a lower Gini index.

All these indicators can be projected using the projections of q_{xt}, obtained from different methodologies, in our study Lee-Carter models (Lee and Carter, 1992; Brouhns et al., 2002; Debón et al., 2008). The errors associated with these estimations can be calculated by means of a block-bootstrap methodology (Liu and Braun, 2010) and a confidence interval can be provided.

The aim of this paper is to evaluate whether different mortality models produce significant differences in the projections of some of these mortality indicators. The reason is that improvement in goodness-of-fit is achieved by increasing the complexity of the models proposed in the literature at the cost of a higher computational complexity. We consider the Lee-Carter model with one or two terms, and in the latter case we consider two independent univariate time series and the case of a bivariate time series.

We want to study if this improvement in goodness-of-fit will be reflected in significant differences between the indicators that we are using, because according to some authors such as Lazar and Denuit (2009) and Debón et al. (2010) there is little change in the ability to forecast life expectancy in comparison with the original Lee-Carter model. In addition, in so far as we are aware, no studies have used formal model selection criteria to compare models based on their projected mortality indicators.

Therefore, the study has a second aspect that we wish to point out: the method used to assess these differences. On the one hand is the use of block-sampling techniques to obtain bootstrap prediction intervals for mortality indicators. Some authors have drawn attention to the narrow intervals obtained by classical bootstrap techniques; however block-bootstrap techniques produce prediction intervals which are more realistic. The second methodological innovation is the use of functional analysis techniques to detect possible differences between the projections of the indicators obtained with the different models. The reasons for the use of functional data are two-fold: firstly, because the projections of one indicator over time are correlated values and proper analysis requires longitudinal data techniques or functional data analysis, as we propose; secondly, because so far the comparison has been carried out by comparing graphs informally, rather than by using objective criteria.

The paper is structured as follows. Section 2 is devoted to describing Dynamic Life Tables and a brief summary of Lee-Carter models. Section 3 presents the definition and properties of the indicators of mortality used in the analysis: life expectancy, the Gini index and modal age at death. Section 4 introduces the block-bootstrap techniques for building prediction intervals. Functional data analysis techniques allowing the comparison of mortality indicators are presented in Section 5. Section 6 devoted to the results of the analysis of Spanish mortality data by means of the above indicators. Finally, Section 7 establishes the conclusions to be drawn from the results in the previous section.

2 Dynamic Life Tables

We consider a set of crude mortality rates $\dot{q}_{xt}$, for age $x \in [x_1, x_k]$ and calendar year $t \in [t_1, t_n]$, which we use to produce smoother estimates, $\hat{q}_{xt}$, of the true but unknown mortality probabilities q_{xt}. A crude rate at age x and time t is typically based on the corresponding number of deaths recorded, d_{xt}, relative to those initially exposed to risk, E_{xt}.

According to Arias (2010), there are two types of life tables: the cohort (or generation) and the period (or current) tables. The cohort life table presents the mortality experience of a particular birth cohort, it reflects the mortality experience of an actual cohort from birth until no lives remain in the group table. It therefore requires data over many years, so instead we normally use the period life table. The period life table presents what would happen to a hypothetical (or synthetic) cohort if it experienced the mortality conditions of a particular time period throughout its entire life.

A dynamic life table is a rectangular mortality data array (q_{xt}), where x denotes age and t denotes calendar time. Each column in this array represents the constituents of the period life table for year t.

2.1 Lee-Carter models

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. There is no consensus among the various authors about the measure of mortality to be modelled in (1), with both m_{xt} and μ_{xt} being used. We prefer to use q_{xt} (with the logit transformation-see later) on account of the good results obtained in our previous work (Debón et al., 2005). Thus, this paper directly models the probability of death, q_{xt} which has the advantage that most actuarial calculations involve life table, although the results can be extended to the force of mortality, μ_{xt} or m_{xt} central mortality rates at age x and time t.

Problems with the estimation of q_{xt} (Lee, 2000) can be avoided by modelling the logit death rates. It is for this reason that we apply this model to the logit transformation of the 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 $r = 2$, and consequently the corresponding models will be denoted $LC1$ and $LC2$, respectively.

The model is well known and will not be considered further in this presentation. A detailed description of the model and its adjustment by different methods can be found in Debón et al. (2008).

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

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

where c is a constant and u_t is 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.

Renshaw and Haberman (2003b) have applied separate univariate ARIMA processes to the first two-period components, the underlying assumption being that the two time series corresponding to k_t^1 and $k_t^{(2)}$ are independent. Applied to Spanish mortality data this assumption also made by Debón et al. (2008) represents a potential 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 (2003a). Lazar and Denuit (2009) have used multivariate time series techniques to forecast age-specific death rates and life expectancies, specifically the dynamic factor analysis and the Johansen cointegration methodology. Futhermore, a comparison with the classical Lee-Carter approach is performed.

3 Mortality Indicators

3.1 Life expectancy

Life expectancy at different ages can be calculated from a dynamic life table. For a year t, the hypothetical number of people alive at the beginning of each age interval $[x, x+1)$ is given by the iterative formula $l_{(x+1)t} = l_{xt}(1 - q_{xt})$, with an arbitrary value $l_{0t} = 100000$. This allows us to calculate the number of deaths $d_{xt} = l_{xt} - l_{(x+1)t}$, and the corresponding number of person-years $L_{xt} = l_{(x+1)t} + a_{xt}d_{xt}$, where a_{xt} is the average time in years that people dying at age x live in $[x, x+1)$ (Chiang, 1960, 1968, 1972). When micro-data of mortality are not available, $a_{xt} = 1/2$. The total number of person-years that would be lived after the beginning of the age interval x to $x+1$ by the synthetic life table cohort is $T_{xt} = \sum_{i \geq x} L_{it}$ (Anderson, 1999). The life expectancy for individuals with of x is given by

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

Life expectancy calculated in Section 6 refers to life expectancy at birth and at 65, e_{0t} and e_{65t}, respectively.

3.2 Modal age at death

The modal age at death is the age associated with the maximum frequency of death. In industrialized countries where infant mortality has decreased dramatically, the modal age of the distribution of deaths is found at older ages. This shift to the advanced ages has been denoted as *expansion* and has a collateral effect on the survival curve, which adopts the form of a rectangle, a phenomenon that has been denoted as *rectangularization*, and is related to an increase in life expectancy.

Additionally, high and dispersed mortality rates are also present in young and intermediate ages, particularly for men. This phenomenon is known as the *accident hump*, as some authors associate it with traffic accidents.

The choice of this indicator is justified by two points outlined by Canudas-Romo (2008),

1. the modal age at death is strongly dependent on the force of mortality prevailing at older ages, and

2. changes in infant mortality are indirectly related to the modal age at death, by having an effect on the modal number of deaths.

It follows that modal age at death may be less robust than life expectancy and therefore can reflect changes in the probability of death q_{xt}, which are not detected with life expectancy.

3.3 Gini index and Lorenz curve

The increase in life expectancy is a consequence of the improved living conditions of individuals, perhaps being the most important improvement achieved for health. However, life expectancy does not provide any information about whether the improvement applies equally to different age groups.

The Gini index is the most common statistical index used in social science for measuring inequality or diversity. It has also been used to measure the contribution of different ages to mortality over time (Llorca et al., 1998). The Gini index is related to the Lorenz curve, which is the curve obtained when we represent the cumulative proportion of a population on the x-axis and the cumulative proportion of years lived by this population on the y-axis. The curve is obtained by joining these points and it is always below the diagonal. The Gini index is twice the area that lies between the diagonal and Lorenz curve, and its value varies from 0 (perfect equality) to 1 (perfect inequality). The value 0 is obtained when all individuals die at the same age, while the value 1 is achieved if the entire population dies at 0 years and one individual dies at an infinite age.

