



Mathematical Statistics  
Stockholm University

**Mortality predictions for longevity  
analysis and annuity valuation**

Peter Wohlfart

**Examensarbete 2006:11**

**Postal address:**

Mathematical Statistics  
Dept. of Mathematics  
Stockholm University  
SE-106 91 Stockholm  
Sweden

**Internet:**

<http://www.math.su.se/matstat>



Mathematical Statistics  
Stockholm University  
Examensarbete 2006:11,  
<http://www.math.su.se/matstat>

# Mortality predictions for longevity analysis and annuity valuation

Peter Wohlfart\*

May 2006

## Abstract

In life annuity business it is of great importance to have accurate mortality predictions for calculating the value of annuity contracts. In this paper we are considering the issue of predicting future mortality. We are using a non-parametric counting process approach combined with a kernel smoothing- and a bias correction technique for estimation of the past mortality. By using the Lee-Carter method we are adapting a mortality model and generating forecasts of the future mortality. We are using population data from Sweden and Denmark during 1900–2004 and the results are evaluated by backtesting. We are also evaluating the consequences of varying the length of the estimation period in the Lee-Carter model, and it appears that it may have a great impact on the predictions.

---

\*Postal address: Dept of Mathematical Statistics, Stockholm University, SE-106 91 Stockholm, Sweden. E-mail: [pwohlfart@deloitte.dk](mailto:pwohlfart@deloitte.dk). Supervisor: Anders Martin-Löf.

## **Foreword**

This paper is a 20 credits Master's thesis, performed at Stockholm University in cooperation with the Actuarial and Insurance Solutions group at Deloitte in Copenhagen. The work has been performed during November 2005 to May 2006. First of all, I would like to thank Peter Fledelius, my supervisor at Deloitte in Copenhagen, for excellent supervising. Peter has been given me great support and a lot of rewarding discussions during my work. I would also like to thank professor Steven Haberman and Arthur Renshaw at City University in London for spending time answering questions concerning the Lee-Carter method, and for sending me their forthcoming publication. Finally, I would also like to thank my supervisor at Stockholm University, Anders Martin Löv.

# Contents

<b>1</b>	<b>Introduction</b>	<b>5</b>
1.1	The valuation problem . . . . .	5
1.2	Problems with lack of information . . . . .	6
1.3	Comparing Sweden and Denmark . . . . .	7
1.4	Purpose . . . . .	8
<b>2</b>	<b>Previous Research</b>	<b>10</b>
<b>3</b>	<b>Model and Data</b>	<b>13</b>
3.1	Mortality Model . . . . .	13
3.2	Time Dependent Mortality Model . . . . .	14
3.2.1	Lexis Diagrams . . . . .	15
3.2.2	Estimating the Mortality Rate . . . . .	16
3.3	Data . . . . .	18
3.3.1	Expected Lifetime . . . . .	18
3.3.2	Death Counts and Exposure . . . . .	19
3.3.3	Problems with the Counting Process Technique . . . . .	20
3.3.4	Mortality Rate Estimations . . . . .	21
<b>4</b>	<b>Kernel Smoothing Technique</b>	<b>23</b>
4.1	Local Constant Kernel Smoothing . . . . .	24
4.1.1	Kernel Smoothing at Boundaries . . . . .	26
4.1.2	Bias Problem at Boundaries . . . . .	27
4.2	Non-Smoothed and Smoothed Mortality Rates . . . . .	27
4.3	Optimal Bandwidth - Bias vs Smoothness . . . . .	28
4.4	Bias Correction Technique . . . . .	29
<b>5</b>	<b>The Lee-Carter Method</b>	<b>32</b>
5.1	The Model . . . . .	32
5.2	Fitting the Model . . . . .	33
5.2.1	Second Stage Estimation . . . . .	34

5.2.2	Studying the Mortality Index . . . . .	35
5.3	Mortality Forecasts . . . . .	36
<b>6</b>	<b>Procedure</b>	<b>38</b>
6.1	Backtesting . . . . .	38
6.2	Period and Cohort Expected Lifetime . . . . .	39
6.3	Evaluation . . . . .	40
6.4	Varying the Length of the Estimation Period . . . . .	44
6.4.1	Finding the Optimal Length . . . . .	45
<b>7</b>	<b>Results</b>	<b>48</b>
7.1	35 Years Estimation Periods . . . . .	48
7.1.1	Mortality Rates . . . . .	48
7.1.2	Cohort Expected Lifetime . . . . .	50
7.2	2-75 Years Estimation Periods . . . . .	53
7.2.1	Cohort Expected Lifetime . . . . .	54
<b>8</b>	<b>Analysis</b>	<b>58</b>
8.1	Length of the Estimation Period . . . . .	58
8.2	Size of the Prediction Error . . . . .	61
8.3	Underprediction Problem . . . . .	62
<b>9</b>	<b>Future Forecasts</b>	<b>66</b>
9.1	Comparison to other Predictions . . . . .	68
<b>10</b>	<b>Summary and Concluding Remarks</b>	<b>71</b>
10.1	Summary . . . . .	71
10.2	Concluding Remarks . . . . .	73
<b>11</b>	<b>Figures</b>	<b>80</b>

# 1 Introduction

## 1.1 The valuation problem

One of the main problems regarding life annuity business is how to evaluate different contracts. The main part of the cash flows belonging to the contracts often takes place a long time ahead, and the task is to identify these future cash flows and to determine their present value. The value of a contract depends partly on the interest rate, and partly on the remaining lifetime of one or several insured individuals. Since both the interest rate and the lifetime of individuals can be regarded as stochastic, one cannot determine the exact value of a contract. Nevertheless, when determining the size of the company's reserves, it is very important to have a reasonable understanding for the expected value of a specific contract. To estimate the expected value of a specific contract one has to consider both the expected investment return and the expectation of the underlying mortality. Concerning the investment return, one has to predict the development of the interest rate during the period from today to the time when the contract expires. Even if this prediction is not an easy task, one can often overcome the problem (or at least reduce it) by hedging with financial contracts. This way it is sometimes possible for the company to create a deterministic interest rate<sup>1</sup>. Concerning the mortality prediction, the problem is more complicated. There is no financial market, like for the interest rate, where it is possible to trade and hedge with contracts regarding mortality. This means that we have to rely on our mortality predictions in a much greater way than we had for the interest rate predictions. The future mortality has to be predicted by statistical methods, and usually the predictions are based on historical observations of the mortality. A problem which complicates this kind of prediction

---

<sup>1</sup>However, in life annuity business, where the contracts' time to maturity may be very long, it is a risk that there is no financial contract with the desired time to maturity. Nevertheless, even if this is the case, the risk connected with the insecurity of the interest rate can be substantially reduced

is that the mortality does not tend to be constant over time. The general trend is that the mortality rate has decreased over time, i.e. the expected lifetime for individuals have increased (see for example Velfaerdskommissionen (2004)). If you completely rely on historical data when predicting the expected remaining lifetime, and assumes that the future mortality will be the same as the mortality today, you will run the risk of overestimating the mortality, and therefore underestimate the remaining expected lifetime. This underestimation may result in that the company will face costs, regarding their contracts, that differ a lot from their expectations. These unexpected costs may strike the company in opposite directions. An annuity contract that pays money to the insured as long as he or she is alive tends to become more expensive for the company, while a contract that pays money in the case of death tends to become cheaper. For the company it is primary the first type of contract that it has to worry about, but it is important for the company to have accurate predictions of all their future costs. In the long run an incorrect mortality prediction can be a fateful mistake, risking the whole company's survival.

## **1.2 Problems with lack of information**

A problem that may arise when basing future mortality predictions on historical mortality observations, may be the lack of data. One way of estimating the mortality rate for a specific age is to compare the number of deaths to the total number of individuals at that age. The mortality rate is estimated as the ratio between these two quantities. When considering low- or middle aged individuals this estimation seldom causes any problem since there is a lot of information available for individuals at these ages. However, the higher ages you are considering, the less information is available, since the number of individuals to study become smaller as people pass away. When there is not much of information available, the estimated mortality rates are



more uncertain. This may be a problem since it is desirable to plan the company's reserves on more reliable estimations. A way to handle the problem with lack of information for high ages is to use a smoothing technique, where you "borrow" information from adjacent ages. For example, if you want to estimate the mortality rate for a 90 year old man, and you do not have so much demographic information available, you may use the information for men aged 88-92. In the same way you may use the data for men aged 89-93 to estimate the mortality rate for a 91 year old man. By this smoothing technique you will get more stable mortality estimations, since the estimations are based on more data. Unfortunately, this smoothing technique also has its drawbacks. The smoothing technique may result in a bias problem since it is highly probable that the mortality rate is not constant in the interval 88-92. When using this smoothing technique it is important to have an awareness of this bias problem, and to use suitable methods to overcome it.

### **1.3 Comparing Sweden and Denmark**

As mentioned above, the mortality rate does not tend to be constant over time. Further, there are noticeable differences between men and women and between different countries. Comparing the expected lifetime in Sweden and Denmark, using information from Velfaerdskommissionen (2004), we find some interesting facts. In year 1960 the expected lifetime was similar, with a little advantage for Sweden, and both countries were among those with the highest life expectancy among the world's developed countries. This was the case for both men and women. From that on, the overall global trend is showing an increase in expected lifetime, which is also the case for both Sweden and Denmark. However, studying the situation in year 2000, the increase in Sweden has been much larger than the increase in Denmark. This has resulted in that Denmark has dropped several places in the ranking at the expected lifetime table, while Sweden is still in a top position. The pattern is

similar for both men and women; the difference in expected lifetime between Sweden and Denmark has increased from less than a year (in 1960) to almost three years (in 2000), see Velfaerdskommissionen (2004). These properties are illustrated in figure 1, where we have plotted the expected lifetime for males and females in Sweden and Denmark during 1950-2004.

## 1.4 Purpose

In this Master thesis we are dealing with the problem of estimating and predicting the mortality rate and the expected lifetime. We are comparing Sweden and Denmark, and we are investigating how the mortality has developed. We are considering the force of mortality as a two dimensional function of age and chronological time. We are using a non-parametric counting process approach, combined with a kernel smoothing technique, to estimate the mortality. We will then use the Lee-Carter (LC) method to adapt a parametric model and to produce predictions of the future mortality rates. The LC method, proposed in Lee and Carter (1992) has become the leading statistical model, for estimating and predicting mortality, in the demographic literature (Deaton and Paxson (2001)).

For the estimation we will only use a part of the data available. The rest of the data is then used for evaluating the performance of the predictions from the LC method. The observed mortality is compared to the predicted mortality produced by the LC method. In the evaluation of the LC method we are mainly comparing the observed and the predicted remaining life expectancy. Finally, we are trying to improve the performance of the predictions by varying the length of the estimation period when adapting the model. We are trying to find the optimal length (in years) of the estimation period in terms of producing as good predictions as possible of the future mortality. The issue of finding an optimal length of the estimation period is for us an unfamiliar

concern in the actuarial literature, and is therefore the main contribution of this thesis.

In section 2 we are presenting some previous studies of the area of interest. In section 3 we introduce the basic notation and some general formulas. The data used in this thesis will then be presented. We will illustrate and discuss the differences between Sweden and Denmark regarding the expected lifetime. We will also study the mortality rate and the way it can be estimated, using a counting process technique. Further, we will discuss some of the problems that may arise due to lack of information. In section 4 we will present suggestions for solving these kind of problems. We will present and explain the non-parametric kernel smoothing technique, and we will present a bias correction technique for solving the bias problem from the local constant kernel estimator. In section 5 we are presenting the two dimensional Lee-Carter (1992) model. We will discuss the model and its properties, and the way the parameters are being estimated. We will also look at how future mortality predictions can be generated by the model. In section 6 we are discussing the importance of validating when choosing a statistical model. Different ways of measuring the life expectancy are illustrated. We are then describing the procedure for the evaluation of the Lee-Carter method, and how to find the optimal length of the estimation period. In section 7 we are presenting the result from the evaluation of the Lee-Carter method. The results are analyzed in section 8. In section 9 we are using the prediction method to produce forecast of the future mortality. In section 10 we will present a summarization and give some concluding remarks. All figures are presented in section 11.

## 2 Previous Research

Lee and Miller (2001) are evaluating the performance of the LC method for forecasting mortality. Just as Lee and Carter they are using U.S. data from the 20th century. They are using the LC method to predict the mortality (expected lifetime) in 1998. They are constructing forecast with different "jump-off" years, by pretending they only had data up to that point. They are then comparing their forecasts with the actual outcomes. They are also comparing the forecasts with forecasts made by the Social Security. There are several interesting conclusions. The hypothetical forecasts tend to be too low, but they are still fairly close to the actual mortality in 1998. The earlier forecasts, produced by using data up to the 1920s and the 1930s, are on average 5 years below the true value. But the forecasts from 1946 and forward are always within 2 years from the actual value in 1998. Compared to the forecasts from the Social Security, which also have systematically under-predicted the mortality, the LC forecasts have substantially lower prediction errors and mean squared errors. Unlike Lee and Miller, who are comparing the predictions performance from different "jump-off" years, we will instead focus on how the predictions are affected by varying the number of estimation years for adapting the LC model.

Girosi and King (2005) are considering the Lee-Carter method, and they show that the method is a special case of a multivariate random walk with drift (RWD). They also show that LC predictions will become less smooth after some time, and that this nonsmoothness will continue forever.

Lindbergson (2001) is investigating the mortality in Sweden among elderly for the years 1988-1997. The annual increase for the observed mortality rate is increasing with the age, but it levels out at higher ages. Lindbergson therefore suggest that the Makeham function, which assumes an exponential growth, is replaced by an straight line at very high ages. She therefore fits a "Modified Makeham" function, which is a combination of a Makeham func-

tion and a straight line. The estimated mortality rates for very high ages will then be lower than for the ordinary Makeham function.

Fledelius, Guillen, Nielsen and Petersen (2002) are studying the mortality for old-aged (90 and above) people in Sweden during 1988-1997. They are using the same data as Lindbergson (2001) and they are applying a smooth two-dimensional kernel hazard estimation to the data. They apply both a local constant- and a local linear estimator, and they compare the result to the work of Lindbergson (2001). They show that the local constant estimator is not suitable for elderly people since it significantly underestimates the mortality. This is due to the rapid increase in mortality and the rapid decrease in exposure at these old ages. These properties are similar to those experienced during working with this thesis.

The overall trend has been a decreasing mortality over time but Willets (2004) mention that the mortality decrease for adults stopped decline and even slightly increased in the 1930's due to higher deaths from lung cancer and heart disease. These trends were also spotted for elderly in the Swedish and Danish population during this thesis.

Renshaw and Haberman (2003b) have observed that the traditional Lee-Carter age-period model not always fits empirical data well. They are therefore extending the model (Renshaw and Haberman (2006-forthcoming publication)) to also include the cohort effect. This makes the modeling a bit more complicated due to the relationship

cohort = period - age

For estimating the parameters they are first setting some suitable starting values, and are then using an iterative process. They are considering the England and Wales mortality, and when studying residual plots they found that the extended Lee-Carter (age-period-cohort) model better fits empirical data than the original LC version.

Booth, Maindonald and Smith (2002) have adapted the Lee-Carter model to take account for departures from the linearity in the time component. They have expanded the model by including terms of second and higher order. However, these terms are not easily incorporated into the forecast. There are a lot of different mortality models. A problem when having a large number of parameters in the model is that the potential for extrapolation is limited.

### 3 Model and Data

#### 3.1 Mortality Model

Let the lifetime for an individual be denoted by  $T$ , which can be regarded as a non-negative continuous stochastic variable. If the probability distribution function for  $T$  is denoted by  $F$ , we get

$$F(z) = P(T \leq z), \quad z \geq 0. \quad (1)$$

This is the probability that the lifetime of an individual will be shorter than  $z$  years. It is often convenient to study the survival function  $l(z)$ , which is defined as

$$l(z) = 1 - F(z) = P(T > z), \quad z \geq 0. \quad (2)$$

The mortality rate can be interpreted as an intensity at which individuals die. The mortality rate is defined as

$$\mu(z) = \frac{f(z)}{1 - F(z)}, \quad z \geq 0. \quad (3)$$

where  $f(z) = F'(z)$ . The mortality rate can also be expressed in terms of the survival function

$$\mu(z) = \frac{-l'(z)}{l(z)}, \quad z \geq 0. \quad (4)$$

The remaining lifetime of an individual aged  $x$  is denoted by  $T_x$ . Just like the variable  $T$ ,  $T_x$  can be considered as a non-negative continuous stochastic variable. We have that

$$P(T_x \leq z) = P(T \leq x + z | T > x) = \frac{F(x + z) - F(x)}{1 - F(x)}, \quad z \geq 0. \quad (5)$$

This is the probability that a person aged  $x$  dies within  $z$  years. The probability that a person aged  $x$  will live for, at least, another  $z$  years can then be expressed as

$$l_x(z) = P(T_x > z) = \frac{l(x+z)}{l(x)}, \quad z \geq 0. \quad (6)$$

The density function for  $T_x$  is  $\frac{-l'(x+z)}{l(x)}$  for  $z \geq 0$ , and the expected remaining lifetime for an individual aged  $x$  can be expressed as

$$E(T_x) = e_x = \int_0^\infty \frac{l(x+z)}{l(x)} dz \quad (7)$$

Equation (7) assumes continuous time. In this thesis however, we are working in discrete time, with annual information about the individuals. We are therefore using an approximation for calculating the expected remaining lifetime

$$e_x = \left( \sum_{i=0}^{\infty} \frac{l(x+i)}{l(x)} \right) - 0.5 \quad (8)$$

### 3.2 Time Dependent Mortality Model

In all the formulas presented this far, we have not considered the chronological time as an independent variable. The only independent variable has been the age of the individual. However, in this thesis, we will consider the mortality rate to be dependent on not just the age ( $x$ ), but also on the chronological time ( $t$ ). This is because of the overall trend of an increasing life expectancy over time for individuals. The formulas presented above are still relevant though. For each time  $t$ , we will have an unique density function,



and therefore also an unique intensity function. Consequently, the difference will be that, instead of just having one specific mortality rate for a certain age,  $\mu(x)$ , we will now for each calender year have an unique mortality rate for that age,  $\mu(x, t)$ .

### 3.2.1 Lexis Diagrams

Lexis diagrams are often used in demographical studies, and they are a simple way to graphically present population dynamics. There are three demographic co-ordinates of special interest; the chronological time, the age and the moment of birth of the individual. Using these co-ordinates, the lifetime of individuals can be illustrated in a three dimensional space. However, it is sufficient to know two of these three co-ordinates, since the third then is given by the other two. This means that the lifetime of individuals can be illustrated by a coordinate system, where the axes can be chosen as two of these three co-ordinates (se for example Vandeschrick (2001)). Here we are using Lexis diagrams with the chronological time and the age on the two axes. In figure 2 we give an example of a Lexis diagram. The age of the individual is represented on the vertical axis, while the horizontal axis represents the calender year. The lifetime for an individual is represented by a 45 degrees straight line. Here we are only studying three different individuals, just to describe the way a Lexis diagram can be used. Each line starts at the time when the individual first get exposed to risk (for example at birth), and ends when the individual no longer is exposed to risk (for example in the case of death). Here we are considering ages from 0 to 110 and the calender years from 1900 to 2004. Individual *A* first got exposed to risk at birth in year 1901. As time goes by, the line representing *A* stretches up to the right. In year 2003 individual *A* dies at the age of 101, and is therefore no longer exposed to risk. Individual *B* gets exposed to risk at the age of 17 in year 1900 and dies at the age of 67 in year 1950. Finally, individual *C* is born in

1936. At the end of 2004,  $C$  is still exposed to risk, at an age of 68.

### 3.2.2 Estimating the Mortality Rate

The mortality rate can also be expressed in terms of the remaining lifetime

$$\mu(x) = \lim_{z \rightarrow 0} = \frac{P(T_x \leq z)}{z} \quad (9)$$

The probability that a person aged  $x$  will die at an age within the interval  $(x, x + dx)$  is therefore approximately  $\mu(x)dx$  for small  $dx$ . A natural way to get an estimation for  $\mu(x)$  is to therefore study the number of individuals aged  $x$  that passes away at an age within an interval  $(x, x + dx)$  compared to the total number of individuals exposed to risk during the same interval. The mortality rate can be estimated as the ratio between these two quantities. The more information available, i.e. the more individuals exposed to risk, the more stable and reliable the estimation will be, since the randomness of the estimation will be reduced.

In this thesis we are considering the chronological time as an independent variable, and we then have to get estimates of the mortality rates, not just for all ages  $x$ , but also for each calendar year  $t$ . By including this extra independent variable, we are also adding one extra dimension for the intensity (mortality rate) function. Instead of considering the mortality function as a curve, depending on the just the age, in a two dimensional space, it now has to be considered as a surface in a three dimensional space, depending on both the age and the calendar year. Accordingly, the exposure for the individuals now has to be split up with respect to both the age and the calendar year. This will substantially reduce the available information for the estimations (the exposure for a certain age and a certain calendar year will become substantially smaller than the information that could have been used if just the

age, and not the year was considered). The lack of information may lead to some difficulties which will be discussed later.

To get the estimates of  $\mu(x, t)$ , we now have to study the number of individuals that passes away at an age "near"  $x$ , and at a time "near"  $t$ , compared to the individuals exposed to risk inside the same interval. In other words, we have to study the number of individuals that dies at an age and at a time within the two dimensional interval  $((x, t), (x + dx, t + dt))$ , compared to the total exposure inside the same interval. In this thesis we are only studying integer ages on annual basis, so the interval  $((x, t), (x + dx, t + dt))$  will therefore be symmetric with a length and a width of 1 year. Accordingly, we are not taking care of when exactly during a calender year the deaths occur. Neither are we considering the exact age of the individuals (we are just considering integer ages).

Let the "occurrence matrix",  $O_{x,t}$ , represent the number of deaths for individuals aged  $x$  in year  $t$ . Further, let the "exposure matrix",  $E_{x,t}$ , represent the total exposure for individuals aged  $x$  in the same period. The mortality rate  $m_{x,t}$  for an individual aged  $x$  in time  $t$  can then easily be estimated as the ratio between these two quantities.

$$\hat{m}(x, t) = \frac{O_{x,t}}{E_{x,t}} \quad (10)$$

The occurrence matrix and the exposure matrix can be calculated using a Lexis diagram as a starting point. Each cell in the Lexis diagram represents a specific age and a specific calender year. By summing up the number of deaths in each cell we will get the elements in the  $O_{x,t}$  matrix. In the same way we can sum up all the exposure in each cell and get the elements in the  $E_{x,t}$  matrix.

This counting process technique is a non-parametric approach for the estima-

tion of the mortality rate, and the technique requires no assumptions about the distribution of the underlying mortality. (Fledelius, Guillen, Nielsen and Vogelius (2004)). However, we will assume that the mortality surface is smooth. Consequently, we assume that there is no large "jumps" in the mortality rate between adjacent ages and years.

### 3.3 Data

All data are collected from the *Human Mortality Database*<sup>2</sup> (HMD). For Sweden we have data for the period 1751-2004, and for Denmark we have data from 1835-2004. The data consists of the deaths and the exposure for individuals ages 0-110.

#### 3.3.1 Expected Lifetime

Looking at the data for the expected lifetime, we note that there is a very clear increase in the expected lifetime during time. Roughly speaking we can say that the expected lifetime in 2004 has doubled since 1835. The data also indicates that there is an obvious difference between men and women. During the whole time period the life expectancy for men is smaller than for women. In figures 3 and 4 where we are illustrating the development of the expected lifetime during 1835-2004 for men and women in both countries. During the first half of the time period, the curves are rather unsmooth. The big drop in year 1918 is due to the flu epidemic. The expected lifetime has over time been slightly higher in Sweden than in Denmark for both men and women. Furthermore, as mentioned in the introduction, we also observe an increase in the differences between the two countries during the last 50 years.

---

<sup>2</sup>Human Mortality Database is the work of three teams of researchers in USA, Germany and Canada.

This observation is better illustrated in figure 1, where the expected lifetime for the period 1950-2004 are plotted. The difference in expected lifetime between Sweden and Denmark was less than a year in 1950, (0.7 years for males and 0.9 years for females). In 2004 the difference between the both countries has increased to about 3 years (3.2 years for males and 2.9 years for females). The expected lifetime in 1950 and in 2004 are presented in the tables below.

	1950	2004
Male Sweden	69.83	78.35
Male Denmark	69.10	75.14
Difference	0.73	3.21

	1950	2004
Female Sweden	72.44	82.66
Female Denmark	71.52	79.80
Difference	0.92	2.86

### 3.3.2 Death Counts and Exposure

Our data is grouped annually, giving us death counts,  $O_{x,t}$ , and exposure,  $E_{x,t}$ , for each year and each integer age. Both the number of deaths and the exposure for different ages in different years can be illustrated as a surface in a three dimensional space. The number of deaths at different ages during 1900-2004 for Swedish women are illustrated in figure 5. The figure tells us that there was a high infant mortality at the beginning of the 20th century. During the last 100 years, the number of infant deaths has decreased considerably, and we instead recognize an increase in the number of deaths for elderly people. These characteristics are the same for Swedish men and for both men and women in Denmark. In figure 6 the exposure for Swedish females during 1900-2004 are illustrated in the same way. The figure shows that the exposure is decreasing with the age. This is completely in line with our expectations; the exposure will be less when more individuals have passed away. The figure also indicates some increase in the exposure with the chronological time. This can be explained by an increasing population, and by the fact that the expected lifetime has increased. However, it is the

age, rather than the chronological time, that has the main influence of the exposure.

### 3.3.3 Problems with the Counting Process Technique

When the occurrence and the exposure are known, the mortality rates can easily be calculated according to (10). We assume that the underlying mortality surface is smooth and that there is no large "jumps" between adjacent ages and years. This assumption causes no problem when the information, that our estimations are based on, is large. When the exposure is large, the estimations are rather stable, and the surface describing the estimated mortality rate is therefore rather smooth. These favorable properties are fulfilled for most ages and years. However, for higher ages, this is necessary not true, and we might therefore run into problems with a very small (and sometimes even zero) exposure for some old ages at a certain year. When the exposure is small the randomness in the estimations will increase. When the randomness increases the estimations may become unstable and insecure. As we have seen earlier, the exposure is decreasing rapidly for higher ages. Problems that then may come up are the following.

- There might be situations where the number of deaths for a certain (old) age and year is zero (and the exposure is very small). This will make the mortality rate estimation equal to zero, which, of course, is an unrealistic assumption.
- Another difficult situation is when the exposure is very small (almost zero), but we still have some death(s). Then the mortality rate estimation may be unrealistic high compared to the adjacent estimations.
- A third problematic situation that might come up is when the exposure, and therefore also the number of deaths, is zero. In this case we will

have no mortality estimation at all since ratio between zero and zero is not defined.

### 3.3.4 Mortality Rate Estimations

These are all problematic situations, which need to be taken care of, when trying to describe the development of the mortality rate. Looking at our data of the mortality rates, we recognize these problems for higher ages. In figure 7 the estimations of the mortality rate in Sweden and Denmark are illustrated. To make the illustration more clear, we have only included the estimations for individuals up to an age of 99. The estimations for people aged 100 and above are not suitable for graphical presentation, since they are very unsmooth. The estimations are varying a lot between zero and very high values. We also have some "missing values" when the exposure is zero. All these events occur according to the problematic situations described above. Except the obvious fact that the mortality in general is increasing with the age, we can, by looking at figure 7, see that the mortality rate at birth was rather high at the beginning of the 20th century, but that it has decreased considerably after that. This is in line with the observations from figure 5. We also notice that the mortality at higher ages has decreased over the years, which also is in line with our, up to now, received experiences. It is difficult to draw any big conclusions by just studying this figure, but we recognize the estimations to be rather unsmooth at the highest ages (and this is the case even when individuals aged 100 above are excluded).

The non-parametric approach has the advantage that we do not need to make a lot of assumptions about the distribution of the mortality. For example, we do not have to rely on the Gompertz Makeham model. Further, we do not have to make assumptions of how the mortality changes over time (see Fledelius et. al. (2004)). By using annually grouped data and this non-parametric technique, we are able to get a unique estimate of the mortality rate for each year and each age.

In figure 8 the mortality rate estimations are illustrated in an alternative graphical approach. The figures show a contour plot where equal mortality rate estimations are connected with level curves. The conclusions from the figure are about the same as before. It is clear from the figure that, when looking at old- or middle aged people, the mortality for a certain age has decreased over time.

The figures presented in this section only give us an general picture of the demographic development. Here we have only made the illustration for Swedish women, but the main demographic trends are the same for Swedish men and for Danish men and women. To be able to discuss differences between the two countries, we have to study the data more detailed.



