REVISTA DE MÉTODOS CUANTITATIVOS PARA LA ECONOMÍA Y LA EMPRESA (16). Páginas 47–67.
Diciembre de 2013. ISSN: 1886-516X. D.L: SE-2927-06.
URL: http://www.upo.es/RevMetCuant/art.php?id=78

A Comparison between General Population Mortality and Life Tables for Insurance in Mexico under Gender Proportion Inequality

Ornelas, Arelly
Department of Econometrics, Riskcenter-IREA
Universitat de Barcelona (España)
Correo electrónico: arelly.ornelas@ub.edu

Guillén, Montserrat
Department of Econometrics, Riskcenter-IREA
Universitat de Barcelona (España)
Correo electrónico: mguillen@ub.edu

ABSTRACT

We model the mortality behavior of the general population in Mexico using data from 1990 to 2009 and compare it to the mortality assumed in the tables used in Mexico for insured lives. We fit a Lee-Carter model, a Renshaw-Haberman model and an Age-Period-Cohort model. The data used are drawn from the Mexican National Institute of Statistics and Geography (INEGI) and the National Population Council (CONAPO). We also fit a Brass-type relational model to compare gaps between general population mortality and the mortality estimates for the insured population used by the National Insurance and Finance Commission in Mexico. As the life tables for insured lives are unisex, i.e. they do not differentiate between men and women, we assume various sex ratios in the mortality tables for insured lives. We compare our results with those obtained for Switzerland and observe very similar outcomes. We emphasize the limitations of the mortality tables used by insurance companies in Mexico. We also discuss the bias incurred when using unisex mortality tables if the proportion of male and female policyholders in an insurance company is not balanced.

Keywords: mortality rates; Lee-Carter; longevity dynamics; Brass-type model; insured population.
JEL classification: I13; I14; I18; C13; C18.
MSC2010: 62-07; 62H12; 62J05; 62P05.

Artículo recibido el 5 de junio de 2013 y aceptado el 10 de octubre de 2013.

Una comparación entre la mortalidad de la población general y las tablas de vida de los seguros en México ante porcentajes desiguales de género

RESUMEN

Interesados en conocer las diferencias entre la mortalidad general y la de un subgrupo de la población, como son los asegurados en una compañía de seguros, hemos ajustado un modelo relacional Brass-Type. Para ello, en primer lugar, hemos modelado el comportamiento de la mortalidad de la población general de México entre los años 1990 y 2009. Hemos ajustado un modelo Lee-Carter, un modelo Renshaw-Haberman y un modelo edad-período-cohorte. Los datos utilizados proceden del Instituto Nacional de Estadística y Geografía (INEGI) y el Consejo Nacional de Población (CONAPO). Una vez estimadas las tasas de mortalidad se han comparado con la mortalidad asumida por las compañías aseguradoras mexicanas. Estas tasas de mortalidad han sido calculadas por la Comisión Nacional de Seguros y Finanzas de México. Como las tablas de mortalidad del seguro de vida son *unisex*, es decir, que no distinguen entre hombres y mujeres, hemos creado diferente escenarios modificando el porcentaje de hombres y mujeres en las tablas de mortalidad. Comparamos los parámetros estimados con los parámetros obtenidos en un análisis con la población Suiza y se observan resultados muy similares. Finalmente, hacemos hincapié en las limitaciones de las tablas de mortalidad utilizadas por las compañías de seguros en México y se analiza el sesgo cuando la proporción de los asegurados masculinos y femeninos en una compañía de seguros no está equilibrada.

Palabras clave: tasas de mortalidad; modelo Lee-Carter; modelo Brass-Type; población asegurada.
Clasificación JEL: I13; I14; I18; C13; C18.
MSC2010: 62-07; 62H12; 62J05; 62P05.

1 Introduction

Insurance policyholders are widely held to be wealthier than the average citizen on the grounds that they can, in fact, afford insurance. Similarly, members of the insured subpopulation are believed to invest more in prevention, resulting in their presenting lower mortality rates than those of the general population for the same age and gender groups. In this article we model the mortality behavior of the general population in Mexico, drawing on data for the period 1990 to 2009, and compare these patterns to the unisex mortality tables used by insurance companies in Mexico. Since this table are unisex we could studied the gender effect in mortality. We conclude that the overly crude life tables used by Mexican insurers may overestimate policyholder mortality. The fact that, in many countries insurance regulators dictate that only unisex life tables are used to price life insurance products raises several questions: First, does the sex ratio have a direct impact on mortality rates? Does the gap in the mortality rates of the general population and the insured population differ between the sexes? Typically, when implementing the unisex rule in a life insurance portfolio, the insurer does not have the same gender and age composition as in the unisex table being applied. In the application reported here we assume different sex ratios in the insured population and conclude that the life tables used in Mexico for pricing insurance products appear to be based on a sex ratio biased in favor of men. The obligation in Europe to avoid gender discrimination when pricing life insurance led us to undertake this analysis, in which we seek to contribute to a better understanding of mortality models in this context. Guillén (2012) discusses the possible effects of unisex pricing in life insurance.

1.1 Background

During the 20th century, there was a substantial decrease in mortality, attributable in the main to the reduction in infant mortality from infectious diseases. In Mexico, for example, life expectancy at birth has increased from 57.04 years in 1960 to 76.47 in 2009, equivalent to an average year-on-year growth of 0.39 years. If we only consider the last two decades, average life expectancy has risen by 0.27 and 0.33 years for females and males, respectively.

The deceleration in the life expectancy rate can be attributed to the fact that while some causes of death, such as accidents and violence-related deaths, have been reduced, there has been an increase in such chronic conditions as heart-related disease. Moreover, among the Mexican male population, the changes can also be attributed to emigration (typically to the United States), a factor that constantly modifies the composition and structure of the population (Camposortega, 1997).

In short, during the second half of the 20th century, mortality has fallen markedly in early childhood, while it has only slightly increased at later ages. These changes have given rise to the need for further research in the insurance sector, and above all in the field of risk management, insofar as life expectancy affects mortality forecasting, longevity risk, reserve calculations, annuities, pension plan design and premiums for life products.

Over recent decades, various models have been proposed to describe mortality. Yet as early as the 19th century, reference was already being made in the literature to such concepts as mortality, annuities and adverse selection risk. Indeed, the first life tables to consider alterations in mortality were created in the United Kingdom, constructed with data for insurers in the period 1900 to 1920. Pitacco et al. (2009) provides an interesting review of the historical background to early mortality tables.

