UNIVERSIDAD CARLOS III DE MADRID working papers

UC3M Working Papers
Statistics and Econometrics
15-10
ISSN 2387-0303
May 2015

Departamento de Estadística
Universidad Carlos III de Madrid
Calle Madrid, 126
28903 Getafe (Spain)
Fax (34) 91 624-98-48

Hierarchical Lee-Carter model estimation through data cloning applied to demographically linked countries

Benchimol, A.[a], Albarrán, I.[a], Marín, J.M[a], Alonso González, P.J.[b]

Abstract

Some groups of countries are connected not only economically, but also social and even demographically. This last fact can be exploited when trying to forecast the death rates of their populations. In this paper we propose a hierarchical specification of the Lee-Carter model and we assume that there is a common latent mortality factor for all of them. We introduce an estimation procedure for this kind of structures by means of a data cloning methodology. To our knowledge, this is the first time that this methodology is used in the actuarial field. It allows approximating the maximum likelihood estimates, which are not affected by the prior distributions assumed for the calculus. Finally, we apply the methodology to some France, Italy, Portugal and Spain data. The forecasts obtained using this methodology can be considered as very satisfactory.

Keywords: Bayesian inference, data cloning, hierarchical model, Lee-Carter model, longevity risk, projected life tables.

[a] Department of Statistics, Universidad Carlos III de Madrid.
[b] Department of Economics, Universidad de Alcalá.

Hierarchical Lee-Carter model estimation through data cloning applied to demographically linked countries

Benchimol, A.[1], Albarrán, I.[2], Marín, J.M.[1], and Alonso-González, P.J.[3]

(1) and (2) Statistics Department. Universidad Carlos III de Madrid.
(3) Economics Department. Universidad de Alcalá. *

Abstract

Some groups of countries are connected not only economically, but also social and even demographically. This last fact can be exploited when trying to forecast the death rates of their populations. In this paper we propose a hierarchical specification of the Lee-Carter model and we assume that there is a common latent mortality factor for all of them. We introduce an estimation procedure for this kind of structures by means of a data cloning methodology. To our knowledge, this is the first time that this methodology is used in the actuarial field. It allows approximating the maximum likelihood estimates, which are not affected by the prior distributions assumed for the calculus. Finally, we apply the methodology to some France, Italy, Portugal and Spain data. The forecasts obtained using this methodology can be considered as very satisfactory.

Keywords: Bayesian inference, data cloning, hierarchical model, Lee-Carter model, longevity risk, projected life tables.

AMS subject classification: 62P05, 62F15.

1 Introduction

Insurance companies are usually facing risks of a quite different nature. Some of them have financial roots (such as investment returns or the impact of inflation), other are related to biometric items (such as mortality or longevity). In other cases, risks are linked to insureds and their personal situations (such as the likely existence of disabilities), either decisions taken by them (such as surrenders, unpaid premiums or the possibility of fraud) or due to unexpected evolution in some variables such as overheads or any other kind of expenses.

When considering biometric risks linked to survival, it is possible to distinguish amongst three different cases (Pitacco et al. (2009)):

(a) An individual may live a number of years around the average lifetime of his/her population. That is, if we consider the number of annual deaths in a population at a certain age, it can be seen that mortality fluctuates around a value with no systematic deviations across the time. This kind of mortality is referred to individuals and it is related to random fluctuations. This phenomenon is habitual in the life insurance business and its impact can be reduced enlarging the policies portfolio with similar contracts. *Ceteris paribus*, when the portfolio is large enough, it becomes a negligible risk but if this were not the situation, this risk could be reduced or hedged using some specific instruments, such as reinsurance or retrocession contracts.

(b) There can be differences between recorded and expected lifetimes in a population across the years. That is, the number of observed deaths may systematically be greater or lower than the expected. This may be the result of a wrong mortality model specification, or a biased estimation of the

*Authors' address: (1) Statistics Department, Universidad Carlos III de Madrid, C/ Madrid 126, 28903 Getafe, Spain (2) Statistics Department, Universidad Carlos III de Madrid, Avda. Gregorio Peces-Barba 22, 28270 Colmenarejo, Spain. (3) Economics Department, Universidad de Alcalá, Pl. Victoria 2,28802, Alcalá de Henares, Spain. E–mails: A. Benchimol, abenchim@est-econ.uc3m.es, I. Albarrán, irene.albarran@est-econ.uc3m.es, J.M. Marín jmmarin@est-econ.uc3m.es, P.J. Alonso-González, pablo.alonsog@uah.es

relevant parameters of the model. An example of this situation is the longevity risk, situation in which the insureds live more years than the expected. Unlike the former case of mortality, this one is concerned with the whole population. So, the solution of increasing the portfolio size has no effect in hedging against this risk. Some authors (see Cox et al. (2010), Blake et al. (2006) and Lin and Cox (2005)) suggest the use of specific financial instruments, such as mortality derivatives, to protect insurance companies against the economic effects of these contingencies.

(c) It is possible that death rates may suffer a sudden increase due to the presence of some critical life conditions, such as epidemics, natural disasters, severe weather conditions or wars. This is called catastrophic risk and it has a strong impact in the short term. As longevity risk, it is linked to the whole population, but the scope in time is quite different because longevity risk is referred to the medium and long term. Regardless the nature of the disaster, the companies require an appropriate protection against the effects of them. Increasing the size of the insured portfolio is not an effective measure to hedge against these contingencies. The right decision is diversification. Due to its non-individual scope, it can be obtained through risk transfer agreements. This solution is another difference between this risk and that related with longevity because this last one affects the whole portfolio.

The first kind of risk is linked to random fluctuations around a certain value whereas the last one is concerned with the effects of unexpected events. Given a specific environmental conditions, both of them can not be predicted. However, the second situation is referred to errors in the life tables estimation process that can be improved. These tables are frequently used by actuaries in their professional activity in life insurance business. As they can be used in a great number of tasks (such as calculating premiums, annuities, pension benefits or reserves), it is paramount to be sure that they are rightly calibrated and that the estimated probabilities reflected in them do not include any possible bias. If it is the case, the consequences will be passed to the financial statements resulting in lesser profits or even bankruptcy. A good life tables generator should produce sets of probabilities with the ability of incorporating the decreasing trend in death probabilities at every age. That is, it would be desirable that increasing life expectancies could be deduced from them. Although the enlargement of the life span has been one of the most relevant achievement made by the mankind during the last century, it is also one of the most disturbing financial challenges for the next decades. The insurance companies have been forced to allocate more capital to sustain their existing annuity business to prevent from adverse effects on reserves and profitability, and therefore to protect the solvency of companies. These effects are related to the fact that the longer the life is, the greater the number of years in retirement. Of course, everybody wants to live this extra time as good as possible. However, an increasing amount of funds is needed to get this goal. This is the root of the problem: neither public nor private pensions schemes are nowadays able to attend these financial requirements. Public Social Security systems are facing with the consequences of the population ageing process, whereas the private systems must take into account that the amounts reserved for those future payments were calculated in the past using previous life tables. The life expectancies derived of those tables are systematically lesser than those really recorded. This fact supposes not only a negative effect in the liquidity but also into the solvency of the undertakings devoted to this business. This situation is called longevity risk. As some authors suggest, this risk arises when a population enjoys of survival time greater than expected (Cairns et al. (2006)).

