UNIVERSITY OF
LIVERPOOL

Modelling Mortality Dynamics in Heterogeneous Human Populations

Thesis submitted in accordance with the requirements of the University of Liverpool for the degree of Doctor in Philosophy by

Demetris Avraam

December 2015

Contents

List of Publications

- Avraam Demetris, de Magalhaes Joao Pedro, Vasiev Bakhtier (2013). A mathematical model of mortality dynamics across the lifespan combining heterogeneity and stochastic effects. *Experimental Gerontology* vol 48 issue 8 pp 801-811.

- Avraam Demetris, de Magalhaes Joao Pedro, Arnold (-Gaille) Séverine, Vasiev Bakhtier (2014). Mathematical study of mortality dynamics in heterogeneous population composed of subpopulations following an exponential law. *Book devoted to the 3rd Stochastic Modeling Techniques and Data Analysis (SMTDA2014) International Conference.*

- Avraam Demetris, Arnold (-Gaille) Séverine, Jones Dyfan, Vasiev Bakhtier (2014). Time-evolution of age-dependent mortality patterns in mathematical model of heterogeneous human population. *Experimental Gerontology* vol 60 pp 18-30.

- Avraam Demetris, Arnold (-Gaille) Séverine, Vasieva Olga, Vasiev Bakhtier (2015). On the heterogeneity of human populations as reflected by mortality dynamics. (Submitted).

List of Presentations

- 24 November 2011: *Modelling age-related mortality patterns*. Oral presentation at a lab meeting of Integrative Genomics of Ageing Group, Institute of Integrative Biology, University of Liverpool.

- 14 December 201: *A mathematical model of mortality dynamics combining heterogeneity and stochastic effects*. Oral presentation at a seminar of Applied Mathematics Cluster, Department of Mathematical Sciences, University of Liverpool.

- 21 January 2013: *A mathematical model of mortality dynamics combining heterogeneity and stochastic effects*. Oral presentation at a workshop organised by the Institute for Financial and Actuarial Mathematics (IFAM), Department of Mathematical Sciences, University of Liverpool.

- 28 January 2013: *A mathematical model of mortality dynamics combining heterogeneity and stochastic effects*. Oral presentation at the Perspectives on Actuarial Risks in Talks of Young Researches (PARTY2013) winter school, Ascona, Switzerland.

- 11 April 2013: *Mathematical modelling of mortality dynamics*. Poster presentation at the University of Liverpool Poster Day.

- 03 October 2013: *Time-evolution of human mortality dynamics.* Oral presentation at a lab meeting of Integrative Genomics of Ageing Group, Institute of Integrative Biology, University of Liverpool.

- 11 June 2014: *Mathematical analysis of the role of population heterogeneity in the evolution of human mortality dynamics*. Oral presentation at the 3rd Stochastic Modeling Techniques and Data Analysis International Conference and Demographics Workshop (SMTDA2014), Lisbon, Portugal.

- 02 July 2015: *Time-evolution of age-dependent mortality patterns in mathematical model of heterogeneous human population*. Oral presentation at the 16th Conference of the Applied Stochastic Models and Data Analysis (ASMDA2015) and Demographic Analysis and Research International Workshop (Demographics2015), Piraeus, Greece.

Acknowledgments

I would like to express my sincere gratitude to my supervisor Dr Bakhtier Vasiev for his guidance and support throughout my PhD project. I would also like to thank Dr Joao Pedro de Magalhaes, Dr Séverine Arnold (-Gaille), Dr Olga Vasieva and Mr Dyfan Jones for their useful discussions on that research.

My special thanks go to my family for their continuous support and their lifelong encouragement.

Abstract

The mortality patterns of human populations reflect several inherent biological attributes and other external factors including social, medical and environmental conditions. Mathematical modelling, in addition to experiments and simulations, is an important tool for the analysis of those patterns. One of the main observed characteristics of mortality patterns in human populations is the age-specific increase in mortality rate after sexual maturity. This increase is predominantly exponential and satisfies the well-known Gompertz law of mortality. Although the exponential growth in mortality rate is observed over a wide range of ages, it excludes early- and late-life intervals.

The heterogeneity of human populations is a common consideration in describing and validating their various age-related features. In this study we develop a mathematical model that combines (i) the assumption of heterogeneity within each human population, where different subpopulations are distinguished for their certain mortality dynamics and (ii) the assumption that the mortality of each constituent subpopulation increases exponentially with age (in the same way as described by the Gompertz law). The proposed model is used to fit available observational data in order to analyse the dynamics of mortality across the lifespan and the evolution of mortality patterns over time.

We first explore the effects of the variation of the model parameters to the dynamics of mortality and use the model to fit actual age-specific mortality data. We show that the model successfully reproduces the entire age-dependent mortality patterns explaining the peculiarities of mortality at young and very old ages. In particular, we show that the mortality data on Swedish populations can be reproduced fairly well by a model comprising of four subpopulations. Besides the confirmation that heterogeneity can explain the irregularities of mortality patterns at young ages and the deceleration of mortality at extremely old ages, we analyse the influence of stochastic effects on mortality and we conclude that evident effects due to stochasticity are manifested at the age intervals (early and late life ages) where only few individuals contribute to mortality.

We then analyse the evolution of mortality patterns over time by fitting the proposed model to (Swedish) mortality data of consecutive periods across the 20$^{\text{th}}$ century. The evolution of mortality is described in terms of the changes of model parameters estimated by fitting the model to data from different time periods. We show that the evolution of model parameters confirms the applicability of the compensation law of mortality to each constituent

subpopulation separately. The compensation law states an inverse relationship between the scale and the shape parameter of Gompertz law. Our analysis also indicates a change in the structure of this population over time in a way that the population tends to become more homogeneous by the end of the 20th century. This change in structure is reflected in changes to the initial proportions of the constituent subpopulations. These two observations, namely the validity of the compensation effect and the homogenisation of the population, imply that the alteration of model parameters (which reflect demographic terms) can explain the decrease of the overall mortality over time. It is shown that the decrease in mortality across the 20th century is mainly due to changes in the structure of the population, and to a lesser extent, to a reduction in mortality for each of the subpopulations.

The outcomes of our research show that the consideration of heterogeneity is efficient for the description of various features of a population's mortality. The idea of "pure" subpopulations, such that in each of them exponential law is held for all ages, has been used as a convenient mathematical constraint which allows very accurate reproduction of the entire mortality patterns. This provides a justification for the deviation of mortality from its exponential increase at young and very-old ages and for the decrease of mortality over time. In the last part of this thesis we propose that the proposed heterogeneity is not only a convenient tool for fitting mortality data but indeed reflects the true heterogeneous structure of the population. Particularly we demonstrate that the model of a heterogeneous population fits mortality data better than most of the other commonly used models if the data are taken for the entire lifespan and better than all other models if we consider only old ages. Also, we show that the model can reproduce seemingly contradicting observations in late-life mortality dynamics like deceleration, levelling-off and mortality decline. Finally, assuming that the differences between subpopulations reflect genetic variations within the population and using the Swedish mortality data for the 20th century, we show that evolutionary processes resulting in changes of allele frequencies, can explain the homogenisation of the population as predicted by the model.

