Netspar

Network for Studies on Pensions, Aging and Retirement

Netspar THESES

Gintautas Jakstas

Quantifying Longevity Risk

MSc Thesis 2014-036

Quantifying Longevity Risk

By Gintautas Jakstas, ANR: 335567

A thesis submitted in partial fulfillment of the requirements for the degree of

Master of Science in Quantitative Finance and Actuarial Science

Tilburg School of Economics and Management

Tilburg University

Supervisor: prof. dr. Bertrand Melenberg

August 2014

Abstract

Longevity risk is the risk that people on average live longer than expected; it is the negative effect of changes in mortality rates on the risk measures of companies involved in pensions. According to Eurostat, in the Baltic countries: Lithuania, Latvia, and Estonia, the degree of population (65+) living at risk of poverty or social exclusion is above the European Union average, which implies that changes in the pension system have to be done. One of the most important parts in designing pension system is the retirement age which is based on life expectancy. In this Thesis the life expectancy analysis and predictions are performed to the Baltic countries: Lithuania, Latvia, and Estonia. The Lee-Carter and the Cairns-Blake-Dowd models are used to derive projections. We found that Lithuania and Latvia show very similar patterns of mortality, except, it seems that Lithuania faces more difficulties on the way to male mortality reduction. Projected life expectancy in Estonia is significantly higher than in the other two Baltic countries.

1. Table of Contents

1. Introduction

Life insurance companies and pension funds have been using projected life tables for decades. But the problem concerning actuaries is that people have been living much longer than they were expected to according to the life tables being used for actuarial computations. Such potential risk attached to higher than predicted life expectancy of pensioners and policy holders, which can transform in higher than expected pay-out-ratios for many pension funds and insurance companies, is called longevity risk. Traditionally, longevity risk is viewed as non-hedgeable, a market for mortality-linked derivatives is developing. In order to reduce longevity risk, actuaries are trying to develop better models for mortality improvement by being more aware of levels of uncertainty involved in the forecasts.

Since pension funds and life insurance companies are interested in life annuity premiums and reserves, their main concern is long-term projections of mortality indicators. According to Pitacco *et al.* (2009), the Lee-Carter (LC) model introduced by Lee-Carter (1992) is not intended to produce accurate predictions of death rates for a particular age but it is originally used for long-term forecasts. In the recent past, the LC have become more and more popular and have been applied for life expectancy and age specific mortality rates predictions for many countries and time periods. In this Thesis the LC model is fitted to historical data of the Baltic countries Lithuania, Latvia, and Estonia. Box-Jenkins and ARIMA methods are used to model and project a time varying parameter obtained from the LC model. Future death rates are derived from prediction of mortality level.

Additionally to the LC model, the Cairns-Blake-Dowd (CBD), proposed by Cairns *et al.* (2006), for mortality forecasting is fitted to historical data of the Baltic countries. Estimated parameters of the CBD model are then projected using a bivariate time series model. From this forecast of the future intercept and the slope parameters, the future one-year death probabilities are computed in combination with the linear age effect.

Both the LC and the CBD models are based on statistical modelling procedures for the time series which allow to diminish the role of subjective judgment. However, some decisions have to be made, such as how many past years to include in the time series or which ages to use to fit models.

Estimations, performed by the LC and the CBD models, are beyond the original observation data, which is called extrapolation. Thus, the main assumption is that information about the future is contained in the past and projections will not capture explosive increase in life expectancy caused by discovery of medical science and will not capture sudden decrease caused by some new epidemics or diseases.

Based on the fact that the retirement age in Lithuania is increasing gradually every year until the year 2026, we estimate expected life time in the retirement to assess how this change will affect people retiring in this period.

As the first finding in this Thesis we indicate that the LC and the CBD models produce similar forecasts of the life expectancy in the Baltic countries. This finding makes the forecasts more reliable since different methods generates similar outcome. We expected to find that life expectancy patterns in the Baltic countries are very similar, but we found that life expectancy in Estonia is increasing much faster than in the other two countries. We also found that life expectancy for the Lithuanian males is increasing very slowly, which may be the result of the very high emigration from Lithuania, since very often emigrants are the working-age men. Finding the reasons why these differences in three similar countries exists, could be an interesting topic for the future research. During the investigation of the effect of the changes in the retirement age in Lithuania, we found that, according to our estimates, government is increasing retirement age too fast and it causes negative effects for retirees. This finding could be important for Lithuanian government, which could continue this investigation and consider how to compensate for the loosing cohorts.

To perform statistical computing and graphics free open code software R was chosen. In the appendix we present main functions used in this Thesis. In addition to our own code, we have found the *demography* package (Hyndman, 2014) very beneficiary. This package contains various demographic analyses and it has implemented the LC model by single value decomposition (SVD) method.

The paper is organized as follows. Firstly we describe the data which are used in this Thesis, which includes brief description of the database and the main variables which are used in the following chapters. Next, in the Sections 4 and 5 we describe, give estimation algorithms and fit the Lee-Carter and the Crains-Blake-Dowd models to the mortality data of the Baltic countries. In order to avoid irregularities in the life tables, in the Section 6 we apply smoothing method to the age specific patterns of mortality change in the Lee-Carter model. Since we work

in a regression framework, in Section 7 we assess models' performance on terms of the randomness of the residuals. Later in the paper, we estimate mortality projections and prediction intervals. In section 10 we estimate and present the life expectancies for Lithuania, Latvia and Estonia. We finish he paper by providing conclusions and giving recommendations for the related future researches.

2. Data description

The data source used for the studies made in this paper is The Human Mortality Database (HMD) (Wilmoth, et al., 2000).

The Human Mortality Database began in the year 2000 and was launched in May 2002 after its first phase of development. HMD received financial and logistical support from its two sponsoring institutions — the Department of Demography at the University of California, Berkeley in US and the Max Planck Institute for Demographic Research in Rostock, Germany. It also has financial support from the National Institute on Aging, USA, and received technical advice and assistance from many other international collaborators.

The database provides detailed mortality and population data to students, researchers, policy analysts, journalists, and others interested in the historical data of human longevity and this database is free of charge. Currently, it contains detailed data of 37 countries.

The information provided by HMD is standardized and includes the following types of data:

- Births (annual counts of live births by sex);
- Deaths (death counts);
- Population size (annual estimates of population size on January 1st);
- Exposure-to-risk (estimates of the population exposed to the risk of death during some age-time interval);
- Death rates (death rates are always a ratio of the death count for a given age-time interval divided by an estimate of the exposure-to-risk in the same interval);
- Life tables (HMD files are organized by age, sex and time. Size of population is given for one-year and five-year age groups).

In this paper, the data we will use is the death rate period and the population size is one-year age groups.

In the HMD, the mortality series for Lithuania begins in 1959. For the years 1959-1988 data source is Statistical Office of the Lithuanian Soviet Socialist Republic (SSR) and for the later years 1999-2011 – Statistics Lithuania. HMD warns users that the quality of the data for 1959-1979 is lower than in later years and these years should be used with caution.

3. Mortality rates

Usually force of mortality ($\mu_{x,t}$) at age x in year *t* is defined as follows:

$$\mu_{x,t} = \lim_{\Delta x \searrow 0} \frac{\mathbb{P}\left[x < T_{0,(t-x)} \leq x + \Delta x | T_{0,(t-x)} > x\right]}{\Delta x} \qquad 3.1$$

Here $T_{0,(t-x)}$ is the remaining life time of individual born at time $t - x$. This individual will die at age $x + T_{0,(t-x)}$ in year $t + T_{0,(t-x)}$.

Force of mortality ($\mu_{x,t}$) estimation is based on the assumption:

$$\mu_{x+\xi_1,t+\xi_2} = \mu_{x,t}, \quad 0 \leq \xi_1, \xi_2 < 1 \qquad 3.2$$

Formula (3.2) can be best illustrated by a Lexis diagram with calendar time as abscissa and age as coordinate, for more details about the Lexis diagram, see Carstensen (2007). Assumption (3.2) assumes that the mortality rate is constant within each square, but allows varying between squares, which basically means that mortality rates are constant for a calendar year. Maximum likelihood estimator of mortality rates is obtained by dividing the number of deaths recorded at age x in year *t* ($D_{x,t}$) by the corresponding exposure to risk ($N_{x,t}$) which is the estimates of the age x population exposed to the risk of death in year *t*.

$$\mu_{x,t} = \frac{D_{x,t}}{N_{x,t}} \qquad 3.3$$

The forces of mortality $(\mu_{x,t})$ and the central death rates $(m_{x,t})$ coincide under assumption 3.2), thus $m_{x,t} = \mu_{x,t}$.

To get an idea about how the mortality rates develop over time, mortality rates for different age, male and female, are presented in graphs (Figure 3.1, Figure 3.2, Figure 3.3, and Figure 3.4).

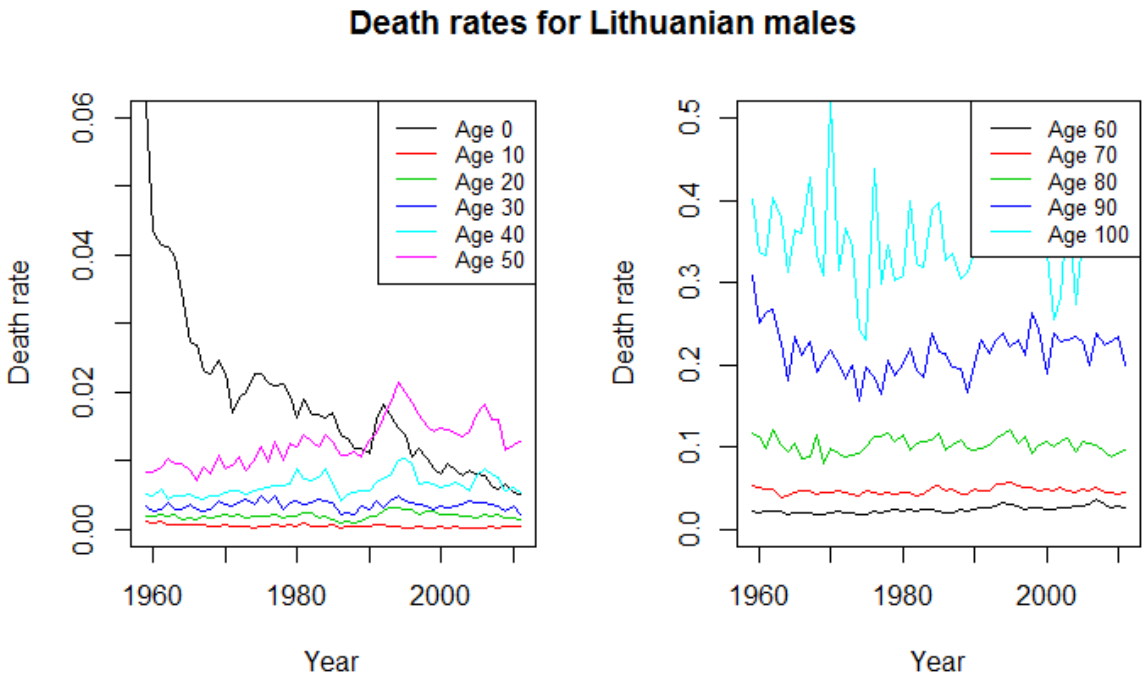

Figure 3.1 Death rates for Lithuanian males. On the left-hand side age 0-50, on the right-hand side age 60-100

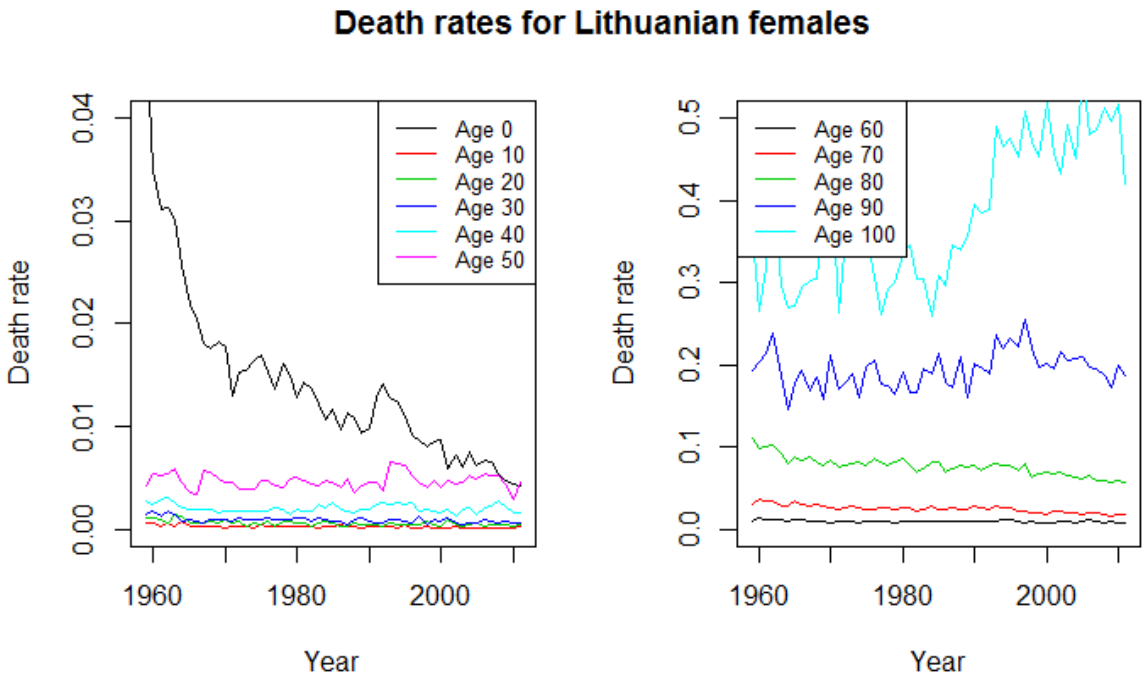

Figure 3.2 Death rates for Lithuanian females. On the left-hand side age 0-50, on the right-hand side age 60-100

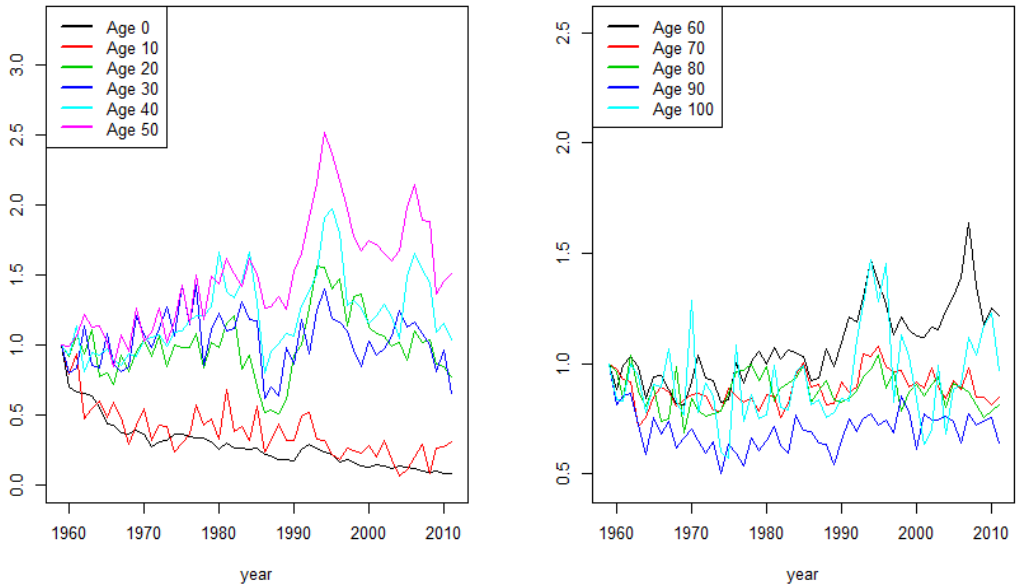

Figure 3.3 Relative changes in death rates for Lithuanian males (compared to the starting year). On the left-hand side age 0-50, on the right-hand side age 60-100

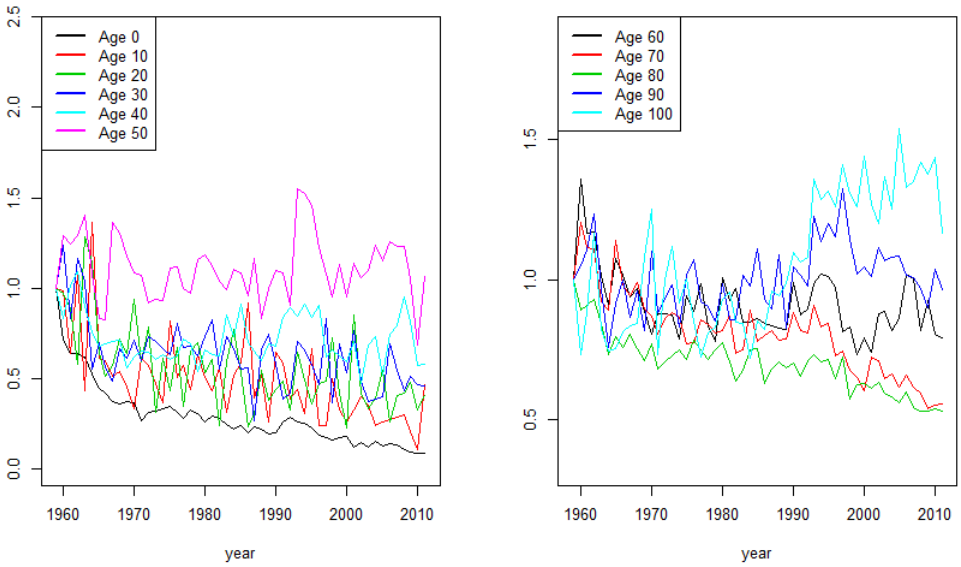

Figure 3.4 Relative changes in death rates for Lithuanian females (compared to the starting year). On the left-hand side age 0-50, on the right-hand side age 60-100

In general, female mortality rates are lower than male. Exception is for citizens of 100 years of age, but it is only because of a too small population size of 100 years old people. The most significant changes of mortality rates (in absolute scale) in time are for infants. The main reasons why infants mortality has decreased are improved sanitation and medicine, most of

child births are given in maternal hospitals. But still death rates for the first year after birth are relatively high and decrease rapidly to the lowest point around age 10. There is a slight decrease of mortality rates for female around ages 20-30 which reflects decline in childbearing mortality.

