Case 2

Modeling and forecasting Chile survival probabilities

Solange Gonçalves
Bruno Komatsu
Elias Cavalcante Filho
Juliana Tessari

2016

Universidade de São Paulo

1 Introduction

The study of mortality rates and the life expectancy of a population are of interest to a number of fields, including health, demography, social security and the planning of many public policies. Firstly, as mortality rates reduce, life expectancy increases and, with it, a series of issues arise concerning social security systems, including disability benefits, healthcare and old-age pensions, are affected by mortality trends (Brouhns et al. (2002)). For example, will the working population still be able to earn what is needed to accommodate the growing number of retirees (in a pay as you go scheme)? Must people work more hours during their career and retire at older ages? Should pensions be cut, or premiums increased to keep benefits affordable? Should people be encourage to take personal initiative to ensure their retirement?

Related to social security issues, both the return of savings as the pricing of insurance and pensions supplied by private industry depend on the stock market returns, the interest rates and the future mortality rates (Currie et al. (2004)).

The trend for the human mortality has been predominantly downward for the populations of many developed countries over the last century (Haberman and Renshaw (2011)). But the same pattern may be observed in some developing countries. In the Latin American, the demographic transition started with a decline in mortality, but the mortality levels and trends are very dissimilar among countries and the situation in some countries is very similar to that of the developed countries. Chile is one of the main examples of this case. According to World Health Organization (WHO), Chile has presented the highest growth of the life expectancy of the Latin American. The Table 1 below shows the evolution of mortality rates for male and female in Chile. It is possible to observe that the trend is decreasing for both genders and to total of population.

As a country develops the life expectancy of its population increases and, as a consequence, the dependency rates rise as well. Figure 1 illustrates this point with ratios of old people to working age population, the dependency rates. Our case of interest is Chile, one of the most developed countries of Latin America. Figure 1 shows that while its dependency rate is lower as compared to a developed country as Germany, it is ahead of many other Latin American countries and growing at a faster pace. Depending on the future development of this trends, pay as you go retirement schemes may become unfordable.

This paper aims to model the mortality forecast and the distribution of individual survivor probabilities, using the benchmark of the Lee and Carter (1992) model (L-C Model). Moreover, we seek to perform simulations scenarios of possible developments of survivor probabilities and calculations of the life expectancy for some age groups. We use data from the Human Mortality Database (HMDB) for the period 1992-2005.

As expected we verified that younger ages have typically higher survival rates, compared to older ages, and women have higher rates than men. We alfo found that the survivor rates show a increasing trend for all ages and for both sexes. These results reflects in the differential pattern of the cumulative

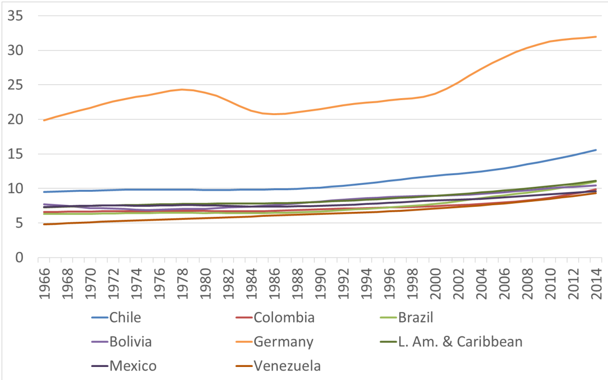

Figure 1: Evolution of Dependency Rates for Old People

Source: World Development Indicators/World Bank

probabilities of survival, which are almost steady up to age of 80, but show a very steep decline for those 65-year-olds in 2015.

This paper is organized into four sections, following this Introduction. Section 2 briefly presents a literature review about mortality forecast. Section 3 describes the L-C Model that we have used for forecasting mortality and the distribution of individual survivor probabilities. Section 3 presents the methodology to generate the forecasts. Section 4 describes the results and has some closing remarks.

2 Literature Review

There has been a growing trend of the literature of stochastic mortality models in the last decade, as a result of the increasing demand of measures for risk management and insurers and pension funds (Plat, 2009).