The early study of mortality led to the development of graduation models, which have been used to smooth crude mortality data and to analyze mortality behavior. This line of study in the actuarial literature has drawn heavily on the so-called Mortality Laws of Heligman-Pollard, Gompertz-Makeham, among others. With the advent of new statistical methods, mortality analysis for insurance portfolios was revolutionized. Recently introduced models include Poisson models, which fit the number of deaths given the number of individuals exposed to risk. These models also include a trend component that captures the evolution of mortality over time given a fixed gender and age. Survival models, such as the Cox proportional hazards model, fit the probability that an individual survives to a particular age; hence, it can model the duration of life conditional to various characteristics, including lifestyle and health. In this case, age is considered a continuous variable (Cox and Oakes, 1984). All these models can include risk factors other than gender in their analyses of mortality. Obvious examples include: time trends and other regressors that can incorporate interaction effects between variables. In non-parametric models, such as generalized additive models (GAM), specification provides the potential for better data fits than those provided by other methods. This is attributable to the fact that the effects induced by risk factors can allow more flexible forms (Hastie and Tibshirani, 1990). Other approach in non-parametric models, focused in older ages, is described in Fledelius et al. (2004).

Contemporary predictive mortality modeling usually takes into account cohort (i.e., a group of the population born in the same period, typically, a specific year) effects. The advantage of modeling cohort mortality is that individuals in the same cohort evolve together and are, therefore, exposed to the same external phenomena, such as epidemics. This approach facilitates mortality prediction because the effects of ageing can vary from one cohort to another.

Among the models that take cohort or age effects into account, we find: the Lee-Carter model (1992), the Renshaw-Haberman model (2006), the Currie model (2006) and the Cairns, Blake and Dowd model (2006) and its generalizations. In the section that follows we discuss these models and then adopt them in order to analyze and predict mortality rates for the Mexican population. Lee-Carter model is a reference model in mortality, it has been used by several authors, for instance, Brouhns et al. (2002), Renshaw and Haberman (2003), Delwarde et al. (2003), Debón et al. (2008) and more recently Betzuen (2010) compared Spanish mortality rates with the ones estimated by the Lee-Carter model.

2 Methodology

In this section we describe various models that are currently being used to estimate demographic trends in the general population. We also present a model for comparing a reference mortality population (for example, the general population) with another population (for example, a population of insured policyholders).

Using the following notation, the log crude rate of mortality $m(x,t)$ can be calculated as:

$$\log m(x,t) = \log \frac{d(x,t)}{l(x,t)}, \tag{1}$$

where $l(x,t)$ is the number of living individuals and $d(x,t)$ the number of deaths at time t of individuals aged x.

2.1 Mortality Models

2.1.1 Lee-Carter model

This model combines time series to adjust decreasing mortality rates exponentially. Mortality is represented by $m(x,t)$, which is the crude mortality rate at time t of individuals aged x . A single index guides the trend and it is used to obtain both predictions and confidence intervals (Lee and Carter, 1992). The model is specified as follows:

$$\log[m(x,t)] = \beta_x^{(1)} + \beta_x^{(2)}\kappa_t + \epsilon_{x,t}, \tag{2}$$

where κ_t is an index that varies over time and which represents variations in mortality per year. $\hat{\beta}_x^{(1)}$ and $\hat{\beta}_x^{(2)}$ are parameters to be estimated and the residuals $\epsilon_{x,t}$ is a set of independent random variables. $\hat{\beta}_x^{(1)}$ represents the time average for $log[m(x,t)]$ and $\frac{\partial log(m(x,t))}{\partial t} = \beta_x^{(2)}\frac{\partial \kappa_t}{\partial t}$ is the rate change over time. A negative $\beta_x^{(2)}$ is expected for certain ages, indicating that at these ages mortality increases. The error term is assumed to have zero mean and constant variance σ_ϵ^2.

As specified above, the model parameters are not identified, i.e. given the date, the solution of the system of equations when fitting equation (2) by maximum likelihood (assuming that the error term is normally distributed) is not unique. Consequently, constraints have to be imposed to guarantee identifiability:

$$\sum_x \beta_x^{(2)} = 1 \qquad \text{and} \qquad \sum_t \kappa_t = 0. \tag{3}$$

$\sum_t \kappa_t = 0$ implies that the parameter $\hat{\beta}_x^{(1)}$ is in fact the empirical average over the time of the age profile in age x. Given that $\beta_x^{(2)}$ and κ_t have restrictions, they are estimated by using singular value decomposition applied to the difference $log[m(x,t)] - \hat{\beta}_x^{(1)}$. The estimates of κ_t can be modeled as an ARIMA time series by using the Box-Jenkins methodology.

2.1.2 Renshaw-Haberman model

The model proposed by Renshaw and Haberman (2006) seeks to identify and represent cohort effects; thus, clearly establishing generational changes. As such, the term $\beta_x^{(3)}\gamma_{t-x}^{(3)}$ is added to the Lee-Carter model:

$$\log[m(x,t)] = \beta_x^{(1)} + \beta_x^{(2)}\kappa_t^{(2)} + \beta_x^{(3)}\gamma_{t-x}^{(3)} + \epsilon_{x,t}. \tag{4}$$

Here $\beta_x^{(3)}$ is the rate change over time for the cohort effect $\gamma_{t-x}^{(3)}$. As in the Lee-Carter model, this new model also presents identification problems, so a further restriction is added to those already included in (3):

$$\sum_x \beta_x^{(3)} = 1 \qquad \text{and} \qquad \sum_{x,t} \gamma_{t-x}^{(3)} = 0. \tag{5}$$

The parameters are estimated using an iterative process. Details can be found in Renshaw and Haberman (2006).

2.1.3 Age-Period-Cohort model

The Age-Period-Cohort model was proposed by Currie (2006) and usually refers to the following model:

$$\log[m(x,t)] = \beta_x^{(1)} + \frac{1}{n_a}\kappa_t^{(2)} + \frac{1}{n_a}\gamma_{t-x}^{(3)} + \epsilon_{x,t}, \tag{6}$$

where n_a is the total number of ages. Currie used p-splines to estimate the parameters. The following constraints are required:

$$\sum_t \kappa_t^{(2)} = 0 \qquad \text{and} \qquad \sum_{x,t} \gamma_{t-x}^{(3)} = 0. \tag{7}$$

2.1.4 Model Selection Criteria

The following two criteria can be used to compare the mortality models fitted. One of them is the Bayesian information criterion (BIC), which penalizes models that are over parameterized. It is defined as:

$$BIC_k = \hat{l}_x - \frac{1}{2}n_k \log N, \tag{8}$$

where n_k is the number of parameters, $\hat{l}_x$ is the maximum of the likelihood function and N is the number of observations.