In the Figure 3.3 and the Figure 3.4 we can observe relative changes in mortality rates. We can see that for older ages (90, 100), variation in mortality rates is much bigger compared to younger. We can also see that mortality rates for many age groups of Lithuanian males are not decreasing, only for the very young ages is an obvious decrease. In Figure 3.4 we can observe, that mortality rates for women of age 50 are not decreasing, which may be related to the high numbers of deaths from cancer.

In the following pictures (Figure 3.5), the pattern of logarithmic death rates according to age are presented for male, female, and the total population, respectively.

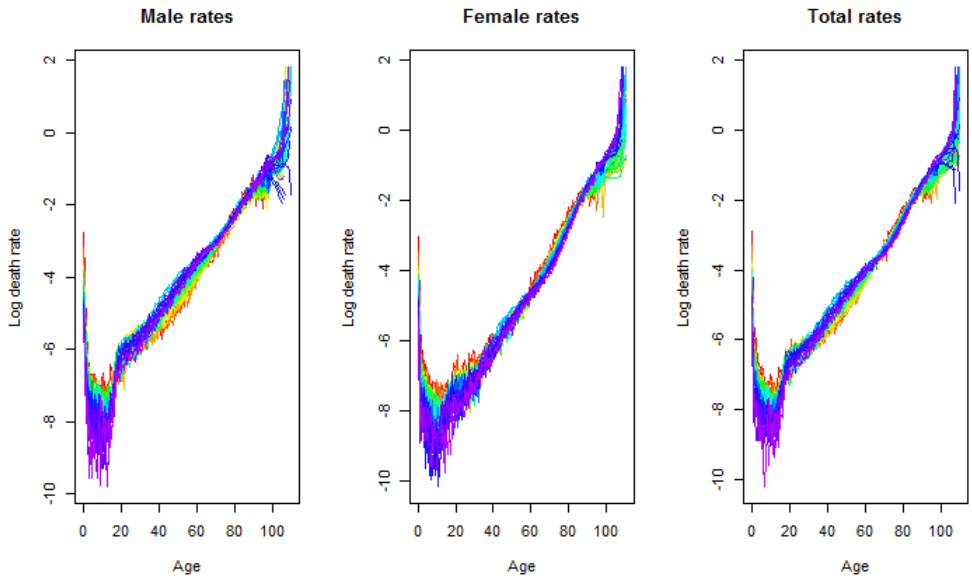

Figure 3.5 Logarithm of death rates for Lithuanian males, females and total population

Logarithmic death rates are falling as kids are growing (at age range 0-14), then for teenagers death rates are more or less the same (age group 14-18) and for adults logarithmic death rate is rising nearly linearly. At age 18 there is a significant increase in death rates (especially for males), mostly because of young adults acquire new rights (driving license, right to use alcohol) and are less protected by parents.

In Figure 3.6 logarithm of death rates are presented, we can see that Lithuanian mortality is falling rapidly for young ages, but almost no changes for older people (80 years old and more).

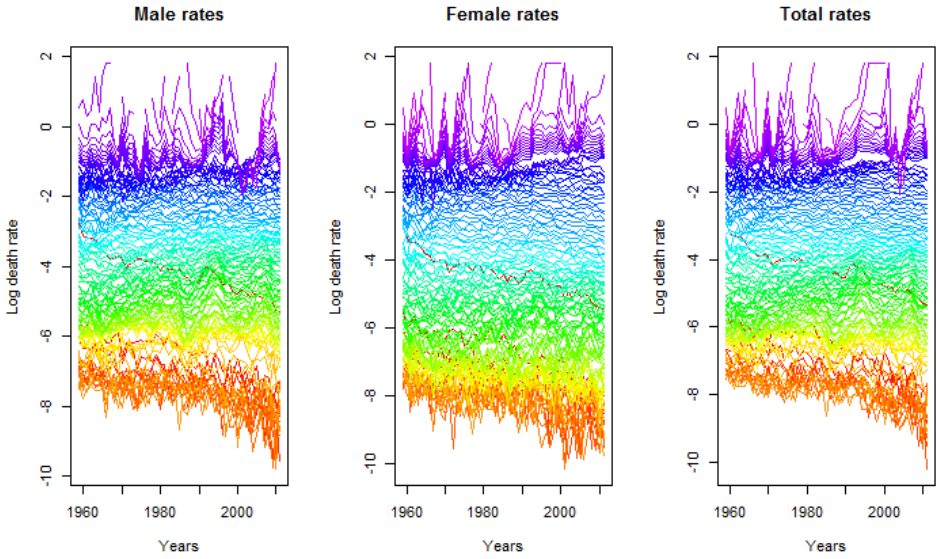

Figure 3.6 Logarithm of death rates for Lithuanian males, females and total population, time component for all ages

4. Lee-Carter (LC) model.

Lee and Carter (1992) proposed a model for prediction of the time trends and age patterns in mortality. They specified a log-bilinear model for mortality rate $m_{x,t}$:

$$\ln(m_{x,t}) = \alpha_x + \beta_x k_t + \varepsilon_{x,t} \qquad 4.1$$

Here $\varepsilon_{x,t}$ – homoskedastic centered residual term at age x and time t, with mean 0 and variance σ^2. α_x – the age-specific pattern of mortality averaged over time; β_x – age specific patterns of mortality change; k_t – a time-trend index of general mortality level. $\mathrm{Exp}(\alpha_x)$ is the general shape of the mortality scheme while actual mortality rate changes in accordance with mortality index k_t, adjusted by the age response β_x.

The model expressed by 4.1 cannot be fitted by simple regression, since α_x, β_x and k_t are not observable variables. Least squares estimates can be found by using singular value decomposition. The parameter estimation is based on a matrix of available mortality rates.

The LC model (4.1) suffers from identifiability problem. To see that, we use revised parametrisation:

$$\ln(m_{x,t}) = \tilde{\alpha}_x + \tilde{\beta}_x \tilde{k}_t$$
$$\text{where } \tilde{\alpha}_x = \alpha_x + b\beta_x \qquad 4.2$$
$$\tilde{\beta}_x = \frac{\beta_x}{a}$$
$$\tilde{k}_t = a(k_t - b)$$

Thus, $\ln(m_{x,t}) = \alpha_x + b\beta_x + \frac{\beta_x}{a} a(k_t - b)$ is identical to 4.1 with any values of a and b. To avoid this problem we need to impose two constrains on the parameters: $\sum_t k_t = 0$ and $\sum_x \beta_x = 1$.

Let's assume, that, the sample range for the time component is $t = t_1, \dots, t_n$ and for the age component – $x = x_1, \dots, x_m$.

$$\sum_{t=t_1}^{t_n} \ln(m_{x,t}) = \sum_{t=t_1}^{t_n} (\alpha_x + \beta_x k_t + \varepsilon_{x,t}) = (t_n - t_1 + 1)\alpha_x + \beta_x \sum_{t=t_1}^{t_n} k_t + \sum_{t=t_1}^{t_n} \varepsilon_{x,t} \qquad 4.3$$

Since we assumed that the residual term average is 0, it vanishes. Due to the constraint for k_t: $\beta_x \sum_{t=t_1}^{t_n} k_t = 0$ and

$$\sum_{t=t_1}^{t_n} \ln(m_{x,t}) = (t_n - t_1 + 1)\alpha_x. \qquad 4.4$$

To the estimate Lee-Carter model, singular value decomposition (SVD) method may be used.

SVD method algorithm:

1) In order to minimize the sum of squares of the residuals, by using SVD method, estimator of the α_x is obtained:

$$\hat{\alpha}_x = \frac{1}{n} \sum_{t=t_1}^{t_n} \ln(m_{x,t}) \qquad 4.5$$

2) Create a matrix $\mathbb{Z} = \{\mathbb{Z}_{x,t}\}_{\substack{x=x_1,\dots,x_m \\ t=t_1,\dots,t_n}}$ for estimating β_x and k_t , here $\mathbb{Z}_{x,t} = \ln(m_{x,t}) - \hat{\alpha}_x = \beta_x k_t$.

3) Apply the Singular Value Decomposition to the matrix $\mathbb{Z}$, which decomposes the matrix of $\mathbb{Z}$ into the product of three matrices : $ULV' = SVD(\mathbb{Z}) = L_1 U_{x1} V_{t1} + \cdots + L_X U_{xX} V_{tX}$, where U sands for the the age component, L – the singular values, and V represents the time component. For more details about SVD method in general see a tutorial by Baker (2005).
4) Approximate a new matrix $\hat{\mathbb{Z}}$ by the product of estimated parameters $\hat{\beta}_x$ and $\hat{k}_t$ to get $\hat{\mathbb{Z}}_{x_1,t_1} = \hat{\beta}_{x_1}\hat{k}_{t_1}$.

$$\hat{\mathbb{Z}} = \begin{bmatrix} \hat{\mathbb{Z}}_{x_1,t_1} & \cdots & \hat{\mathbb{Z}}_{x_1,t_n} \\ \vdots & \ddots & \vdots \\ \hat{\mathbb{Z}}_{x_m,t_1} & \cdots & \hat{\mathbb{Z}}_{x_m,t_n} \end{bmatrix} \quad \begin{matrix} \hat{\beta}_x \\ \\ \\ \hat{k}_t \end{matrix}$$

5) Estimate the logarithm of the death rate, $\ln(\hat{m}_{t,x}) = \hat{\alpha}_x + \hat{\mathbb{Z}} = \hat{\alpha}_x + \hat{\beta}_x\hat{k}_t$

Lee-Carter model estimation by using SVD method is implemented in the R-package *demography* (Hyndman, 2014).

4.1.Estimation by Newton–Raphson method

A stepwise alternative to the SVD method is an iterative Newton-Raphson method which can be applied to find the estimations for the parameters α_x, β_x, and k_t.

Objective function:

$$F(\alpha,\beta,k) = \sum_{x=x_1}^{x_m}\sum_{t=t_1}^{t_n}\left(\ln(m_{x,t}) - \alpha_x - \beta_x k_t\right)^2 \qquad 4.6$$

Partial derivatives:

$$\frac{\partial F(\alpha,\beta,k)}{\partial \alpha_x} = \sum_{t=t_1}^{t_n} 2\left(\ln(m_{x,t}) - \alpha_x - \beta_x k_t\right) = 0, \qquad x = x_1, \dots, x_m \qquad 4.7$$

$$\frac{\partial F(\alpha,\beta,k)}{\partial k_t} = \sum_{x=x}^{x_m} 2\beta_x\left(\ln(m_{x,t}) - \alpha_x - \beta_x k_t\right) = 0, \qquad t = t_1, \dots, t_n \qquad 4.8$$

$$\frac{\partial F(\alpha,\beta,k)}{\partial \beta_x} = \sum_{t=t_1}^{t_n} 2k_t\left(\ln(m_{x,t}) - \alpha_x - \beta_x k_t\right) = 0, \qquad x = x_1, \dots, x_m \qquad 4.9$$

The idea of the Newton–Raphson method is to start by guessing $x^{(0)}$ for a root of the function $f(x)$, than use the approximation of the following estimate $x^{(1)}$:

$$x^{(1)} = x^{(0)} - \frac{f\left(x^{(0)}\right)}{f'(x^{(0)})} \tag{4.10}$$

The procedure is repeated until convergence (the differences between estimates becomes very small) by:

$$x^{(k+1)} = x^{(k)} - \frac{f\left(x^{(k)}\right)}{f'(x^{(k)})} \tag{4.11}$$

Newton-Raphson method for estimating the Lee-Carter model:

Step 1. We need to choose starting values for all parameters, e.g. $\hat{\alpha}_x^{(0)} = 0, \hat{\beta}_x^{(0)} = 1, \hat{k}_t^{(0)} = 0.$

Step 2. Update parameter $\hat{\alpha}_x$:

$$\hat{\alpha}_x^{(k+1)} = \hat{\alpha}_x^{(k)} + \frac{\sum_{t=t_1}^{t_n}\left(\ln\left(m_{x,t}\right) - \hat{\alpha}_x^{(k)} - \hat{\beta}_x^{(k)}\hat{k}_t^{(k)}\right)}{t_n - t_1 + 1} \tag{4.12}$$

$$\hat{\alpha}_x^{(k+1)} := \hat{\alpha}_x^{(k+1)} + \hat{\beta}_x^{(k)} E_t\left(\hat{\beta}_x^{(k)}\right)$$

Step 3. Update parameter $\hat{k}_t$:

$$\hat{k}_t^{(k+1)} = \hat{k}_t^{(k)} + \frac{\sum_{x=x_1}^{x_m}\hat{\beta}_x^{(k)}\left(\ln\left(m_{x,t}\right) - \hat{\alpha}_x^{(k+1)} - \hat{\beta}_x^{(k)}\hat{k}_t^{(k)}\right)}{\sum_{t=t_1}^{t_n}\left(\hat{\beta}_x^{(k)}\right)^2} \tag{4.13}$$

$$\hat{k}_t^{(k+1)} := \left(\hat{k}_t^{(k+1)} - E_t\left(\hat{k}_t^{(k+1)}\right)\right)\sum_x \hat{\beta}_x^{(k)}$$

Step 4. Update parameter $\hat{\beta}_x$:

$$\hat{\beta}_x^{(k+1)} = \hat{\beta}_x^{(k)} + \frac{\sum_{t=t_1}^{t_n}\hat{k}_t^{(k)}\left(\ln\left(m_{x,t}\right) - \hat{\alpha}_x^{(k+1)} - \hat{\beta}_x^{(k)}\hat{k}_t^{(k+1)}\right)}{\sum_{t=t_1}^{t_n}\left(\hat{k}_t^{(k+1)}\right)^2} \tag{4.14}$$

$$\hat{\beta}_x^{(k+1)} := \frac{\hat{\beta}_x^{(k+1)}}{\sum_x \hat{\beta}_x^{(k+1)}}$$

Repeat Steps 2-4 until convergence. In this Thesis we choose to stop procedure when a relative increase in objective function (4.6) is smaller than 10^{-6}.

Here we do not use adjustment of the $\hat{k}_t$'s by refitting to the total observed deaths, since we wanted to do estimations in the case of heaving only mortality rates data available. To exploit additional available data (total number of deaths and exposure to risk) we use the Poisson maximum likelihood estimation method, which is described in the following section.

4.2. Poisson maximum likelihood estimation

Poisson maximum likelihood estimation method better exploits the available information and does not assume that the variability of mortality rates is the same for all age groups which is unrealistic since mortality rates are much more volatile at older ages. In this method, the assumption is that the number of deaths at age x at year t follows a Poisson distribution. It is assumed that the remaining lifetimes of the observed individuals are independent and identically distributed. In this method instead of $m_{x,t}$, both $D_{x,t}$ and $N_{x,t}$ are used.

$$D_{x,t} \sim \text{Poisson}\left(N_{x,t} m_{x,t}\right) \tag{4.15}$$

In order to define the parameter $N_{x,t}m_{x,t}$, (Brouhns, et al., 2002) assume a log-bilinear force of mortality:

$$\ln\left(m_{x,t}\right) = \alpha_x + \beta_x k_t \tag{4.16}$$

Objective function:

$$F(\alpha, \beta, k) = \sum_{x=x_1}^{x_m} \sum_{t=t_1}^{t_n} \left[\left(D_{x,t} \cdot (\alpha_x + \beta_x k_t\right) - N_{x,t} \cdot \exp(\alpha_x + \beta_x k_t)\right] \tag{4.17}$$

Partial derivatives:

$$\frac{\partial F(\alpha, \beta, k)}{\partial \alpha_x} = \sum_{t=t_1}^{t_n} \left[D_{x,t} - N_{x,t} \cdot \exp(\alpha_x + \beta_x k_t)\right] = 0 \tag{4.18}$$

$$\frac{\partial F(\alpha, \beta, k)}{\partial k_t} = \sum_{x=x_1}^{x_m} \left[D_{x,t} - N_{x,t} \cdot \exp(\alpha_x + \beta_x k_t)\right]\beta_x = 0 \tag{4.19}$$

$$\frac{\partial F(\alpha, \beta, k)}{\partial \beta_x} = \sum_{t=t_1}^{t_n} \left[D_{x,t} - N_{x,t} \cdot \exp(\alpha_x + \beta_x k_t)\right]k_t = 0 \tag{4.20}$$

Algorithm to update is the same as it was in previous section:

Step 1. We need to choose starting values for all parameters, e.g. $\hat{\alpha}_x^{(0)} = 0, \hat{\beta}_x^{(0)} = 1, \hat{k}_t^{(0)} = 0$.

Step 2. Update parameter $\hat{\alpha}_x$:

$$\hat{\alpha}_x^{(k+1)} = \hat{\alpha}_x^{(k)} + \frac{\sum_{t=t_1}^{t_n} \left[D_{x,t} - N_{x,t} \cdot \exp\left(\hat{\alpha}_x^{(k)} + \hat{\beta}_x^{(k)} \hat{k}_t^{(k)}\right)\right]}{\sum_{t=t_1}^{t_n} \left[N_{x,t} \cdot \exp\left(\hat{\alpha}_x^{(k)} + \hat{\beta}_x^{(k)} \hat{k}_t^{(k)}\right)\right]}$$

$$\hat{\alpha}_x^{(k+1)} := \hat{\alpha}_x^{(k+1)} + \hat{\beta}_x^{(k)} E_t\left(\hat{k}_t^{(k)}\right)$$

Step 3. Update parameter $\hat{k}_t$:

$$\hat{k}_t^{(k+1)} = \hat{k}_t^{(k)} + \frac{\sum_{x=x_1}^{x_m}\left[D_{x,t} - N_{x,t}\cdot\exp\left(\hat{\alpha}_x^{(k+1)} + \hat{\beta}_x^{(k)}\hat{k}_t^{(k)}\right)\right]\hat{\beta}_x^{(k)}}{\sum_{t=t_1}^{t_n}\left[N_{x,t}\cdot\exp\left(\hat{\alpha}_x^{(k+1)} + \hat{\beta}_x^{(k)}\hat{k}_t^{(k)}\right)\right]\left(\hat{\beta}_x^{(k)}\right)^2}$$

$$\hat{k}_t^{(k+1)} := \left(\hat{k}_t^{(k+1)} - E_t\left(\hat{k}_t^{(k+1)}\right)\right)\sum_x \hat{\beta}_x^{(k)}$$

Step 4. Update parameter $\hat{\beta}_x$:

$$\hat{\beta}_x^{(k+1)} = \hat{\beta}_x^{(k)} + \frac{\sum_{t=t_1}^{t_n}\left[D_{x,t} - N_{x,t}\cdot\exp\left(\hat{\alpha}_x^{(k+1)} + \hat{\beta}_x^{(k)}\hat{k}_t^{(k+1)}\right)\right]\hat{k}_t^{(k+1)}}{\sum_{t=t_1}^{t_n}\left[N_{x,t}\cdot\exp\left(\hat{\alpha}_x^{(k+1)} + \hat{\beta}_x^{(k)}\hat{k}_t^{(k+1)}\right)\right]\left(\hat{k}_t^{(k+1)}\right)^2}$$

$$\hat{\beta}_x^{(k+1)} := \frac{\hat{\beta}_x^{(k+1)}}{\sum_x \hat{\beta}_x^{(k+1)}}$$

Repeat Steps 2-4 until convergence. As in the previous section, we choose to stop procedure when a relative increase in objective function (4.17) is smaller than 10^{-6}.