In practice, actuaries work with discrete data from a life table, thus approximate expressions for the abscissas and ordinates of the Lorenz curve can be obtained by means of

$$f_{xt} = \frac{l_{0t} - l_{xt}}{l_{0t}} = 1 - \frac{l_{xt}}{l_{0t}}, \tag{4}$$

and

$$g_{xt} = \frac{T_{0t} - T_{xt} - xl_{xt}}{T_{0t}}, \tag{5}$$

respectively. The situation of perfect equality arises if all individuals die at the same age x_{0t}. In this case, the line consists of only two end-points: $f_{xt} = 0, g_{xt} = 0,\ \forall x \neq x_{0t}$ and $f_{xt} = 1, g_{xt} = 1$, for $x = x_{0t}$.

One of the most widely used approaches for the Gini index is

(Martín-Pliego, 1994),

$$I_{G_t} = \frac{\sum_{x=0}^{(\omega-1)} (g_{xt} - f_{xt})}{\sum_{x=0}^{(\omega-1)} f_{xt}},$$

where w is the last age observed.

The Gini index summarizes the degree of concentration collected by the Lorenz curve in a single value. This feature is certainly an advantage, but has the disadvantage that different concentration configurations, equivalent to different Lorenz curves, can provide the same index value. Hence there is the need to use both the Lorenz curve and Gini index to adequately describe inequality in the length of life.

Other indices such as the Interquartile range (IQR), which also allow the measurement of this unequal contribution, do not have the three desirable basic properties for any measure of inequality (Shkolnikov et al., 2003):

1. *population-size independence*, the index does not change if the overall number of individuals changes with no change in the proportions of years lived,

2. *mean and scale independence*, the index does not change if everyone's years lived changes by the same proportion, and

3. *the Pigou-Dalton condition*, any transfer from an older to a younger individual that does not reverse their relative ranks reduces the value of the index.

4 Block-bootstrap prediction intervals

Mortality predictions are not normally accompanied by measures of sensitivity and uncertainty. Some authors, Pedroza (2006) among others, argue that such measures are necessary and suggest the construction of prediction intervals for the estimations obtained. One way to combine all these sources of uncertainty is to use bootstrapping procedures as Brouhns et al. (2005) and Koissi et al. (2006) do.

In the case of Spain this methodology was used by Debón et al. (2008), who obtained prediction intervals for the predictions provided by the Lee-Carter model with one or two terms. Parametric and non-parametric bootstrap techniques are used, in both cases based on the binomial distribution assumption, as distinct from the work by

Brouhns et al. (2005), Koissi et al. (2006) and Renshaw and Haberman (2008), who employ the Poisson distribution. Another difference to point out is the residuals sampled in the non-parametric case, while Debón et al. (2008) sample over the residuals given by expression (6), Koissi et al. (2006) and Renshaw and Haberman (2008) sample over the deviance residuals.

The narrow prediction intervals obtained by classical bootstrap techniques have attracted the attention of other researchers in this field, Lee and Carter (1992), Lee (2000), Booth et al. (2002) and Koissi et al. (2006), who provide different explanations.

In the case of spatial (two-dimensions: age, time) dependence of residuals ordinary bootstrap is not valid Liu and Braun (2010). Therefore, we aim at obtaining prediction intervals for the mortality indicator by using a residual-based block-bootstrap, as Liu and Braun (2010) propose for deviance residuals, because this technique partially retains the underlying dependence structure in the residuals and generates more realistic resamples (Efron and Tibshirani, 1993). In this paper, block-bootstrap prediction intervals for the mortality indicator are going to be obtained using logit residuals and binomial distribution.

The procedure used is the following. We start with the logit residuals, $\hat{\epsilon}_{xt}$ which are obtained from the original data,

$$\hat{\epsilon}_{xt} = logit(\dot{q}_{xt}) - \widehat{logit(q_{xt})}, \tag{6}$$

ordered in a rectangular array $(\hat{\epsilon}_{xt})$, where x denotes age and t denotes calendar time. To set up a new artificial set of residuals $\hat{\epsilon}^n_{xt}$, we start with an empty rectangular array which has the same dimensions as the original matrix of residuals. The empty array is then partitioned into smaller rectangular blocks. Each block is replaced by a block of the same size, which is randomly selected from the original matrix. This block consists of all residuals in the rectangle to the southeast of the randomly selected element from the original matrix.

Estimated logit rates, $\widehat{logit(q_{xt})}$, are then set and the observed logit rates, for the n$^{\text{th}}$ element of the sample, are obtained from the inverse expression

$$logit(\dot{q}_{xt})^n = \widehat{logit(q_{xt})} + \hat{\epsilon}^n_{xt}.$$

With these new sampled logit rates, a new adjustment of the model is obtained which provides new estimations of the parameters. The process is repeated for the N bootstrap samples, which in turn provide a sample of size N for the set of model parameters, and the k_t's are then projected on the basis of an $ARIMA$ model, obtaining predictions for the mortality rates and the corresponding life expectancy and mortality indicators for the desired

future years. The prediction intervals are obtained from the percentiles, $IC_{95} = [p_{0.025}, p_{0.975}]$.

5 Functional data analysis

The experimental design we have carried out in order to test whether different modeling mortality ratios produce significant differences in the projections of mortality indicators is a two-way design, whose structure is shown in Table 1. It is a balanced design with the same number of repetitions in each cell, n_B, which is equal to the number of block-bootstrap samples. Each one of these repetitions is a set of 20 values, the projected mortality indicator corresponding to the years 2011 to 2030. The factor *model* has three categories, the Lee-Carter with one time parameter, *LC1*, and the Lee-Carter with two time parameters forecasted by means of two independent time series *LC2ind* or by means of bivariate one *LC2bi*. In turn, the factor *residual* also has two categories reflecting the origin of the residuals used in the bootstrap process, *RLC1* and *RLC2*, according to whether they were obtained from the adjustment of the original data with the *LC1* or *LC2* model.

		model		
		LC1	LC2ind	LC2bi
residual	RLC1	n_B	n_B	n_B
	RLC2	n_B	n_B	n_B

Table 1: Experimental design for comparison of functional indicators

As mentioned at the beginning, our aim is to check differences among the projections obtained with different models and different residuals. A classic ANOVA method could be applied with univariate observations, but this is not the case. Each repetition can be considered as a discrete function evaluated in each year of the forecast period. We have to deal with functional data and the use of methods of functional data analysis is allowed. Several authors have dealt with the development of ANOVA techniques for these kinds of data, but not all of them can be used in a design with two fixed factors. Some of them address a single factor (Cuevas et al., 2004; Martínez-Camblor and Corral, 2011) and others address mixed effects (Abramovich and Angelini, 2006; Antoniadis and Sapatinas, 2007). Therefore, we will resort to the method proposed by Cuesta-Albertos and Febrero-Bande (2010) which is based on the analysis of randomly chosen one-dimensional projection of functional data. This method can perform a two way ANOVA with interactions. Following Cuesta-Albertos and Febrero-Bande (2010) we can write,

$$X_i^{mod,res}(t) = m(t) + f^{mod}(t) + g^{res}(t) + h^{mod,res}(t) + \epsilon_i^{mod,res}(t), \quad (7)$$

where m is non-random and describes the overall shape of the projections, $i = 1, \ldots, n_B$, and the functions f^{mod}, g^{res} and $h^{mod,res}$ account for the main effect and interaction of model and residual. Finally, $\epsilon_i^{mod,res}$ are independent random trajectories centered on the mean.

