Temporal evolution of some mortality indicators. Application to Spanish data.

A. Debón[1], F. Martínez-Ruiz[2], F. Montes[3]

[1]Dpto. de Estadística e I. O. Aplicadas y Calidad.

Universidad Politécnica de Valencia. Spain

[2] Oficina de Estadística. Ayuntamiento de Valencia. Spain

[3] Dpt. d'Estadística i I. O.. Universitat de València. Spain

Abstract

In Spain, as in other developed countries, there have been significant changes in mortality patterns during the 20th and 21st centuries. One reflection of these changes is life expectancy, which has improved in this period, though the robustness of this indicator prevents these changes from being of the same order as those for the probability of death, q_{xt}. 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.

These indicators are applied to Spanish mortality data for the period 1981-2008, for the age range 0 to 99. To study its future evolution, the mortality ratios have to be projected using an adequate methodology, namely the Lee-Carter model (Lee and Carter, 1992; Brouhns et al., 2002; Debón et al., 2008b). With the aim of incorporating the uncertainty measures suggested by Pedroza (2006) into predictions, confidence intervals are obtained for these predictions. These intervals can be calculated using the

methodology which Lee and Carter apply in their original article for expected lifetime confidence intervals, but they only take into account the error in the prediction of the mortality index k_t obtained from the ARIMA model adjusted to its temporal series, excluding other sources of error such as that introduced by estimations of the other parameters in the model. That is why bootstrap procedures are preferred, as used in Koissi et al. (2006), permitting the combination of all sources of uncertainty.

Keywords: Lee-Carter model, mortality indicators, bootstrap.

Contents

List of Figures

List of Tables

1 Introduction

Important changes in mortality pattern have taken place in the last decades. Olivieri (2001) and Pitacco (2004) point out the existence of two types of process: that of *expansion*, which consists of the movement of the survivors curve towards older age groups, which in turn becomes a movement of the mode of the deaths curve towards those same age groups, and that of *rectangularization*, which is a reference to the rectangular form the survivors curve takes as a consequence of the increase in the concentration of deaths around the mode. Another concept related to rectangularization is the *compression of mortality*, which means a state in which mortality from exogenous causes is eliminated and the remaining variability in the age at death is caused by genetic factors (Fries, 1980). More recently, Canudas-Romo (2008) studied the shifting mortality hypothesis by assessing the changes in the late modal age at death. The study of modal age at death provides a different perspective of the changes in the distribution of deaths and an explanation for the change in mortality at older ages (Cheung et al., 2005).

Regarding evolution of mortality in Spain during the period 1981-2008, Debón et al. (2009) show that, as in other developed countries, infant mortality has undergone a dramatic reduction, mortality for the middle-aged has increased in the last decade and mortality for the higher age groups has stabilized, or shows a slight increase, due to the growth in the ageing population which occurred in recent years (Horiuchi and Wilmoth, 1998).

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, the Lorenz curve, the Gini index, the Interquartile range and the modal age at death. All these indicators can be projected using the projections of q_{xt}, obtained from an adequate methodology, in our study the Lee-Carter model (Lee

and Carter, 1992; Brouhns et al., 2002; Debón et al., 2008b). The errors associated with these estimations can be calculated by means of a bootstrap methodology (Renshaw and Haberman, 2007) and a confidence interval can be provided.

The aim of this paper is the study of Spanish mortality data for the period 1981-2008 and for the age range 0 to 99, by means of a descriptive analysis of the evolution of life expectancy and other indicators. The paper is structured as follows: Section 2 presents the definition and properties of the indicators of mortality used in the analysis: life expectancy, Lorenz curve, Gini index, modal age at death, standard deviation above modal age and shortest age interval for the 50% of deaths; and it introduces a brief summary of Lee-Carter´s model and the bootstrap techniques for building confidence intervals. Section 3 is devoted to the results of the analysis of Spanish mortality data by means of the above indicators. Finally, Section 4 establishes the conclusions to be drawn from the results in the previous section.

2 Description and Prediction of Mortality Rates with Time

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). 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. Therefore it requires data over many years so, instead, normally we 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 of this array represents the period life table for a year t.