4.3. Application to Baltic countries

After applying the Lee-Carter model to Lithuanian data we are able to present $\hat{\alpha}_x$, $\hat{\beta}_x$ and $\hat{k}_t$ (Figure 4.1). The Lee-Carter model is applied separately to male, female and the total population, by considering a maximum age equal to 100 and the starting year is 1980. In Figure 4.1 we can see that patterns of the $\hat{\beta}_x$'s looks similar for males and females, while the time-trend index of general mortality level ($\hat{k}_t$'s) for females are decreasing faster for females. We can observe, that the $\hat{\alpha}_x$'s have a similar pattern to logarithmic death rates in Figure 3.5, which is natural since $\hat{\alpha}_x$'s represent the age-specific pattern of mortality averaged over time.

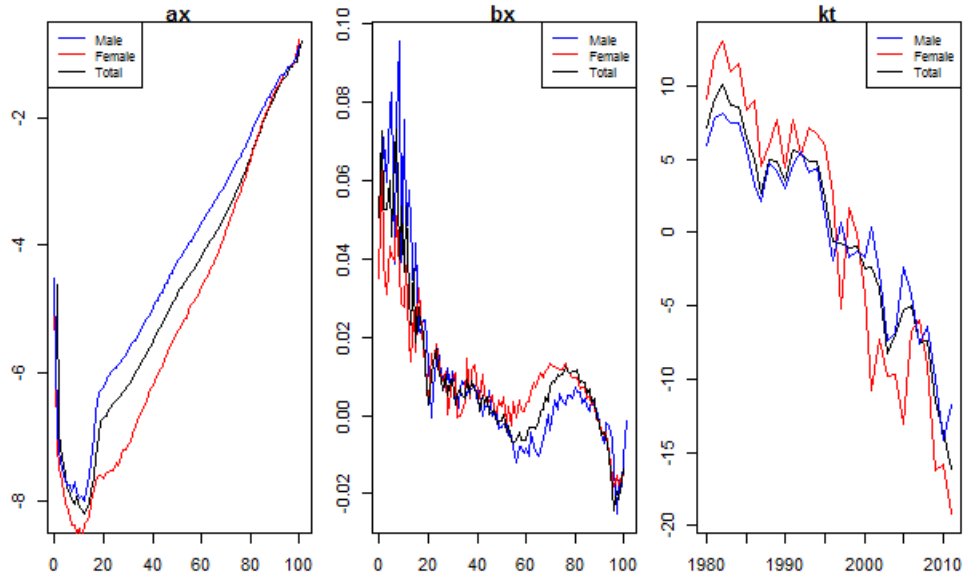

Figure 4.1 The Lee-Carter model fitted for Lithuanian males, females and total population from 1980 to 2011 years and from age 0 to 100. In the graph on the left-hand side $\hat{\alpha}_x$ is presented, in the middle - $\hat{\beta}_x$ and on the right-hand side - $\hat{k}_t$.

In the Figure 4.2, Figure 4.3 and Figure 4.4 we can see graphical comparison of the fitted LC model for Baltic countries (males, females and the total population, respectively).

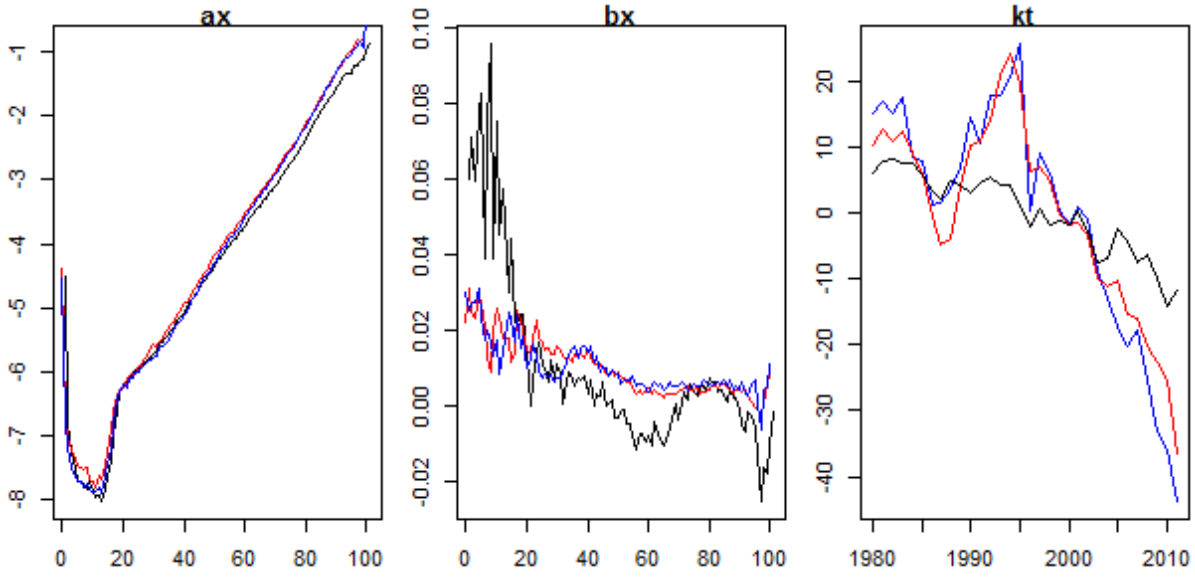

Figure 4.2 The Lee-Carter model fitted for Lithuanian, Latvian and Estonian males, from 1980 to 2011 years and from age 0 to 100. In the graph on the left-hand side $\hat{\alpha}_x$ is presented, in the middle - $\hat{\beta}_x$ and on the right-hand side - $\hat{k}_t$. Blue line stands for Estonia, red line – Latvia, black – Lithuania.

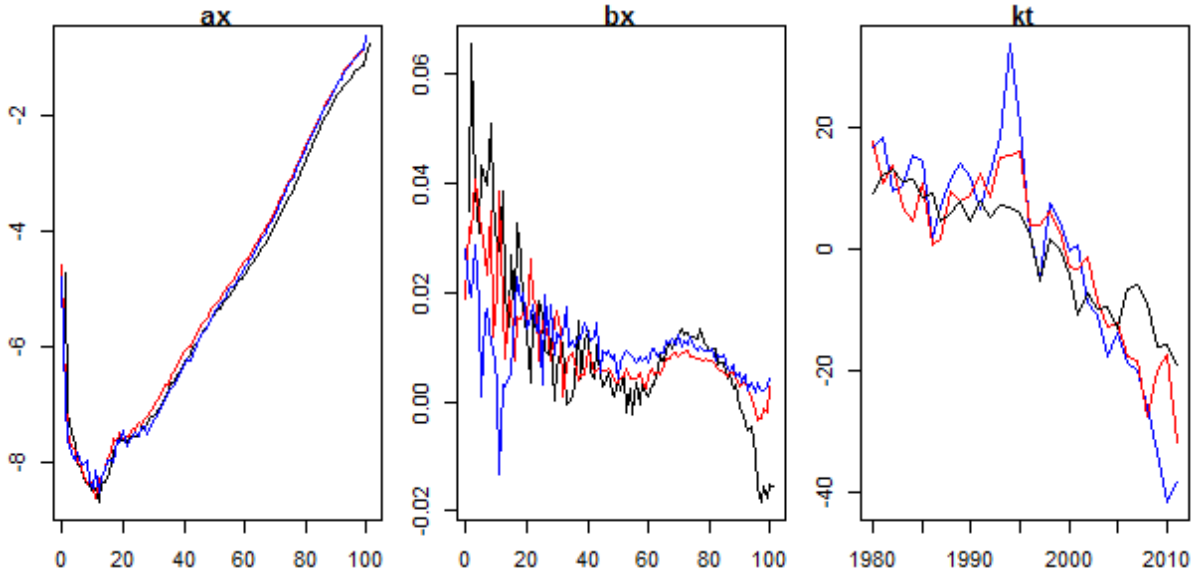

Figure 4.3 The Lee-Carter model fitted for Lithuanian, Latvian and Estonian females, from 1980 to 2011 years and from age 0 to 100. In the graph on the left-hand side $\hat{\alpha}_x$ is presented, in the middle - $\hat{\beta}_x$ and on the right-hand side - $\hat{k}_t$. Blue line stands for Estonia, red line – Latvia, black – Lithuania.

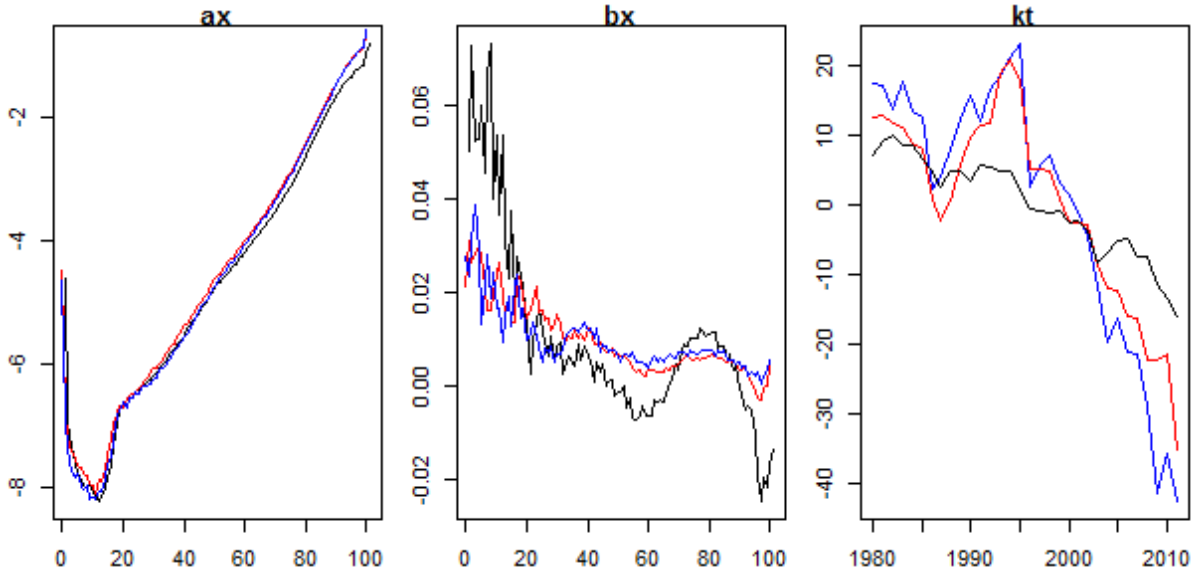

Figure 4.4 The Lee-Carter model fitted for Lithuanian, Latvian and Estonian total population, from 1980 to 2011 years and from age 0 to 100. In the graph on the left-hand side $\hat{\alpha}_x$ is presented, in the middle - $\hat{\beta}_x$ and on the right-hand side - $\hat{k}_t$. Blue line stands for Estonia, red line – Latvia, black – Lithuania.

We can see that general pattern of mortality by age is almost identical to all Baltic countries, while the relative speed of change at each age (β_x) and the index of the level of mortality at time *t* significantly differs between countries. The $\hat{k}_t$'s, when the model is fitted for Lithuanian males, are decreasing much slower compared to Latvia and Estonia. The interpretation behind

it is that mortality rates for Lithuanian male are decreasing much slower than in the other two countries, which may happened because of high level of working-age males emigration from Lithuania. According to Kingdom (2014), high value of $\hat{\beta}_x$'s at young ages and negative values at old ages for Lithuanian males may occur because younger-age mortality showed most improvement in the earlier years while in later years it was older-age mortality which improved the most.

In Figure 4.5 and Figure 4.6 the Lee-Carter model fitted for Lithuanian population by using Newton–Rapshon and single value decomposition methods is presented. By using Newton–Rapshon estimation method, there was implemented an additional constrain for $\hat{\beta}_x$'s ($\hat{\beta}_x \geq 0$). As we can see in these graphs, it has no effect on $\hat{\alpha}_x$'s, however it caused some changes in $\hat{k}_x$'s.

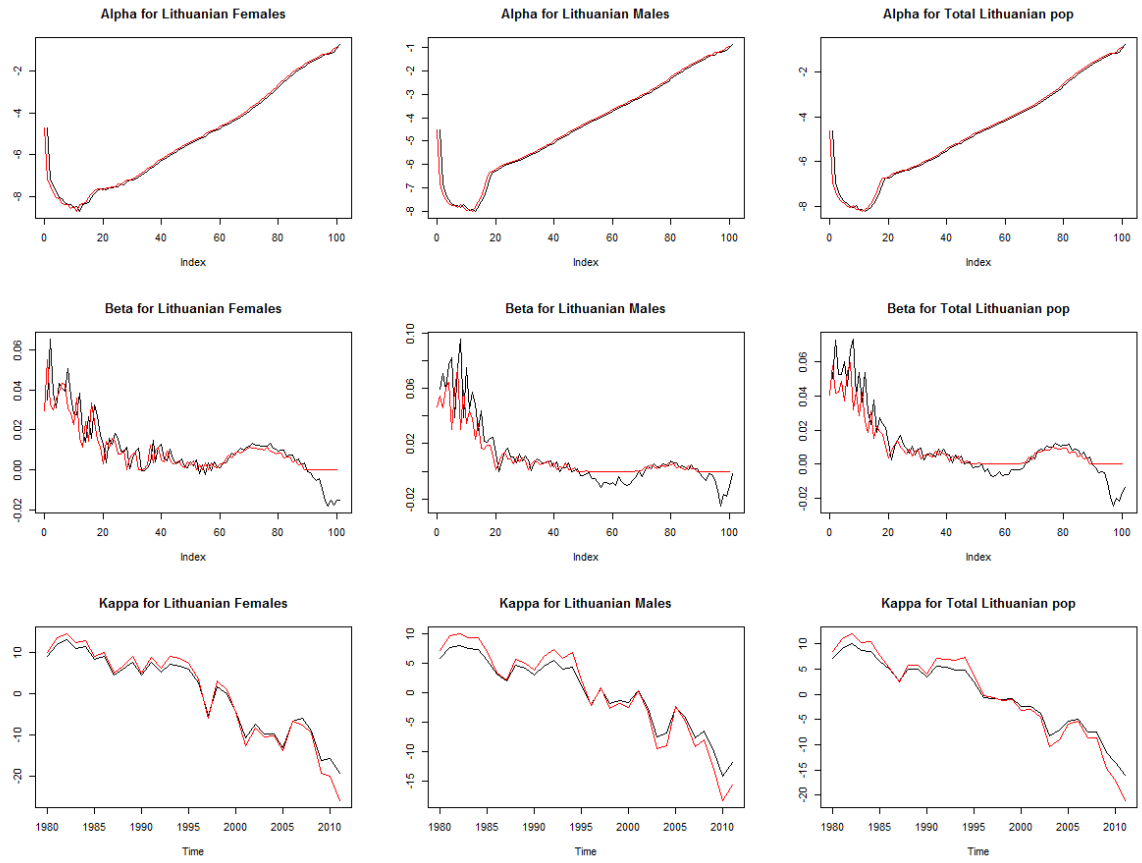

Figure 4.5 The Lee-Carter model fitted for Lithuanian population (from the left: female, male, total population). Red line presents Newton–Rapshon method with minimum $\hat{\beta}_x$ set to 0, black line – SVD method.

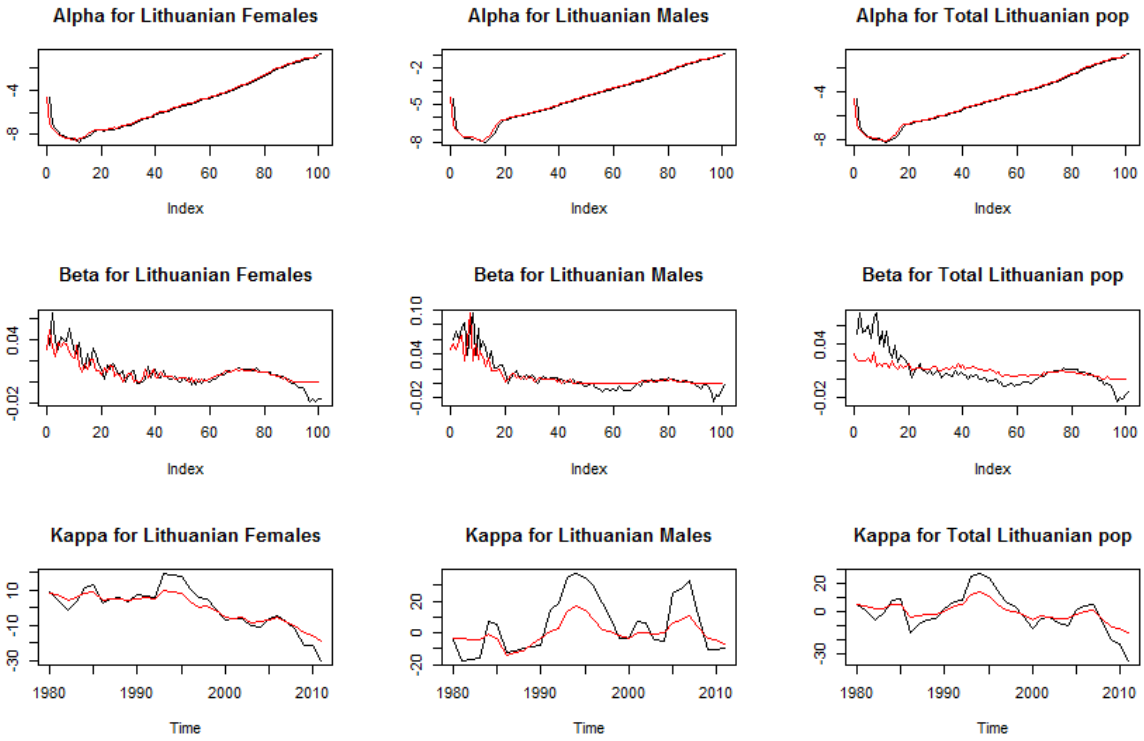

Figure 4.6 The Lee-Carter model, Poisson maximum likelihood estimation, fitted for Lithuanian population (from the left: female, male, total population). Red line presents Newton–Rapshon method with minimum $\hat{\beta}_x$, black line – SVD method.

