THE LEE-CARTER MODEL UNDER THE CONDITION OF VARIABLES AGE-SPECIFIC PARAMETERS

Marie-Claire Koissi*   Arnold F. Shapiro**

ABSTRACT

In recent years, unexpected level of mortality improvement has become an increasing challenge for life annuities business. As a result, the need for robust and reliable models for mortality projection has become a growing issue among actuaries and policies makers. The model proposed by Lee and Carter in 1992 (Lee and Carter, 1992) seems to be generally accepted, because first it produced satisfactory fits and forecasts of mortality rates for various countries. Secondly, the structure of the Lee-Carter model allows the construction of confidence intervals related to mortality projections. To improve the performance of the Lee-Carter model several extensions to the original version have been proposed.

In this paper, we propose a modification of the Lee-Carter model that accommodates variations in age-specific parameters. The Lee-Carter assumption of constant age-specific pattern of mortality over the year is known to be unrealistic (Gutterman and Vanderhoof, 1999; Tuljarpurkar and Boe, 1998). The paper also proposes an extended weighted least square approach to find the model parameters. Finally, the paper investigates the horizon beyond which forecasts conditioned on past observations are no longer relevant. The economics notion of content function is used for this purpose.

Key words: Mortality projection; Lee-Carter method; Singular Value Decomposition; Weighted Least Square methods; Time series modeling; Forecast horizon.

*Department of Mathematics, Western Illinois University, Macomb, IL 61455, USA. mk122@wiu.edu
**Smeal College of Business, Penn State University, University Park, PA 16802, USA. afs1@psu.edu

1. INTRODUCTION

Mortality forecasts have a long history in demography and actuarial science. Demographers used mortality forecasts for population projections, and actuaries used mortality forecasts for cash flow projections and the assessment of premium and reserves in life insurance and pension annuities. Official agencies also used mortality forecasting to support their policy decisions. Earlier models for mortality rates were deterministic, and include the Gompertz curve (1825), which provides a satisfactory fit to adult mortality, but overestimates death rates at ages greater than 80. The Perks (1932) logistic model (a generalization of the Gompertz curve) gives a relatively good fit to mortality rates over the entire adult range. The Heligman and Pollard (1980) curve provides a relatively good fit to mortality rates over all ages. Progresses in computational algorithms help handling complexes models, and the number of parameters is no longer an issue. Reviews of earlier contributions to mortality forecast and recent models are provided by Pitacco (2004), Wong-Fupuy and Haberman (2004) and Tuljapurkar and Boe (1998). Some recent studies shown that the mortality rates predicted from the classic parametric formulas were erratic (Stoto 1983; Murphy 1995). Stochastic models seem more appealing, because they associate a confidence error to each estimate.

In 1992, Lee and Carter presented a stochastic model, based on a factor analytic approach, to fit and predict mortality rates for United States. Since then, because of its simplicity and relatively good performance, the Lee-Carter (LC) model has been widely used for demographic and actuarial applications in various countries. For example, the LC model was used for Japan (Wilmoth 1993), the seven most economically developed countries (G7) (Tuljapurkar et al. 2000), Austria (Carter and Prskawetz 2001), Australia (Booth et al. 2002), Belgium (Brouhns et al. 2002), and the Nordic countries (Lundström and Qvist 2004; Koissi et al. 2006b). Despite its reasonable performance, the LC model

had several limitations (Lee 2000) which caused negative reactions from some actuaries (Gutterman and Vanderhoof 1999).

1.1 Lee-Carter model

The Lee-Carter (1992) model is as follows

$$\ln(m_{x,t}) = \alpha_x + \beta_x \kappa_t + \xi_{x,t}, \qquad \sum_t \kappa_t = 0, \qquad \sum_x \beta_x = 1 \qquad (1.1)$$

where $m_{x,t}$ is the matrix of the central death rates at age x ($x = x_1,...,x_N$) in year t ($t = t_1, t_1 + 1,..., t_1 + T - 1$). The term $\xi_{x,t}$ represents the deviation of the model from the observed log-central death rates and is expected to be Gaussian $\xi_{x,t} \sim N(0, \sigma_\varepsilon^2)$. Model (1.1) is underdetermined because it is invariant under the transformations $(\alpha_x, \beta_x, \kappa_t) \rightarrow (\alpha_x, \beta_x / c, c\kappa_t)$ and $(\alpha_x, \beta_x, \kappa_t) \rightarrow (\alpha_x - c\beta_x, \beta_x, \kappa_t + c)$, for any constant c. The constraints insure a unique solution for (1.1).

By summing over the years, t, and using $\sum_t \kappa_t = 0$, an estimate of the age-dependent parameter, $\hat{\alpha}_x = (1/T)\sum_t \ln(m_{x,t})$, is obtained. The parameter $\hat{\alpha}_x$ is interpreted as the average pattern of mortality at age x. By summing both sides of (1.1) over the ages, and using $\sum_x \beta_x = 1$, an estimate of κ_t, $\hat{\kappa}_t = \sum_x (\ln(m_{x,t}) - \hat{\alpha}_x)$, is obtained. The coefficient κ_t represents the general level of mortality at time t.

An estimate for β_x, $\hat{\beta}_x = (\partial \ln(m_{x,t})/\partial t)/(\partial \hat{\kappa}_t / \partial t)$, is obtained by differentiating both sides of (1.1) with respect to time. Then, the parameter β_x captures the relative sensitivity of the logarithm of the central death rates to change in the mortality index κ_t. The function β_x moderates the time-dependent element κ_t by age.

1.2 Motivations for proposed extensions.

The extensions proposed in the current paper address the following issues:

1. Variations in age-specific parameters. The LC model assumes that the combined effect of past process will remain the same in the future (Lee 2000:85). As a consequence, the age-specific pattern of mortality change is kept constant over the year. Such an assumption, however, is not realistic (Tuljarpurkar and Boe 1998; Gutterman and Vanderhoof 1999:135). We extend the investigation of Carter and Prskawetz (2001) and introduce a modification of the Lee-Carter model that accommodates variations in age-specific parameters. With such formulation, predicting mortality rates requires forecasting not only the mortality index, but also the age-specific terms.

2. Relevant length of forecast. Wong-Fupuy and Haberman (2004:80) and Alho and Spencer (1985:314) mentioned the issue of a reasonable length for a forecast and questioned the reliability of long-term forecasts. In economics, the forecast content function and content horizon (Galbraith 2003) are used to set the horizon beyond which forecasts conditioned on past observations are no more relevant. These notions are adapted to the present model.

The paper is organized as follows. Section 2 presents the standard methods for estimating the parameters of the LC model. Some previous extensions of the model are also summarized. Section 3 deals with the extensions we propose. First, we present a modification of the LC which accommodates variations in the age-specific parameters. Then, we discuss the length of reasonable forecast horizon. The paper ends with some concluding remarks.

2. ESTIMATING THE PARAMETERS OF THE LC MODEL AND MORTALITY FORECASTING

2.1 Singular Values Decomposition of $[\ln(m_{x,t}) - \hat{\alpha}_x]$

Lee and Carter used Singular Values Decomposition (SVD) (Lawson and Hanson 1974) of the matrix $Y_{x,t} = [\ln(m_{x,t}) - \hat{\alpha}_x]$ to obtain estimates of β_x and κ_t:

$$SVD(Y_{x,t}) = \sum_{i=1}^{r} \rho_i U_{x,i} V_{i,t}, \tag{2.1}$$

where $r = rank(Y_{x,t})$, $\{\rho_1 \geq \rho_2 \geq \cdots \geq \rho_r\}$ are the ordered singular values of $Y_{x,t}$, $U_{x,i}$ and $V_{t,i}$ are the left and right singular vectors. The SVD code is available in standard mathematical software, such as MATLAB (Appendix A.1). By using the theorem of low rank approximation (first stated and proved by Eckart and Young (1936)), the rank h least square approximation of (2.1) is obtained

$$\hat{Y}_{x,t}^{(h)} = \sum_{i=1}^{h} \rho_i U_{x,i} V_{i,t} = \sum_{i=1}^{h} \beta_x^{(i)} \kappa_t^{(i)}, \quad h \leq r \tag{2.2}$$

where $\beta_x^{(i)} \kappa_t^{(i)} = \rho_i U_{x,i} V_{i,t}$. Then, the rank h residuals associated with (2.1) are

$$\xi_{xt} = \sum_{i=h+1}^{r} \rho_i U_{x,i} V_{i,t}. \tag{2.3}$$

The corresponding rank-h approximation least square errors is

$$\xi_h^2 = \sum_{i=h+1}^{r} \rho_i^2, \tag{2.4}$$

which implies that the errors have similar variance. However, this assumption is violated for mortality data: the variance of the log-central death rate is approximately $Var[\ln(m_{x,t})] \approx 1/d_{x,t}$ (Wilmoth 1993:2), which is smaller at younger ages than at older ages (because there are fewer deaths at old ages).