The literature on this topic can be organized into three approaches to the prediction problem: extrapolation, expectation and explanation approaches (Booth and Tickle, 2008). The oldest one is the extrapolation from previous statistical data to construct predictions about mortality rates in the future, using linear extrapolation methods. The expectation approach consists in calculations of the best guess, with an addition of a high and a low case deterministic scenarios, incorporating theoretical knowledge into the model, or based on experts opinions. The explanations approach aims to elaborate on identifying determining factor for the life expectancy (Peters et al., 2012).

The aim of this paper has a closer relation to the extrapolation approach, although recent methodologies combine more than one of those approaches in their calculations (Peters et al., 2012). The extrapolations models can generally be classified according to the number of factors of the Lexis Surface they take into account, up to three of them (Booth and Tickle, 2008). Zero factor models and one factor models have serious drawbacks regarding predictions, as the former may lead to inconsistent age profiles and the latter estimates mortality rates separately for each period, and as a consequence it is less clear how they could be used to make predictions (Peters et al., 2012).

The most prominent model is probably that of Lee and Carter (1992), which is referred to as the golden standard of forecasting (Li and Chan, 2007). It is a relatively simple model of age and a time trend (two factor model), and is able to explain a large proportion of the variance of the mortality rates Booth and Tickle (2008). The main idea of this method is to factor the surface of mortality rates into its principal components: the age profile, a time trend and the specific age deviations of mortality changes over the period, using a Singular Value Decomposition method. Lee and Carter (1992) estimate and project the time trend as a random walk with a drift and, as a consequence, it is able to explicitly model mortality over time as a stochastic process, with uncertainty bounds around the estimate Peters et al. (2012).

However, the Lee and Carter (1992) method has a number of drawbacks, and a range of studies was developed to deal with those disadvantages in the last decade. Peters et al. (2012) highlight that Lee and Carter (1992) assumed fixed age-specific levels of improvement of survival for the entire period of analysis, which was very unlikely in actual past data. Relaxing that assumption, Cairns et al. (2006) elaborate a projection model with two stochastic factor: one that affects equally mortality at all ages, and a second that is proportional to the age.

A series of another disadvantages of the Lee and Carter (1992) model is pointed out by Plat (2009): trivial correlation structure of the model, poor fit to historical data, lack of smoothness in the estimated age effect, and the strong dependency of the estimated mortality rates on that effect, which may lead to too low uncertainty of future death rates.

More recently, Pitacco et al. (2009) developed a family of regression-based models. Moreover, a range of multi-factor models accounting for an additional cohort effect (three factor models)(Renshaw and Haberman, 2003; Cairns et al., 2009; Plat, 2009), which allowed for more precise predictions (Peters et al., 2012).

A wide and recent literature intends to compare the one-factor Lee and Carter (1992) model with its extensions and with another stochastic mortality models. In this context, Booth and Tickle (2008) considered five variants or extensions of the Lee-Carter model to perform comparisons. Cairns et al. (2009)examined eight different stochastic mortality models, including the Lee-Carter model, using the qualitative criteria based on Bayesian Information Criterion (BIC). Dowd et al. (2010) chose six of the models discussed by Cairns et al. (2009) and applied a set of quantitative and qualitative criteria to assess each model?s ability to explain historical patterns of mortality. It is important to mention the article of Cairns et al. (2011), which focused on the ex-ante plausibility and robustness of forecasts for comparison of twelve extrapolation models. Another comparisons among variants of the Lee-Carter model are: Lee and Miller (2001), Booth et al. (2002), Booth et al. (2001) and Booth et al. (2005).

The main conclusions of this literature which compares the Lee-Carter model with its extensions and with another mortality models are: 1) in order to assess whether any mortality model is a good model or not, it is important to consider certain criteria to evaluate the model; 2) some models perform better under some criteria than others, but no single model is superior under all the criteria considered; 3) it is important to conduct tests and methodologies of comparison to identify suitable forecast models for specific datasets and it is recommended that decisions be based on more than one approach; 4) more e sophisticated methods are not a guarantee for better results; 5) a good fit to historical data does not guarantee sensible forecasts.

3 Methodology

The model of age-specific death rates proposed by Lee and Carter (1992) has been used as a benchmark by many researchers and international statistical bureaus. This model was developed specifically for U.S. mortality data, for the period 1933-1987.