4.4. Break point test

In the Figure 4.2, Figure 4.3 and Figure 4.4, we can see that at the first fitted years $\hat{k}_t$'s ranges at the same level or even have an increasing trend, while at the later years $\hat{k}_t$'s are obviously decreasing. It suggests to test for structural change in mortality series. To test for a structural break we use the Chow test invented by Gregory Chow (Chow, 1960). The idea of this test is to test whether the coefficients of two regressions are the same. So, to check if there is a structural break in the time trend of k_t, data at the structural break point have to be split and two regressions have to be run. The Chow test assumes that the residuals of the models are independent and identically distributed.

$$\begin{aligned} k_t &= c + \gamma t + \varepsilon \\ k_t^{[1]} &= c^{[1]} + \gamma^{[1]} t^{[1]} + \varepsilon^{[1]} \\ k_t^{[2]} &= c^{[2]} + \gamma^{[2]} t^{[2]} + \varepsilon^{[2]} \end{aligned} \qquad 4.21$$

The null hypothesis of the Chow test is that $c^{[1]} = c^{[2]}$ and $\gamma^{[1]} = \gamma^{[2]}$. As we can see, to use Chow test we need to know the breakpoint.

In order to compute the optimal breakpoints with given the number of breaks we use the algorithm based on a Bellman principle in dynamic programming approach. The key step is computing a triangular RSS matrix, from which the residual sum of squares for a segment can be obtained. More about this method in Prinz, *et al.,* (2007). The set of such breakpoints is chosen based on Bayesian information criterion (Schwarz, 1978). This method is implemented in the package '*strucchange'* and the name of the procedure is *breakpoints.*

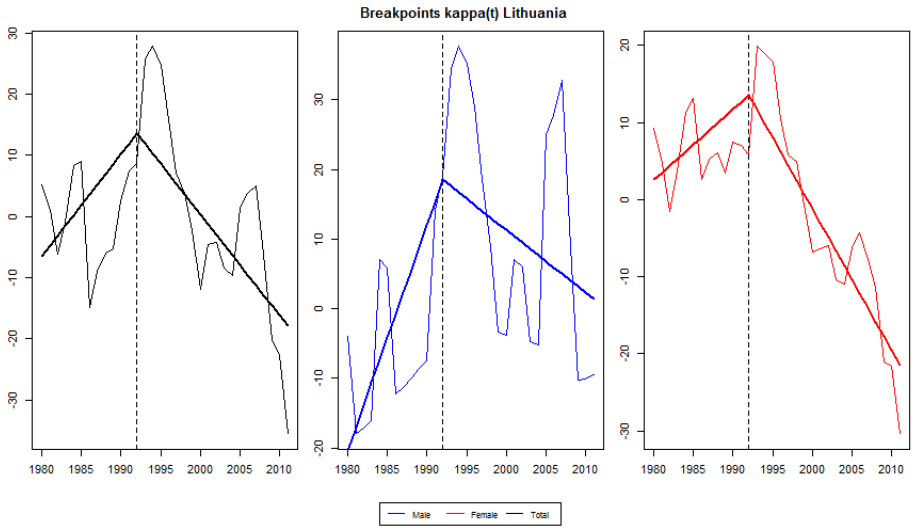

Figure 4.7 $\hat{k}_t$'s and trends from the fitted LC model for Lithuanian males, females and total population.

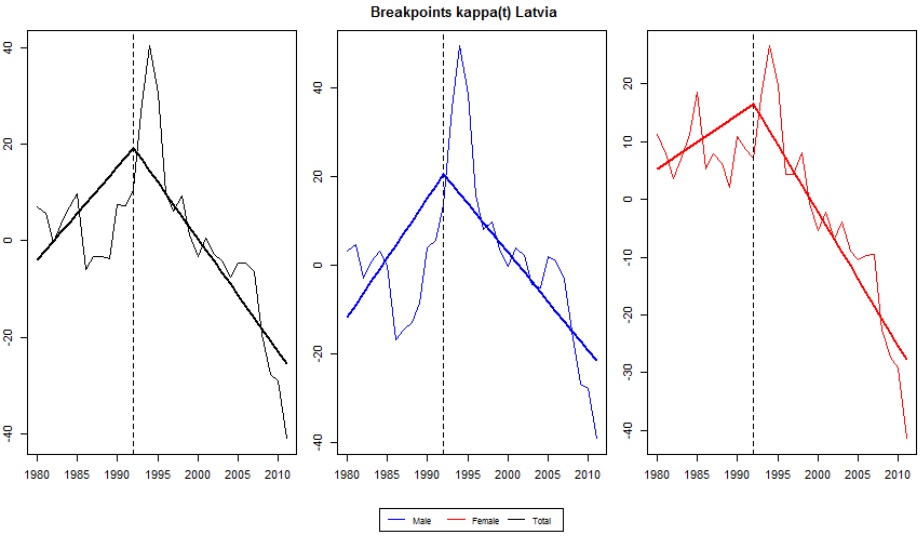

Figure 4.8 $\hat{k}_t$'s and trends from the fitted LC model for Latvian males, females and total population.

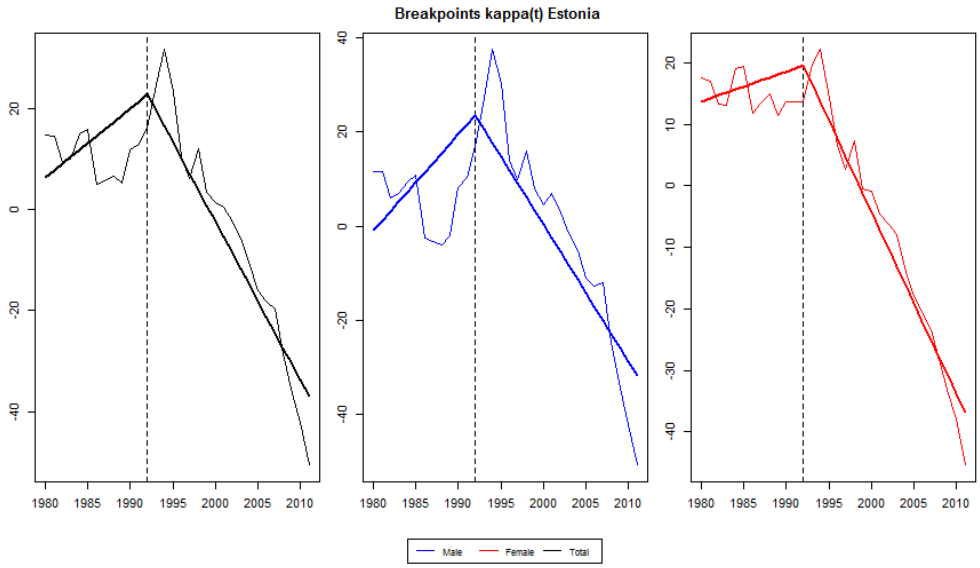

Figure 4.9 $\hat{k}_t$'s and trends from the fitted LC model for Estonian males, females and total population.

Application of algorithm for computing optimal breakpoint led to the result, that the year 1992 is the optimal choise for breakpoint. After applying the Chow test we got that differences between trends, before year 1992 and after, are signifficant. For the Chow test estimation R function *sctest* from the package *'strucchange'* is used. According to Pitacco, *et al*. (2009), the length of data series is not critical as long as it is more than about 10-20 years, so in our case 20 years should be enough.

Table 4.1 Chow test statistics for $\hat{k}_t$'s obtained by the Lee-Carter model fitted to Lithuanian, Latvian, and Estonian males, females and the total population.

$\hat{k}_t$'s obtained from	Chow test F-statistics	p-values
Lithuanian male	10.8891	$3.2 \cdot 10^{-4}$
Lithuanian female	24.7840	$6.4 \cdot 10^{-7}$
Total Lithuanian population	15.3115	$3.2 \cdot 10^{-5}$
Latvian male	33.6650	$3.6 \cdot 10^{-8}$
Latvian female	27.1857	$2.8 \cdot 10^{-7}$
Total Latvian population	33.4712	$3.8 \cdot 10^{-8}$
Estonian male	53.6743	$2.6 \cdot 10^{-10}$
Estonian female	80.7722	$2.4 \cdot 10^{-12}$
Total Estonian population	68.8206	$1.6 \cdot 10^{-11}$

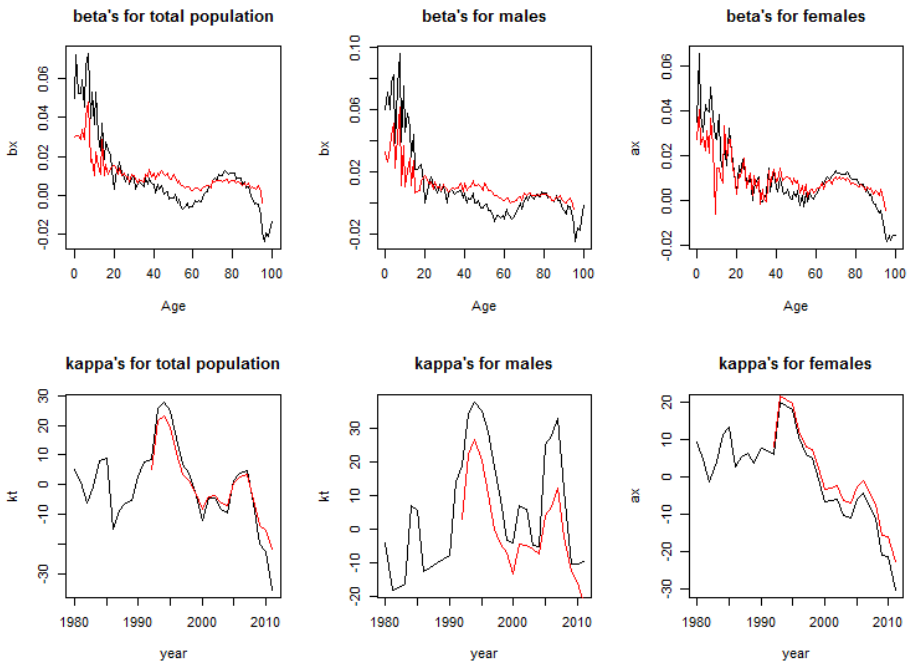

Figure 4.10 Comparison of the $\hat{k}_t$'s and the $\hat{\beta}_t$'s obtained by the Lee-Carter Poisson maximum method, fitted to Lithuanian males, females and the total population. Red line – fitting period after structural break 1992-2011; black line – 1980-2011.

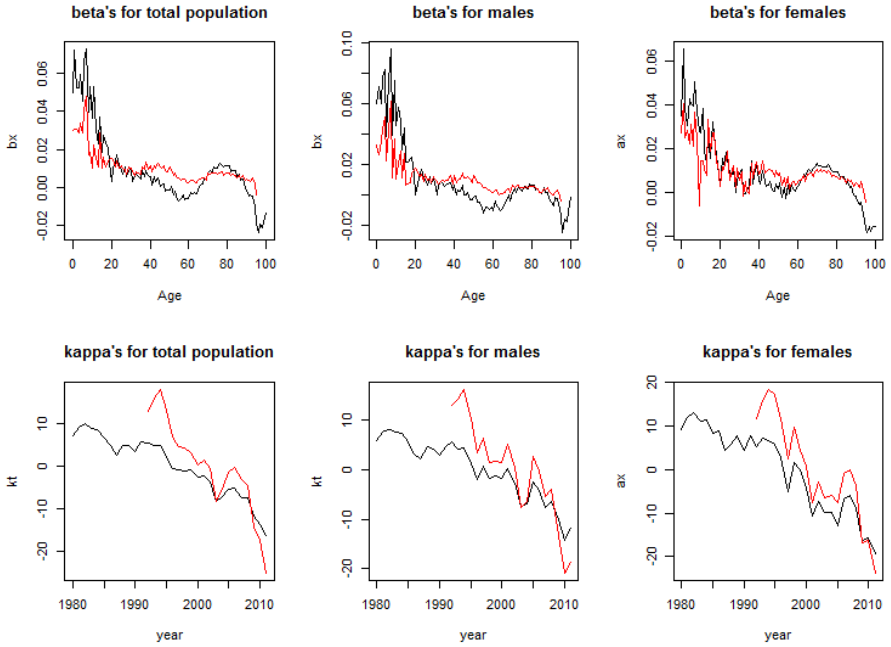

Figure 4.11 Comparison of the $\hat{k}_t$'s and the $\hat{\beta}_t$'s obtained by the Lee-Carter model, fitted to Lithuanian males, females and the total population. Red line – fitting period after structural break 1992-2011; black line – 1980-2011.

In the Figure 4.10 and Figure 4.11 is compared the outcome of the LC model fitted on the years 1980-2011 and fitted only after the structural break (1980-2011).We can see, that reducing fitting period affects $\hat{k}_t$'s and $\hat{\beta}_t$'s. Almost all $\hat{\beta}_t$'s become positive when the model is fitted to

years 1992-2011 and the slope of $\hat{k}_t$'s in the Figure 4.11 becomes bigger, which means that after year 1992 mortality rates are decreasing faster.

5. Cairns-Blake-Dowd (CBD) model

An alternative mortality forecasting method was proposed by Cairns et al. (2006). This model includes two time factors with a smoothing of the age effects using a logit transformation of the one-year death probabilities. Specifically, the logit of the one-year death probabilities is modelled as a linear function of age, with intercept and slope parameters following some stochastic process. Compared to the LC approach, the CBD model includes two time factors, which allow the model to capture the imperfect correlation in mortality rates at different ages from one year to the next.

Let's assume that $q_{t,x} = 1 - \exp(-m_{t,x})$ – death probability, $p_{t,x} = 1 - q_{t,x}$ – survival probability, and $\tilde{x} := x - \frac{\sum_{i=1}^{m} x_i}{x_m - x_1 + 1}$.

$$\ln \frac{q_{t,x}}{p_{t,x}} = \text{logit}(q_{t,x}) = k_t^{[1]} + k_t^{[2]} \cdot \tilde{x} + \varepsilon_{x,t} \leftrightarrow$$

$$q_{t,x} = \frac{\exp\left(k_t^{[1]} + k_t^{[2]} \cdot \tilde{x}\right)}{1 + \exp\left(k_t^{[1]} + k_t^{[2]} \cdot \tilde{x}\right)} + \varepsilon'_{x,t} \qquad 5.1$$

Here $\varepsilon_{x,t}$ – homoskedastic centered residual term at age x and time t, with mean 0 and variance σ^2. This specification does not have any identifiability problems so that no constrains are needed. $\varepsilon'_{x,t} \neq \varepsilon_{x,t}$, but we do not discuss assumptions on $\varepsilon'_{x,t}$, since we do not use this specification of the CBD model.

We see that in the CBD model age is treated as a continuous covariate and centered age enters the model in a linear way on the logit scale. The intercept $k_t^{[1]}$ and the slope $k_t^{[2]}$ parameters make up a bivariate time series whose future path governs the projected life tables. If $k_t^{[1]}$ is decreasing, it means that future mortality rates have been decreasing over time. If $k_t^{[2]}$ is increasing, it means that during the fit period mortality improvements have been greater at lower ages.

Estimates of the parameters can be obtained by the least squares method. In order to get maximum likelihood estimates Cairns (2007) provides an algorithm to update parameters:

Step 1. Set starting parameters as estimated by the least squares method

Step 2. Unlike in the Lee-Carter model parameters estimation algorithm, for the CBD model, both parameters $k_t^{[1]}$ and $k_t^{[2]}$ can be updated in the same step:

Step 2-1. Define variables which will be used to update both $k_t^{[1]}$ and $k_t^{[2]}$

$$Z^{[1]} := \ln\left(1 + \exp\left(k_t^{[1](k)} + k_t^{[2](k)} \cdot \tilde{x}\right)\right)$$

$$Z^{[2]} := \frac{1}{1 + \exp\left(-k_t^{[1](k)} - k_t^{[2](k)} \cdot \tilde{x}\right)}$$

$$f' := \frac{d_{t,x}}{Z^{[1]}} Z^{[2]} - e_{t,x} \cdot Z^{[2]}$$

Step 2-2. Choose small dt, e.g. $dt = 10^{-6}$

$$dZ_1^{[1]} = \ln\left(1 + exp\left(\left(k_t^{[1](k)} + dt\right) + k_t^{[2](k)} \cdot \tilde{x}\right)\right)$$

$$dZ_2^{[1]} = \frac{1}{1 + exp\left(-\left(k_t^{[1](k)} + dt\right) - k_t^{[2](k)} \cdot \tilde{x}\right)}$$

$$dZ_1^{[2]} = \ln\left(1 + exp\left(k_t^{[1](k)} + \left(k_t^{[2](k)} + dt\right) \cdot \tilde{x}\right)\right)$$

$$dZ_2^{[2]} = \frac{1}{1 + exp\left(-k_t^{[1](k)} - \left(k_t^{[2](k)} + dt\right) \cdot \tilde{x}\right)}$$

$$df'^{[1]} = \frac{d_{t,x}}{dZ_1^{[1]}} dZ_2^{[1]} - e_{t,x} \cdot dZ_2^{[1]}$$

$$df'^{[2]} = \frac{d_{t,x}}{dZ_1^{[2]}} dZ_2^{[2]} - e_{t,x} \cdot dZ_2^{[2]}$$

$$f'' = \frac{df' - f'}{dt}$$

$$k_t^{[1](k+1)} = k_t^{[1](k)} - \frac{f'}{f''^{[1]}}$$

$$k_t^{[2](k+1)} = k_t^{[2](k)} - \frac{f'}{f''^{[2]}}$$

Repeat Step 2 until convergence. In this Thesis we choose to stop procedure when a relative increase in log-likelihood function is smaller than 10^{-6}.