## 4 Kernel Smoothing Technique

As mentioned in previous sections, there might appear problems when estimating the mortality for higher ages. The reason is that there is less information available for people at higher ages. This fact is illustrated in figure 9, where we have plotted the 25 years (1980-2004) aggregated exposure for different ages in Sweden and Denmark. As we can see from the figure, the exposure decreases rapidly for people aged about 75 and above. We also see that the total exposure is greater in Sweden than in Denmark, due to a larger population.

The mortality rate is estimated as the ratio between the number of deaths and the total exposure at a specific age and time (see (10)). In section 3 we saw that the estimations could become rather unstable for higher ages when the exposure were small. This resulted in a very unsmooth mortality surface, which is not consistent with the smoothness assumption. A way to overcome this problem is to borrow information from adjacent ages and times. Suppose for example that we are interested in estimating the mortality rate for a 90 year old man in year 2000. Instead of just using the information for the current age and the current year, we may instead, for example, use all information available for men in the age interval 88-92 during the years 1998-2002. The idea is illustrated in the Lexis diagram in figure 10. The dark shaded cell in the middle is representing the age (90) and the year (2000), for which we want to estimate the mortality. For the estimation we are using information from cells in a local neighborhood of this origin cell. In this case we are using information from 25 (5x5) different cells, each representing a specific age and a specific year. These cells are the ones that are shaded in the figure. The impact on the different cells depends on the distance to the origin. In the figure, the distance for each cell is indicated by two numbers, the first representing the absolute distance in the age dimension, while the second represents the absolute distance in the calendar year dimension. The closer distance to the origin, the more weight is added to the information

from that cell. By using this smoothing technique, we are using more information, and we will therefore have a larger exposure. This way the estimates will become more stable. However, this smoothing technique may result in a bias problem since it is unlikely that the mortality should be the same for a 88 year old than for a 92 year old. It is also possible that there might be mortality rate differences over time during the period 1998-2002. But these potential differences in the "time" dimension are probably small compared to the differences due to the age. To optimize the result of the smoothing technique, one has to handle this bias problem in one way or another.

#### 4.1 Local Constant Kernel Smoothing

We consider the mortality rate as a two-dimensional function of age and chronological time. Using the same notations as before we have that  $O_{x,t}$  and  $E_{x,t}$  are the occurrence (number of deaths) and the exposure for people aged  $x$  in year  $t$ . To calculate the mortality rate we will smooth occurrence and exposure separately. The smoothed occurrence and exposure for the same age and year are denoted by  $\bar{O}_{x,t}$  and  $\bar{E}_{x,t}$ .

The smoothed, local constant, mortality rate is then defined by

$$\bar{m}_{x,t} = \frac{\bar{O}_{x,t}}{\bar{E}_{x,t}} \quad (11)$$

where the smoothed occurrence and exposure are defined as (see Fledelius, Lando and Nielsen (2004)).

$$\bar{O}_{x,t} = \sum_{x_1=1}^N \sum_{t_1=1}^T K_{b_1}(x - x_1) * K_{b_2}(t - t_1) * O_{x_1,t_1} \quad (12)$$

$$\bar{E}_{x,t} = \sum_{x_1=1}^N \sum_{t_1=1}^T K_{b_1}(x - x_1) * K_{b_2}(t - t_1) * E_{x_1,t_1} \quad (13)$$

where  $N$  and  $T$  are the highest age and the latest year we are considering.

$\bar{O}_{x,t}$  and  $\bar{E}_{x,t}$  can be regarded as weighted averages of the occurrence and the exposure in the two-dimensional interval  $(x \pm b_1, t \pm b_2)$ , where  $b_1$  and  $b_2$  are the bandwidths.  $b_1$  decides the number of different ages to be included, while  $b_2$  decides the number of calendar years. The weights are determined by the kernel functions  $K_{b_1}$  and  $K_{b_2}$ . Here we use the Epanechnikov kernel function.

$$K_{b_1}(x - x_1) = 0.75 * I_{[|x-x_1|<b_1]} * \left[ 1 - \left( \frac{x - x_1}{b_1} \right)^2 \right] \quad (14)$$

$$K_{b_2}(t - t_1) = 0.75 * I_{[|t-t_1|<b_2]} * \left[ 1 - \left( \frac{t - t_1}{b_2} \right)^2 \right] \quad (15)$$

where  $I$  is an indicator function taking value one or zero.

The weight for an observation is a function of the distance to the point of interest. The closer the distance the more weight is added to the observation<sup>3</sup>. An observation outside the interval  $(x \pm b_1, t \pm b_2)$  will receive zero weight.

The smoothing technique is adapted to the occurrence and the exposure separately. This means that the idea presented in the Lexis diagram in figure 10 should be performed twice, generating smoothed values of both the occurrence ( $\bar{O}_{x,t}$ ) and the exposure ( $\bar{E}_{x,t}$ ). The smoothed mortality rate are then calculated as the ratio between ( $\bar{O}_{x,t}$ ) and ( $\bar{E}_{x,t}$ ), and (hopefully) the randomness in this estimation is strongly reduced compared to the non-smoothed estimation.

---

<sup>3</sup>Compare with the uniform kernel function, where all observations have the same weight, and the smoothed value is just the empirical average of the observations inside the interval  $[x \pm b_1, t \pm b_2]$ .

#### 4.1.1 Kernel Smoothing at Boundaries

The idea with the Kernel smoothing technique is to borrow information from the cells in the neighborhood of the origin cell to get a more stable mortality estimation. The bandwidths  $b_1$  and  $b_2$  determine the number of extra cells to be used for the estimation. The Epanechnikov kernel are defined in such a way that we will use the same number of "extra" cells on the right hand side of the origin cell as on the left hand side. Further, we will use the same number of cells above as below the origin cell. However, for some areas in the Lexis diagram, we have to make exceptions from this symmetric. For example, if we want to estimate the smoothed mortality for individuals aged  $x$  in year 2004, (i.e.  $\bar{m}_{x,2004}$ ), it is not possible to use information from calendar years later than 2004, since we only have data available until 2004. We then have to be content with using information to the left of the origin cell in the Lexis diagram. This is illustrated in the Lexis diagram in figure 11, where we are estimating the mortality for an individual aged 90 in year 2004. Here we are only using information from 15 (3x5) cells (instead of 25 as we did in figure 10). The technique explained here may also be used for certain ages, where we are not able to get information for some adjacent ages. If we for example want to estimate the mortality at the age of 110, and we have no data for ages above 110 we will have a similar problem.

Comparing the situations in figures 10 and 11, we will see that the total sum of the weights, created from the Epanechnikov kernel in equation 14 and 15, will not be the same for the both cases. This is because we have no restrictions on the weights. In figure 10, the sum consists of 25 individual weights, while the sum in figure 11 only consists of 15 individual weights. Therefore, the sum of the weights are considerably greater for the former sum. This is not a problem though, and a normalization of the weight will not be necessary. The reason is that we are smoothing the occurrence and the exposure individually, and the sum of the weights will always be the same

for the occurrence matrix as for the exposure matrix.<sup>4</sup>

#### 4.1.2 Bias Problem at Boundaries

In cases, like in figure 11, when we are not able to use the fully kernel there might appear bias problem at the boundaries. This boundary bias problem for the kernel estimator is well known in the literature, see for example Hall and Park (2002) who are proposing methods for solving these kind of problems. In figure 11 the bias problem appears since we are using data from the years 2002-2004 for estimating the mortality in 2004.<sup>5</sup> If the overall trend is a decreasing mortality we may therefore overestimate the mortality at these right endpoints.

## 4.2 Non-Smoothed and Smoothed Mortality Rates

The effect of the kernel smoothing technique is illustrated in figure 12. The dotted line describes the observed mortality rates for Danish females aged 65 during 1900-2004. The full line represents the smoothed mortality rates.

The effect of the smoothing technique can also be illustrated in a three dimensional graph, as in figure 13, where we are comparing the non-smoothed to the smoothed mortality rates for Swedish females aged 65-99 during 1900-2004. From the figure we see that the non-smoothed mortality rate surface are much more thorny than the smoothed surface, especially for the higher ages. The bandwidths used are  $b_1 = 4$  and  $b_2 = 4$ . The way the Epanechnikov kernel is defined implies that the endpoints in the interval will receive

---

<sup>4</sup>If we instead would have adapted the smoothing to the mortality rate directly, it would have been necessary with a normalizing constraint.

<sup>5</sup>If the mortality would have been independent of the chronological time, then we would not have a boundary bias problem in this situation.

zero weights<sup>6</sup>. In this case, Since the data only consist of integer ages and years, the smoothed occurrence (exposure) will be a weighted average of the observations  $[x \pm 3, t \pm 3]$ . In figure 14 we are illustrating the same comparison for Danish females. In the graph for the non-smoothed mortality rates we see an example of the problems arising when the exposure is small. For some years in the beginning of the 20th century, we have a mortality rate estimation of zero, due to lack of exposure for that certain age and year. As discussed in section 3 this is of course unrealistic. But looking at the smoothed graph, we see that the problem is solved, and the surface is much smoother. In figures 15 and 16 the same illustration are made for Swedish and Danish males respectively. The problems with lack of exposure is even more obvious for males than for females, since males have a shorter life expectancy. The problem is also more obvious for Denmark than for Sweden.

### 4.3 Optimal Bandwidth - Bias vs Smoothness

As mentioned earlier, the local constant kernel smoothing estimator is biased. It will significantly underestimate the true mortality for high ages. The reason is a rapidly increasing mortality and rapidly decreasing exposure for those ages. (see for example Fledelius et al. (2002)). When performing the smoothing technique for the exposure for a certain (high) age  $x$ , we will have a situation where the exposure for ages just lower than  $x$  is much higher than the exposure for ages just higher than  $x$ . These cells with a relatively high exposure will have a great impact when summing up the the exposure for the cells included in the "smoothing area". The denominator in (11) will therefore be "too" large, and the ratio will become "too" small.

The bandwidth vector  $b = (b_1, b_2)$  determines the smoothness of the surface.

---

<sup>6</sup>when dealing with continuous distributions, it does not matter whether the endpoints are included or not

A larger bandwidth will increase the smoothness, but it will also increase the bias. A smaller bandwidth will reduce the bias, but will instead increase the variance and lower the smoothness. This property is clearly illustrated in figure 17 where we are presenting the mortality for Danish females in year 2004. As we can see from the figure, the raw mortality is not very smooth for higher ages. When selecting a bandwidth of  $b = (10, 10)$ , we have a very smooth curve, but we are underestimating the mortality substantially, i.e. we have a very large bias. When choosing a bandwidth of  $b = (4, 4)$ , the curve seems to fit the raw data better. The curve is rather smooth but we still seem to have a bias problem for higher ages, since the most of the raw data points lie above the smoothed curve. This fact may be a problem when performing estimations and predictions.<sup>7</sup>

#### 4.4 Bias Correction Technique

In figure (17) we saw that even when choosing a rather small bandwidth, we still had a bias problem. To be able to get reliable mortality estimations this bias problem needs to be taken care of. Here we are applying a bias correction technique used in Fledelius et al. (2002). The technique can be explained in several steps. First we are creating a preliminary estimator of the mortality. This estimator, which we call  $\tilde{\alpha}_{x,t}$ , is created by using the local constant smoothing technique with very a large bandwidth. Here we are using a bandwidth of  $b = 10, 10$ .

$$\tilde{\alpha}_{x,t} = \frac{\bar{O}_{x,t}}{\bar{E}_{x,t}} \quad (16)$$

The idea is that this preliminary estimator should be very smooth. We are

---

<sup>7</sup>Here we have only illustrated the bias problem for Danish females in year 2004, but the situation is similar for both genders, in both countries, and for different years.

then performing another local constant smoothing. The difference this time is that the exposure  $E_{x,t}$  in the denominator is replaced by  $E_{x,t} * \tilde{\alpha}_{t,x}$ , i.e. the exposure multiplied by our preliminary estimator of the mortality. The numerator is, just as before, the raw occurrence. This second smoothing is performed on the numerator and the denominator separately and we are here using a bandwidth of  $b = (4, 4)$ . This second smoothing gives us an estimator which we call  $g_{x,t}$ <sup>8</sup>.

$$g_{x,t} = \frac{\bar{O}_{x,t}}{E_{x,t} * \tilde{\alpha}_{x,t}} \quad (17)$$

The final step is to multiply the estimator  $g_{x,t}$  with the preliminary estimator of the mortality  $\tilde{\alpha}$ .

$$\tilde{m}_{x,t} = g_{x,t} * \tilde{\alpha} \quad (18)$$

The idea of the way this correction technique works may seem a little bit unclear, but the effect of using it is exceptional. In figure 18 this bias correction technique is illustrated. The figure is an extension of figure 17, where we have included an extra curve for the estimation with the bias correction applied. This estimation fits the raw data better than the ordinary local constant kernel smoothing estimation, and there is no visible bias problem.

Even if the bias correction technique described above is applied, we may still have a bias problem at the boundaries (see section 4.1.2). The bias correction technique is not taking care of the problem that may arise at boundaries, where we are not able to use the full kernel. The boundary bias problem may arise both at certain ages and at certain years. However, in this thesis we will not have a problem in the age dimension. As we will explain in section 6 we will only consider the ages 65-99, and since the data covers the ages 0-110

---

<sup>8</sup> $g_{x,t}$  should not be considered as an estimator of the mortality.



we will be able to use the full kernel for the smoothing procedure in the age dimension. Nevertheless, the problem is inevitable at the right boundary for the calendar year dimension. For the left boundary there will be no problem since we are able to use information from prior years. For the right boundary the problem appears, since we are not able to use future information (see figure 11).

## 5 The Lee-Carter Method

In this section we will describe the Lee-Carter method (Lee and Carter (1992)), which consists of a mortality model and a methodology of fitting the model. The method also contains a time series model of the dominant parameter, which can be used for forecasting (Booth et. al. (2002)).

### 5.1 The Model

In section 3 we studied the development of the expected lifetime and the mortality rate, and it is obvious that neither of these has been constant over time. When trying to construct a model describing the mortality rate it is therefore desirable to have a model depending on time. The Lee-Carter model, proposed in Lee and Carter (1992), has become the leading statistical mortality model in the demographic literature (Deaton and Paxson, 2001). It is a parsimonious demographic model combined with statistical time-series methods. Let  $m(x, t)$  denote the central mortality rate for age  $x$  in year  $t$ . The method proposed by Lee and Carter models the logarithm of the mortality rate:

$$\ln(m_{x,t}) = a_x + b_x k_t + \varepsilon_{x,t} \quad (19)$$

The different parameters can be explained as follows

- $a_x$  is a set of age-specific constants describing the general pattern of mortality at different ages.
- $k_t$  is an index describing the general level of mortality at different times.  $k_t$  captures the main trend in death rates at all ages. Since the overall trend is a decreasing mortality, one can expect the index to be decreasing as well<sup>9</sup>

---

<sup>9</sup>This statement is dependent on that  $b_x > 0$ . If  $b_x < 0$  then the overall trend of a decreasing mortality corresponds to an increasing  $k$  index. If  $b_x = 0$ , then the mortality does not depend on time at all.

- $b_x$  is a set of age-specific constants describing the relative speed of mortality changes, at each age, when  $k_t$  changes.  $b_x$  modifies the main time trend according to whether the change at a particular age is faster or slower than the main trend, and if the change is in the same or the opposite direction. The model allows for both negative and positive values of  $b_x$ . A negative value of  $b_x$  indicates that the mortality for that age is rising with an increase in time<sup>10</sup>. However, in practice, this does not seem to occur in the long run. So, when the model is fitted over fairly long periods, then all  $b_x$  have the same sign (Lee and Miller 2001). When  $b_x$  is large for some  $x$ , then the mortality rate at that age differs a lot for different times. When  $b_x$  is small, the mortality rate at age  $x$  shows small changes for variations in the main trend, which is often the case with mortality for older ages (Lee and Carter 1992).
- $\varepsilon_{x,t}$  denotes the error term, which is not captured in the model.

## 5.2 Fitting the Model

The parameters to be estimated in the model are  $a_x$ ,  $b_x$  and  $k_t$ . For a given matrix of mortality rates,  $m_{x,t}$  we therefore seek the least square solution to (19). The model is overparameterized, and there is therefore no unique solution. The parametrization is invariant under either of the transformations (Renshaw and Haberman (2003b))

$$\begin{aligned} \{a_x, b_x, k_t\} &\mapsto \{a_x, b_x/c, ck_t\} \\ \{a_x, b_x, k_t\} &\mapsto \{a_x - cb_x, b_x, k_t + c\} \end{aligned}$$

for any constant  $c$ . This means that  $a_x$  is only determined up to an additive constant,  $b_x$  is determined up to a multiplicative constant, and  $k_x$  is determined up to a linear transformation<sup>11</sup>. This is not a problem though; it only means that the likelihood function associated with the model has an infinite

<sup>10</sup>Conditioning on that the time index is decreasing.

<sup>11</sup>These observations are connected with the discussion above about the sign of the  $b_x$ 's.

number of equivalent maxima, and all of them would produce identical forecasts (Giroi and King (2005)). To obtain a unique solution we just have to impose some restrictions. Lee and Carter are letting the  $b_x$ 's sum to unity and the  $k_t$ 's sum to 0.<sup>12</sup>

$$\sum_x b_x = 1 \quad \sum_t k_t = 0 \quad (20)$$

The implication of the second restriction is that  $a_x$  is the empirical average over time for the logarithmic mortality rate for people aged  $x$  (Giroi and King (2005)). This means that, even though the  $a_x$ 's are time independent, the estimations of the  $a_x$ 's will depend on the historical period we are using for our estimations. If we change the historical period, the average of the logarithmic mortality will change, and so will the  $a_x$ 's.

Since we only have parameters on the right hand side of (19), and no observed variable, we cannot use ordinary regression methods for solving the model (Lee and Carter 1992)<sup>13</sup>. Maximum likelihood methods can be used, but the multiple maxima or the constraints will make standard optimization programs work poorly (Giroi and King (2005)). To find the least square solution, Lee and Carter suggest to use the singular value decomposition (SVD) method, applied to the logarithmic mortality matrix after the  $a_x$ 's have been subtracted. This way we will get estimates of the  $b_x$ 's and the  $k_t$ 's.

### 5.2.1 Second Stage Estimation

When the parameters are known, we are able to calculate the theoretical mortality rates. Further, applying these rates to the population data we can

---

<sup>12</sup>These normalization together with the overall trend of decreasing mortality will probably lead to that the  $b_x$ 's are positive and that the time index is increasing over time.

<sup>13</sup>However, the estimates of the  $a_x$ 's, as the empirical average over time, are still least square estimates due to the normalizing constraints in (20) (Renshaw and Haberman (2003)).

derive the number of deaths generated by the model, for each age and each year. The theoretical number of deaths for each year will in general not be equal to the actual number of deaths (Lee and Carter (1992)). The reason is that all ages have received the same weight in the SVD, regardless of the size of the mortality rate. This means that the error terms,  $\varepsilon_{x,t}$ , corresponding to small  $x$ -values will have the same weights as the error terms corresponding to large  $x$ -values, yet the contribution to the total number of deaths is much smaller from youth than for older people. Lee and Carter therefore suggest a reestimation of the  $k_t$ 's, where the values of the  $a_x$ 's and  $b_x$ 's from the first estimation are fixed, so that the actual number of deaths for each year  $t$  will equal the deaths generated by the model for that year. This second stage estimation is done by an iterative search.

In this thesis however, we are not applying this reestimation of the  $k$ -index. The reason is that our intention is not to fit the number of deaths exactly for each year. We have assumed the mortality surface to be smooth, and to fulfill this assumption we are applying the kernel smoothing technique. This way we have reduced some of the randomness included in the raw mortality rate estimations, especially for higher ages when the exposure is small. The raw mortality rate estimations are not very smooth, and the intention is not to fit the exact mortality rate for each year and each age. Instead we want to have a smooth model that "on average" captures the main trend of the mortality. For the same reasons we are not interesting in a model that for each year exactly generates the actual number of deaths for that year.

## 5.2.2 Studying the Mortality Index

When Lee and Carter developed their model they used U.S. data from 1900-1989. Studying the estimations of the mortality index, the  $k_t$ 's, they found some interesting facts.

- During 1900-1989 the mortality index was decreasing roughly linearly. The decline in the first half of the estimation period was about the same as the decline in the second half. This was interesting since the development in life expectancy during the same period was definitely not linear<sup>14</sup>. This appeared to be an advantage by modeling the death rates instead of modeling the life expectancy.
- They also found that the short-run fluctuations in  $k$  was about the same during the whole time period<sup>15</sup>.

The linearity and the relatively constant variance are very convenient features for prediction purposes.

### 5.3 Mortality Forecasts

In order to predict the future mortality, we only have to predict the evolution of the mortality index<sup>16</sup>. Lee and Carter predict the mortality index by an univariate time series model. They try several ARIMA specifications but suggest that a random walk with drift describes the index well. They suggest the following model:

$$k_t = k_{t-1} + c + e_t \quad (21)$$

where  $c$  is the drift and  $e_t$  is the error term. The only parameter to be estimated is the drift term ( $c$ ).  $c$  is estimated by calculating the slope of the

---

<sup>14</sup>During 1900-1944 the life expectancy at birth in USA rose by 17.6 years. The increase in life expectancy during 1944-1988 was 9.9. The decline in the time index for the same periods was 15.8 and 16.7 respectively.

<sup>15</sup>Except for the influenza epidemic in 1918.

<sup>16</sup>Both the  $a_x$ 's and the  $b_x$ 's are age specific (do not depend on time) and therefore, they do not need to be predicted.

line that is drawn through the first and the last observation of the  $k$ -index. When this is done, we are able to predict forecast of the future mortality rates. The predictions of the future mortality index is then produced by extrapolation of this line. This is illustrated in figure 19 where the mortality index for Swedish females aged 65-99 is estimated during 1970-2004 and predicted during 2005-2039.

To handle the influence epidemic in 1918, Lee and Carter introduce a dummy variable for that year. Due to the influence epidemic in 1918, the number of deaths during that year was unusually high. When using the kernel smoothing technique, the effect of this disease has been diminished, since the mortality estimations are then based on information during more than one calendar year. Therefore, we do not find it necessary to introduce a dummy variable for that year.

## 6 Procedure

### 6.1 Backtesting

We have mentioned the great importance for life insurance companies of having accurate forecast of the future mortality. Unfortunately, you cannot determine the performance of the predictions until afterwards. Nevertheless, there is still reasons for carefully considerations of the way you are generating the mortality predictions. There are a lot of possible ways for producing forecasts. The common procedure is to consider the historical mortality and try to fit a model to the historical data. Predictions can be produced by extrapolation of the parameters in the estimated model. It is very important to be aware of the fact that a good historical fit (between empirical data and the estimated model) does not guarantee good future predictions. By applying more complicated models and including new parameters it is often possible to improve the historical fit. However, when going on to the prediction purposes, there may come up problems. First, as Booth et al. (2002) points out, a large number of parameters will limit the potential for extrapolation, and therefore complicate the predictions. Second, even if extrapolation is possible and predictions are made, the predictions may be unstable and unrealistic, with large prediction errors (see Fledelius (2003)). A good fit to the past does not guarantee a good fit in the future, and our objective is exactly a good fit in the future. So, when choosing a mortality model it is the prediction capacity (rather than the estimation capacity) of the model that should be of primary interest.

Since the future is unknown, one can never know for sure if a certain model will be good at predicting the future mortality. However, by evaluating the historical prediction performance of the model, we may get a hint of its capacity. If a model historically has performed well, one has stronger reasons for believing that it will perform well also in the future. In this thesis we will examine the historical prediction performance of the LC method. We



will use a part of the data for estimating the model, and the rest part for comparison to the predictions generated by the estimated model. By this backtesting technique we are able to get a hint of the potential prediction capacity of a certain model. Backtesting is therefore an important tool for validating a prediction model.

## 6.2 Period and Cohort Expected Lifetime

In figures 1 and 3-4 we are illustrating the expected lifetime for men and women in Sweden and Denmark. The expected lifetime for a certain year is calculated by studying the age of the individuals that die in that year. This way of measuring is called the period life expectancy, and should not be mixed up with the cohort life expectancy, i.e. the expected lifetime for individuals born in a certain year. To calculate the expected lifetime for individuals born in a certain year we have to follow the whole cohort of individuals that were born in that specific year, and the value of the expected lifetime can therefore not be known exactly in more than 100 years afterwards. So, it is important to distinguish between expected lifetime by year of death (period expected lifetime) and expected lifetime by year of birth (cohort expected lifetime)<sup>17</sup>. In the HMD database, from where we have taken our data, both these measures of the expected lifetime are available. As mentioned before, we have data for the period life expectancy up to year 2004. But for the cohort life expectancy, data are only available up to year 1913. In figure 20 the difference between the both life expectancy measures are illustrated. In

---

<sup>17</sup>In this thesis the "cohort expected lifetime" and "expected lifetime by year of birth" will have the same meaning. Nevertheless, in this case, it may feel a little bit strange to call it "expected lifetime by year of birth", since we are not studying the birth year at all. We are only studying the calendar year and the age of the individuals, and by knowing these two co-ordinates we get information also about the birth year according to the relationship: cohort = period - age. However, the birth year cannot be known exactly since we are working with discrete time.

all graphs, the upper full line represents the expected lifetime by the year of birth (cohort), while the lower dashed line represents expected lifetime by year of death (period). The time period in the figure is 1875-1913. As we can see from the graphs, the trend is similar for males and females and for Sweden and Denmark, and there is quite a big difference between the two lines.

A problem when investigating the cohort expected lifetime is that you need to have information for a very long time ahead. The expected lifetime can be calculated from the mortality rates, according to (7) and (8) in section 3, but if you want to calculate the cohort expected lifetime for a newly born individual you need to have predictions of the mortality rates for about 100 years ahead. This may be a hard problem in practice since it may be difficult to have accurate predictions such a long time into the future.

### **6.3 Evaluation**

In this thesis we are evaluating the performance of the Lee-Carter method in predicting future mortality rates. We have focused on individuals aged 65-99, since we think that these ages are extra relevant in the perspective of an insurance company. For the evaluation process we are using all data from year 1900 and forward. The idea is to first use data only from a certain limited time period, which we will call the estimation period. For this data we are then performing the local constant kernel smoothing technique with bias correction. By doing so we will get smooth estimates for the mortality rate for every year in the estimation period and for every single age in the age interval 65-99. We will then adapt the Lee-Carter model to this smoothed data. The parameters in the LC model will be estimated by the Singular Value Decomposition (SVD) technique.

When all the parameters in the LC model are estimated it is easy to create

forecasts of the mortality index, by just extrapolating from the latest year in the estimation period. Since the mortality index is the only parameter that needs to be predicted, we will then also easily get forecasts for the future mortality rates. For a time period, called the prediction period, and starting at the year immediately after the end of the estimation period, we will perform predictions of the future mortality rates.

The final step in the LC evaluation process is to compare the predicted mortality rates with the actual mortality rates observed. To minimize the randomness, we are applying the smoothing technique (with bias correction) to the observed mortality rates before we do the comparison. This is because we want to minimize the effect of the inherent variation of the mortality for individual observations. By doing this comparison we are able to say something about the historical performance of the LC method. It is clear that the closer the predicted mortality rates are to the observed rates the better the model is<sup>18</sup>. A natural thought that appear is how far away from the observed rates the predicted rates are allowed to be in order to call it a "good" prediction. It may be difficult to determine the performance of the model by just studying the mortality rates themselves. The mortality rates are actually intensities and may sometimes be hard to interpret.<sup>19</sup> Just like Lee and Miller (2001), who are evaluating the performance of the LC method on American data, we will mainly study the expected lifetime to evaluate the LC performance. Since we are only studying the ages 65-99 we will consider the expected remaining lifetime for individuals at ages in this interval.

Let's say, for example, that the estimation period ends at calendar year  $t$ , i.e. we have used data up to year  $t$  to estimate the parameters in the LC model.

---

<sup>18</sup>at least in prediction purposes.

<sup>19</sup>In this case, however, since we are studying the mortality at discrete, one year interval, time points, the mortality rates can be seen as a probabilities for dying during a certain year. This fact will make it easier to interpret the meaning of the intensity. But it may still be difficulties to explain differences between predicted and observed mortality by just looking at the mortality rates.

The prediction period does then start at year  $t+1$ , and we are predicting the future mortality rates from year  $t+1$  on. By using the predicted mortality rates it is possible to calculate the expected remaining lifetime. The life expectancy is calculated by summing up the contribution in lifetime for each single age up to the age of 99. If we want to calculate the cohort expected remaining lifetime for an individual aged 65 in year  $t+1$  we therefore have to study the following:

- the mortality rate for a 65 year old in year  $t+1$
- the mortality rate for a 66 year old in year  $t+2$
- ...
- the mortality rate for a 99 year old in year  $t+35$

This means that we need predictions 35 years ahead to perform a prediction of the cohort expected remaining lifetime of an individual aged 65 in year  $t+1$ .<sup>20</sup> If we would like to calculate the cohort expected remaining lifetime of a 66 year old in year  $x+1$ , the procedure is the same. We will have to study

- the mortality rate for a 66 year old in year  $t+1$
- the mortality rate for a 67 year old in year  $t+2$
- ...
- the mortality rate for a 99 year old in year  $t+34$

In this case we "only" need 34 years future predictions. For calculating the life expectancy for a 67 year old we need 33 future predictions and so on.