2.1 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 age x is given by

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

A tool to measure mortality improvement is the temporary life expectancy from age x to age $x+n$, ${}_ne_{xt}$, defined as person-years lived within a specific age interval, per person alive at the start of the interval (Arriaga, 1984). The expression for ${}_ne_{xt}$ is,

$${}_ne_{xt} = \frac{\sum_{i=x+1}^{x+n} L_{it}}{l_{xt}}.$$

2.2 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*, which 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 measurement is justified by two points outlined by Canudas-Romo (2008),

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

2. changes in infant mortality are indirectly related to the modal age at death, by having an effect in the modal number of deaths. Improvement of infant survival increase the life expectancy at age 0.

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.

Life expectancy calculated in Section 3 refers to expected life time at birth, actually e_{0t}.

2.3 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 in 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 population on

the x-axis and the cumulative proportion of years lived by this population on the y-axis. The curve is obtained joining these points and it is always below the diagonal. The index is twice the area that lies between the diagonal and Lorenz curve, its value varies between 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 die 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

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

and

$$g_{xt} = \frac{T_{0t} - T_{xt} - xl_{xt}}{T_{0t}}, \tag{2}$$

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

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 is certainly an advantage, but has the disadvantage that different concentration configurations, equivalent to different Lorenz curves, can provide the same index value. Hence the need to use, both, the Lorenz curve and Gini index to adequately describe the inequality in 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 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. *Pigou-Dalton condition*, any transfer from a older to a younger individual that does not reverse their relative ranks reduces the value of the index.

2.4 Compression of mortality

In this section we outline some of the compression of mortality measures which are proposed by Kannisto (2000). Specifically:

1. standard deviation of the age at death above de mode, $sd(m+)$, and

2. the shortest age interval in which the 50% of deaths take place C50.

Detailed information about these indicators are available in the above mentioned paper by Kannisto. Wilmoth and Horiuchi (2003) suggest the use of the Interquartile range, but "C50 is a better compression indicator than IQR because it consistently points out a narrower age interval for the same number of deaths - in other words, greater compression. IQR is particularly ill suited to measure compression in high-mortality populations. C50 has the additional advantage that it can be supplemented by other C-indicators which together give a more complete and many-sided picture of compression" (Kannisto, 2000).

2.5 Lee-Carter model and bootstrap confidence intervals

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

$$\ln(m_{xt}) = a_x + b_x k_t + \epsilon_{xt}. \tag{3}$$

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.

Problems with the estimations of q_{xt} (Lee, 2000) 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},

$$\ln\left(\frac{q_{xt}}{1-q_{xt}}\right) = a_x + b_x k_t + \epsilon_{xt}. \tag{4}$$

The model is sufficiently 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. (2008b).

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. 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 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.

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 confidence intervals for the estimations obtained. A 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. (2008b), who obtained confidence 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 turning to the binomial distribution, as distinct from the work by Brouhns et al. (2005) and Koissi et al. (2006) who employ the Poisson distribution. Another difference to point out are the residuals sampled in the non-parametric case, while Debón et al. (2008b) sample over the residuals given by expression (5), Koissi et al. (2006) do so over the deviance.

The procedure used is the following. Starting from the logit residuals, $\hat{\epsilon}_{xt}$, obtained by the original data,

$$\hat{\epsilon}_{xt} = logit(\dot{q}_{xt}) - \widehat{logit(q_{xt})}, \tag{5}$$

a bootstrap sample is drawn, estimated logit rates, $\widehat{logit(q_{xt})}$, are set and the observed logit rates, for the $n-th$ 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 provides 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 mortality ratio and the corresponding life expectancy and mortality indicators for the desired future years. The confidence intervals are obtained from the percentiles, $IC_{95} = [p_{0.025}, p_{0.975}]$.