Reactions to the problem of longevity risk are mainly two. On one hand, actuaries have been trying to create better models to reflect the improvements in mortality, paying more attention to the levels of uncertainty of predictions (see Lee and Carter (1992), Cairns et al. (2006) and Pedroza (2006)). On the other hand, actuaries are trying to do the same as Capital Markets do to share risks, through the development of financial instruments linked to mortality (see Cox et al. (2010). Blake et al. (2006) and Lin and Cox (2005)). This paper is devoted to introduce a model able to capture these improvements. The final motivation obeys to actuarial purposes: to protect insurance companies in respect of mortality improvements. Actuaries in insurance companies must rely on tables that include future trends in mortality. This is known as *projected and dynamic life tables*.

As it has been pointed out before, mortality has been reducing across years during the last decades, sometimes faster, sometimes slower. So, this process can be treated from a statistical point of view and the evolution of this variable can be reflected using a stochastic model. We can distinguish between two different kinds of models: classical vs. Bayesian. The first group uses techniques based on more or less complex regressions. They use sample information without including any prior information. Within this group, Lee-Carter model (Lee and Carter (1992)) is one of the most popular among users. It considers a relation between death rates independent of years and a mortality index that is variable across the time

and independent of ages. An extension of this model is due to Renshaw and Haberman (2006). They included an extra variable that tries to reflect the cohort effect in mortality. More complex models are those based on the use of smoothing techniques. The first one was introduced by Currie et al. (2004). They suggested the use of B-splines and P-splines for fitting mortality surfaces. Afterwards, the APC model (stands for Age-Period-Cohort) was introduced by Currie et al. (2006), that can be considered as a particular case of the Renshaw-Haberman model. The difference between both models is that the APC was estimated using P-splines in order to assure a smooth result. Finally, models including two or more factors are a subset within the classical group. The seminal work was the CBD model (stands for Cairns, Blake and Dowd) due to Cairns et al. (2006). Its main assumption is that the effects linked to age and period are different between them and they both affect the future death rate. Further development of this basic model can be found in Cairns et al. (2009) and in Dowd et al. (2010). They include cohort effect and quadratic terms.

Relevant information based on historical information or in skilled opinions are used in Bayesian models to improve the estimation. From an actuarial point of view, a projected life table can be used as a starting point to achieve better estimations of the future mortality. Pedroza (2006) is a good example of this methodology. She proposed a Bayesian approach to the Lee-Carter model considering the uncertainty in the age parameters, as well as in the usually forecast mortality index. Expert-based probabilistic population projections have been produced by Lutz et al. (1998), Lutz et al. (2004) and Lutz et al. (2008). However, this method is not explicitly based on available data, and instead relies on a collection of experts and their ability to specify probabilistic bounds, that may or may not be accurate (Alho (2005)). Besides, Dellaportas et al. (2011) used a Bayesian version of the Heligman and Pollard (1980) to predict mortality figures.

As it has been pointed out before, econometric models are one of the most used statistical tools to analyse and project death rates. General Linear Models (GLM) can be applied assuming that age, cohort and/or period are explanatory variables. Estimates can be obtained after using maximum likelihood techniques which can be difficult to obtain in complex models. An alternative estimation methodology is the data cloning technique that attempts to estimate a GLM model but using a MCMC based algorithm. The two seminal papers about this methodology are Lele et al. (2007) and Lele et al. (2010), with applications in complex ecological models. After that, the technique has been applied in other fields like Finance (see Marín et al. (2015)). The algorithm is based on the technique called *simulated annealing* (Brooks and Morgan (1995)). This is a computationally intensive method, but the algorithm can be easily implemented with the R library `dclone` (Sólymos (2009)).

This paper deals with the estimation of a hierarchical model through data cloning. Our goal is to estimate a global Lee-Carter model (Lee and Carter (1992)) for a group of countries that are supposed to be close related. Therefore, the corresponding parameters must be linked among them. In such case, we can assign prior distributions that are related by means of a given set of hyperprior distributions with some specific parameters (hyperparameters). As far as our knowledge reaches, this is the first paper that tries to use this methodology in the actuarial field.

The remainder of this paper is organized as follows. Section 2 gives an overview of the Lee-Carter model, its fundamentals and some developments derived from it. Section 3 shows the foundations of the Bayesian approach to the former model. Section 4 describes and introduces the data cloning methodology. Section 5 describes the validation of the estimation procedure followed in the paper. Section 6 reflects the use of this algorithm to real data. Finally, Section 7 summarizes the main conclusions and final remarks.

2 Lee-Carter Model (LC)

Lee and Carter (1992) proposed a model to forecast mortality as a function of a time-varying index. This paper was the seminal work for further developments in the estimation of future mortality, such as Lee (2000), Booth et al. (2002), Li and Lee (2005) and Pedroza (2006).

The LC model deals with $m_x(t)$, the central death rate for age x in year t (it is calculated as the ratio of deaths to mid-year population size for a given interval of age and time). The LC model can be expressed as:

$$\log\left[m_x(t)\right] = \alpha_x + \beta_x\kappa_t + \varepsilon_{x.t}. \tag{1}$$

where α_x parameters describe the pattern of the average mortality at each age, while the parameters β_x describe deviations from this average pattern when κ_t varies. Both set of parameters, $\{\alpha_x\}$ and $\{\beta_x\}$ are independent from time. The variable, κ_t, is an index. It can be expressed as a time series process

and describes the change in the level of mortality over time. It is important to notice that this index is an unobservable variable, so it must be estimated. Finally, $\varepsilon_{x.t}$, the error term, denotes the random deviations unexplained by the model, and it is assumed to have mean 0 and constant variance σ_ε^2.

The expression $\beta_x\kappa_t$ indicates that it is assumed that there is no interaction between age and time variables: β_x parameters are fixed over time and κ_t parameters are fixed throughout ages. This means that β_x will be the same during all the years and κ_t will be the same for all ages. The LC model is quite useful in cases when the only available information is that referred to central death rates, $m_x(t)$. These figures are easily get from census data.

When a frequentist approach is used, the parameters can be estimated by least squares using the first element of the singular value decomposition of the following matrix:

$$Z = \begin{pmatrix} \log\left[\widehat{m}_{x_1}(t_1)\right] - \widehat{\alpha}_{x_1} & \cdots & \log\left[\widehat{m}_{x_1}(t_n)\right] - \widehat{\alpha}_{x_1} \\ \vdots & \ddots & \vdots \\ \log\left[\widehat{m}_{x_m}(t_1)\right] - \widehat{\alpha}_{x_m} & \cdots & \log\left[\widehat{m}_{x_m}(t_n)\right] - \widehat{\alpha}_{x_m} \end{pmatrix}. \tag{2}$$

Since the system of equations of the LC model is underdetermined, two additional constraints on the parameters β_x and κ_t have to be imposed. The indeterminacy of the LC model is due to the likelihood associated with the model, which has an infinite number of equivalent maxima, each of which produces the same estimations. Lee and Carter (1992) proposed as constraints $\sum_x \beta_x = 1$ and $\sum_t \kappa_t = 0$. Under this set of constraints, α_x is the average value of the logarithm of the central death rate for age x in the considered sample, whereas β_x is the percent change of the natural logarithm of the central death rate at a certain age due to changes in the mortality index in a certain year.