The first step of the L-C model consists of modeling these mortality rates as follow:

$$m(x,t) = \exp[a_x + b_x k_t + \varepsilon_{x,t}]; \tag{1}$$

where $m(x,t)$ denotes the central death rate in age group x and time t for a specific country; a_x describes the general age shape and b_x is the tendency of mortality at age x to change when the general level of mortality k_t changes. The k_t parameter is a time-varying index, $\varepsilon_{x,y}$ is a set of random disturbances with mean 0 and variance σ_ε^2 and reflects particular age-specific historical influences not captured by the model.

As our database have information for several years, parameters a_x, b_x and k_t were estimated simultaneously, for more efficiency (Lee and Carter, 1992).

The second step of this method is to refit k_t on the number of deaths to assure a better convergence between estimated and observed deaths. The values of k_t are adjusted so as to match the actual number of deaths for each period t, $D(t) = \sum [N(x,t) \exp(a_x + b_x k(t))]$, where $N(x,t)$ is the population age distribution.

The forecast consists of an extrapolation of the estimated parameters and a time series formed by the values of k_t. This can be estimated using standard statistical methods following an ARIMA model, as described below:

$$k_t = k_{t-1} + c + e_t \tag{2}$$

where c is the drift term and e_t are the deviations from this path. The projected k can be used in Equation (1) together with the estimated a_x and b_x to calculate forecasts of the age-specific death rates.

Besides the modeling of mortality rates, another distribution function which may be informative to public policies, specially social security and pensions are the survival probabilities. The survival function gives the probability of being alive just before duration t, or the probability that the death has not occurred by duration t.

The one-year survival probability at age x during the year t is expressed by:

$$p_{x,t} = \exp\left(-m_{x,t}\right) \tag{3}$$

We use the freely available data from the Human Mortality Database (HMDB) for the period 1992-2005 to estimate the mortality forecast and the distribution of individual survivor probabilities. The HMDB provides open, international access to original data and calculations of death rates and life tables for national populations (38 countries or areas), as well as the input data used in constructing those tables (see: www.mortality.org).

The method proposed by Lee and Carter presents a number of appealing features. Although its use involves a number of steps, the procedure is simple and largely nonparametric. Besides its simplicity, the model incorporates particular features of a given populations mortality pattern. In general, the applications of the method show that the historical trend in k has always been highly linear with time, and the random walk with drift has given a good fit (Lee and Miller (2001)).

4 Results

Figure 2 shows the estimated survival probabilities, estimated as in equation 3 for males and females. For each period, younger ages (represented by darker colors) have typically higher survival rates, compared to older ages, and women have higher rates than men. In spite of that, we can observe a nonlinear effect for age zero, which shows a lower survival rate compared to ages 5 to 45. This fact stems from the natural vulnerability of the newborn. The survivor rates start to drop from age 55 onward, reaching, in 1992, 77% for 95-year-old males and 75% for 95-year-old females.

The survivor rates show an increasing trend for all ages and for both sexes. The rates for age 95 are very unstable in the estimation period and their predicted trajectories are not reliable. For the remaining ages, they converge to values above 95% as of 2045.

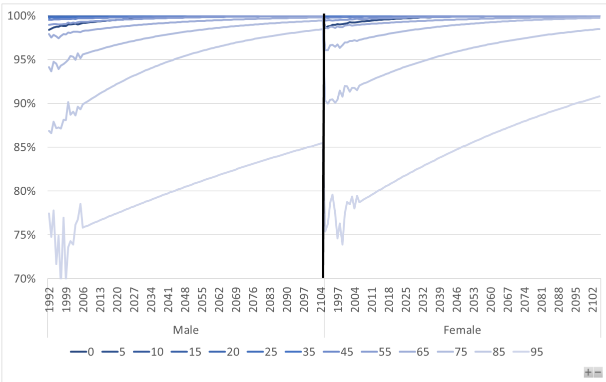

Figure 2: Individual Survivor Probabilities of Males and Females

The forecast of life expectancies for newborns and for age 65 are shown in Figure 3 and figure 4. Life expectancy is higher for women, however increases more sharply for men. In 2006 men had a life expectation of 87 years, reaching 94 years in 2106. As for women, in 2006, the life expectation was 91 years in 2006, and 96 years in 2106. The increases are aproximately of 0.84 moths per year for men, and 0.6 months per year for women, estimates below historical increases.