Models were also evaluated based on standardized residuals,

$$Z(x,t) = \frac{d(x,t) - l(x,t)\hat{m}(x,t;\phi)}{\sqrt{l(x,t)\hat{m}(x,t;\phi)}}, \tag{9}$$

where $d(x,t)$ is the number of individuals aged x who die at time t and $l(x,t)$ is the number of people still living at age x and time t . Residuals $Z(x,t)$ are assumed to be distributed identically and independently as a standard normal law. More details can be found in Cairns et al. (2009).

2.2 Brass-type relational model

The Brass-type relational model is a mortality model that links the survival probabilities of a given group with a set of benchmark survival probabilities (Brass, 1971). In others words, the model assumes that the mortality rates of two different populations might be related in terms of their survival probabilities. For instance, it allows the mortality of insured policyholders to be compared with the mortality of the general population. The association is established through a linear or logistic regression so as not to omit information from the reference rates (Brouhns et al., 2002).

In the actuarial sciences, the Brass-type relational model can also be used to measure adverse selection. If the mortality rates of the insured population are higher than those of the general population in adult ages, then individuals may have private information that is not disclosed to the company and they underwrite insurance because they know they have a greater risk of mortality than individuals of the same age. If there is no adverse selection, the opposite phenomenon is likely to be found.

As mentioned above, as insurance policyholders tend to have higher incomes than those of the average citizen, they tend to spend more money and resources on healthcare and, consequently, their risk of mortality can be expected to be lower than that of a similar person in the uninsured population. This situation is known as counter adverse selection or positive selection.

The Brass-type relational model is specified as follows:

$$f(m(x,t)) = \beta_0 + \beta_1 f(m(x,t)_{ref}) + \epsilon_{x,t}, \tag{10}$$

where $m(x,t)_{ref}$ refers to the crude mortality rate of a reference group. Link function $f(\cdot)$ can be the logarithm or the logistic function. This model can be extended to incorporate interactions between mortality rates and time t , as Gatzert and Wesker (2011) have proposed, or to incorporate interactions between mortality rates and age x. Their model includes a time trend component t:

$$\log[m(x,t)] = \beta_0 + \beta_1 \log[m_{ref}(x,t)] + \beta_2 \log[m_{ref}(x,t)] \cdot t + \epsilon_{x,t}. \tag{11}$$

Since the mortality rates available are unisex we vary the proportion w of women in the population, at a given age x and time t, and calculate unisex mortality rates for the general population: $m_w(x,t) = wm_f(x,t) + (1-w)m_m(x,t)$, where $m_f(x,t)$ and $m_m(x,t)$ are the crude mortality rates for the female and male populations, respectively. We assume that the proportion is stable and constant for all adult age groups. This assumption can be relaxed, as we know that after the age of 65 the proportion of women in each cohort increases substantially due to their longer life expectancy.

When comparing a general population and an insured population, the simplest Brass-type relational model is:

$$\log[m_{ins}(x,t)] = \beta_0 + \beta_1 \log[m_w(x,t)] + \epsilon_{x,t}, \tag{12}$$

where $m_{ins}(x,t)$ is the crude mortality rate for insured individuals aged x at time t.

Parameter β in equations (10), (11) and (12), should be interpreted as the parameters in linear regression. Consequently β_0 is the intercept estimate,then, it represents the mean outcome when $log(m(x,t)_{ref}) = 0$. Finally β_1 is the slope coefficient, so that, it represent the expected increment or decrement (depending on the sign) in the response per unit change in $log(m(x,t)_{ref})$.

3 Data description

The data for this study were collected from the Mexican National Institute of Statistics and Geography (INEGI) and the National Population Council's (CONAPO) website. Mortality rates $m(x,t)$ were computed as the ratio between the number of people alive and the number of registered deaths of individuals aged x in year t (see equation (12)). The data reflect the demographic behavior of the 32 Mexican states between 1990 and 2009, and are arranged by sex for ages ranging from 0 to 100 years.

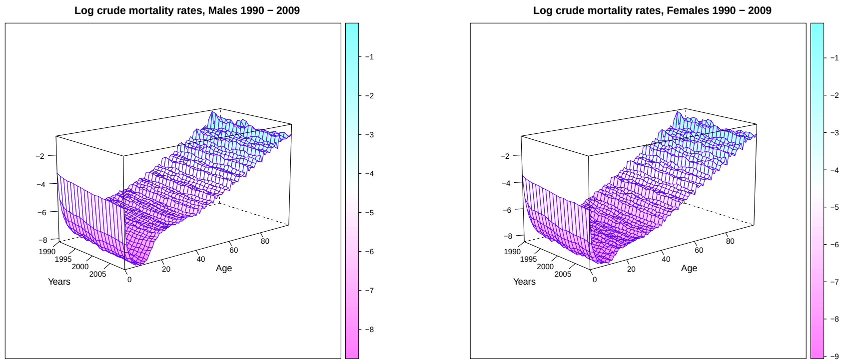
In order to avoid any biased information, data without any description of sex and/or age were excluded. The omission of these details is usually the result of death having occurred at a very early age and so the individual's sex is not recorded, or, in the case of the elderly, the exact age at death is sometimes not known. The percentage of such data is small.

We only have information for those living in the years when a census was held. Information for the years without a census was completed by linear interpolation in accordance with the method described by Delwarde and Denuit (2003).

From information disaggregated by sex and age, the interpolated $l(x,t)$ data for the periods t between 1992 and 1994, 1996 and 1999, 2001 and 2004 and finally between 2006 and 2009 were obtained.

4 Results

We analyzed the evolution in mortality rates separately by gender over t years and for different ages x. Figure 1 shows that the log mortality rate curves for men and women are very similar during childhood and adulthood, the largest differences being found in young adulthood when men are more likely to be involved in accidents. This phenomenon, known as the accident hump, has not changed greatly



(a) Males (b) Females

Figure 1: Log crude mortality rates for Mexican population from 1990 to 2009: (a) Males, (b) Females

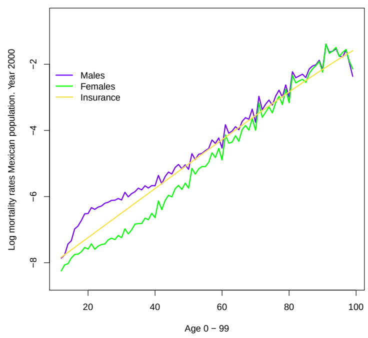


Figure 2: Log mortality rates for Mexican population in 2000. Log insured mortality rates (CNSF) Mexico.

over the two decades studied here.