Once the values of α_x, β_x and κ_t have been estimated, it is needed a pattern for the future evolution of the mortality index. LC model supposes that κ_t can be modelled as a random walk with drift:

$$\kappa_t = \kappa_{t-1} + \theta + \omega_t.$$

where θ is the drift parameter which models a linear trend and ω_t is an error term. Thus, the forecast values of the unobservable process κ_t for r periods after the last observation at time n is calculated as follows:

$$\widehat{\kappa}_{n+r} = \widehat{\kappa}_n + r\widehat{\theta}.$$

Finally, the forecast death rates at age x for year $n+r$ can be obtained including this last estimation in the equation. The results for $\widehat{m}_x(n+r)$ can be expressed as:

$$\log(\widehat{m}_x(n+r)) = \widehat{\alpha}_x + \widehat{\beta}_x\widehat{\kappa}_{n+r}$$

The variance of the forecast mortality index $\widehat{\kappa}_{n+r}$ r periods ahead can be expressed as (see Pedroza 2006):

$$\widehat{\sigma}^2_{\kappa_{n+r}} = r^2\widehat{\sigma}^2_\theta + r\widehat{\sigma}^2_\omega$$

where $\widehat{\sigma}^2_\theta$ is the estimate of the variance of θ and $\widehat{\sigma}^2_\omega$ is the estimate of the variance of the error term ω. Thus, the LC $(1-\alpha)\times 100\%$ prediction intervals for the log-central death rates are calculated as follows:

$$\log(\widehat{m}_x(n+r)) \pm z_{\frac{\alpha}{2}}\widehat{\sigma}_{\kappa_{n+r}}$$

3 Bayesian approach to the LC model

This section shows a Bayesian statistical methodology for making inferences about the parameters of a LC model for several related populations by means of a hierarchical Bayesian model. Originally, the Bayesian approach of the LC model applied to one population was studied by Pedroza (2006) who analysed U.S. mortality data.

Hierarchical modelling is used when information about observational units is available on different levels. The hierarchical form of analysis and organization is quite useful regarding multi-parameter problems (see Gelman et al. (2014)). This paper is concerned to a case that considers several populations with some social-economic characteristics in common. For this reason, the parameters of the LC model can be treated as connected in some way, implying that the dependence among them can be reflected with a joint probability model.

Let be $j = 1. \ldots, J$ populations, in such a way that $m_x^{(j)}(t)$ is the central death rate for an age x at time t and population j, such that the corresponding LC model can be expressed as:

$$\begin{aligned} \log\left[m_x^{(j)}(t)\right] &= \alpha_x^{(j)} + \beta_x^{(j)}\kappa_t + \varepsilon_{x.t}^{(j)} \\ \kappa_t &= \kappa_{t-1} + \theta + \omega_t. \end{aligned}$$

where

$$\begin{aligned} \varepsilon_{x.t}^{(j)} &\sim N\left(0; \sigma_\varepsilon^{2(j)}\right). \\ \omega_t &\sim N\left(0; \sigma_\omega^2\right). \end{aligned}$$

As usual, we observe $m_x^{(j)}(t)$ whereas κ_t is unobserved. The goal is to estimate the set of parameters $\alpha_x^{(j)}$, $\beta_x^{(j)}$ and κ_t and use them to forecast the central death rates at each age for each generation in each population.

We assume proper prior distributions with a hierarchical structure. We consider normal distributions for $\alpha_x^{(j)}$ parameters and Dirichlet distributions for $\beta_x^{(j)}$ ones. in order to include the restriction about $\sum_x \beta_x = 1$, we assume inverse-gamma distributions for variances as it is a conjugate distribution in a normal model, allowing to deal with most of real cases about prior information regarding the parameters of the model.

In particular, we assume as prior distributions for the parameters (for $j = 1. \ldots .J$):

$$\begin{aligned} \alpha_x^{(j)} &\sim N(\mu_x^{(j)},\ \sigma_x^{2(j)}). \\ \beta_x^{(j)} &\sim \text{Dirichlet}(1, 1, \ldots, 1). \\ \sigma_\varepsilon^{2(j)} &\sim InvGamma(\gamma_1,\ \gamma_2). \\ \kappa_1 &\sim N(0,\ \sigma_\omega^2). \\ \theta &\sim N(0,\ 100). \\ \sigma_\omega^2 &\sim InvGamma(1,\ 1). \end{aligned} \tag{3}$$

And the hyperprior distributions of the hyperparameters are

$$\begin{aligned} \mu_x^{(j)} &\sim N(0,\ 10). \\ \sigma_x^{2(j)} &\sim InvGamma(1,\ 1). \\ \gamma_1 &\sim Gamma(1,\ 1). \\ \gamma_2 &\sim Gamma(1,\ 1). \end{aligned}$$

By other hand, the likelihood function for this model is:

$$L\left(m_x^{(j)}(t); \boldsymbol{\Theta}\right) = \prod_{j=1}^{J}\prod_{t=1}^{n}\prod_{x=1}^{\omega} \frac{1}{\sqrt{2\pi}\sigma_\varepsilon^{(j)}} \exp\left[-\frac{1}{2}\left(\frac{\log\left[m_x^{(j)}(t)\right] - \alpha_x^{(j)} - \beta_x^{(j)}\kappa_t}{\sigma_\varepsilon^{(j)}}\right)^2\right].$$

where $\boldsymbol{\Theta} = \left(\alpha_x^{(j)}, \beta_x^{(j)}, \kappa_1, \mu_x^{(j)}, \sigma_x^{2(j)}, \sigma_\varepsilon^{2(j)}, \theta, \sigma_\omega^2, \gamma_1, \gamma_2\right)$ for $j = 1, \ldots, J$ and $x = 1, \ldots, \omega$.

The joint posterior distribution for all the parameters is obtained by multiplying the likelihood function by the corresponding prior distributions (3). In general, the full set of conditional distributions is required to implement a MCMC algorithm. Then, the conditional posterior distributions of each parameter is easily obtained from the joint posterior distribution, considering only the proportional terms to each parameter.

4 The data cloning methodology

The data cloning method is a simulation technique to compute maximum likelihood estimates of parameters along with their asymptotic variances, by using a MCMC methodology (see Lele et al. (2007) and Lele et al. (2010)). It uses the simplicity of Monte Carlo algorithms to calculate maximum likelihood estimations in models whose complexity make necessary the use of high-dimensional integration to obtain them. The methodology is based on the basic idea of repeating an experiment several times conditioned to obtain always the same data.

Let us denote $\log[m_x(t)]$ as y_t for $(t = 1, \ldots, n)$ in such a way that $\mathbf{y} = (y_1, \ldots, y_n)$. In a MCMC procedure, once data $\mathbf{y}$ has been observed, the samples from the posterior distribution $\pi(\boldsymbol{\Theta}|\mathbf{y})$ are generated. This posterior distribution is proportional to the product of the likelihood function $L(\Theta|\mathbf{y})$ and a given proper prior distribution $\pi(\Theta)$. Then, in data cloning, samples are generated from the posterior distribution, $\pi^{(K)}(\Theta|\mathbf{y})$, that is proportional to the Kth power of the likelihood, $[L(\boldsymbol{\Theta}|\mathbf{y})]^K$, multiplied by a proper prior distribution, $\pi(\Theta)$.