6 Analysis of Mortality Data from Spain

The data used in this analysis come from the Spanish National Institute of Statistics (INE) (see their official web site at http://www.ine.es). In particular, we have worked with published life tables. The crude estimates of q_{xt}, necessary for the models under study, were obtained with the new methodology recently proposed by Spanish National Institute of Statistics (INE) (2009) based on Elandt-Johnson and Johnson (1980), who explain that given complete, continuous-time observations of all births and deaths for all people in a population exposed to the risk of mortality, it is possible to produce direct estimates of the central mortality rates, m_{xt}, by means of

$$\dot{m}_{xt} = \frac{d_{xt}}{1/2P_{xt} + 1/2P_{x(t+1)} + \sum_i \delta_{xti}}, \quad (8)$$

where d_{xt} are deaths in the year t at age x, and P_{xt} and $P_{x(t+1)}$ are the population that are x years old on December 31$^{\text{st}}$ of year t and year $t+1$, respectively. Finally, δ_{xt} is defined as the difference, in years, between the date of death and the birthday in year t, of each individual i who dies in year t with age x. We can obtain $\dot{q}_{xt}$ from,

$$\dot{q}_{xt} = \frac{\dot{m}_{xt}}{1 + (1 - a_{xt})\dot{m}_{xt}}, \quad (9)$$

where a_{xt} is the average number of years that people dying in year t have lived between ages x and $x+1$,

$$a_{xt} = \frac{\sum_{i=1}^{d_{xt}} a_{xti}}{d_{xt}},$$

where a_{xti} is the time in years that individual i, dying in year t at age x, lived between ages x and $x+1$.

The more recent Spanish data set such as the life tables from INE (Spanish National Statistical Institute) are computed with the most suitable methodology and are more accurate than data in the Human Mortality Data base (Muriel et al., 2010).

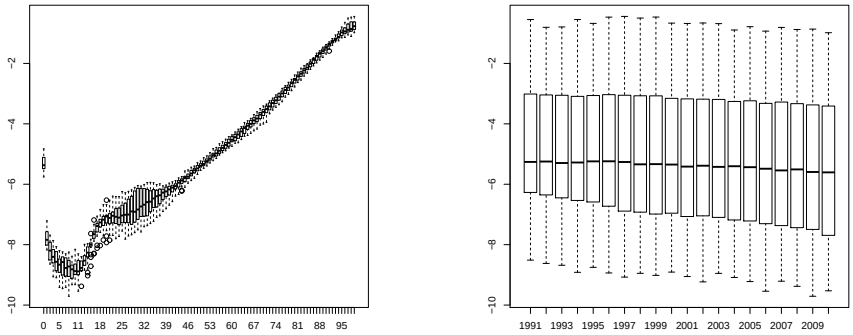


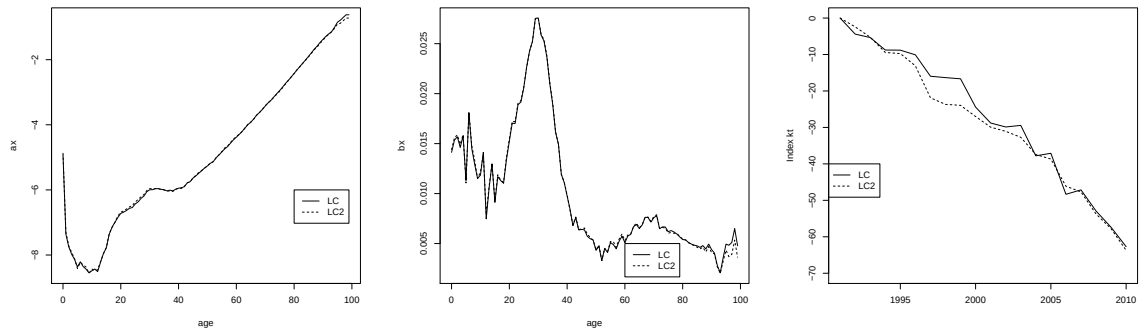
Figure 1: Behaviour of the logit of the crude mortality rates for age and time.

6.1 Model adjustment

The models described in Section 2.1 with the expression (3) for $r = 1$ and $r = 2$ have been used to adjust mortality data in Spain for the period 1991-2010 and a range of ages from 0 to 99. The estimation of the parameters is carried out by means of maximum likelihood methods using the gnm library published by Turner and Firth (2006) of R R Development Core Team (2005), as Debón et al. (2010) propose. The adjustments have been made separately for women and men. The findings for women are similar to those obtained for men, and for the sake of brevity are not reproduced here. Figure 1 shows the behaviour of the logit of the crude mortality rates according to age x and year t.

As the number of parameters estimated in the $LC1$ model is high, $100 \times 2 + 20 = 220$, we prefer to present them in the form of a graph in Figure 2. As a general comment, we must point out that the differences between the a_x estimations obtained for the $LC1$ and $LC2$ models are very small and are not appreciable for most ages. More specifically, the comparison of parameter a_x for both models shows that mortality for men reveals an increase in the age range from 20 to 40 that some authors (Guillen and Vidiella-i-Anguera, 2005) attribute to accident mortality. The hump in Figure 2(a) is slightly more pronounced for model $LC2$ than model $LC1$.

The values of parameter b_x are positive for all ages for both models $LC1$ and $LC2$, indicating that mortality rates in these age groups decrease over time when the values of k_t are negatives. The comparison of parameter b_x for both models for advanced ages indicates that mortality in these age groups is corrected in $LC2$ with the second factor. For k_t, differences are appreciable in general but more specially for the calendar years in the middle of the fitting

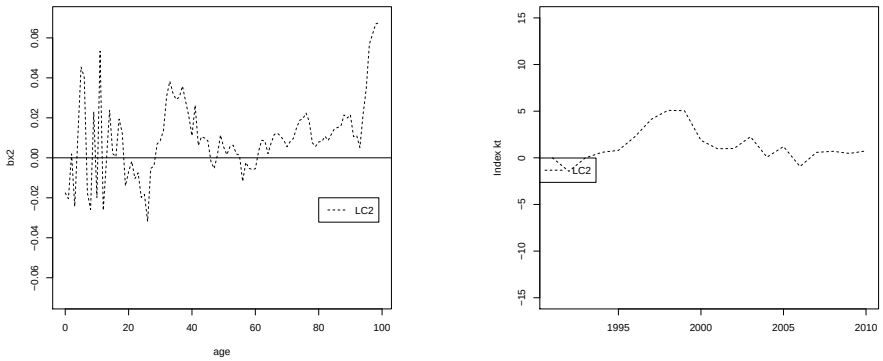


(a) a_x (b) b_x (c) k_t

Figure 2: Estimated values for $LC1$ (solid line) and for $LC2$ (dotted line) for men.

period.

Figure 3(a) shows the estimations obtained with the $LC2$ model corresponding to the second term $b_x^{(2)}$, showing larger values for some low ages (3, 4 and 7), middle ages (25-45) and old ages (over 97), which implies that the effect of adding a second term acts more specifically on these age groups. With regards to $k_t^{(2)}$ Figure 3(b), values are low in general but specially higher for years in the middle (around 1999), indicating that mortality in these years is corrected in the $LC2$ model by the inclusion of the second factor.



(a) $b_x^{(2)}$ (b) $k_t^{(2)}$

Figure 3: Estimated values men (dotted line) for the $LC2$ model.

Renshaw and Haberman (2006) suggest carrying out diagnostic checks on

the fitted model by plotting residuals. These are done in Figures 4 and 5 with the logit residuals for expression (6). Figure 6 is used to show the underlying dependence structure in the logit residuals for both models.

The ordinary bootstrap is not valid in the case of spatial dependence of residuals since it is based on simple random sampling, so we are going to use block-bootstrap suggested by Liu and Braun (2010). Now our main problem is how to select the block size. In the absence of firm theoretical guidance, Liu and Braun (2010) plot a correlogram and a contour map of the original raw residuals and compare these with the resampled residuals. If these plots match reasonably well, this gives confidence in the choice of block size as it supports the fact that they have a similar underlying dependence structure.