Thesis Outline

This thesis is comprised of five chapters:

Chapter 1 presents the main mortality-related observations, outlines several mathematical models of mortality and describes our motivation and methodology. Here, the focus is mainly on age-dependent observations represented by three phenomena known as the "three laws of mortality". The first law is the *Gompertz law* which describes the exponential increase of mortality rate for a certain range of ages. The second law is the *compensation effect* which states that high initial mortality rate in a population is compensated by low rate of mortality change with age. The third law known as *late-life mortality deceleration* states that the mortality rate in later life increases at a lower rate than the exponential increase during a wide portion of lifespan as described by the Gompertz law. In addition to age-dependent observations, we describe the compression of mortality and the "rectangularization" of survival curve which are both indicators of the evolution of mortality over time. Moreover, we outline a number of mathematical models that have been developed to describe mortality-related observations and we specify the modelling approaches that take into account the heterogeneity of populations. The Chapter ends with the research aims which include the development of a model able to describe and analyse the observations of mortality and with the description of fitting procedure and selection criterion applied in the current study.

In Chapter 2 we present the results of a study published in the 48th volume of *Experimental Gerontology* journal in 2013. In this Chapter we develop the mathematical model that will be used in the analysis of the dynamics of mortality across the lifespan. The model combines the consideration of population heterogeneity with the assumption that the mortality dynamics of each constituent subpopulation follow an exponential law. The model attains to generate a mortality trajectory that accurately fits an entire dataset of actual mortality rates and particularly to reproduce the peculiarities of mortality patterns that are observed at early and late life intervals. We also analyse the influence of stochastic effects on the mortality dynamics to show that they play a role only at young and very old ages, when only a few individuals contribute to mortality. We conclude that the deviations from the exponential law at young ages can be explained by the heterogeneity of populations, while the deviations at old ages can be viewed as fluctuations and explained by stochastic effects.

Chapter 3 is based on a work published in the 60th volume of *Experimental Gerontology* journal in 2014. In that Chapter we use the model of a heterogeneous population (described in

Chapter 2) to fit actual mortality data of one-century period in order to analyse the evolution of mortality dynamics over time. The analysis is performed by examining the changes of optimum model parameters that fit data from consecutive periods. This study shows that the evolution of model parameters validates the applicability of the compensation law of mortality to each subpopulation separately. Furthermore, the analysis indicates that the population's structure changes in a way that the population tends to become more homogeneous over time. We finally show that the decrease of the overall mortality in the Swedish population over the 20th century is mainly due to the homogenisation of the population and to a lesser extent to a reduction of mortality of each of the constituent subpopulations, the latter being represented by an alteration to the scale and shape parameters of Gompertz functions.

In Chapter 4 we present the results submitted for publication in *Experimental Gerontology* journal. In that Chapter we present a few arguments to say that the heterogeneity in human populations is not only a convenient constraint used to accurately fit actual mortality data but indeed reflects the real structure of the population. We first show that the model of a heterogeneous population fits mortality data better than most of the other models if the data from the entire lifespan are used and better than all other models if we consider only old ages. Also, we show that the model can reproduce seemingly contradictory observations in late-life mortality dynamics which include deceleration, levelling-off and mortality decline. Furthermore, assuming that the existence of subpopulations reflects the genetic variations in the population, we show that the homogenisation of the Swedish population over the 20th century, as predicted by the model, can be associated with the evolution of allele frequencies.

Chapter 5 discusses the findings of this research and provides directions for future work. The discussion includes an analysis of the model by summarising its advantages, demonstrating further applications (continuous model of mortality, probability density and survival function in heterogeneous populations) and indicating its limitations. We then describe four directions for future research that provides an extension of the present study and we close the Chapter (and thesis) with our final conclusions.

Chapter 1.

Introduction

Studies of ageing and senescence attract the attention of scientists from a variety of disciplines including biology, medicine, actuarial mathematics and epidemiology. One of the central problems addressed in these studies is to understand what fails in the organisms with age and how to prevent and delay these changes (Olshansky, 1998). With a dramatic increase in human longevity, the process of biological ageing and mortality becomes also a central issue for many practical concerns related to the development of modern society. Ageing is defined as an inevitable progressive decline of physiological function with increasing age that is associated with a decline in fecundity and an increase of mortality (Rose, 1991, Bronikowski and Flatt, 2010). Demographically, ageing is manifested by the acceleration of mortality with age and mathematically expressed by the Gompertz law of mortality. Therefore, while the death rate is a proxy indicator of ageing, a number of studies attempt to analyse the dynamics of mortality in order to understand the processes underlying human ageing.

In a similar way, the premise of our research is to use mathematical techniques to analyse the dynamics of mortality over age and time in human populations. In this Chapter which is the introduction to our research we present the main mortality-related age and time-dependent observations (Section 1.1), outline an overview of several mathematical models of mortality (Section 1.2) and present our motivation and methodology (Section 1.3).

1.1. Basic mortality-related observations

Mortality–related data such as death rates, survival probabilities and probabilities of death have specific characteristics at different ages and at different periods. Some of those characteristics are well understood while others still look paradoxical and remain biologically unexplained. For example the exponential increase of mortality rate with age, observed after sexual maturity, has been known for two centuries but still remains to be an empirical observation without a uniform biological justification. Another paradoxical observation is the phenomenon known as the compensation effect which describes the decrease of the relative differences in death rates with age, between different populations. This is observed when higher initial death rates in disadvantaged populations are compensated by a lower rate of mortality increase with age (Gavrilov and Gavrilova, 1979). The main age and time dependent observations related to mortality are described in the following sections.

1.1.1. Lexis diagram and mortality rates

The analysis of human mortality dynamics over the lifespan and their evolution over time requires an introduction of some common notations and definitions. These definitions include the measurements that are tabulated in life tables. The life tables provide essential information on the age-related structure of a certain population and on the survivorship of its individuals at specific ages (Preston et al., 2000). The life tables contain probabilities of death and survival, numbers of survivors and numbers of deaths, life expectancies, death rates and other mortality-related quantities at each age.

There are two types of life tables, period and cohort, which correspond to two different ways the data are recorded. Data recording deaths of individuals born in a certain year (or time interval of another length) form the cohort data. Data recording deaths in a certain population occurring during a specific year form the period data. Both kinds of data can be illustrated by a Lexis diagram. The Lexis diagram (Lexis, 1875) outlines the stocks and flows of a population and the occurrence of demographic events (such as deaths) over age and time. It is a two-dimensional graph (Figure 1.1) where the vertical axis represents age and the horizontal axis represents time, both measured in the same units (e.g. years). Deaths occurring in a parallelogram formed by two diagonal lines in Figure 1.1 contribute to cohort mortality, while the deaths occurring in a rectangle outlined by two vertical lines in Figure 1.1 refer to period mortality.