5.1.Application for Baltic countries

The CBD model is fitted to the Lithuanian, Latvian and Estonian males, females and the total population. In Figure 5.1 are presented resulted parameters when model was fitted for Lithuanian population at ages from 0 to 95. Note that the CBD modes was not designed to cover all ages and since most concerned age period for most actuaries are from age around 60, we also present outcomes from the CBD model fitted on the ages 60-95 (Figure 5.2). As we will see in the analysis of residuals, when the CBD model applied for the ages from 0 to 95, the error terms for the lower ages are higher.

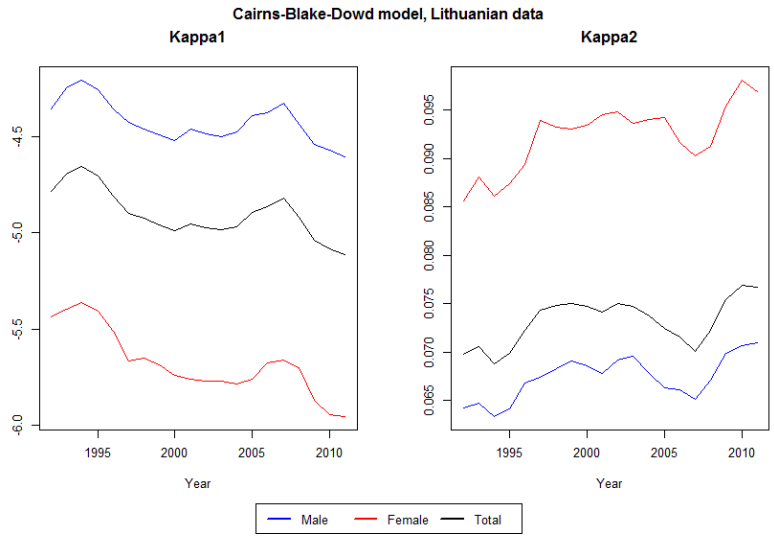

Figure 5.1 $\hat{k}_t$'s from the fitted CBD model for Lithuanian males, females and total population.

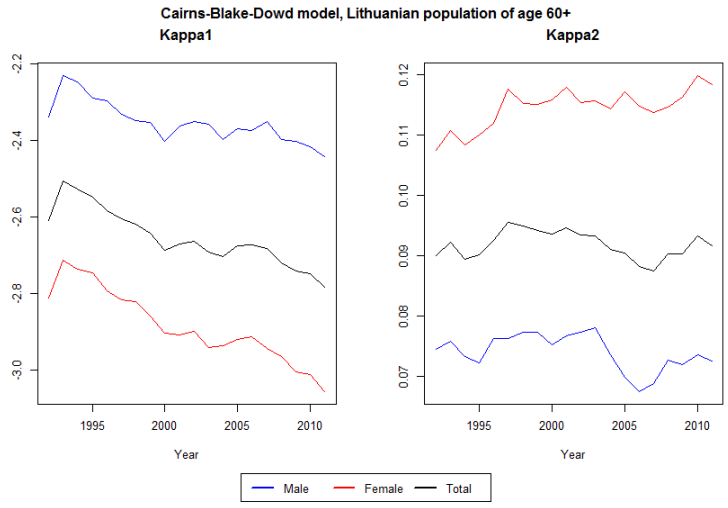

Figure 5.2 $\hat{k}_t$'s from the fitted CBD model for Lithuanian males, females and total population of age between 60 and 95.

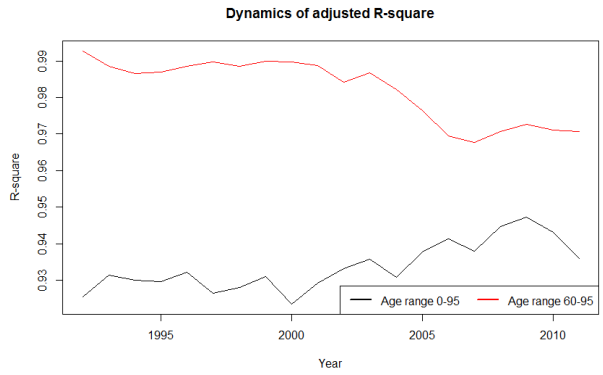

Figure 5.3 Coefficient of determination for the CBD model fitted for total Lithuanian population and for the age range from 60 to 95

When the CBD model is fitted for data of all ages, $k_t^{[1]}$'s has a decreasing tendency, but fluctuates around the trend more widely and the slope is smaller compared to the model fitted for the age range from 60 to 95. We can also see that the slope of the downward trend is the highest for the female part of the population, which means that life expectancy for the females is increasing faster than for the males and thus making the difference between gender life expectancies even bigger. When the model is fitted on the ages 0-95, $k_t^{[2]}$ is increasing over time which indicates that mortality improvements have been comparatively greater for the younger part of population. When only the restricted age group is included in the model, $k_t^{[2]}$ is more or less fluctuating around a constant which means that for the population of age 60-95 mortality improvements were more or less the same.

The value of the adjusted R-square coefficient indicates that the CBD model applied to the total population explains about 93.4% of the variance, while the model applied to the restricted population about 98.2%. The amount of explained variance increased with restricted population, because the CBD model takes advantage of the approximate linearity in age at higher ages to provide a representation of one-year death probabilities even with lack of empirical data. The adjusted coefficient of determination close to 1 indicates, that the CBD model is good to describe mortality in Lithuania.

6. Smoothing

Actuaries use projected life tables in order to compute life annuity prices, life insurance premiums as well as reserves that have to be held by insurance companies to enable them to be able to pay the future contractual benefits. Any irregularities in these life tables would then be passed on to the price list and to the balance sheets, which is not desirable. Therefore, as long

as these irregularities do not reveal particular features of the risk covered by the insurer, but are likely to be caused by sampling errors, actuaries prefer to resort to statistical techniques to produce life tables that exhibit a regular progression, in particular with respect to age.

Durban et al. (2004) have smoothed death rates with P-splines in the context of a Poisson model. The P-spline approach is an example of a regression model. Regression models take a family of basis functions, and choose a combination of them that best fits the data according to some criterion. The P-spline approach uses a spline basis, with a penalty function that is introduced in order to avoid oversmoothing.

As can be seen in Figure 4.5, the estimated β_x's exhibit an irregular pattern. This is undesirable from an actuarial point of view, since the resulting projected life tables will also show some unstable variation across ages.

We only need to smooth the β_x's in order to get projected life tables with mortality varying smoothly across the ages. In order to smooth the estimated β_x's we can use the least-squares approach objective function, proposed in Delwarde, *et al.,* (2007):

$$\begin{aligned} F(\alpha,\beta,k) &= \sum_{x=x_1}^{x_m}\sum_{t=t_1}^{t_n}\left(\ln \hat{m}_{t,x} - \alpha_x - \beta_x k_t\right)^2 \\ &\quad + \pi_\beta \sum_{x=x_1}^{x_m} (\beta_{x+2} - 2\beta_{x+1} + \beta_x)^2 \\ &= \sum_{x=x_1}^{x_m}\sum_{t=t_1}^{t_n}\left(\ln \hat{m}_{t,x} - \alpha_x - \beta_x k_t\right)^2 + \beta'\pi_\beta \Delta'\Delta\beta\,, \end{aligned} \qquad 6.1$$

here

$$\Delta = \begin{bmatrix} 1 & -2 & 1 & 0 & 0 & 0 & \dots \\ 0 & 1 & -2 & 1 & 0 & 0 & \dots \\ 0 & 0 & 1 & -2 & 1 & 0 & \dots \\ 0 & 0 & 0 & 1 & -2 & 1 & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}$$

Here π_β is the smoothing parameter. The first term is the Lee-Carter model objective function and the second is to penalize irregular β_x's. Second-order differences penalize deviations from the linear trend. The smoothing parameter π_β controls the smoothing level of the original data.

If $\pi_\beta = 0$, smoothed data perfectly fits the original data. In the limit, $\pi_\beta \to \infty$, we obtain a linear fit.

The estimation algorithm remains the same as it was without smoothing, the only difference is in updating the parameter β_x.

Step 4. Update parameter $\hat{\beta}_x$:

$$\hat{\beta}_x^{(k+1)} = \frac{\left(\sum_{t=t_1}^{t_n}\left(\hat{k}_t^{(k+1)}\right)^2 + diag\left(\pi_\beta \Delta'\Delta\right) - \pi_\beta \Delta'\Delta\right)\hat{\beta}_x^{(k)}}{\left(\sum_{t=t_1}^{t_n}\left(\hat{k}_t^{(k+1)}\right)^2 + diag\left(\pi_\beta \Delta'\Delta\right)\right)}$$

$$+ \frac{\sum_{t=t_1}^{t_n} \hat{k}_t^{(k)}\left(\ln\widehat{(m_{t,x})} - \hat{\alpha}_x^{(k+1)} - \hat{\beta}_x^{(k)}\hat{k}_t^{(k+1)}\right)}{\left(\sum_{t=t_1}^{t_n}\left(\hat{k}_t^{(k+1)}\right)^2 + diag\left(\pi_\beta \Delta'\Delta\right)\right)}$$

The choice of the smoothing parameter π_β is crucial, since too small π_β causes under-smoothing and too large over-smoothing.

In order to estimate optimal value of the smoothing parameter π_β, firstly we have to define the prediction errors $\varepsilon_{xt,\pi_\beta}$:

$$\varepsilon_{xt,\pi_\beta} = \ln\left(\hat{m}_{t,x}\right) - \hat{\alpha}_{x,\pi_\beta} - \hat{\beta}_{x,\pi_\beta}\hat{k}_{t,\pi_\beta} \qquad 6.2$$

Here $\hat{\alpha}_{x,\pi_\beta}$, $\hat{\beta}_{x,\pi_\beta}$ and $\hat{k}_{t,\pi_\beta}$ are the least-squares estimators obtained from penalized function. The optimal value of $\hat{\pi}_\beta$ is obtained by solving:

$$\hat{\pi}_\beta = \arg\min_{\pi_\beta}\left[\sum_{x=x_1}^{x_m}\sum_{t=t_1}^{t_n}\varepsilon_{xt,\pi_\beta}^2\right] \qquad 6.3$$

6.1.Application to Baltic countries

We fit the penalized Lee-Carter model for the Baltic countries and in Figure 6.1 we present graphs for the Lithuanian data. Applied smoothing method does impact neither the estimated $\alpha_x{}'$s nor the $k_t{}'$s. However, the impact on β_x's is obvious and now appears to behave very

regularly with age. P-spline procedure for smoothing is implemented in R package *'pspline'* function *smooth.Pspline*.

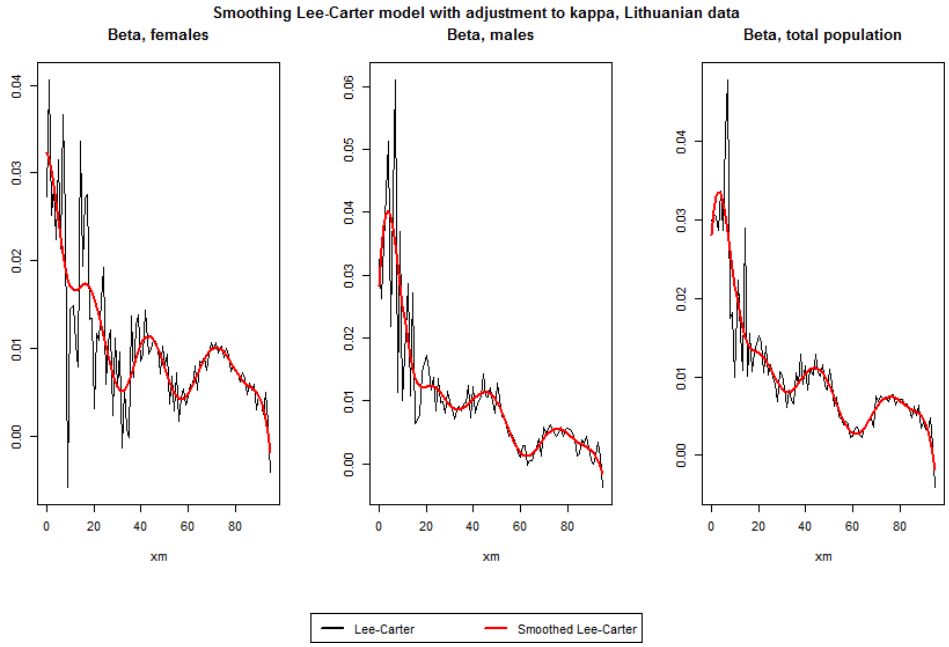

Figure 6.1 β_x's obtained from the Lee-Carter model, Poisson maximum likelihood estimation, fitted for Lithuanian population (from the left-hand side: female, male, total population). Red line presents the smoothed LC method, while black – without smoothing.

7. Analysis of residuals

Since we work in a regression framework, it is essential to inspect the residuals. Model performance is assessed in terms of the randomness of the residuals. A lack of randomness would indicate the presence of systematic variations, such as age-time interactions. We note that the adjustment of k'_ts in the Lee-Carter case may have introduced systematic changes to the residuals so that the examination of model performance is in fact based on the residuals computed with the adjusted k'_ts.

When the parameters are estimated by least-squares, Pearson residuals have to be inspected (Pitacco, et al., 2009). In the Lee-Carter case, these residuals are given by:

$$r_{xt} = \frac{\hat{\varepsilon}_{x,t}}{\sqrt{\frac{1}{(x_m - x_1)(t_n - t_1 - 1)} \sum_{x=x_1}^{x_m} \sum_{t=t_1}^{t_n} \left(\hat{\varepsilon}_{x,t}\right)^2}}, \qquad 7.1$$

where $\hat{\varepsilon}_{x,t} = \ln\left(\hat{m}_{t,x}\right) - \hat{\alpha}_x - \hat{\beta}_x \hat{k}_t$. In the CBD case $\hat{\varepsilon}_{x,t} = \text{logit}\left(\hat{q}_{t,x}\right) - \hat{k}_t^{[1]} - \hat{k}_t^{[1]} x$ and

$$r_{xt} = \frac{\hat{\varepsilon}_{x,t}}{\sqrt{\frac{1}{(x_m - x_1 - 1)(t_n - t_1 + 1)} \sum_{x=x_1}^{x_m} \sum_{t=t_1}^{t_n} \left(\hat{\varepsilon}_{x,t}\right)^2}} \qquad 7.2$$

If residuals exhibit some regular pattern, this means that the model is not able to describe all of the phenomena appropriately.

7.1. Application to Baltic countries

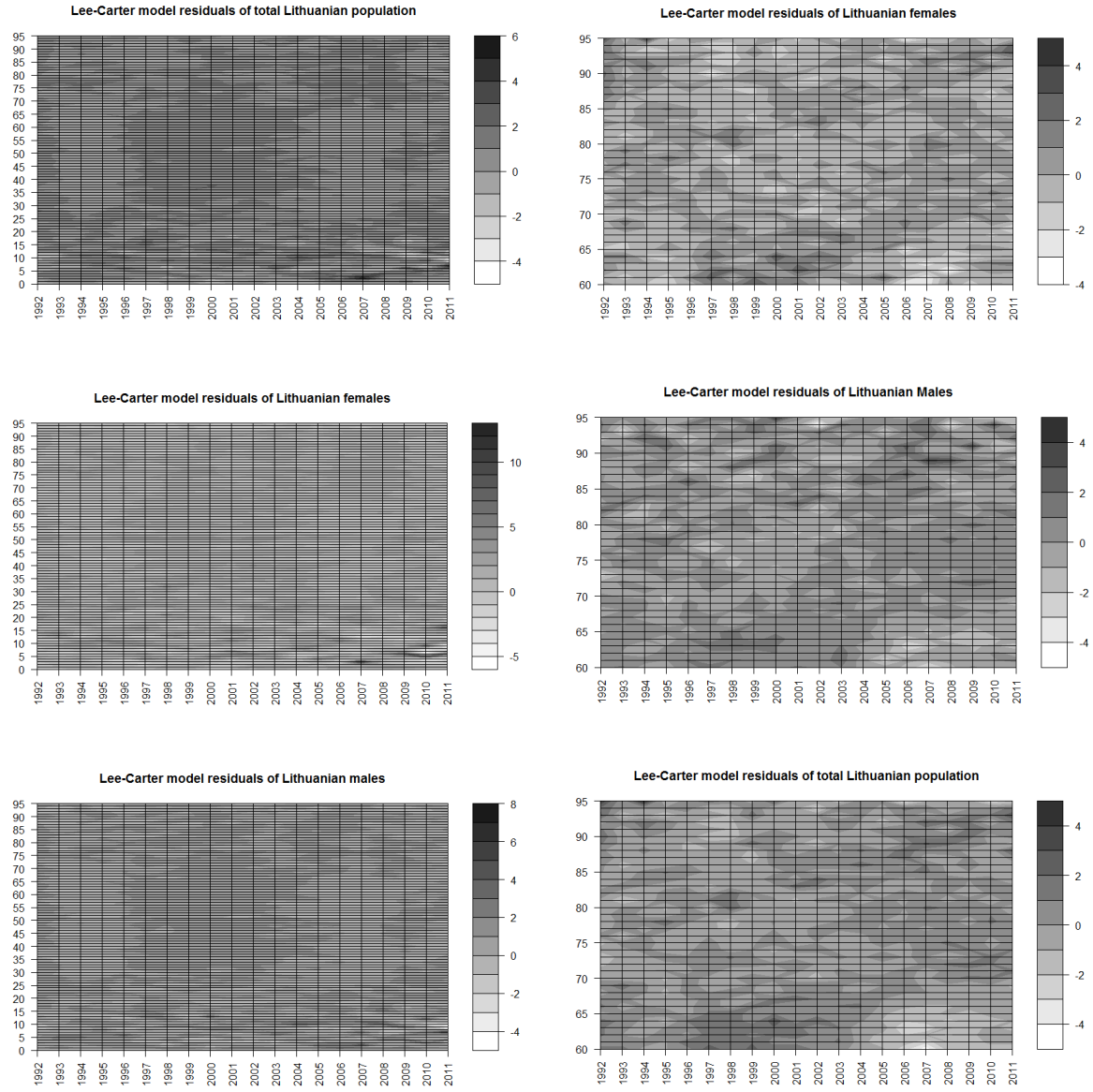

Figure 7.1 Residuals of the Lee-Carter model fitted to Lithuanian data. Left-hand side fitting ages 0-95, right-hand side – 60-95.