Some studies discussed the choice of a suitable value for h (Renshaw and Haberman 2003; Booth et al. 2002), although the first order approximation ($h = 1$) is used in Lee and Carter (1992), and most applications of the model. The proportion of variance explained by the ith term ($\rho_i U_{x,i} V_{i,t}$) of the decomposition (2.1) is given by $\rho_i^2 / \sum_{j=1}^{r} \rho_j^2$, and the total variance explained by a rank-h approximation is $\sigma_h^2 = \sum_{i=1}^{h} \rho_i^2 / \sum_{j=1}^{r} \rho_j^2$. It is clear that $0 \leq \sigma_h^2 \leq 1$ and the closer this value is to 1, the better is the approximation. For US data, Lee and Carter (1992) restrained the SVD approximation to the first order $\hat{Y}_{x,t}^{(1)} \approx \rho_1 U_{x,1} V_{1,t} = \beta_x^{(1)} \kappa_t^{(1)}$ and obtained an explained variance $\sigma_1^2 = 92.7\%$. Tuljapurkar et al. (2000) also used a rank-1 SVD approximation for G7 countries and found an explained variance greater than 94%. Li et al. (2004) used the standard LC for South Korea, with $\sigma_1^2 < 85\%$.

A second stage in standard LC method involves finding a modified $\tilde{\hat{\kappa}}_t^{(i)}$, which adjusts the total number of death $\sum_x d_{x,t}$ to the estimated number of deaths as follows:

$\sum_x d_{x,t} = \sum_x E_{x,t} \exp(\hat{\alpha}_x + \sum_i \hat{\beta}_x^{(i)} \tilde{\hat{\kappa}}_t^{(i)})$, $\forall t$, where $E_{x,t}$ and $d_{x,t}$ are the exposure to risk and the actual number of deaths at age x and time t. This step is necessary because the LC model fits the logarithm of the death rates instead of the current death rates (Bell 1997; Lee 2000).

Predicting mortality with the LC model is reduced to forecasting the index κ_t using time series approaches (Brockwell and Davis 1996). In general an ARIMA(0,1,0) with drift, $\hat{\kappa}_t = \hat{\kappa}_{t-1} + c + \xi_t$, is found suitable, though other ARIMA forms provided better fit to some data (Brouhns et al. 2002).

2.2 Previous Extensions of the LC Model

2.2.1 Weighted Least Squares

The parameters of the LC model can also be estimated using the Weighted Least Squares (WLS) suggested by Wilmoth (1993):

$$\text{Minimize } \sum_{x=1}^{X} \sum_{t=1}^{T} w_{x,t} [\ln(m_{x,t}) - \alpha_x - \beta_x \kappa_t]^2 \tag{2.5}$$

The inverse of the sample variance is commonly chosen as weights (Barlow 1989:93). Since $\text{Var}[\ln(m_{x,t})] \approx 1/d_{x,t}$ (Wilmoth 1993:2), a suitable choice is $w_{x,t} = d_{x,t}$, where $d_{x,t}$ is the observed number of deaths for age-group x in year t.

The use of the term "weight" may suggest a summation to one, which does not hold here. An alternative choice is $w_{x,t} = d_{x,t}/(\sum_x \sum_t d_{x,t})$ for $d_{x,t} \neq 0$ and $w_{x,t} = 0$ for $d_{x,t} = 0$, which gives $\sum_x \sum_t w_{x,t} = 1$ as expected. The WLS problem becomes:

minimize $D \sum_{x=1}^{X} \sum_{t=1}^{T} d_{x,t} [\ln(m_{x,t}) - \alpha_x - \beta_x \kappa_t]^2$, $\quad (D = 1/(\sum_x \sum_t d_{x,t}) > 0)$ which is equivalent to (2.4). Wilmoth's WLS approach gave satisfactory fit to data from Japan, 1951-1990 (Wilmoth 1993), Austria, 1947-1999 (Carter and Prskawetz 2001), and the Nordic countries, 1955-1999 (Koissi et al. 2006b).

2.2.2 Maximum Likelihood Estimation

Wilmoth (1993) and Alho (2000) proposed using Maximum Likelihood Estimation to find the parameters in the LC model (1.1). This approach is based on a Poisson approximation of the number of death $D_{x,t}$ presented by Brillinger (1986):

$$D_{x,t} \sim Poisson(m_{x,t} E_{x,t}), \text{ where } m_{x,t} = \exp(\alpha_x + \beta_x \kappa_t). \tag{2.6}$$

The coefficients α_x, β_x and κ_t are estimated by maximizing the full log-likelihood (Wilmoth 1993:5)

$$l = \sum_{x=x_1}^{x_A} \sum_{t=t_1}^{t_1+T-1} [D_{x,t} \ln(m_{x,t} E_{x,t}) - E_{x,t} \exp(\alpha_x + \beta_x \kappa_t) - \ln(D_{x,t}!)] \qquad (2.7)$$

The Maximum Likelihood Estimation (MLE) allows non-additive heteroscedastic errors (Renshaw and Haberman 2003:255), and avoids the assumption of errors with constant variance present in the SVD approach (Lee and Carter 1992:660). The MLE formulation of the LC model is often referred to as the Poisson log-bilinear model from the paper by Brouhns et al.(2002), which provides a detail algorithm to minimize (2.7). The MLE was used to model death rates from Belgium (Brouhns et al. 2002), UK and Wales (Renshaw and Haberman 2003) and the Nordic countries (Koissi et al. 2006b).

2.2.3 Expanded SVD

Booth et al. (2002) investigated the performance of the LC model on Australian data after the second term of the SVD was incorporated in the approximation:

$$\hat{Y}_{x,t}^{(2)} \approx \rho_1 U_{x,1} V_{1,t} + \rho_2 U_{x,2} V_{2,t} = \beta_x^{(1)} \kappa_t^{(1)} + \beta_x^{(2)} \kappa_t^{(2)}. \qquad (2.8)$$

Their result suggested that the second term $\rho_2 U_{x,2} V_{2,t}$, which represented 1.3 percent of the total variance in the log-central death rates, explained the cohort-period effect present in the data. As a result, the corresponding second order residuals exhibited less systematic pattern than the first order residuals (Booth et al. 2002:332).

Similarly, Renshaw and Haberman (2003) introduced up to the 5th terms of the SVD to fit mortality from England & Wales. This extension improves the fitting results in the sense that the additional terms (from the 2nd to the 5th) account for more than 4 percent (and nearly 13 percent for one group of data) of the total variance explained by the approximation. However, the standard LC model is still found attractive, because it limits the forecasting computations to the study of only one mortality index $\hat{\kappa}_t$ (Li et al. 2004).

2.2.4 Dealing with Uncertainties

The original LC model incorporates uncertainty arising from all three parameters: $\text{Var}(\hat{\alpha}_x)$ is computed through the variance of the death rates $m_{x,t}$ over time, $\text{Var}(\hat{\beta}_x)$ can be computed by bootstrap, and $\text{Var}(\hat{\kappa}_t)$ is derived from the time series model associated with the index $\hat{\kappa}_t$ (Lee and Carter 1992:670).

The bootstrap approach was the basis for likelihood-based methods of estimating the LC parameters. These methods assume that the observed number of deaths follow a Poisson distribution with mean equal to the expected number of death under the LC model. For the likelihood-based models, bootstrap and Bayesian technique were used to compute confidence intervals for all the LC parameters.

Brouhns et al. (2005) used a pair bootstrap procedure to compute the confidence interval of the LC parameters, and related predicted demographic and actuarial rates. The approach relies on sampling the matrix of number of deaths, $D_{x,t}$, from a Poisson distribution: $D_{x,t}^b \sim \text{Poisson}(D_{x,t})$, b=1,..., B, where B is a large number of samplings and $D_{x,t}^b$ is the matrix of death obtained at the bth sampling. Then, for each b=1,..., B, the LC parameters are obtained by replacing $D_{x,t}$ by $D_{x,t}^b$ in (2.7).

Koissi et al. (2006b) used a residual bootstrap method. This approach consisted in sampling with replacement the matrix of deviance residuals r_D, which resulted from fitting observed central death rates with the maximum likelihood formulation of the LC model: $r_D = sign(D_{x,t} - \hat{D}_{x,t})[D_{x,t} \ln(D_{x,t}/\hat{D}_{x,t}) - (D_{x,t} - \hat{D}_{x,t})]^{1/2}$, where $\hat{D}_{x,t} = E_{x,t}\hat{m}_{x,t}$ are the fitted number of deaths. A large number B of replications $\{r_D^b, b = 1, \cdots, B\}$ is generated, leading to corresponding B sets of LC parameters $\{\hat{\alpha}_x^b, \hat{\beta}_x^b, \hat{\kappa}_x^b\}$. Then, the bootstrap percentile approach provides the desired confidence interval.

Czado et al. (2005) applied Bayesian technique to the LC Poisson log-bilinear model. A probability distribution is assigned to α_x, β_x and κ_t. The parameters of these laws are obtained by Markov Chain Monte Carlo simulations. The uncertainties associated with each estimate can then be obtained. A particularly attractive feature of this method is that the step consisting in modelling and predicting κ_t is incorporated in the simulation.