For men and women at age 65 the magnitude of increases are compara-

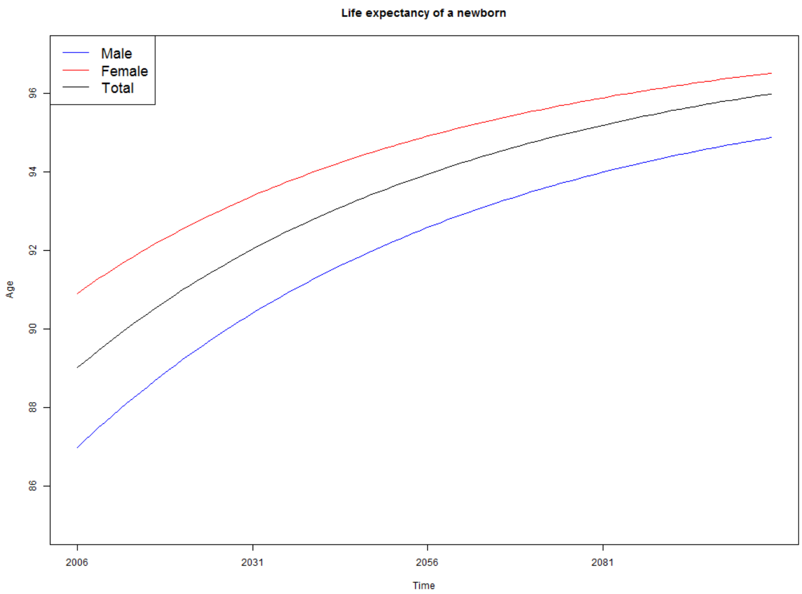

Figure 3: Life Expectancy of a Newborn Male and Female

tively larger, ranging from 18 to 28 years for males in one hundred years (1.2 months per year), and from 22 to 30 years in the same spam of time (0.96 months per year).

Figures 5 to 7 show the cumulative probability of survivor for newborns and 65-year-olds at dates 2015, 2025 and 2035. In 2015 newborns show almost steady curves for the cumulative probabilities, in contrast with the steep decline after the age of 80. This means that the mortality rates increases strikly after this age. In contrast, the cumulative probabilities for 65-year-olds show a sharp decline from the start, reflecting the fact that life expectation for those born in 1950 (65 years before 2015) increased greatly. This pattern does not change in the other periods.