The expression $[L(\boldsymbol{\Theta}|\mathbf{y})]^K$ is the likelihood for K copies of the original data and, for K large enough, $\pi^{(K)}(\Theta|\mathbf{y})$ converges to a multivariate normal distribution with mean equal to the maximum likelihood estimates of the parameters, and covariance matrix equal to $1/K$ times the inverse of the Fisher information matrix for the maximum likelihood estimates (see Lele et al. (2007)).

Once the samples have been obtained with a MCMC procedure, sample means are computed based on the posterior distributions of the parameters. They provide an approximation of the maximum likelihood estimates of the parameters.

As a summary, the data cloning algorithm follows the next steps:

Step 1: Create K-cloned data set $\mathbf{y}^{(K)} = (\mathbf{y}, \mathbf{y}, \ldots, \mathbf{y})$, where the observed data vector is repeated K times.

Step 2: Using an MCMC algorithm, generate random numbers from the posterior distribution that is based on a prior $\pi(\Theta)$ and the cloned data vector $\mathbf{y}^{(K)} = (\mathbf{y}, \mathbf{y}, \ldots, \mathbf{y})$, where the K copies of $\mathbf{y}$ are assumed to be independent of each other. In practice, any proper prior distribution can be used.

Step 3: Compute the sample mean and variances of each of the individual values of the vector of parameters Θ (for M iterations of the MCMC algorithm) generated from the marginal posterior distribution. The maximum likelihood estimates of Θ and the approximate variances of the maximum likelihood estimates are those referred to the posterior mean values and to K times the posterior variances. respectively.

The algorithm has been programmed using the package `dclone` (Sólymos (2009)) from the `R` project (R Core Team. R Foundation for Statistical Computing (2012)). The optimal number of clones has been established considering some statistics computed in the package `dclone` (Sólymos (2009)), such as the maximum eigenvalue of the posterior covariance matrix, the minimum squared error and the squared error (see Lele et al. (2010)).

5 Validating the estimation procedure

Before applying the procedure to real data, we check it by means of a validation study. The validation has been done by simulating an array of log-central death rates, assuming 4 populations, 41 consecutive ages and 50 consecutive calendar years each.

In order to simulate data we took as a starting point the estimated parameters for some French data taken from the *Human Mortality Database* (see `www.mortality.org`) for ages between 60 and 100, and for calendar years between 1960 and 2009, including both ends in both cases. The estimation was undertaken by the standard frequentist procedure based on the singular value decomposition raised by Lee and Carter (1992) using `R` .

For population 1, we generated each parameter $\alpha_x^{(1)}$ as the respective French $\hat{\alpha}_x$ parameter plus a random component following a N(μ=0;σ=0.001) distribution. For populations 2, 3 and 4, $\alpha_x^{(j)}$ parameters were generated taking 95%, 90%, and 105% of French $\hat{\alpha}_x$ parameters, plus a random component following N(μ=0;σ=0.002). N(μ=0;σ=0.003) and N(μ=0;σ=0.004) distributions, respectively. Regarding $\beta_x^{(j)}$ parameters for each population, they were simulated by four vectors of dimension 41, each of them following a Dirichlet distribution with all its parameters equal to 1, in order to fulfill the constraint $\sum_x \beta_x = 1$.

On the other side, regarding the mortality index κ_t, adjusting for France estimates, parameters were assumed to be as $\hat{\kappa}_0 = -11.2075$, $\hat{\theta} = -0.5728$ and $\hat{\sigma}_\omega^2 = 5$. Thus, κ_t was fixed as a vector of 50 values from $t = 1960$ up to $t = 2009$, where each κ_t was a random value generated from a normal distribution with mean $\kappa_{t-1} + \hat{\theta}$ and variance $\hat{\sigma}_\omega^2$.

Finally, the data set of log-central death rates $\log\left[m_x^{(j)}(t)\right]$ (four matrices of dimensions 50×41, one for each population) was simulated by means of four normal distributions with means $\alpha_x^{(j)} + \beta_x^{(j)}\kappa_t$ and variances $\sigma_y^{2(1)}$ =0.0010, $\sigma_y^{2(2)} = 0.0015$, $\sigma_y^{2(3)} = 0.0020$ and $\sigma_y^{2(4)}$=0.0025, respectively.

We used the package `dclone` (Sólymos (2009)) from the `R` project (R Core Team (2015)) in order to program the algorithm. We checked the optimal number of clones by means of some statistics computed in the package `dclone` (Sólymos (2009)), such as the maximum eigenvalue of the posterior variance, the minimum squared error and the squared error (see Lele et al. (2010)). There were not relevant improvements in their values when the number of clones was larger than 5, and therefore we worked with 5 clones to estimate the parameters. Besides, we used the same prior distributions as in Section 3.

Then, the parameters of this hierarchical model were estimated applying our algorithm with the data cloning technique, on the same simulated data set 100 times, generating thus 100 replicates, and the results were assessed.

Let us consider the Pearson's coefficient of variation (CV) and the the relative mean squared error ($RMSE$) respectively as $CV(\hat{\theta}) = \sigma_{\hat{\theta}}/|\bar{\hat{\theta}}|$ and $RMSE(\hat{\theta}) = \sqrt{E[(\hat{\theta}-\theta)^2]}/|\theta|$.

We calculated both relative dispersion measures for the 8200 (50 calendar years $\times$ 41 ages $\times$ 4 populations) estimators of the log-central death rates, $\log\left[\widehat{m}_x^{(j)}(t)\right]$. The mean CV for the whole sample was 0.048, and the mean RMSE was 0.047. In the case of the CV, 154 out of 8200 were larger than 0.2. while in the case of the RMSE, 171 out of 8200 were larger than this threshold (1.87% and 2.08% of the sample, respectively). These results suggest that the procedure seems to be unbiased and stable.

In Figures 1a and 1b the respective histograms of both measures for the whole sample of 8200 estimators are shown.

FIGURES 1a and 1b ROUND HERE

6 Application to real data

The data cloning methodology has been applied to a set of European countries mortality data. We have taken the central death rates from France, Italy, Portugal and Spain located in the *Human Mortality Database* (see `www.mortality.org`).

We have focused a time span between years 1960 and 2009 and ages between 60 and 100 years old. These countries present a similar social development and their populations enjoy parallel welfare states. More specifically, they show similar standard demographic indices such as:

- *Life expectancy at birth* (LEB): it is the average number of years that a newborn can expect to live, according to the mortality conditions at his/her birth time.
- *Life expectancy at age 65* (LE65): it is the average number of years that a person age 65 can expect to live, according to the mortality conditions at the time he/she attains that age.
- *Median age of the population* (MAP): it is the median of ages of the alive people in a country at a certain time. It shows the progressive ageing in the population as a result of the low birth and death rates, yielding in a high life expectancy. This index is also affected by the immigration flows.
- *Old-dependency ratio* (ODR): it is the proportion between the number of people over 65 years old and the number of people between 16 and 64 years old. Namely, it is the ratio between retired and active people (the last one is the sum of employed and unemployed workers). ODR index is used to measure the pressure of the elderly people over the productive population.