It is worth briefly mentioning that the two principal sources of uncertainties in a regression model are randomness and fuzziness, although the later generally is neglected. Koissi and Shapiro (2006a) proposed a fuzzy formulation of the LC model which incorporates randomness and fuzziness. The LC parameters α_x, β_x and κ_t are formulated as random fuzzy numbers, and their distributions are computed using Bayesian technique.

Many extensions to the original LC model were proposed leading to improvements in the model's assumptions, the computation methods and the statistical properties of the estimated parameters. There are still few issues related to the LC model which raised criticisms among researcher. Next, we discuss some of these issues: the LC model under the condition of variations in age-specific parameters, and the search for a reasonable length of forecast.

3. PROPOSED EXTENSIONS CRITICAL ANALYSIS OF LC

3.1 LC under Variables Age-Specific Parameters

In the LC model, the age-specific vectors α_x and β_x are treated as invariant in the sense that the age parameters obtained when fitting the data are kept unchanged during the prediction process. The hypothesis is that the relative change in mortality pattern by age observed in the past will remain unchanged. The possibility that such a "heroic

assumption" (Gutterman and Vanderhoof 1999:135) will be violated in the future has been discussed in several studies (Lee 2000; Gutterman and Vanderhoof 1999). In what follows, we conduct two studies which justify the need for reformulation of the β_x parameters.

a- Preliminary studies

Application of LC to four Eastern European countries

The LC model is applied to mortality data of four eastern European countries: Bulgaria (1955-97), Hungary (1955-99), Lithuania (1960-2001) and Russia (1970-99). Figure 1 depicts the resulting age-specific parameter β_x. High values of β_x indicates that the rate of improvement in mortality at these ages is faster than in general (Lee, 2000:85), while the negative values at some ages mean that mortality is increasing. As the figure suggests, in Bulgaria, from 1955 to 1997, the mortality rates for women aged 75 years and older has not improved. In Hungary, over the period from 1955 to 1999, the mortality rate for females aged 35 to 60 has worsened. In Lithuania, during the period 1960 to 2000, similar conclusions were drawn for women aged 40 to 60 and women aged 75 years and older,. The graph also shows that mortality among young Russian girls aged 12 years and less has not improved (relatively to mortality at other ages) during the period from 1970 to 1999. It is reasonable to assume that the variation in mortality observed during these past periods will not be the same in future years. Especially, the rate of improvement for the period prior to the year 1991 (when the Russian federation was dislocated) is likely to be different from the age-specific values after 1992. In addition, the progress in medical research and political changes (such as entry in European Union, for some countries) are more likely to bring changes in the shape of the β_x function.

Figure 1 is about here.

Application of LC to Finland, with variable periods

The LC model is applied to Finnish (women) mortality rates over 35 periods, all starting from year 1955: $Period(i) = [1955;1963 + i]$, $i = 1,...,35$. Figure 2 depicts, for selected age-groups, x, the β_x coefficients by period. The years in the abscises represent the end of the observation period. The results show that for each age-group, the β_x values are not constant over time. No regular pattern emerges from these graphs. However, the fluctuations in β_x are less "unpredictable" when the last twenty periods only are used. The straight line that represents β_x is obtained by taking the entire period 1955-1998.

Figure 3 depicts the coefficients α_x for selected ages of Finish women. The graphs suggest that α_x values decrease linearly over the observation periods. Since the parameter α_x is a measure of the overall force of mortality at age x, the negative trend in α_x is in accord with the observed improvement in mortality. For each period, the mortality index κ_t has the expected (almost) linear decreasing trend (Figure 4). In the following, we focus on the age-specific terms α_x and β_x.

Figures 2-4 are about here.

b- Modeling the age-specific coefficients $\beta(x,s)$ and $\alpha(x,s)$

The variation in the coefficients α_x and β_x over the years, for each age x, suggests the use of a LC reformulation that includes the observation period. Carter and Prskawetz (2001:6) studied the trends in α_x and β_x over variable time intervals with identical length. When using data from Austria (1947-1999), Carter and Prskawetz (2001:9) noted some differences (from 1.5 to 3.5 years) between observed and estimated life expectancies by observation period.

Following the lead of Carter and Prskawetz (2001:8), we investigated an expanded LC model with a component related to the observation period. For example, for 35 observation periods, all starting from 1955,

$$\ln(m(x,t,s)) = \alpha(x,s) + \beta(x,s)k(t) + \xi(x,t), \qquad (3.1)$$

where $s = 1,...,35$ and $t(1) = 1955 - 1964$, $t(2) = 1955 - 1965$,..., $t(35) = 1955 - 1998$.

Our aim is to model and predict $\alpha(x,s)$ and $\beta(x,s)$, for each age-group $x = \{1, \cdots, 24\}$.

Modeling $\alpha(x,s)$

The previous result suggests using a linear relation between α_x and the period $t(s)$

$$\alpha(x,s) = a_0(x) + a_1(x) \times s, \qquad (3.2)$$

where $s = 1,...,35$ are the number of "ex-post" periods. Table 1 shows the $a_0(x)$ and $a_1(x)$ values obtained using the Finnish data previously described. The slope a_0 ranges between -0.0279 and -0.0043. Larger values (in absolute value) are obtained for younger ages.

Modeling $\beta(x,s)$

For each age-group x, the age-specific vector $\beta(x,s)$ is modeled using polynomial interpolation. The analysis of the residuals suggests that a polynomial of order 4 is suitable.

$$\beta(x,s) = \beta_0(x) + \beta_1(x) \times s + \beta_2(x) \times s^2 + \beta_3(x) \times s^3 + \beta_4(x) \times s^4 \qquad (3.3)$$

The coefficients are found using MATLAB software. The last sixth columns of Table 1 display the $\beta_i(x)$ values obtained for each age-group. $\beta_4(x)$ and $\beta_3(x)$ are very small compare to the other coefficients. The last column of Table 1 depicts $\sum_{i=0}^{4} \beta_i(x)$. Then, the LC model with variables age-specific parameters is as follows

$$\ln(m(x,t,s)) = \alpha(x,s) + \beta(x,s)\kappa(t) + \xi(x,t) , \qquad \xi(x,t) \sim N(0,\sigma_\varepsilon^{2}) \qquad (3.4)$$

where $\alpha(x,s) = a_0(x) + a_1(x) \times s$

$$\beta(x,s) = \beta_0(x) + \beta_1(x) \times s + \beta_2(x) \times s^2 + \beta_3(x) \times s^3 + \beta_4(x) \times s^4 ,$$

$$\kappa_s(t) = \kappa_s(t-1) + k_{s0} + \varepsilon_t .$$

$$s = 1,...,35 ; \ t(s) = \left[\ 1955;1963 + s \ \right]; \ x = x_1,...,x_{24} .$$

Table 1 and Figure 5 are about here.

c- Comparison with Standard LC model

The standard LC (1.1) is used to fit Finnish (women) central death rates, from 1955 to 1998. The estimates $\hat{\alpha}_x , \hat{\beta}_x$ and $\hat{\kappa}_t$ are obtained using SVD. Note that the parameters κ_t are similar in both models. We only introduce variations in α and β. For a period from 1955 to (1998+n), we get

$$\alpha(x, j) = a_0(x) + a_1(x) \times j , \text{ where } j = 35 + n . \qquad (3.5)$$

$$\beta(x, j) = \beta_0(x) + \beta_1(x) \times j + \beta_2(x) \times j^2 + \beta_3(x) \times j^3 + \beta_4(x) \times j^4 \qquad (3.6)$$

Figure 6 displays the results obtained with the standard LC (denoted by "LC") and Model (3.4) denoted by "period-LC". The parameters $\beta(x,s)$ and $\alpha(x,s)$ are forecasted over 5 years. Almost no variation is observed between $\alpha(x,s)$ and $\hat{\alpha}_x$, whereas $\beta(x,s)$ fluctuate. In fact, it can be shown that, under (3.5), the extended model is equivalent to the standard LC[1]. This result suggests first applying graduation (Whittaker for example, see Verrall 1996 and Shiu 1986) to the original series. Secondly, restricting the ex-post study to ten years backward may result in more accurate $\beta(x,s)$

---

[1] $\ln(m(x,t,s)) = \alpha(x,s) + \beta(x,s) \, k(t) + \xi(x,t) = a_0(x) + a_1(x) \times j + \beta(x,s) \, k(t) + \xi(x,t) = A(x) + B(x) \cdot K(x)$

Figure 6 is about here.