With respect to the block sizes, our initial guesses are based on the dependence structure observed for the raw residuals in Figure 6. Firstly, we begin with the raw residuals of model $LC1$ in Figure 6 (left), and as an example we only show the results obtained for block choices 20×5, 5×4 and the ordinary bootstrap i.e. 1×1. Contour maps are plotted in Figures 7, 10 and 13. These plots show the occurrence of large patches of large-positive and large-negative residuals, except for the last one corresponding to the ordinary bootstrap. The other two are highly suggestive of spatial dependence but the structure of the pattern for block size 5×4 is closer to the raw residual pattern shown in Figure 6 (left). In addition, the corresponding prediction intervals for estimated values for the parameters $LC1$ and $LC2$ models (Figures 11 and 12) contain the estimated values for the $LC1$ and $LC2$ parameters from the original data, while the prediction intervals for the remaining block-bootstrap choices in Figures 8, 9, 14 and 15 fail in the estimation of the second terms of the $LC2$ model.

Secondly, we considered the raw residuals for the $LC2$ model in Figure 6 (right) and again, as an example we only show the results obtained for choices 20×5, 5×4 and ordinary bootstrap i.e. 1×1. Contour maps are plotted in Figures 16, 19 and 22. These plots show the occurrence of large patches of large-positive and large-negative residuals, except for the last one corresponding to the ordinary bootstrap. The other two are highly suggestive of spatial dependence, but the structure of block size 5×4 is closer to the raw residual pattern shown in Figure 6 (right). In addition, the corresponding prediction intervals for the estimated values for the parameters for the $LC1$ and $LC2$ models (Figures 20 and 21) contain the estimated values for the $LC1$ and $LC2$ parameters from the original data, while prediction intervals for the remaining block-bootstrap in Figures 17, 18, 23 and 24 fail in the estimation of the second terms of $LC2$ model. Our final choice is to use a 5×4 block size.

Therefore, we have three independent 50 block (5×4)-bootstrap samples

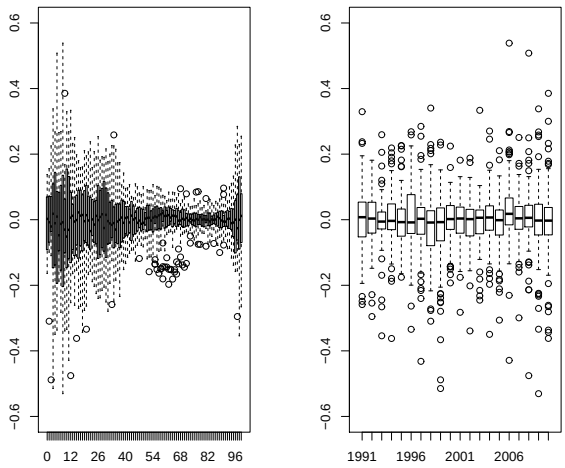


Figure 4: Residuals for the $LC1$ model for men.

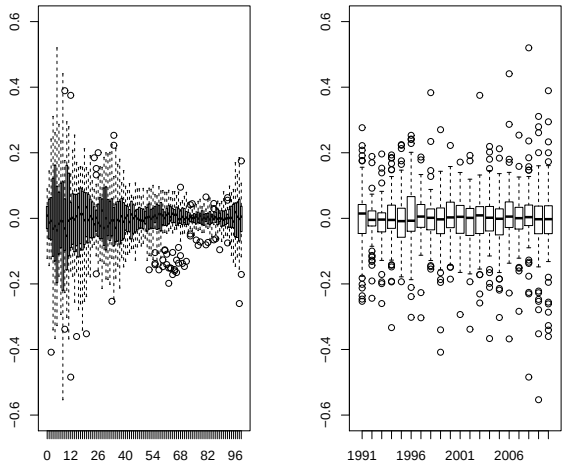


Figure 5: Residuals for the $LC2$ model for men.

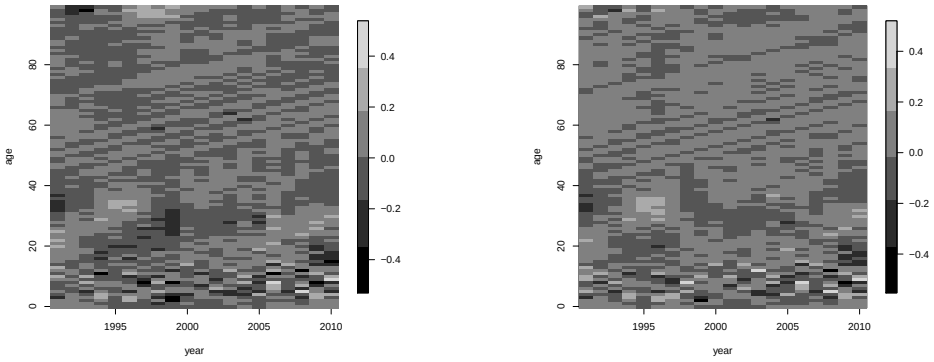


Figure 6: Residuals for age-period for $LC1$ (left) and $LC2$ (right) for men.

obtained from logit $LC1$ residuals and the same from logit $LC2$ residuals, which we are going to call *RLC1* and *RLC2*, reflecting the origin of the residuals used in the bootstrap process and according to whether they were obtained from the adjustment of original data with the *LC1* or *LC2* model as we mentioned in Section 5. And we are going to predict mortality indicators by the Lee-Carter with one time parameter, *LC1*, and the Lee-Carter with two time parameters forecasted by means of two independent time series *LC2ind* or by means of the bivariate model *LC2bi*.

6.2 Prediction intervals for mortality indicators on block-bootstrap

Life expectancy remains the most familiar measure of longevity among demographers, and although it reflects the changes in mortality over time, it does so in a smooth way due to its robustness. This is the reason why other indicators are studied in this paper: modal age at death and the Gini index. The Gini index is a measure of compression or dispersion of the mortality measure and the other two forecasted indicators are measures of the central tendency of mortality, life expectancy and modal age of death. All these indicators share the advantage of summarizing information about mortality independently of the age structure. An indicator of the evolution of mortality with time, widely used by actuaries, is the residual life expectancy at age 65 in year t, e_{x65}. The importance of this indicator arises because 65 is the normal retirement age.

Forecasted mortality indicators for the period 2011-2030 were calculated

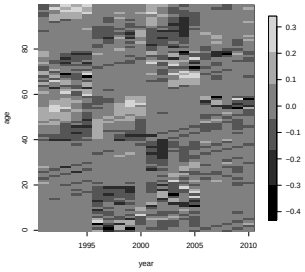
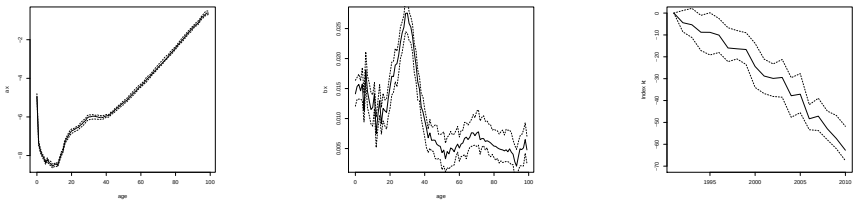
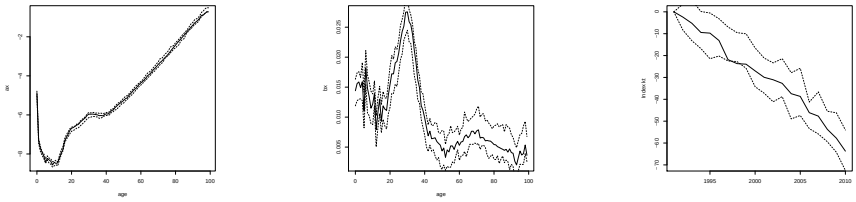


Figure 7: Contour map of block-bootstrap (20×5) residuals from $LC1$ model for men.