For calculating the "usual" expected remaining lifetime in year  $t+1$ , i.e. the life expectancy for individuals that dies in year  $t+1$ , we are only using the

---

<sup>20</sup>We are here ignoring the contribution in life expectancy for ages higher than 99. However, this contribution should have no large impact.

mortality rates in year  $t+1$ . Except the fact that we are only using mortality rates from the same year ( $t+1$ ), the procedure for calculating the life expectancy is the same as for the cohort life expectancy.

The formulas (7) and (8) for calculating the life expectancy in section 3 were only considering the age as an independent variable. Since we are also considering the calendar year as an independent variable, we have to introduce the "time" dimension in the formulas. The period life expectancy for an individual aged  $x$  in year  $t$  can be calculated as

$$e_{x,t} = \left( \sum_{i=0}^{\infty} \frac{l(x+i,t)}{l(x,t)} \right) - 0.5 \quad (22)$$

The mortality rates used in the calculation are all from the same calendar year (year  $t$ ). The cohort life expectancy for an individual aged  $x$  in year  $t$  can be calculated as

$$e_{x,t} = \left( \sum_{i=0}^{\infty} \frac{l(x+i,t+i)}{l(x,t)} \right) - 0.5 \quad (23)$$

Here we are using mortality rates for different years in the calculation.

The predicted remaining lifetime should then be compared to the observed remaining lifetime. The observed remaining lifetime are calculated by using the observed mortality rates in the same way as for the predicted remaining lifetime.<sup>21</sup> To be able to compare the predicted cohort expected remaining lifetime for a 65 year old to the observed remaining lifetime we cannot use data later than 1969 for the estimation period, since we need 35 years predictions, and we only have data up to 2004 available. For the study of a 66 years old we can at most use data up to 1970 and so on. When studying the

---

<sup>21</sup>By using the same procedure as before, the concern that we are ignoring contributions from individuals aged 100 and above can be neglected.

expected lifetime by the year of death, we do not have these restrictions, and can therefore use more recent data for the estimation.

Since we would like to base our LC evaluation on as much information as possible, we will use several different estimation periods, where each of them will give an unique prediction.

Before starting the evaluation process we also have to decide the length of the estimation period. We decided to first use an estimation period of 35 years.<sup>22</sup> This resulted in the following estimation/prediction periods:

- Period 1: Estimation: 1900-1934 Prediction: 1935-1969
- Period 2: Estimation: 1901-1935 Prediction: 1936-1970
- ...
- Period 71: Estimation: 1970-2004 Prediction: 2005-2039

Thus, when using a length of 35 years of the estimation period, we will get 71 different periods. Each of them will have an unique set of estimates for the parameter in the LC model. Each of them will also have an unique set of predictions based on the estimated parameters. But, as mentioned above, we will not be able to use all predicted mortality rates in the evaluation process since we only have data until 2004. The last period (Period 71) for example cannot be used for evaluation at all since the first prediction year in this case is 2005.

## 6.4 Varying the Length of the Estimation Period

This far the length of the estimation periods has been fixed to 35 years. By using the model estimated from these 35-years estimation periods, we have predicted the future mortality. The quality of the predictions should then

---

<sup>22</sup>Since we are supposed to perform predictions 35 years ahead it may seem natural to also have an estimation period of 35 years.

be evaluated by comparing the predicted mortality to the observed mortality. An interesting thought is how the performance of the predictions would have changed if we had changed the length of the estimation period<sup>23</sup>. If it is possible to improve the performance of the predictions by changing the estimation period, it may be of great interest to find the optimal length of the estimation period.<sup>24</sup>.

### 6.4.1 Finding the Optimal Length

We now have to introduce some further notation for the remaining life expectancy. Earlier we defined the remaining life expectancy for an individual aged  $x$  in year  $t$  as  $e_{x,t}$ . At this stage it is suitable to distinguish between predicted,  $e_{x,t}^p$ , and observed,  $e_{x,t}^o$ , remaining life expectancy. For the predictions we have to introduce an additional index, indication the number of years,  $z$ , for the estimation period. So, the predicted remaining life expectancy for an individual aged  $x$  in year  $t$ , with an estimation period of  $z$  years, will then be denoted  $e_{x,t,z}^p$ .<sup>25</sup> We will let the length of the estimation periods vary between 2 and 75 years. In all cases the first prediction year (in Period1) will be 1935, just as before. We will try to find the optimal length of the estimation pe-

---

<sup>23</sup>This approach is different from the approach performed by Lee and Miller (2001). They are predicting the mortality in 1998 by using different "jump-off" years. They are always using 1900 as the first year for the estimation period, while they are varying the last year for the estimation period. When the estimation period is increased, the prediction period is decreased. As the length of the prediction period decreases, the uncertainty of the predictions will be lower. From this approach it is difficult to evaluate the optimal length of the estimation period.

<sup>24</sup>By optimal in this context, we mean the length that historically has been the optimal length. There is absolutely no guarantee that this length will be optimal also for future predictions. But, as mentioned in section 6.1, if the procedure has been successful historically one may have stronger expectation about the performance in the future as well

<sup>25</sup>This  $z$ -index is only relevant for the predictions. For the observed life expectancy there is no estimation period at all.

riod for each single age between 65 and 99. In the case of a 65 year old, we therefore has to compare the predictions  $(e_{65,1935,z}^p, e_{65,1936,z}^p, \dots, e_{65,1970,z}^p)$  to the observed  $(e_{65,1935}^o, e_{65,1936}^o, \dots, e_{65,1970}^o)$  (36 values to compare). In the case of a 66 year old, we instead has to compare  $(e_{66,1935,z}^p, e_{66,1936,z}^p, \dots, e_{66,1971,z}^p)$  to  $(e_{66,1935}^o, e_{66,1936}^o, \dots, e_{66,1971}^o)$  (37 values to compare). For higher ages the same procedure is applied. When we increase the age by one year, we get one further value that can be used in the evaluation.

To find the optimal length we will have to define some technique for measuring the performance of the predictions. As mentioned before, we are mainly using the life expectancy, rather than just the mortality rates, for evaluation of the performance. We therefore think it is natural to also base the performance measuring on the life expectancy. We will use three different measures, which are presented below, to find the optimal length of the estimation period. For the illustration we will consider the case of a 65 year old. For each of them we will find the  $z$  that minimizes the following quantities.

### 1. Mean Absolute Error

$$\frac{|(e_{65,1935,z}^p - e_{65,1935}^o)| + \dots + |(e_{65,1970,z}^p - e_{65,1970}^o)|}{36} \quad (24)$$

### 1. Mean Square Error

$$\frac{(e_{65,1935,z}^p - e_{65,1935}^o)^2 + \dots + (e_{65,1970,z}^p - e_{65,1970}^o)^2}{36} \quad (25)$$

### 1. Max Error

$$Max((e_{65,1935,z}^p - e_{65,1935}^o), \dots, (e_{65,1970,z}^p - e_{65,1970}^o)) \quad (26)$$

One can discuss which one of these three measures that is most appropriate



for determining the optimal length of the estimation period, but hopefully the result will be similar for all the three of them. This procedure will be implemented for all ages (65-99), for both genders and for both countries.<sup>26</sup>

---

<sup>26</sup>These are just three possible ways, out of many, for measuring the performance. An alternative approach is to use different weights for the prediction errors. One could then, for example, use the exposure for determining the weights, A large exposure will correspond to a large weight an so on.

## 7 Results

In this section we are presenting the result together with some clarifications about the procedure. We will give some comments to the result, but the main comments, analysis and conclusions will be given in section 8. The first part of this section (7.1) includes the result when we had a fixed length (35 years) of the estimation periods. In the second part of this section (7.2) we are presenting the result when having different lengths of the estimation periods. From these result we are then trying to determine the optimal length of the estimation period.

### 7.1 35 Years Estimation Periods

In this section we have applied a fixed length of the estimation periods. We are using 35 years for estimating the LC model, and we are then predicting the future mortality rates for the next 35 years. According to the procedure described in section 6 we have 71 different time periods, each consisting of 70 calender years, where the first 35 years are for estimation and the 35 last years are for prediction. (For the first time period, 1935 is the first prediction year, while for the last time period, 2005 is the first prediction year.) For each of the 71 time periods we have produced a matrix of mortality rates. The dimension of each matrix is 35x70, where the 35 rows are representing the ages from 65 to 99, and the 70 columns are representing the 70 calender years. Thus, columns 1-35 are LC estimations, and columns 36-70 are LC predictions.

#### 7.1.1 Mortality Rates

As mentioned in section 6 it may be difficult to determine the prediction performance of the LC method by just studying the predicted mortality rates. Our main evaluation tool in this thesis will instead be the study of the life ex-

pectancy. However, we will still present some graphs illustrating the comparison between predicted and observed mortality rates. A problem for this kind of graphical presentations is that it is difficult to get an extensive overview of the situation without presenting too many graphs. The reason is that there are several dimensions to take care of. We both have to consider different ages and different years. Further, in this case we have applied the LC method 71 times (for each of the 71 different time periods). It is impossible to make a complete illustration in a single graph since it would require four dimensions, and then we have not even considered which gender and which country we are studying.

In figure 21 we are comparing the observed smoothed mortality rates to the rates estimated and predicted by the LC method. The comparison is made for Swedish females, and we have chosen the time period 1935-2004. There are six graphs, and each graph illustrates a fixed age (65, 70, 75, 80, 85 and 90). The dotted line represents the observed smoothed mortality. The full line represents the estimated (1935-1969) and the predicted (1970-2004) mortality. The vertical line separates the estimation and the prediction period. In all six graphs, the estimated (1935-1969) mortality seems to fit pretty well to the observed mortality, at least for most years. However, in all graphs, except for the age of 65, there is a visible discrepancy between estimated and observed mortality just at the last 2-3 estimation years. It seems like the mortality is overestimated at those years. Further, the overestimation at those years seems to increase with the age. We suspect that this overestimation may occur due to the boundary bias problem discussed in section 4.1.2. We will return to the discussion of this problem in section 8. Going on to the prediction performance, we see that for the age of 65 the result looks very good. The predicted mortality rates are very close to the observed mortality rates. When studying higher ages, it seems like the LC method is considerably overpredicting the true mortality. The overprediction seems to increase with the age.

The situation for Danish females, Swedish males and Danish males are illustrated in figures 22-24. For Swedish males we see a clear tendency that the LC method is overestimating the mortality, at least at the later part of the prediction period. For Danish females and males the situation is not that obvious. For the highest ages however, it still seems like we have an over-prediction problem. A recurring property is, like for Swedish females, that there is a discrepancy between estimated and observed mortality for the last years of the estimation period.

### **7.1.2 Cohort Expected Lifetime**

As discussed in section 6, the way you measure the expected lifetime is of great importance. There are significant differences between the year-of-birth (cohort)- and the year-of-death (period) expected lifetime (se figure 20). Our focus in this thesis is on the life expectancy by year of birth, and we will in this section present some result concerning the cohort expected lifetime.

In section 7.1.1 we compared the predicted and the observed mortality by studying the mortality rates. We observed that on some occasions the result looked good, while sometimes the predicted mortality where far from what was observed. As mentioned in section 6.3, it may be difficult to evaluate the performance of the predictions by just comparing the mortality rates. It may also be difficult to determine the consequences of a poor prediction. If we are studying the life expectancy instead, it may be easier to evaluate the performance of the predictions. When considering the difference (in years) in life expectancy between predicted and observed mortality we get a good understanding of the actual performance of the LC method. The life expectancy is also easier (than mortality rates) to understand in everyday speech.

The life expectancy is calculated by summing up survival probability contri-

butions from different ages, according to what was described in section 6.3. The expected lifetime for a 65 year old was calculated by studying 35 different mortality rates. The difference in life expectancy (between predicted and observed) for a 65 year old can therefore be seen as a summary of the performance of these 35 different predicted mortality rates. The considering of the life expectancy is therefore a more convenient way for presenting foreseeable results compared to the considering of the mortality rates directly. Of course, there may be some disadvantages enclosed in this procedure. For example, we could have a situation where the predicted mortality rates are far away from the observed mortality. If some of the predictions are too high and some of them are too low, these prediction errors may compensate each other, and we may see a good result when considering the life expectancy. However, on the whole, we think that it is more appropriate to consider the life expectancy than just the mortality rates for the LC evaluation process.

### **Female, Sweden**

In figure 25 we are comparing the predicted and the observed remaining life expectancy for Swedish females. The figure consists of six graphs, where each graph represents a fixed age (65, 70, 75, 80, 85 and 90 years). The full line in each graph represents the predictions and the dotted line represents the observed values. For the remaining life expectancy for a 65 year old we only have observations until 1970. As explained earlier, this is because we need 35 future years for the calculation of the remaining life expectancy for a 65 year old. For ages higher than 65 we need less than 35 years for the calculation, and we are therefore able to make the comparison also for years later than 1970. For the age of 65 the predicted remaining life expectancy is considerably lower than the true expected lifetime during the whole time period. For higher ages, we still have an underprediction problem, but the problem seems to decrease with the age. For the 85 and 90 year old, the predictions for the latest years is above the true remaining life expectancy. For the years around 1945 we notice a peak for the predictions.

### **Female, Denmark**

Figure 26 illustrates the same graphs for Danish females. Comparing figure 25 and 26, we notice some similarities. Also for Danish females we tend to have an underprediction problem. The problem seems to decrease with the age. For the highest ages, we even notice some years where the predictions are above the observed remaining life expectancy.

### **Male, Sweden**

For Swedish males, figure 27, the graphs look different compared to figures 25 and 26. The increase in remaining life expectancy over time has been much smaller than for females and the underprediction problem that we observed earlier is almost disappeared. We are experiencing some underprediction during the first part of the time period, but not at the same size as in the previous cases. Even if the predictions are a bit unsmooth at some years, it seems like the predictions overall are much closer to the observed values for all ages. The increase in remaining life expectancy for Swedish males seems to be smaller than for Swedish and Danish females.

### **Male, Denmark**

The same graphs for Danish males are illustrated in figure 28. Here we notice some interesting properties. The increase in the observed remaining life expectancy over time is smaller than in all previous cases. During some years we even have a decrease in the remaining life expectancy. Studying the predictions we see that these are rather unsmooth. Also, we do not seem to notice the underprediction problem that we have observed earlier.

### **Summary**

By studying figures 25-28 we are able to draw some conclusions. The increase in remaining life expectancy over time has been greater for females than for males. We also notice a greater increase in Sweden than in Denmark. The unsmoothness in the predictions is greater in Denmark than in Sweden and for males than for females. For Swedish and Danish females we tend to have

an underprediction problem for the lowest ages. For males, we do not seem to have this problem. In the following table the performance of the LC predictions are summarized. The table shows the mean absolute difference (in years) between the LC predicted and the observed life expectancy.

Age	65	70	75	80	85	90
Female, Sweden	0.94	0.58	0.36	0.21	0.13	0.09
Female, Denmark	0.96	0.60	0.36	0.21	0.13	0.08
Male, Sweden	0.20	0.15	0.14	0.12	0.09	0.08
Male, Denmark	0.36	0.28	0.23	0.16	0.11	0.09

The result indicate that the mean absolute difference is decreasing with the increase in age. This is natural for two reasons. First, when we increase the age for which we consider the remaining life expectancy, the number of prediction years decrease, and the uncertainty therefore also decrease. Second, when we increase the age, the remaining life expectancy (both actual and predicted) decreases and the magnitude of the prediction error therefore also decreases. Another interesting property is that the prediction error is on average much greater for females than for males.

## 7.2 2-75 Years Estimation Periods

In section 7.1 we presented some results concerning the prediction performance of the LC method. We had a fixed length (35 years) of the estimation period. In some cases we saw tendencies of an overestimation of the mortality, which is connected to an underestimation of the remaining life expectancy. It may be interesting to examine whether the performance could be improved by varying the length of the estimation period. For example, if we want to predict the mortality from year 1950 (and ahead), we have this far used

the mortality during 1915-1949 <sup>27</sup> One can argue that it may be unreasonable to attach too great importance of the mortality during the beginning of this period. Maybe the result would have been better if we instead have used information only from the later part of the period. In order to examine whether it is possible to improve the results we will here perform the same procedure as before, but the length of the estimation period will not be fixed to 35 years. We will let the length of the estimation period vary between 2 and 75 years. In this section we will not consider the individual mortality rates, but only the remaining life expectancy.

### 7.2.1 Cohort Expected Lifetime

By using the three measures (24)-(26) in section 6.4.1 we will try to find the optimal length of the estimation period. We will just consider the case of the remaining life expectancy for a 65 year old. Below we present the result for Swedish females, Danish females, Swedish males and Danish males.

#### Female, Sweden

In section 7.1.2 we saw that the mean absolute difference between predicted and observed remaining life expectancy was 0.94 for Swedish females aged 65. We also noticed that during the whole time period, for which we were making predictions (1935-1970), the predictions were below the observed

---

<sup>27</sup>To clarify the procedure; When we say that we are using the mortality during 1915-1949 we mean that we are using the smoothed mortality rates during this period. However, the smoothing technique used here means that the smoothed mortality for 1915 actually is calculated by using data not only from 1915, but also from the adjacent years to 1915. This means that we are using data for more than 35 year backwards in time. Regarding the smoothed mortality rate for 1949 we have been careful to not include information for later years than 1949. This is because we want our examination to be as realistic as possible. Since our purpose is to predict the mortality for year 1950 and ahead, it would not be far to consider information for these years when calculating the smoothed mortality for year 1949.



values. In figure 29 we have illustrated the consequences of different lengths of the estimation periods. The figure includes three different graphs, corresponding to the values of the three measures (24)-(26), for different lengths of the estimation period. The first graph illustrates the mean absolute difference. The graph indicates a clear minimum for a length of about 10-15 years of the estimation period. For this length the mean absolute error is reduced to 0.72 years (minimum is 0.718 for a length of 12 years). The second graph, illustrating the mean square error, is similar to the first one, indicating a minimum for the same length of the estimation period (minimum is 0.598 for a 12 years length of the estimation period.) The third graph, illustrating the minimal maximal error between predicted and observed remaining life expectancy is not as clear as the previous ones. The graph shows a local minimum of 1.25 years for a 13 years period length. For greater lengths of the estimation periods the curve seems some irregular, and it shows a global minimum of 1.20 years error for a 26 years period length.

Taking all three graphs into consideration, we think that the optimal (historically) length of the estimation period is somewhere between 10 and 15 years.

### **Female, Denmark**

Figure 30 is presenting the same graphs as in figure 29, but for females aged 65 in Denmark. The first graph, representing the mean absolute difference, shows a minimum of 0.692 years for a period length of 9 years. This should be compared to the mean absolute difference of 0.96 years for the 35 years period length. The first graph also shows a local minimum 0.705 for a 15 years period length. The second graph, representing the mean square error, is similar to the first one. It indicates a minimum of about 0.6 and an optimal length of 9-16 years (minimum is 0.596 for a 13 years length). The third graph, representing the minimal maximal error, shows a minimum of 1.23 years for a 15 years period.

The graphs for Danish females were more or less similar to the graphs for Swedish females. It is therefore natural that also the conclusions are similar. We think that the historically optimal length of the estimation period is somewhere between 9 and 16 years.

### **Male, Sweden**

When moving on to Swedish males, we find the corresponding graphs in figure 31. Compared to figures 29 and 30 this figure looks completely different. Instead of finding a minimum for a 10-15 years estimation period, it now seems like it is more appropriate to have a much longer estimation period. An interesting observation is that the magnitude of the prediction error is much smaller than for Swedish and Danish females. No matter what length of the estimation period we choose, we still have a lower mean absolute error and mean square error than we optimally had for Swedish and Danish females. For both the mean absolute error and the mean square error, we find that the historically optimal length of the estimation period has been 72 years. This length gives a minimum of the mean absolute difference of 0.157 years. The minimum of the mean square error is 0.039. For the minimal maximal error, the optimal length is 73 years, which gives a minimum of 0.433 years.

### **Male, Denmark**

Finally, the corresponding graphs for Danish males are presented in figure 32. A quick glance at the graphs tells us that they are quite similar to the graphs for Swedish males, indicating a much longer optimal length of the estimation period compared to Swedish and Danish females. The graphs are indicating two different minimums. The first minimum appears for an estimation period of about 50 years, while the second minimum indicates an optimal length of about 70-75 years. When considering the mean absolute difference, we find a minimum of 0.326 for a length of 74 years. For the mean square error the minimum is 0.152 and it appears for a 49 years period length. The third measure, the min-max error, shows a minimum of 0.914

years, for a length of 50 years. Just like for Swedish males, the magnitude of the prediction errors is substantially smaller than for Swedish and Danish females. No matter what length of the estimation period we choose, we still would have a smaller mean absolute error and mean square error than for both Swedish and Danish females.

### Summary

In the following table the result is summarized.

	Mean Absolute Error	Mean Square Error	Max Error
Female, Sweden	0.72 (12 Years)	0.60 (12 Years)	1.20 (26 Years)
Female, Denmark	0.69 (9 Years)	0.60 (13 Years)	1.23 (15 Years)
Male, Sweden	0.16 (72 Years)	0.04 (72 Years)	0.43 (73 Years)
Male, Denmark	0.33 (74 Years)	0.15 (49 years)	0.91 (50 Years)

The result is similar for Swedish and Danish females and for Swedish and Danish males. We also observed that the magnitude of the error is greater for females than for males.

## 8 Analysis

In the previous section we observed that the historically optimal length of the estimation period was considerably shorter, and the prediction error was considerably greater, for females than for males. We also observed an underprediction problem of the remaining life expectancy for females. In this section we will try to find out and explain the reasons for this result. We will consider three different concerns:

- 1. The length of the estimation period.
- 2. The size of the prediction error.
- 3. The underprediction problem

First we will consider the consequences of different lengths of the estimation period.

### 8.1 Length of the Estimation Period

The shorter length of the estimation period we choose, the more weight (for our estimations) will be added to more recently observed data. In cases when we are facing new demographical trends, these trends will be easier captured when we are using a shorter estimation period. This is an advantage by using a shorter length. At the same time, a short estimation period will make the predictions more sensitive to randomness and temporary trends in the observed mortality during the estimation period. A few years with unusually high (or low) mortality during the estimation period may result in a misspecified model and large prediction errors. Even if some of the mortality randomness should have been erased by the smoothing technique, we still may have temporary mortality trends complicating the estimation (and therefore also the prediction). In figure 33 we are comparing the predictions (of the remaining life expectancy for a 65 year old) received when

using a 10 years (full line) and a 50 years (dotted line) estimation period. These predictions are compared to the observed remaining life expectancy (plotted points). The four graphs tell us that the variation of the predictions is greater when using a 10 years length of the estimation period. The 50 years estimation period gives us more stable predictions. This is in line with the discussion above. The fact that a shorter estimation period easier will capture new demographical trends may also lead to an exaggerated respect to temporary trends, which sometimes seems to be the case, especially for males, in figure 33.

After having a basic understanding of the consequences of varying the length of the estimation period, we now go on to analyze the observed result. In section 7.2.1 we found out that the historically optimal length of the estimation period for Swedish females was about 12 years, while the optimal length for Swedish males was more than 70 years. In order to try to understand these results we have to look at development of the mortality. In figure 34 the smoothed mortality rates for Swedish males and females aged 65 are illustrated. The upper line represents males, while the lower line is for females. The vertical dotted lines illustrate the period (1935-1970) for which our predictions of the remaining life expectancy are made. Considering females, we see that there is a sharp break on the curve right before the year 1940. The negative slope of the curve, indicating a decreasing mortality, has been much higher after 1940 than before. It should be clear that if we put too much weight of the mortality before 1940 (for estimating the model) we will risk to get predictions that are not capturing this decreasing mortality trend. In this case, we will risk to overpredict the mortality, and consequently, underpredict the remaining life expectancy. By choosing a short estimation period we will quicker get rid of the "misleading" mortality that was observed before 1940. This is confirmed in figure 33, where the predictions in general are better (the underprediction problem is smaller) for the 10 years estimation period than for the 50 years period.

Studying the corresponding curve for Swedish males (figure 34), we observe that this curve is somewhat different to the curve for Swedish females. Also for males we notice a break, with a rapid decrease in the mortality, right before 1940. But the curve is then quickly leveling out and we observe a very small decrease in the mortality right to the end of the 1970's where we observe a rapid and continuing mortality decrease. Compared to Swedish females, the decrease in mortality for Swedish males has been rather small during 1900-1980. In broad outline, one can say that, except for some temporary mortality trends, the decrease in the mortality during 1900-1980 has been fairly constant (at least compared to Swedish females). Therefore, it should be no surprise that the best historically predictions have been derived from a very long estimation period, when the temporary trends and the randomness in the mortality have a small effect of the future predictions. There has been no need for a short estimation period, that quickly captures new demographical trends since the mortality trend has been fairly constant.

However, the fact that a long estimation period has been optimal historically, do not necessary mean that it should be optimal also for future predictions. The fact that we have a rapid and continuing decrease in the mortality during the last 25 years indicates a potential change in the demographical trend. Whether this trend will continue also in the future is almost impossible to say, but, if the method applied here should be used for predicting the future mortality, we think it is highly probable that is better to use a shorter estimation period than what has been optimal historically.

The corresponding curves for Danish females and males are drawn in figure 35. As noticed before, the unsmoothness of the mortality has been much higher in Denmark than in Sweden. For females we observe a great and continuing mortality improvement starting around 1940. For males, we observe a mortality improvement during the last 20 years. These mortality trends are in some way similar to the trends observed for Swedish individuals. The discussion above about concerning Swedish individuals can be applied also

for Danish individuals. It is then possible to get an understanding of the result also for Denmark, even if the conclusions are not that obvious as for Sweden.

Before moving on the next section, we will make a clarification. One should be aware of that the study of figures 34 and 35 may not be enough for a complete analysis. In these figures we are only illustrating the mortality for a 65 year old. When studying the remaining life expectancy, one has to consider the mortality at all ages during 65-99. Nevertheless, we think that these figures are still relevant for the understanding, due to two reasons. First, main mortality trends are often similar at different ages, especially when a smoothing technique is applied. Second, the mortality at the age of 65 is the mortality that contributes most when calculating the remaining life expectancy. This is due to the way the life expectancy is calculated (see (22) and (23)). The contribution from a certain age depends on not just the mortality at that age, but also on the mortality at all previous ages. Therefore, the mortality at age 65 affects the contribution in expected lifetime for all other forthcoming ages (66-99).

## 8.2 Size of the Prediction Error

We noticed in section 7.2.1 that the prediction error for females were much higher than corresponding errors for males, and this was true for both Sweden and Denmark. We will present two different explanations for this feature. First, we have measured the prediction errors in absolute (and not in relative) terms. It is well known that the remaining life expectancy is greater for females than for males. It is therefore natural that the absolute prediction error in general is greater when the size of the quantity which we want to predict is greater. The second explanation, which we also think is the most relevant, is connected with the overall trend for the life expectancy. The increase in the remaining life expectancy has been higher for females than for males. During the years for which we are making predictions (1935-1970),

the development of the remaining life expectancy for a 65 year old is presented in the table below.

	1935	1970	Difference
Female, Sweden	13.81	17.96	+ 4.15
Female, Denmark	13.38	17.46	+ 4.08
Male, Sweden	13.15	14.08	+ 0.93
Male, Denmark	12.87	13.41	+ 0.54

The increase is more than four years for females, and less than one year for males. Compared to the females, the remaining life expectancy for males has been almost constant. It seems natural that it should be easier to predict something that is almost constant than something that is not. Therefore, it is no surprise that the prediction errors are much greater for females than for males.

