Fitting and Forecasting Mortality Rates for Nordic Countries Using the Lee-Carter method

Marie Claire Koissi[†⋆] Arnold Shapiro [⋆] Göran Högnäs [†]

[†]Department of Mathematics, Åbo Akademi University, Finland.

[⋆] Smeal College of Business, Penn State University, USA

Abstract

The Lee and Carter (LC) method is getting widely adopted for long-run forecasts of age specific mortality rates. That popularity is first due to the model simplicity. Moreover, the LC model has given successful results for various countries (e.g. U.S., Chile, G7 countries). However, some difficulties arose when using population data from certain countries such as UK and Australia. In the present paper, the LC model is applied to data from four Nordic countries: Denmark, Finland, Norway and Sweden. The Singular Value Decomposition (SVD), the Weighted Least Square (WLS) and the Maximum Likelihood estimate (MLE) are used. These approaches give satisfactory results. The appropriate fitting period needs, however, to be well chosen. The properties of the model's parameters are studied using a bootstrap simulation.

Key words: Lee-Carter model; Mortality Forecasting; Nordic countries.

1 Introduction

"A population forecast is a projection in which certain assumptions are considered to yield a realistic picture of the future development of a population" (United Nations [20]). These assumptions concern the expected trend in fertility, mortality and migration. In the present work, we are interested in modeling and forecasting mortality and life expectancy, which describes the average length of life within a population under the assumption of constant living conditions.

Traditionally, a parametric curve is fitted to annual mortality rates. Then, graduation is used to obtain projected rates. The attempt to find the appropriate mortality curve has a long history in demography and actuarial sciences. The most ancient known formula is due to the French mathematician De Moivre (1725) who wrote the survival function in the form $S(x) = 1 - x/w$ (w is the maximum age. The deaths are assumed uniformly distributed.). Later on, the British actuary Gompertz (1825) observed that for the age grouping of between 20 and 60 years, the force of mortality increased almost exponentially with age ($\mu(x) = ae^{bx}$, $a, b \in \mathbb{R}$). Gompertz observation was based on population data from England, Sweden and France. Since then, many studies confirmed that the Gompertz law even works for various other countries, although lack of fit was observed for particular ages (Olshansky et al. ,1997 [16]). In order to correct the weakness of that law (especially at old age), many expressions was proposed: amongst others Makeham (1867), Thiele (1872), Wittstein (1883), Pearson (1895), Perks (1932), Brilinger (1961), Heligman and Pollard (1980), Hannerz (1999).

However, studies conducted in the last twenty years (Stoto, 1983 [18]) revealed many errors in the forecasts. Keilman (1998 [8]) reported that the earlier forecasts have missed some important events such as the post second World War baby-boom and the decline in fertility in Greece and Spain after 1985. Old-age mortality decline was also underestimated and increases in life expectancy under-projected.

Stochastic models provide a realistic estimate of the expected error of forecasts (Alho, 1998 [1]). In their 1992' paper (1992 [10]), Lee and Carter (henceforth LC) presented a stochastic method based on cohort component approach to model and forecast the age specific mortality rates of the US population. Despite its simplicity, the LC model has proved to give good results in fitting mortality from diverse countries: Canada (Lee and Nault, 1993 [12]), Chile (Lee and Rofman, 1994 [13]), Japan (Wilmoth ,1996 [22]), the seven most economically developed nations (G7) (Tuljapurkar et al., 2000 [19]), Belgium (Brouhns et al., 2002 [6]).

Some difficulties arose when applying the LC method to population data from certain countries: the assumption of invariant age component over time was violated when applied to Australia (Booth et al., 2002 [4]). The method was also less straightforward when using historical data from U.K. that presented a strong cohort effect (Renshaw and Haberman, 2003 [17]). To ensure good model performance, alternative approaches to the original Singular Value Decomposition (SVD) were proposed: Wilmoth (1993 [21]) proposed the Weighted Least Square (WLS) approach to obtain an estimation of the LC model parameters. The Maximum Likelihood Estimation (MLE) gives optimal solution of the LC equation under a Poisson model (for details see Wilmoth, 1993 [21]).

This paper aims at comparing tree different methods of estimating the model's parameters: the Singular Value Decomposition, the (Wilmoth-) Weighted Least Square method and the Maximum Likelihood Estimate. Data from the following four Nordic countries are used: Denmark, Finland, Norway and Sweden. Section 2 presents the LC method and briefly describes the three methodologies for deriving the model's parameters. In section 3, forecasted death rates and life expectancies are presented. Section 4 is devoted to a simulation study made to compare the performance of the different estimation methods. Section 5 gives a summary of the results.

2 The data and method

2.1 The data

The "Human Mortality Database" (www.mortality.org) is the main source of data. However, in the case of Finland, data published by Statistics Finland are also used. The data, given by sex and by five-year age group, cover the following period for each country: Denmark (1921-1999), Finland (1878-1999), Norway (1846-1999) and Sweden (1861-1999).

2.2 The Lee-Carter method

Lee and Carter (1992 [10]) wrote the logarithm of the matrix of central death rates $(m_{x,t})$ as followed

$$\ln(m_{x,t}) = a_x + b_x k_t + \varepsilon_{x,t}. \tag{2.1}$$

The parameter a_x describes the average age-specific pattern of mortality and k_t represents a time-trend index of general mortality level. Decline in mortality at particular age x is captured by b_x. For model identification, the following constraints are imposed: $\sum_t k_t = 0$ and $\Sigma_x b_x^2 = 1$. Thus the parameter vector a_x is computed as the average over time of the logarithm of the central death.

2.3 Estimation approaches

The Singular Value Decomposition

In their original paper Lee and Carter (1992) computed the parameters of equation 2.1 using the Singular Value Decomposition (SVD): first, the parameter vector a_x is computed as the average over time of the logarithm of the central death. Then, the Singular Value Decomposition, applied to matrix $Z = \ln(m) - \hat{a}$, produces the matrices $PdQ' = SVD(Z_{xt}) = d_1 P_{x1} Q_{t1} + ... + d_X P_{xX} Q_{tX}$. Approximation to the first term gives the estimates $\hat{b}_x = P_{x1}$ and $\hat{k}_t = d_1 Q_{t1}$. The simplicity of the SVD computation is a reason for its popularity. However, as noticed by Alho (2000 [2]), that method is not optimal and the Maximum Likelihood Estimation (MLE) may produce better solutions.

The Maximum Likelihood Estimation

The random variable $D_{x,t}$, representing the number of deaths within the people of age x at time t, can be satisfactorily modeled by a Poisson distribution with parameter $\lambda_{x,t} = m_{x,t}E_{x,t}$ (Brillinger 1986 [5]). In this expression, $E_{x,t}$ denotes the exposure to risk at age x and time t. The parameters of the LC equation are found by maximizing the corresponding full likelihood function

$$l = \sum_x \sum_t [d_{x,t} \ln(\lambda_{x,t}) - \lambda_{x,t} - \ln(d_{x,t}!)].$$

under the condition $m_{x,t} = e^{\hat{a}_x + \hat{b}_x \hat{k}_t}$. A third method consists in assigning to each cell of the matrix of death rates (m_{xt}), a weight equal to the observed number of deaths (d_{xt}) for age x at time t.