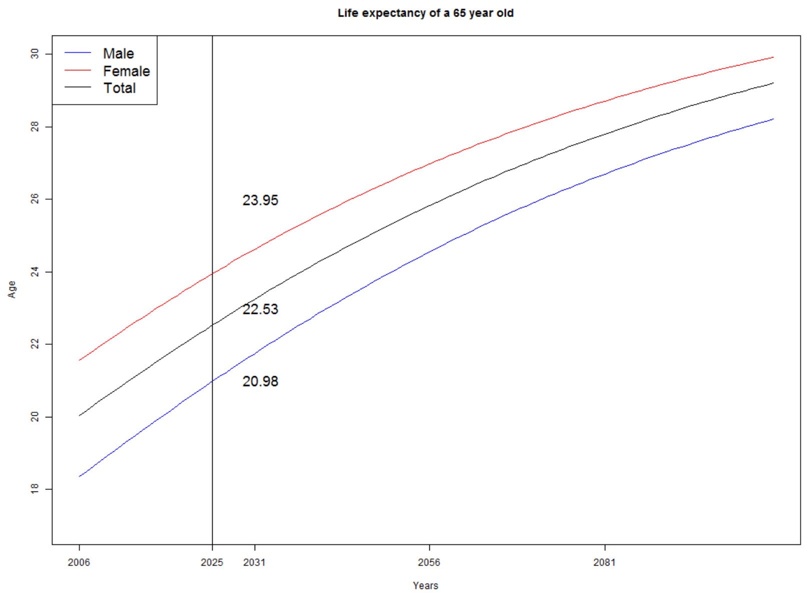

Figure 4: Life Expectancy of a 65-Year-Old Male and Female

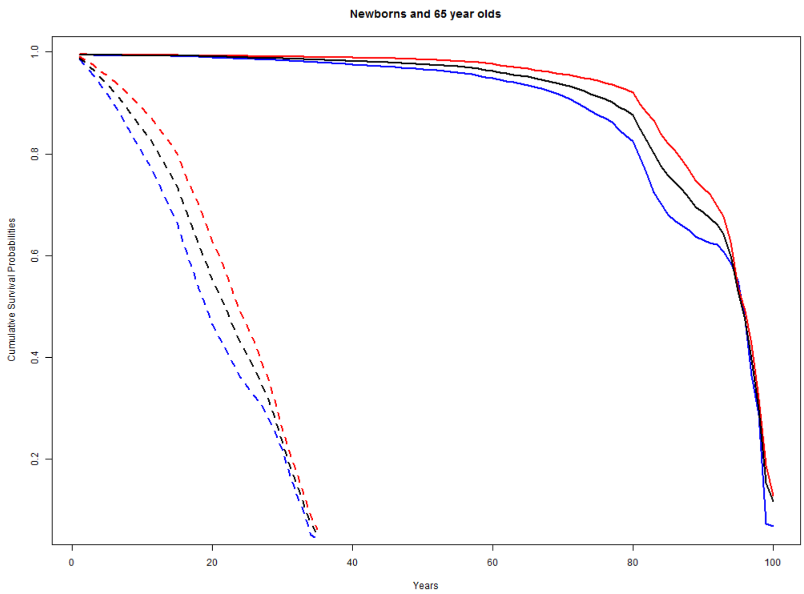

Figure 5: Cumulative Probability of Survivor 2015

5 Conclusion Remarks

We found that in Chile survivors probabilities increase for both sexes and more sharply for the older in the long run. Life expectancy for newborns also