2
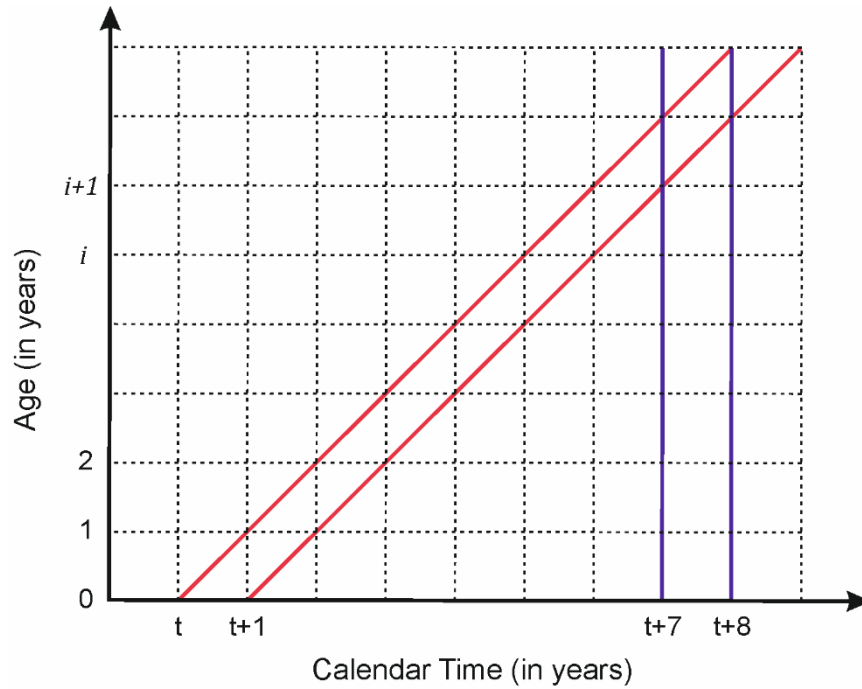
Figure 1.1: Lexis diagram. *The diagram illustrates demographic events as distributed over age and time. Cohort mortality rates refer to the deaths in a cohort that are occurring in a parallelogram formed by two diagonal lines. Period mortality rates refer to deaths occurring within a period outlined by two vertical lines.*

Two of the basic quantities of interest that are tabulated in a human life table are the probability of death and the mortality rate. Probability of death, q_i, is the probability that an individual aged i will die before they reach age $i + 1$. The probability of death is expressed as the ratio of the number of deaths of people aged i, ΔN_i, divided by the number, N_i, of individuals who have reached age i in the cohort data or alive individuals at exact age i in the period data. Death (or mortality) rate m_i is defined as the number of deaths of people aged i divided by the number of person years of age i:

$$m_i = \frac{\Delta N_i}{PY_i} \approx \frac{\Delta N_i}{0.5(N_i + N_{i+1})},$$

(1.1)

where the number of deaths of people aged i is represented as $\Delta N_i = N_i - N_{i+1}$. The number of person years can be approximated by the average number of survivors within a one-year age interval which roughly coincides with the number of survivors at the centre of the interval and therefore the mortality rate is commonly referred to as the central death rate.

When mortality is considered as a continuous process, age is defined by a real number x (continuous age) rather than by the integer number i (discrete age). In this case, the instantaneous mortality (force of mortality), $\mu(x)$, at age x is defined by

$$\mu(x) = \lim_{\Delta x \to 0} \frac{-\Delta N(x)}{N(x)\Delta x} = \frac{-1}{N(x)}\frac{dN(x)}{dx}. \tag{1.2}$$

Discrete and continuous descriptions of mortality are almost identical (the discrete form can be derived from the continuous one) and mathematical modelling can be performed in both forms. The characteristics of mortality dynamics across the life span and observations on mortality-related measurements over time are described in Sections 1.1.2 and 1.1.3.

1.1.2. Observations on age-specific mortality dynamics

The mortality rate in human populations increases exponentially with age for a significant part of the lifespan, starting from the period of reproductive maturity (age ~35) up to extreme old ages (age ~100), satisfying the Gompertz law of mortality (Gompertz, 1825). Mathematically, the Gompertz law is expressed as

$$m_i = m_0 e^{\beta i}, \tag{1.3}$$

where parameter m_0 can be considered as initial mortality at age $i = 0$ and parameter β defines the rate of change of mortality with age (usually called rate of ageing, mortality coefficient or Gompertz slope).

Graphically, mortality data are most frequently presented in semi-logarithmic graphs (logarithm of mortality versus age) and therefore their exponential increase with age, as expressed by the Gompertz law (equation (1.3)), is represented with a straight line. Figure 1.2 shows period mortality rates of the 2010 Swedish population and a Gompertz line (solid line) showing their exponential increase over a big portion of lifespan. Although the exponential growth of mortality dominates across a wide range of ages, some deviations from it are observed at young (before 35) and extremely old (after 100) ages.
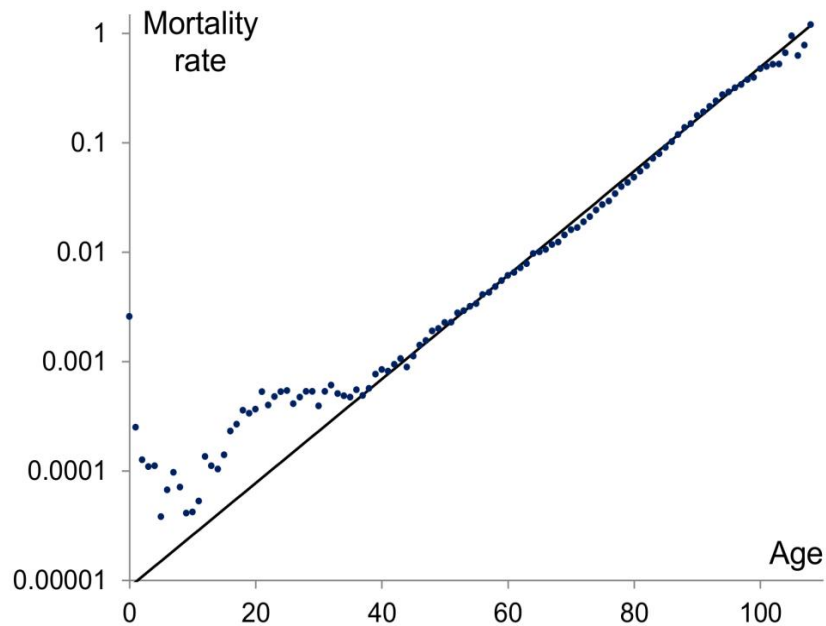
Figure 1.2: The Gompertz law of mortality. *The mortality rate increases exponentially from sexual maturity to extremely old ages. The presented data are the mortality rates of the 2010-period Swedish population set in a semi-logarithmic scale. The data are taken from the Human Mortality Database (http://www.mortality.org). The Gompertz function with parameters $m_0 = 8.7 \cdot 10^{-6}$ and $\beta = 0.109$ fits the data very well after the age of 35. Deviations from the exponential growth are observed at young (before 35) and considerably old (after 100) ages.*