The Weighted Least Square

The Weighted Least Square (WLS) estimates of the LC parameters a_x, b_x and k_t are obtained by minimizing the following squared errors (see Wilmoth, 1993 [21], Carter & Prskawetz, 2001 [7])

$$\sum_{x=1}^{X} \sum_{t=1}^{T} d_{xt}[\ln(m_{x,t}) - a_x - b_x k_t)]^2. \tag{2.2}$$

The WLS is worth using, since it has better properties than the SVD method (Lee, 2000 [9]). The three approaches are used on data from the Nordic countries.

2.4 Fitting the LC model

After preliminary computations, the period from 1955 to 1999 was chosen to ensure the linear decreasing trend in the mortality index k. Figures 1 and 2 show the estimates $\hat{a}_x$, $\hat{b}_x$ and $\hat{k}_t$ with the SVD, the MLE and the WLS. The following comments can be made:

- No variation is observed for the parameter vector $\hat{a}$ with the three approaches.

- For parameter $\hat{b}$, the values obtained through WLS and the MLE are quite identical.

- The mortality index $\hat{k}$ has a common (almost) linear decreasing trend in the four countries with the three methods. The WLS and the MLE also give quite identical values.

Corresponding fitted death rates are computed. Then, abridged life tables are constructed, from which life expectancies are derived. Figure 3 shows the observed and fitted life expectancies at birth, both sexes combined, for the four countries. The fitting errors, displayed on table 1 and figure 4 prove that the WLS and the MLE provide better fit. The (small) error magnitude shows that the three approaches however give good results.