3 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 micro mortality data reporting individual dates of birth and dates of death. The crude estimates of q_{xt}, necessary for the models under study, were obtained with the new methodology recently proposed by the 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

$$\dot{m}_{xt} = \frac{d_{xt}}{1/2P_{xt} + 1/2P_{x(t+1)} + \sum_i \delta_{xti}}, \tag{6}$$

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^{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 (6),

$$\dot{q}_{xt} = \frac{m_{xt}}{1 + (1 - a_{xt})m_{xt}}, \tag{7}$$

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 with age x, have lived between ages x and $x + 1$.

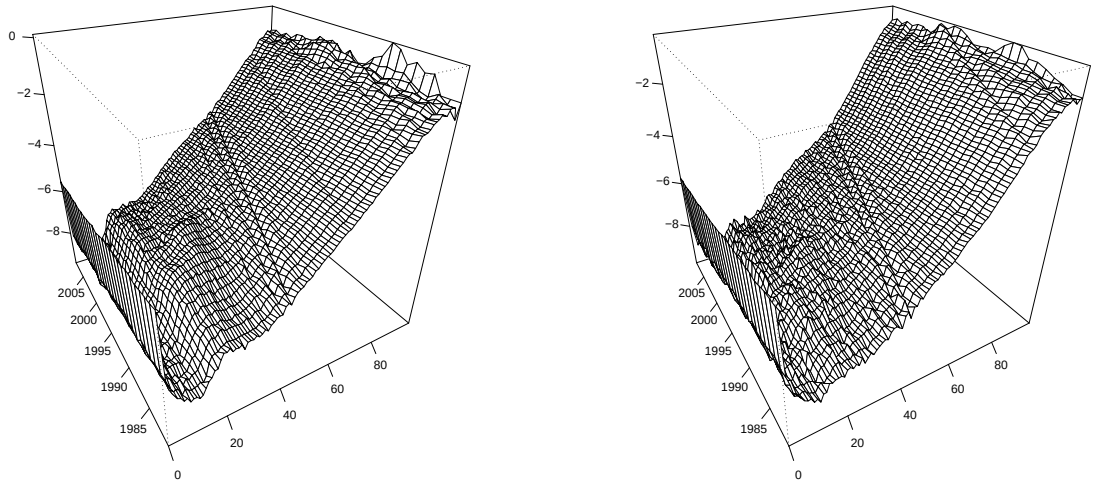


Figure 1: Logit of probabilities of death for men (left) and women (right).

3.1 Period 1981-2008

Figure 1 shows the surface of the logit of the Spanish crude mortality rates for ages from 0 to 99 and period from 1981 to 2008, for men and women.

The mortality indicators presented in the above sections will be used to describe the temporal evolution of mortality in Spain for the age and period ranges mentioned above. The analysis has been done separately for women and men. Tables 1 and 2 summarize these indicators, and Figure 2 (first row) shows the Lorenz curve for 1981 and 2008 and (second row) the difference between the increments of proportions of years lived and population for each age, obtained from (1) and (2) by the following expression,

$$d_{xt} = \Delta g_{xt} - \Delta f_{xt} = [g_{xt} - g_{(x-1)t}] - [f_{xt} - f_{(x-1)t}], \quad x = 1, \ldots, 99.$$

All indicators in Tables 1 and 2 show an improvement in mortality over the period studied, behavior which is consistent with that of other countries in the socio-geographical

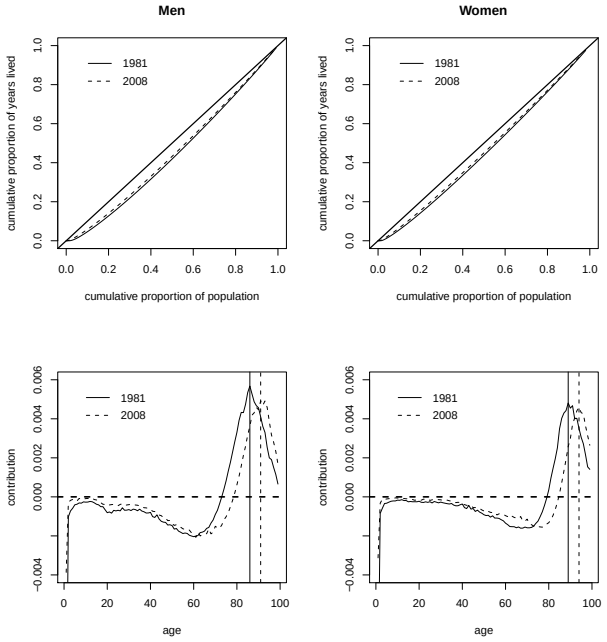


Figure 2: Evolution of Lorenz curve for men (left) and women (right).

environment. This affirmation is partially true for temporary life expectancy, given that the improvement is only seen in the intervals for old ages (66-84) and very old (over 85).