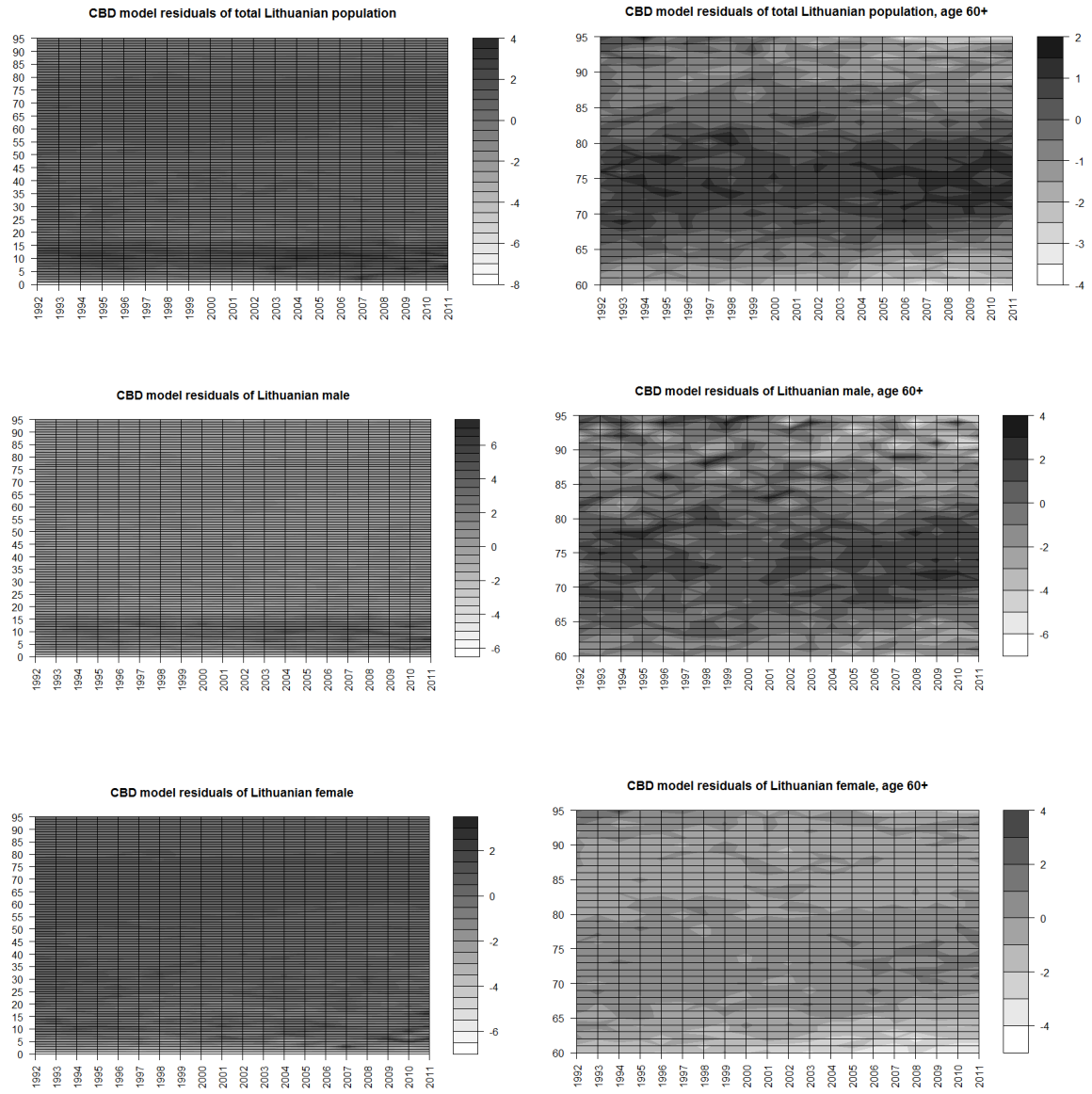

Figure 7.2 Residuals of the Crains-Blake-Dowd model fitted to Lithuanian data. Left-hand side fitting ages 0-95, right-hand side – 60-95.

In the Figure 7.1 residuals of the Lee-Carter approach fitted to the total Lithuanian population, separately to the males and females are presented. We see that the residuals for the young ages are higher than for the older. When the model is fitted to the ages 0-95 we can observe a little impact which seems to be a cohort affect, which is not visible when the model is fitted for the ages 60-95. According to Pitacco et al. (2009), adjustment to the k_t's carried out using Poisson maximum likelihood estimation, may have introduced systematic changes to the residuals.

In the Cairns-Blake-Dowd approach (Figure 7.2) high values of the residuals are even more obvious than in the Lee-Carter case. This information may mean that both the CBD and the LC

models are not suitable for younger ages. Since older ages are the most relevant in pension and annuity applications, we restrict the further analysis to ages 60 and over.

8. Mortality projections

Both the Lee-Carter model and the Cairns-Blake-Dowd model have time factors which are considered to be stochastic processes. $k_t^{[1]}$'s , $k_t^{[2]}$'s and k_t's can be modelled and forecasted by using the Box-Jenkins technique. By using these forecasts life expectancies, age-specific mortality rates, and single premiums for life annuities can be estimated.

8.1. Lee-Carter model

The dynamics of k_t's (estimated by using the Lee-Carter model) is usually explained by using the classical autoregressive integrated moving average (ARIMA(p, d, q)) models if it is (at least weakly) stationary and

$$\nabla^d k_t = \varphi_1 \nabla^d k_t + \cdots + \varphi_p \nabla^d k_t + \xi_t + \psi_1 \xi_{t-1} + \cdots + \psi_q \xi_{t-q}, \qquad 8.1$$

With $\varphi_p \neq 0, \psi_q \neq 0$, and where ξ_t is a Gaussian white noise process.

In order to fit ARIMA(p, d, q) model, firstly we need to identify the orders: d – differencing order, q – moving average order and p – autoregressive order. If the series has a unit root, it can be removed be using a first difference. Preliminary orders of p and q can be found by investigating the autocorrelation function (ACF) and partial autocorrelation function (PACF) of the time series. Usually orders do not exceed 5 and the best fitted set of orders leads to lowest BIC.

Weak stationarity. The process k_t is said to be weakly stationary if

- $E[k_t^2] < \infty, \forall t \in t_1, \dots, t_n$
- $E[k_t] = \text{const}, \ \forall t \in t_1, \dots, t_n$
- $\gamma_X(s, t) = \gamma_X(s + h, t + h), \ \forall s + h, t + h, s, t \in t_1, \dots, t_n.$

Two most common cases when time series is non-stationary:

- Presence of a deterministic trend. Time series can be transformed to stationary by removing this time trend.

- Presence of a unit root. In such a case the process is transformed into a stationary one by applying first order differences.

By a first look to the estimated k_t's (for all Baltic countries) we can see that they have a downward trend, tending to decrease gradually over time. In Figure 8.1 we can observe that ACF is decreasing in a linear rate as the lag is increasing, which is the classic behavior of a non-stationary process.

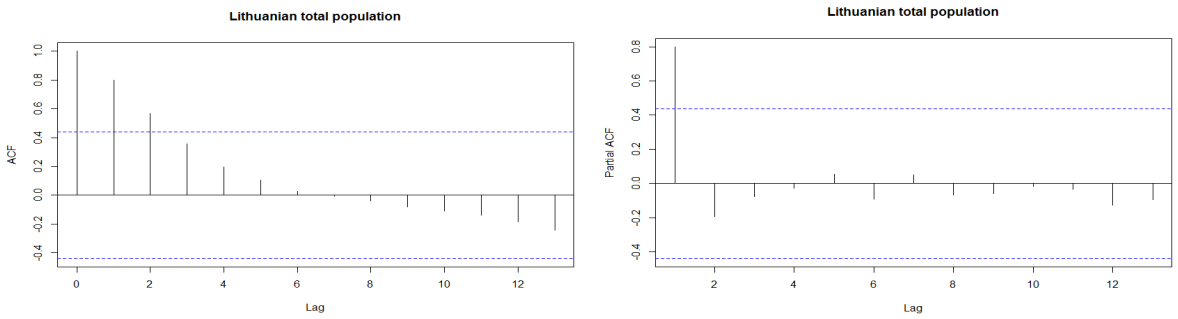

Figure 8.1 Autocorrelation and partial autocorrelation functions fitted for k_t's of total Lithuanian population.

In addition to the graphs, we use the Kwiatowski-Philips-Schmidt-Shin (KPSS) test, see Kwiatkowski et al. (1992), which rejects both level and trend stationarity. To test if the time series has a unit root we use the Phillips-Perron test (Phillips & Perron, 1988), which accepts the null hypothesis that time series has a unit root in all the cases, see Table 8.1. After differentiating the time series the KPSS test does not reject that time series is level stationary, see Table 8.2.

Table 8.1 Values of the Phillips-Perron test statistics and p-values. Test used for k_t's obtained from the Lee-Carter model fitted to the Baltic countries total population, males and females.

Country	Gender	Phillips-Perron	
		statistics	p-value
Lithuania	Total population	-2.5061	0.3796
Lithuania	Male	-2.2452	0.4790
Lithuania	Female	-3.1951	0.1171
Latvia	Total population	-3.0863	0.1585
Latvia	Male	-3.0291	0.1803
Latvia	Female	-3.1360	0.1396
Estonia	Total population	-2.8974	0.2305
Estonia	Male	-2.6180	0.3369
Estonia	Female	-3.2166	0.1089

Table 8.2 Values of the Kwiatowski-Philips-Schmidt-Shin test statistics and p-values. Test used for k_t's and differentiated k_t's obtained from the Lee-Carter model fitted to the Baltic countries total population, males and females.

Country	Gender	KPSS test before differencing the time series		KPSS after differencing the time series	
		statistics	p-value	statistics	p-value
Lithuania	Total population	0.7161	0.012	0.1421	>0.1
Lithuania	Male	0.4779	0.046	0.1265	>0.1
Lithuania	Female	0.9257	<0.01	0.1908	>0.1
Latvia	Total population	0.9007	<0.01	0.1737	>0.1
Latvia	Male	0.8545	<0.01	0.1462	>0.1
Latvia	Female	0.9592	<0.01	0.2227	>0.1
Estonia	Total population	1.0192	<0.01	0.2910	>0.1
Estonia	Male	0.984	<0.01	0.2970	>0.1
Estonia	Female	1.047	<0.01	0.2940	>0.1

After differencing the time series neither ACF, nor PACF are significantly different from zero, an ARIMA(0,1,0) process seems to describe the behavior of k_t's the best. This result is also supported by the minimum BIC. So, the dynamics of the estimated k_t's for all three Baltic countries are given by

$$k_t = k_{t-1} + d + e_t, \tag{8.2}$$

where *d* is drift parameter and e_t's are assumed to be independent and normally distributed with average 0 and variance σ^2.

The point forecast of the time index for *h* steps a head is

$$\dot{k}_{t_n+h} = \mathrm{E}\left(k_{t_n+h} | k_{t_1}, \dots, k_{t_n}\right) = k_{t_n} + hd \tag{8.3}$$

It means that the forecast is a straight line with a slope *d.* The conditional variance of the forecast is

$$\mathrm{Var}\left(k_{t_n+h} | k_{t_1}, \dots, k_{t_n}\right) = h\sigma^2 \tag{8.4}$$

As the forecasting horizon *h* increases, the conditional standard errors of the forecast increase by the $\sqrt{h}$.

The dynamics of ARIMA(0,1,0) ensures that $k_t - k_{t-1}$ are independent and normally distributed with mean *d* and variance σ^2. Sample mean and variance $k_t - k_{t-1}$ indeed are the ML estimators of *d* and σ^2.

$$\hat{d} = \frac{1}{t_n - t_1}\sum_{t=t_2}^{t_n}\left(\hat{k}_t - \hat{k}_{t-1}\right) = \frac{\hat{k}_{t_n} - \hat{k}_{t_1}}{t_n - t_1}$$

$$\sigma^2 = \frac{1}{t_n - t_1}\sum_{t=t_2}^{t_n}\left(\hat{k}_t - \hat{k}_{t-1} - \hat{d}\right)^2$$

Table 8.3 The Lee-Carter model fitted for Baltic countries, fitting period 1992-2011

Country	**Gender**	$\hat{d}$ **Age [0,95] (s.e.)**	σ^2 **Age [0,95]**	$\hat{d}$ **Age [60,95] (s.e)**	σ^2 **Age [60,95]**
Lithuania	Male	-1.3214 (0.65)	67.002	-0.132 (0.07)	1.269
Latvia	Male	-2.3019 (0.92)	71.721	-0.363 (0.15)	1.935
Estonia	Male	-3.1783 (1.05)	37.726	-0.694 (0.2)	1.128
Lithuania	Female	-1.6065 (0.61)	23.202	-0.387 (0.15)	1.216
Latvia	Female	-2.6153 (0.87)	49.192	-0.558 (0.18)	1.825
Estonia	Female	-3.2761 (1.02)	16.867	-0.944 (0.28)	1.170

This approach connects the first and last points of the available data and then extends the resulting line further in order to produce a forecast. Lee and Carter (1992) did not impose the random walk with drift model for all situations. However, this model has been judged to be appropriate in very many cases (e.g. Belgium, see Pitacco et al., (2009); G7 countries, see Tuljapurkar et al. (2000)). According to Pitacco et al., (2009), even when the more complex model is applied, it is found to give results which are close to those obtained with the random walk with drift.

8.2. Cairns-Blake-Dowd model

When the CBD model is fitted to the data, the changes in the time index $k_t = [k_t^{[1]}, k_t^{[2]}]^T$ has a linear trend which means that a bivariate random walk with drift may be a proper model to describe dynamics of k_t.

$$k_t = \begin{cases} k_{t-1}^{[1]} + d^{[1]} + \xi_t^{[1]} \\ k_{t-1}^{[2]} + d^{[2]} + \xi_t^{[2]} \end{cases},$$

where $d = \left[d^{[1]}, d^{[2]}\right]^T$ are drift parameters and $\xi = \left[\xi^{[1]}, \xi^{[2]}\right]^T$ are independent bivariate Normally distributed pairs, with mean 0 and covariance matrix $\sum$

$$\sum = \begin{pmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{21} & \sigma_2^2 \end{pmatrix}$$

$$\hat{d} = \begin{pmatrix} \frac{\hat{k}_{t_n}^{[1]} - \hat{k}_{t_1}^{[1]}}{t_n - t_1} \\ \frac{\hat{k}_{t_n}^{[2]} - \hat{k}_{t_1}^{[2]}}{t_n - t_1} \end{pmatrix}$$

$$\Sigma = \begin{pmatrix} \frac{1}{t_n - t_1}\sum_{t=t_2}^{t_n}\left(\hat{k}_t^{[1]} - \hat{k}_{t-1}^{[1]} - \hat{d}^{[1]}\right)^2 & \frac{1}{t_n - t_1}\sum_{t=t_2}^{t_n}\sum_{t=t_2}^{t_n}\left(\hat{k}_t^{[2]} - \hat{k}_{t-1}^{[2]} - \hat{d}^{[2]}\right)\left(\hat{k}_s^{[2]} - \hat{k}_{s-1}^{[2]} - \hat{d}^{[2]}\right) \\ \frac{1}{t_n - t_1}\sum_{t=t_2}^{t_n}\sum_{t=t_2}^{t_n}\left(\hat{k}_t^{[2]} - \hat{k}_{t-1}^{[2]} - \hat{d}^{[2]}\right)\left(\hat{k}_s^{[2]} - \hat{k}_{s-1}^{[2]} - \hat{d}^{[2]}\right) & \frac{1}{t_n - t_1}\sum_{t=t_2}^{t_n}\left(\hat{k}_t^{[2]} - \hat{k}_{t-1}^{[2]} - \hat{d}^{[2]}\right)^2 \end{pmatrix}$$

Application of these formulas gives the following estimates for data of Baltic Countries

Table 8.4 Cairns-Black-Dowd model fitted for Baltic countries, fitting period 1992-2011, age range from 0 to 95.

Country	Gender	$\hat{d}^{[1]}$ (s.e.)	$\hat{d}^{[2]}$ (s.e.)	$\hat{\sigma}_1^2$	$\hat{\sigma}_2^2$	$\hat{\sigma}_{12}^2 = \hat{\sigma}_{21}^2$
Lithuania	Male	-0.0131 (0.005)	0.0004 (0.0001)	0.00373	$1.6 \cdot 10^{-6}$	$0.0 \cdot 10^{-35}$
Latvia	Male	-0.0245 (0.009)	0.0006 (0.00018)	0.00707	$3.2 \cdot 10^{-6}$	$0.0 \cdot 10^{-35}$
Estonia	Male	-0.0318 (0.011)	0.0005 (0.0002)	0.00570	$5.4 \cdot 10^{-6}$	$0.0 \cdot 10^{-34}$
Lithuania	Female	-0.0272 (0.008)	0.0006 (0.0002)	0.00387	$3.7 \cdot 10^{-6}$	$0.0 \cdot 10^{-17}$
Latvia	Female	-0.0288 (0.01)	0.0005 (0.00018)	0.00867	$5.1 \cdot 10^{-6}$	$0.0 \cdot 10^{-17}$
Estonia	Female	-0.0441 (0.015)	0.0006 (0.0003)	0.00680	$5.5 \cdot 10^{-6}$	$0.0 \cdot 10^{-17}$

Table 8.5 Cairns-Black-Dowd model fitted for Baltic countries, fitting period 1992-2011, age range from 60 to 95.