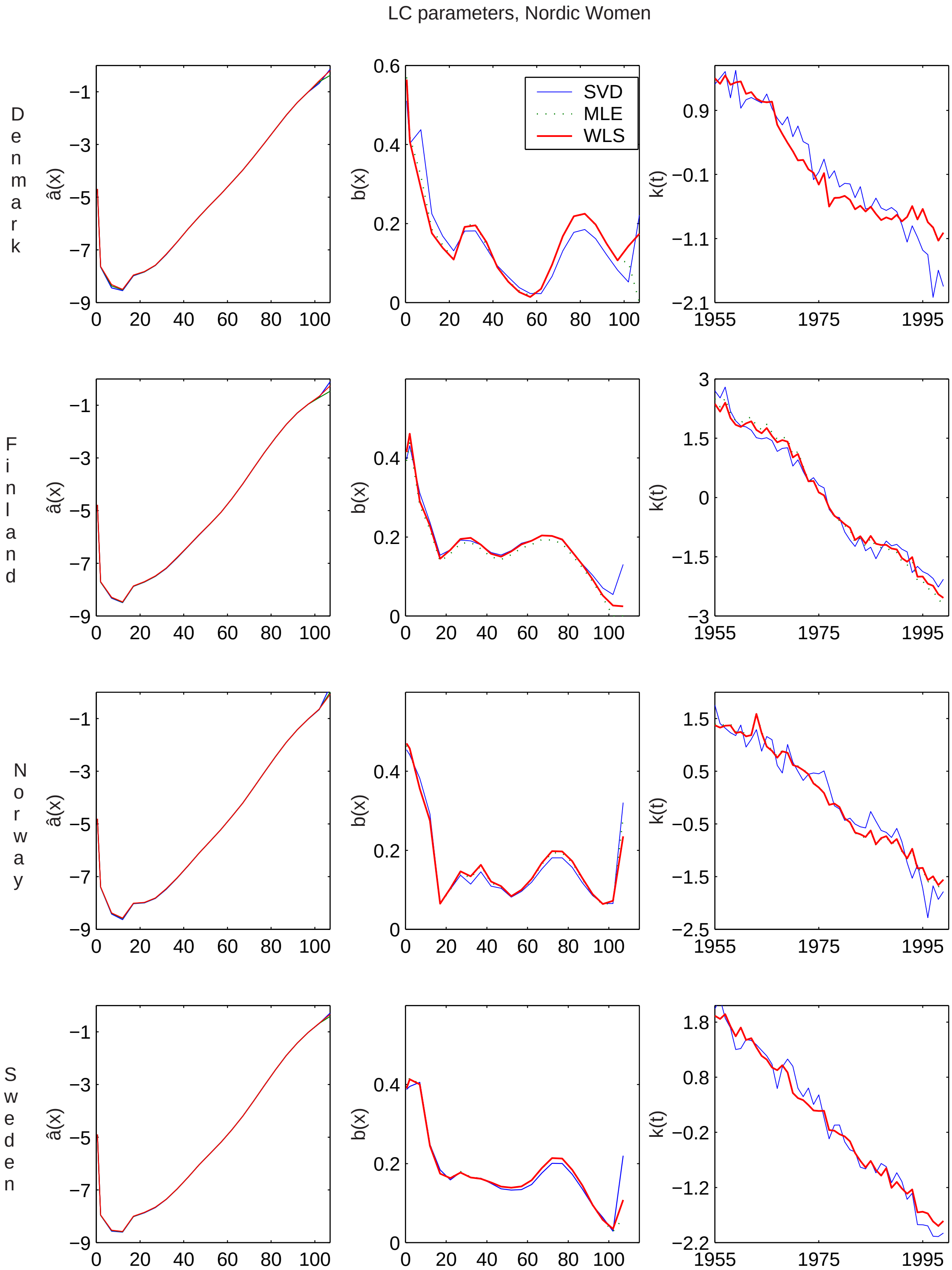

Figure 1: LC parameters, Women

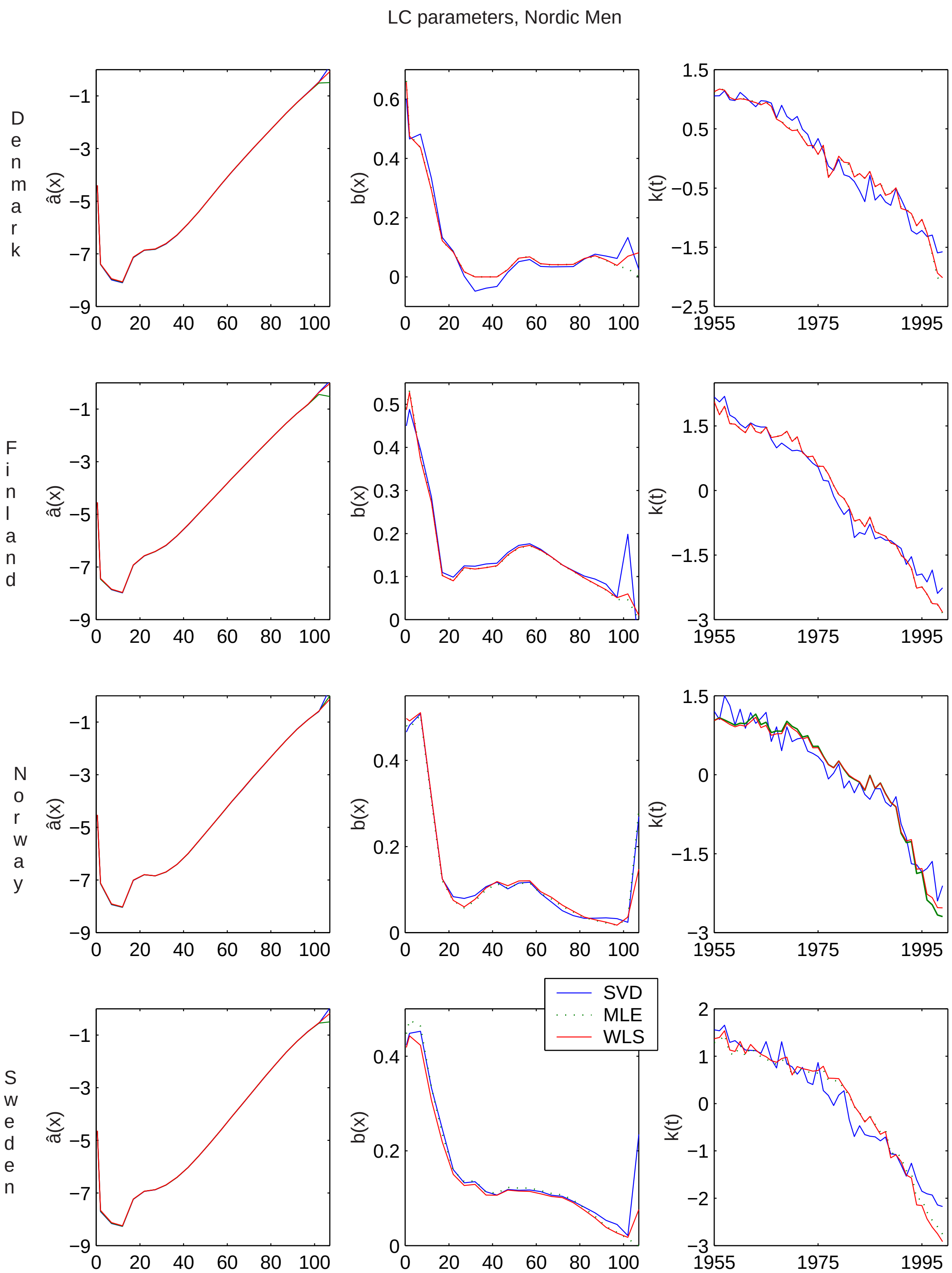

Figure 2: LC parameters, Men

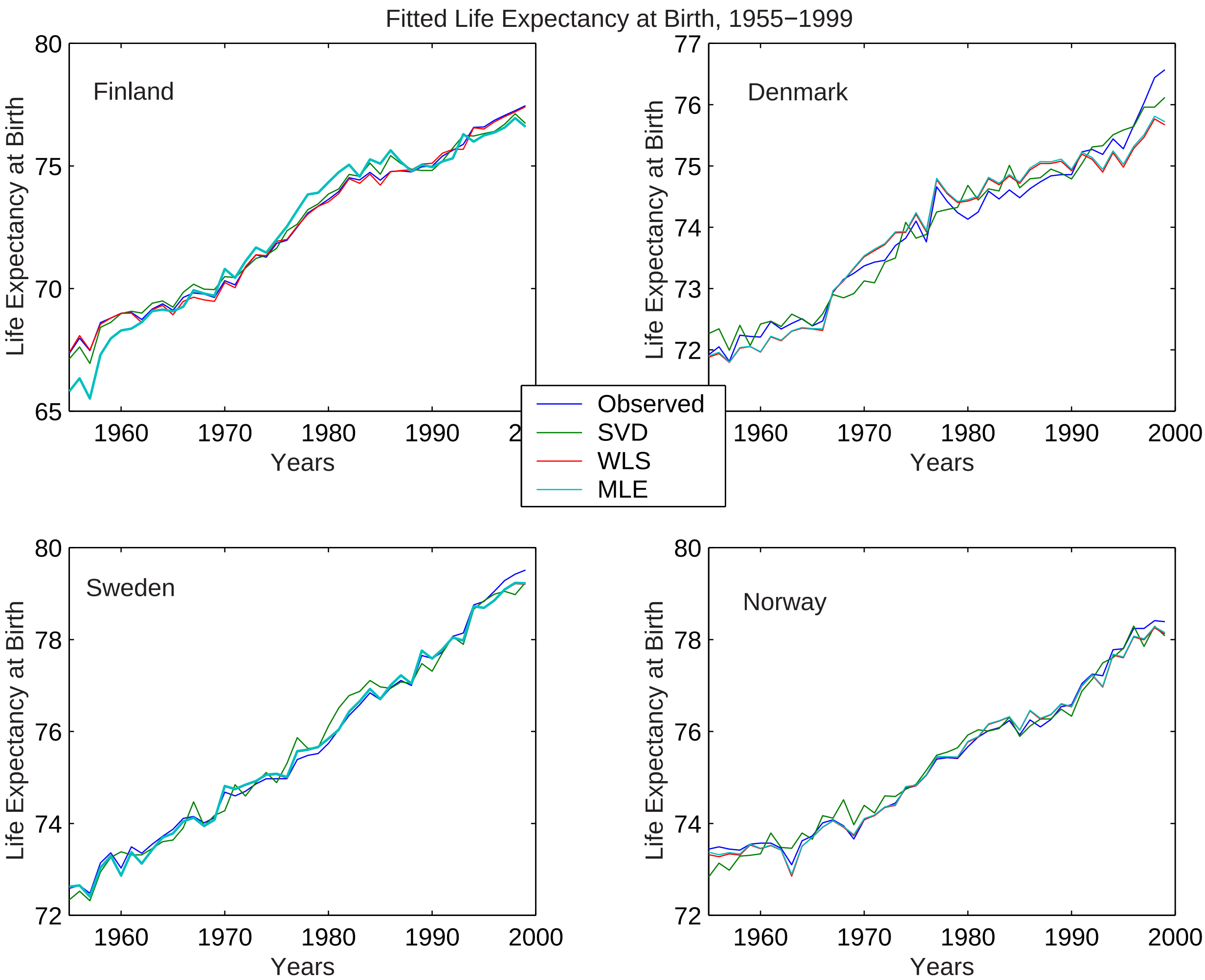

Figure 3: Fitted vs. Observed Life Expectancy at Birth, Both sexes combined, 1955-1999.

Countries	Mean Error			MSE		
	SVD	WLS	MLE	SVD	WLS	MLE
Denmark	-0.0143	0.0240	0.0019	0.0519	0.0707	0.0682
Finland	-0.0269	0.0391	0.0168	0.0778	0.0099	0.0082
Norway	-0.0062	0.0373	0.0222	0.0530	0.0144	0.0126
Sweden	-0.0038	0.0227	0.0105	0.0547	0.0139	0.0135

Table 1: Mean Error in Fitted Life Expectancy at Birth, Both sexes combined.