### 8.3 Underprediction Problem

Finally, we are considering the underprediction problem of the remaining life expectancy, that mainly is observed for females. One should be aware of the fact that the predictions are performed 35 years ahead. If the mortality during this 35 years period is decreasing more rapidly than it has during the estimation period, we will probably have an underprediction problem. As observed, in figure 34 the mortality for Swedish females has been decreasing more rapidly after 1940 than before. When using data prior to 1940 we may underpredict the future mortality, according to the discussion in section 8.1. However, by reducing the length of the estimation period, one may argue that this underprediction problem should disappear. For example, when using a 10 years length of the estimation period, the predictions for year 1950 and thereafter will not be affected by the mortality prior to 1940. However, studying figure 33, we see that the underprediction problem follow us all



the way until 1970. We believe that some of this underprediction problem depends on a boundary bias problem. The problem comes up in the kernel smoothing technique at points close to the right boundary in the time dimension, as discussed in section 4.1.2. Even if we are using the bias correction technique described in section 4.4, we still may have a bias problem near this boundary. The bias problem may result in that the smoothed mortality at this boundary is incorrect. When the mortality trend is decreasing (which is often the case) we will tend to overestimate the smoothed mortality near the boundary during the estimation period. The smoothed mortality during the estimation period underlies the LC estimations, which includes estimates of the mortality index  $k_t$ . An overestimated smoothed mortality at the endpoints will therefore automatically be incorporated into the LC estimations. The way the LC predictions are generated means that the predictions of the future mortality depends solely on just the first and the last value of the estimated mortality index, since the predicted future mortality index is generated by extrapolation of the line that is drawn through these two points (see figure 19). If the last estimated point of the mortality index is too high, the negative slope of the mortality index may be too low, which result in future mortality predictions that are too high. This is illustrated in figure 36. The plotted points represent the LC estimated mortality index during a 10 years estimation period. The vertical line separates the estimation period (1960-1969) and the prediction period (1970-2004). The full line represents the LC predictions of the mortality index. As mentioned before, the predictions are produced by extrapolation of the line that connects the first and the last points of the mortality index. Now assume that the last points of the estimated mortality index are too high, due to the boundary bias problem. If we ignore the last three points and instead produce the predictions by connecting the first and the fourth last points of the estimated mortality index, we would then have predictions according to the dotted line. Comparing the two different "prediction curves", we see that there is a fairly big difference between them, especially for the latest prediction years, since the

gap between the curves increases as time goes by. The potential boundary bias problem was observed already in figures 21-24, where we observed a discrepancy between observed and estimated mortality just at the last years of the estimation period.

It is difficult to determine in what degree the result has been affected by the possible boundary bias problem. The starting point when analyzing the performance of the predictions is to compare the predicted to the observed remaining life expectancy. This is fairly straightforward, but if we want to analyze the result in more details, we have to consider the individual mortality rates, where we compare our predictions to the observed. The fact that we are considering the cohort expected lifetime instead of the period life expectancy makes the graphical analyzing even more complex. To determine the remaining life expectancy for a 65 year old we have to consider the mortality rates for individuals aged 65-99. For these 35 ages, the mortality rates should be considered during 35 different calendar years, which makes it difficult to graphically analyze the influence of the potential boundary bias problem. It should be clear however that the boundary bias problem will be smaller when we increase the length of the estimation period. When the length is increased, the distance (in years) between the first and the last point of the mortality index will become greater. The possible bias of the boundary point will then have a smaller impact on the slope of the prediction curve. At the same time, the potential problem becomes more essential when we decrease the length of the estimation period, which may have been the case for Swedish and Danish females.

The potential boundary bias problem may probably be reduced by using methods for boundary bias corrections. It is also possible that a local linear-instead of a local constant kernel smoothing technique would have been more appropriate. An interesting and alternative approach for reducing the problem is to change the way the predictions of the mortality index is produced. As explained before, the predictions depends only on the first and the last ob-

ervation of the mortality index. If the predictions instead would have been produced by a linear regression, then the impact of the last (and the first) observation would have been reduced. The predictions would then instead be affected by all earlier observations of the mortality index.

## 9 Future Forecasts

At this point we are able to present predictions of the future remaining life expectancy for individuals in Sweden and Denmark. Before making the predictions we have to decide the number of years to be used for the LC estimation. The issue of finding the optimal length of the estimation period was investigated in section 7.2.1. For females (in both Sweden and Denmark) we found a historically optimal length of about 10-15 years. For Swedish and Danish females, we have decided to use a 12 years estimation period when producing our future predictions.

For males, the result, in terms of finding the optimal length of the estimation period, was not as obvious as for females. We found that a very long (50-75 years) estimation period historically has been most successful. In section 8.1, when analyzing the result, we studied the development of the mortality and had a discussion about the optimal length for producing future predictions. We had some doubts whether it maybe would be more appropriate to have a shorter estimation period to capture the more recent mortality trend for males (see figure 34 and 35). When producing the future predictions for males we will therefore present to different predictions, using two different lengths of the estimation period. First, we will use a "long" estimation period, according to what has been optimal historically. We will use a length of 72 years for Swedish males and a length of 50 years for Danish males. In the second approach we will use a "short" estimation period (15 years) for both Swedish and Danish males.

In figure 37 we are illustrating the predicted remaining cohort life expectancy during 1935-2005 for individuals aged 65. For the prediction for a certain year, we have used the data available until that year. The predictions of the remaining life expectancy are calculated as before, using the future 35 years predicted mortality rates. Accordingly, the prediction for 2005 are calculated from the predicted mortality rates during 2005-2039, where we have used the

data until 2004 for estimating the model. As far as it is possible (until 1970), we have included the observed remaining expected lifetime in the graphs, which is illustrated by the full line. For Swedish and Danish males two different predictions are presented. The dotted line illustrates the predictions generated from the "long" estimation period (72 years for Sweden and 50 years for Denmark). The dashed line corresponds to the predictions generated by the "short" (15 years) estimation period. For females, we are only presenting one prediction (12 years estimation period), which is illustrated by the dotted line. For females, the predicted remaining life expectancy in 2005 is 21.09 (for Sweden) and 19.25 (for Denmark). The corresponding predictions for males depends on which of the approaches we are using. If we are using the "long" estimation period, the predicted remaining life expectancy is 17.49 (Sweden) and 15.58 (Denmark). If we instead are applying the "short" estimation period the corresponding values are 18.27 (Sweden) and 16.34 (Denmark). The result is summarized in the following table, where we also have included the observed remaining life expectancy in 1970.

	Observed 1970	Predicted 2005	Difference
Female, Sweden	17.96	21.09	3.13
Female, Denmark	17.46	19.25	1.79
Male, Sweden	14.08	17.49 (18.27)	3.41 (4.19)
Male, Denmark	13.41	15.58 (16.34)	2.17 (2.93)

By studying the result we find some interesting observations:

- We observe that the remaining life expectancy in 2005 for males depends strongly on whether we are using the "long" or the "short" estimation period. For both Sweden and Denmark, the difference between the two approaches is about 0.8 years. It is no surprise, according to earlier discussions, that we get a higher value of the prediction when we are using a shorter estimation period. When applying a shorter period, the estimated model better captures the more recent demographical trend of a decreasing mortality. The fact that the both approaches

differ that much confirms the importance of choosing an "appropriate" length of the estimation period. Compared to other prediction methods, the LC method is known to involve less subjective judgments (Lee and Miller (2002)). However, we have here highlighted that one still has to take care of the issue of choosing the estimation period. The importance of choosing the length of the estimation period is a concern that has not been discussed very much earlier in the actuarial literature.

- Another interesting observation is that the increase in remaining life expectancy (the difference between the 2005 prediction and the 1970 observation) is greater for males than for females. This should be compared to the development during 1935-1970 (see table on page 62) where the increase was much greater for females.
- According to the predictions, the difference in remaining life expectancy between Sweden and Denmark tend to increase, compared to the situation in 1970. In 1970 the difference was 0.43 years for females and 0.28 years for males. The predicted difference in 2005 is 1.84 years for females and 1.91 years ("long" estimation period) or 1.93 years ("short" estimation period) for males.

## 9.1 Comparison to other Predictions

Finally we will compare our predictions to some other measures of the remaining life expectancy for a 65 year old. The comparison is illustrated in the bar charts in figure 38. The different names of the bars should be interpreted as follows:

- **P**: The predicted remaining life expectancy (for females) using a 12 years estimation period.
- **P1**: The predicted remaining life expectancy (for males) using a "long" estimation period (72 years for Sweden and 50 years for Denmark.)

- **P2**: The predicted remaining life expectancy (for males) using a "short" estimation period (15 years for both Sweden and Denmark.)
- **ITP**: The mortality according to the Swedish ITP<sup>28</sup>-plan.
- **1964**: The mortality according to the "1964 års grunder" (M64)
- **SCB**: The predicted remaining life expectancy in 2005 according to the Swedish "Statistiska centralbyrån".
- **T**: The mortality used for traffic insurance by a Swedish insurance company.
- **G82**: The mortality according to the Danish G82, which is the standard mortality table in Denmark.

The mortality from SCB and the traffic insurance (T) are, just like our predictions (P, P1 and P2) based on the whole population, while the rest of the mortalities are based on insured individuals. The result is summarized in the following table:

	P	P1	P2	ITP	1964	SCB	T	G82
Female, Sweden	21.09	-	-	22.41	18.77	20.40	18.89	-
Female, Denmark	19.25	-	-	-	-	-	-	17.85
Male, Sweden	-	17.49	18.27	18.22	15.84	17.25	15.78	-
Male, Denmark	-	15.58	16.34	-	-	-	-	15.11

For Swedish females, the predicted remaining life expectancy (21.09) is far below (1.32 years) the life expectancy according to the ITP-plan (22.41), but far above the other measures. For Swedish males, the conclusion depends on which predictions we are considering. When using a "long" (72 years) estimation period the prediction (17.49) is 0.73 years below ITP (18.22). If we instead are using the "short" (15 years) estimation period, the prediction (18.27) actually is 0.05 years greater than ITP. One should remember

---

<sup>28</sup>ITP=Industrins och handelns tilläggspension.

that our predictions are based on the whole population whereas the ITP-predictions are adapted to insurance.

For Denmark we only compare to the G82 standard mortality table. For both males and females, our predictions are greater than G82. For females, the difference is 1.40 years. For males, the difference depends on which predictions we are using. For the "long" estimation period the difference is 0.47 years, while the difference is 1.23 years for the "short" estimation period. The difference to G82 is large, but one should be aware of that G82 was created in 1982, and it is well known that it is overestimating the mortality (i.e. it underestimates the expected lifetime). However, G82 is adapted to insured individuals, which makes the comparison harder to interpret.



## 10 Summary and Concluding Remarks

### 10.1 Summary

In this thesis we are dealing with the problem of estimating and predicting the mortality for old aged (65-99) individuals in Sweden and Denmark during 1900-2004. When an insurance company settle the reserves for their life annuity contracts they has to rely on forecasts of the future mortality. It is well known that the mortality depends on both the gender and the age of the individual. It is also clear that the mortality has not been constant over time. In our modeling we are studying each gender and each country (Sweden and Denmark) separately, and we are considering both the age of the individual and the chronological time. For estimating the mortality rate we are using the ratio of the number of deaths and the exposure at a certain age and time. This is a non parametric counting process technique and requires no assumptions about the distribution of the mortality. However, we assume the underlying mortality to be smooth. A problem with this technique is the potential lack of information (exposure) at old ages. As long as individuals passes away, the exposure will become less. When the exposure is less, the estimations tend to become unstable, which is not consistent with the smoothness assumption for the counting process technique. To get rid of this problem, we are applying a kernel smoothing technique. This way we will have a greater exposure and, therefore, more stable mortality estimations. Unfortunately, we will also have a bias problem due to the rapidly decreasing exposure and the rapidly increasing mortality at old ages. To handle the bias problem we are using a bias correction technique. However, we still may have a bias problem, at the right boundary in the time dimension, for the estimation period.

The next step is to adapt a parametrical model to our non-parametric mortality estimations. We are applying the Lee-Carter method to produce estimates of the mortality. From the LC estimates we are then generating predictions

of the future mortality. By only using a part of the data for the estimation, we are able to compare our LC predictions to the actual and observed mortality. We will then be able to evaluate the performance of the LC prediction method. The evaluation is mainly performed by comparing the predicted and the observed remaining life expectancy for a 65 year old. We notice a clear underprediction tendency for females in both Sweden and Denmark. When using a 35 years estimation period, the mean absolute error between predicted and observed remaining life expectancy for a 65 year old female was almost a year in both Sweden (0.94) and Denmark (0.96). For males, the predictions were in general better than for females, and there were no clear underprediction tendency. The mean absolute difference for males were 0.20 (for Sweden) and (0.36) years (for Denmark).

We were then trying to improve the results by varying the length of the estimation period. By reducing the length of the estimation period, new demographical trends will easier be captured by the predictions. At the same time the predictions will be more sensitive to temporary and random mortality trends. A longer estimation period will produce more stable predictions, but requires longer time to incorporate new demographical trends into the predictions. For females, we found that the historically optimal length of the estimation period was 10-15 years. The mean absolute error between predicted and observed remaining life expectancy could then be reduced to about 0.7 years. For males, the historically optimal length has been 50-75 years. The mean absolute error could then be reduced to 0.16 years for Sweden and 0.33 years for Denmark. The reason for the difference (in optimal length of the estimation period) between males and females is analyzed by studying the historically development of the remaining life expectancy for a 65 year old. During the evaluation period (1935-1970), the increase in remaining life expectancy for females has been much greater than for males. The increase for females between 1935 and 1970 is more than 4 years, while the increase for males is less than one year. For later years, we have seen tendencies of a more rapidly decreasing mortality also for males, and we believe

that is appropriate to shorten the estimation period for future predictions, in order to capture this potential decreasing demographical trend.

Finally, we have produced forecasts of the future remaining life expectancy. An interesting observation is the fact that the future predictions may depend substantially of the length of the estimation period in the LC model. The LC method is known to have less subjective judgments than other methods (Lee and Miller (2002)). However, we have shown that it is important to consider the length of the estimation period before producing mortality predictions. The mortality in Sweden has general been lower than in Denmark over time. During the last 50 years, the difference between the both countries has increased. Our predictions of the future mortality is not indicating that the difference will decrease. The predicted remaining life expectancy for a 65 year old in 2005 is 21.1 (19.2) for Swedish (Danish) females and 18.3 (16.3) for Swedish (Danish) males. Comparing females and males, the improvements in remaining life expectancy has been greater for females over time. However, our predictions are indicating a more rapidly future improvement for males than for females.

Our predictions of the future remaining life expectancy have been compared to other theoretical measures of the remaining life expectancy in Sweden and Denmark. For both countries and genders, our predictions often are greater than these theoretical values.

## **10.2 Concluding Remarks**

Throughout this Master thesis we have discussed the importance of having a lot of information when making estimations and predictions about the mortality. A small exposure will make the estimations unstable, and it is obvious that this is not desirable. In this paper we have used information from the whole Swedish and Danish populations, which can be considered as

a rather large information amount. Nevertheless, we still have the problem with insufficient information, and therefore also the problem with unstable mortality estimations.

Considering the valuation problem from the point of view of an insurance company, the task is to perform accurate mortality predictions concerning the individuals which the contracts' value are based on. A natural question to ask is whether the mortality of these individuals can be considered as the same as the mortality of the whole population. Statistical surveys have shown that this is not the case (see Ajne and Ohlin (1990)). The mortality of an insured individual is in general lower than the whole populations' mortality, which is due to the medical check-up that is made in connection with the signing of the contract. Further, when looking at insured individuals, the time since the contract was signed is also of importance when considering the mortality. In figure 38 and in section 9 we noticed that the LC predictions of the future remaining life expectancy (for a 65 year old) often were greater than other theoretical values of the remaining life expectancy. The LC predictions were based on population data. If we assume that the mortality of the population actually is higher than the mortality of insured individuals, the difference between our predictions and the theoretical mortality should be even greater.

The fact that there is a difference between population and insurance mortality will complicate things for the insurance company. By using population data it will run the risk of making incorrect mortality predictions, which could be an expensive issue for the company. An alternative is to create mortality estimations based on historical observations only from its own customers. This way the predictions may be unbiased, but instead they may be very unstable since the information available is much more less. Since this was a problem already for the whole population case, one can expect the problem to be even greater this time.

When looking at the prediction problem from the perspective of an insurance

company, there is another complicating issue that has not been mentioned this far. In this thesis we have considered the task of predicting the mortality after the age of 65. We have assumed an age of 65 for the individual, and we have then predict the future mortality during the following 35 years. In practice, it is often the case that the age of an individual signing a life insurance contract is less than 65. If we assume an age of 50 of the individual signing the contract, the prediction issue for the insurance company is more complicated. Instead of considering the mortality at ages 65-99 starting from today (as in this thesis), it instead has to consider the mortality at ages 65-99 starting 15 years from now. Apart from that, it also has to consider the mortality for the ages 50-65 for the following 15 years. Consequently, the insurance company has to consider the mortality 50 years ahead (instead of 35). For individuals younger than 50, the situation is even more complicated since further prediction years are then necessarily.

Taking this difficulties into consideration, combined with the fact that the mortality prediction already has appeared to be a difficult task, one realize that the mortality predictions that an insurance company face may be very difficult.

## References

- [1] Ajne, B. and Ohlin, J. *Livförsäkringsmatematik*. Kompendium, Maj 1990.
  
- [2] Booth, H., Maindonald, J. and Smith, L. *Age-time interactions in mortality projection: Applying Lee-Carter to Australia*. Working Papers in Demography. Revised 2 August 2002.
  
- [3] Deaton, A. and Paxson, C. *Mortality, Income, and Income Inequality Over Time in Britain and the United States*. Working Paper 8534, October 2001. Available at <http://papers.nber.org/papers/w8534.pdf>
  
- [4] Fledelius, P. *Validating the performance and economic consequences of mortality projections for the Danish populations*. Biometrika.(2003).
  
- [5] Fledelius, P., Guillen, M., Nielsen, J.P. and Petersen, K.S. *A Comparative Study of Parametric and Non-parametric Estimators of Old-Age Mortality in Sweden*. Journal of Actuarial Practice.(2002).
  
- [6] Fledelius, P., Guillen, M., Nielsen, J.P. and Vogelius, M. *Two-dimensional Hazard Estimation for Longevity analysis*. Scandinavian Actuarial Journal (2004).

- [7] Fledelius, P., Lando, D. and Nielsen, J.P. *Non-parametric analysis of rating transition and default data*. Journal of Investment Management. (2004).
- [8] Giroi, F. and King, G. *A Reassessment of the Lee-Carter Mortality Forecasting Method*. Working paper, 3 March 2005.  
Available at <http://gking.harvard.edu/files/lc.pdf>
- [9] Hall, P. and Park, B.U. *New methods for bias correction at endpoints and boundaries*. The Annals of Statistics, 2002, Vol. 30, No. 5, 1460-1479.
- [10] *Human Mortality Database*. University of California, Berkeley (USA), and Max Planck Institute for Demographic Research (Germany). Available at [www.mortality.org](http://www.mortality.org). (Data downloaded on 24th November 2005).
- [11] Karunamuni, R.J. and Alberts, T. *On boundary correction in kernel density estimation*. Statistical Methodology 2 (2005) 191-212.
- [12] Lee, Ronald D. and Carter, Lawrence R. *Modelling and Forecasting U.S Mortality*. Journal of the American Statistical Association; September 1992; 87, 419.

- [13] Lee, Ronald D. and Miller, T. *Evaluating the Performance of the Lee-Carter Method for Forecasting Mortality*. Demography, Volume 38-Number 4, November 2001: 537-549.
- [14] Lindbergson, M. (2001) *Mortality Among the Elderly in Sweden 1988-1997*. Scandinavian Actuarial Journal. 2001; 1: 79-94.
- [15] Renshaw, A.E. and Haberman, S. (2003a) *Lee-Carter mortality forecasting with age-specific enhancement* Insurance:Mathematics and Economics 33 (2003), 255-277.
- [16] Renshaw, A.E. and Haberman, S. (2003b) *On the forecasting of mortality reduction factors* Insurance:Mathematics and Economics 32 (2003), 379-401.
- [17] Renshaw, A.E. and Haberman, S. (2005) *Mortality reduction factors incorporating cohort effects* Actuarial research paper no. 160
- [18] Renshaw, A.E. and Haberman, S. (2006-forthcoming publication) *A cohort-based extension to the Lee-Carter model for mortality reduction factors* Insurance:Mathematics and Economics.
- [19] *Statistiska centralbyrån*. Available at [www.scb.se](http://www.scb.se). 2006).



- [20] Vandeschrick, C (2001). *The Lexis diagram, a misnomer*. Demographic Research, volume 4, Article 3, March 2001.
- [21] Velfærdskommissionen *Fremtidens velfærd kommer ikke af sig selv*. Analyserapport. Maj 2004, 249-256.

## 11 Figures

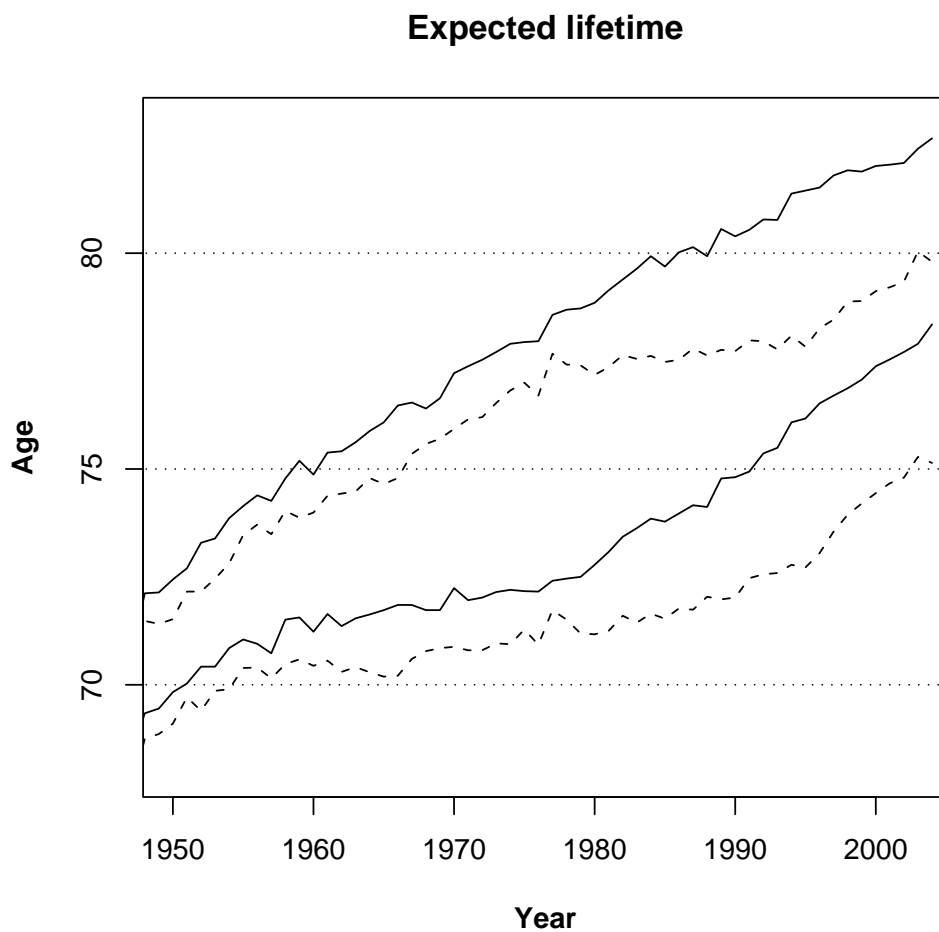


Figure 1: The figure shows the development of the life expectancy during 1950-2004. The two lines on top represent women, while the lower two lines represent men. The full lines represent Sweden and the dashed lines represent Denmark.

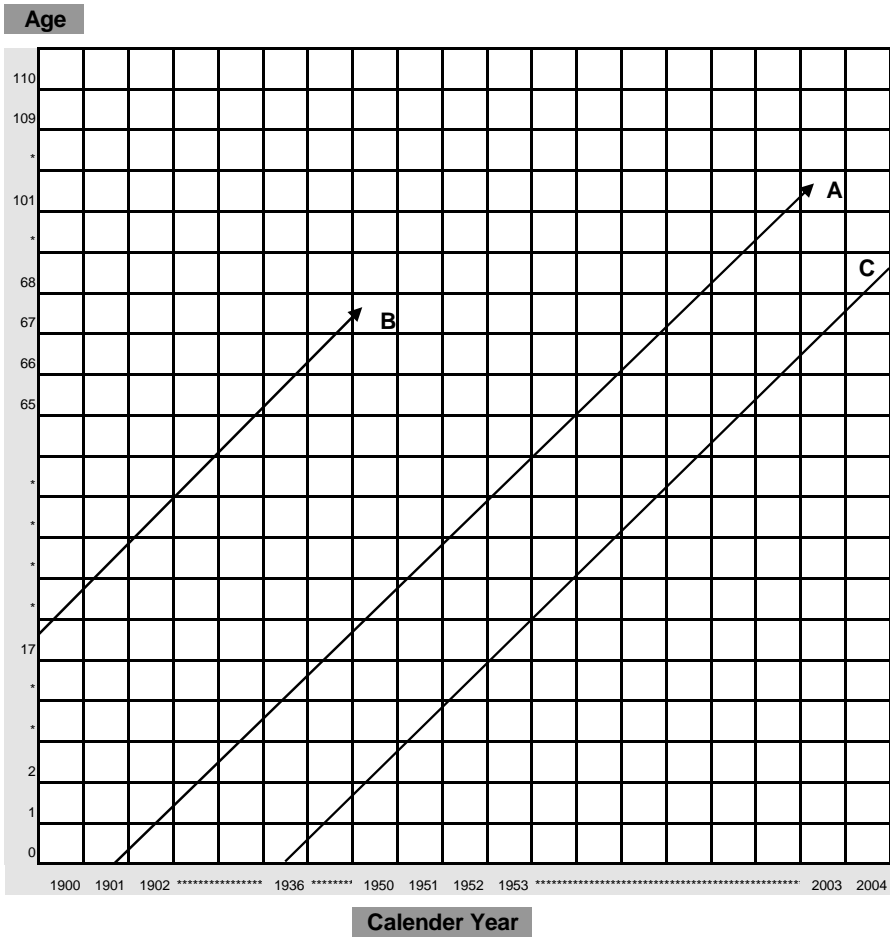


Figure 2: Lexis diagram

### Life expectancy for men 1835–2004

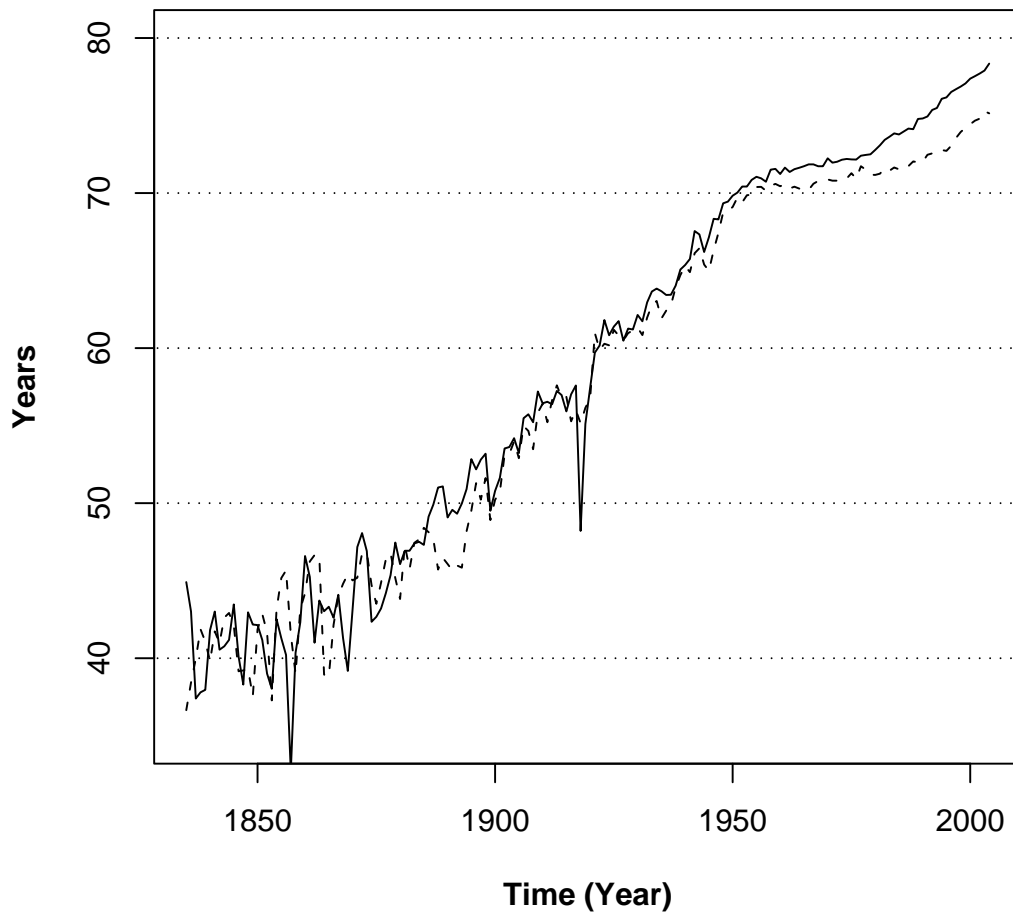


Figure 3: The figure shows the development of the life expectancy for men during 1835-2004. The full line represents Sweden and the dashed line represents Denmark.