The evolution of Lorenz curves in Figure 2 show a slight trend toward the diagonal, which together with the low values of the Gini index reflects a relative equality between the different ages. In order to appreciate the contribution of each age to the life expectancy at birth we can consider the graphs in the second row. Advanced ages show positive values indicating that the difference between the proportion of deaths and the proportion of years lived by individuals in this age group is in favor of the latter and, therefore, their contribution is greater than could be expected in a completely balanced distribution. It is interesting to note the different behavior of men and women, the maximum positive contributions occur at older ages in women than in men (vertical lines), this fact is consistent with their greater life expectancy. Another interesting aspect is the greater deficits that men show, particularly around 60 years, an age that health experts say is dangerous for men.

year	e_{0t}	${}_{19}e_{1t}$	${}_{44}e_{21t}$	${}_{18}e_{65t}$	${}_{14}e_{85t}$	mode	sd(m+)	C50	Gini Index
1981	72.49	19,90	42.37	12.71	4.77	80	7.70	16	0.1565
1982	73.12	19.91	42.46	12.94	4.95	81	7.51	16	0.1546
1983	72.88	19.90	42.41	12.85	4.73	80	7.71	16	0.1519
1984	73.12	19.91	42.40	12.97	4.94	79	8.38	16	0.1523
1985	73.05	19.91	42.37	12.92	4.77	80	7.77	16	0.1483
1986	73.33	19.91	42.39	13.09	4.92	80	8.02	17	0.1528
1987	73.44	19.92	42.33	13.19	5.08	80	8.15	17	0.1566
1988	73.41	19.92	42.24	13.21	4.94	80	8.09	17	0.1546
1989	73.33	19.92	42.14	13.24	4.97	82	7.24	17	0.1579
1990	73.32	19.92	42.12	13.23	4.87	82	7.12	17	0.1578
1991	73.40	19.92	42.06	13.30	4.95	82	7.19	17	0.1584
1992	73.78	19.93	42.10	13.46	5.12	83	6.88	17	0.1593
1993	73.97	19.93	42.18	13.44	5.03	81	7.71	17	0.1550
1994	74.30	19.94	42.19	13.59	5.10	83	6.92	17	0.1549
1995	74.34	19.94	42.17	13.61	5.07	81	7.79	16	0.1522
1996	74.46	19.94	42.24	13.64	5.07	82	7.38	16	0.1531
1997	75.08	19.94	42.50	13.73	5.11	82	7.44	16	0.1462
1998	75.20	19.94	42.61	13.69	5.05	83	6.90	16	0.1419
1999	75.23	19.95	42.63	13.68	5.00	84	6.41	16	0.1396
2000	75.75	19.95	42.69	13.92	5.25	82	7.55	16	0.1435
2001	76.10	19.95	42.78	14.02	5.33	84	6.68	16	0.1417
2002	76.21	19.95	42.82	14.05	5.32	84	6.67	16	0.1416
2003	76.25	19.95	42.85	14.06	5.23	81	8.00	15	0.1397
2004	76.82	19.96	42.94	14.25	5.51	83	7.25	15	0.1413
2005	76.86	19.96	43.01	14.26	5.42	83	7.19	15	0.1379
2006	77.55	19.96	43.10	14.53	5.72	84	6.97	15	0.1408
2007	77.60	19.96	43.15	14.48	5.62	85	6.47	15	0.1370
2008	78.01	19.96	43.23	14.66	5.69	86	6.07	15	0.1373

Table 1: Evolution of mortality indicators in Spain for men during the period 1980-2008.