Upon analysis of various causes of death, Gompertz concluded that if mortality rate is taken for only intrinsic causes of death, the exponential law is valid between ages 20 and 80 for humans. The Gompertz law has been shown to be a fundamental mortality law and verified by demographic observations across different countries, different time periods, and even different species (Gavrilov and Gavrilova, 1991). Analysis of available data on mortality rates for various diseases indicates that for most diseases there is a considerably wide age range where the mortality rate is also increasing exponentially (Jones, 1956, Finch, 1994).

The deviation of mortality from the exponential increase at very old ages is a phenomenon called "late-life mortality slow-down". Late-life mortality slow-down corresponds to the deceleration of mortality increase so that it happens at a slower pace compared to the exponential growth observed during adulthood or even saturates to a constant rate (the phenomenon known as mortality level-off or mortality plateau).

Late-life mortality deceleration is observed in human populations (Greenwood and Irwin, 1939, Manton et al., 2008) and other non-human species such as invertebrates; for

example nematodes, flies, beetles, etc. (Economos, 1979, Curtsinger et al., 1992, Carey et al., 1992, Vaupel et al., 1998, Tatar et al., 1993) and in other mammalian populations such as rats, mice, horses, sheep, etc. (Lindop, 1961, Economos, 1979, Economos, 1980, Sacher, 1966). However this phenomenon is not considered as a universal law of mortality and is still debated by scientists. For example, Gavrilov and Gavrilova (Gavrilova and Gavrilov, 2015) have compared the goodness of fitting exponential and logistic functions to mortality data at advanced ages for humans, laboratory mice and laboratory rats and they found that the Gompertz model fits the data in all these three species, significantly better than the logistic model. They concluded that the observations of mortality deceleration or levelling-off, which is described by the logistic function, could be invalid due to various reasons including the inaccuracy of data at very old ages.

Another observation related to the dynamics of mortality is a phenomenon known as the compensation law of mortality or the compensation effect. This law states that high initial mortality (parameter m_0 in Gompertz law) in a population is compensated by a low mortality coefficient (parameter β in Gompertz law) and respectively low initial mortality is compensated by a high mortality coefficient (Gavrilov and Gavrilova, 1979, Gavrilov et al., 1978, Gavrilov, 1984). This phenomenon was observed not only in human populations but also in other species such as fruit flies (Gavrilov and Gavrilova, 1991, Gavrilov and Gavrilova, 2006). The compensation law points to an inverse relationship between the Gompertz parameters and is expressed mathematically by the equation

$$\ln(m_0) = \ln(M) - \beta X, \qquad (1.4)$$

where X is the target lifespan (the age when the last remaining survivors die) and M is the target mortality (the level of mortality when the last survivors die). The inverse relationship between Gompertz parameters (equation (1.4)) was observed first by Strehler and Mildvan and is therefore called the Strehler-Mildvan correlation (Strehler and Mildvan, 1960, Strehler, 1978).

The compensation law of mortality in its strongest form refers to the convergence of mortality trajectories from different populations of a given species at a certain point having coordinates the target lifespan and the target mortality. This late-life mortality convergence is illustrated by the convergence of mortality trajectories in a single point at

6

old ages, obtained by fitting the Gompertz function to mortality rates of different populations (for example data from different countries) and plotting the trajectories onto the same semi-logarithmic graph. The compensation effect is also observed on mortality trajectories for data on the same country but for different periods (see Chapter 3).

The exponential increase of mortality across a wide range of lifespan (Gompertz law), the late-life mortality deceleration and the compensation effect together compose the so-called "*three laws of mortality*". The explanation for the existence and validation of these laws is the main goal of most of the developed theories of ageing and longevity (Gavrilov and Gavrilova, 2001). Furthermore, there are many other interesting observations based on the evolution of mortality over time. These include the compression of mortality and the rectangularization of the survival curve, both of which are described in the following section.

1.1.3. Observations on the evolution of mortality over time

The evolution of mortality over time and primarily the reduction of mortality rate in all age-intervals is the result of several factors such as medical improvements, health and safety, hygiene developments and many others. The mortality evolution can be expressed in terms of mortality compression, rectangularization of survival curve, compensation effect, increase in life expectancy and others. The evolution of the mortality process was also analysed through Epidemiological Transition Theory which explains how some phases of historical changes caused the transition of birth and death rates to different levels and influenced the growth of human populations (Omran, 2005, Olshansky and Ault, 1986, Rogers and Hackenberg, 1987, Robine, 2001).

In this section we present two observations on the evolution of mortality, the compression of mortality and the rectangularization of the survival curve. Mortality compression refers to the decrease in variance of the age-related distribution of deaths over time. Due to the reduction of mortality over time, a larger number of deaths occurs at later ages and these deaths are concentrated in a certain age-range. The phenomenon of mortality compression is shown in Figure 1.3A. The blue and red colours denote the distribution of deaths over ages as taken from the 1900 and 2000 Swedish period life tables respectively. It is observed that over a period of 100 years the number of deaths for ages below age 65 is significantly reduced. These deaths were redistributed across the lifespan and concentrated more at older ages (for the 2000 data this concentration is