### Life expectancy for women 1835–2004

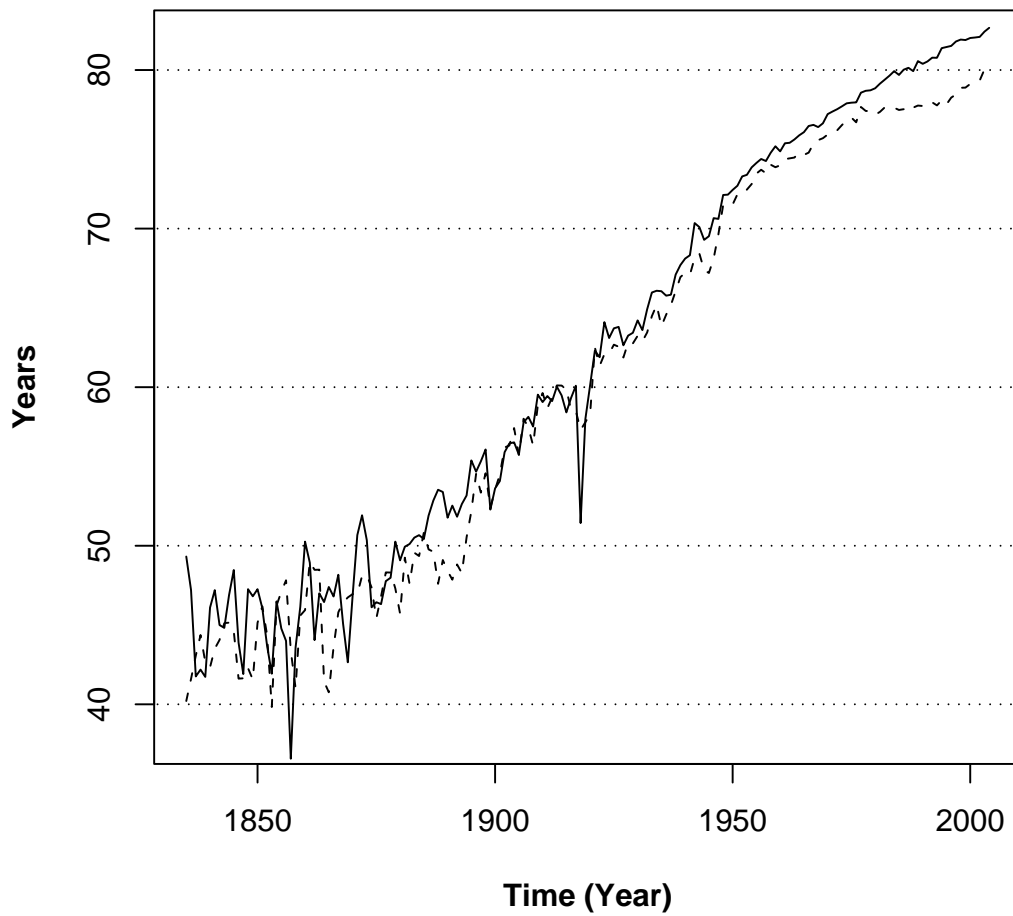


Figure 4: The figure shows the development of the life expectancy for women during 1835-2004. The full line represents Sweden and the dashed line represents Denmark..

### Deaths for Swedish women in 1900–2004

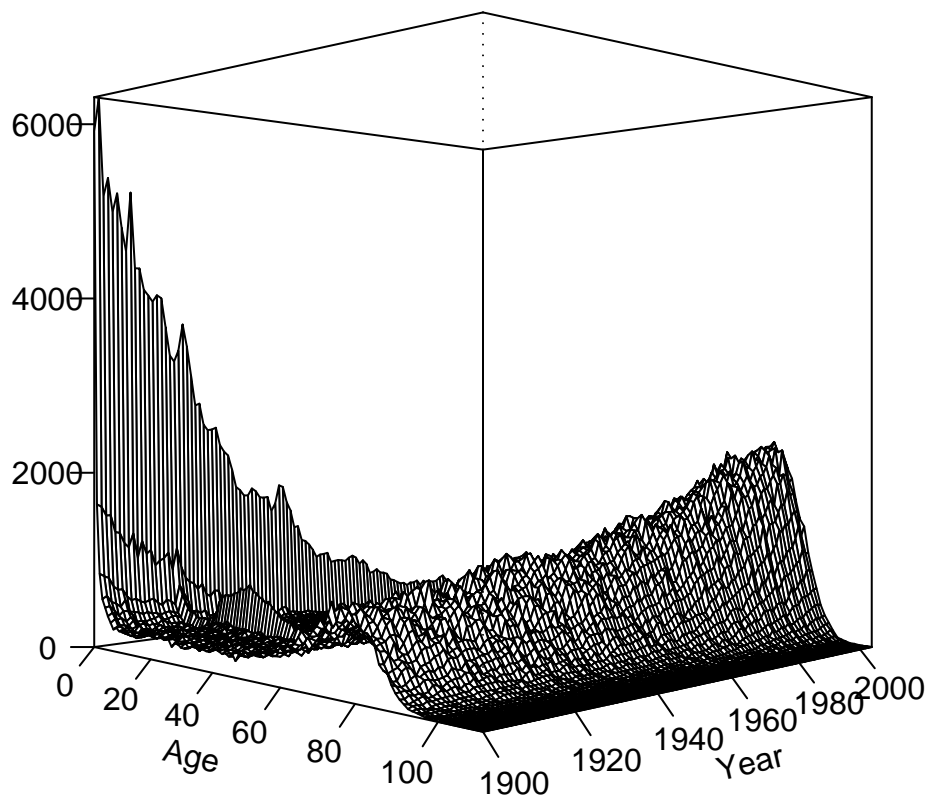


Figure 5: The figure shows the number of deaths for Swedish women for different ages (0-110) and years (1900-2004).

### Exposure for Swedish women in 1900–2004

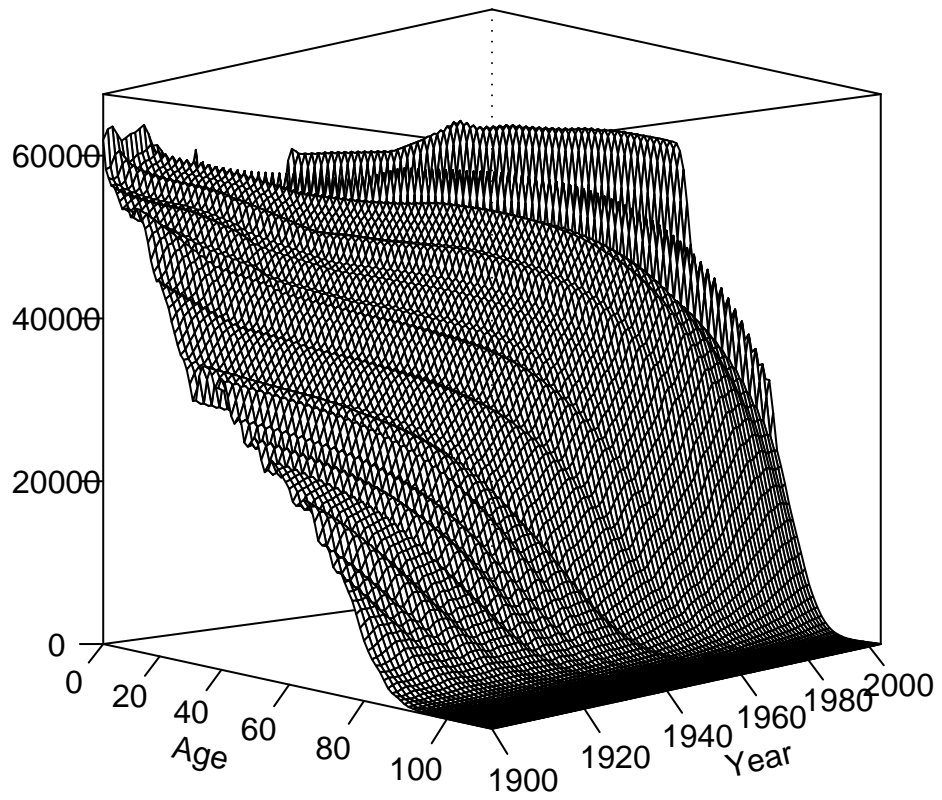


Figure 6: The figure shows the exposure for Swedish women for different ages (0-110) and years (1900-2004).

### Mortality rate for Swedish women in 1900–2004

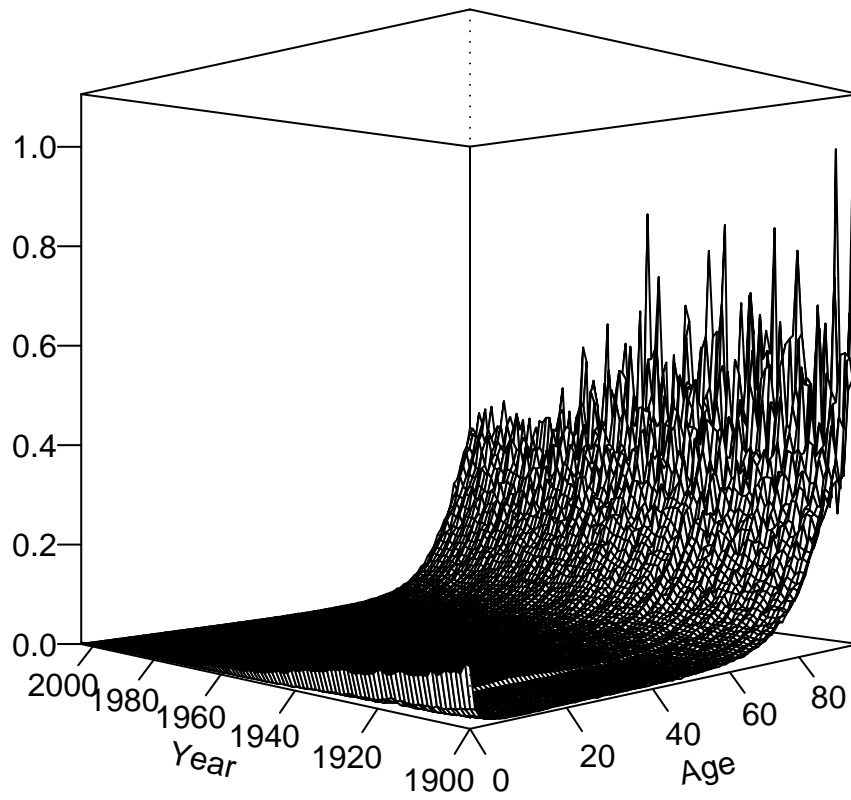


Figure 7: The figure shows the mortality rate estimations for Swedish women for different ages (0-99) and different years (1900-2004).



### Mortality rates for Swedish women 1900–2004

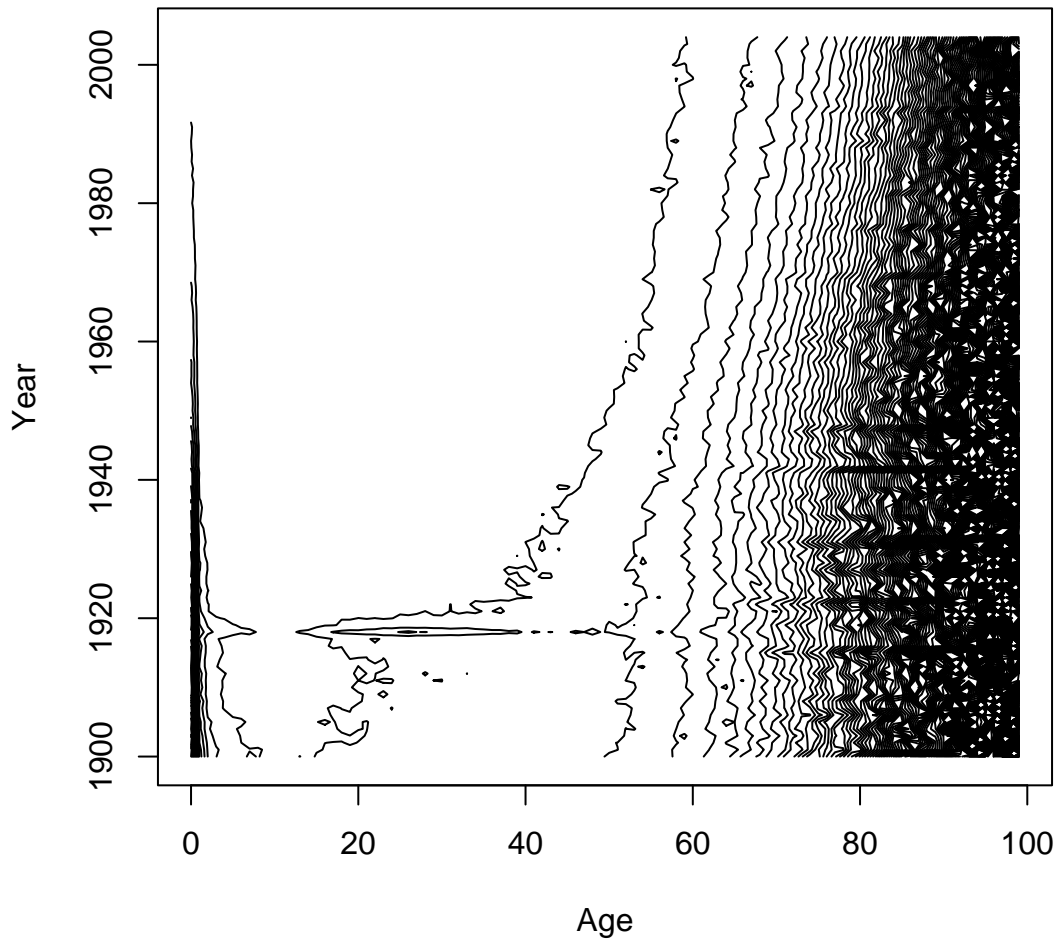


Figure 8: The contour plot shows the mortality rate estimations for Swedish women for different ages and years. Equally sized estimations are connected with level curves.

### Total exposure in Sweden and Denmark 1989–2004

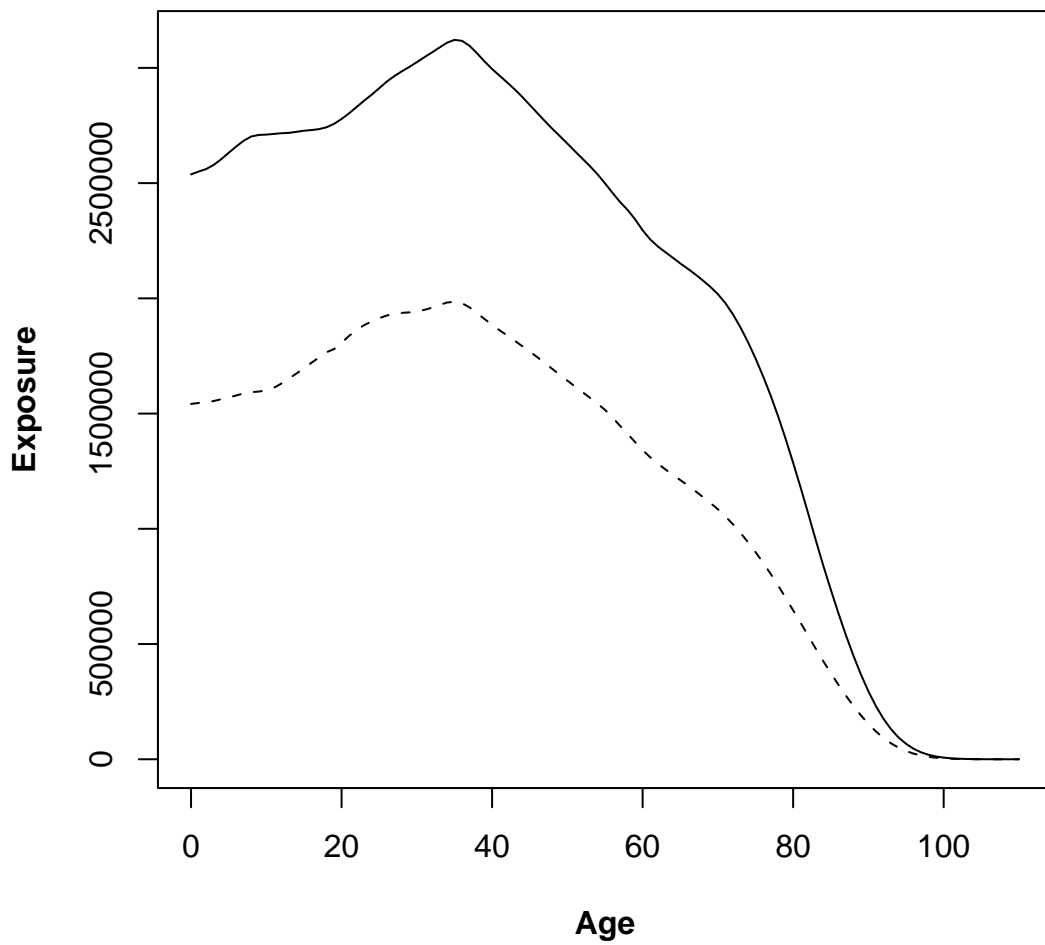


Figure 9: The figure shows 25 years accumulated exposure in Sweden and Denmark for different ages. The full line represents Sweden and the dashed line represents Denmark. The exposure is accumulated for men and women during the period 1980-2004.

Age

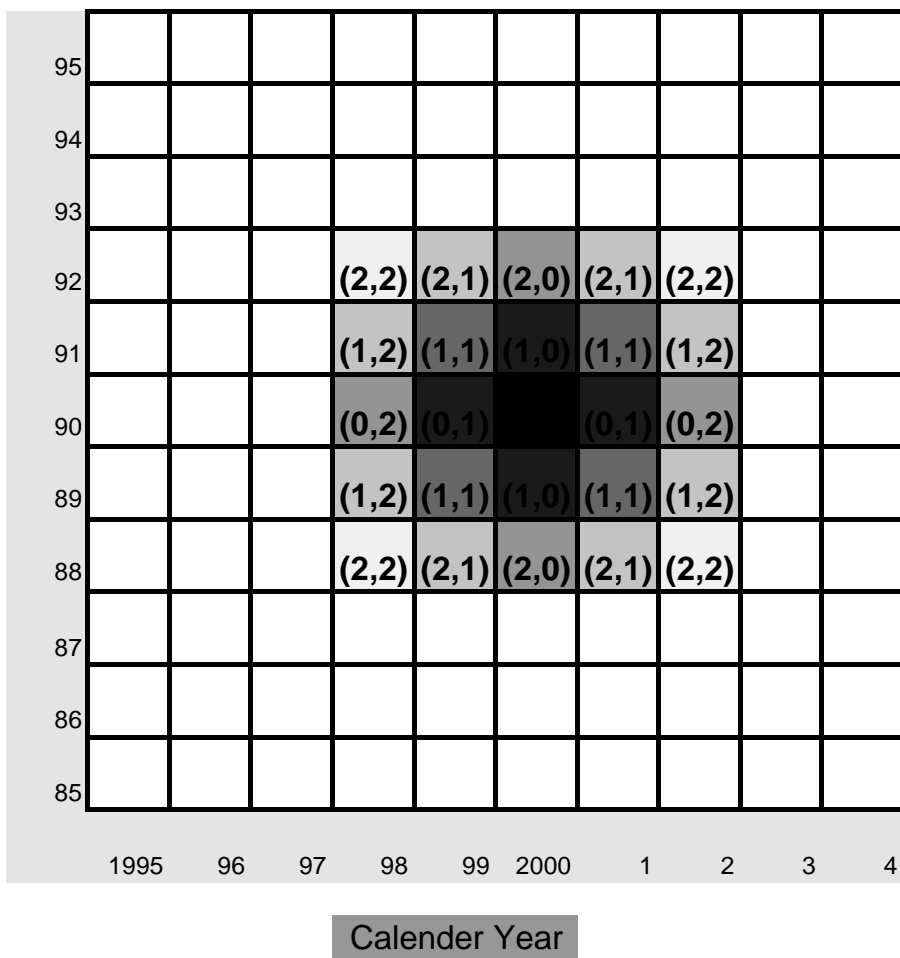


Figure 10: The counting process technique. Estimation of  $\bar{m}_{90,2000}$

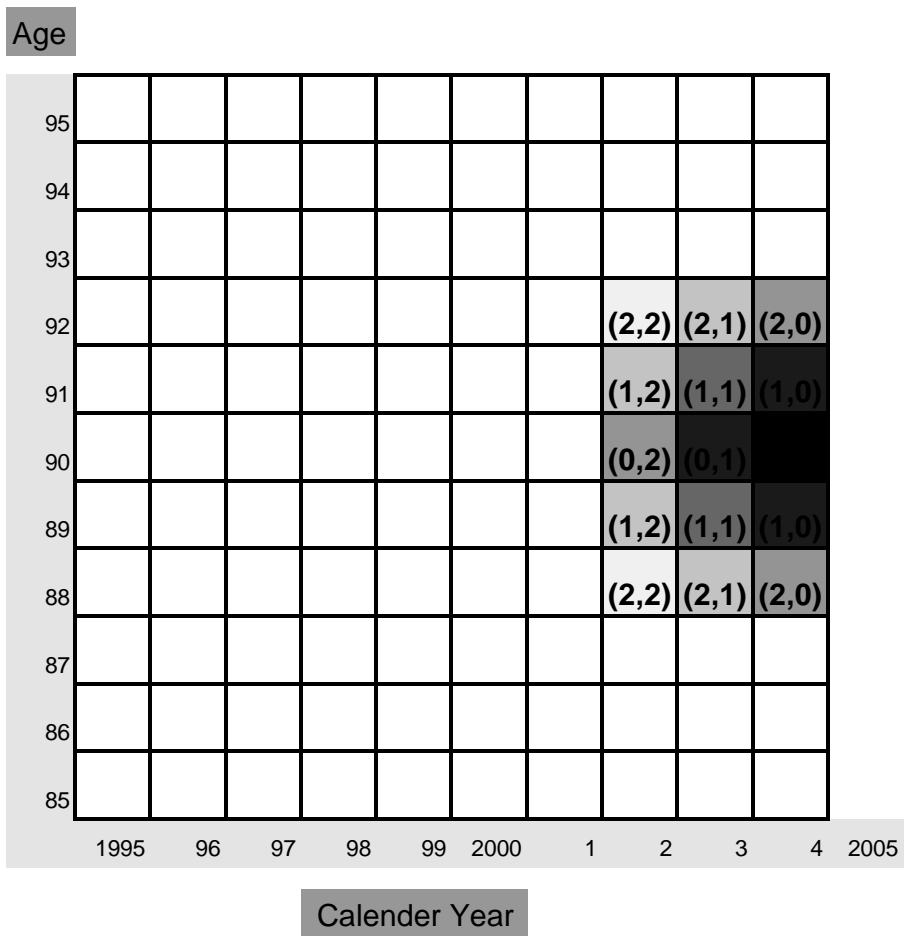


Figure 11: The counting process technique. Estimation of  $\bar{m}_{90,2004}$

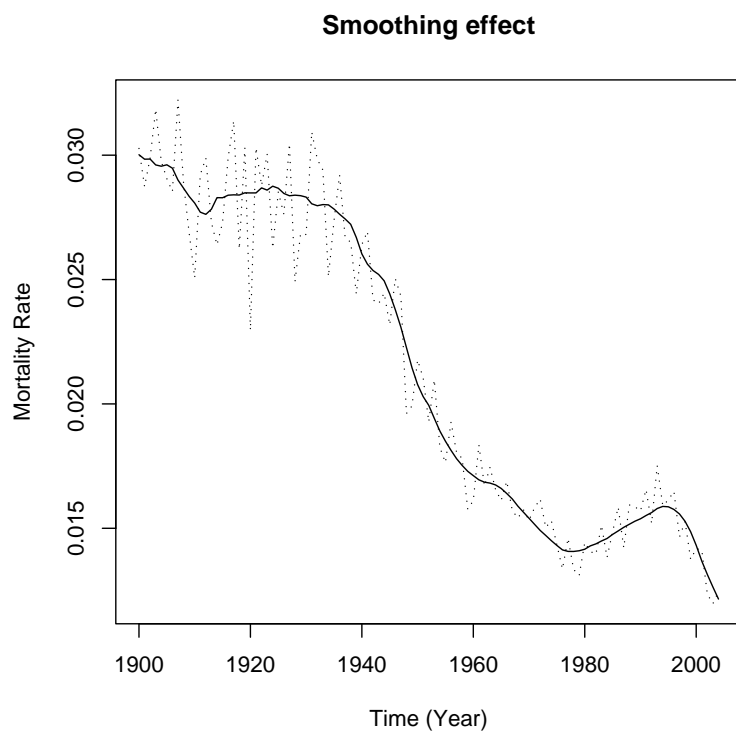
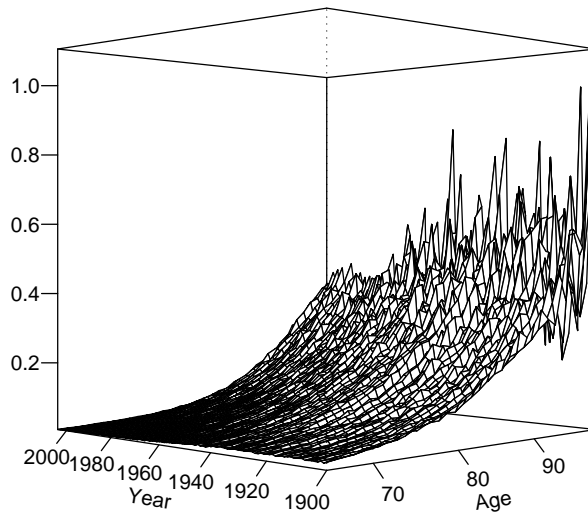


Figure 12: The effect of the smoothing technique for Danish females aged 65.

**Swedish females (non smoothed)**



**Swedish females (smoothed)**

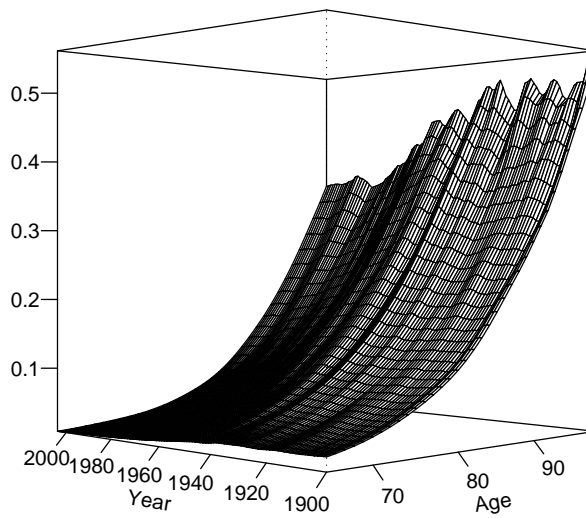
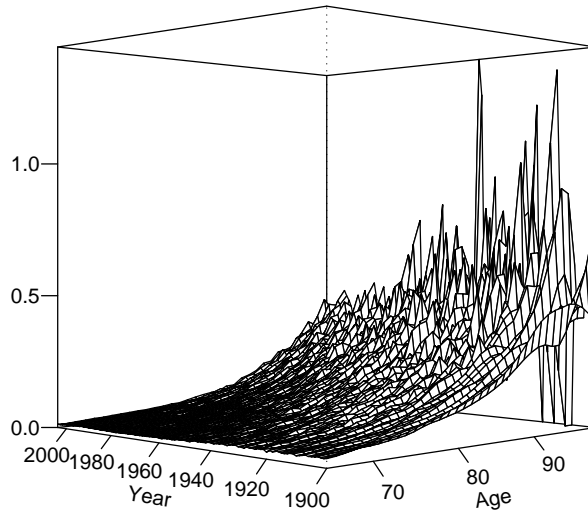


Figure 13: The figure shows the non-smoothed and the smoothed mortality rate estimations for Swedish females aged 65-99 during 1900-2004.