year	e_{0t}	${}_{19}e_{1t}$	${}_{44}e_{21t}$	${}_{18}e_{65t}$	${}_{14}e_{85t}$	mode	sd(m+)	C50	Gini Index
1981	78.68	19.93	43.78	14.87	5.29	84	6.51	13	0.1505
1982	79.24	19.94	43.83	15.02	5.56	86	5.91	13	0.1514
1983	79.01	19.93	43.82	14.95	5.29	86	5.73	13	0.1472
1984	79.55	19.94	43.87	15.15	5.52	83	7.21	13	0.1468
1985	79.55	19.94	43.88	15.11	5.31	84	6.61	13	0.1401
1986	79.75	19.94	43.87	15.21	5.43	85	6.31	13	0.1443
1987	80.06	19.95	43.85	15.35	5.59	86	6.00	13	0.1465
1988	80.11	19.94	43.85	15.39	5.53	87	5.56	13	0.1459
1989	80.32	19.94	43.87	15.46	5.56	88	5.17	13	0.1451
1990	80.39	19.94	43.88	15.50	5.57	86	5.99	13	0.1457
1991	80.58	19.95	43.87	15.57	5.61	85	6.45	12	0.1442
1992	81.06	19.95	43.91	15.73	5.79	86	6.12	12	0.1464
1993	81.13	19.95	43.92	15.75	5.74	87	5.61	12	0.1438
1994	81.47	19.95	43.92	15.86	5.87	87	5.70	12	0.1428
1995	81.59	19.95	43.93	15.90	5.85	86	6.18	12	0.1425
1996	81.74	19.95	43.94	15.95	5.87	88	5.28	12	0.1415
1997	82.07	19.96	44.03	16.02	5.94	87	5.79	12	0.1387
1998	82.15	19.96	44.07	16.02	5.89	87	5.75	12	0.1348
1999	82.18	19.96	44.07	16.04	5.86	86	6.23	12	0.1329
2000	82.59	19.96	44.10	16.15	6.10	87	5.87	12	0.1362
2001	82.92	19.97	44.10	16.24	6.24	88	5.47	12	0.1358
2002	82.99	19.97	44.13	16.30	6.19	87	5.88	12	0.1346
2003	82.87	19.96	44.12	16.28	6.05	88	5.33	11	0.1335
2004	83.45	19.97	44.16	16.42	6.38	89	5.05	12	0.1355
2005	83.42	19.97	44.18	16.43	6.27	87	5.93	11	0.1309
2006	84.00	19.97	44.21	16.59	6.60	88	5.60	11	0.1349
2007	83.97	19.97	44.22	16.60	6.51	89	5.09	11	0.1343
2008	84.13	19.97	44.23	16.67	6.53	88	5.57	11	0.1336

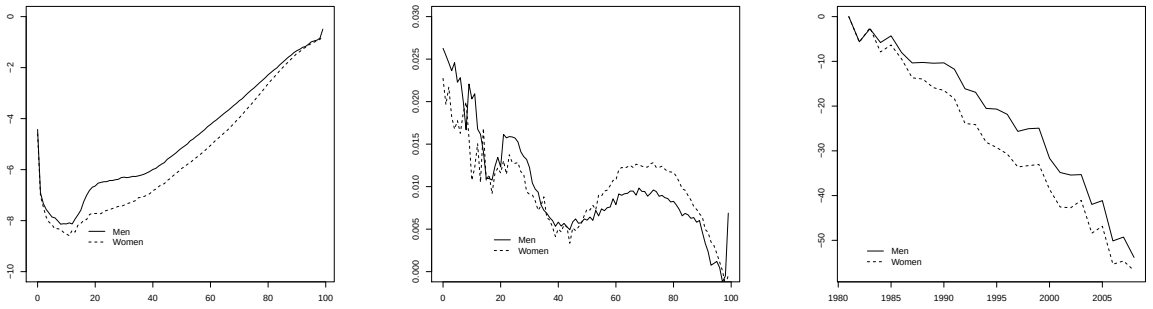
Table 2: Evolution of mortality indicators in Spain for women during the period 1980-2008.