The corresponding values of these indices for years 2012-2013 for each country are shown in table 1.

TABLE 1 ROUND HERE

In order to validate the predictive performance of the model, the data set was split into two groups: the first one (training sample) includes data from 1960 to 1999, whereas the second one (validation sample) includes data from 2000 to 2009.

We complete the analysis of the hierarchical model by applying the data cloning technique. We have programmed the algorithm using package `dclone` (Sólymos (2009)) from the `R` project (R Core Team (2015)). We check what is the optimal number of clones, regarding some statistics computed in the package `dclone` (Sólymos (2009)), such as the maximum eigenvalue of the posterior variance, the minimum squared error and the squared error (see Lele et al. (2010)). There are not relevant improvements in their values when the number of clones is larger than 5, therefore we use this number of clones to analyse the data. We have used the same prior distributions as in Section 3.

Mean and standard deviation of parameters were estimated for the training sample using 5 clones after 50,000 iterations of the MCMC algorithm. The estimated mean for the $\{\alpha_x^{(j)}\}$ and $\{\beta_x^{(j)}\}$ parameters for each country are shown in table 2.

TABLE 2 ROUND HERE

On the other hand, error variances $\sigma_x^{2(j)}$ estimated for each country are shown in table 3.

TABLE 3 ROUND HERE

Using the estimations of $\alpha_x^{(j)}$. $\beta_x^{(j)}$ for each age x and each country, along with the forecasted values κ_t, the log-central death rates $\log\left[m_x^{(j)}(t)\right]$ were estimated for the period 2000-2009.

The 95% prediction intervals for the predicted values, based on the Wald approximation, and the actual values of the log-central death rates are shown in tables 4a, 4b, 4c and 4d for France, Italy, Portugal and Spain, respectively, for ages $x = 60$, 70, 80, 90 and 100, and for a prediction horizon $t = 2000, \ldots, 2009$. Notice that all intervals include the actual values of the parameters.

TABLES 4a, 4b, 4c and 4d ROUND HERE

In Figures 2a, 2b, 2c and 2d, we represent the mortality surfaces for France, Italy, Portugal and Spain, respectively. Here, we show first the observed central death rates $m_x^{(j)}(t)$ (height of the surface) for $t = 1960, \ldots, 2009$. And then to the right, the projected central death rates for $t = 2000$, ..., 2009 forecasted by our model with the data in the training sample. As it can be seen, the projected surfaces suggest a proper fit.

FIGURES 2a, 2b, 2c and 2d ROUND HERE

7 Concluding remarks

We have introduced a hierarchical Lee-Carter model to forecast the death rates of a set of demographically linked countries. Although it has been assumed that each country has its own specific characteristics, the model is based on the existence of a common and latent mortality structure. This idea is quite interesting to estimate the parameters of the model because it allows to take advantage of the whole set of information, that is, the forecasts of a certain country are calculated not only based on its death rates but also in those of the rest of the considered linked populations. Bayesian methodology is a very effective way to deal with hierarchical models. However, this scheme is limited by the fact that it is often necessary that the analyst determines the prior distributions for all parameters and hyperparameters of the model. Data cloning is an alternative to surpass the previous limitation and it allows to approximate the maximum likelihood estimates. In this scenario, the role of the prior distributions is not determinant.

Although this methodology has been applied to areas such as Ecology and recently in Finance, it is the first time that data cloning has been used in the Actuarial field, to our knowledge extent. The set of information includes the central death rates of France, Italy, Portugal and Spain. In order to check the validity of the forecasts, the sample has been divided into two sets. The first one is devoted to estimate the parameters, whereas the second one is used to contrast the accuracy of the results. The model is able to rightly predict the central death rates rates in all cases, using 95% approximated prediction intervals. All of these results can be directly used in the management of private and/or public pension systems, as nowadays one of the most relevant problems for the insurance industry is connected with wrong estimations of survival probabilities.

Future research will involve the implementation of this methodology with other kind of models used to forecast death rates, such as Cairns-Blake-Dowd (*CBD*) stochastic mortality models, or those based on P-splines.

References

Alho, J. (2005). *Statistical Demography and Forecasting.* Springer.

Blake, D., A. Cairns, K. Dowd, and R. MacMinn (2006). Longevity bonds: Financial engineering, valuation, and hedging. *Journal of Risk and Insurance 73(4)*, 647–672.

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

Brooks, S. and B. Morgan (1995). Optimization using simulated annealing. *Statististician 44*, 241–254.

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

Cairns, A., D. Blake, K. Dowd, G. 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*, 1–35.

Cox, S., Y. Lin, and H. Pedersen (2010). Mortality risk modeling: Applications to insurance securitization. *Insurance: Mathematics and Economics 46*, 242–253.

Currie, I., M. Durbán, and P. Eilers (2004). Smoothing and forecasting mortality rates. *Stochastic modeling 4*, 279–298.

Currie, I., M. Durbán, and P. Eilers (2006). Generalized linear array models with applications to multidimensional smoothing. *Journal of the Royal Statistical Society, Series B 68*, 259–280.

Dellaportas, P., A. Smith, and P. Stavropoulos (2011). Bayesian analysis of mortality data. *Journal of the Royal Statistical Society, Series A 164(2)*, 275–289.

Dowd, K., A. Cairns, D. Blake, G. Coughlan, D. Epstein, and M. Khalaf-Allah (2010). Backtesting stochastic mortality models: An ex-post evaluation of multi-period ahead density forecasts. *North American Actuarial Journal 14*, 281–298.

Gelman, A., J. B. Carlin, H. S. Stern, and D. B. Rubin (2014). *Bayesian data analysis*, Volume 2. Taylor & Francis.

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

Lee, R. (2000). The lee-carter method for forecasting mortality, with various extensions and applications. *North American Actuarial Journal 4(1)*, 80–93.

Lee, R. and L. Carter (1992). Modeling and forecasting U.S. mortality. *Journal of the American Statistical Society 87*, 659–675.

Lele, S., B. Dennis, and F. Lutsche (2007). Data cloning: easy maximum likelihood estimation for complex ecological models using Bayesian Markov Chain Monte Carlo methods. *Ecology letters 10*, 551–563.

Lele, S., K. Nadeem, and B. Schmuland (2010). Estimability and likelihood inference for generalized linear mixed models using data cloning. *Journal of the American Statistical Society 105*, 1617–1625.

Li, N. and R. Lee (2005). Coherent mortality forecasts for a group of populations and extension of the lee-carter method. *Demography 42*, 575–594.

Lin, Y. and S. Cox (2005). Securitization of mortality risks in life annuities. *Journal of Risk and Insurance 72(2)*, 227–252.

Lutz, W., W. Sanderson, and S. Scherbov (1998). Expert-based probabilistic population projections. *Population and Development Review 24*, 139–155.

Lutz, W., W. Sanderson, and S. Scherbov (2004). *The end of World population growth in the 21th century: new challenges for human capital formation and sustainable development*. Sterling.

Lutz, W., W. Sanderson, and S. Scherbov (2008). IIasa's 2007 probabilistic world population projections.