Country	Gender	$\hat{d}^{[1]}$ (s.e.)	$\hat{d}^{[2]}$(s.e.)	$\hat{\sigma}_1^2$	$\hat{\sigma}_2^2$	$\hat{\sigma}_{12}^2 = \hat{\sigma}_{21}^2$
Lithuania	Male	-0.0548 (0.002)	-0.0001 (0.00016)	0.00129	$5.2 \cdot 10^{-6}$	$0.0 \cdot 10^{-36}$
Latvia	Male	-0.0116 (0.004)	$-8.9 \cdot 10^{-5}$ (0.0001)	0.00134	$1.4 \cdot 10^{-5}$	$0.0 \cdot 10^{-35}$
Estonia	Male	-0.0206 (0.0057)	$9.8 \cdot 10^{-5}$ (0.0002)	0.00096	$9.8 \cdot 10^{-6}$	$0.0 \cdot 10^{-34}$
Lithuania	Female	-0.0128 (0.005)	0.0005 (0.0002)	0.00109	$5.6 \cdot 10^{-6}$	$0.0 \cdot 10^{-17}$
Latvia	Female	-0.0170 (0.005)	0.0002 (0.0002)	0.00174	$1.5 \cdot 10^{-5}$	$0.0 \cdot 10^{-17}$
Estonia	Female	-0.0282 (0.0086)	0.0006 (0.0003)	0.00108	$8.1 \cdot 10^{-6}$	$0.0 \cdot 10^{-17}$

As we can see in the tables above, autocorrelation coefficients between $\hat{k}_t^{[1]}$ and $\hat{k}_t^{[2]}$ are close to zero.

9. Prediction intervals

In the current application, it is impossible to derive the relevant prediction intervals analytically. The reason for this is that two very different sources of uncertainty have to be combined: sampling errors in the parameters α_x, β_x and k_t, and the forecast errors in the projected k_t's. An additional complication is that the measures of interest – life expectancies or life annuities premiums and reserves – are complicated non-linear functions of the parameters α_x, β_x and k_t and of the ARIMA parameters. The key idea behind the bootstrap is to resample from the original data in order to create replicated data sets, from which the variability of quantities of interest can be assessed. Because this approach involves repeating the original data analysis procedure with many replicate sets of data, it is sometimes called a computer-intensive method. Bootstrap techniques are particularly useful when, as in our problem, theoretical calculation with the fitted model is too complex.

To avoid any Normality assumption, we use the bootstrap percentile interval to construct confidence interval for the predicted life expectancy. The bootstrap procedure yields B samples of α_x, β_x and k_t parameters, denoted as α_x^b, β_x^b and $k_x^b, b = 1,2, \dots, B$.

One of the possible ways to carry on bootstrap procedure, described in the previous paragraph, is a Poisson bootstrap. Starting from the observations $(N_{x,t}, D_{x,t})$, Brouhns, et al. (2005) create B bootstrap samples $(N_{x,t}, D_{x,t}^b)$, b=1,2,…,B, where the $D_{x,t}^b$'s are realizations from the Poisson distribution with mean $N_{x,t}\hat{m}_{x,t} = D_{x,t}$. The bootstrapped death counts $D_{x,t}^b$ are obtained by applying a Poisson noise to the observed numbers of deaths. The α_x's, β_x's and k_t's are estimated for each bootstrap sample. The paper by Brouhns, et al. (2005) provides more details about this method.

9.1.Application to Baltic countries

In the approach proposed by Carter and Lee (1992), future age-specific death rates are obtained inserting the point wise projections $\dot{k}_{t_n+h}$ into the formulas:

$$\ln(\dot{m}_{x,t_n+h}) = \hat{\alpha}_x + \hat{\beta}_x \dot{k}_{t_n+h} \leftrightarrow \dot{m}_{x,t_n+h} = \exp(\hat{\alpha}_x + \hat{\beta}_x \dot{k}_{t_n+h}) \qquad 9.1$$

In this case, the rates from the last year of fitting period are fitted rates. Such rates are called jump-off rates. Bell (1997) criticized the Lee-Carter method because discontinuity in the jump-off year is possible and suggested that the forecast could be started with observed mortality rates instead of fitted. Such a change would help to avoid possible jump-off if fitted mortality rates do not precisely fit observed ones. The forecast mortality rates are adjusted to the latest available empirical data:

$$\dot{m}_{x,t_n+h} = m_{x,t_n} \exp\left(\hat{\beta}_x\left(\dot{k}_{t_n+h} - k_{t_n}\right)\right) \qquad 9.2$$

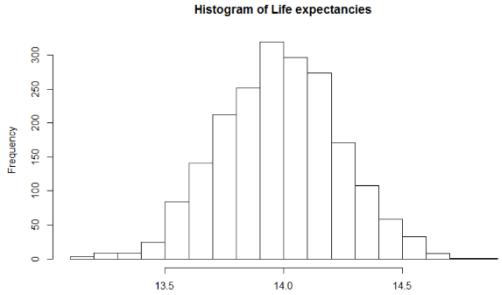

Figure 9.1 Histogram for the 2000 bootstrapped values of the life expectancies at the age 65 in year 2012 for Lithuanian male population.

The bootstrap confidence interval for 65 years male at year 2012 with the confidence level 95% is (13.32, 14.52) with mean 13.92. To estimate bootstrap, *lifemetrics* code is used, see Cairns (2007).

10. Forecasting life expectancies

Life expectancies are projected at a retirement age 65 for males and females separately in the three Baltic countries. Computation is obtained from the Lee-Carter and the Crains-Blake-Dowd models, by replacing mortality rates with their forecast values.

Life expectancies are often used by demographers to measure the evolution of mortality. Specifically, $\bar{e}^{\uparrow}_{x,t}$ is the average number of years that an x-aged individual in year t will survive, allowing for the evolution of mortality rates with time after t. We, thus, expect that this person will die in year $t + \bar{e}^{\uparrow}_{x,t}$ at age $x + \bar{e}^{\uparrow}_{x,t}$. The formula giving $\bar{e}^{\uparrow}_{x,t}$ is:

$$\begin{aligned} \bar{e}^{\uparrow}_{x,t} &= \int_{\omega \geq 0} \exp\left(-\int_0^{\omega} m_{x+h,t}\, dh \right) d\omega \\ &= \frac{1-\exp(-m_{x,t})}{m_{x,t}} + \sum_{i \geq 1} \left(\prod_{j=0}^{i-1} \exp(-m_{x+j,t}) \right) \frac{1-\exp(-m_{x+i,t})}{m_{x+i,t}} \end{aligned} \qquad 10.1$$

In 10.1 formula, the ratio $\frac{1-\exp(-m_{x,t})}{m_{x,t}}$ is the average fraction of the year lived by an individual alive at age $x + i$, and the product $\prod_{j=0}^{i-1} \exp(-m_{x+j,t})$ is the probability ${}_ip^{\uparrow}_{x,t}$ of reaching age $x + i$. Life expectancies obtained by the Lee-Carter model are compared with the Cairns-Blake-Dowd forecast.

To estimate the forecasts we use R function *forecast* from the package *'forecast'*.

Figure 10.1 displays the values of the life expectancies at age 65 and 95% prediction interval for the Lee-Carter model. The values obtained with different models are very close with a slightly larger values coming from the Lee-Carter approach for females projection of life expectancies. We can see that in the case of male life expectancy confidence intervals are very wide and the lower 95% confidence interval is decreasing, which means that it is more than 5% probability that life expectancy can decrease. Confidence intervals for female are a little narrower compared to male and the lower confidence border is not decreasing.

In Figure 10.2 life expectancies projected by the Lee-Carter and the Cairns-Blake-Dowd models for Latvian population are presented. We can see that life expectancies are very similar to Lithuanian, except the slopes are a little bigger, which means that life expectancy is increasing faster.

Figure 10.3 presents life expectancies for the Estonian population. We can see that life expectancies in Estonia significantly differ compared to other two Baltic states. Life expectancies in Estonia are increasing much faster and the 95% confidence intervals are much narrower for both males and females.

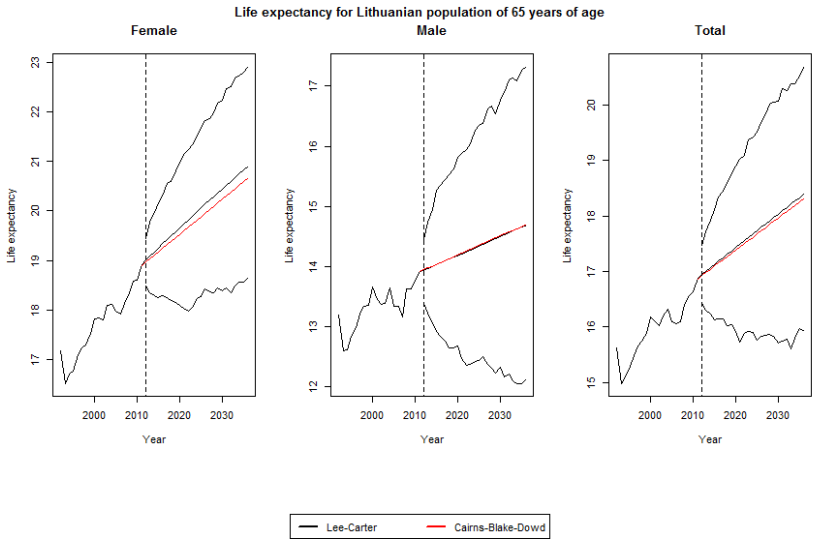

Figure 10.1 Comparison of life expectancy $e_{65}^{\uparrow}(2011,\ldots,2035)$ projected by the Lee-Carter and the Cairns-Blake-Dowd models fitted for Lithuanian population of age 60-95.

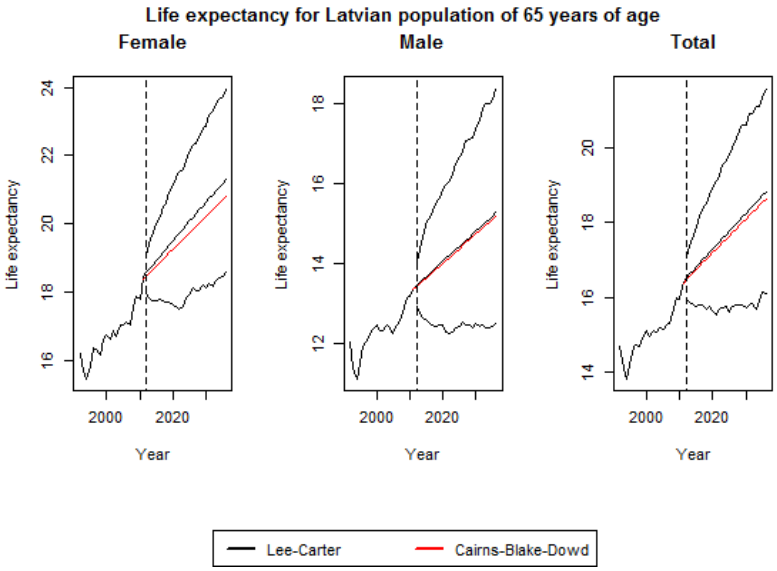

Figure 10.2 Comparison of life expectancy $e_{65}^{\uparrow}(2011,\ldots,2035)$ projected by the Lee-Carter and the Cairns-Blake-Dowd models fitted for Latvian population of age 60-95.

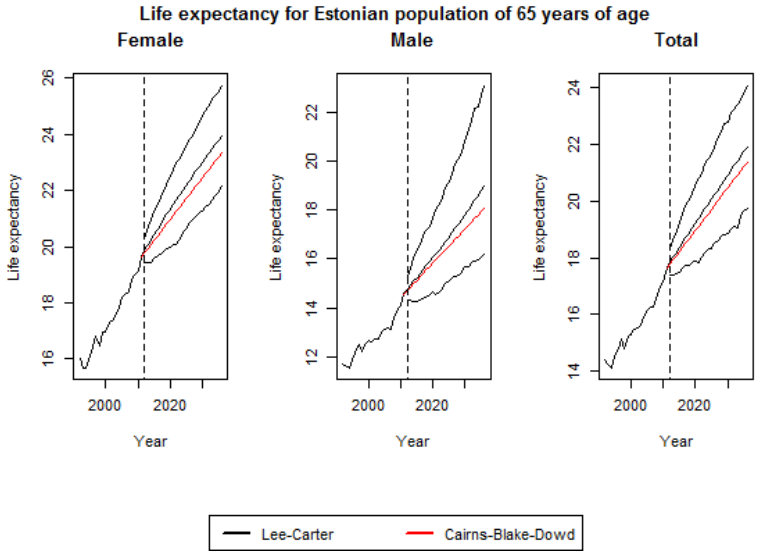

Figure 10.3 Comparison of life expectancy $e^{\uparrow}_{65}(2011,\ldots,2035)$ projected by the Lee-Carter and the Cairns-Blake-Dowd models fitted for Estonian population of age 60-95.

In the year 2014, the retirement age in Lithuania for males is 63, while for females – 61. According to Lithuanian government, in the year 2026 both males and females retirement age will be 65. In order to achieve this, male retirement age increases by 2 months every year and for female by 4 months. In Figure 10.4 life expectancy at the retirement age and corresponding 95% confidence intervals are presented. We can see that the new retirement age system decreases life expectancy in the retirement for both male and female population. According to the estimated life expectancy, difference of the life time in the retirement for females who retires in 2012 and who retires in 2026 is 2.46 years, while for males – 0.51 years. Thus, by the new retirement age system females suffer more than males and cohort which will be 65 years old at the year 2026 will be negatively affected the most.

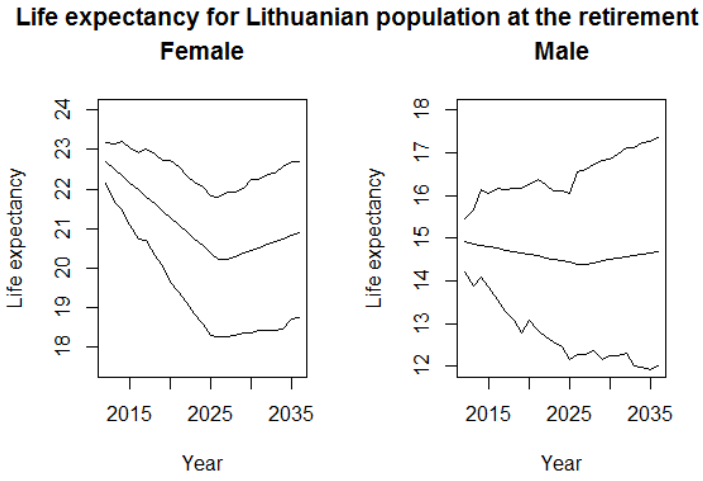

Figure 10.4 Life expectancy and 95% confidence intervals at the retirement age, projected by the Lee-Carter model fitted for Lithuanian population of age 60-95.

10.1. Fan charts

In this chapter fan charts for $e^{\nearrow}_{65}(2011,\dots,2035)$ based on residuals bootstrap (with B=300) are presented. Fan charts present central projections of the forecasted variables and the bounds around showing the probabilities that the variable will be between these ranges. In Figure 10.5 the heaviest shading presents central 10% prediction interval surrounded by 20%, 30%,…,90%. Darkness of the shadow can be interpreted as reflection of the likelihood outcome: the darker shade is – the more likely the outcome.

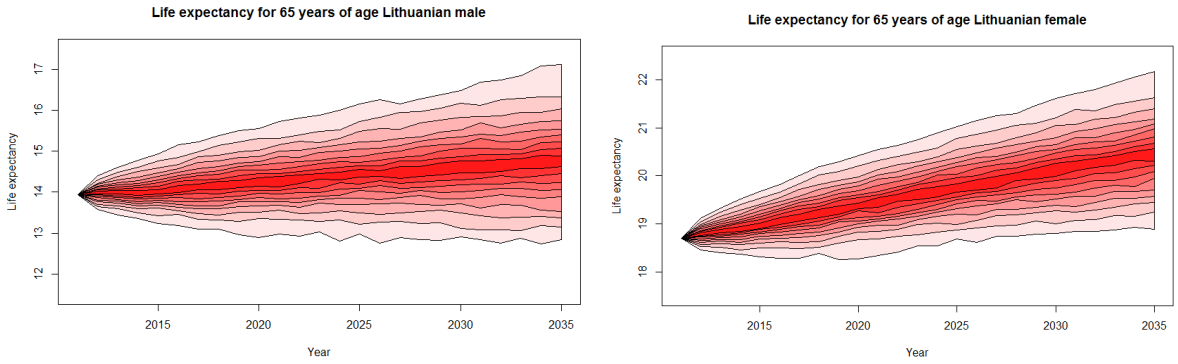

Figure 10.5 Fan chart for period life expectancies at age $e^{\nearrow}_{65}(2011,\dots,2035)$ for Lithuanian males and females

In Figure 10.5 6 Fan chart for period life expectancies for the Lithuanian males and females are presented. We can see that life expectancy for females is increasing, and the variation is smaller than the variation of the Lithuanian male, even the lower 90% prediction interval is increasing, while in the male case it is decreasing.

10.2. Actuarial Present Value

Now we investigate the impact of longevity risk on public pension plans by calculating the actuarial present value (APV) of a life annuity. We assume that at retirement each annuitant gets one unit of money per year, conditionally on survival, with a constant risk-free interest rate, of 2%. Survival is modelled by the Lee-Carter model fitted on years 1992-2011 and ages 60-95.

In Table 10.1 we can see the dynamics, how APV is expected to change in the Baltic States. We can see that at 2005-2015 APV for Lithuania and Latvia is very similar for both males and females, but for later years the difference between APV for Lithuanian and Latvian males are getting bigger. The reason for that are the slowly decreasing mortality rates for Lithuanian males. Much bigger slope of increasing life expectancies trend in Estonia is recognizable in the APV, which stands out compared to the other Baltic countries.