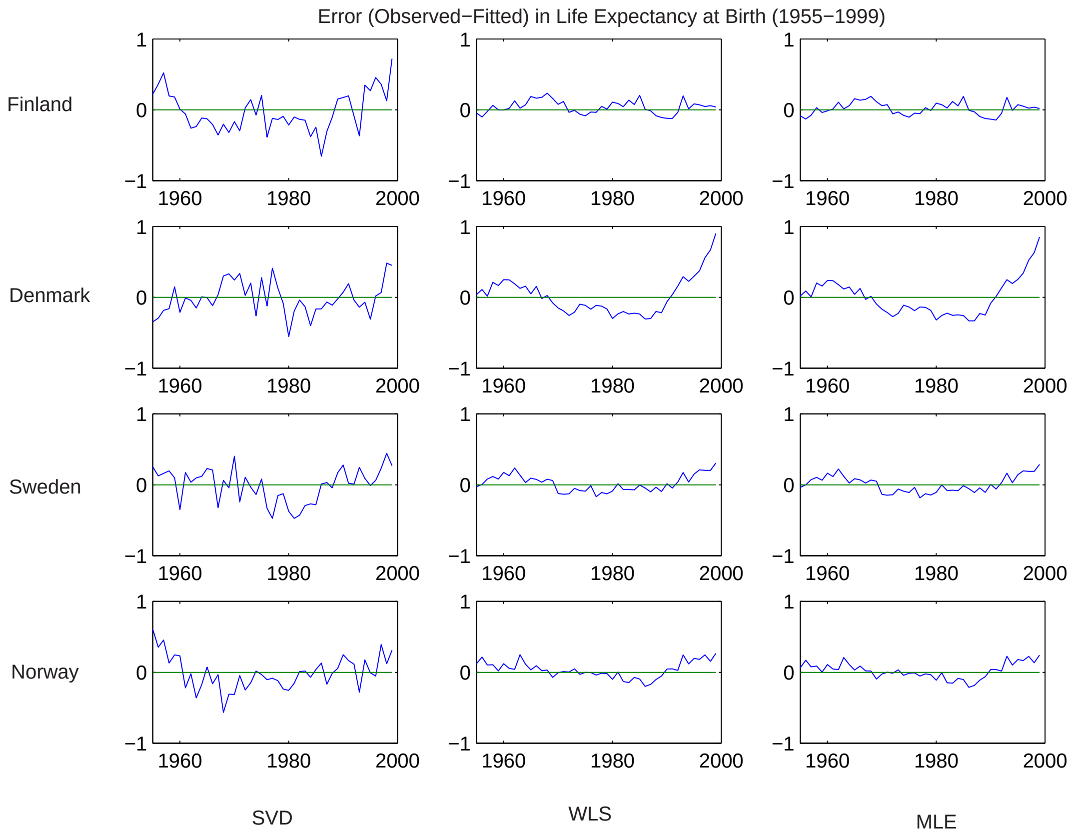

Figure 4: Error in Life Expectancy at Birth, Both sexes combined, 1955-1999.

3 Forecasted death rates and life expectancies

One good property of the LC approach is that, once the data are fitted to the model and the values of the vectors $\hat{a}$, $\hat{b}$ and $\hat{k}$ are found, only the mortality index $\hat{k}_t$ needs to be predicted. Appropriate ARIMA models are obtained for each time series $\hat{k}_t$. An ARIMA(0,1,0) was suitable for each $\hat{k}_t$ series, sexes combined. The partial autocorrelation function (PACF) and the Q-Q plots were used to check for the stationarity and normality of the series of first order difference. Figure 5 shows the fitted and forecasted mortality index, obtained with the WLS approach.

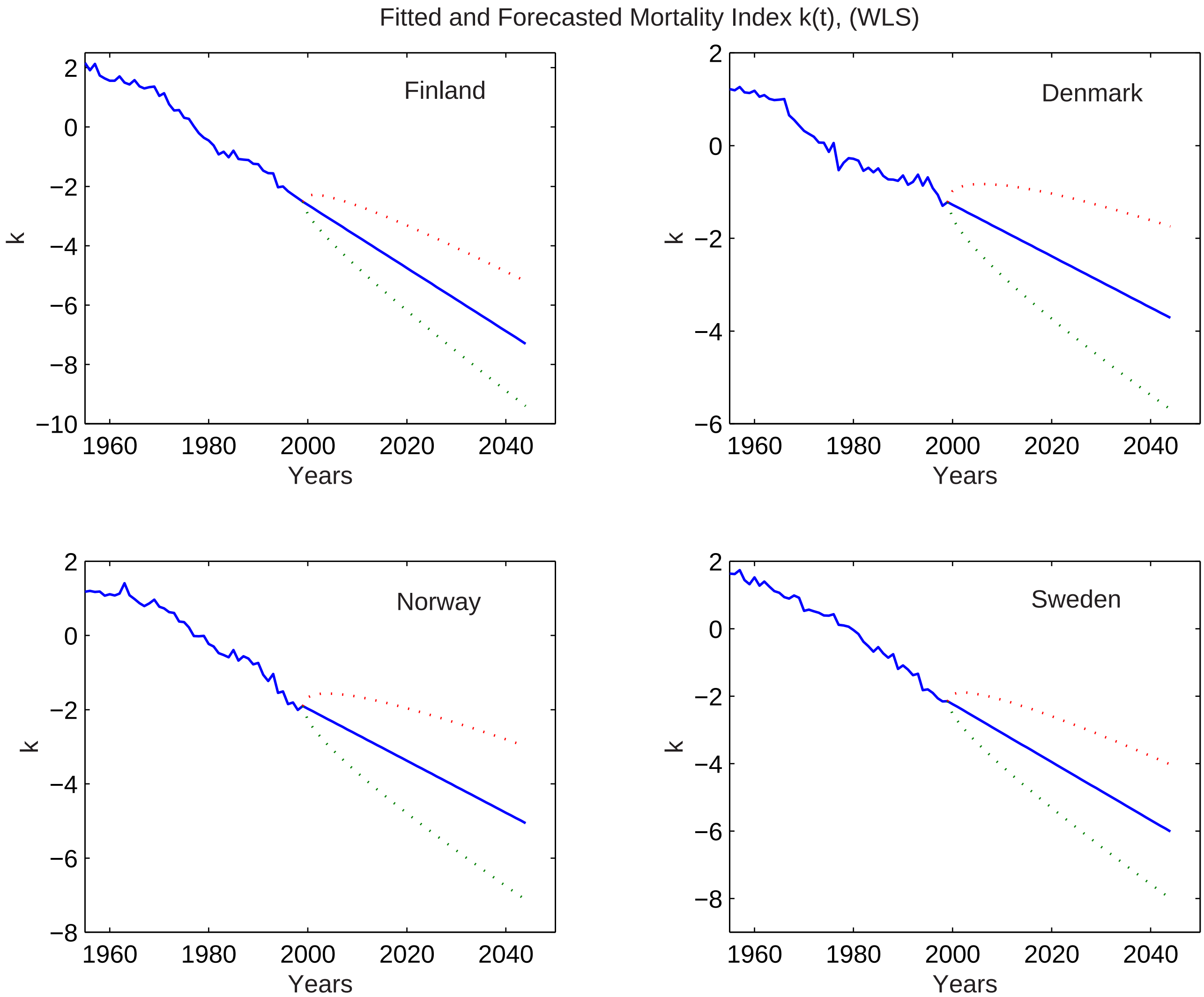

Figure 5: Fitted and forecasted (WLS) values of $\hat{k}$ (Both sexes combined), with 95% prediction lines.