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

Figure 8: Estimated values for $LC1$ parameters for block-bootstrap (20×5) from $LC1$ model for men.



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

(d) $b_x^{(2)}$ (e) $k_t^{(2)}$

Figure 9: Estimated values for $LC2$ parameters for block-bootstrap (20×5) from $LC1$ model for men.

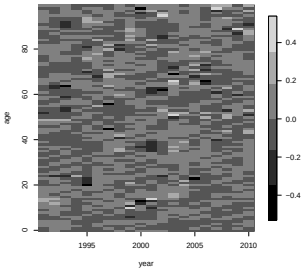
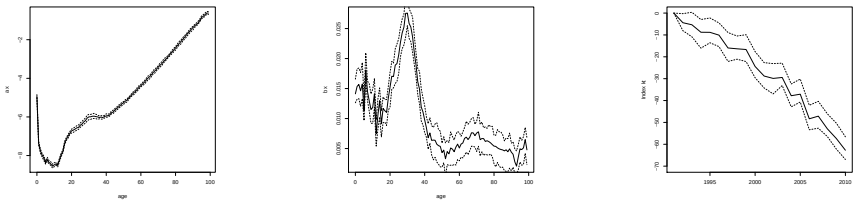
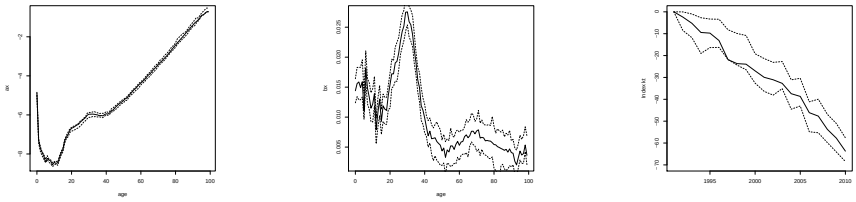


Figure 10: Contour map of block-bootstrap (5×4) residuals from $LC1$ model for men.



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

Figure 11: Estimated values for $LC1$ parameters for block-bootstrap (5×4) from $LC1$ model for men.



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

(d) $b_x^{(2)}$ (e) $k_t^{(2)}$

Figure 12: Estimated values for $LC2$ parameters for block-bootstrap (5×4) from $LC1$ model for men.

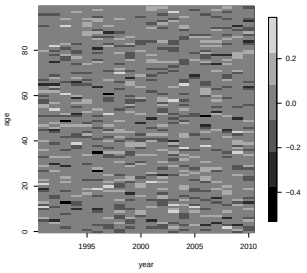
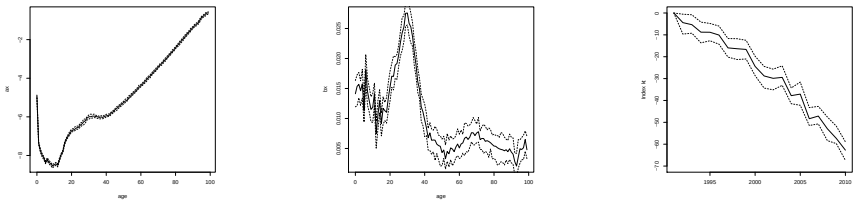
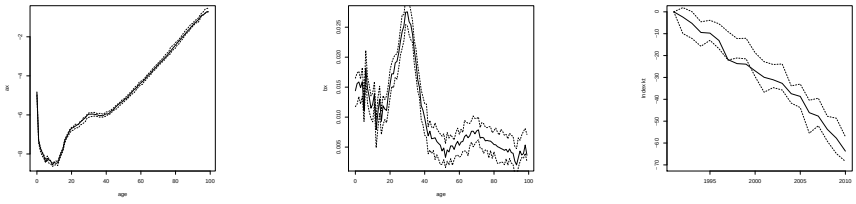


Figure 13: Contour map of ordinary bootstrap (1×1) residuals from $LC1$ model for men.



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

Figure 14: Estimated values for $LC1$ parameters for ordinary bootstrap (1×1) from $LC1$ model for men.



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

(d) $b_x^{(2)}$ (e) $k_t^{(2)}$

Figure 15: Estimated values for $LC2$ parameters for ordinary bootstrap (1×1) from $LC1$ model for men.

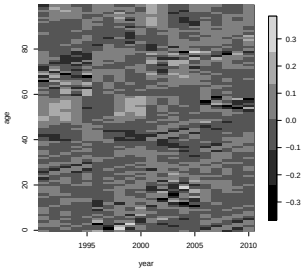
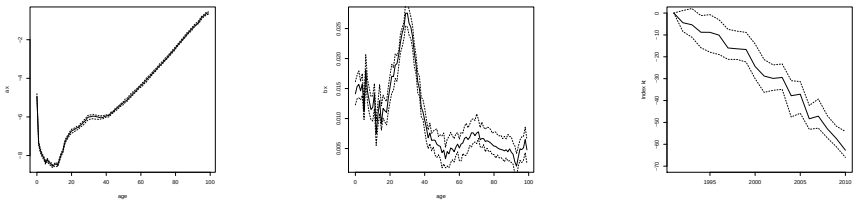
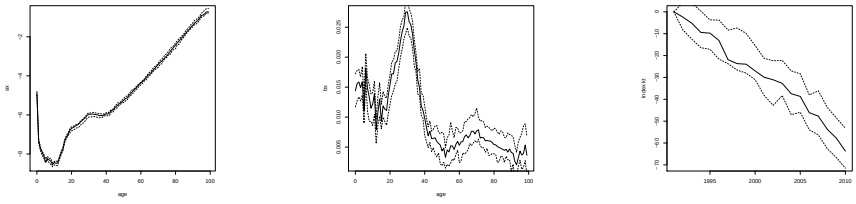


Figure 16: Contour map of block-bootstrap (20×5) residuals from the $LC2$ model for men.



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

Figure 17: Estimated values for $LC1$ parameters for block-bootstrap (20×5) from the $LC2$ model for men.



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

(d) $b_x^{(2)}$ (e) $k_t^{(2)}$

Figure 18: Estimated values for $LC2$ parameters for block-bootstrap (20×5) from $LC2$ model for men.

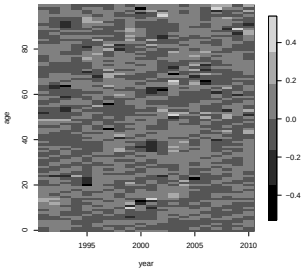
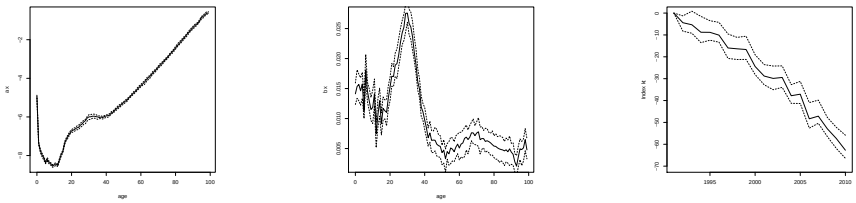
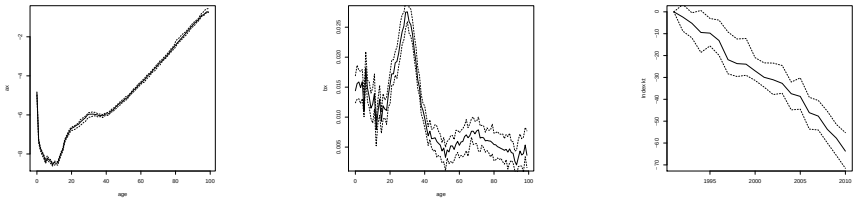


Figure 19: Contour map of block-bootstrap (5×4) residuals from the $LC2$ model for men.