3.2 Forecast Horizon

   The horizon of a forecast obviously is a relevant issue. Nonetheless, very few studies on this topic can be found in the literature related to mortality models (Alho and Spencer 1985; Booth et al. 2002), compared to economic forecast (Galbraith 2003; Oke and Öller 1999; Diebold and Kilian 2001). For economic variables predicted with autoregressive models, Galbraith (2003) defined the "forecast content function" and the "content horizon" beyond which forecasts conditioned on past observations are no longer relevant. In this section, this approach is adapted to the LC model.

a- Definitions (Galbraith, 2003)

Given a sequence $\{y_t, t = 1,...,T\}$ of observations on a covariance stationary, ergodic, scalar process y, denote by $\{\hat{y}_{T+s|T}, s = 1,..., H\}$ the conditional forecasted values of y up to horizon H and denote by $\bar{y} = (1/T)\sum_{t=1}^{T} y_t$ the sample mean. At a certain time, the observed series $\{y_t, t = 1,...,T\}$ will be obsolete and will no longer, improve the forecasts. Hence $E(\hat{y}_{T+s|T} - y_{T+s})^2 \geq E(\bar{y}_T - y_{T+s})^2$, i.e. the forecasts based on unconditional mean will perform better. The relative variation between the conditional and unconditional mean squared errors is measured by the forecast content function

$$C(s) = 1 - \frac{E(\hat{y}_{T+s|T} - y_{T+s})^2}{E(\bar{y}_T - y_{T+s})^2} = 1 - \frac{MSE(\hat{y}(s))}{MSE(\bar{y}(s))}, \quad s = 1,..., H . \tag{3.7}$$

where $MSE(\hat{y}(s))$ is the mean squared error of the conditional predictions and $MSE(\bar{y}(s))$ is the mean squared error for predictions based on the sample mean. When $MSE(\hat{y}(s)) < MSE(\bar{y}(s))$, the conditional forecasts $\hat{y}_{T+s|T}$ is said to have a positive

content and the past observations improve the predictions. The content horizon as defined in Galbraith (2003) is the smallest forecast horizon s_0 verifying $C(s) \leq 0$ for $s \geq s_0$. A δ-level content horizon is such that $C(s) \geq \delta$ for $s \leq s_\delta$ and $C(s) < \delta$ for $s > s_\delta$.

b- Computation of the forecast content function

For autoregressive series: $y_t = \alpha_0 + \sum_{j=1}^{p} \alpha_j y_j + e_t, e_t \sim N(0, \sigma^2)$, the characteristic equation is given by $\lambda^p - \sum_{j=1}^{p} \alpha_j \lambda^{p-j}$. The forecast content function differs as the largest root, λ_{max}, of the characteristic equation verifies $\lambda_{max} \leq 1$ (Fuller and Hasza 1981). We compute the forecast content function of the LC parameter κ_t, for Finish women.

The mortality index κ_t is modeled as a random walk with drift: $\kappa_{t+1} = u + \kappa_t + \varepsilon_t$, where $\varepsilon_t \sim N(0, \sigma^2)$. The associated characteristic function has a unique non-zero root equal to one. The observations are $\{\kappa_t, t = t_0, \ldots, t_0 + T - 1\}$, where $t_0 = 1955$ and $t_0 + T - 1 = 1999$. Let $t_0 \equiv 1$, then the sequence of observations are indexed as follows $\{\kappa_t, t \equiv 1, \ldots, T\}$.

The forecasts verify $\hat{\kappa}_{T+s|T} = u + \hat{\kappa}_{T+s-1}$. Define $K_{T+s} = (\kappa_{T+s}, 1)'$. Therefore (Fuller and Hasza 1981, Theorem 3.1) the conditional mean squared error (MSE) for K is given by

$$MSE(K) = \sigma^2 \sum_{j=0}^{s-1} A^j M A'^j \; ; \; M = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \; A = \begin{pmatrix} 1 & u \\ 0 & 1 \end{pmatrix} \text{ such that } \begin{pmatrix} \kappa_{t+1} \\ 1 \end{pmatrix} = A \begin{pmatrix} \kappa_t \\ 1 \end{pmatrix} + \begin{pmatrix} e_t \\ 0 \end{pmatrix}.$$

Then, the MSE for κ_{T+s} is the first element of the matrix $B = MSE(K)$, denoted by $B(1,1)$.

The mean squared error for predictions based on the sample mean $\bar{\kappa}_T (= T^{-1} \sum_{t=1}^{T} \kappa_t)$ is $MSE(\bar{\kappa}(s)) = E(\bar{\kappa}_T - \kappa_{T+s})^2$. An analytical expression of this mean squared error for autoregressive process is given by Galbraith (2003, Appendix A).

c- Results

Figure 7 depicts the forecast content function of the LC Kappa parameter for Finish women, based on data from 1955 to 1999. For horizon years greater than seven, $C(s)$ is less than zero. This means that the sample mean $\bar{\kappa}_T (= T^{-1} \sum_{t=1}^{T} \kappa_t)$ provides a sufficient approximation to the forecasts κ_{T+s}, beyond the seventh forecast horizon years. As a consequence, this result also suggests that the forecasted death rates and life expectancies produced with LC method are useful up to seventh year. This content horizon is shorter than the maximum horizons provided in various application of the LC model. Alho and Spencer' (1985) advice of not to exceed 15 years for forecast horizons was not pessimistic.

Figure 7 is about here.

4. CONCLUDING REMARKS

In this paper, we proposed a modification of the Lee-Carter which incorporates variations in the age-specific parameters. The parameters were computed for different periods. The time series obtained from this process are then modeled using polynomial interpolation. Such model is a realistic alternative to the standard LC assumption of constant age parameters, although it is less simple than the original model.

The paper also discusses the accuracy of the forecast length, using the forecast content function (Galbraith 2003). The results suggest that, beyond the seventh year, the past values of the mortality index are not fully relevant for predictions. Thus, forecasts horizons might preferably not exceed ten years, to keep the relevance of the information they contain.

APPENDIX

A.1 Singular Value Decomposition

Denote by A, a $m \times n$ matrix of rank k. Then, there is an $m \times m$ orthogonal matrix U, an $n \times n$ orthogonal matrix V and a $m \times n$ diagonal matrix such that (Lawson and Hanson 1974:18):

$$A = USV',\qquad\qquad\qquad\qquad (A.1)$$

where $V' = (v_{ji})$ is the transpose of matrix $V = (v_{ij})$. (A.1) has the following matrix representation (for $m < n$)

$$A = \begin{bmatrix} u_{1,1} & \cdots & \cdots & u_{1,m} \\ \vdots & \ddots & \vdots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ u_{m,1} & \cdots & \cdots & u_{m,m} \end{bmatrix} \times \begin{bmatrix} \rho_1 & 0 & \cdots & & 0 \\ 0 & \ddots & \ddots & & \vdots \\ \vdots & \ddots & \rho_{m-1} & \ddots & \vdots \\ 0 & \cdots & 0 & \rho_m & 0 \end{bmatrix} \times \begin{bmatrix} v_{1,1} & \cdots & v_{1,m} & \cdots & v_{1,n} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ v_{m,1} & \ddots & v_{m,m} & \ddots & v_{m,n} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ v_{n,1} & \cdots & \cdots & \cdots & v_{n,n} \end{bmatrix}$$

In particular, for $A = [\ln(m_{x,t}) - \alpha_x]$, x= 1,...,X, t= 1,..., T, the rank-1 approximation

$\hat{Y}_{x,t}^{(1)} = \rho_1 U_{x,1} V_{1,t} = \beta_x^{(1)} \kappa_t^{(1)}$ gives $\hat{\beta}^{(1)} = \begin{pmatrix} u_{1,1} & u_{2,1} & \cdots & u_{X,1} \end{pmatrix}'$

and $\hat{\kappa}^{(1)} = \rho_1 \times \begin{pmatrix} v_{1,1} & v_{2,1} & \cdots & v_{T,1} \end{pmatrix}$. By using the LC constraints (2.2), the estimates

coefficients are finally $\hat{\beta}^{(1)} = (1/\sum_x u_{x,1}) \begin{pmatrix} u_{1,1} & u_{2,1} & \cdots & u_{X,1} \end{pmatrix}'$ and

$\hat{\kappa}^{(1)} = (\sum_x u_{x,1}) \times s_1 \times \begin{pmatrix} v_{1,1} & v_{2,1} & \cdots & v_{T,1} \end{pmatrix}$.