Forecasted mortality index are used to compute forecasted death rates

$$\ln(m_{x,1999+t}) \approx \hat{a}_x + \hat{b}_x \hat{k}_{1999+t}. \tag{3.1}$$

Then, forecasted life expectancies are deduced from corresponding abridged life tables. Figure 6 shows the fitted and forecasted life expectancies at birth, obtained with the WLS. For selected years, the results are compared with the values published by the Population Division of US Census Bureau (available online), see table 2. For Denmark and Norway, the official forecasts are higher than the ones we obtained in all the three methods. In the case of Finland, the values we computed are larger than the ones obtained by the US Census Bureau. For Sweden there is no strict order. Although there are differences in the forecasted life expectancy, the gap does not exceed one year (except in the case of Denmark where it reaches almost 4 years).

Countries		SVD	WLS	MLE	Official		SVD	WLS	MLE	Official
Denmark	*2010*	76.87	76.41	76.46	78.47	*2020*	77.54	77.06	77.12	79.94
	2030	78.19	77.69	77.75	81.12	*2040*	78.84	78.31	78.38	82.06
Finland	*2010*	78.69	79.40	79.43	79.13	*2020*	80.38	81.10	81.13	80.47
	2030	81.99	82.70	82.73	81.54	*2040*	83.53	84.20	84.25	82.39
Norway	*2010*	79.11	79.08	79.10	80.08	*2020*	80.00	79.91	79.93	81.23
	2030	80.86	80.71	80.73	82.14	*2040*	81.69	81.48	81.5	82.87
Sweden	*2010*	80.67	80.56	80.58	80.97	*2020*	81.91	81.74	81.76	81.94
	2030	83.10	82.86	82.88	82.71	*2040*	84.24	83.93	83.95	83.31

Table 2: Comparison of Life Expectancy at Birth, Both sexes combined.

In next section, the properties of the estimates of the model parameters provided by the three estimation methods are compared on the basis of a simulation study.

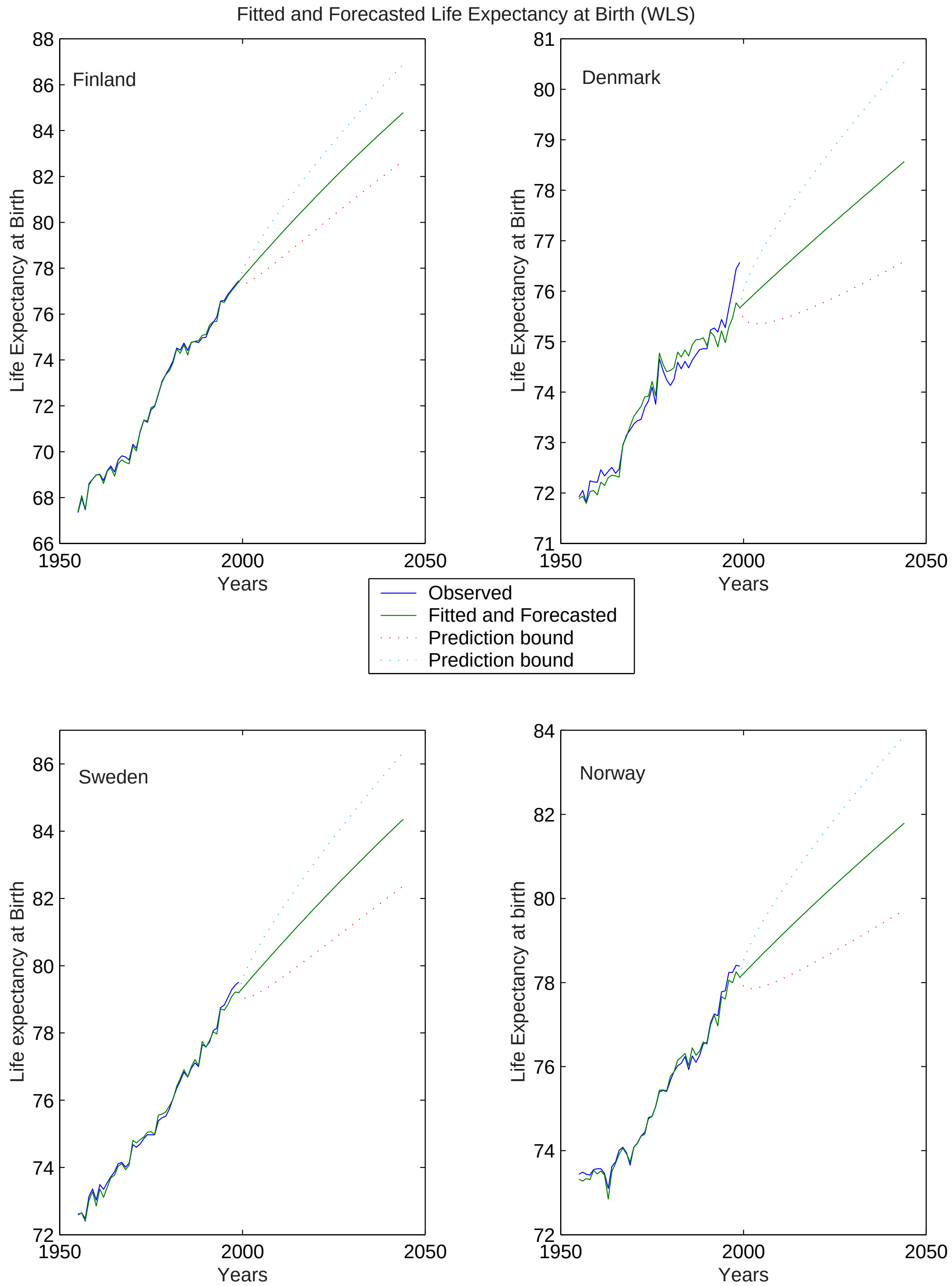

Figure 6: Fitted and Forecasted Life Expectancy at Birth, WLS, Both sexes combined (95% prediction lines are in dash).

4 Simulation studies

The idea behind the simulation is to check if a small fluctuation in the error term leads to an important change in the values of the model parameters. From the error matrix

$$\hat{\varepsilon}_{x,t} = \ln(m_{x,t}) - (\hat{a}_x I_t + \hat{b}_x \hat{k}_t), \tag{4.1}$$

new random matrices $[\hat{\varepsilon}_{x,t}]_i$ of errors are generated using **bootstrap** technique. Then, new matrices of death rates are obtained:

$$[\ln(m_{x,t})]_i = \ln(m_{x,t}) - [\hat{\varepsilon}_{x,t}]_i, \tag{4.2}$$

where $[\hat{\varepsilon}_{x,t}]_i$ $(i = 1, ..., R)$ is the error matrix obtained after the i^{th} re-sampling. One hundred bootstraps re-samplings are made ($R = 100$), from which one hundred values of the parameters $\hat{a}$, $\hat{b}$ and $\hat{k}$ were obtained (see figures 7-9). The starting point is the matrix of death rates for Finland, 1955-1999, both sexes combined. The sum of the mean squared errors (MSE), in table 3, shows that the MLE gives better result for the parameters a and b, while the smaller MSEs for k are obtained with the SVD. Figure 10 also confirms those findings. Additionally, a non-negligible advantage of the SVD is its short computation time.