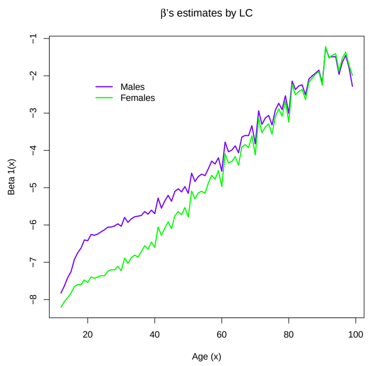
In the following sections we present the results obtained from the Lee-Carter model, the Renshaw-Haberman model and the Age-Period-Cohort model. The estimation was performed separately for men and women, from the ages of 0 to 99 and between the years 1990 to 2009. We only report estimates for ages 12 to 99 years, because the minimum age for underwriting a contract in most insurance companies is twelve. To obtain these estimates we used Lifemetrics software implemented in R version 2.13.

Since the mortality rates for the insured population are available, Figure 2 compares the logarithm of the mortality rate for an insured policyholder and the logarithm of the mortality rate for the Mexican population in 2000. After the age of 70, the mortality rate of an insured individual seems to be lower than that of the general population. If the mortality for insured individuals is lower than that of the general population, then we would expect the longevity of insured individuals to be much longer than that observed in the general population. This may give rise to some bias in the reserve calculations or in the number of expected claims if the general population's mortality is considered in the forecasts. Later we address this problem by studying the Brass-type relational model described in equation (12). The CNSF rates do not represent the real mortality shape.

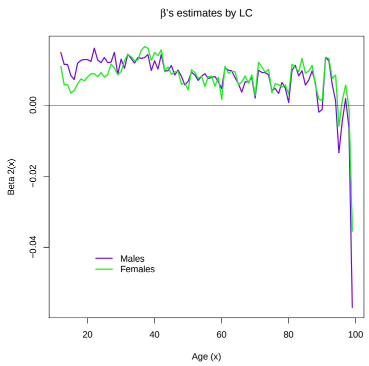
4.1 Results of the general mortality models

4.1.1 Lee-Carter model

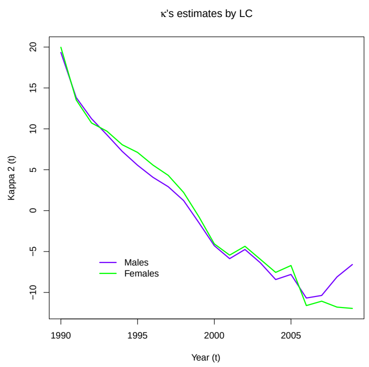
Figure 3 presents the parameter estimates for the Lee-Carter model for the Mexican population data. The number of estimated parameters is 88 for both $\hat{\beta}_x^{(1)}$ and $\hat{\beta}_x^{(2)}$ and 20 for $\hat{\kappa}_t$, resulting in a total of 196 parameters.



(a) $\hat{\beta}_x^{(1)}$



(b) $\hat{\beta}_x^{(2)}$



(c) $\hat{\kappa}_t^{(2)}$

Figure 3: Estimated parameters for the Lee-Carter model. Male and Female aged 12 to 99 from 1990 to 2009. Mexican population.

Note that $\hat{\beta}_x^{(1)}$ estimates for women are lower than those for men (Figure 3), except at ages above 80 . Applying exponential function yields, crude mortality rates have a median of 0.007 for women (interquartile range, denoted as IQR, is 0.001,0.045) and a median of 0.010 for men (IQR 0.003,0.054).

A further feature is that $\hat{\beta}_x^{(2)}$ has a positive sign for both sexes, and for all ages ranging from 12 to 78. Changes in the logarithm of mortality rates are very low because $\hat{\beta}_x$ parameter estimates are around 0.01. The median of $\hat{\beta}_x^{(2)}$ is 0.009 (IQR 0.006,0.011) for women and 0.009 (IQR 0.006,0.012) for men. Substantial differences between men and women are shown for ages between 12 and 40, when the female logarithmic mortality rates grow steadily. After the age of 40, the curves do not differ by sex.

Inspection of the values of $\hat{\kappa}_t$ show that they decrease for both genders. If we analyze the differences between male and female values of $\hat{\kappa}_t$, we note that from 1995 to 2005 , women presented a slight increase in their mortality level, approaching that of men for 2000. However, after 2005 the opposite occurs with the male $\hat{\kappa}_t$ increasing.

4.1.2 Renshaw-Haberman model

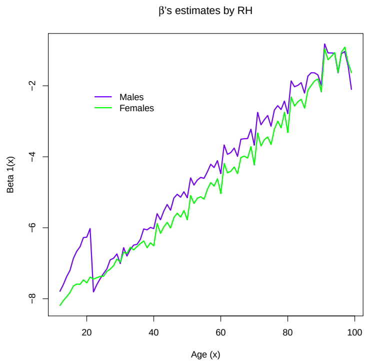
Parameter estimates were obtained for the Renshaw-Haberman model using an iterative process. Figure 4 shows the values for both the male and female population aged 12 to 99 years. The number of estimated parameters is 88 for every $\hat{\beta}_x$ (3×88), 20 for $\hat{\kappa}_t$ and 119 for $\hat{\gamma}_{t-x}$, resulting in a total of 403 estimated parameters. In general, these estimated parameters are similar to those estimated using the Lee-Carter model. There is no perceived difference between the ages of 22 and 40, though the most notable difference is just before the age of 22 . Mortality rates are higher for men than they are for women.

Values of $\hat{\beta}_x^{(2)}$ for women have a median of 0.008 (IQR 0.003,0.011), while for men this median is equal to 0.007 (IQR 0.005,0.014). Estimates for men are above zero. These estimates increase at the age of 23 up to 0.033, followed by a gradual drop until the age of 60. From this age until 90, we can observe a much more stable behavior, with some occasional fluctuations. After the age of 90, estimated parameters have negative values. Estimates are more stable between the ages of 12 and 40 for women than they are for men. For young women, there is a slight fall and these parameters eventually become negative. Between the ages of 45 and 60, values oscillate and are very close to zero. From the age of 61 onward, values begin to rise steadily, reaching their peak at the age of 91. After this age, observed parameters have similar values, with the exception of ages 95 and 99.