Marín, J. M., M. T. Rodríguez-Bernal, and E. Romero (2015). Data cloning estimation of GARCH and COGARCH models. *Journal of Statistical Computation and Simulation 85(9)*, 1818–1831.

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

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

R Core Team (2015). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing.

Renshaw, A. and S. Haberman (2006). A cohort-based extension to the Lee-Carter model for mortality reduction factors. *Insurance: Mathematics and Economics 38*, 556–570.

Sólymos, P. (2009). dclone: Data cloning and MCMC tools for maximum likelihood methods. *R package version*, 1–0.

	LEB		LE65		MAP		ODR
	Male	Female	Male	Female	Male	Female	General
France	78.7	85.4	19.1	23.4	38.2	41.2	27.5
Italy	79.8	84.8	18.5	22.1	43.0	45.6	32.7
Portugal	77.3	83.6	17.6	21.3	37.6	41.9	29.4
Spain	79.5	85.8	18.7	22.8	40.1	42.9	26.3

Source: Eurostat 2012, 2013.

Table 1: Demographic indices

	France		Italy		Portugal		Spain	
x	α_x	β_x	α_x	β_x	α_x	β_x	α_x	β_x
60	-4.2041	0.0308	-4.3192	0.0354	-4.2101	0.0292	-4.3799	0.0302
61	-4.1258	0.0311	-4.2155	0.0344	-4.1261	0.0262	-4.3003	0.0258
62	-4.0550	0.0326	-4.1221	0.0340	-4.0465	0.0304	-4.2187	0.0331
63	-3.9676	0.0319	-4.0354	0.0338	-3.9441	0.0296	-4.1232	0.0311
64	-3.8957	0.0325	-3.9293	0.0320	-3.8708	0.0302	-4.0293	0.0324
65	-3.7844	0.0294	-3.8367	0.0320	-3.7692	0.0292	-3.9388	0.0332
66	-3.7493	0.0343	-3.7390	0.0306	-3.6857	0.0293	-3.8473	0.0308
67	-3.6656	0.0342	-3.5864	0.0232	-3.5832	0.0294	-3.7519	0.0321
68	-3.5808	0.0343	-3.5521	0.0291	-3.4903	0.0311	-3.6611	0.0333
69	-3.5056	0.0347	-3.4624	0.0295	-3.3903	0.0292	-3.5676	0.0305
70	-3.4138	0.0347	-3.3750	0.0299	-3.2894	0.0332	-3.4699	0.0345
71	-3.3321	0.0349	-3.2760	0.0279	-3.1739	0.0267	-3.3762	0.0286
72	-3.2398	0.0338	-3.1840	0.0290	-3.0996	0.0324	-3.2755	0.0367
73	-3.1532	0.0344	-3.0938	0.0286	-2.9929	0.0310	-3.1636	0.0322
74	-3.0647	0.0346	-2.9978	0.0282	-2.8906	0.0316	-3.0673	0.0335
75	-2.9568	0.0331	-2.9048	0.0280	-2.7885	0.0317	-2.9760	0.0327
76	-2.8762	0.0342	-2.8027	0.0267	-2.6748	0.0301	-2.8400	0.0287
77	-2.7751	0.0333	-2.7063	0.0263	-2.5766	0.0303	-2.7669	0.0300
78	-2.6717	0.0322	-2.6167	0.0269	-2.4658	0.0307	-2.6723	0.0339
79	-2.5771	0.0324	-2.5120	0.0256	-2.3615	0.0287	-2.5625	0.0278
80	-2.4660	0.0304	-2.4158	0.0273	-2.3062	0.0269	-2.4695	0.0308
81	-2.3543	0.0285	-2.3168	0.0260	-2.2115	0.0246	-2.3669	0.0241
82	-2.2371	0.0263	-2.2226	0.0257	-2.1235	0.0271	-2.2670	0.0298
83	-2.1387	0.0253	-2.1167	0.0247	-2.0080	0.0243	-2.1734	0.0260
84	-2.0372	0.0242	-2.0232	0.0240	-1.9233	0.0250	-2.0716	0.0291
85	-1.8042	0.0001	-1.9248	0.0231	-1.8301	0.0248	-1.9815	0.0251
86	-1.8385	0.0230	-1.8302	0.0224	-1.7338	0.0227	-1.8778	0.0244
87	-1.7166	0.0183	-1.7336	0.0218	-1.6204	0.0197	-1.7799	0.0222
88	-1.6406	0.0194	-1.6409	0.0210	-1.5465	0.0215	-1.6890	0.0233
89	-1.5488	0.0186	-1.5542	0.0210	-1.4672	0.0219	-1.5916	0.0191
90	-1.4528	0.0174	-1.4605	0.0204	-1.3737	0.0189	-1.5028	0.0193
91	-1.3577	0.0171	-1.3760	0.0190	-1.3090	0.0182	-1.3907	0.0033
92	-1.2756	0.0150	-1.2849	0.0187	-1.1934	0.0152	-1.3331	0.0138
93	-1.1823	0.0125	-1.2042	0.0168	-1.1482	0.0173	-1.2694	0.0116
94	-1.0953	0.0106	-1.1280	0.0171	-1.0862	0.0175	-1.1708	0.0118
95	-1.0279	0.0105	-1.0624	0.0160	-0.9946	0.0148	-1.1056	0.0117
96	-0.9569	0.0099	-0.9878	0.0149	-0.9286	0.0143	-1.0321	0.0106
97	-0.8869	0.0091	-0.8986	0.0117	-0.8617	0.0133	-0.9621	0.0096
98	-0.8193	0.0081	-0.8515	0.0131	-0.7910	0.0115	-0.8964	0.0088
99	-0.7537	0.0070	-0.7912	0.0127	-0.7387	0.0116	-0.8278	0.0073
100	-0.6875	0.0055	-0.7292	0.0116	-0.6632	0.0085	-0.7704	0.0070

Table 2: Point estimates of $\alpha_x^{(j)}$ and $\beta_x^{(j)}$ parameters

Parameter	Point Estimate
$\sigma_{\varepsilon}^{2(France)}$	0.0022
$\sigma_{\varepsilon}^{2(Italy)}$	0.0023
$\sigma_{\varepsilon}^{2(Portugal)}$	0.0025
$\sigma_{\varepsilon}^{2(Spain)}$	0.0014