Table 10.2 presents APV for males and females at retirement age in Lithuania. We can see that the government's plan to equalize the retirement age for males and females has an impact on the APV at retirement age. Since the retirement age for females is increasing twice as fast as for males, the impact to the APV is also much bigger. These results suggests that increments in the retirement age are too big and cohorts in the period until the year 2026 suffer from these changes, because increase in the life expectancy is lower than increase in the retirement age.

To estimate APV we use R package '*lifecontingencies'* function *axn.* We assume that annuities are monthly.

Table 10.1 The actuarial present values for 65 years old ($APV_{65}^{(t)}$) male and female in Lithuania, Latvia and Estonia at year *t.*

Country	Lithuania		Latvia		Estonia	
Gender	Male	Female	Male	Female	Male	Female
$APV_{65}^{(2005)}$	3.28	5.95	3.31	5.98	4.28	7.09
$APV_{65}^{(2010)}$	3.37	6.18	3.51	6.31	4.71	7.58
$APV_{65}^{(2015)}$	3.45	6.40	3.7	6.62	5.12	8.05
$APV_{65}^{(2020)}$	3.53	6.61	3.9	6.93	5.55	8.48
$APV_{65}^{(2025)}$	3.60	6.82	4.1	7.24	5.97	8.9
$APV_{65}^{(2030)}$	3.68	7.02	4.3	7.54	6.39	9.29
$APV_{65}^{(2035)}$	3.76	7.21	4.51	7.83	6.81	9.65

Table 10.2 The actuarial present values for *x* years old ($APV_{x}^{(t)}$) male and female in Lithuania at year *t.*

APV for Female		APV for Male	
$APV_{61}^{(2014)}$	8.23	$APV_{63}^{(2014)}$	3.99
$APV_{62}^{(2017)}$	7.90	$APV_{63.5}^{(2017)}$	3.90
$APV_{63}^{(2020)}$	7.56	$APV_{64}^{(2020)}$	3.80
$APV_{64}^{(2023)}$	7.22	$APV_{64.5}^{(2023)}$	3.71
$APV_{65}^{(2026)}$	6.86	$APV_{65}^{(2026)}$	3.62

11. Conclusions and recommendations and summary

In this paper we use the Lee-Carter model and the Crains-Blake-Dowd model to investigate mortality patterns in Lithuania, Latvia and Estonia. We derived projections of life expectancies and actuarial present values. Structural break in mortality pattern was found in all three countries at the same year 1992, which may be the result of the re-established independence of the Soviet Union in the early 90's.

During analysis of the residuals we noticed that the residuals for the young ages are higher than for the older, and since the older ages are the most relevant in pension and annuity applications, in most of the further analysis we limited fitting period for the years 60-95. We also noticed that there may be some signs of a cohort effect, but these were too weak and therefore neglected.

For all three Baltic States, we made projections of the life expectancy 25 years ahead from the year 2012. To project life expectancy we used both the Lee-Carter model and the Cairns-Blake-Dowd model and compared predictions of these two methods. It turned out that in most of the cases the differences between the projections are very small, but in some cases life expectancies estimated by using the Cairns-Blake-Dowd approach are little lower than the Lee-Carter's.

Since the government, based on the aging population, in Lithuania decided to increase the retirement age gradually to 65 years in 2026, we checked the impact for cohorts retiring during this period. According to our estimations, life expectancy is increasing much slower than retirement age, and both male and female retiring in this period will be affected negatively. Moreover, we found that life expectancy in Lithuania (especially for males) are very volatile, and in the 95% confidence interval even negative projections are possible.

The methods and findings discussed in this Thesis illustrate just a small part of the research area in longevity risk. Considering the limitations of this Thesis, it is useful to propose recommendations for future work. Therefore, this section combines the limitations of studies as well as potential future research based on this Thesis.

As possible further research, models with cohort affects may be applied to ascertain whether these affects can be neglected as we assumed. We also do not consider continuous-time models for mortality, which are inherited from the credit risk and interest rate literature. As a possible literature for such investigation we refer Biffis *et. al* (2006).

Based on this research, further studies in managing and transferring the longevity risk may be done. For identification of the risks affecting an insurer we refer IAA (2001).

In the Section 4, the classical log-bilinear Lee-Carter (LC) model and the Poisson maximum likelihood estimation procedures are discussed. Later in the chapter, the LC model is fitted to the mortality data of the Baltic countries. Possible breakpoints were detected and tested in the outcome of the LC model, which led to reducing the fitting period.

Section 5 presents the Cairns-Blake-Dowd (CBD) model. Firstly, specification and the estimation procedure is discussed. Later in the chapter, the CBD model is fitted to the different age groups of the Baltic countries.

Undesirable irregularities where noticed in the age specific patterns of mortality change obtained by the Lee-Carter model. In order to get projected life tables with mortality varying smoothly across the ages, smoothing method P-spline were applied in the Section 6.

In the Section 7, we discussed the method of estimation of the residuals term. Model performance were assessed in terms of the randomness of the residuals and produce the graphical representation of the residual terms for different models fitted to different countries.

Section 8 examines the mortality projections. Firstly, the Lee-Carter model time index was transformed to achieve stationarity. Afterwards, the point forecast and variance were derived, and estimates for different countries and genders presented in the table. Subsequently, mortality projections for Cairns-Blake-Dowd model were derived.

Since it is impossible to derive the relevant prediction intervals analytically, in the Section 9, the prediction intervals were estimated based on the Poisson bootstrap procedure.

Section 10 covers estimation of life expectancies and presents projections for the Baltic countries. This section analyzes life expectancy in Lithuania at the retirement age, based on how retirement age will change every year until the year 2026. At the end of the section, the Fun charts for Lithuanian life expectancy are presented.

12. References

Baker, K., 2005. *Singular Value Decomposition Tutoria.* [Online] Available at: http://www.ling.ohio-state.edu/~kbaker/pubs/Singular_Value_Decomposition_Tutorial.pdf

Bell, W., 1997. Compaing and assessing time series methods for forecasting age-specific fertility and mortality rates. *Journal of Official Statistics,* Volume 13, pp. 279-303.

Biffis, E. & Denuit, M., 2006. A bidimentional approach to mortality risk. *Decisions in Economics and Finance,* Volume 29, pp. 71-94.

Blake, D., Cairns, A. & Dowd, K., 2006. Living with mortality: longevity bonds and other mortality-linked securities. *British Actuarial Journal,* pp. 153-228.

Brouhns, N., Denuit, M. & Keilegom, I. v., 2005. Bootstrapping the Poisson log-bilinear model for mortality forecasting. *Scandinavian Actuarial Journal,* Volume 3, pp. 212-224.

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

Cairns, A., 2007. *LifeMetrics Open Source R code for Stochastic Mortality Modelling.* [Online] Available at: http://www.macs.hw.ac.uk/~andrewc/lifemetrics/

Cairns, A., Blake, D. & Dowd, K., 2006. A two-factor model for stochastic mortality with parameter uncertainty: theory and calibration. *The Journal of Risk and Insurance,* pp. 687-718.

Cairns, A., Blake, D. & Dowd, K., 2006. Pricing death: frameworks for the evaluation and securitization of mortality risk. *ASTIN Bulletin,* pp. 79-120.

Carstensen, B., 2007. Age–period–cohort models for the Lexis diagram. *Statist. Med.,* Volume 26, pp. 3018-3045.

Chow, G. C., 1960. Tests of equality between sets of coefficients in two linear regressions. *Econometrica,* pp. 591-605.

Delwarde, A., Denuit, M. & Eilers, P., 2007. Smoothing the L-Carter and Poisson log-bilinear models for mortality forecasting. *Statistical Modeling,* Volume 7, pp. 29-48.

Durban, I., Currie, M. & Eilers, P., 2004. Smoothing and forecasting mortality rates. Volume 4, pp. 279-298.

Hyndman, R. J., 2014. *Package 'demography'.* [Online] Available at: http://cran.r-project.org/web/packages/demography/demography.pdf

IAA, 2001. *A global framework for insurer solvency assessment. Reasearch Report of the Insurer Solvency Assessment Working Party, International Actuarial Association..* sl:sn

Kingdom, J., 2014. *Longevity: Mortality improvement.* [Online] Available at: http://www.theactuary.com/archive/old-articles/part-6/longevity-3A-mortality-improvement/

Kwiatkowski, D., Phillips, P. C., Schmidt, P. & Shin, Y., 1992. Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root?. Volume 54, pp. 159-178.

Lee, R. & Carter, L., 1992. Modelling and foecasting US sex differentials in mortality. *International Journal of Forecasting,* pp. 393-411.

Phillips, P. C. & Perron, P., 1988. Testing for a Unit Root in Time Series Regression. Volume 75(2), pp. 335-346.

Pitacco, E., Denuit, M., Haberman, M. & Olivieri, A., 2009. *Modelling Longeviy dynamics for Pensions and Annuity Business.* sl:Oxford.

Prinz, A., Steenge, A. E. & Schmidt, J., 2007. *The Rules of the Game: Institutions, Law and Economics.* Berlin: sn

Schwarz, G. E., 1978. Estimating the dimension of a model. *Annals of Statistics,* Volume 6 (2), pp. 461-464.

Tuljapurkar, S., Li, N. & Boe, C., 2000. Universal Pattern of Mortality Decline in the G7 Countries. *Nature,* Volume 405, pp. 789-792.

Wilmoth, J., Andreev, K., Jdanov, D. & Glei, D., 2007. *www.mortality.org,* s.l.: s.n.

Wilmoth, J. R., Shkolnikov, V. & Barbieri, M., 2000. *The Human Mortality Database.* [Online] Available at: http://www.mortality.org/

13. Appendix

Download data from the HMD.

```
library(demography)
LTUDemo<-hmd.mx(country="LTU", username="jakstas.gintautas@gmail.com",
                password="*****", label="Lithuania")
load(file="mortalityDatasets.RData")
```

The Lee-Carter model estimation function described in the section 4.1.

```
LC1=function(mtx,xv,yv){
  n=length(xv)  # number of ages
  m=length(yv)  # number of years
    # Stage 1. initialize parameter vectors
    alpha=(1:n)*0
    beta=(1:n)*0
    kappa=(1:m)*0
    beta[1]=1
    dD=1
    while(abs(dD) > (10^(-6))) {
    # Stage 2 optimize over the alpha
    for(j in 1:n)
    {
      s1=sum(log(mtx[,j])-alpha[j]-beta[j]*kappa)
      s2=yv[length(yv)]-yv[1]+1
      alpha[j]=alpha[j]+s1/s2
    }
    #apply constrain on alpha's
    alpha=alpha+beta*mean(kappa)
    #estimate objection function
    mhat=mtx*0
    for(i in 1:m)
    {
      mhat[i,]=(alpha+beta*kappa[i])
    }
    D=sum((log(mtx)-mhat)^2)
    cat(D," ->")

    # Stage 3 optimize over the kappa
    for(i in 1:m)
    {
      s1=sum(beta*(log(mtx[i,])-alpha-beta*kappa[i]))
      s2=sum((beta)^2)
      kappa[i]=kappa[i]+s1/s2
    }
    #apply constrain on kappa's
    kappa=(kappa-mean(kappa))*sum(beta)
    #estimate objection function
    mhat=mtx*0
    for(i in 1:m)
    {
      mhat[i,]=(alpha+beta*kappa[i])
    }
    Du=sum((log(mtx)-mhat)^2)
    cat(Du," ->")

    # Stage 4 optimize over the beta(x)
    for(j in 1:n)
    {
      s1=sum(kappa*(log(mtx[,j])-alpha[j]-beta[j]*kappa))
      s2=sum(kappa*kappa)
      beta[j]=beta[j]+s1/s2
    #  if(beta[j]<0){beta[j]=0} #additional constrain for beta(x) to ensure that
    #  beta(x)'s are non-negative
    }
    #apply constrain on beta's
    beta=beta/sum(beta)
    #estimate objection function
    mhat=mtx*0
```

```
    for(i in 1:m)
    {
      mhat[i,]=(alpha+beta*kappa[i])
    }
    Du=sum((log(mtx)-mhat)^2)
    cat(Du," ->")
    dD=D-Du
    cat(dD,"\n")
  }           # end while loop
    list(ax=alpha,bx=beta,kt=kappa,x=xv,y=yv)
}
```

The Lee-Carter model estimation function described in the section 4.2.

```
LM2=function(mtx,xv,yv,etx,dtx)
  {
  n=length(xv)  # number of ages
  m=length(yv)     # number of years

# Stage 1. initialize parameter vectors
  alpha=(1:n)*0
  beta=(1:n)*0
  kappa=(1:m)*0
  for(j in 1:n)
  {
    beta[j]=1/n
  }
kappa=(m:1)-(m+1)/2
# Start loop for 2-4 stages
 dD=1
  while(abs(dD) > (10^(-6)))
  {
# Stage 2 optimize over the alpha
    for(j in 1:n)
    {
      s1=sum(dtx[,j]-etx[,j]*exp(alpha[j]+beta[j]*kappa))
      s2=sum(etx[,j]*exp(alpha[j]+beta[j]*kappa))
      alpha[j]=alpha[j]+s1/s2
    }
    #apply constrain on alpha's
    alpha=alpha+beta*mean(kappa)
    #estimate objection function
    mhat=mtx*0
    mhat2=mtx*0
    for(i in 1:m)
    {
      mhat[i,]=alpha+beta*kappa[i]
      mhat2[i,]=exp(alpha+beta*kappa[i])
    }
    D=sum(dtx*mhat-etx*mhat2)
    cat(D," ->")

# Stage 3 optimize over the kappa
    for(i in 1:m)
    {
      dv=dtx[i,]   # actual deaths
      ev=etx[i,]   # exposure
       s1=sum(beta*(dv-ev*exp(alpha+beta*kappa[i])))
           s2=sum(ev*exp(alpha+beta*kappa[i])*beta*beta)
      kappa[i]=kappa[i]+s1/s2
```

```
    }
    #apply constrain on kappa's
    kappa=(kappa-mean(kappa))*sum(beta)
    #estimate objection function
    mhat=mtx*0
    mhat2=mtx*0
    for(i in 1:m)
    {
      mhat[i,]=alpha+beta*kappa[i]
      mhat2[i,]=exp(alpha+beta*kappa[i])
    }
    Du=sum(dtx*mhat-etx*mhat2)
cat(Du," ->")

# Stage 4 optimize over the beta
    for(j in 1:n)
    {
      dv=dtx[,j]  # actual deaths
      ev=etx[,j]  # exposure
       s1=sum(kappa*(dv-ev*exp(alpha[j]+beta[j]*kappa)))
           s2=sum(ev*exp(alpha[j]+beta[j]*kappa)*kappa*kappa)
      beta[j]=beta[j]+s1/s2
      #posible restriction for beta's to be non-negative.
       #if(beta[j]<0){beta[j]=0}
    }
    #apply constrain on beta's
    beta=beta/sum(beta)
    #estimate objection function
    mhat=mtx*0
    mhat2=mtx*0
    for(i in 1:m)
    {
      mhat[i,]=alpha+beta*kappa[i]
      mhat2[i,]=exp(alpha+beta*kappa[i])
    }
    Du=sum(dtx*mhat-etx*mhat2)
    cat(Du," ->")
    dD=D-Du
    cat(dD,"\n")
  }           # end while loop
  list(ax=alpha,bx=beta,kt=kappa,x=xv,y=yv)
}
```

For the SVD method R function *lca* is used.

```
library(demography)
lca(LTUDemo,series="male",max.age=95,ages=seq(60,95,1),years=seq(1992,2011,1),inte
rpolate = TRUE,adjust="none")#adjust=”none” is for estimating the method described
in section 4.1 and adjust=”dt” – for the method described in section 4.2. For more
details use ?lca.
```

To estimate the CBD model we use function provided by *lifemetrics,* see (Cairns, 2007).

Mortality projections estimation for the Lee-Carter model (Table 8.3):

```
####estimate d
dhat<-function(kt){(kt[length(kt)]-kt[1])/(length(kt)-1)}
####estimate sigma
sigmahat<-function(kt,d){
  S=0
  for (i in 2:length(kt)){
```

```
    S=S+(kt[i]-kt[i-1]-d)^2
  }
  S/(length(kt)-1)
}
#estimate s.e. of d
se<-function(kt){sd(kt)/length(kt)}
```

Mortality projections estimation for the CBD model (Table 8.4 and Table 8.5):

```
dhati<-function(kt){(kt[length(kt)]-kt[1])/(length(kt)-1)}
####estimate sigma
sigmahati<-function(kt,d){
  S=0
  for (i in 2:length(kt)){
    S=S+(kt[i]-kt[i-1]-d)^2
  }
  S/(length(kt)-1)
}
sigmahat12<-function(kt1,kt2,d1,d2){
  S=0
  for (i in 1:(length(kt1)-1)){
  for (j in 1:(length(kt2)-1)){
    S=S+(kt1[i+1]-kt1[i]-d1)*(kt2[j+1]-kt2[j]-d2)
  }
  }
  S/(length(kt1)-1)
}
sigmahat12<-function(kt1,kt2,d1,d2){
        ((sum(diff(kt1)-d1))*(sum(diff(kt2)-d2)))/(length(kt1)-1)
}
```

Estimation of the residuals for the Lee-Carter model.

```
n=yt[length(yt)]-yt[1]-1 # number of years
m=xt[length(xt)]-xt[1] # number of ages
mhat=mtx*0 #create mhat as a copy of mortality rates matrix
for(i in 1:length(yt))
{
  mhat[i,]=exp(ax+bx*kt[i])
}
epsilon=log(mhat)-log(mtx)
residual=epsilon/(sqrt(1/(n*m)*sum(epsilon^2)))
```

Estimation of the residuals for the CBD model.

```
n=yt[length(yt)]-yt[1]+1 # number of years
m=xt[length(xt)]-xt[1]-1 # number of ages
mhat=mtx*0 #create mhat as a copy of mortality rates matrix
for(i in 1:length(yt))
{
  mhat[i,]=exp(ax+bx*kt[i])
}
epsilon=log(mhat)-log(mtx)
residual=epsilon/(sqrt(1/(n*m)*sum(epsilon^2)))
```