**Danish females (non smoothed)**



**Danish females (smoothed)**

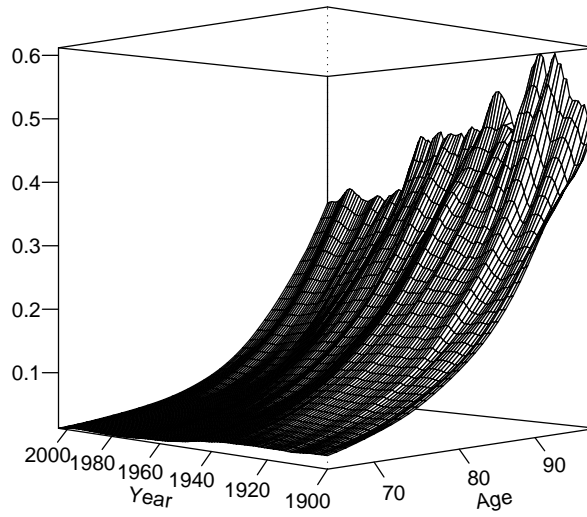
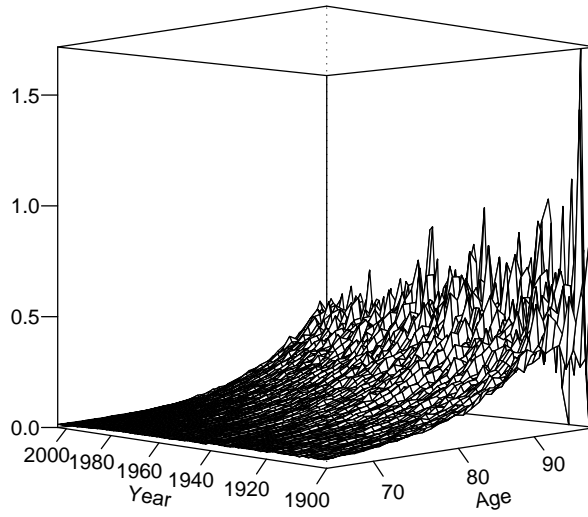


Figure 14: The figure shows the non-smoothed and the smoothed mortality rate estimations for Danish females aged 65-99 during 1900-2004.

**Swedish males (non smoothed)**



**Swedish males (smoothed)**

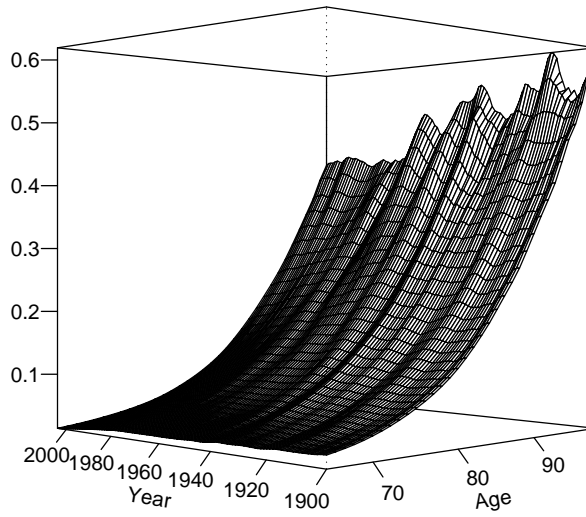


Figure 15: The figure shows the non-smoothed and the smoothed mortality rate estimations for Swedish males aged 65-99 during 1900-2004.



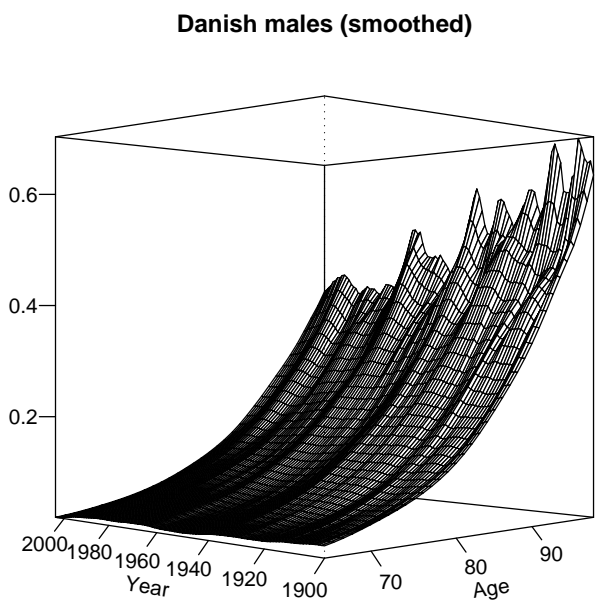
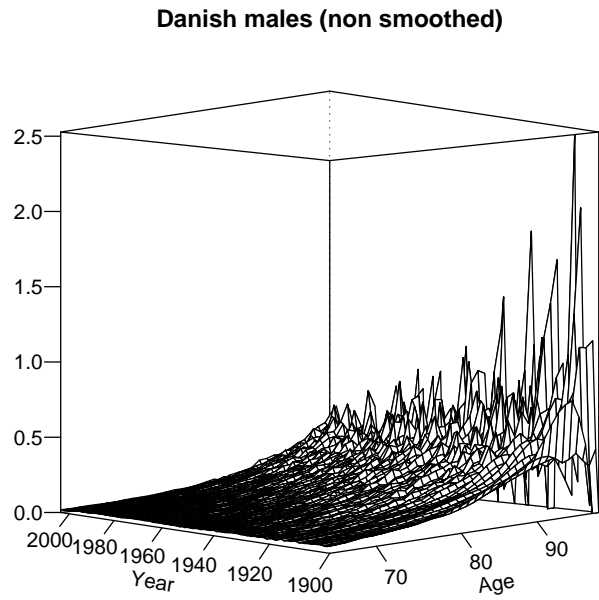


Figure 16: The figure shows the non-smoothed and the smoothed mortality rate estimations for Danish males aged 65-99 during 1900-2004.

### Danish Females 2004

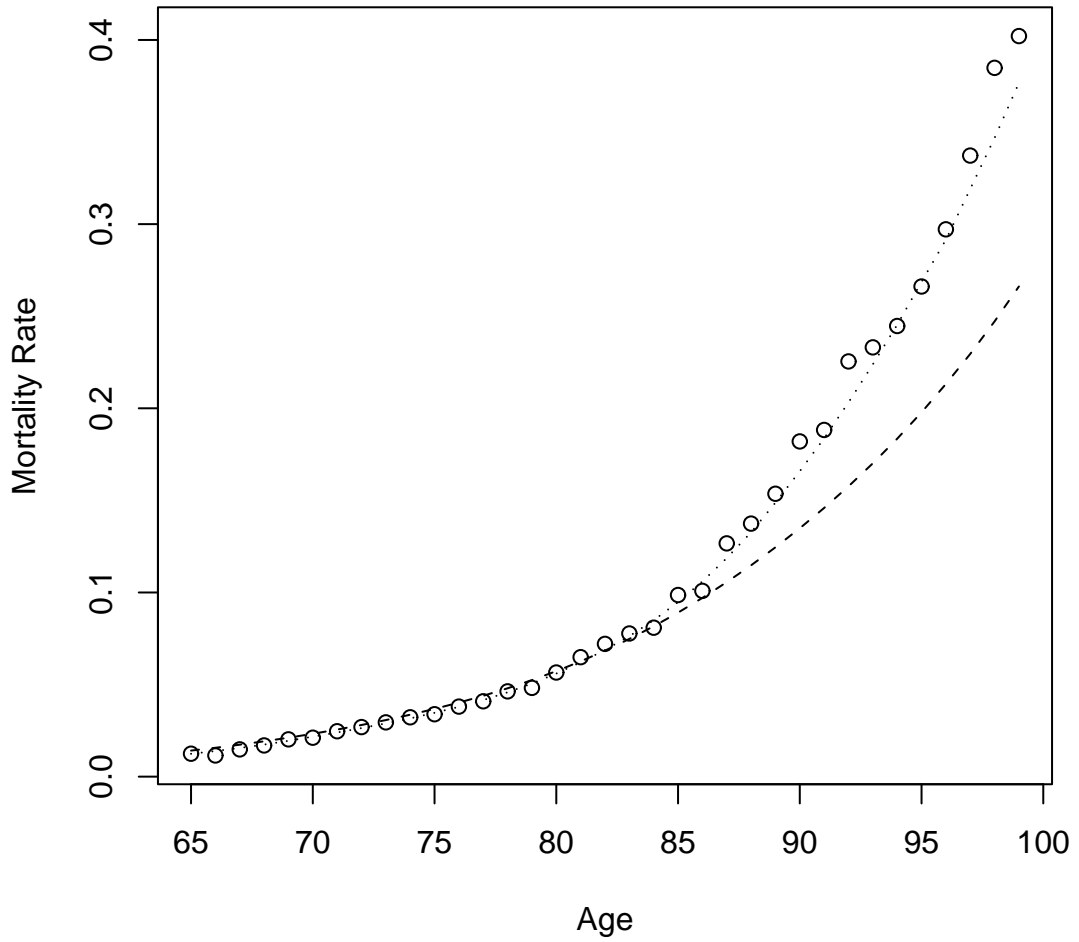


Figure 17: The figure illustrate the mortality for Danish females aged 65-99 in year 2004, and the effect of the bandwidth selection. The plotted points represent the raw mortality (non-smoothed) For the dashed lower line (smoothed) the bandwidth is  $b = (10, 10)$ . For the dotted upper line (smoothed) the bandwidth is  $b = (4, 4)$ .

### Danish Females 2004

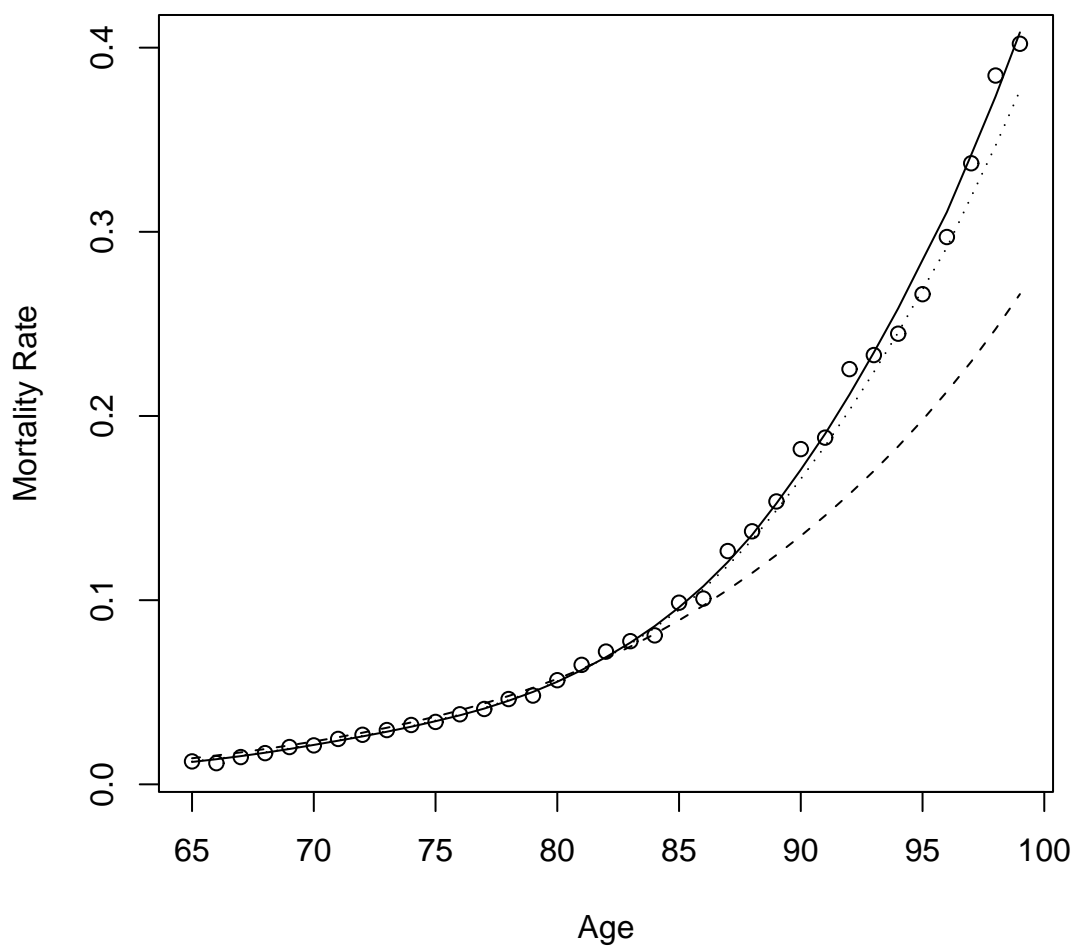


Figure 18: The figure illustrate the effect of the bias correction technique for Danish females aged 65-99 in year 2004. The points represent the raw mortality (non-smoothed) For the dashed lower line the bandwidth is  $b = (10, 10)$ . For the dotted upper line the bandwidth is  $b = (4, 4)$ . The full line represents the mortality estimates when the bias correction technique is applied.

### Estimated and Predicted Mortality Index

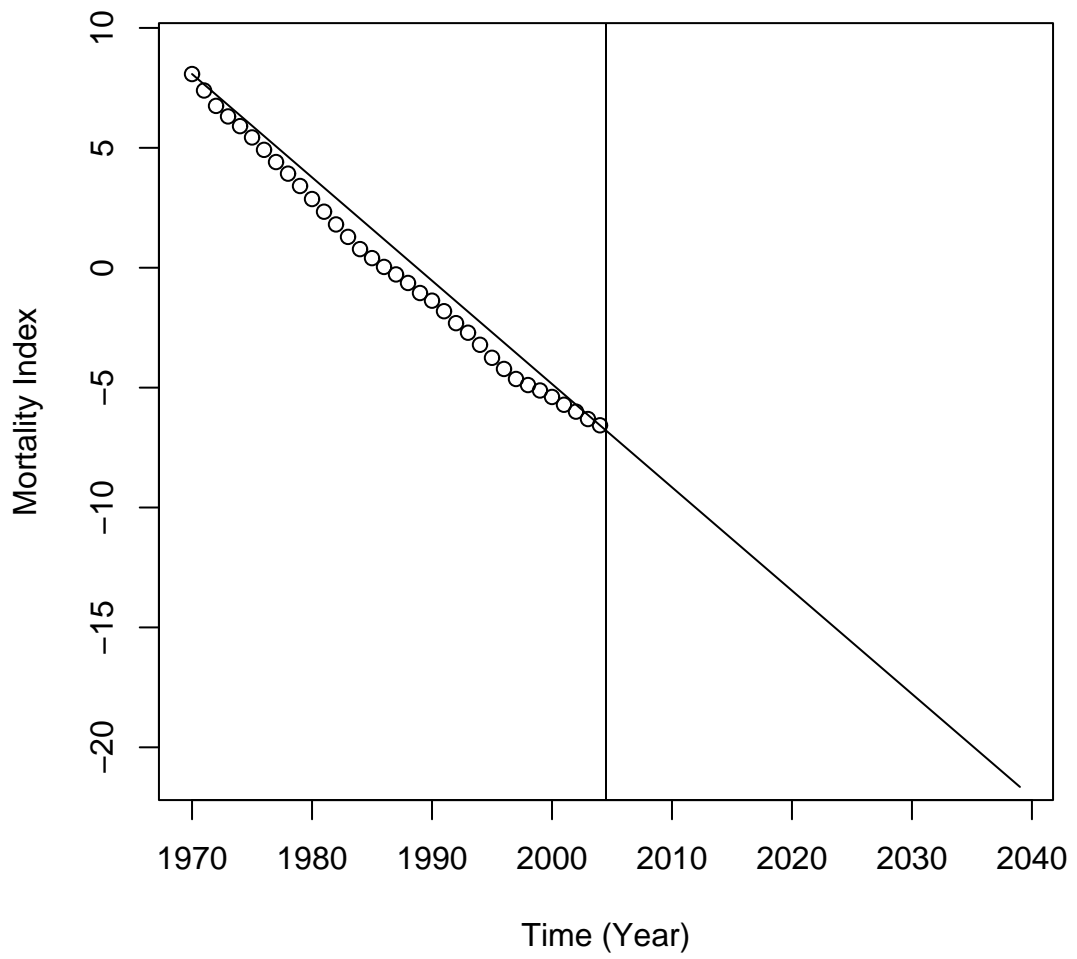


Figure 19: The figure shows estimations (1970-2004) and predictions (2005-2039) of the mortality index for Swedish females aged 65-99.

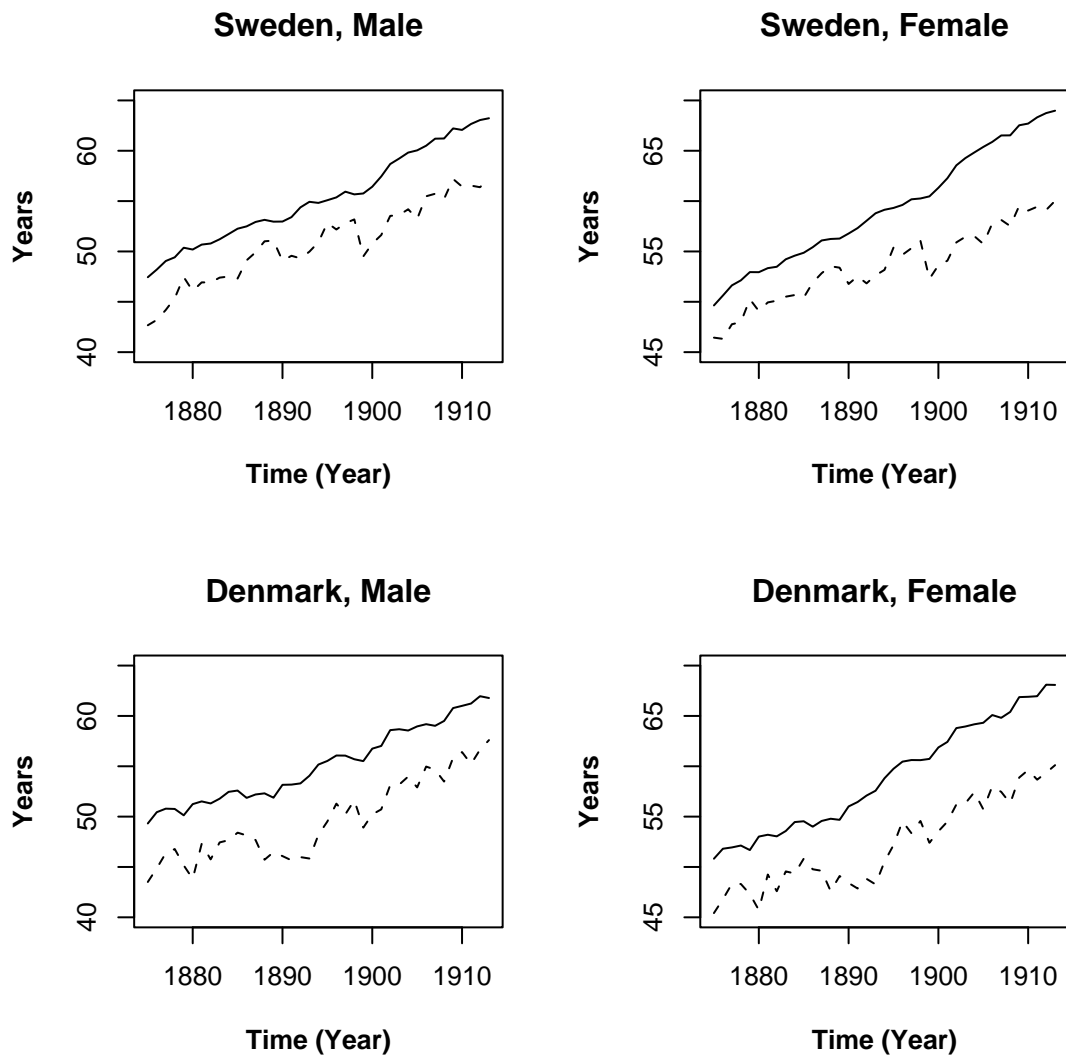


Figure 20: The figure shows the expected lifetime. The full line represent the period life expectancy, while the dashed line represents the cohort expected lifetime.

## Females, Sweden

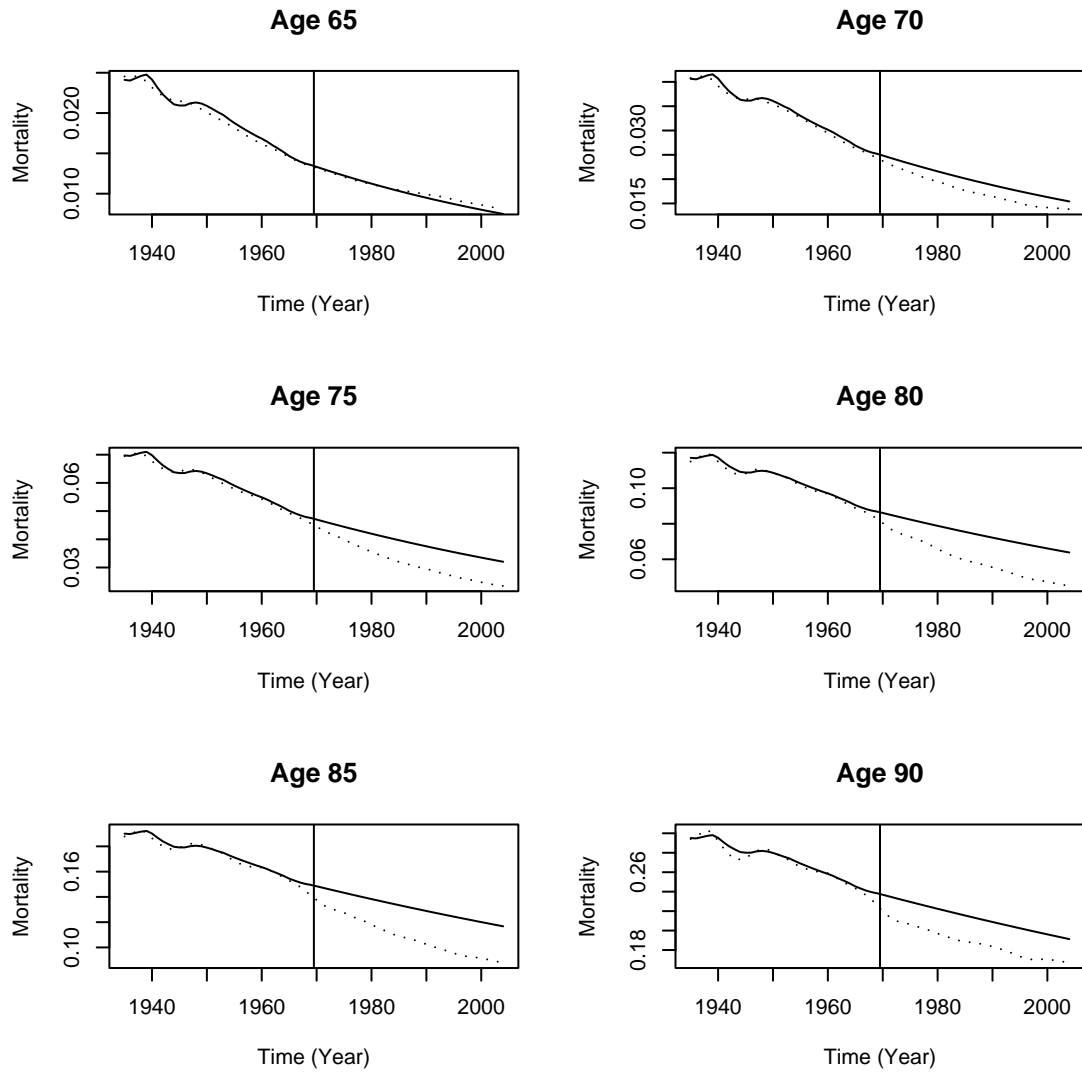


Figure 21: The figure shows the LC estimated and predicted mortality mortality compared to the observed mortality for Swedish females. The dotted line represents the observed smoothed mortality. The full line represents the estimated and predicted mortality. The vertical line separates the estimation and the prediction period.

## Females, Denmark

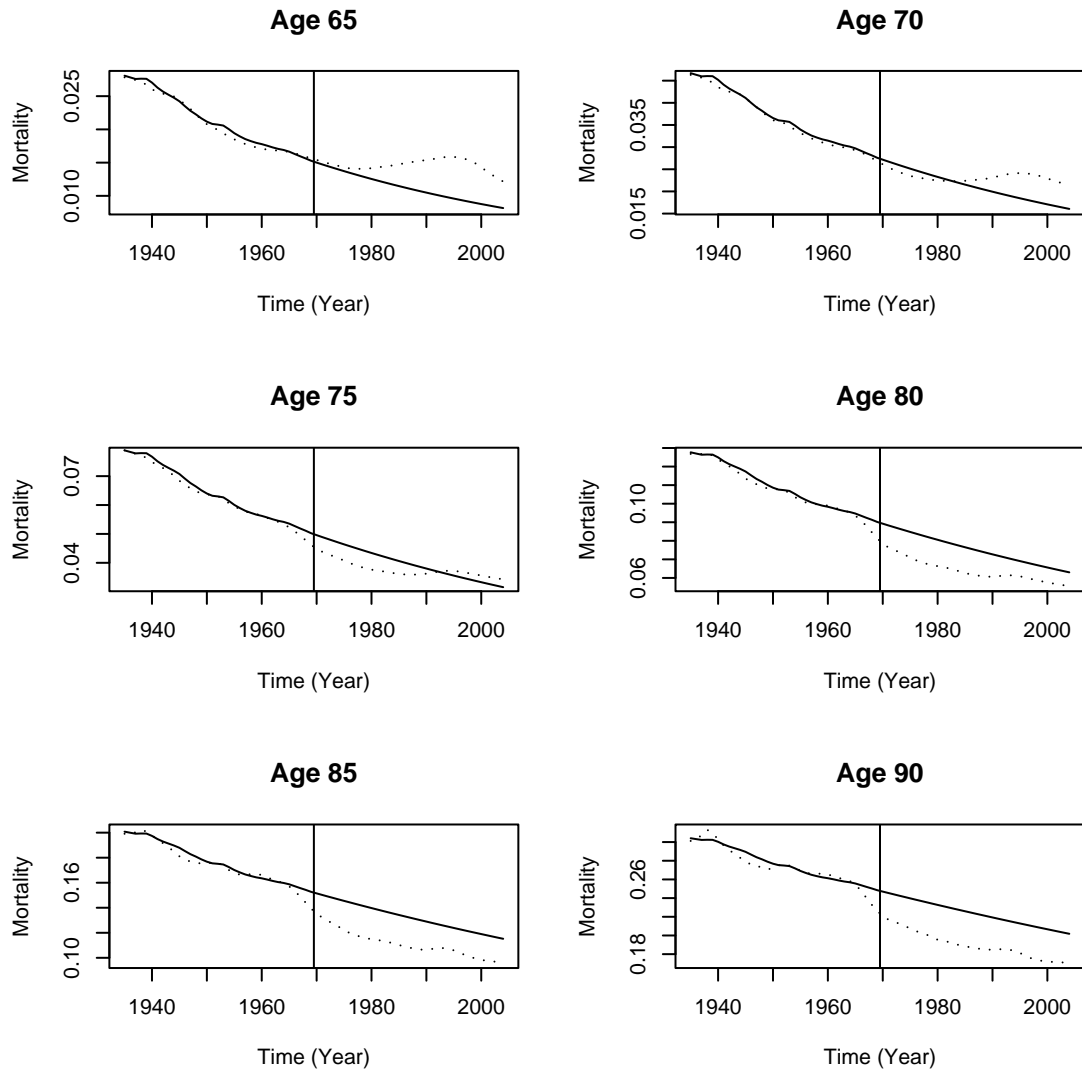


Figure 22: The figure shows the LC estimated and predicted mortality mortality compared to the observed mortality for Danish females. The dotted line represents the observed smoothed mortality. The full line represents the estimated and predicted mortality. The vertical line separates the estimation and the prediction period.