Table 3: Error variances estimates

(a) France

	x=60			x=70			x=80		
t	95% PI		Actual value	95% PI		Actual value	95% PI		Actual value
	Up B.	Low B.		Up B.	Low B.		Up B.	Low B.	
2000	-4.7128	-3.9458	-4.4295	-3.9740	-3.1368	-3.6005	-2.9710	-2.2088	-2.5806
2001	-4.8500	-3.8354	-4.5117	-4.1305	-3.0090	-3.6439	-3.1051	-2.1003	-2.6520
2002	-4.9636	-3.7463	-4.4671	-4.2609	-2.9074	-3.6412	-3.2169	-2.0137	-2.6533
2003	-5.0685	-3.6671	-4.4559	-4.3786	-2.8188	-3.6893	-3.3199	-1.9357	-2.6526
2004	-5.1645	-3.5965	-4.5236	-4.4888	-2.7362	-3.7474	-3.4149	-1.8647	-2.7465
2005	-5.2567	-3.5288	-4.5153	-4.5919	-2.6616	-3.7872	-3.5042	-1.8004	-2.7589
2006	-5.3435	-3.4674	-4.5394	-4.6924	-2.5902	-3.8231	-3.5915	-1.7383	-2.7926
2007	-5.4298	-3.4073	-4.5413	-4.7886	-2.5226	-3.8467	-3.6755	-1.6796	-2.8166
2008	-5.5138	-3.3487	-4.5631	-4.8817	-2.4572	-3.8839	-3.7573	-1.6232	-2.8341
2009	-5.5926	-3.2943	-4.5516	-4.9730	-2.3955	-3.9001	-3.8348	-1.5706	-2.8810

	x=90			x=100		
t	95% PI		Actual value	95% PI		Actual value
	Up B.	Low B.		Up B.	Low B.	
2000	-1.8009	-1.2453	-1.5477	-0.9304	-0.4895	-0.7227
2001	-1.8655	-1.1954	-1.5622	-0.9413	-0.4828	-0.7344
2002	-1.9237	-1.1512	-1.5732	-0.9512	-0.4780	-0.7267
2003	-1.9790	-1.1110	-1.5299	-0.9617	-0.4719	-0.6907
2004	-2.0279	-1.0758	-1.6495	-0.9730	-0.4645	-0.7840
2005	-2.0791	-1.0391	-1.5507	-0.9839	-0.4583	-0.7511
2006	-2.1249	-1.0080	-1.7740	-0.9956	-0.4512	-0.7869
2007	-2.1721	-0.9752	-1.6663	-1.0048	-0.4465	-0.7803
2008	-2.2179	-0.9427	-1.6750	-1.0183	-0.4371	-0.7791
2009	-2.2609	-0.9147	-1.7800	-1.0291	-0.4317	-0.7737

(b) Italy

	x=60			x=70			x=80		
t	95% PI		Actual value	95% PI		Actual value	95% PI		Actual value
	Up B.	Low B.		Up B.	Low B.		Up B.	Low B.	
2000	-4.8909	-4.0362	-4.6628	-3.8806	-3.1145	-3.5972	-2.8838	-2.1691	-2.4969
2001	-5.0499	-3.9076	-4.6356	-4.0097	-3.0106	-3.6276	-3.0013	-2.0753	-2.6571
2002	-5.1816	-3.8045	-4.7128	-4.1192	-2.9252	-3.6861	-3.0990	-1.9990	-2.6116
2003	-5.3034	-3.7119	-4.6450	-4.2206	-2.8484	-3.7132	-3.1909	-1.9292	-2.5919
2004	-5.4134	-3.6305	-4.7028	-4.3126	-2.7804	-3.7491	-3.2751	-1.8678	-2.6863
2005	-5.5203	-3.5526	-4.7689	-4.4024	-2.7153	-3.7828	-3.3551	-1.8101	-2.6757
2006	-5.6238	-3.4781	-4.7879	-4.4877	-2.6554	-3.8509	-3.4315	-1.7571	-2.7243
2007	-5.7188	-3.4124	-4.7855	-4.5708	-2.5979	-3.8849	-3.5070	-1.7033	-2.7186
2008	-5.8175	-3.3427	-4.8258	-4.6518	-2.5407	-3.8709	-3.5804	-1.6518	-2.7292
2009	-5.9069	-3.2828	-4.8757	-4.7274	-2.4889	-3.9394	-3.6488	-1.6054	-2.7952

	x=90			x=100		
t	95% PI		Actual value	95% PI		Actual value
	Up B.	Low B.		Up B.	Low B.	
2000	-1.8463	-1.2411	-1.5906	-1.0218	-0.5293	-0.7425
2001	-1.9261	-1.1781	-1.5580	-1.0580	-0.5028	-0.7473
2002	-1.9971	-1.1236	-1.5955	-1.0905	-0.4797	-0.7606
2003	-2.0629	-1.0755	-1.5318	-1.1223	-0.4577	-0.7106
2004	-2.1228	-1.0319	-1.6455	-1.1541	-0.4348	-0.8115
2005	-2.1829	-0.9880	-1.6189	-1.1834	-0.4151	-0.7651
2006	-2.2381	-0.9500	-1.6573	-1.2137	-0.3945	-0.8031
2007	-2.2939	-0.9112	-1.6724	-1.2431	-0.3753	-0.7857
2008	-2.3487	-0.8733	-1.6676	-1.2702	-0.3571	-0.7717
2009	-2.3990	-0.8390	-1.7730	-1.2988	-0.3380	-0.7741

(c) Portugal

	x=60			x=70			x=80		
t	95% PI		Actual value	95% PI		Actual value	95% PI		Actual value
	Up B.	Low B.		Up B.	Low B.		Up B.	Low B.	
2000	-4.7066	-3.9522	-4.4006	-3.8358	-3.0136	-3.4477	-2.7732	-2.0580	-2.4174
2001	-4.8325	-3.8508	-4.4030	-3.9827	-2.8943	-3.4950	-2.8888	-1.9652	-2.4902
2002	-4.9381	-3.7692	-4.5044	-4.1050	-2.7993	-3.5153	-2.9855	-1.8910	-2.4653
2003	-5.0367	-3.6946	-4.4533	-4.2187	-2.7128	-3.4940	-3.0754	-1.8231	-2.5002
2004	-5.1270	-3.6272	-4.5181	-4.3214	-2.6370	-3.6001	-3.1572	-1.7632	-2.5275
2005	-5.2140	-3.5650	-4.4892	-4.4198	-2.5651	-3.6227	-3.2372	-1.7058	-2.5157
2006	-5.2962	-3.5075	-4.4954	-4.5145	-2.4988	-3.6885	-3.3123	-1.6530	-2.5828
2007	-5.3785	-3.4490	-4.4733	-4.6057	-2.4354	-3.7206	-3.3865	-1.6010	-2.5627
2008	-5.4562	-3.3954	-4.5403	-4.6952	-2.3725	-3.7285	-3.4590	-1.5502	-2.5732
2009	-5.5290	-3.3465	-4.5708	-4.7800	-2.3146	-3.7243	-3.5266	-1.5045	-2.6045

	x=90			x=100		
t	95% PI		Actual value	95% PI		Actual value
	Up B.	Low B.		Up B.	Low B.	
2000	-1.7562	-1.1486	-1.4443	-0.9781	-0.4296	-0.6758
2001	-1.8262	-1.0944	-1.4653	-0.9989	-0.4157	-0.6769
2002	-1.8884	-1.0482	-1.4877	-1.0189	-0.4023	-0.6744
2003	-1.9470	-1.0049	-1.4203	-1.0408	-0.3875	-0.6642
2004	-2.0037	-0.9639	-1.5946	-1.0614	-0.3738	-0.7622
2005	-2.0542	-0.9286	-1.4342	-1.0813	-0.3609	-0.6674
2006	-2.1069	-0.8914	-1.5836	-1.1012	-0.3485	-0.7414
2007	-2.1592	-0.8559	-1.5311	-1.1219	-0.3347	-0.7582
2008	-2.2075	-0.8224	-1.4946	-1.1418	-0.3217	-0.7629
2009	-2.2542	-0.7905	-1.6661	-1.1615	-0.3099	-0.7741