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

Figure 20: Estimated values for $LC1$ parameters for block-bootstrap (5×4) from the $LC2$ model for men.



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

(d) $b_x^{(2)}$ (e) $k_t^{(2)}$

Figure 21: Estimated values for $LC2$ parameters for block-bootstrap (5×4) from the $LC2$ model for men.

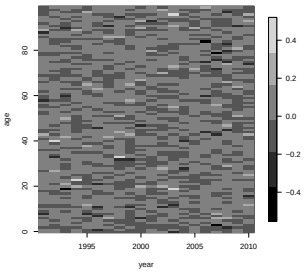
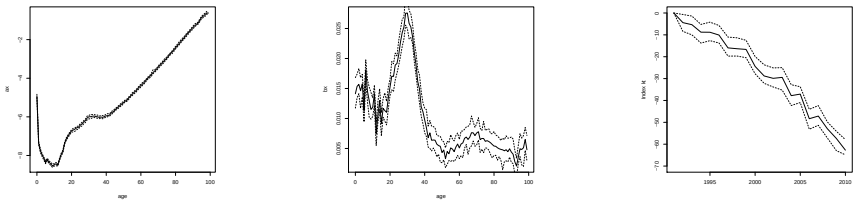
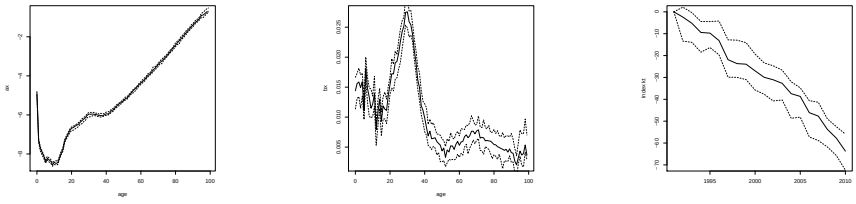


Figure 22: Contour map of ordinary bootstrap (1×1) residuals from the $LC2$ model for men.



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

Figure 23: Estimated values for $LC1$ parameters for ordinary bootstrap (1×1) from the $LC2$ model for men.



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

(d) $b_x^{(2)}$ (e) $k_t^{(2)}$

Figure 24: Estimated values for $LC2$ parameters for ordinary bootstrap (1×1) from the $LC2$ model for men.

using the block-bootstrap technique described in Section 4. Figures 25, 26, 27 and 28 summarize their behavior for all models ($LC1$, $LC2ind$ and $LC2bi$) and six independent samples for each one (three from *RLC1* and three from *RLC2*). As seen in Figures 25, 26, 27 and 28, all the predicted measurements and prediction intervals for all the models and samples behave in a similar way, which makes choosing one model over others difficult. Figures 25, 26, 27 and 28 show the estimated values and the corresponding prediction intervals for the corresponding mortality indicators. Life expectancy, residual life expectancy at age 65 and the modal age at death continue to increase, and the forecast Gini index continues to decrease.

With regard to the width and symmetry of the prediction intervals for life expectancy (Figure 25), the first feature to highlight is the extended width but greater symmetry for the combination $LC1$ model and *RLC1* (Figure 25 left) and for $LC2ind$ and *RLC2* (Figure 25 middle), and narrowness for $LC2bi$ and *RLC2* (Figure 25 right), the explanations for it being the interaction between them. This comment is valid for residual life expectancy at age 65 in Figure 26 and we expect this comment to be valid for the other three indicators also. But the intervals obtained for modal age at death show wider and more irregular intervals (Figure 28), which are narrower and more symmetric for the combination $LC1$ model and *RLC2* (Figure 28 left), $LC2ind$ and *RLC2* (Figure 28 middle) or $LC2bi$ and *RLC2* (Figure 28 right). In the case of the Gini index, we note (Figure 27) the narrowness and symmetry of the prediction intervals for the combination $LC1$ model and *RLC1* (Figure 27 left), for $LC2ind$ and *RLC2* and for $LC2^{\dagger}bi$ and *RLC2*.

6.3 ANOVA for functional data analysis

Figure 29 (left) shows the graphs for the forecasted life expectancy at birth obtained by block-bootstrap for the six combinations of models and residuals. Some possible outliers can be detected, for example $LC2bi \times RCL1$ and $LC2bi \times RCL2$. The ANOVA method used to contrast the effects in model (7) is sensitive to the presence of outliers and can produce erroneous results. It is therefore desirable to detect and suppress them, which is we have done in the graphs on the right. This cleaning process was also carried out with the forecasting of other indicators.

Table 2 summarizes the results for the functional ANOVA. According to

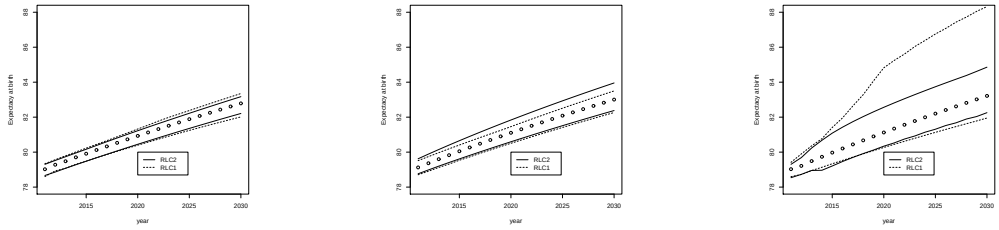


Figure 25: Predicted values for life expectancy at birth for $LC1$ (left), $LC2ind$ (middle) and $LC2bi$ (right) for men

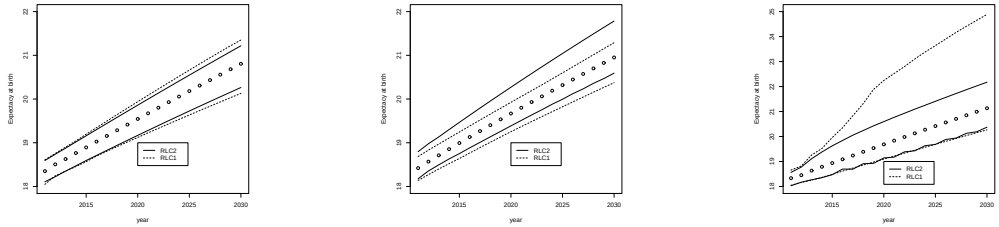


Figure 26: Predicted values for life expectancy at 65 for $LC1$ (left), $LC2ind$ (middle) and $LC2bi$ (right) for men

model 7 the following null hypotheses have been tested,

$$
\begin{aligned}
&H_0^{mod} : f^{LC1} = f^{LC2ind} = f^{LC2bi} = 0 \\
&H_0^{res} : g^{RLC1} = g^{RLC2} = 0 \\
&H_0^{mod,res} : h^{LC1,RCL1} = h^{LC2in,RCL1} = h^{LC2bi,RCL1} = 0 \\
&\qquad\qquad h^{LC1,RCL2} = h^{LC2in,RCL2} = h^{LC2bi,RCL2} = 0.
\end{aligned}
$$

Indicator	Model	Residual	Model×Residual
e_{0t}	R	A	A
e_{65t}	R	A	A
Gini	A	A	A
Modal age	R	A	A

Table 2: Rejection (R) or acceptation (A) of null hypothesis with functional ANOVA.

From Table 2 we conclude that only the main effect of the model is significant, and there is a model effect for all indicators except Gini index.