## Males, Sweden

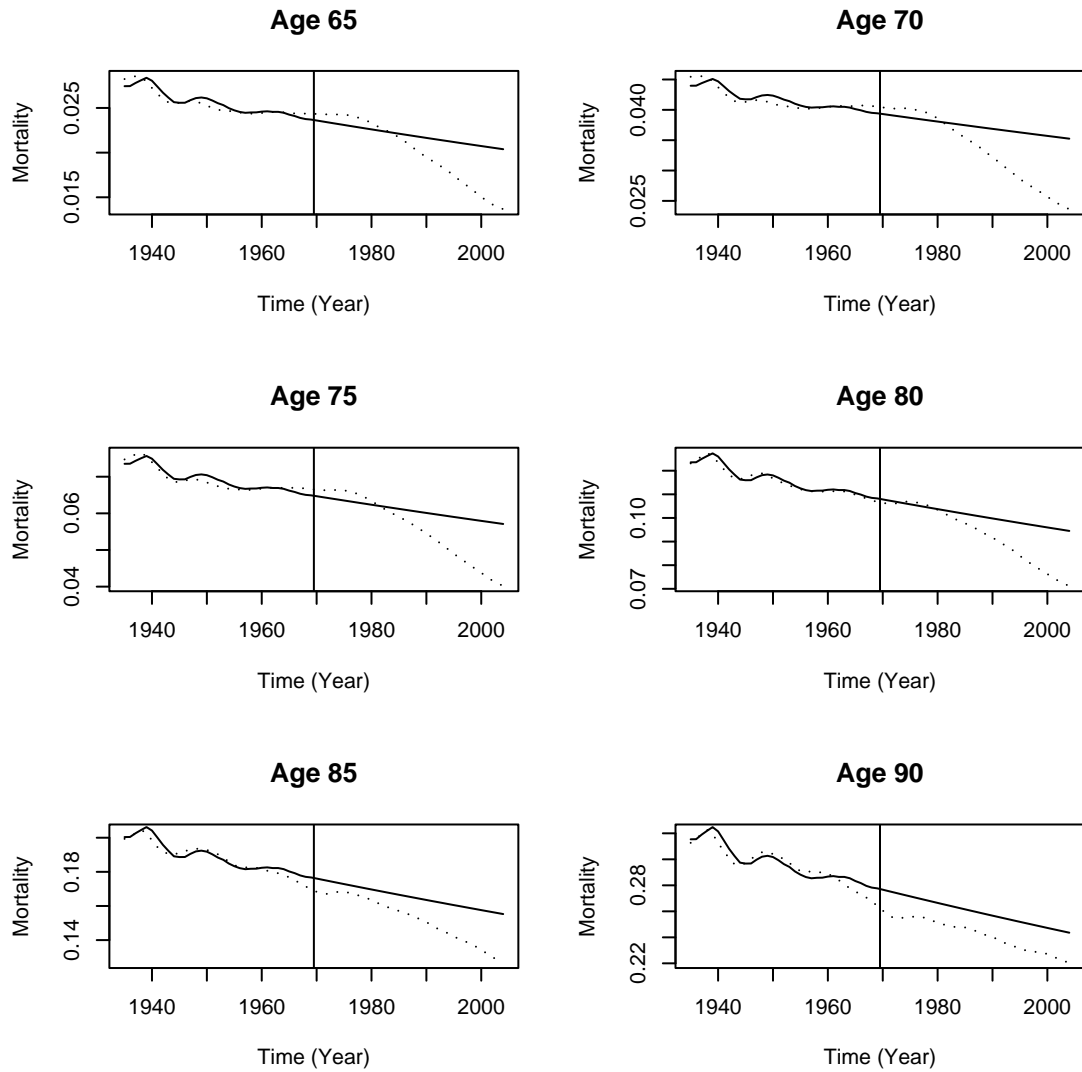


Figure 23: The figure shows the LC estimated and predicted mortality mortality compared to the observed mortality for Swedish males. The dotted line represents the observed smoothed mortality. The full line represents the estimated and predicted mortality. The vertical line separates the estimation and the prediction period.



## Males, Denmark

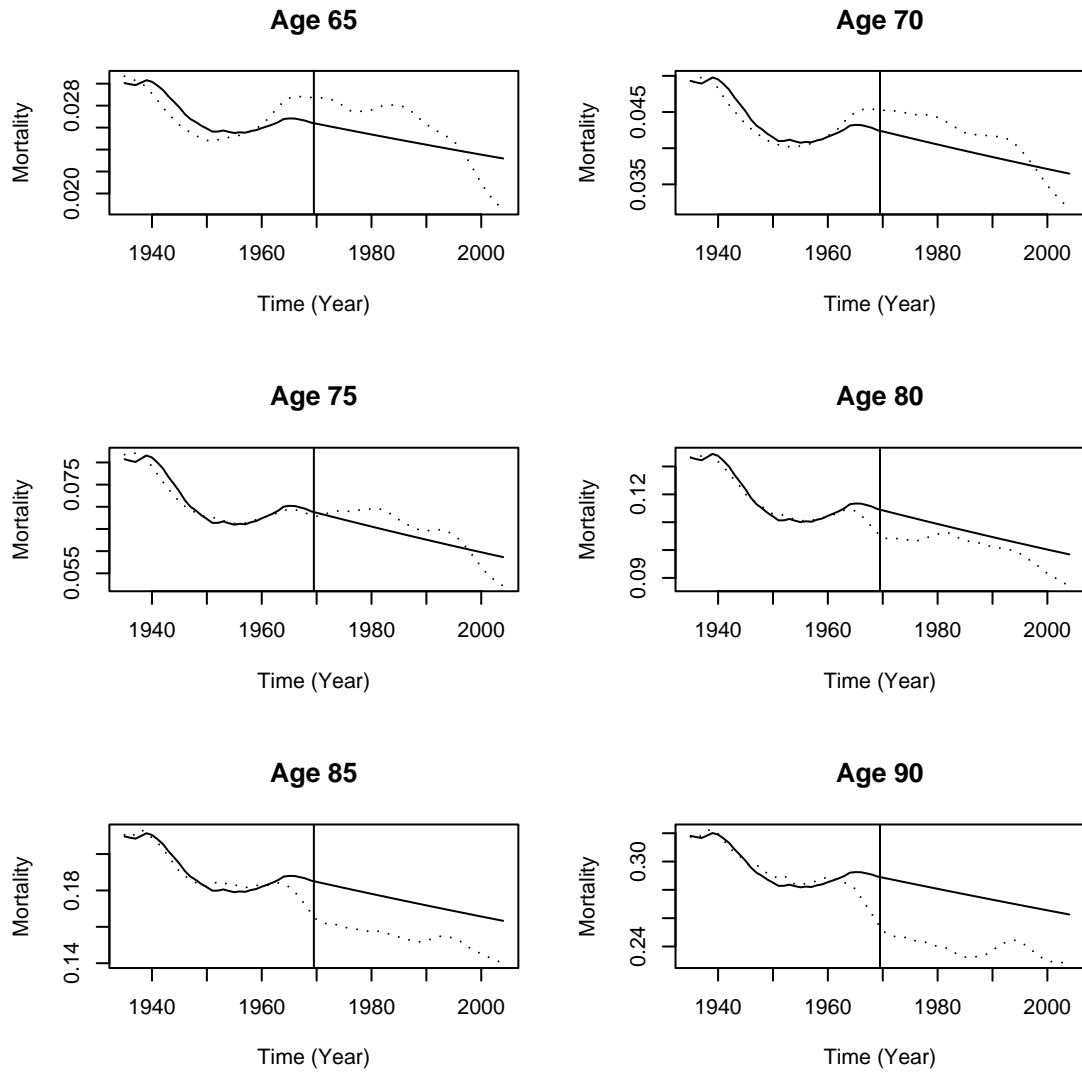


Figure 24: The figure shows the LC estimated and predicted mortality mortality compared to the observed mortality for Danish males. The dotted line represents the observed smoothed mortality. The full line represents the estimated and predicted mortality. The vertical line separates the estimation and the prediction period.

## Female, Sweden

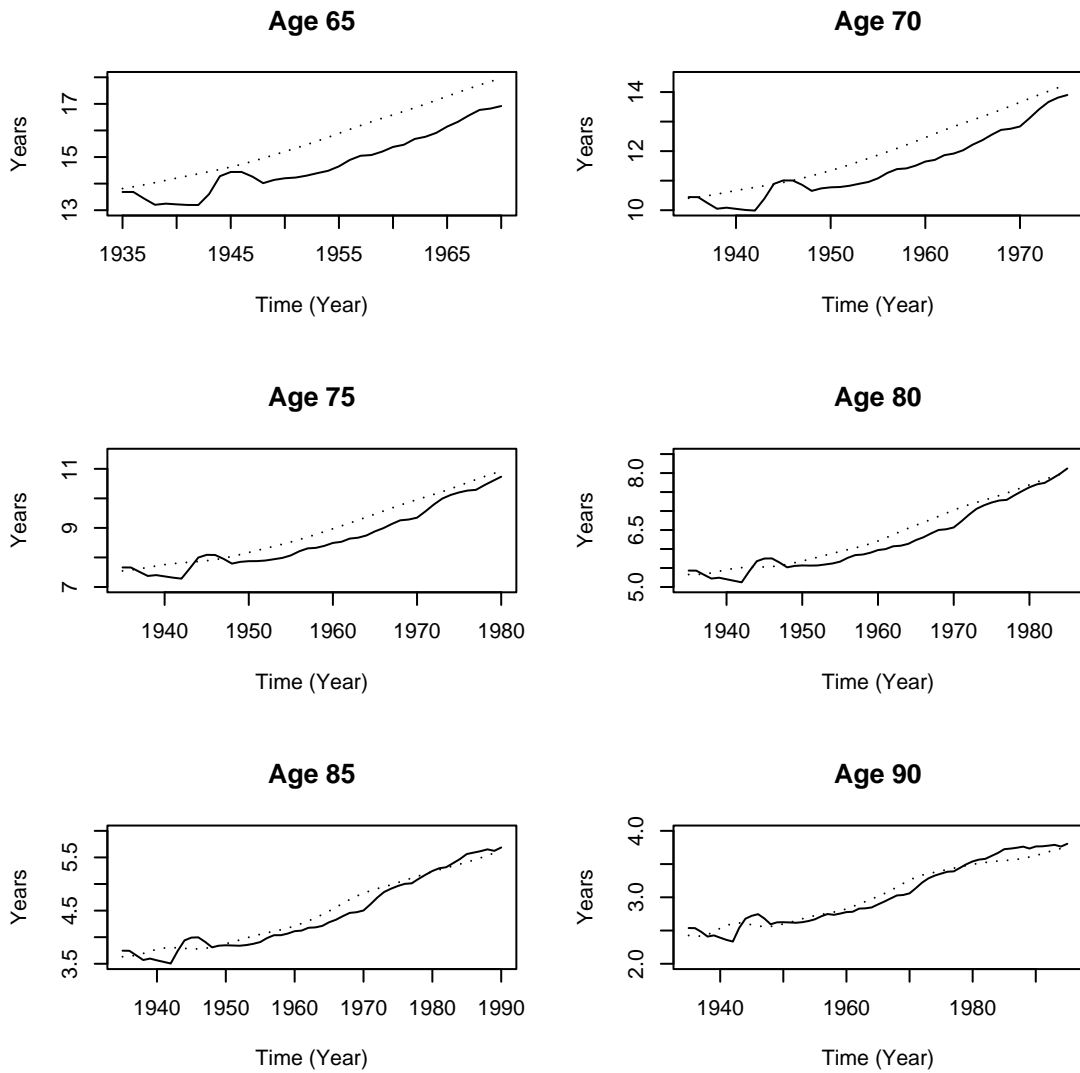


Figure 25: Comparison between the predicted and the observed remaining life expectancy (cohort) for Swedish females. The dotted line represents the observed values, while the full line represents the predictions.

## Female, Denmark

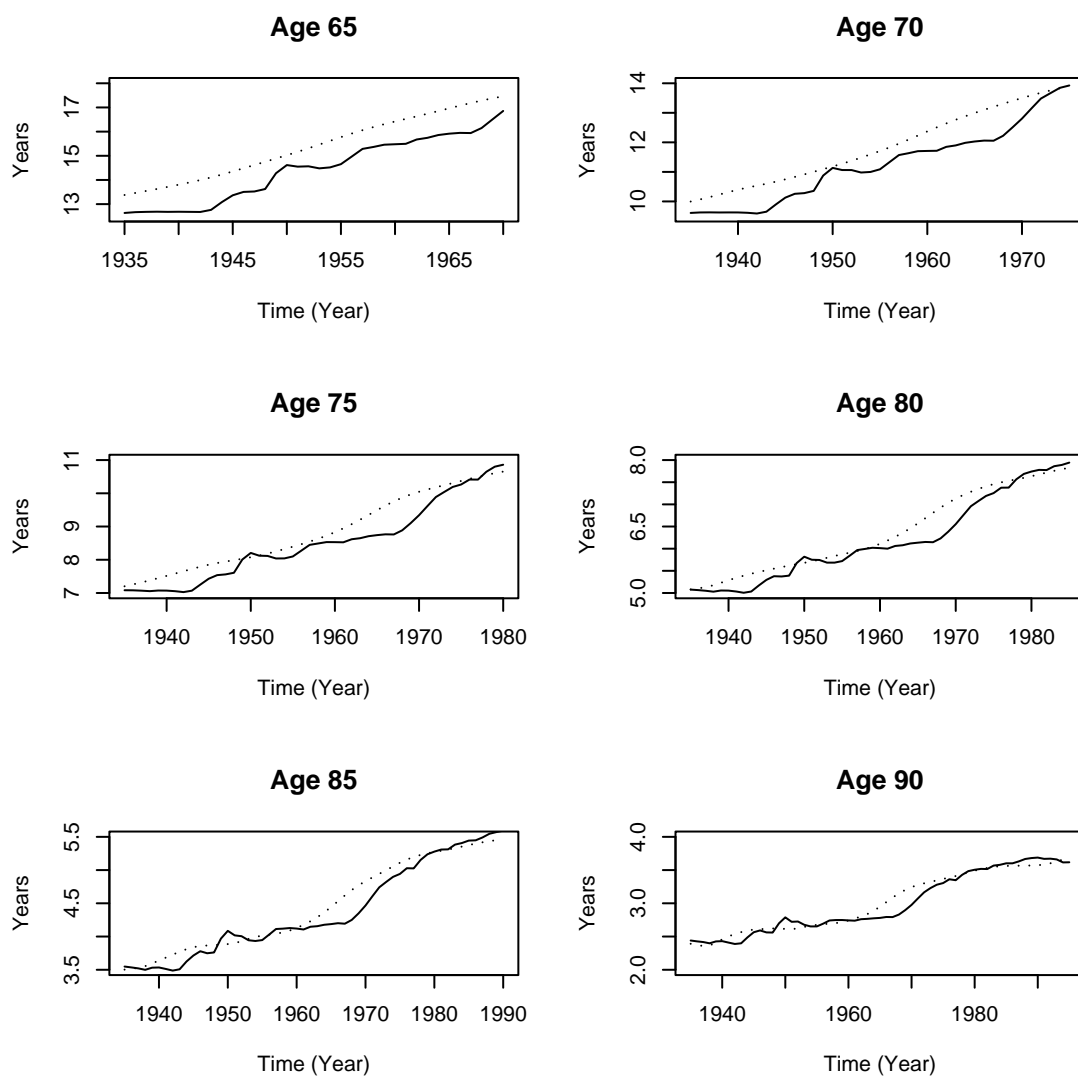


Figure 26: Comparison between the predicted and the observed remaining life expectancy (cohort) for Danish females. The dotted line represents the observed values, while the full line represents the predictions.

## Male, Sweden

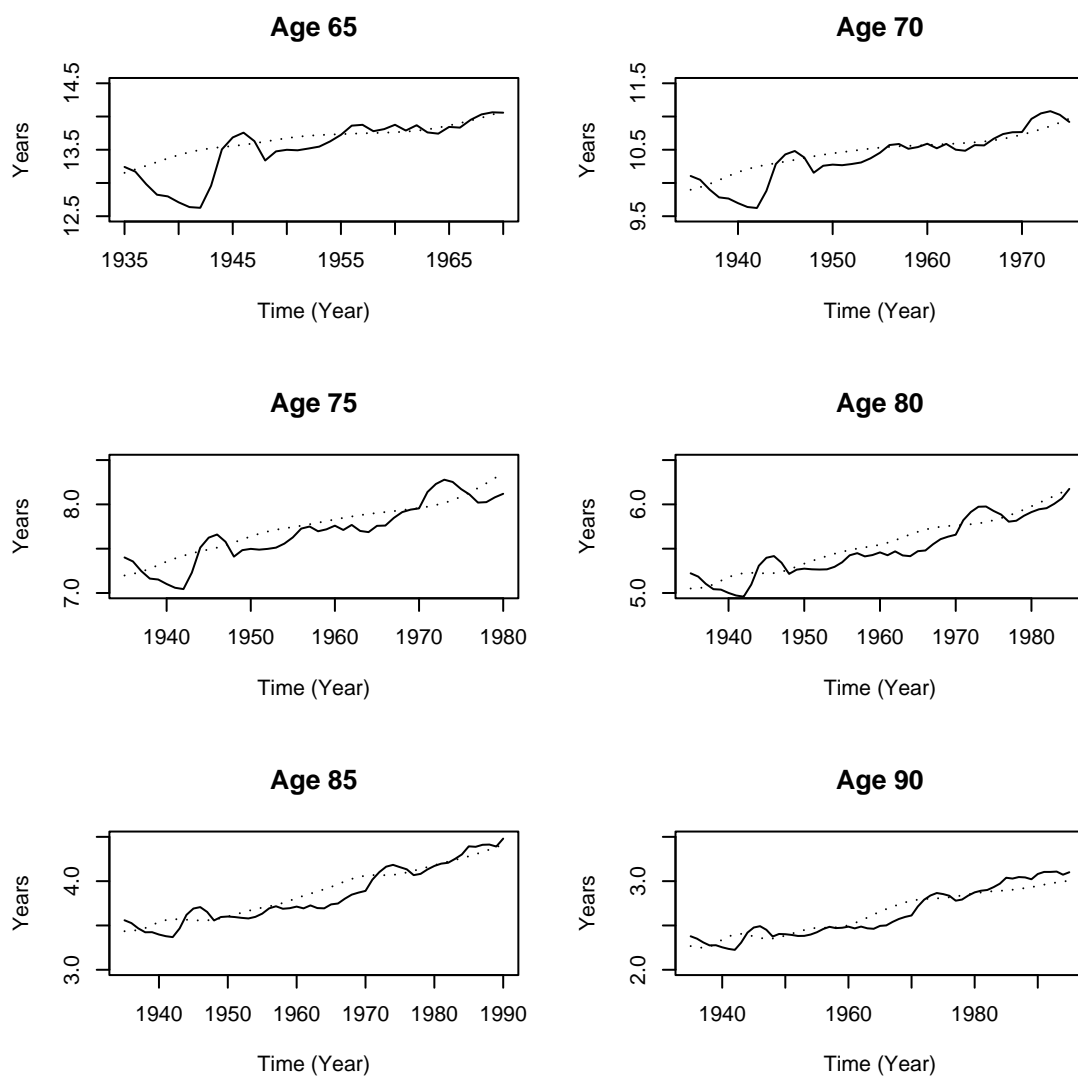


Figure 27: Comparison between the predicted and the observed remaining life expectancy (cohort) for Swedish males. The dotted line represents the observed values, while the full line represents the predictions.

## Male, Denmark

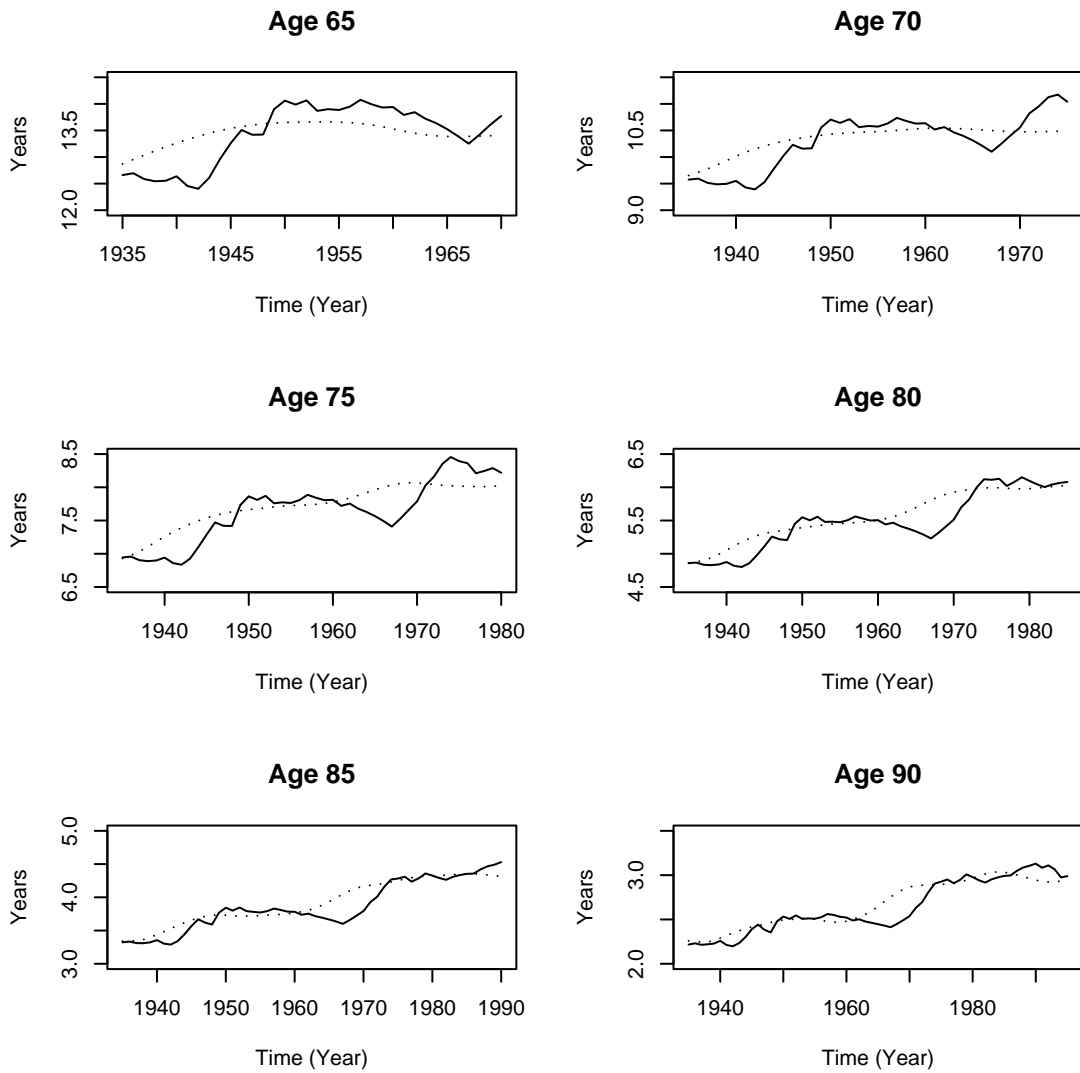


Figure 28: Comparison between the predicted and the observed remaining life expectancy (cohort) for Danish males. The dotted line represents the observed values, while the full line represents the predictions.

Female, Sweden

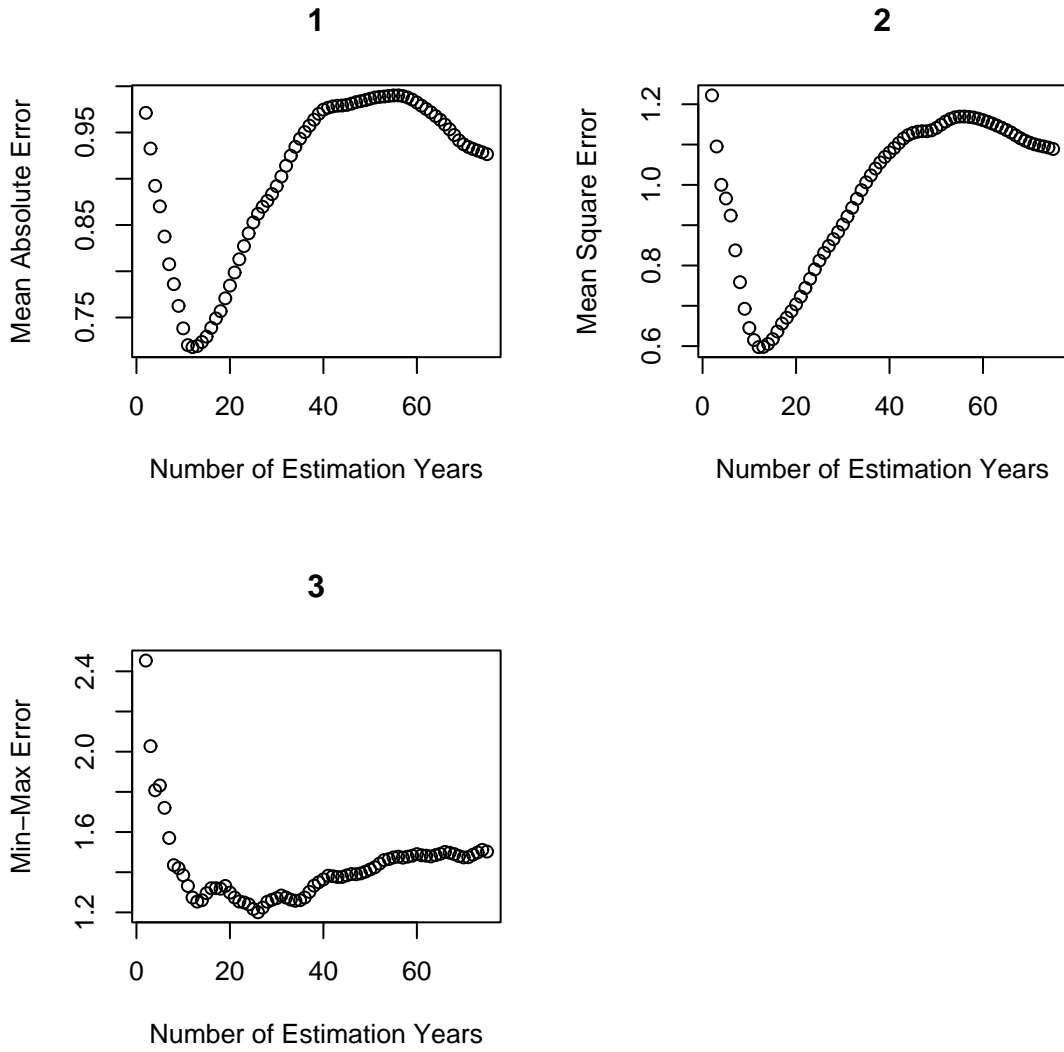


Figure 29: Finding the optimal length of the estimation period.

## Female, Denmark

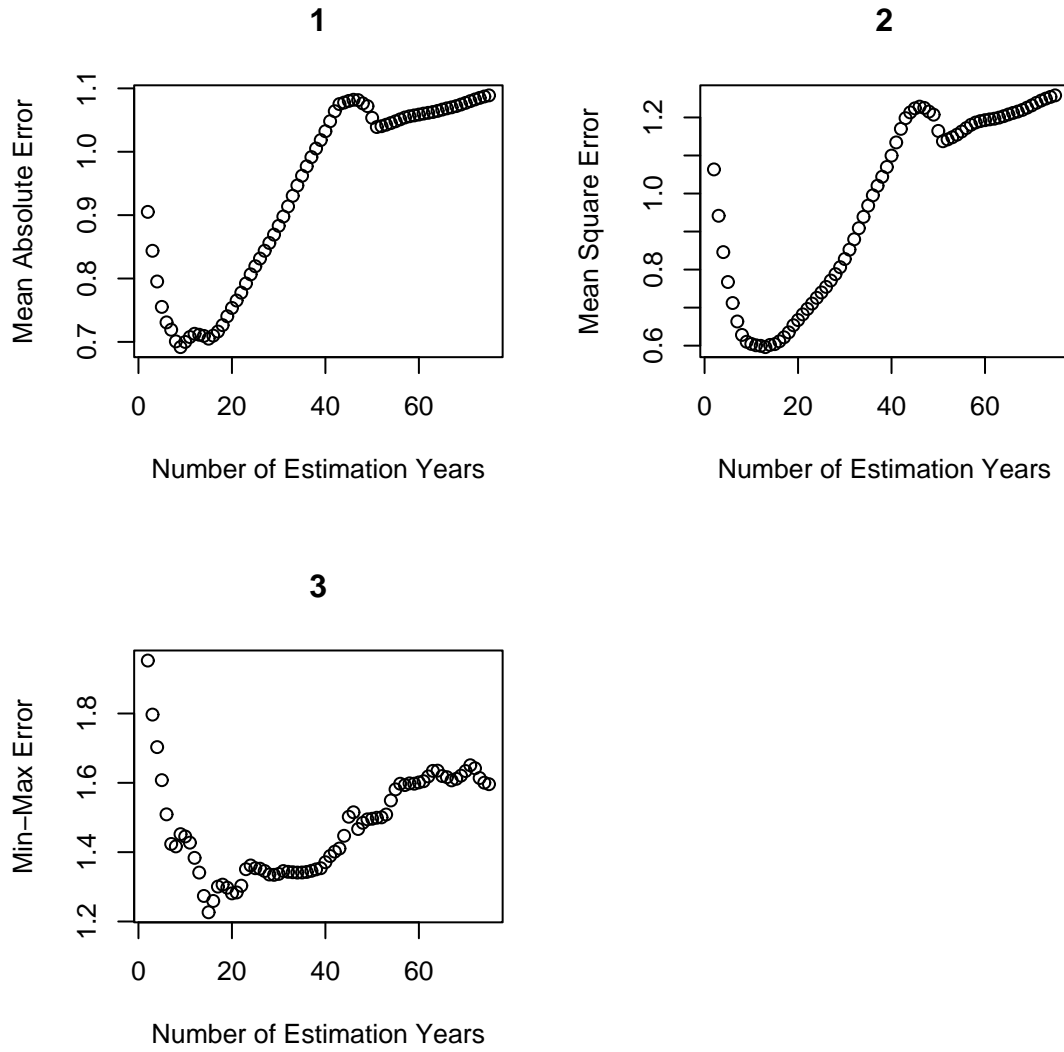


Figure 30: Finding the optimal length of the estimation period.