3.2 Period 2009-2028

3.2.1 Model adjustment

The high number of parameters estimated in the *Lee-Carter* model, $100 \times 2 + 28 = 228$ for men and women are presented in the form of a graph in Figure 3.

The comparison of parameter a_x for both sexes shows that mortality for women is lower than for men. The hump in Figure 3(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 accidental mortality (accident hump). The positive values of parameter b_x for all ages, Figure 3(b), indicate that mortality decreases with time. In Figure 3(c), time parameter k_t shows a clearly decreasing trend for both sexes.



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

Figure 3: Estimated parameters for the Lee-Carter model.

Renshaw and Haberman (2006) suggest carrying out diagnostic checks on the fitted model by plotting residuals which has been done in Figure 4 with logit residuals.

3.2.2 Bootstrap confidence intervals for mortality indicators

As seen in Tables 1 and 2, all the measurements behave in a similar way, which makes choosing some over others dificult. The Gini index satisfies the basic properties mentioned in Section 2.3 and is an intuitively meaningful measure, for that reason and for the sake of

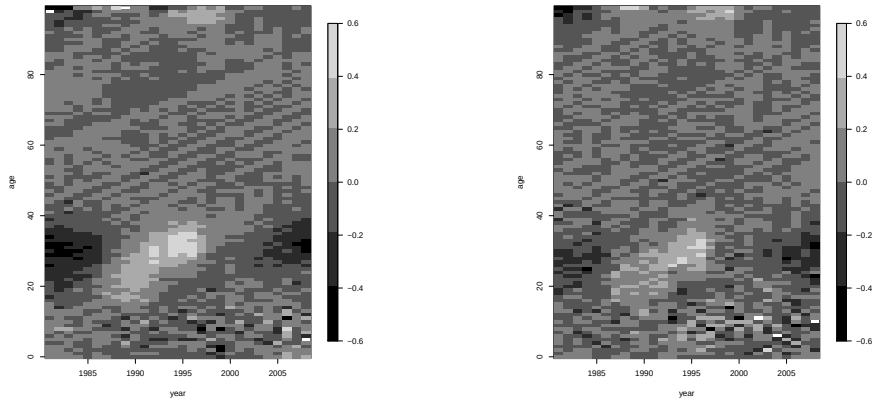


Figure 4: Logit residuals for the Lee-Carter model for men (left) and women (right).

simplicity we have chosen it over compression measures. The other two forecasted indicators are measures of central tendency, life expectancy and modal age of death.

Forecasted mortality indicators for the period 2009-2028 were carried out using the bootstrap technique described in Section 2.5. Table 3 shows the estimations and the corresponding confidence intervals for both sexes. Life expectancy and the modal age at death continue to increase, though more slowly than they did in the previous period.

With regard to the width of the confidence intervals for life expectancy, the first feature to highlight is its narrowness, this fact has attracted the attention of other authors (Lee and Carter, 1992; Lee, 2000; Booth et al., 2002; Koissi et al., 2006) who offer different explanations for it. In the paper by Li et al. (2006) the phenomenon is attributed to the rigidity of the Lee-Carter model structure and to avoid it they relax the structure by incorporating the heterogeneity from each age-period cell. This comment is also valid for the other indicators. The intervals obtained for modal age at death show wider and more irregular intervals. When compared with similar studies carried out using Spanish mortality data, our results on life expectancy are slightly higher than those obtained by Guillen and Vidiella-i-Anguera (2005), Debón et al. (2008b) and Debón et al. (2008a). No comparison is possible for other indicators, because as we point out in the Introduction, we have not found similar studies

for these indicators.