Sum of MSE	$\hat{a}$	$\hat{b}$	$\hat{k}$
MLE	0.0055	0.0015	1.2857
SVD	0.0063	0.0032	0.6168
WLS	0.0073	0.0021	2.1068

Table 3: Sum of the Mean Square Errors (MSE)

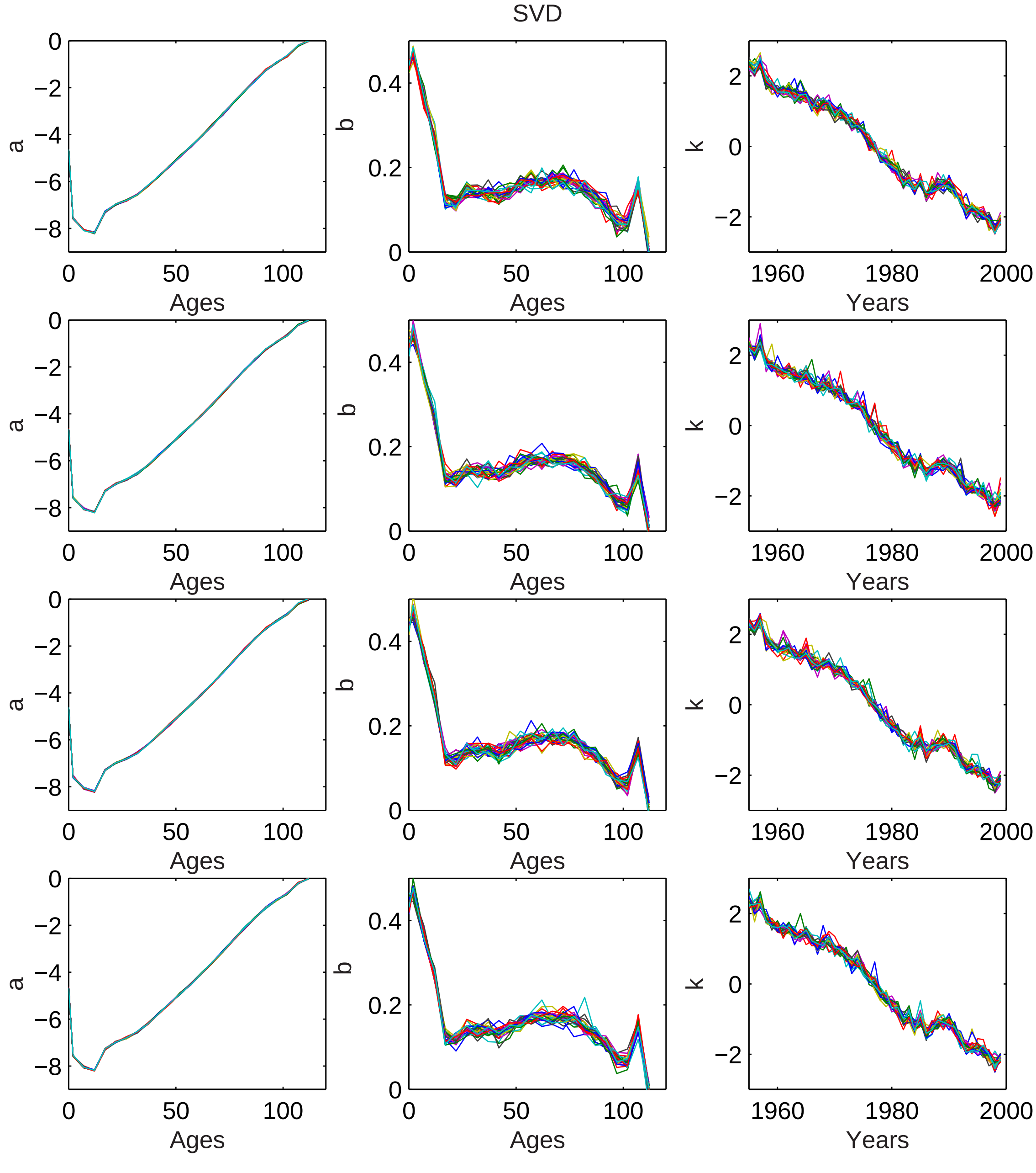

Figure 7: SVD, 100 Bootstraps resampling. The 1-25th simulations are on the first row, the 26-50th are on the 2nd row, the 51-75th on the 3rd row and the 76-100th simulations are on the 4th row.