The MATLAB command "svd" produces the singular vectors U and V, and the eigenvalues ρ_i of a given matrix. The algorithm to solve the standard LC follows.

```
% Construction of Lee-Carter coefficients SVD method;
% Input: matrix of central death rates;      Outputs: LC parameters;


dratefin;                        %gives the matrice death rates Y;
[X,T]=size(Y);                   %gives nber of row(ages) and columns(years) in Y;
Y=log(Y);                        %takes the logarithm;


% Construction of a(x)=(1/T)sum(Y);


A=zeros(1,X);  A=sum(Y');        %initialization and transposee matrix;
Asvd=(1/T).*A;                   %Asvd= Alpha in LC;
Asvd=Asvd';                      %to obtain a matrix with X rows and 1 column;


% Estimation of b(x) and k(t) by SVD under LC constraints;


Z=zeros(X,T);                    %matrix ln(m(x,t))-alpha(x);
  for i=1:X  for j=1:T  Z(i,j)=Y(i,j)-Asvd(i);   end;  end;


[U,S,V]=svd(Z);                  %SVD of ln(m(x,t))-alpha(x);
u1=U(1:X,1:1);                   %u1 first column of U,lenght X=ages;
v1=V(1:T,1:1);                   %v1 first column of V,lenght T=years;
Bsvd=(1/sum(u1))*u1;             %Beta in LC;


ss=sum(S);                       %largest value of S is ss(1);
Ksvd=ss(1)*v1;
Ksvd=Ksvd.*sum(u1);              %Kappa in LC;


% Second stage estimation of k(t);


deathfinf;                       %Read nber of deaths d(24,45);
exposurfinf;                     %Read nber of exposures to risk of deaths e(24,45);

% using Newton-Raphson theorem;
err=1e-5; kadjff=zeros(30,45);       %20=nber of assumed iterations needed;


kadjff(1,:)=Kff;                         %initial value of k adjusted;
for j=1:45    kadjff(2,j)=kadjff(1,j)-(sum(eff(:,j).*exp(Aff+Bff*kadjff(1,j)))-
sum(d(:,j)))/(sum(eff(:,j).*Bff.*exp(Aff+Bff*kadjff(1,j))));
i=2;
while abs(kadjff(i,j)-kadjff(i-1,j))/abs(kadjff(i-1,j))>err
  kadjff(i+1,j)=kadjff(i,j)-(sum(eff(:,j).*exp(Aff+Bff*kadjff(i,j)))-
sum(d(:,j)))/(sum(eff(:,j).*Bff.*exp(Aff+Bff*kadjff(i,j))));
  i=i+1; end; end;


% Readjustement to obtain un k that satisfies the model constraint;
kadjustff=kadjff(i-1,:)-mean(kadjff(i-1,:));
sum(kadjust);                        %to check that constraint is not violated;
```

List of Figures

List of Table

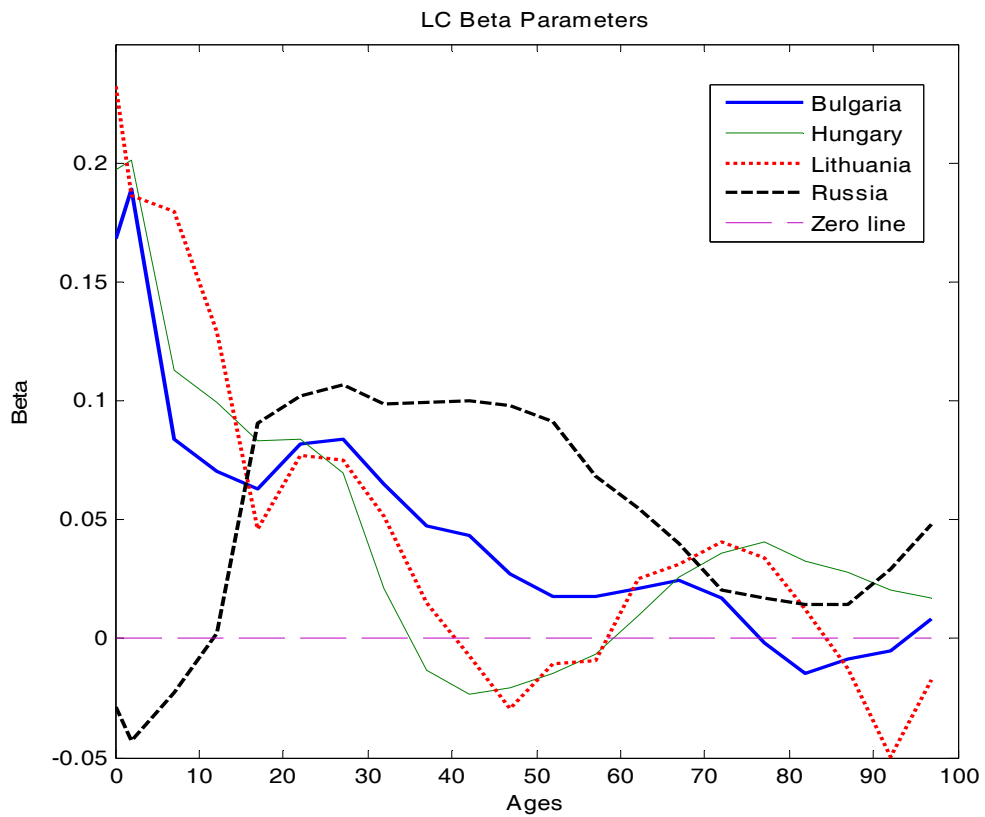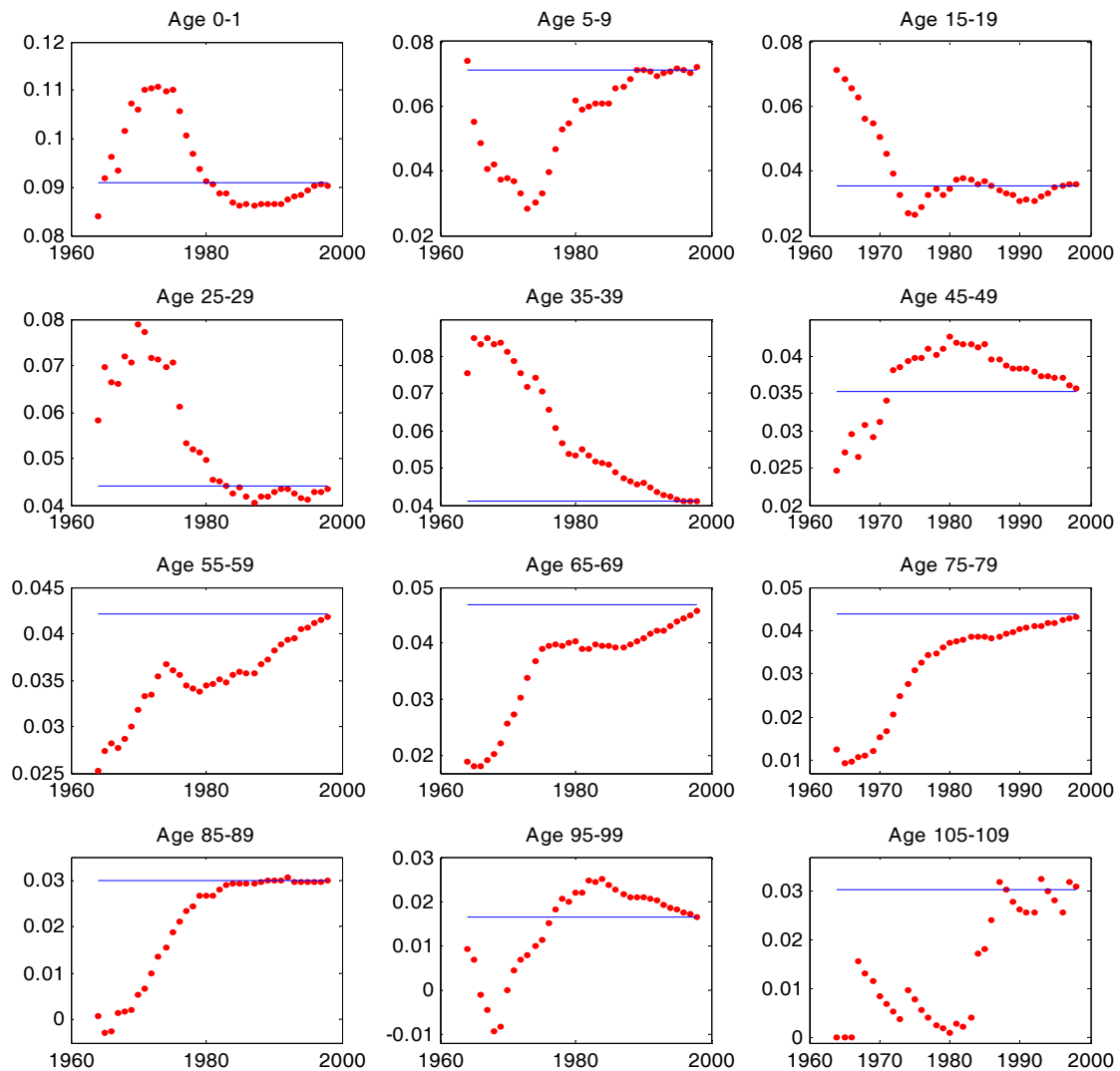Figure 1: β_x parameter, Bulgaria (1955-97), Hungary (1955-99), Lithuania (1960-2001)

and Russia (1970-99), Females (Data source: Human Mortality Database,

www.mortality.org)

LC b(x) Parameter over study period, Finland women.

Figure 2: $\beta(x,s)$ parameter by observation period, Finland women, selected ages.

The years, in abscises, are the end of the observation periods.