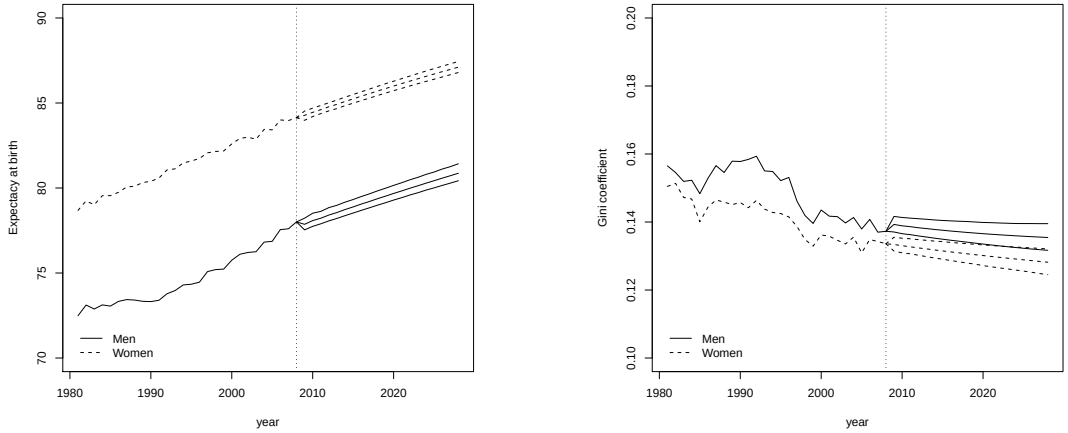


Figure 5: Evolution and bootstrap interval for life expectancy (left) and Gini index (right).

4 Conclusions

Life expectancy remains the most familiar measure of longevity among demographers, and although it reflects the changes in mortality with time, it does it in a smooth way due to its robustness. This is the reason why in the present work other indicators were studied: modal age at death, Lorenz curve and Gini index. Tables 1, 2 and 3 and Figure 2 and 5 summarize their behavior. We can conclude from them,

- mortality in Spain improved in both the observed period, 1981-2008, and the forecast period, 2009-2028,
- the future improvement is 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 tends to the *expansion* and *rectangularization* mentioned in the Introduction,
- the mortality of women is better than men, meaning longer life expectancy and modal age at death, and lower Gini index,

	MEN									WOMEN								
	life expectancy			modal age at death			Gini index			life expectancy			modal age at death			Gini index		
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	$p_{0,975}$
2009	77.54	77.86	78.23	82	85	88	0.1371	0.1392	0.1416	83.98	84.28	84.53	87	89	91	0.1314	0.1333	0.1355
2010	77.74	78.07	78.52	82	85	88	0.1366	0.1388	0.1413	84.20	84.44	84.70	87	89	91	0.1309	0.1330	0.1352
2011	77.89	78.22	78.62	82	85	88	0.1363	0.1386	0.1411	84.37	84.61	84.85	87	89	91	0.1307	0.1327	0.1351
2012	78.07	78.40	78.85	82	85	88	0.1359	0.1383	0.1410	84.54	84.78	85.01	87	89	91	0.1302	0.1323	0.1348
2013	78.21	78.56	78.98	82	85	88	0.1356	0.1380	0.1408	84.67	84.92	85.17	87	90	92	0.1298	0.1321	0.1346
2014	78.37	78.73	79.17	82	85	88	0.1353	0.1378	0.1406	84.84	85.09	85.34	87	90	92	0.1294	0.1317	0.1344
2015	78.53	78.88	79.32	82	85	89	0.1349	0.1376	0.1404	84.99	85.24	85.50	87	90	92	0.1291	0.1315	0.1342
2016	78.68	79.05	79.50	82	86	89	0.1346	0.1373	0.1403	85.14	85.40	85.66	87	90	92	0.1287	0.1312	0.1340
2017	78.83	79.20	79.65	82	86	90	0.1344	0.1371	0.1402	85.28	85.54	85.82	87	90	93	0.1283	0.1309	0.1338
2018	78.99	79.36	79.83	82	86	90	0.1341	0.1369	0.1401	85.44	85.70	85.98	87	90	93	0.1279	0.1306	0.1336
2019	79.13	79.51	79.99	82	86	90	0.1338	0.1367	0.1400	85.58	85.85	86.13	87	91	93	0.1276	0.1303	0.1334
2020	79.29	79.67	80.15	82	86	90	0.1335	0.1365	0.1399	85.73	85.99	86.28	88	91	93	0.1272	0.1301	0.1333
2021	79.43	79.82	80.31	82	86	91	0.1332	0.1364	0.1398	85.86	86.14	86.43	88	91	93	0.1268	0.1298	0.1331
2022	79.59	79.97	80.48	82	86	91	0.1330	0.1362	0.1397	86.00	86.28	86.58	88	91	93	0.1265	0.1295	0.1329
2023	79.72	80.12	80.63	82	87	91	0.1327	0.1361	0.1396	86.14	86.42	86.73	88	91	93	0.1262	0.1293	0.1328
2024	79.88	80.27	80.80	82	87	91	0.1325	0.1359	0.1396	86.27	86.56	86.88	88	91	93	0.1259	0.1290	0.1326
2025	80.01	80.41	80.94	82	87	91	0.1323	0.1358	0.1395	86.40	86.70	87.02	88	91	93	0.1256	0.1288	0.1325
2026	80.15	80.56	81.12	82	87	92	0.1321	0.1357	0.1395	86.53	86.84	87.16	88	91	93	0.1252	0.1286	0.1324
2027	80.29	80.70	81.25	82	87	92	0.1319	0.1355	0.1395	86.66	86.97	87.30	88	91	93	0.1248	0.1283	0.1322
2028	80.43	80.85	81.42	83	88	92	0.1317	0.1354	0.1395	86.79	87.10	87.44	89	91	94	0.1245	0.1281	0.1321