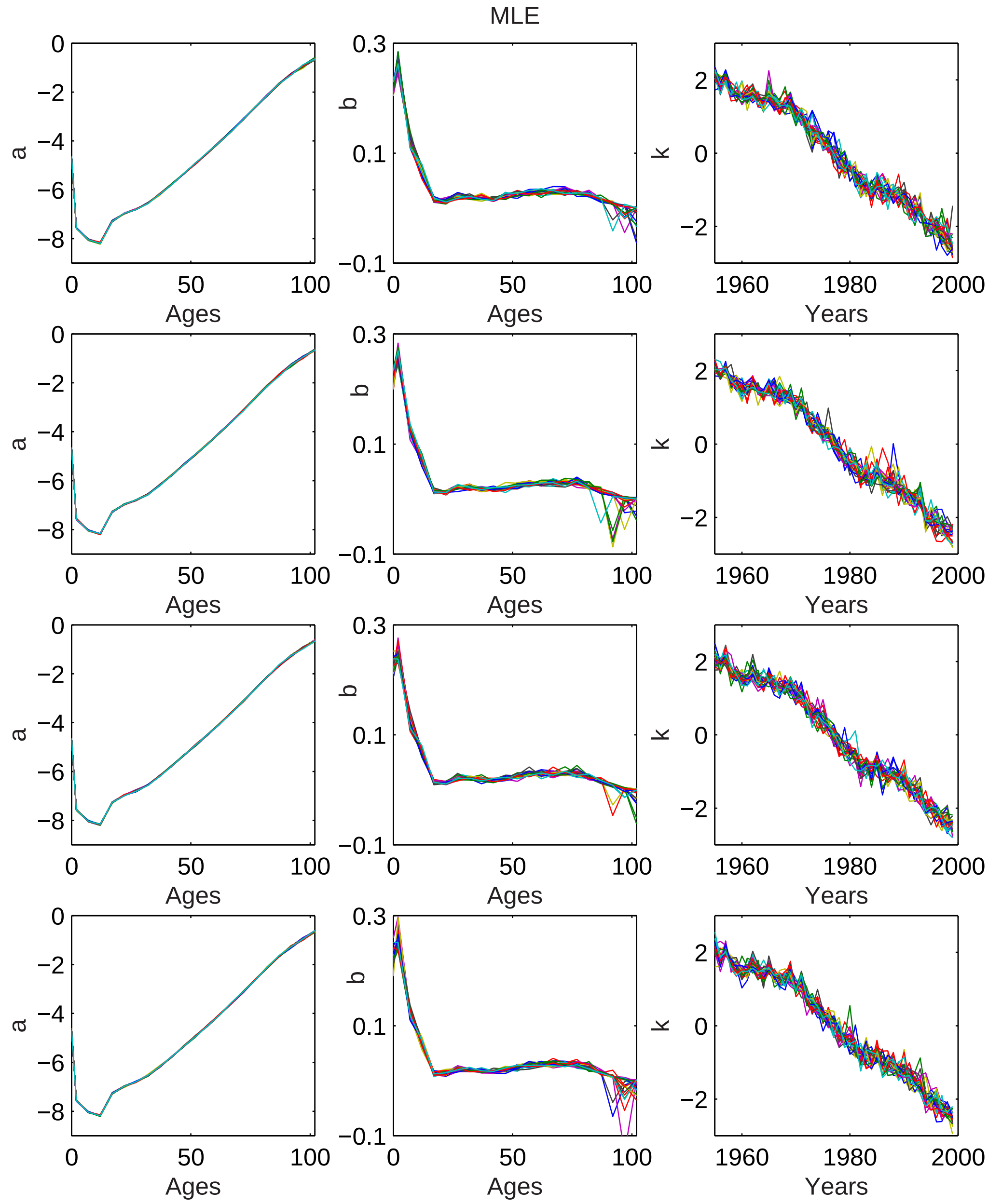

Figure 8: MLE, 100 Bootstraps resampling. The 1-25th simulations are on the first row, the 26-50th are on the 2nd row, the 51-75th on the 3rd row and the 76-100th simulations are on the 4th row.

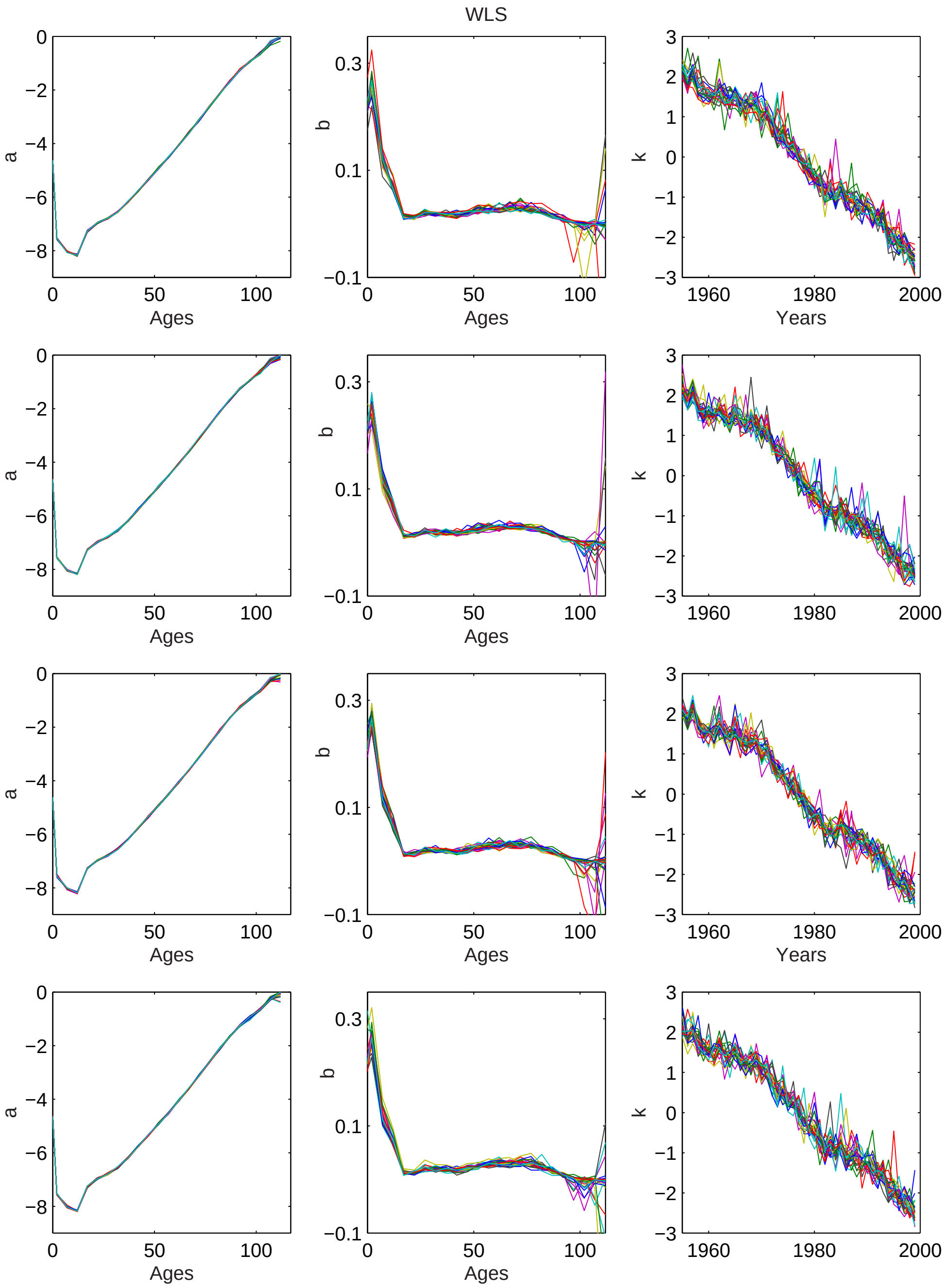

Figure 9: WLS, 100 Bootstraps resampling. The 1-25th simulations are on the first row, the 26-50th are on the 2nd row, the 51-75th on the 3rd row and the 76-100th simulations are on the 4th row.