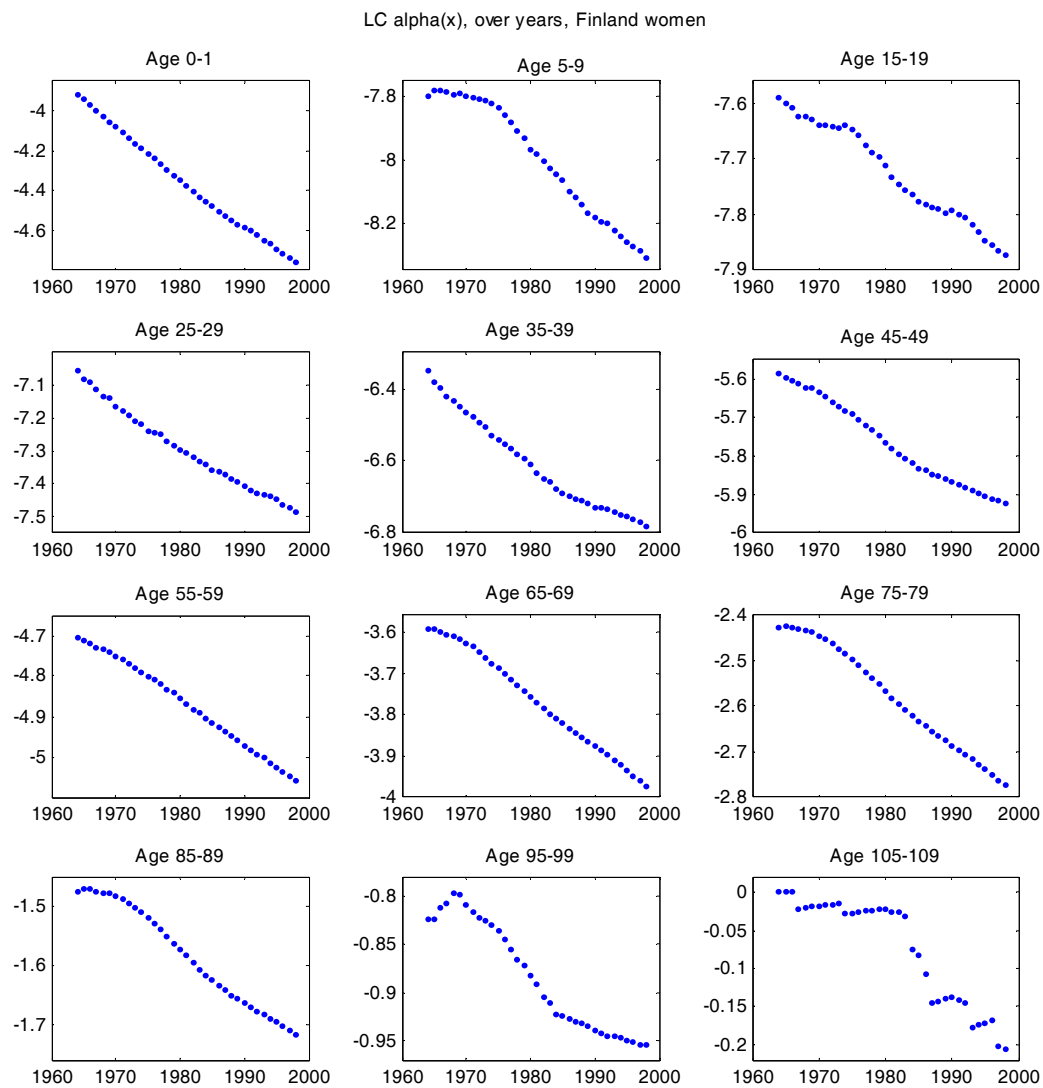The straight line represents the $\beta(x,s)$ value for entire period 1955-1999.

Figure 3: $\alpha(x, s)$ parameter by observation period, Finland women, selected ages.

The years, in abscises, are the end of the observation periods.
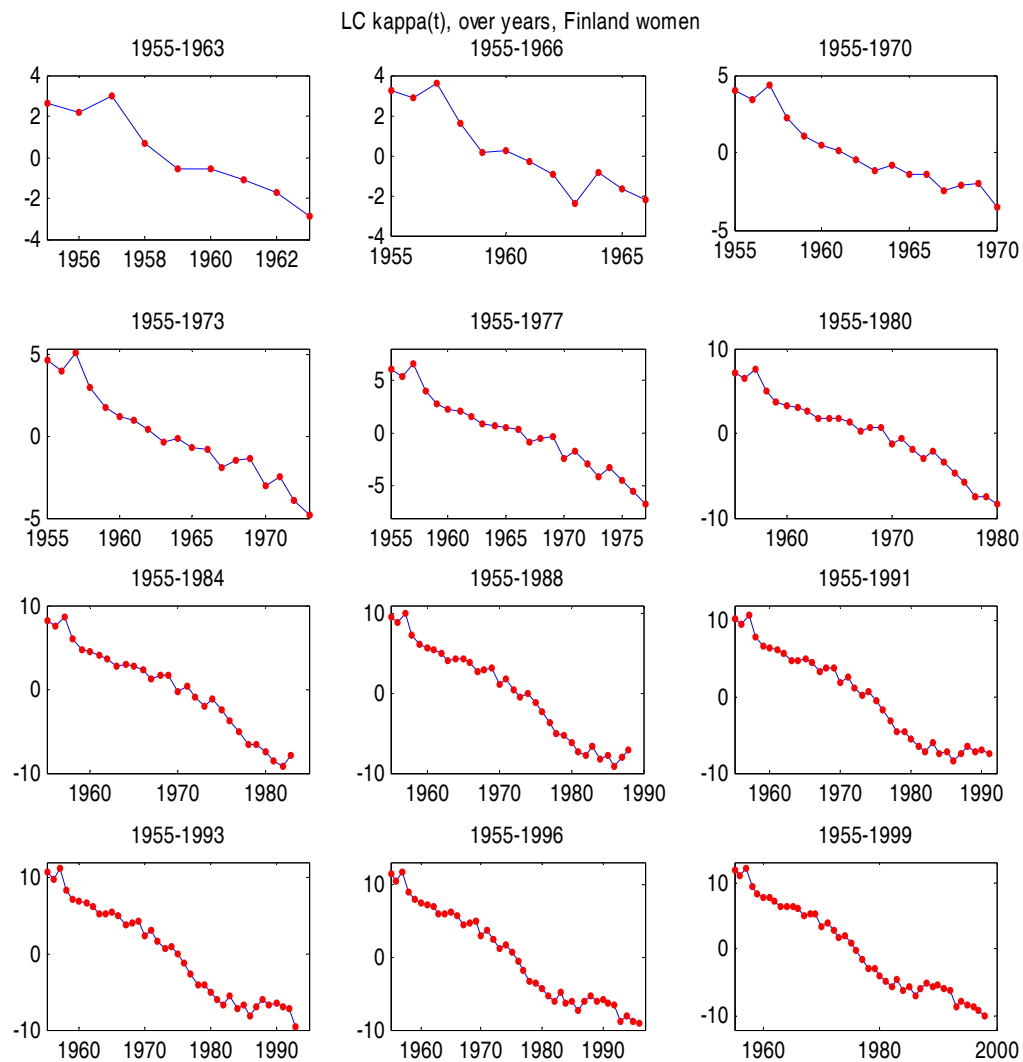
Figure 4: $\kappa(t(s))$ parameter by observation period, Finland women, selected periods.