around the age of 85 as shown by the peak of the red curve in Figure 1.3A). The age where the local maximum number of deaths among adults occurs is called the modal age at adult death (note that another local maximum exists at age 0) (Canudas-Romo, 2008, Cheung et al., 2005). Through time, the compression of deaths occurs in narrower age-ranges.
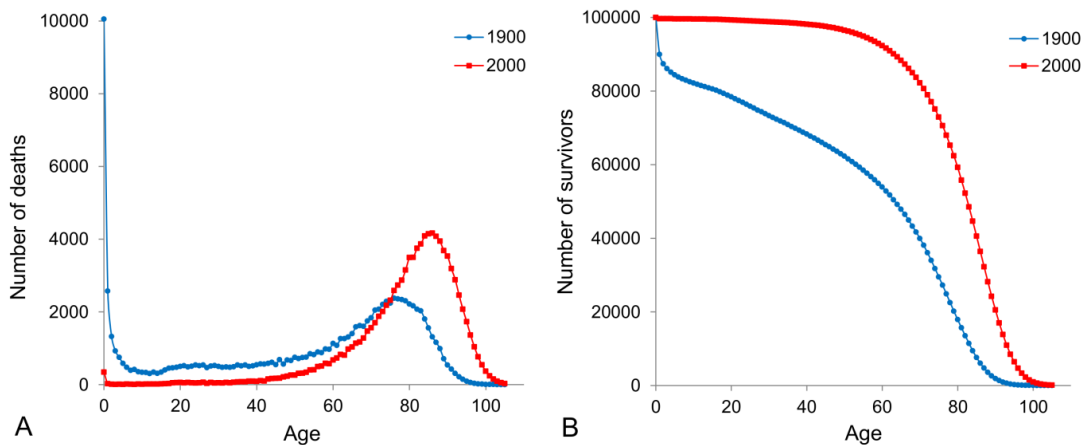
Figure 1.3: Compression of mortality and rectangularization of the survival curve. *Panel A demonstrates the time-evolution of the number of deaths. It is observed that in over long period of time an increasing number of deaths occurs at older ages and in a narrower age-range. Panel B shows the time-evolution of the number of survivors. It is observed that due to the concentration of deaths at older ages, the survival curve becomes more rectangular in shape over time. In both panels the blue and red colours denote the data from the 1900 and 2000 Swedish period life tables respectively. The data are taken from the Human Mortality Database (http://www.mortality.org).*

The compression of mortality over time results in the "rectangularization" of the survival curve. Rectangularization refers to the evolution of the survival curve towards a more rectangular shape due to an increase in survival across the first half of the lifespan and to a concentration of deaths around the modal age of death that is shifted to older ages as time increases (i.e. more individuals survive over childhood and young-adulthood periods but then rapidly die at older ages) (Fries, 1980, Manton and Tolley, 1991, Kannisto, 2001, Weon and Je, 2011). The rectangularization of the survival curve is illustrated in Figure 1.3B. The data represented in blue and red colours are numbers of survivors over ages as taken from the 1900 and 2000 Swedish period life tables respectively.

One of the interesting questions associated with time dependent mortality-related phenomena refers to the existence of an upper limit in human lifespan. In fact, longevity and life expectancy are increasing over time and it is unknown whether a biologically-

related maximum length of life exists (Wilmoth et al., 2000). The existence of a limit in lifespan can be partially explained by a phenomenon known as the Hayflick limit (Hayflick and Moorhead, 1961, Hayflick, 1965) which gives the number of times that a population of human cells proliferates until the end of the cells' division process. During the process of DNA replication which occurs in each cell division, the length of each telomere (the region at the end of each chromosome's chromatid) decreases. When the telomere disappears (or becomes shorter than a certain critical length), the cell cannot divide anymore; this halts the organism's ability to repair tissues bringing it to the death.

Today there is no evidence for a fixed value of a theoretical limit of human lifespan. Some studies state that there is no increase in the maximum limit of human lifespan over time (Olshansky et al., 1990) while others indicate that the length of lifespan is increasing over time and this indicates the reduction of mortality rate and the postponement of senescence (Wilmoth and Lundstrom, 1996). Some other studies use mathematical models to estimate the theoretical maximum lifespan. For example, Weon and Je (Weon and Je, 2009) used a modified Weibull function to model the survival probability and to estimate the maximum human lifespan which they found to be around 125 years.

Studies that aim to analyse human mortality dynamics over the lifespan and their evolution over time are of great importance for many reasons including an understanding of the mechanisms underlying ageing and the development of ways to control and extend the duration of lifespan. Mathematical modelling makes a significant contribution to these studies. Various mathematical techniques are used to model processes associated with ageing and senescence and validate mortality laws and other observations.

A number of theories have been developed in attempts to relate the process of ageing to the dynamics of some (hypothetical) physiological functions. Alternative theories take into account facts such as heterogeneity, stochasticity and homeostasis to model and explain the observations of age-related mortality trajectories (Yashin et al., 2000). Several mathematical models have been developed and used to analyse human mortality dynamics across the lifespan and to explain the deviations in mortality from the exponential growth at young and old ages. For example, the proposed explanations for the late-life mortality plateau include an assumption that the Gompertz law is not valid at those ages and that mortality dynamics should be described by a logistic, a quadratic or

some other mathematical functions (Gavrilov and Gavrilova, 2001, Kannisto et al., 1994, Pham, 2011). In addition, population heterogeneity is a commonly used consideration in modelling mortality dynamics and explaining the deviation of mortality from its exponential increase at young and extremely old ages (Wrigley-Field, 2014, Chen et al., 2013, Drapeau et al., 2000, Steinsaltz, 2005). An overview of mathematical models used to reproduce the dynamics of mortality is presented in Section 1.2. Section 1.2 also introduces mathematical models incorporating the concept of heterogeneity and its effect on mortality dynamics. Section 1.3 provides the aims of our research, which are focused on the modelling and analysis of mortality dynamics in heterogeneous populations, and presents our methodology which includes the fitting procedure and the selection criterion used in this study.

1.2. Mathematical models of mortality dynamics

Although many mortality and survival models have been developed over the last few decades, we should remember that a number of pioneering works were performed in previous centuries. Even before the great work by Gompertz, many scientists and philosophers were trying to investigate the dynamics underlying mortality. Amongst others, John Graunt (1620-1674) was one of the first demographers who, in 1662, analysed the mortality of the population of London. Edmond Halley (1656-1742) was a famous astronomer who, in 1693, published an article on life annuities that included the first known mortality table and had an important influence on the actuarial sciences and mortality statistics. Abraham de Moivre (1667-1754) was a mathematician who, in 1725, proposed the first known model that expresses the number of survivors as a function of age. In addition, works by Antoine Deparcieux (1703-1768), Daniel Bernoulli (1700-1782), Thomas Robert Malthus (1766-1834) and Pierre François Verhulst (1804-1849) were important to the development of population-related studies. However, after the work by Gompertz, a new era on the analysis and modelling of mortality dynamics began. Some reviews and historical overviews on the earliest and most recent mortality models can be found in (Haberman and Sibbett, 1995, Tabeau et al., 2001, Keyfitz, 1982).

Most models of mortality are represented by parameterised functions that attain to reproduce the entire mortality patterns or parts of them. A typical example is the Heligman-Pollard model, an eight parameter function, which accurately reproduces an

entire mortality pattern. This model is composed of three terms, each of which is able to create a specific part of a mortality curve. The first term forms the sharp initial mortality decline at infant ages, the second forms the accidental hump and the third term forms the exponential increase in mortality after sexual maturity. Another example is the Perks model which is a logistic function that can reproduce a wide-range of a mortality pattern, namely, it reproduces fairly well the exponential increase of mortality at adulthood and the mortality plateau at very old ages. In Section 1.2.1 we present a list of commonly used age-dependent parametric mortality models. In Section 1.2.2 we outline a few models of mortality that include both age- and time-dependent components reflecting the cohort and period effects respectively and in Section 1.2.3 we introduce the models of mortality that incorporate the concept of population heterogeneity.