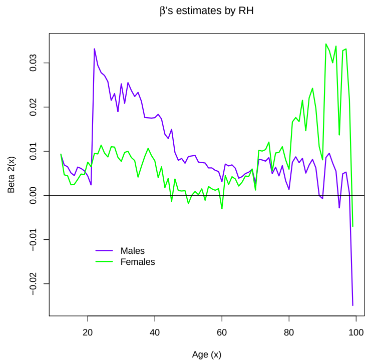
The slope indicates some abrupt increases in the logarithm of the mortality rate as age increases. There is a sudden increase in the mortality of men at the age of 23. This reaffirms the fact that the mortality of young men rises while that of young females does not. Unlike the Lee-Carter model, the Renshaw-Haberman model indicates that there are differences between the mortality of the sexes at older ages. This is most notable in women, whose mortality has slightly increased in recent years.

We know that the parameter $\hat{\beta}_x^{(3)}$ indicates the rate of change for period mortality rates. Using this model, estimated values for women have a median equal to 0.010 (IQR 0.007,0.013), while those for men have a median of 0.009 (IQR

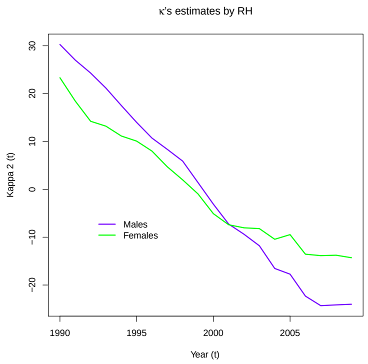
0.007,0.016). We also observe that, for men, oscillations are common around different ages. Meanwhile, females present a very stable behavior, with the most significant difference occurring after the age of 50.



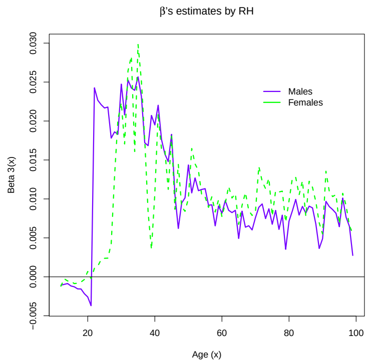
(a) $\hat{\beta}_x^{(1)}$



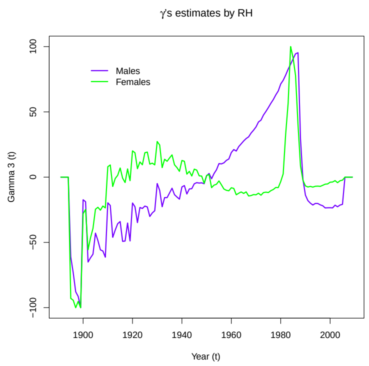
(b) $\hat{\beta}_x^{(2)}$



(c) $\hat{\kappa}_t^{(2)}$



(d) $\hat{\beta}_x^{(3)}$



(e) $\hat{\gamma}_{t-x}^{(3)}$

Figure 4: Estimated Parameters Renshaw-Haberman. Men and women aged between 12 and 99 between 1990 and 2009. Mexican population.

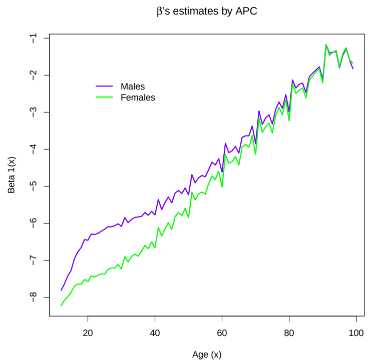
We know that the parameter $\hat{\beta}_x^{(3)}$ indicates the rate of change for period mortality rates. Using this model, estimated values for women have a median equal to 0.010 (IQR 0.007,0.013), while those for men have a median of 0.009 (IQR 0.007,0.016). We also observe that, for men, oscillations are common around different ages. Meanwhile, females present a very stable behavior, with the most significant difference occurring after the age of 50.

The evolution of $\hat{\kappa}_t$ represents the overall path of mortality over time, as if all ages were considered together. The male $\hat{\kappa}_t$ value presents an almost linear behavior up until 2005. From this year on, we observe a long curved path leading to the last year studied here, i.e. 2009. In the case of the women, the slope is less steep, but it maintains a negative value. We see that the decrease in male mortality rates has been faster than it has been for women if we consider all ages. In general, beginning in 2000, the evolution in female rates presents a less steep curve than that of men.

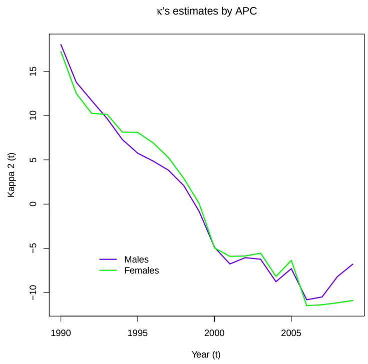
If we focus on the year of birth and treat it as a series with non-constant variance, we note that gender behavior is different for those born between 1903 to 1950, and then for those born from 1980 onward. For these generations born after 1980 the $\hat{\gamma}_{t-x}$ for women is higher than it is for men. Values for men show a linear shape, with a mostly rising trend, while for women the values present a convex and concave shape before and after 1950 respectively. Peaks are reached between 1984 and 1987. A possible explanation for this phenomenon is the strong earthquake that hit Mexico in 1985, affecting several states in the country.

4.1.3 Age-Period-Cohort model

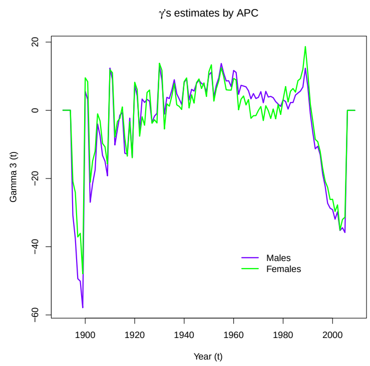
This is a special case of the previous model, with $\hat{\beta}_x^{(1)}$ and $\hat{\beta}_x^{(2)}$ equal to one. The number of estimated parameters is 88 for $\hat{\beta}_x$, 20 for $\hat{\kappa}_t$ and 119 for $\hat{\gamma}_{t-x}$, resulting in a total of 227 estimated parameters. Like the Lee-Carter model, this model emphasizes the differences in mortality by gender among the young. These differences disappear at older ages. This is an expected behavior of the log mortality rates. Figure (5) presents the estimated parameters.