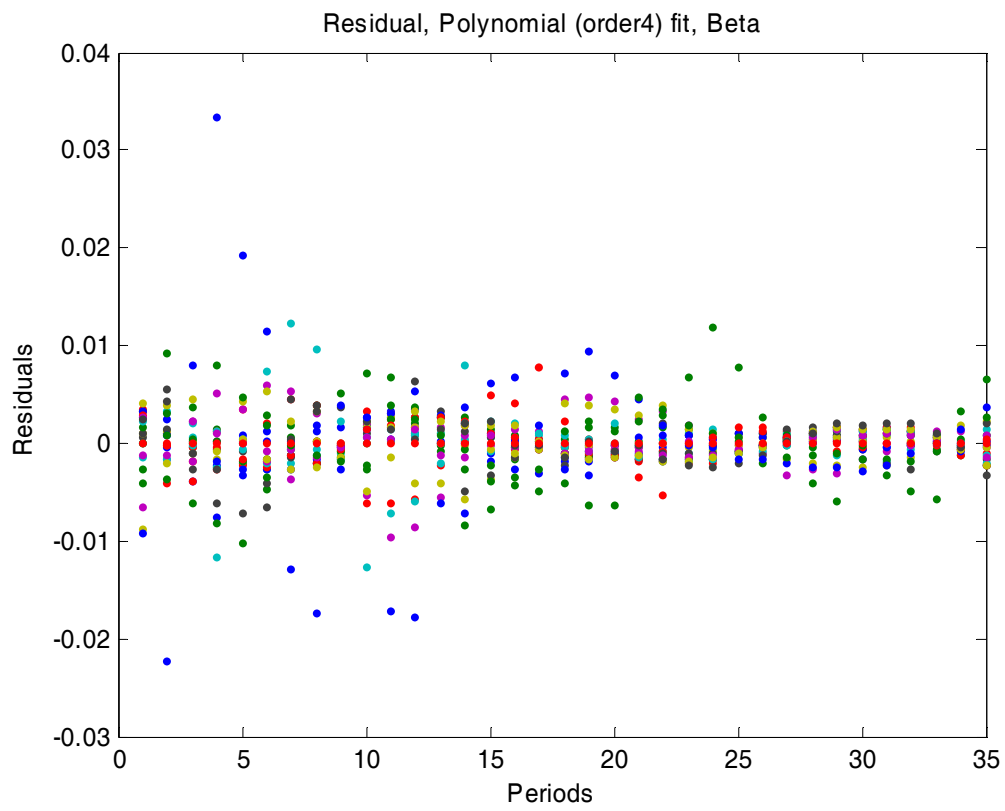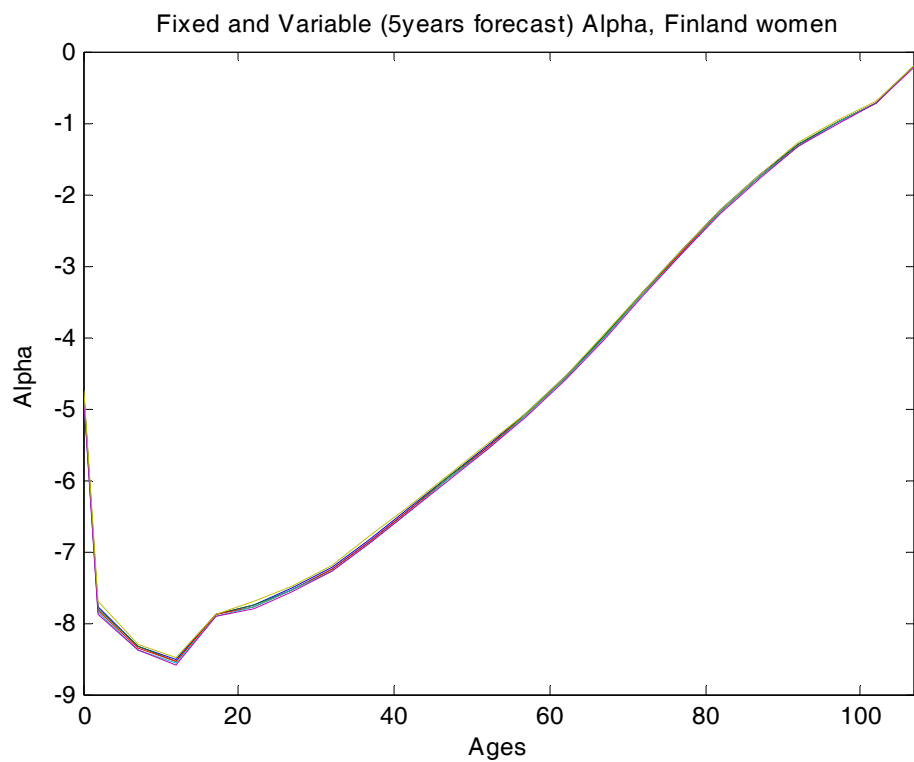The years, in abscises, are the end of the observation periods.

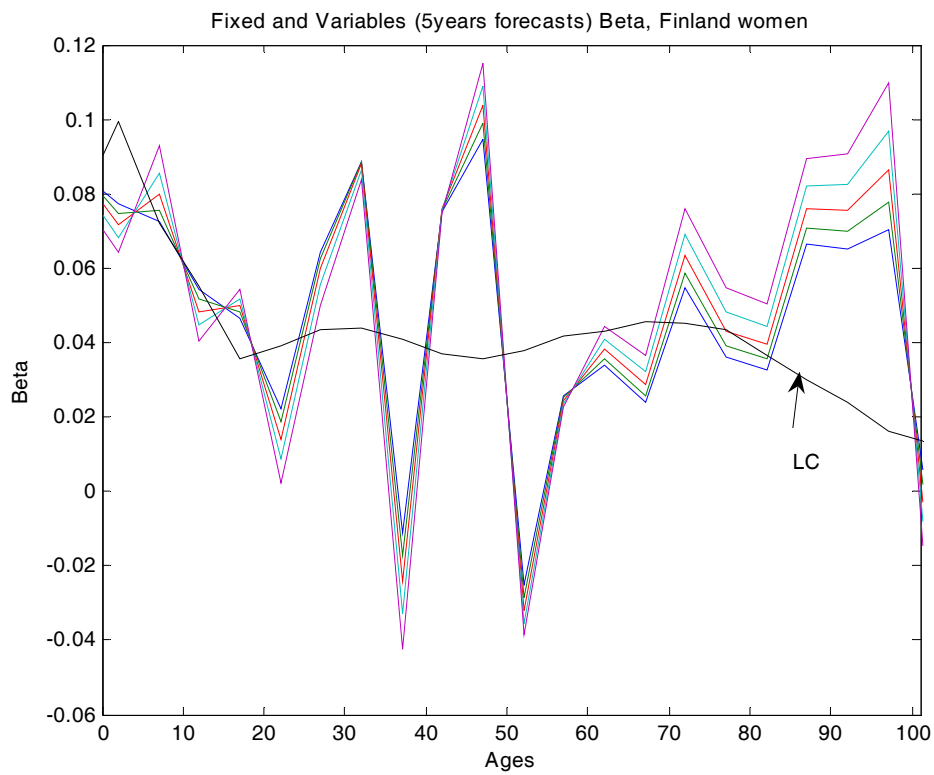Figure 5: Residuals, Polynomial Interpolation of $\beta(x,s)$, Finland Female, 1955-1999.

(a)

(b)

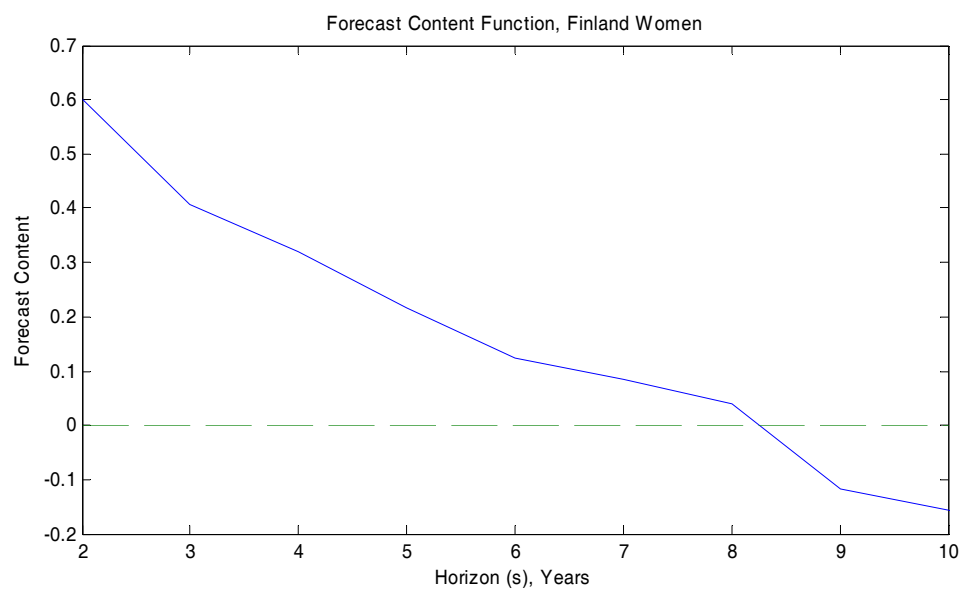Figure 6: Comparison between fix and variable LC age-specific parameters, (a) α (b) β.

Figure 7: Forecast Content Function of LC Kappa Parameter, Finland Women.