1.2.1. Modelling age-dependent mortality

The term "parametric models" refers to models which express mortality-related quantities such as the mortality rate, the force of mortality or the number of survivors as functions of age. Consequently, these models describe the age-specific mortality rates of certain cohorts or certain periods. Table 1.1 is adapted from (Tabeau et al., 2001) and presents the main age-dependent mortality models. The notations that are used in Table 1.1 are the common demographic notations for the variables used in life tables. Hence $\mu(x)$ is the force of mortality at age x, $m(x)$ is the mortality rate at age x, $s(x)$ (equivalent with $l(x)$) is the probability of survival from birth until age x, $q(x)$ and $p(x) = 1 - q(x)$ are the probabilities of dying and surviving respectively between ages x to $x + 1$, $e(x)$ is the expectation of life at age x, ω is the highest attainable age and all other symbols are unknown model parameters.

As shown in Table 1.1, some of the recently developed mortality models represent extensions of older ones and many of them are based on, or related to the Gompertz law. The models in Table 1.1 are separated into groups which are expressed by polynomial and non-polynomial functions. Polynomials often appear in the exponentiation. Many functions can be approximated by polynomials using Taylor expansions and this gives them an advantage in modelling mortality-related data. On the other hand, non-polynomial functions include a set of components, each of which can account for a different cause of death or process underlying mortality, and is able to reproduce a specific part of mortality patterns over a certain age-interval of lifespan. For example

Siler, Thiele, Heligman-Pollard, Rogers-Plank and Kostaki models are represented by functions composed by three terms. In each of these models the terms are different but in all of them the first component describes the decline of mortality in childhood, the second describes the accidental mortality during young-adulthood and the third describes the exponential mortality dynamics at post-reproductive ages.

Author	Publication	Model
OLD NON-POLYNOMIALS		
De Moivre	1725	$\mu(x) = \dfrac{1}{(\omega - x)}$
Gompertz	1825	$\mu(x) = Bc^x$
Makeham	1860	$\mu(x) = A + Bc^x$
		$\mu(x) = \alpha + \gamma x + \beta c^x$
Opperman	1870	$\mu(x) = \dfrac{a}{\sqrt{x}} + b + c\sqrt{x}$
Thiele	1872	$\mu(x) = a_1 e^{-b_1 x} + a_2 e^{-\frac{1}{2}b_2(x-c)^2} + a_3 e^{b_3 x}$
Wittstein	1883	$q(x) = \dfrac{1}{m} a^{-(mx)^n} + a^{-(M-x)^n}$
Steffenson	1930	$\log_{10} s(x) = 10^{-A\sqrt{x}-B} + C$
		$e(x) = \dfrac{1}{A + Bc^x}$
Perks	1932	$\mu(x) = \dfrac{A + Bc^x}{kc^{-x} + 1 + Dc^x}$
Harper	1936	$\log_{10} s(x) = A + 10^{B\sqrt{x}+Cx+D}$
Weibull	1939	$\mu(x) = \alpha x^{\beta - 1}$
Van der Maen	1943	$\mu(x) = A + Bx + Cx^2 + \dfrac{1}{N - x}$
		$\mu(x) = A + Bc^x + \dfrac{c}{N - x}$
POLYNOMIALS		
Unnamed (in Keyfitz, 1982)		$y(x) = e^{(a_0 + a_1 x + a_2 x^2 + \cdots + a_k x^k)}$
		$y(x) = q(x), \dfrac{q(x)}{p(x)}, \mu(x) \text{ or } e(x)$

RECENT NON-POLYNOMIALS

Brillinger	1960	$\mu(x) = \sum_i \left(H_i(x-B_i)^{C_i-1} + \dfrac{A_i}{(b_i-x)^{C_i+1}} + E_i d_i^x \right)$
Beard	1961	$\mu(x) = \dfrac{Be^{ax}}{1+De^{ax}}$
Petrioli	1981	$s(x) = \dfrac{1}{x^a(\omega-x)^{-b}e^{\frac{c}{2}x^2+dx}\frac{1}{k}+1}$
Martinelle	1987	$\mu(x) = \dfrac{A+Be^{kx}}{1+De^{kx}} + ce^{kx}$
British actuaries (in Keyfitz, 1982)	1980s	$\dfrac{q(x)}{p(x)} = A - Hx + bc^x$

RECENT NON-POLYNOMIAL

Siller	1979	$\mu(x) = a_1 e^{-b_1 x} + a_2 + a_3 e^{b_3 x}$
Heligman-Pollard	1980	$\dfrac{q(x)}{p(x)} = A^{(x+B)^C} + De^{-E(\ln x - \ln F)^2} + GH^x$
		$q(x) = A^{(x+B)^C} + De^{-E(\ln x - \ln F)^2} + \dfrac{GH^x}{1+GH^x}$
		$q(x) = A^{(x+B)^C} + De^{-E(\ln x - \ln F)^2} + \dfrac{GH^x}{1+KGH^x}$
		$q(x) = A^{(x+B)^C} + De^{-E(\ln x - \ln F)^2} + \dfrac{GH^{x^K}}{1+GH^{x^K}}$
Brooks et al.	1980	$\mu(x) = \mu_I(x) + \mu_A(x) + \mu_S(x)$
		$\begin{cases} \mu_I(x) = \begin{cases} Q_0 & \text{for } x=0 \\ Q_1^{x^\gamma} & \text{for } x>0 \end{cases} \\[2em] \mu_A(x) = Q_A e^{\frac{(\ln x - \ln x_A)^2}{\delta^2}} \quad \text{for } x \geq 0 \\[2em] \mu_S(x) = \dfrac{Q_S e^{\frac{x}{x_S}}}{1+Q_S e^{\frac{x}{x_S}}} \quad \text{for } x \geq 0 \end{cases}$
Rogers and Planck	1983	$q(x) = A_0 + A_1 e^{-a_1 x} + A_2 e^{-a_2(x-\mu_2)-e^{-\lambda_2(x-\mu_2)}} + A_3 e^{a_3 x}$
Kostaki	1992	$\dfrac{q(x)}{p(x)} = \begin{cases} A^{(x+B)^C} + De^{-E_1^2(\log x/F)^2} + GH^x, & x \leq F \\ A^{(x+B)^C} + De^{-E_2^2(\log x/F)^2} + GH^x, & x > F \end{cases}$

| Rogers and Little | 1993 | $y(x) = a_0 + m_1(x) + m_2(x) + m_3(x) + m_4(x)$ |

where:

$$m_1(x) = a_1 \exp(-a_1 x)$$

$$m_2(x) = a_2 \exp(-a_2(x - \mu_2) - \exp(-\lambda_2(x - \mu_2)))$$

$$m_3(x) = a_3 \exp(-a_3(x - \mu_3) - \exp(-\lambda_3(x - \mu_3)))$$

$$m_4(x) = a_4 \exp(a_4 x)$$

$$y(x) = q(x), \frac{q(x)}{p(x)}, \mu(x)$$

RECENT NON-POLYNOMIAL, PARTIAL AGE FUNCTIONS

| Hartmann | 1981 | ages 0 to 15 years: |

$$y(x) = A_1 + B_1 \ln x$$

ages 15 to 35 years:

$$y(x) = A_2 + B_2 x$$

ages 35 to 60 years:

$$y(x) = A_3 + B_3 c^x,$$

where $y(x) -$ logit of $l(x)$

| Mode and Busby | 1982 | ages 0 to 10 years: |

$$\mu_0(x) = \alpha_0 \beta_0 e^{-\beta_0 x}$$

ages 10 to 30 years:

$$\mu_1(x) = \alpha_1 - \beta_1(x - \gamma_1)^2$$

ages 30 and over:

$$\mu_2(x) = \alpha_2 + \beta_2 \gamma_2 e^{\gamma_2 x}$$

Table 1.1: Main parametric functions of mortality. The table is taken from (Tabeau et al., 2001) where standard demographic notations are used. $\mu(x)$ is the force of mortality at age x, $m(x)$ is the mortality rate at age x, $s(x)$ (equivalent with $l(x)$) is the probability of survival from birth until age x, $q(x)$ is the probability of dying between ages x to $x + 1$ and $p(x)$ is the probability of surviving between this interval ($p(x) = 1 - q(x)$), $e(x)$ is the expectation of life at age x, ω is the highest attainable age and all other symbols are the unknown model parameters.

Some of the functions presented above are used in Chapter 4 where we compare different models by analysis of their fit to mortality data on the entire lifespan and on very old ages. In addition to the age-dependent mortality models, several others have been developed to express mortality as function of age and time. The age and time dependent models of mortality are described in the next section.

1.2.2. Modelling age-time-dependent mortality

"Age-period-cohort" is the name of a class of models that express the mortality rate at a specific age and specific time. In these models, mortality is determined by age, date of follow-up (period) and date of birth (cohort) (Hobcraft et al., 1982, Carstensen, 2007). The age-period-cohort models are mostly used to forecast mortality rates over a certain range of ages and over a certain period of time. The general mathematical form of age-period-cohort models is given by:

$$\eta(m_{x,t}) = \alpha_x + \sum_{i=1}^{N} f^{(i)}(x, \theta_i)k_t^{(i)} + \gamma_{t-x},$$ (1.5)

where η is the link function that transforms the mortality rate $m_{x,t}$ (or other mortality-related measurement such as the probability of dying, $q_{x,t}$) into a suitable form (usually logarithmic or logistic transformations are used), α_x is an age-related function which does not depend on time and specifies the shape of the mortality curve (age effects), $k_t^{(i)}$ are period-trends that express the change in mortality curves over time (period effects), $f^{(i)}(x, \theta_i)$ are age-dependent functions that determine the age-range within each trend i affects, and γ_{t-x} is a function that varies between different cohorts (cohort effects) (Hunt and Blake, 2014).

A simple example using the formulation of equation (1.5) is the Lee-Carter model (Lee and Carter, 1992) which excludes the cohort-related component (γ_{t-x}) but is the most commonly used model for mortality forecasting purposes. Moreover, the models developed by Renshaw and Haberman (Renshaw and Haberman, 2003), Cairns et al. (Cairns et al., 2006, Cairns et al., 2011, Cairns et al., 2009), Yang et al. (Yang et al., 2010) and several others, belong to the age-period-cohort class of mortality models. Table 1.2 presents eight special cases of mortality models that belong to this class as described in (Cairns et al., 2009).

Mathematical models that express mortality as a function of age, or those that take into account both the age and time dependency of mortality, can use various assumptions to model out all possible observed features of mortality patterns. One of the more commonly used considerations is that the population is heterogeneous and composed of several subpopulations having different mortality dynamics (Rossolini and Piantanelli, 2001, Vaupel, 2010). This consideration is made on the basis of research presented here

which is focused on modelling mortality dynamics and analysing their evolution over time (Chapters 2-4). An overview of mortality models that incorporate heterogeneity is presented in Section 1.2.3.

Author	Publication	Model
Lee and Carter	1992	$\log(m(t,x)) = \beta_x^{(1)} + \beta_x^{(2)} k_t^{(2)}$
Renshaw and Haberman	2006	$\log(m(t,x)) = \beta_x^{(1)} + \beta_x^{(2)} k_t^{(2)} + \beta_x^{(3)} \gamma_{t-x}^{(3)}$
Currie	2006	$\log(m(t,x)) = \beta_x^{(1)} + n_a^{-1} k_t^{(2)} + n_a^{-1} \gamma_{t-x}^{(3)}$
Currie, Durban and Eilers	2004	$\log(m(t,x)) = \sum_{i,j} \theta_{ij} B_{ij}^{ay}(x,t)$
Cairns, Blake and Dowd	2006	$\text{logit}(q(t,x)) = k_t^{(1)} + k_t^{(2)}(x - \bar{x})$
Cairns et al.	2009	$\text{logit}(q(t,x)) = k_t^{(1)} + k_t^{(2)}(x - \bar{x}) + \gamma_{t-x}^{(3)}$
Cairns et al.	2009	$\text{logit}(q(t,x)) = k_t^{(1)} + k_t^{(2)}(x - \bar{x}) + k_t^{(3)}((x - \bar{x})^2 - \hat{\sigma}_x^2) + \gamma_{t-x}^{(4)}$
Cairns et al.	2009	$\text{logit}(q(t,x)) = k_t^{(1)} + k_t^{(2)}(x - \bar{x}) + \gamma_{t-x}^{(3)}(x_c - x)$

Table 1.2: Age-period-cohort mortality models. *The table is taken from (Cairns et al., 2009). The functions $\beta_x^{(i)}$, $k_t^{(i)}$ and $\gamma_{t-x}^{(i)}$ represent the age, period and cohort effects respectively. The n_a is the number of ages, $B_{ij}^{ay}(x,t)$ are B-splines with θ_{ij} weights, $\bar{x}$ is the mean age of analyzed age-range and $\hat{\sigma}_x^2$ is the mean value of $(x - \bar{x})^2$.*

1.2.3. Modelling mortality incorporating heterogeneity

Heterogeneity is an important characteristic of populations that should be considered to explain certain observations in mortality dynamics. Heterogeneity has an impact on the dynamics of population's mortality, as individuals of the same age have different biological and physiological characteristics and they face different mortality dynamics even if they belong to the same cohort of a population. The consideration of heterogeneity is made in two different ways as related to the composition of a population. The first way is based on an assumption that the population is composed of a number of homogeneous subpopulations so that within each subpopulation the individuals are identical and follow the same mortality dynamics. This approach is used

16

in this thesis (Chapters 2-5) where we develop a model for the analysis of mortality in heterogeneous human populations. The second way is based on the consideration that each single individual in a population has their own specific traits and faces certain mortality dynamics. In this case, the force of mortality acting upon each individual has a cumulative effect on the mortality process in the entire population.