(a) $\hat{\beta}_x^{(1)}$



(b) $\hat{\kappa}_t^{(2)}$



(c) $\hat{\gamma}_t^{(2)}$

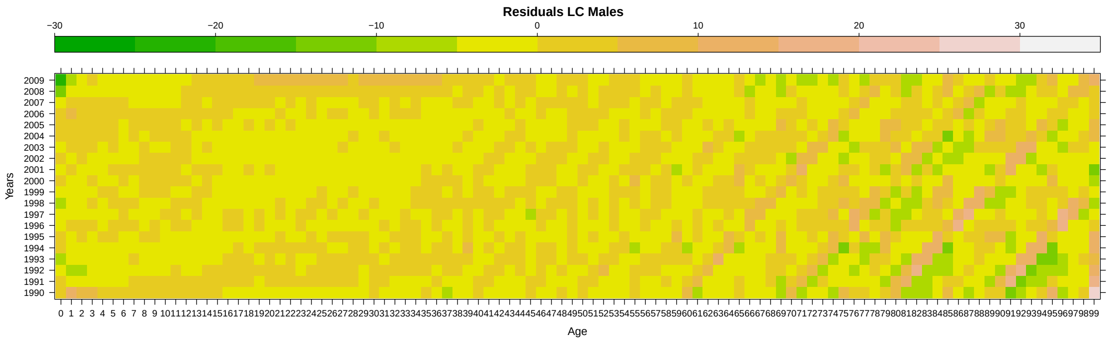
Figure 5: Estimated Parameters Age-Period-Cohort. Male and Female aged 12 to 99 from 1990 to 2009. Mexican population.

Values of the parameters $\hat{\kappa}_t$ were similar to those obtained by the Lee-Carter model especially in men. Prior to 2000, this model predicted lower values, above all between 1995 and 1999. For women, changes are constant and no linear behavior can be identified. If we take the year of birth we see there are no differences between the sexes and the respective behaviors are very similar.

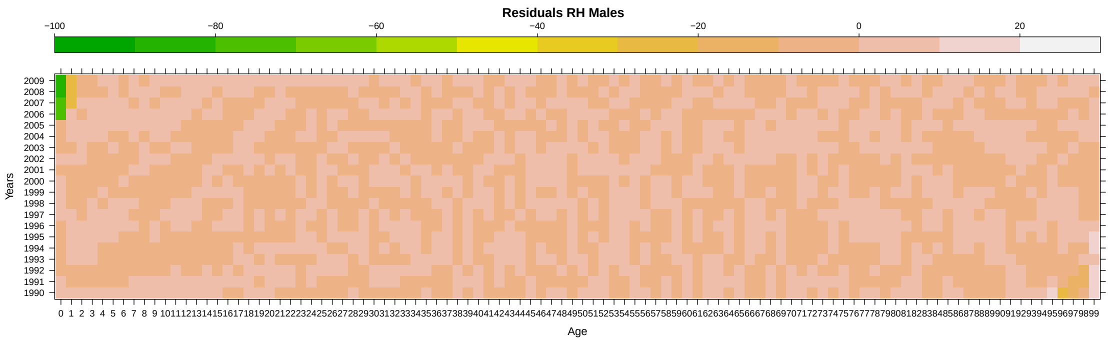
4.1.4 Model comparison

To evaluate the quality of the models, we plotted residuals in order to identify patterns. The x-axis displays age x and the y-axis shows the year t. We would expect residuals to be distributed randomly, i.e. they should be independent and identically distributed.

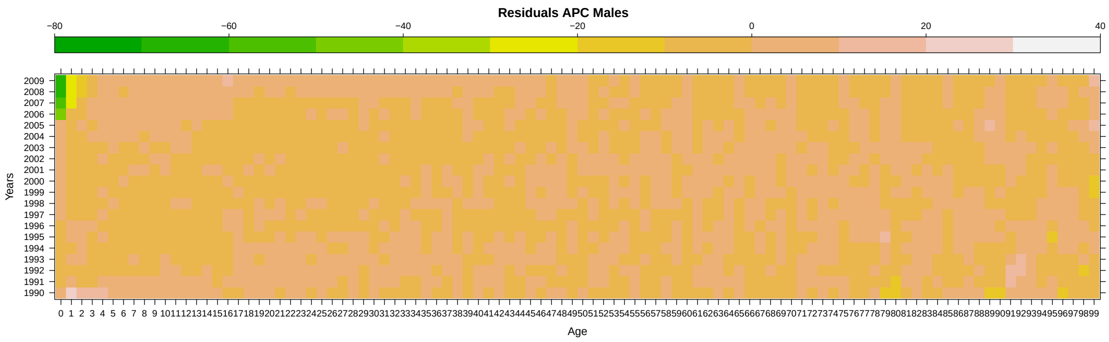
The residuals of the Lee-Carter model show that cohort effects were not captured. There are diagonal lines through years, mainly after 60, for both male and female. Residuals have a median of -0.35 (IQR -1.96,1.60) for men and -0.3 (IQR -1.88,1.52) for women. The largest residuals correspond to the estimated mortality



(a)

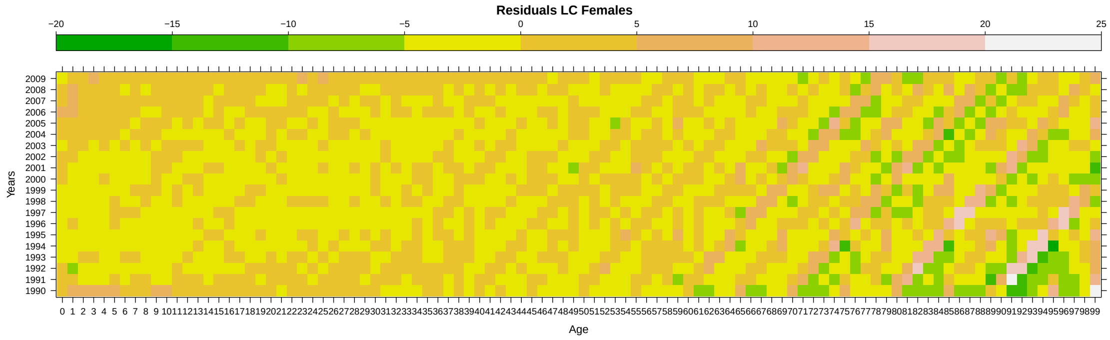


(b)

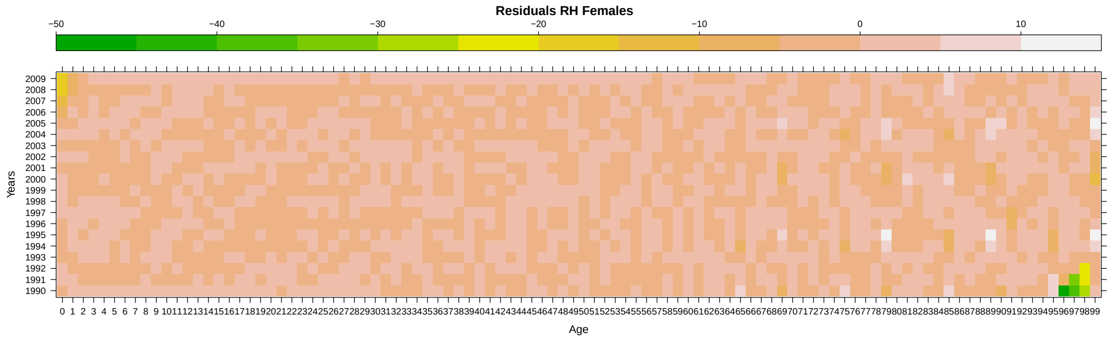


(c)