Table 1

Coefficients in $\alpha(x,s)$ and $\beta(x,s)$ Modeling

	alpha		beta					
Age	alpha0	aplha1	beta4x(^-6)	beta3x(^-4)	beta2	beta1	beta0	sum(beta)
0-1	-3.9107	-0.0250	-0.4026	0.3647	-0.0011	0.0114	0.0708	0.0811
1-4	-6.7649	-0.0279	-0.1430	0.1153	-0.0003	0.0011	0.1286	0.1295
5-9	-7.6812	-0.0177	0.5668	-0.5030	0.0015	-0.0156	0.0849	0.0707
10-14	-7.9833	-0.0149	0.0457	-0.1107	0.0006	-0.0112	0.1197	0.1091
15-19	-7.5730	-0.0084	0.1634	-0.1623	0.0006	-0.0093	0.0866	0.0778
20-24	-7.3336	-0.0116	-0.2047	0.1456	-0.0003	-0.0003	0.0865	0.0859
25-29	-7.0756	-0.0122	-0.4900	0.4111	-0.0011	0.0096	0.0493	0.0578
30-34	-6.7572	-0.0133	-0.3921	0.3356	-0.0009	0.0080	0.0597	0.0667
35-39	-6.3827	-0.0125	-0.3310	0.2633	-0.0007	0.0041	0.0760	0.0795
40-44	-5.9867	-0.0113	-0.1964	0.1730	-0.0005	0.0061	0.0263	0.0319
45-49	-5.5725	-0.0107	0.0917	-0.0555	0.0001	0.0013	0.0233	0.0246
50-54	-5.1417	-0.0106	0.0123	-0.0009	-0.0001	0.0019	0.0196	0.0214
55-59	-4.6789	-0.0108	-0.0640	0.0601	-0.0002	0.0025	0.0219	0.0243
60-64	-4.1488	-0.0110	0.0867	-0.0433	0.0000	0.0023	0.0075	0.0098
65-69	-3.5549	-0.0119	0.1562	-0.0928	0.0001	0.0014	0.0144	0.0158
70-74	-2.9563	-0.0118	0.1931	-0.1240	0.0002	0.0013	0.0031	0.0046
75-79	-2.3773	-0.0113	0.2735	-0.1956	0.0004	-0.0011	0.0103	0.0096
80-84	-1.8873	-0.0095	0.2659	-0.1926	0.0004	-0.0010	0.0023	0.0017
85-89	-1.4355	-0.0082	0.2387	-0.1763	0.0004	-0.0008	-0.0015	-0.0019
90-94	-1.0557	-0.0070	0.2792	-0.2090	0.0005	-0.0024	0.0096	0.0076
95-99	-0.7881	-0.0053	0.5116	-0.3964	0.0010	-0.0069	0.0129	0.0069
100-104	-0.5636	-0.0043	-0.1324	0.0325	0.0002	-0.0080	0.0932	0.0854
105-109	0.0379	-0.0063	-0.5287	0.3613	-0.0008	0.0056	-0.0048	0.0001

Source: Authors calculation based on data from Human Mortality Database (2004)

REFERENCES:

Alho, J., 2000. Discussion of Lee R., 2000. North American Actuarial Journal. 4, 91-93.

Alho, J. and Spencer, B. 1985. Uncertain Population Forecasting. *Journal of the American Statistical Association*, 80, 306-314.

Barlow, R.J. 1989. Statistics: a Guide to the Use of Statistical Methods in the Physical Sciences. John Wiley and Sons Ltd.

Bell, W. R., 1997. Comparing and Assessing Time Series Methods for Forecasting Age Specific Demographic Rates. *Journal of Official Statistics*, 13:279-303.

Booth, H., Maindonald, J., Smith, L., 2002. Applying Lee-Carter Under Conditions of Variable Mortality Decline. *Population Studies*, 56, 325-336.

Brillinger, D. R., 1986. The Natural Variability of Vital Rates and Associated Statistics. *Biometrics* 42(4):693-734.

Brockwell, P., Davis, R. 1996. Introduction to Time Series and Forecasting. Springer Texts in Statistics.

Brouhns, N. Denuit, M., Van Keilegom, I., 2005. Bootstrapping the Poisson Log-bilinear Model for Mortality Forecasting. *Scandinavian Actuarial Journal* 3: 212-224.

Brouhns, N., Denuit, M., Vermunt, J., 2002. A Poisson Log-linear Regression Approach to the Construction of Projected Life Tables. *Insurance: Mathematics and Economics* 31, 373-393.

Carter, L.R. and Prskawetz, A. 2001. Examining Structural Shifts in Mortality Using the Lee-Carter Method. Working Paper WP2001-007, Max Planck Institute for Demographic Research, Rostock, Germany.

Czado, C., Delwarde, A. and Denuit, D. 2005. Bayesian Poisson Log-bilinear Mortality Projections. *Insurance: Mathematics and Economics* 36(3), 260-284.

Diebold, F., Kilian, L. 2001. Measuring Predictability: Theory and Macroeconomic

Applications. *Journal of Applied Econometrics*, 16: 657-669.

Eckart, C. and Young, G. 1936. The Approximation of One Matrix by Another of Lower Rank. *Psychometrika* 1:211-218.

Eeuwijk van, F. 1995. Multiplicative Interaction in Generalized Linear models. *Biometrics*, 51(3): 1017-1032.

Fuller, W., Hasza, D. 1981. Properties of Predictors for Autoregressive Time series. *Journal of the American Statistical Association*. 76 (373): 155-161.

Galbraith, J.W. 2003. Content Horizons for Univariate Time-Series Forecasts. *International Journal of Forecasting*, 19: 23-55.

Gompertz, B. 1825. On the Nature of the Function Expressive of the Law of Human Mortality and on New Mode of Determining the Value of Life Contingencies. *Philosophical Transactions of the Royal Society*, 115:513-585.

Gutterman, S., Vanderhoof, I. 1999. Forecasting Changes in Mortality: a Search for a Law of causes and Effects. *North American Actuarial Journal*, 2 (4): 135-138.

Helligman, L. and Pollard, J. 1980. The Age Pattern of Mortality. *Journal of the Institute of Actuaries*, 107: 49-75.

Horton, R. L. 1978. The General Linear Model, Data Analysis in the Social and Behavioural Sciences. McGraw-Hill, USA.

Koissi, M-C.; Shapiro, A. 2006. Fuzzy Formulation of the Lee-Carter Model for Mortality Forecasting. *Insurance: Mathematics and Economics,* 39:287-309

Koissi, M-C.; Shapiro, A.; Högnäs, G. 2006. Evaluating and Extending the Lee-Carter Model for Mortality Forecasting: Bootstrap Confidence Interval. *Insurance: Mathematics and Economics*, 38:1-20

Lawson, C., Hanson, R., 1974. Solving Least Squares Problems. Prentice-Hall, EngleWood Cliffs, N.J., 1974.

Lee, R. D., 2000. The Lee-Carter Method for Forecasting Mortality, with Various

   Extensions and Applications. *North American Actuarial Journal*, (4) 1, 80-91.

Lee, R. D., Carter, L. R., 1992. Modeling and Forecasting U.S. Mortality. *Journal of

   the American Statistical Association*. (87) 419, 659-675.

Li, N., Lee, R., and Tuljapurkar, S., 2004. Using the Lee-Carter Method to Forecast

   Mortality for Population with Limited Data. *International Statistical Review* 72:19-36.

Loken, E. 2005. Identification Constraints and Inference in Factor Models. *Structural

   Equation Modeling: A Multidisciplinary Journal*, 12 (2): 232-244

Lundström, H., Qvist, J., 2004. Mortality Forecasting and Trend Shifts: an Application

   of the Lee-Carter Model to Swedish Mortality Data. *International Statistical Review*,

   72 (1), 37-50.

Makeham, W. 1867. On the Law of Mortality. *Journal of the Institute of Actuaries*, 13:

   335-340.

Mandel, J. 1961. Non-Additivity in Two-Way analysis of Variance. *Journal of the

   American Statistical Association*, 56: 878-888.

Murphy, M. 1995. The prospect of mortality: England and Wales and the United States of

   America, 1962-1989. *British Actuarial Journal* 1:331-350.

Oke, T., Öller, L.-E. 1999. Testing for Short Memory in VARMA Process. *Journal of

   Forecasting*, 18:477-487.

Perks, W. 1932. On Some Experiments in the Graduation of Mortality Statistics.

   Journal of the Institute of Actuaries, 63: 12-40.

Pitacco, E. 2004. Survival Models in a Dynamic Context: A Survey. *Insurance:

   Mathematics and Economics*, 35: 279-298.

Renshaw, A., Haberman, S., 2003. Lee-Carter Mortality Forecasting with Age-Specific

    Enhancement. *Insurance: Mathematics and Economics* 33: 255-272.

Shiu, E.S.W. 1986. A Survey of Graduation theory. Proceedings of Symposia in Applied Mathematics 35:67-84. American Mathematical Society.

Stewart, G. 1998. Matrix Algorithms: Basic Decompositions, SIAM, 458 pages

Stoto, M., 1983. Accuracy of Population Projections. *Journal of the American Statistical Association*, 78, 13-20.

Thiele, T. 1872. On a Mathematical Formula to Express the Rate of Mortality Throughout Life. *Journal of the Institute of Actuaries*, 16: 313-329.

Tuljapurkar, S., and Boe, C. 1998. Mortality Change and Forecasting: How Much and How Little Do We Know? *North American Actuarial Journal*, 2(4): 13-47.

Tuljapurkar, S., Nan, L. and Boe, C., 2000. A Universal Pattern of Mortality Decline in the G7 Countries. *Nature*, 405, 789-792.

Verrall, R.J. 1996. A Unified Framework for Graduation. Actuarial Research Paper 91, City University, London.

Wilks, D.S. 1996. Statistical Significance of Long-Range Optimal-Climate-Normal Temperature and Precipitation forecasts. *Journal of Climate*, 9: 827-839.

Wilmoth, J. R.1993. Computational Methods for Fitting and Extrapolating the Lee-Carter Model of Mortality Change. Technical Report, University of California, Berkeley. USA.

Wong-Fupuy, C., Haberman, S. 2004. Projecting Mortality Trends: Recent Developments in the United Kingdom and the United States. *North American Actuarial Journal*, 8(2): 56-83.