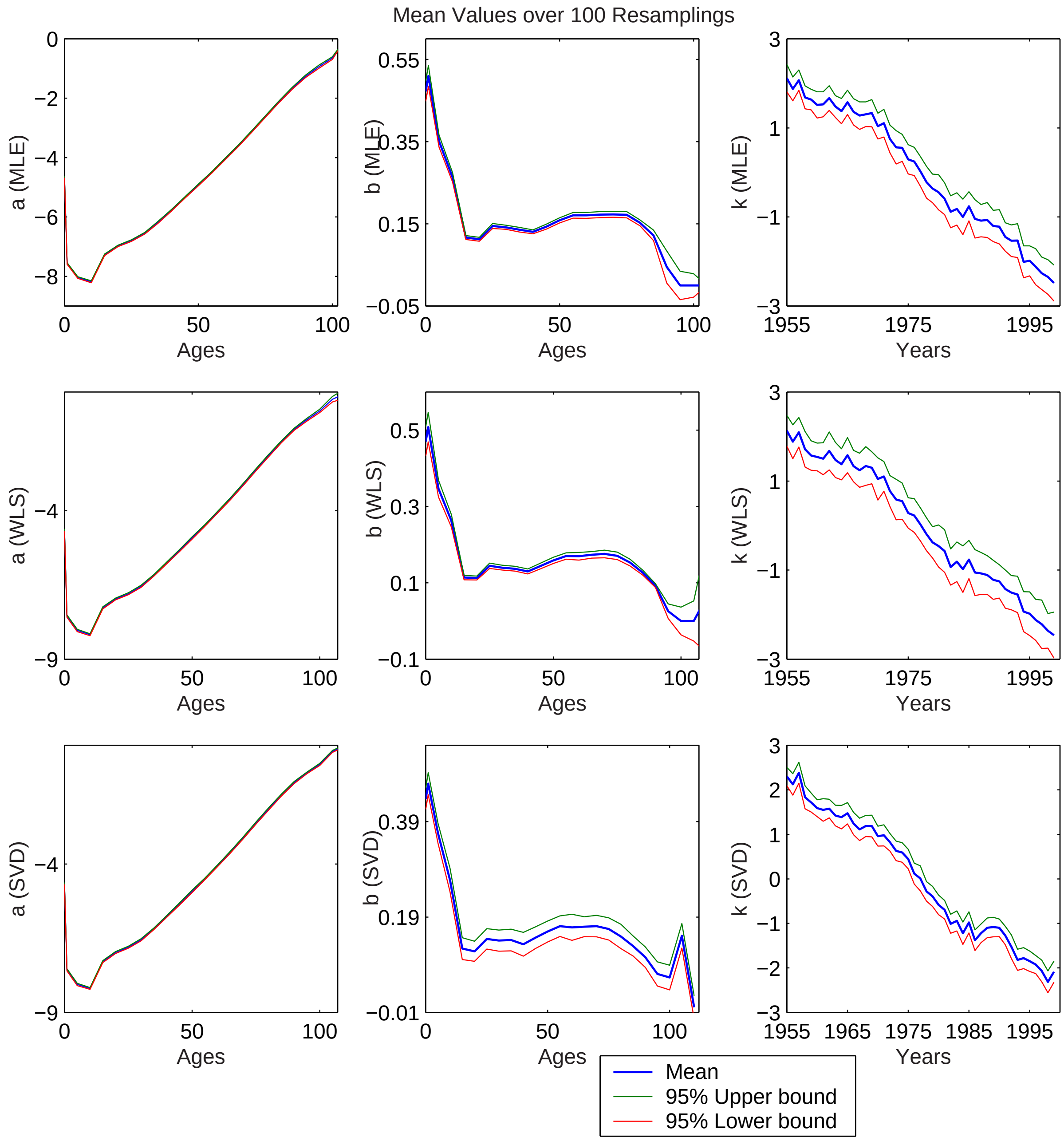

Figure 10: Mean Values over 100 Resamplings (MLE, WLS and SVD)

5 Conclusion

To study the efficiency of the Lee-Carter method, the model's parameters were estimated with the Singular Value Decomposition (SVD), the Maximum Likelihood Estimate (MLE) and the Weighted Least Square (WLS) method. The computations were made using the vital rates (1955 to 1999) from Finland, Sweden, Denmark and Norway.

The results show that, under an appropriately chosen estimation period, the model fit the observed death rates quite well, with the SVD, the MLE or the WLS. The estimates for the age parameters a and b are almost alike, while there is some variation in the estimates of the time-dependent mortality index k. A bootstrap simulation indicates that the use of the MLE results in smaller mean squared errors for the parameters a and b than the use of the two other methods. The SVD is, however, the best alternative for the mortality index k.

References

[1] Alho, Juha (1998). *A Stochastic Forecast of the Population of Finland*. Tilastokeskus, reviews 1998/4.

[2] Alho, J. (2000). Discussion of *Lee Ronald D.(2000). North American Acturial Journal*. Vol.**4**, N0. 1, 91-93.

[3] Alho, Juha (2003). Predictive Distribution of Adjustment for Life Expectancy Change. Finnish Centre for Pensions. Working Papers 3.

[4] Booth, Heather; Maindonald, John and Smith, Len. (2002). Applying Lee-Carter under conditions of variable mortality decline. *Population Studies*, 56, 325-336.

[5] Brillinger David R., (1986). The natural variability of vital rates and associated statistics. *Biometrics* **42**(4):693-734.

[6] Brouhns, Natacha; Denuit, Michel and Vermunt, Jeroen K. (2002). A Poisson log-linear regression approach to the construction of projected lifetables. *Insurance: Mathematics and Economics* 31, 373-393.

[7] Carter, Lawrence R. and Prskawetz Alexia (March 2001). Examining Structural shifts in Mortality Using the Lee-Carter Method. Working Paper. Max Planck Institute for Demographic Research.

[8] Keilman, Nico (1998). How Accurate Are the United Nations World Population Projections? *Population and Development Review*. Vol. **24**, Supplement: Frontiers of Population Forecasting, pp. 15-41. Society of Actuaries, U.S.A.

[9] Lee, Ronald D. (2000). The Lee-Carter method for forecasting mortality, with various extensions and applications. *North American Acturial Journal*. Vol.**4**, N0. 1, 80-91.

[10] Lee, Ronald D. and Carter, Lawrence R. (September, 1992). Modeling and Forecasting U.S. Mortality, *Journal of the American Statistical Association*. **87**,Issue 419 ,659-675.

[11] Lee, Ronald D. and Miller T. (November, 2000). Evaluating the performance of the Lee-Carter mortality forecasts. *Demography*, v. 38, n. 4 pp. 537-549.

[12] Lee, Ronald D. and Nault, Francois. (1993). Modeling and Forecasting Provincial Mortality in Canada. Presented at the World Congress of the International Union for the scientific Study of Population, Montreal.

[13] Lee, Ronald D. and Rofman, R. (1994). Modeling and Forecasting Mortality in Chile. *Notas* **22**, No. 59:182-213.

[14] Li, N.; Lee, R. and Tuljapurkar, S. (April 2004). Using the Lee-Carter method to forecast mortality for population with limited data. *International Statistical Review*. Vol 72(1), 19-36.

[15] Oeppen, Jim and Vaupel, James W. (10 May 2002). Broken Limits to Life Expectancy. *Science* **296**, 1029-1031.

[16] Olshansky, S. Jay and Carnes, Bruce A. (February 1997). Ever Since Gompertz. *Demography*, **Vol. 34**, Number 1.

[17] Renshaw, A.E. and Haberman, S. (2003). Lee-Carter mortality forecasting with age-specific enhancement. *Insurance: Mathematics and Economics* 33, 255-272.

[18] Stoto, M. (1983). Accuracy of population projections. *Journal of the American Statistical Association*, **78**, 13-20.

[19] Tuljapurkar, Shripad; Nan Li and Boe, Carl . (15 June 2000). A universal pattern of mortality decline in the G7 countries. *Nature*, **Vol. 405**, 789-792.

[20] United Nations, 1958. Multilingual Demographic Dictionary. *United Nations Population Studies*, No. 29.

[21] Wilmoth, John R. (Dec. 1993). Computational Methods for Fitting and Extrapolating the Lee-Carter Model of Mortality change. Technical Report, University of California, Berkeley. U.S.A.

[22] Wilmoth, John R. (1996). Mortality Projections for Japan: A comparison of Four Methods. Health and Mortality among Elderly Population, Eds Graziella Caselli and Alan Lopez. New York: Oxford University Press.