## Male, Sweden

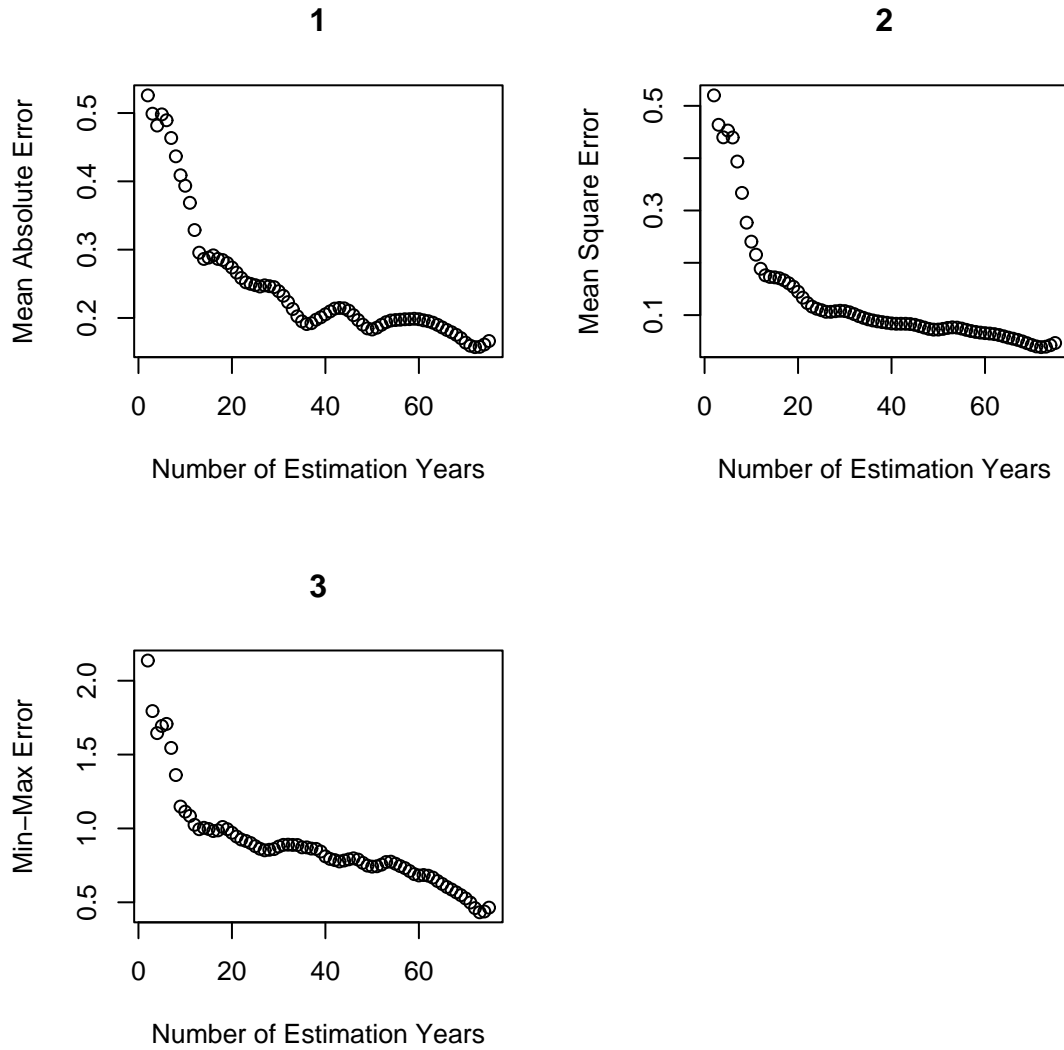


Figure 31: Finding the optimal length of the estimation period.



## Male, Denmark

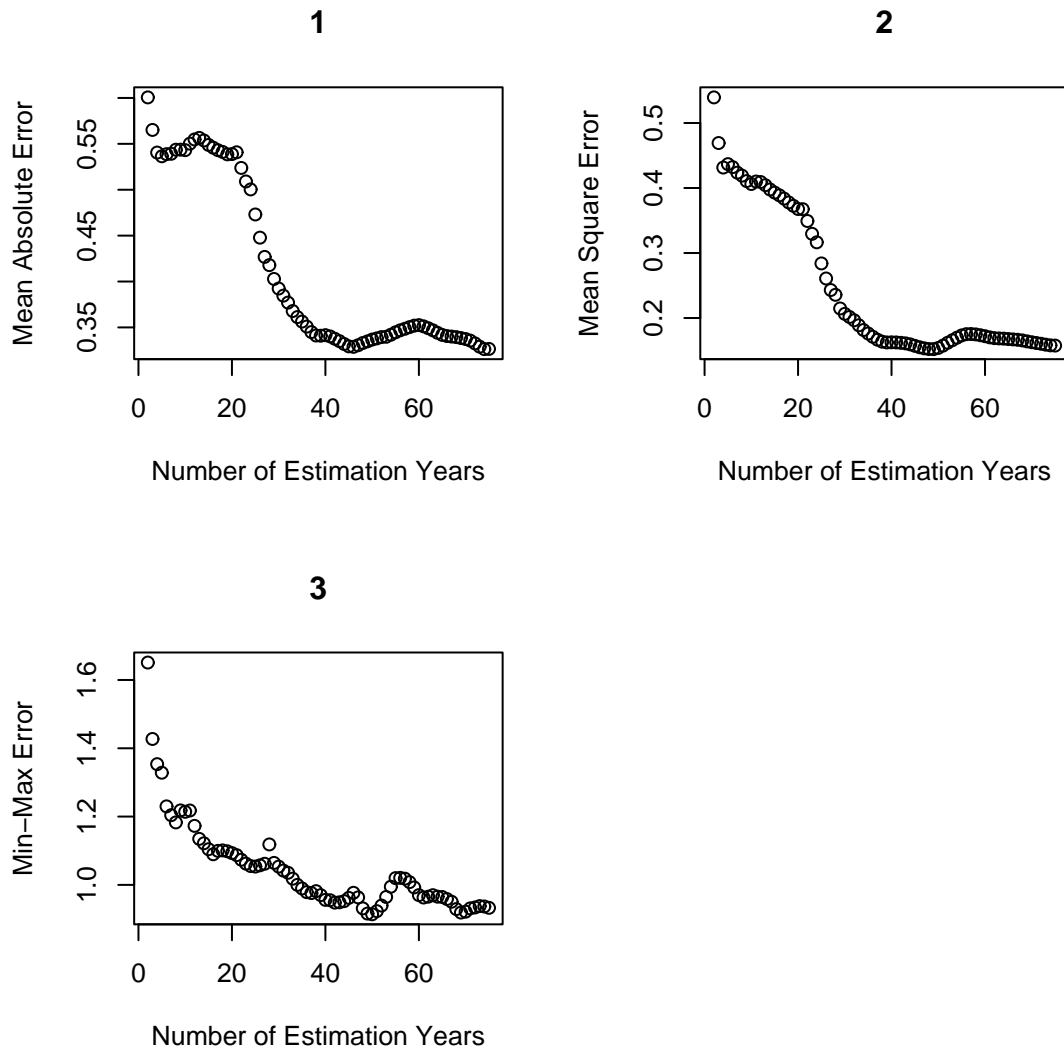


Figure 32: Finding the optimal length of the estimation period.

## Remaining Life Expectancy

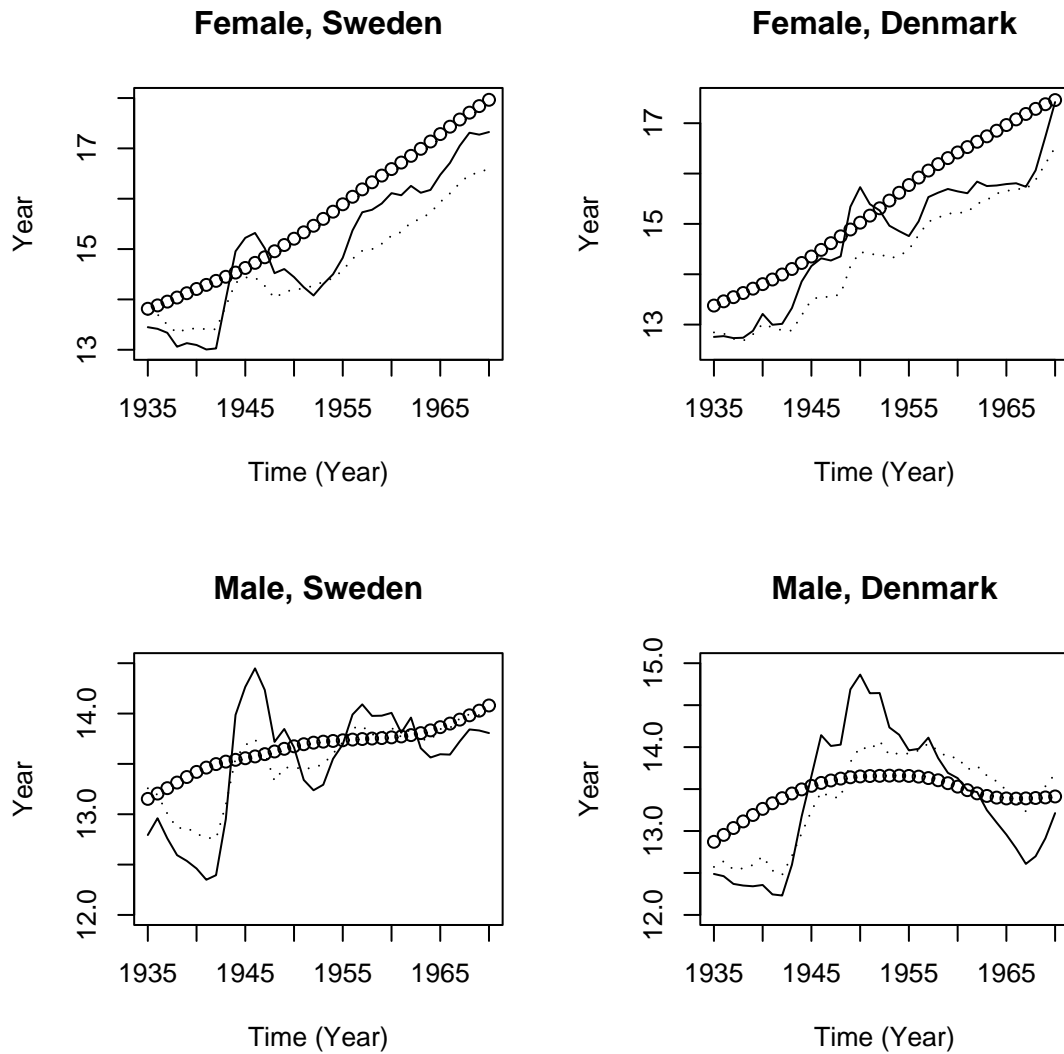


Figure 33:

Points: Observed remaining life expectancy

Full line: Predicted remaining life expectancy (10 Years estimation)

Dotted line: Predicted remaining life expectancy (50 Years estimation)

### Swedish Males and Females, aged 65

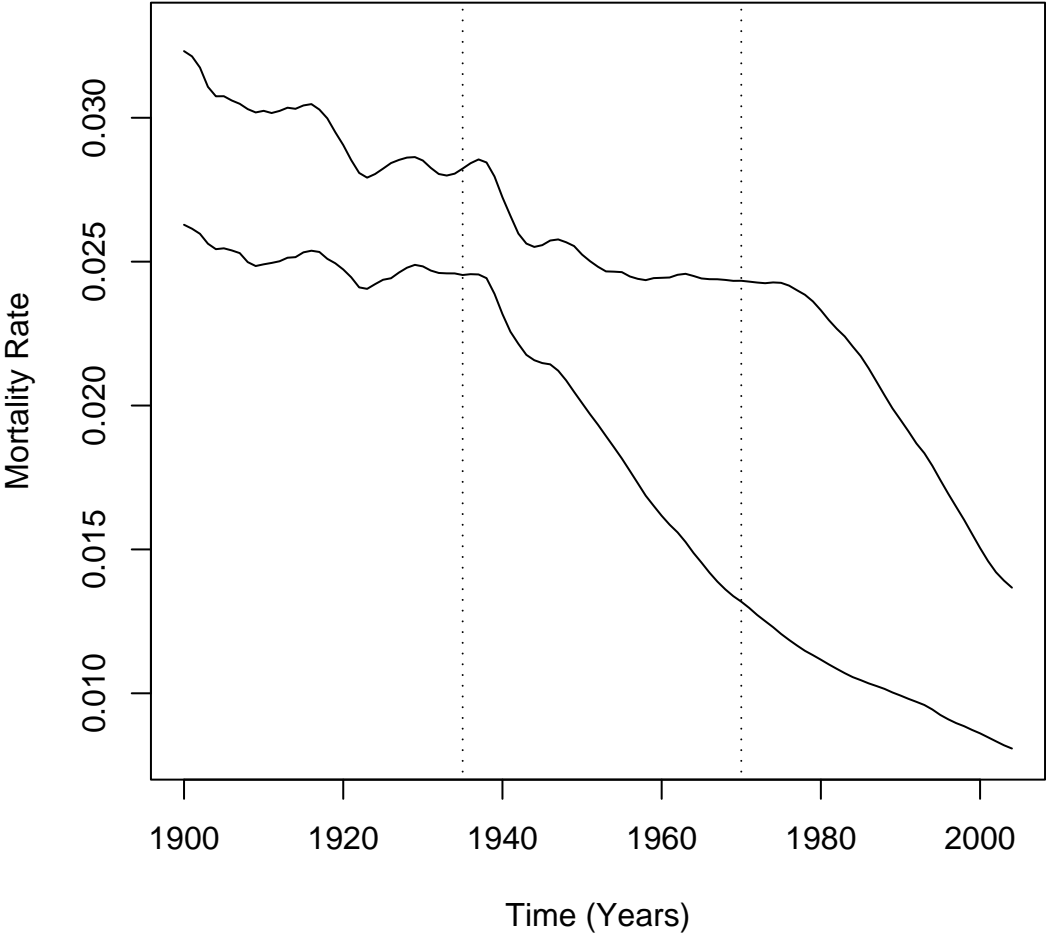


Figure 34: Upper line: Swedish Males. Lower line: Swedish Females.

### Danish Males and Females, aged 65

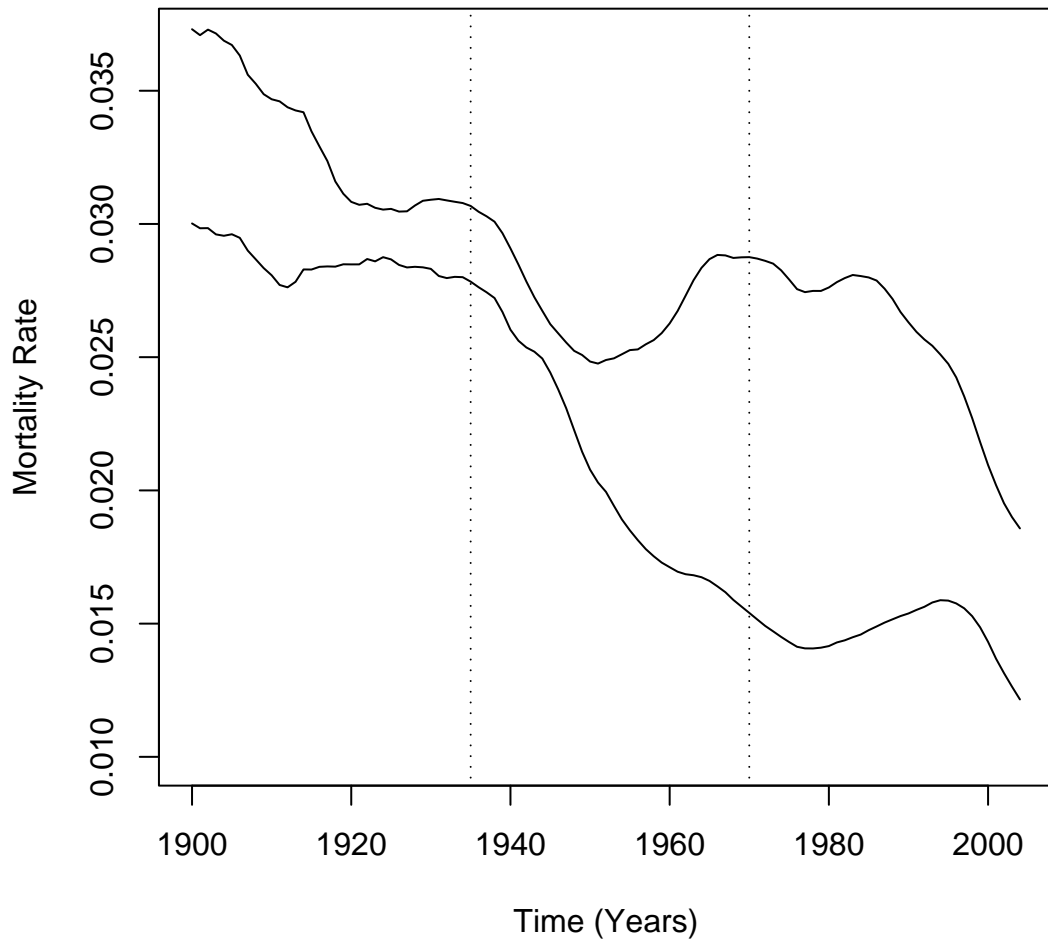


Figure 35: Upper line: Danish Males. Lower line: Danish Females.

### Estimated and Predicted Mortality Index

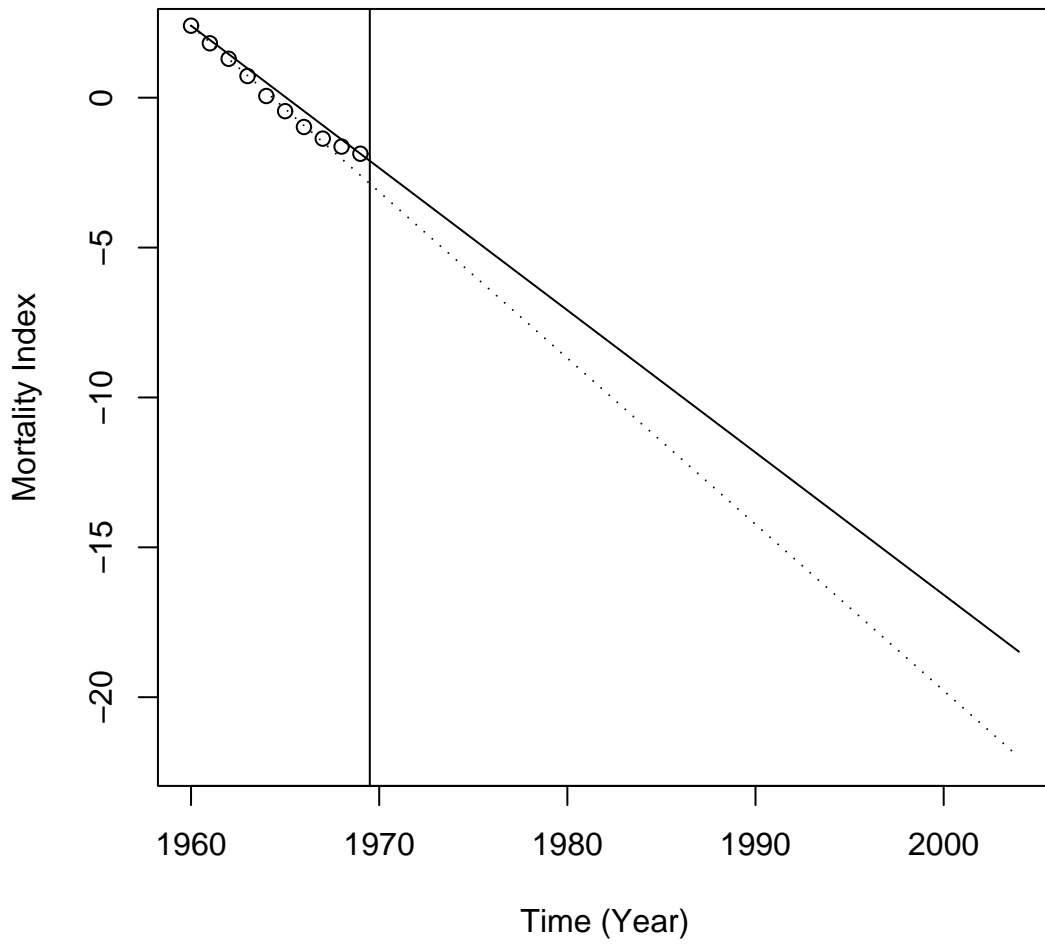


Figure 36: Illustration of the boundary bias problem

## Remaining life expectancy for a 65 year old

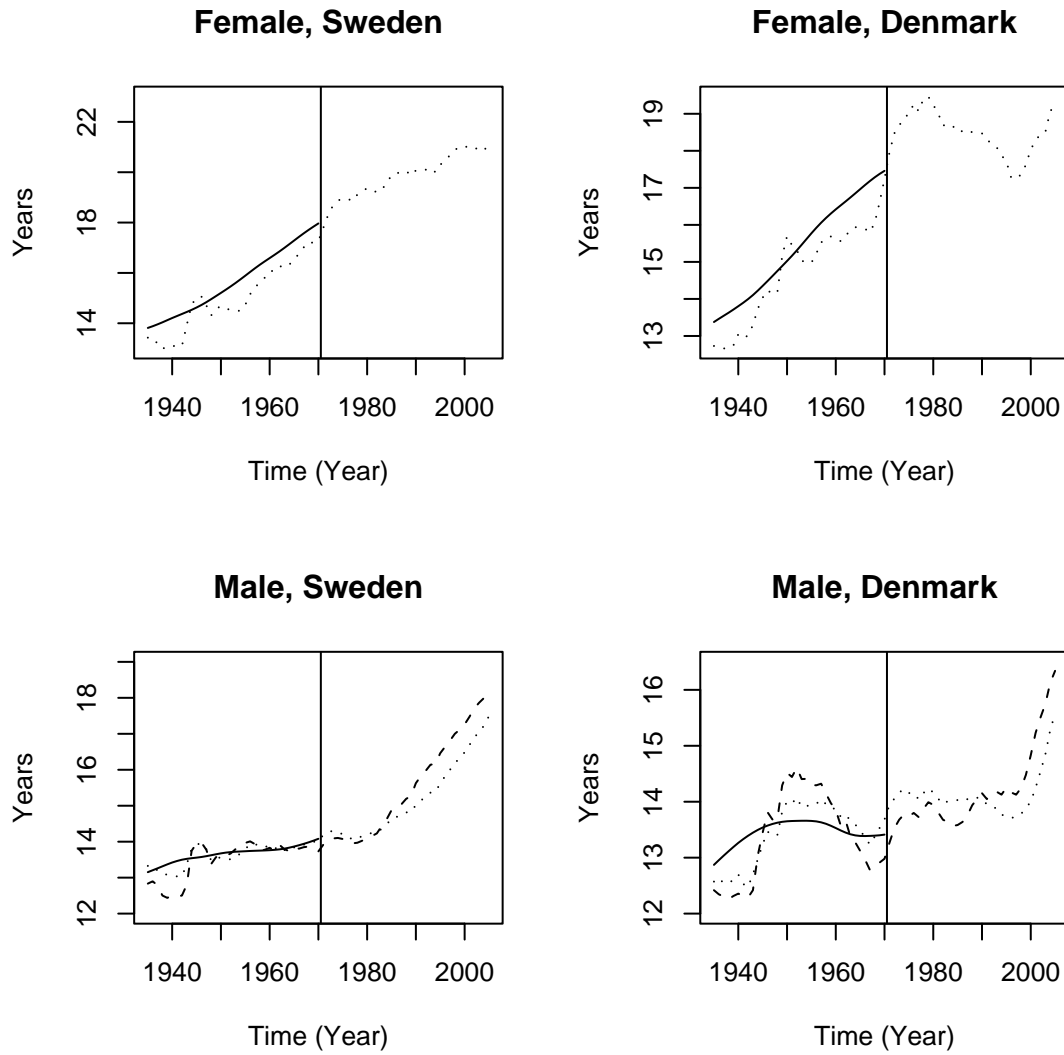


Figure 37:

Dotted line: Predicted remaining life expectancy (alt. 1)

Dashed line: Predicted remaining life expectancy (alt. 2, for males only)

Full line: Observed remaining life expectancy

Horizontal full line: Theoretical remaining life expectancy

## Remaining life expectancy for a 65 year old

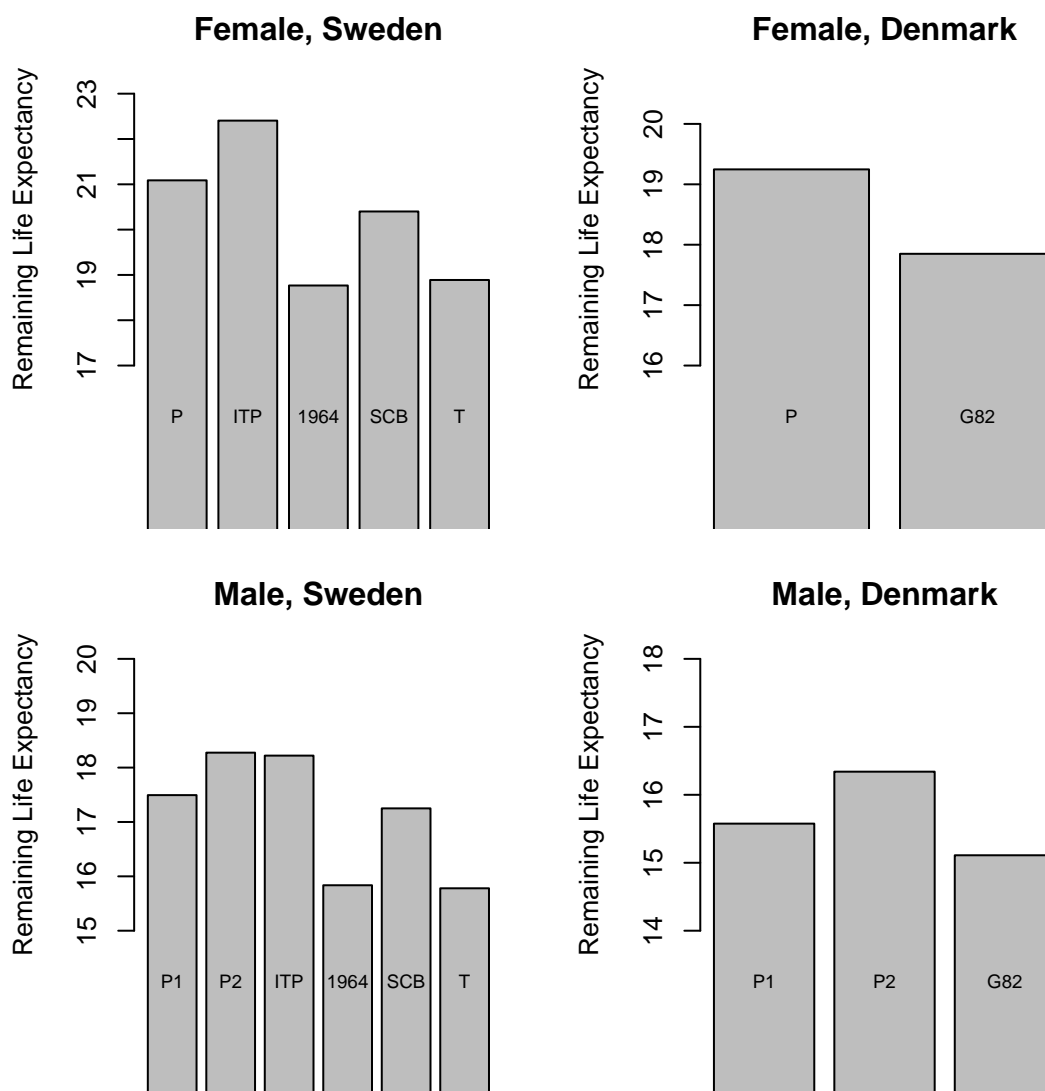


Figure 38:

P: Predicted (12 Years estimation period)

P1: Predicted ("Long" estimation period)    P2: Predicted ("Short" estimation period)

ITP: ITP plan    SCB: Statistiska centralbyrån    1964: 1964 "grunder"    T: Traffic