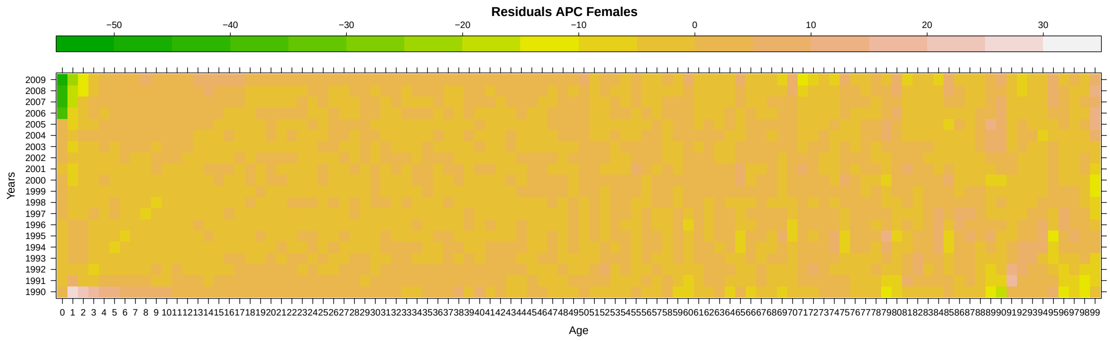
Figure 6: Standardized Residuals $Z(x,t)$ (a)Lee-Carter , (b)Renshaw-Haberman and (c)Age-Period-Cohort models for males. Largest residuals appear in dark $Z(x,t) < 0$ or soft colors$Z(x,t) > 0$



(a)



(b)



(c)

Figure 7: Standardized Residuals $Z(x,t)$ (a)Lee-Carter , (b)Renshaw-Haberman and (c)Age-Period-Cohort models for females. Largest residuals appear in dark $Z(x,t) < 0$ or soft colors$Z(x,t) > 0$

Model	Male	Female
Lee-Carter	$-23,105.68$	$-23,559.63$
Renshaw-Haberman	$-14,963.81$	$-14,302.79$
Age-Period-Cohort	$-19,058.61$	$-18,273.52$

Table 1: Bayes Information Criterion

at older ages.

In contrast, the residuals of the Renshaw-Haberman model do not show a clear pattern, as cohort effects were included in the model and these have a median of -0.06 (IQR -1.13,1.13) for men, and a median of -0.02 (IQR -1.03,1.02) for women. Residuals are larger through ages 96 to 99 , and also in the early nineties. We can also observe that residuals are larger in women than they are in men, most presenting negative values. Nevertheless, overall residuals are close to zero.

Finally, the residuals of the Age-Period-Cohort model have a median equal to -0.22 (IQR -1.61,1.58) for men and -0.06 (IQR -1.43,1.4) for women. No large differences are found at any particular age or year (see Figures 6 and 7). As in the Renshaw-Haberman model, diagonals in residuals do not appear in the Age-Period-Cohort model. Residuals are larger for this particular model than for the previous ones.

According to the Bayes information criterion (BIC), the Renshaw-Haberman model provides a better fit than the other models. This model includes cohort effects and presents a better distribution of the residuals.

4.2 Results for the Brass-type relational model

In this subsection we analyze the results obtained from the Brass-type relational model. We carried out a linear regression to compare the mortality reported in the insurance industry with the mortality in the general population. We consider as reference mortality rates those estimated for the general population in Mexico with the Lee-Carter, the Renshaw-Haberman and the Age-Period-Cohort models, as described in sections 2.1.1, 2.1.2 and 2.1.3. A Brass-type relational model is used to analyze the mortality rates calculated by the National Insurance and Finance Office for insured lives. The fundamental idea is that the insured individuals' mortality behavior is significantly different from the mortality experienced by the general population (Brouhns et al., 2002).

In our setting, we assume that the mortality rates for the insured population are unisex, which is the case of the Mexican mortality tables for the insured population, there being no male-female segregation.

Thus, using the Brass-type relational model, we investigate the influence of the gender composition of insurance portfolios on the management of mortality risk when only a unisex table is available. We see that for insurers, using a single unisex mortality life table can be hazardous.

Estimated parameters and the BIC for this model are shown in Table 2. The first column indicates the percentage of women in the population. This means that we considered different values for weight w in the linear regression. The estimates of the slope are equal to one when the proportion of women is exactly 25%, which means that there is a similarity between the insured population and the general population when we consider that one individual in every four is a woman. The significant negative intercept indicates that the mortality of the insured population is lower than the mortality of the general population.

Examining the logarithm rates adjusted by the Brass-type relational model from ages 12 to 18, the larger the proportion of women, the larger we can expect the difference between the mortality rate of the insured and that of the general population to be. The opposite occurs in the age interval 19-40; at these ages the differences are small.

% Women	Lee-Carter			Renshaw-Haberman			Age-Period-Cohort		
	$\hat{\beta}_0$	$\hat{\beta}_1$	BIC	$\hat{\beta}_0$	$\hat{\beta}_1$	BIC	$\hat{\beta}_0$	$\hat{\beta}_1$	BIC
0%	−0.06	1.04*	33.36	−0.04	1.05*	22.06	−0.01	1.05*	24.17
25%	−0.14*	1.00*	17.88	−0.12	1.00*	3.48	−0.10	1.00*	7.79
50%	−0.22*	0.95*	8.70	−0.20*	0.95*	−7.86	−0.19*	0.96*	−1.18
75%	−0.30*	0.91*	6.00	−0.27*	0.91*	−11.18	−0.27*	0.91*	−2.64
100%	−0.38*	0.87*	8.77	−0.35*	0.87*	−7.58	−0.35*	0.87*	1.96

Table 2: $\hat{\beta}$ by Brass-Type relational model: $\log[m_{ins}(x,t)] = \hat{\beta}_0 + \hat{\beta}_1 \log[m_w(x,t)]$, with different weights w according to the female proportion. * significance $\alpha = 0.05$.