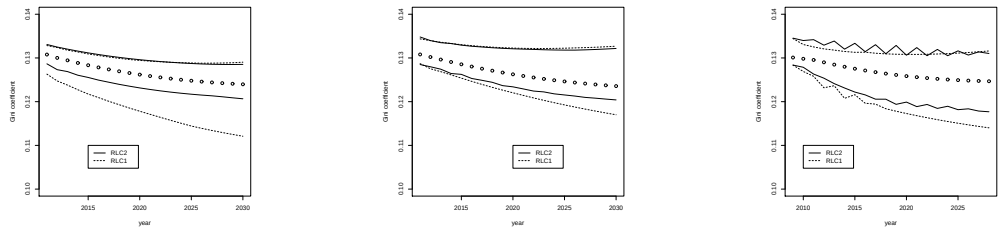


Figure 27: Predicted values for the Gini index for $LC1$ (left), $LC2ind$ (middle) and $LC2bi$ (right) for men.

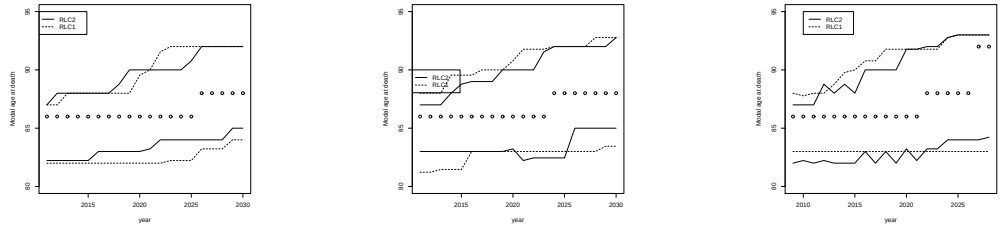


Figure 28: Predicted values for modal age at death for $LC1$ (left), $LC2ind$ (middle) and $LC2bi$ (right) for men.

The results of a multiple comparison between the three models is shown in Table 3. This results are best understood when we look at Figure 30 when the average forecasting functions are shown. Bonferroni correction have been taking into account when obtaining p-values for the multiple comparisons.

7 Conclusions

Predictions for e_{0t} and e_{65t} with Spanish mortality data are larger than those obtained in previous works by the authors (Debón et al., 2008; Debón et al., 2010, 2011). The only reference for comparing the other indicators is Debón et al. (2011) because, as we have pointed out in the Introduction, we found no similar studies in the literature. Larger predictions may be more realistic than those obtained previously and show a better response to the financial challenge that "longevity risk" implies, as noted by the IMF in a recent report (International Monetary Fund, 2012).

One of the aims of this work has been to test the effect of the projection model when forecasting indicators. We have tried to objectify this assessment

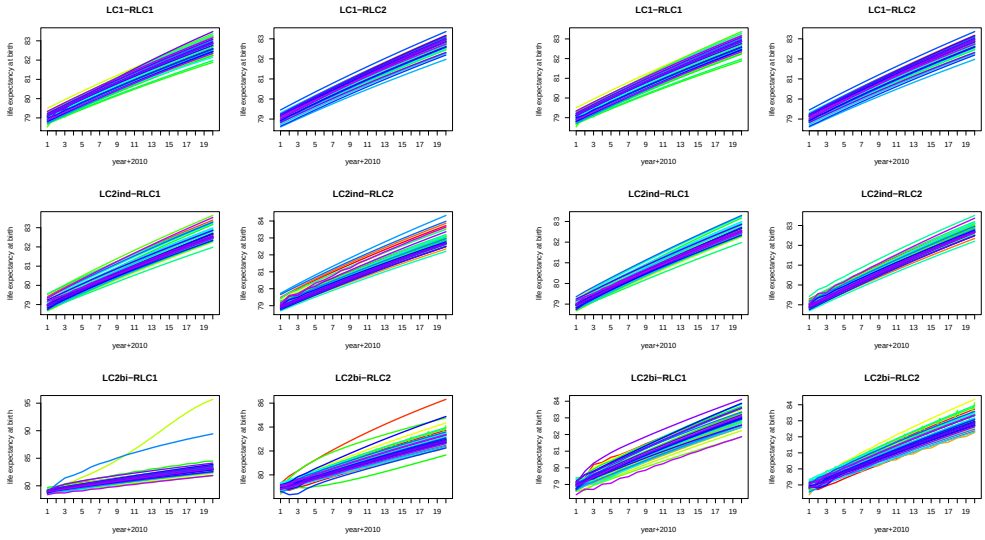


Figure 29: Expected life at birth forecasting with (left) and without (right) outliers.

Indicator	$LC1 = LC2ind$	$LC1 = LC2bi$	$LC2ind = LC2bi$
e_{0t}	A	R	R
e_{65t}	A	R	R
Modal age	A	R	A

Table 3: Rejection (R) or acceptation (A) of null hypothesis with functional ANOVA.

by using techniques associated with functional data analysis combined with a simulation based on block-bootstrap. Specifically, we have used a functional ANOVA to test a two-factor model with interactions (Sections 4 and 5). The result shows that there is only a model effect and subsequent multiple comparisons conclude that the $LC2bi$ model provides higher projections than the others, which would make it more convenient according to the IMF report mentioned above.

In relation to the work of other authors, we should highlight one distinctive feature of the methodology presented here, which is the possibility of comparing the projections of different models with an objective criterion. In short, we propose statistical tools which provide a clear framework for supporting decisions in mortality modelling.

Finally, we follow Gaille (2012) in pointing out that techniques used in

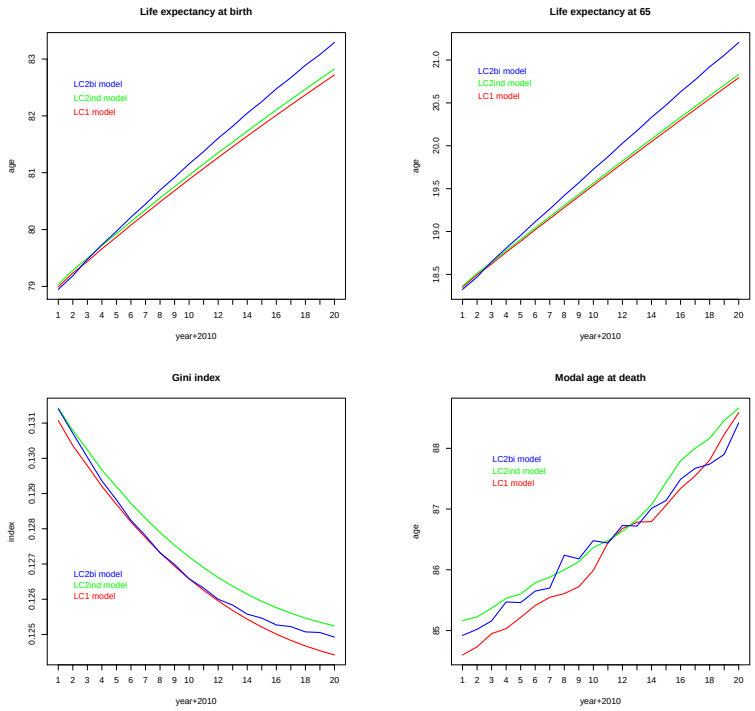


Figure 30: Average forecasting functions for indicators by models.

practice often differ from tools developed in academia. Practitioners tend to simplify models and the presentation of results. Such an approach, according to this author and our results may lead to an underestimation in projected life expectancies which may have important implications for insurance companies and pension funds.

References

Abramovich, F. and Angelini, C. (2006). Testing in mixed-effects fanova models. *Journal of Statistical Planning and Inference*, 136(12):4326 – 4348.

Anderson, R. (1999). Method for constructing complete annual u.s. life tables. *National Center for Health Statistics. Vital and Health Statistics*, 129(2).

Antoniadis, A. and Sapatinas, T. (2007). Estimation and inference in functional mixed-effects models. *Computational Statistics and Data Analysis*, 51(10):4793 – 4813.