Table 3: Confidence intervals for forecasted Spanish mortality indicators for the period 2009-2028.

- the contribution of different ages to life expectancy is unbalanced, with contributions proportionately larger for older ages, drawing attention to the deficit experienced by men around 60 years.

As a final a comment, it should be noted that the asymmetry of confidence interval is due to the use of logit residuals. According to Renshaw and Haberman (2007), the problem can be partially overcome using deviance residuals as these residuals allow the maintenance of the hypothesis of the initial distribution of mortality measurement.

Acknowledgments

This work was partially supported by a grant from MEyC (Ministerio de Educación y Ciencia, Spain, project MTM2008-05152).

References

Anderson RN. (1999). Method for constructing complete annual U.S. life tables. *National Center for Health Statistics. Vital and Health Statistics*, 2(129).

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

Arriaga E. (1984). Measuring and explaining the change in life expectancies. *Demography* 21 :83–96.

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.

Cheung, S., Robine, J., Tu, E., and Caselli, G. (2005). Three dimensions of the survival curve: Horizontalization, verticalization, and longevity extension. *Demography*, 42(2):243–258.

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.

Debón, A., Montes, F., and Sala, R. (2009). *Tablas de mortalidad dinámicas. Una aplicación a la hipoteca inversa en España.* Fundación ICO. Publicaciones de la Universitat de València, valència.

Debón, A., Montes, F., Mateu, J., Porcu, E., and Bevilacqua, M. (2008a). Modelling residuals dependence in dymanic life tables. *Computational Statistics and Data Analysis*, 52(3):3128–3147.

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

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

Fries, J.F. (1980). Aging, natural death and the compression of morbidity. *New England Journal of Medicine*, 303:130-135.

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

Horiuchi, S. and Wilmoth, J. (1998). Decelaration in the age pattern of mortality at older ages. *Demography*, 35:391–412.

Kannisto, V. (2000). Measuring the compression of mortality. *Demographic Research* , 3(6).

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.

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.

Li, S.-H., Hardy, M., and Tan, K. (2006). Uncertainty in mortality forecasting: an extensión to the classical Lee-Carter approach. Technical report, Waterloo University.

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.

Olivieri, A. (2001). Uncertainly in mortality projections: an actuarial perspective. *Insurance: Mathematics & Economics*, 29(4):231–245.

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

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

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. (2007). On simulation-based approaches to risk measurement in mortality with specific reference to poisson Lee-Carter modelling. *Insurance: Mathematics & Economics*.

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.

Wilmoth, J.R. and Horiuchi, S. (1999). Rectangularization Revisited: Variability of Age at Death Within Human Populations. *Demography*, 36(4):475–495.