According to Gatzert and Wesker (2011), $\hat{\beta}_0 = -0.3197$ and $\hat{\beta}_1 = 1.0747$ for the Swiss population. Comparing these estimates with the estimated parameters for the Mexican population using the Brass-type relational model, we see that $\hat{\beta}_0$ is similar to that estimated by the Brass-type relational model when there is a population with more women than men. Meanwhile, in contrast with the previous parameter, $\hat{\beta}_1$ resembles that obtained for Switzerland when considering a low proportion of women.

Figure 8 shows the result of setting the year 2000 for the Mexican mortality rates, and when considering the parameter estimates in accordance with the Swiss Brass-type relational model, we obtained mortality rate estimates that could be used by the insurers.

5 Conclusions

We have described the mortality behavior of the Mexican population and, as in other countries, the pattern of mortality has been shown to be dependent on the factors of age and gender. Gender is a relevant factor in the longevity and mortality risk of the Mexican population, with the main gap between male and female

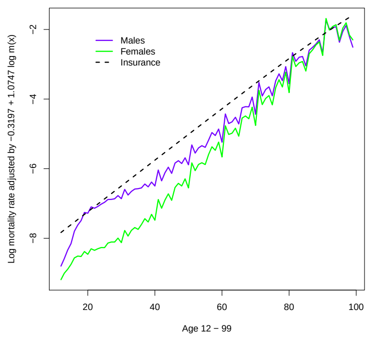


Figure 8: : Solid line represents the log-adjusted death rates obtained with the parameters for the Switzerland population. Dotted line represents the rates of CNSF Mexico year 2000.

mortality rates occurring in the age interval of 20 to 50. According to the Mexican Association of Insurance Institutions (AMIS), this age interval corresponds to adult ages and is precisely where the highest proportion of policyholders is to be found. Thus, using unisex life tables for insurance purposes can lead to a marked bias in the data, particularly if the male-female ratio in any given portfolio is not taken into account.

No historical mortality rates are available for the life insurance market in Mexico, but a unisex life table has been used by insurers. Here, therefore, we have proposed approximating the mortality of the insured population by comparing experiences in other countries. Gatzert and Wesker (2011) have used a Brass-type relational model to fit the insured population mortality rates to the general population mortality rates in Switzerland. This has been done for men and women separately.

We have compared the parameters obtained in the Brass-type relational model for Switzerland and the parameter estimates for the Mexican case, when the mortality of the population mortality is compared to the unisex life table assuming different male-female ratios. We obtained linear regression coefficients that are similar to those for the Swiss population. However, note that as the Mexican life table for insured portfolios is unisex, it should be stressed that we cannot compare the two sets of results directly.

We recommend studying mortality in insurance portfolios separately by gender, since the proportion of men and women in the general population and in a given portfolio can differ substantially. Unisex tariffs should be based on reasonable assumptions regarding the proportion of women in the portfolio, and on the gap between the mortality of the general population and that of the insured population. Unisex life tables for insured portfolios, such as those used by Mexican insurers, implicitly assume a constant male-female equilibrium of risk exposures in the portfolio. Thus, we believe that even though regulations insist that insurance prices should be equal for men and women, mortality analyses for actuarial purposes should be conducted separately. Insurers need to bear in mind the gender composition of their portfolios when setting the price of their particular life insurance product. All in all, a unisex life table for insurers should serve as a benchmark against which their portfolio mortality can be compared.

Acknowledgments

All authors thank support received from The Spanish Ministry: ECO2010-21787-C03-01. Montserrat Guillén thanks ICREA Academia.

References

Betzuen, A. (2010). Un Análisis sobre las posibilidades de prediccion de la mortalidad futura aplicando el modelo Lee-Carter. *Anales del Instituto de Actuarios Españoles : Colegio Profesional*, (16):111–140.

Brass, W. (1971). *On the Scale of Mortality*. Taylor and Francis, London.

Brouhns, N., Denuit, M., and Vermunt, J. K. (2002). A Poisson Log-Bilinear Regression Approach to the Construction of Projected Lifetables. *Insurance: Mathematics and Economics*, 31(3):373–393.

Cairns, A., Blake, D., Dowd, K., Coughlan, G., Epstein, D., Ong, A., and Balevich, I. (2009). A Quantitative Comparison of Stochastic Mortality Models Using Data from England and Wales and the United States. *North American Actuarial Journal*, 13:1–35.

Camposortega, S. (1997). Cien Años de Mortalidad en México. *Demos*, 10:11–13.

Cox, D. R. and Oakes, D. (1984). Analysis of Survival Data. In *Monographs on Statistics and Applied Probability*, volume 21. Chapman and Hall.

Currie, I. (2006). Smoothing and Forecasting Mortality Rates with P Splines. Talk given at the Institute of Actuaries.

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

Delwarde, A. and Denuit, M. (2003). Importance de la Période d'observation et des âges Considérés dans la Projection de la Mortalité selon la Méthode de Lee-Carter. *Belgian Actuarial Bulleti*, 3:1–21.

Delwarde, A., Denuit, M., Guillen, M., and Vidiella, A. (2003). Application of the Poisson Log-Bilinear Projection Model to the G5 Mortality Experience. *Belgian Actuarial Bulletin*, 6:54–68.

Fledelius, P., Guillen, M., Nielsen, J., and Petersen, K. (2004). A Comparative Study of Parametric and Nonparametric Estimators of old age Mortality in Sweden. *Journal of Actuarial Practice*, 11:101–126.

Gatzert, N. and Wesker, H. (2011). Mortality Risk and its Effect on Shortfall and Risk Management in Life Insurance.

Guillén, M. (2012). Sexless and Beautiful Data: from Quantity to Quality. *The Annals of Actuarial Science*, 6(2):231–234.

Hastie, T. J. and Tibshirani, R. J. (1990). *Generalized Additive Models*. Chapman and Hall.

Lee, R. and Carter, L. (1992). Modeling and Forecasting U. S. Mortality. *Journal of the American Statistical Association*, 87:659–671.

Pitacco, E., Denuit, M., Haberman, S., and Olivieri, A. (2009). *Modelling Longevity Dynamics for Pensions and Annuity Business*. United States: Oxford University Press.

Renshaw, A. and Haberman, S. (2003). Lee-Carter Mortality Forecasting: a Parallel Generalized Linear Modelling Approach for England and Wales Mortality Projections. *Journal of the Royal Statistical Society Series C*, 52(1):119–137.

Renshaw, A. and Haberman, S. (2006). A Cohort-Based Extension to the Lee - Carter Model for Mortality Reduction Factors. *Insurance: Mathematics and Economics*, 38(3):556–570.