Arias, E. (2010). United states life tables, 2006. *National vital statistics reports*, 58(21). Hyattsville, MD: National Center for Health Statistics).

Booth, H., Maindonald, J., and Smith, L. (2002). Applying Lee-Carter under conditions of variable mortality decline. *Population Studies*, 56(3):325–336.

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

Canudas-Romo, V. (2008). The modal age at death and the shifting mortality hypothesis. *Demographic Research*, 19(30):1179–1204.

Chiang, C. (1960). A stochastic study of the life table and its applications, ii: Sample variance of the observed expectations of life and other biometric functions. *Human Biology*, 32.

Chiang, C. (1968). *Introduction to Stochastic Processes in Biostatistics.* Wiley, New York.

Chiang, C. (1972). On constructing current life tables. *Journal of the American Statistical Association*, 67.

Cuesta-Albertos, J. and Febrero-Bande, M. (2010). A simple multiway anova for functional data. *TEST*, 19:537–557.

Cuevas, A., Febrero-Bande, M., and Fraiman, R. (2004). An anova test for functional data. *Computational Statistics and Data Analysis*, 47(1):111 – 122.

Debón, A., Martínez-Ruiz, F., and Montes, F. (2010). A geostatistical approach for dynamic life tables: The effect of mortality on remaining lifetime and annuities. *Insurance: Mathematics and Economics*, 47(3):327 – 336.

Debón, A., Martínez-Ruiz, F., and Montes, F. (2011). Temporal evolution of some mortality indicators. application to Spanish data. In *Living to 100 monograph. Mortality Compression*, `http://livingto100.soa.org/monographs.aspx`. Living to 100 Symposium. Society of Actuaries.

Debón, A., Montes, F., and Puig, F. (2008). Modelling and forecasting mortality in Spain. *European Journal of Operation Research*, 189(3):624–637.

Debón, A., Montes, F., and Sala, R. (2005). A comparison of parametric models for mortality graduation. Application to mortality data of the Valencia region (Spain). *Statistics and Operations Research Transactions*, 29(2):269–287.

Efron, B. and Tibshirani, R. (1993). *An introduction to the boostrap*. Chapman & Hall, New York & London.

Elandt-Johnson, R. and Johnson, N. (1980). *Survival Models and Data Analysis*. Wiley, New York.

Gaille, S. (2012). Forecasting mortality: when academia meets practice. *European Actuarial Journal*, pages 1–28.

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

International Monetary Fund (2012). The financial impact of longevity risk. Technical report.

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):1–20.

Lazar, D. and Denuit, M. M. (2009). A multivariate time series approach to projected life tables. *Applied Stochastic Models in Business and Industry*, 25(6):806–823.

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.

Liu, X. and Braun, W. J. (2010). Investigating mortality uncertainty using the block bootstrap. *Journal of Probability and Statistics*, 2010:385–399.

Llorca, J., Prieto, M., Farinas, C., and Delgado-Rodriguez, M. (1998). Age differential mortality in spain, 1900-1991. *Journal of Epidemiology and Community Health*, 52(4):259–261.

Martín-Pliego, J. (1994). *Introducción a la Estadística Económica y Empresarial*. AC, Madrid.

Martínez-Camblor, P. and Corral, N. (2011). Repeated measures analysis for functional data. *Computational Statistics and Data Analysis*, 55(12):3244 – 3256.

Muriel, S., Cantalapiedra, M., and López, F. (2010). Towards advanced methods for computing life tables. Technical report, Instituto Nacional de Estadística.

Pedroza, C. (2006). A bayesian forecasting model: predicting U.S. male mortality. *Biostatistics*, 7(4):530–550.

R Development Core Team (2005). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0.

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

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. (2006). A cohort-based extension to the Lee-Carter model for mortality reduction factors. *Insurance: Mathematics & Economics*, (3):556–570.

Renshaw, A. and Haberman, S. (2008). On simulation-based approaches to risk measurement in mortality with specific reference to poisson lee–carter modelling. *Insurance: Mathematics and Economics*, 42(2):797 – 816.

Shkolnikov, V., Andreev, E., and Begun, A. (2003). Gini coefficient as a life table function: computation from discrete data, decomposition of differences and empirical examples. *Demographic Research*, 8(11):305–358.

Spanish National Institute of Statistics (INE) (2009). Tablas de mortalidad. metodología. Technical report, Instituto Nacional de Estadística.

Turner, H. and Firth, D. (2006). *Generalized nonlinear models in R: An overview of the gnm package*. R package version 0.9-1.

Vaupel, J. W. (2010). Biodemography of human ageing. *Nature*, 464(7288):536 – 542.

Ultimi Contributi di Ricerca CRENoS

I Paper sono disponibili in: http://www.crenos.it

12/14 *Ahmed A.A. Khalifa, Shawkat Hammoudeh, Edoardo Otranto, Sanjay Ramchander,* "Volatility Transmission across Currency, Commodity and Equity Markets under Multi-Chain Regime Switching: Implications for Hedging and Portfolio Allocation"

12/13 *Raffaele Paci, Emanuela Marrocu,* "Knowledge assets and regional performance"

12/12 *Luciano Mauro, Cesare Buiatti, Gaetano Carmeci,* "The Origins of the Sovereign Debt of Italy: a Common Pool Issue?"

12/11 *Roberto Basile, Stefano Usai,* "Analysis of regional endogenous growth"

12/10 *Emanuela Marrocu, Raffele Paci,* "Different tourists to different destinations. Evidence from spatial interaction models"

12/09 *Ahmed A.A. Khalifa, Shawkat Hammoudeh, Edoardo Otranto*, "Volatility Spillover, Interdependence, Comovements across GCC, Oil and U.S. Markets and Portfolio Management Strategies in a Regime-Changing Environment"

12/08 *Mario Macis, Fabiano Schivardi,* "Exports and Wages: Rent Sharing, Workforce Composition or Returns to Skills?"

12/07 *Anna Bussu, Claudio Detotto, Valerio Sterzi,* "Social conformity and suicide"

12/06 *Luca G. Deidda, Dimitri Paolini,* "Education race, supply of skills and the wage skill premium"

12/05 *Giampiero M. Gallo, Edoardo Otranto,* "The Markov Switching Asymmetric Multiplicative Error Model"

12/04 *Oliviero Carboni*, "A Spatial Analysis of R&D: the Role of Industry Proximity"

12/03 *Juan Gabriel Brida, Nicolàs Garrido, Francesco Mureddu,* "Club performance dynamics at Italian regional level"

12/02 *Emanuela Marrocu, Raffaele Paci,* "Regional Development and Creativity"

12/01 *Bianca Biagi, Maria Giovanna Brandano, Claudio Detotto,* "The effect of tourism on crime in Italy: a dynamic panel approach"

11/23 *Rinaldo Brau, Anna Maria Pinna,* Movements of People for Movements of Goods?"

11/22 *Giorgio Garau, Giovanni Mandras, Lucia Schirru,* "A Statistical Information System supporting Environmental Policies"

11/21 *Emanuela Marrocu, Raffaele Paci, Stefano Usai,* "The complementary effects of proximity dimensions on knowledge spillovers"

11/20 *Giuseppe Pulina,* "Tax Evasion and Presumptive Taxation Methods. A Case Study in Italy: Sector Studies"

11/19 *Juan Gabriel Brida, Giacomo Del Chiappa, Marta Meleddu, Manuela Pulina,* "The perceptions of an island community towards cruise tourism: A factor analysis"

11/18 *Fabio Cerina, Francesco Mureddu,* "Structural Change and Growth in a NEG model"

Finito di stampare nel mese di Settembre 2012
Presso **Copy…Right! studio grafico & stampa digitale**
Via Turritana 3/B – Tel. 079.200395 – Fax 079.4360444
07100 Sassari

www.crenos.it

ISBN 978-88-84-67-749-5