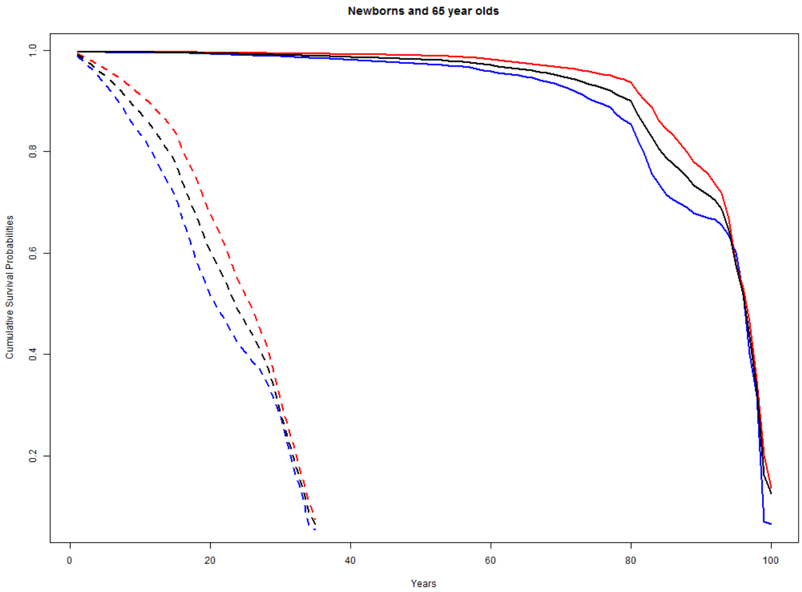

Figure 6: Cumulative Probability of Survivor 2025

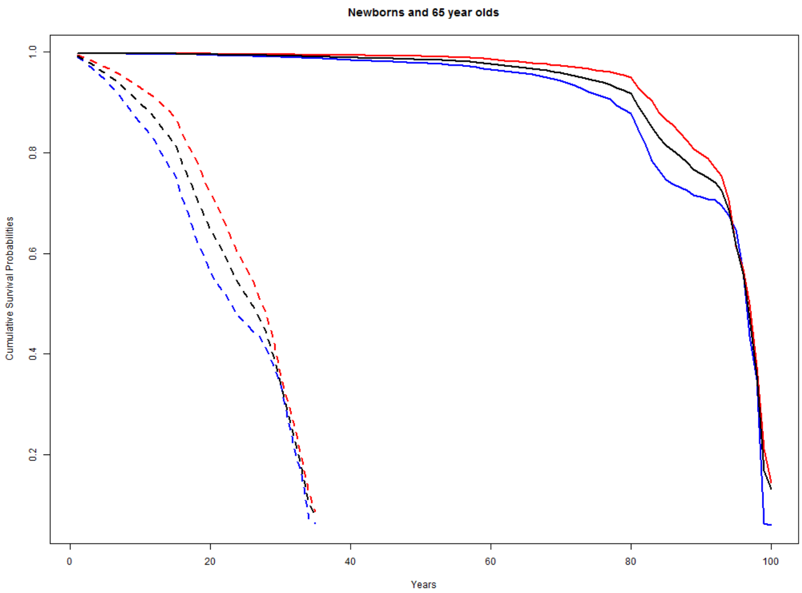

Figure 7: Cumulative Probability of Survivor 2035

rise, however, at a slower pace, com paced to 65-year-old expectancies. These results reflects in the differential pattern of the cumulative probabilities of survival, which are almost steady up to age of 80, but show a very steep decline for those 65-year-olds in 2015.

Our forecasts are not very different from estimates based on observed data (Pulgar and Acosta, 2004). This goes along with the argument of Lee and Miller (2001), who points out that the L-C model shows a good fit in a comparison among eight countries.

References

Technical report.

Booth, H., J. Maindonald, and L. Smith (2002). Applying lee-carter under conditions of variable mortality decline. *Population studies 56*(3), 325–336.

Booth, H., J. Maindonald, L. Smith, et al. (2001). Age-time interactions in mortality projection: Applying lee-carter to australia.

Booth, H. and L. Tickle (2008). Mortality modelling and forecasting: A review of methods. *Annals of actuarial science 3*(1-2), 3–43.

Booth, H., L. Tickle, L. Smith, et al. (2005). Evaluation of the variants of the lee-carter method of forecasting mortality: a multi-country comparison. *New Zealand Population Review 31*(1), 13–34.

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

Cairns, A. J., D. Blake, and K. Dowd (2006). A two-factor model for stochastic mortality with parameter uncertainty: Theory and calibration. *Journal of Risk and Insurance 73*(4), 687–718.

Cairns, A. J., D. Blake, K. Dowd, G. D. Coughlan, D. Epstein, and M. Khalaf-Allah (2011). Mortality density forecasts: An analysis of six stochastic mortality models. *Insurance: Mathematics and Economics 48*(3), 355–367.

Cairns, A. J., D. Blake, K. Dowd, G. D. Coughlan, D. Epstein, A. Ong, and I. Balevich (2009). A quantitative comparison of stochastic mortality models using data from england and wales and the united states. *North American Actuarial Journal 13*(1), 1–35.

Currie, I. D., M. Durban, and P. H. Eilers (2004). Smoothing and forecasting mortality rates. *Statistical modelling 4*(4), 279–298.

Dowd, K., A. J. Cairns, D. Blake, G. D. Coughlan, D. Epstein, and M. Khalaf-Allah (2010). Evaluating the goodness of fit of stochastic mortality models. *Insurance: Mathematics and Economics 47*(3), 255–265.

Haberman, S. and A. Renshaw (2011). A comparative study of parametric mortality projection models. *Insurance: Mathematics and Economics 48*(1), 35–55.

Lee, R. and T. Miller (2001). Evaluating the performance of the lee-carter method for forecasting mortality. *Demography 38*(4), 537–549.

Lee, R. D. and L. R. Carter (1992). Modeling and forecasting us mortality. *Journal of the American Statistical Association 87*(419), 659–671.

Li, S.-H. and W.-S. Chan (2007). The lee-carter model for forecasting mortality, revisited. *North American Actuarial Journal 11*(1), 68–89.

Peters, F., W. Nusselder, and J. Mackenbach (2012). The longevity risk of the dutch actuarial associations projection model. Technical report.

Pitacco, E., M. Denuit, S. Haberman, and A. Olivieri (2009). *Modelling longevity dynamics for pensions and annuity business*. OUP Oxford.

Plat, R. (2009). On stochastic mortality modeling. *Insurance: Mathematics and Economics 45*(3), 393–404.

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