(d) Spain

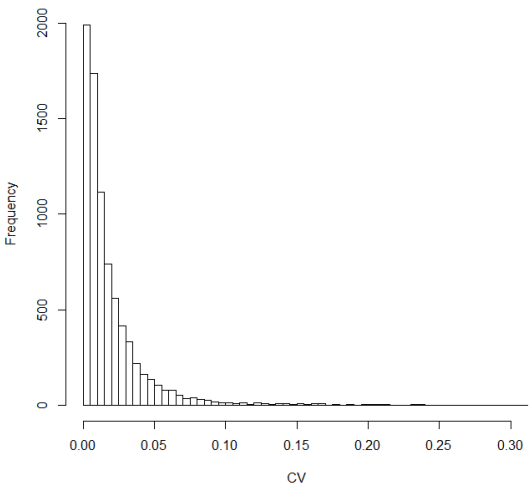
	x=60			x=70			x=80		
t	95% PI		Actual value	95% PI		Actual value	95% PI		Actual value
	Up B.	Low B.		Up B.	Low B.		Up B.	Low B.	
2000	-4.8579	-4.1486	-4.5375	-4.0072	-3.2149	-3.6268	-2.9560	-2.2346	-2.5579
2001	-4.9948	-4.0364	-4.5612	-4.1679	-3.0828	-3.6153	-3.0980	-2.1186	-2.6364
2002	-5.1104	-3.9457	-4.5602	-4.3009	-2.9785	-3.6766	-3.2134	-2.0281	-2.6282
2003	-5.2139	-3.8671	-4.5422	-4.4202	-2.8872	-3.6679	-3.3210	-1.9464	-2.6194
2004	-5.3089	-3.7964	-4.6032	-4.5296	-2.8057	-3.7297	-3.4180	-1.8737	-2.6573
2005	-5.4001	-3.7303	-4.5863	-4.6345	-2.7290	-3.7503	-3.5114	-1.8063	-2.6116
2006	-5.4855	-3.6698	-4.5824	-4.7335	-2.6593	-3.8027	-3.5991	-1.7439	-2.7440
2007	-5.5687	-3.6114	-4.6295	-4.8296	-2.5917	-3.7810	-3.6848	-1.6836	-2.7161
2008	-5.6523	-3.5528	-4.6183	-4.9237	-2.5258	-3.8295	-3.7681	-1.6255	-2.7657
2009	-5.7290	-3.5002	-4.7161	-5.0119	-2.4652	-3.8685	-3.8470	-1.5721	-2.7853

	x=90			x=100		
t	95% PI		Actual value	95% PI		Actual value
	Up B.	Low B.		Up B.	Low B.	
2000	-1.8425	-1.3199	-1.5932	-0.9796	-0.6158	-0.7771
2001	-1.9231	-1.2565	-1.5937	-0.9986	-0.6030	-0.7818
2002	-1.9922	-1.2027	-1.6224	-1.0165	-0.5899	-0.7872
2003	-2.0574	-1.1537	-1.5202	-1.0340	-0.5783	-0.7302
2004	-2.1149	-1.1115	-1.6183	-1.0512	-0.5671	-0.7838
2005	-2.1721	-1.0694	-1.5957	-1.0676	-0.5569	-0.7632
2006	-2.2258	-1.0328	-1.6630	-1.0863	-0.5434	-0.8045
2007	-2.2808	-0.9931	-1.6424	-1.1022	-0.5334	-0.7979
2008	-2.3306	-0.9589	-1.6481	-1.1183	-0.5232	-0.7949
2009	-2.3799	-0.9258	-1.6944	-1.1350	-0.5122	-0.8178

Table 4: 95% prediction intervals and actual values of the log-central death rates

Figure 1: CV and RMSE Histograms

(a) CV Histogram



(b) RMSE Histogram

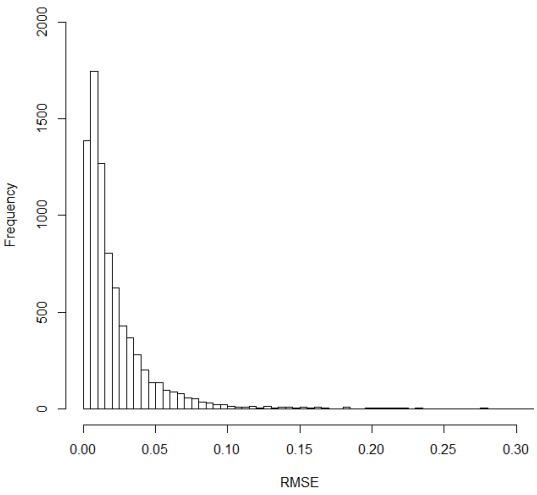
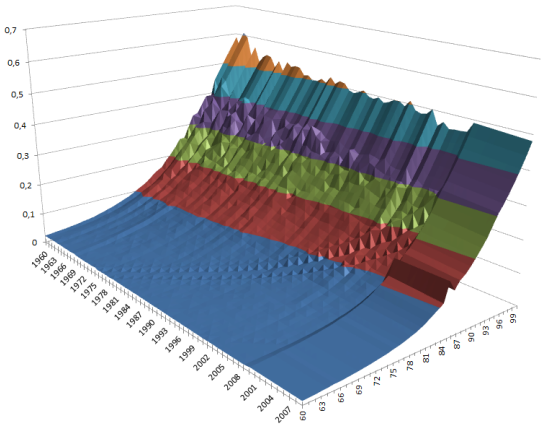
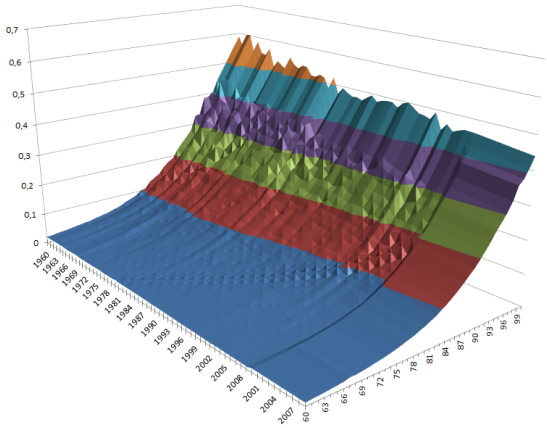


Figure 2: Observed and forecasted mortality surfaces

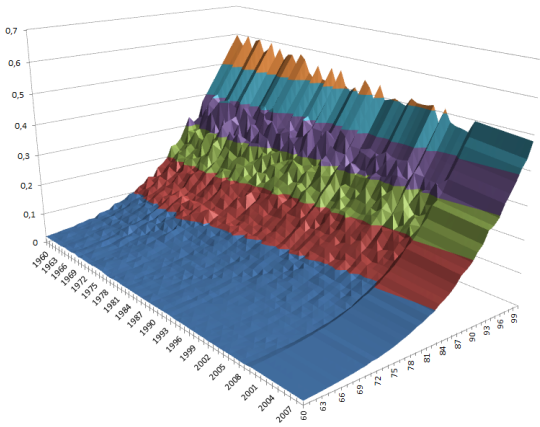
(a) France



(b) Italy



(c) Portugal



(d) Spain