Heterogeneity reflects many observed and unobserved characteristics. Observed characteristics include the life-style and environmental conditions, occupation, sex and many others, while unobserved characteristics mainly refer to genetic differences. A theory of mortality and ageing based on genetic heterogeneity was developed by Szilard (Szilard, 1959) who stated that the differences in the duration of individuals' lifespans are mainly due to the different number of faults they have inherited. The consideration of genetic heterogeneity was also used by Beard (Beard, 1959) who developed a mortality model of heterogeneous populations which was expressed as a logistic curve. In addition, Strehler and Mildvan (Strehler and Mildvan, 1960) mentioned the existence of heterogeneity in populations and used it to explain the deviation in mortality from the exponential growth at very old ages. Other remarkable works were performed by Vaupel et al. (Vaupel et al., 1979), who defined individual frailty as a measure of individual differences in the chances of survival and proposed a model of mortality for heterogeneous populations, and by Vaupel and Yashin (Vaupel and Yashin, 1985a, Vaupel and Yashin, 1985b), who analysed the effects of heterogeneity on the dynamics of mortality including the explanation for late-life mortality deceleration and levelling-off. There are many other interesting works which use heterogeneity to analyse human mortality (Shepard and Zeckhauser, 1980, Manton et al., 1981, Woodbury and Manton, 1977, Keyfitz and Littman, 1979).

A typical model of mortality, which takes into account the heterogeneity among individuals, includes a random variable which is called frailty and reflects individual unobserved heterogeneity on survival. The force of mortality of an individual aged x as affected by its frailty Z can be represented as:

$$\mu(x, Z) = Z\mu_0(x), \tag{1.6}$$

where $\mu_0(x)$ is a pre-fixed function that describes underlying mortality. In this formulation it is assumed that each individual is born at a specific frailty level and remains at that level across their entire lifespan (Vaupel et al., 1979). In that case, frailty

is called fixed or unchanging. In (Vaupel et al., 1979) it is assumed that frailty follows a gamma-distribution with mean 1 and variance σ^2. The gamma distribution is used because frailty cannot be negative and this distribution takes only positive values.

In the fixed frailty model, the average frailty of the population at age x is given by

$$\bar{Z}(x) = \frac{1}{1 + \sigma^2 H(x)},$$ (1.7)

where $H(x) = \int_0^x \mu(u)du$ is the cumulative hazard function. If the underlying mortality is described by the Gompertz model, $\mu_0(x) = ae^{bx}$, the average frailty is

$$\bar{Z}(x) = \left(1 + \sigma^2 \frac{a}{b}(e^{bx} - 1)\right)^{-1},$$ (1.8)

and the average mortality at age x is given by

$$\bar{\mu}(x) = ae^{bx}\left(1 + \sigma^2 \frac{a}{b}(e^{bx} - 1)\right)^{-1}.$$ (1.9)

The average mortality of surviving individuals at age x is equal to the mortality of the entire heterogeneous population that the individuals belong to. This model can be used to show that individuals' mortality could increase faster than the population's mortality. Indeed, as frail individuals with high values of frailty, die at young ages, the average frailty of the entire population decreases while age increases.

The fixed frailty model (equation (1.9)) expresses the mortality of a population as a logistic function of age. The same function can be derived in a model of changing frailty (also known as acquired heterogeneity or debilitation) (Yashin et al., 1994). The result that the same mortality trajectory can be reproduced either by the fixed-frailty or changing-frailty model, indicates that the use of survival data alone is not enough to distinguish between different mechanisms that are able to generate the actual mortality patterns (Yashin et al., 2000) and that the modelling of frailty requires more complex considerations. The latter is indeed true, as in reality frailty is never fixed across the lifespan and also changes in heterogeneity are not always debilitative (for example if an individual has an increased probability of survival after recovering from a disease) (Yashin et al., 2000).

18

The second approach in modelling mortality by incorporating heterogeneity is based on the decomposition of a population into subgroups of individuals. The individuals in each subgroup are considered to be identical in terms of their exposure to mortality dynamics. Thus, a heterogeneous population can be seen as a composition of a number of homogeneous subpopulations. To examine the patterns of mortality in a simple case of a heterogeneous population composed by only two subpopulations, Vaupel and Yashin (Vaupel and Yashin, 1985b) used a mathematical model that expresses the mortality of the entire population, $\bar{\mu}(x)$, as a function of mortality rates, $\mu_1(x)$ and $\mu_2(x)$, of the two subpopulations:

$$\bar{\mu}(x) = \pi(x)\mu_1(x) + \left(1 - \pi(x)\right)\mu_2(x), \qquad (1.10)$$

where $\pi(x) = \pi(0)p_1(x)/\left[\pi(0)p_1(x) + \left(1 - \pi(0)\right)p_2(x)\right]$ is the proportion of survivors at age x of the first subpopulation to the entire population and $p_j(x) = \exp\left(-\int_0^x \mu(u)du\right)$ for $j = 1, 2$ are the survival functions for the two subpopulations. In their work, Vaupel and Yashin have shown how different age-dependent functions of mortality rates for the two subpopulations (different functions for $\mu_1(x)$ and $\mu_2(x)$) affect the mortality trajectory of the entire population. Formulation similar to that expressed by equation (1.10) is used in our research as denoted in our motivation (following section) and described in detail in Chapters 2, 3 and 4.

1.3. Motivation and methodology

1.3.1. Motivation

This study is focused on mathematical modelling of mortality dynamics in human populations. The model that we propose is based on two main assumptions. The first assumption is that each population is heterogeneous and consists of a number of distinct subpopulations. The factor that distinguishes our model from other existing models of heterogeneous populations is our second assumption which states that the mortality dynamics in each constituent subpopulation increases exponentially with age in a way similar to what is described by the Gompertz law. In other words, our model has the same formulation as equation (1.10) but in the general case of a population composed by n subpopulations where the mortality rate $\mu_j(x)$, for each subpopulation $j = 1, \cdots, n$, is described by the exponential function.

The model that we developed is described in Chapter 2 and is used for the analysis of mortality dynamics across the lifespan and their evolution over time as presented in Chapters 2, 3 and 4. We should point out that the existence of the subpopulations in our study is highly hypothetical; the subpopulations are introduced as a convenient mathematical constraint which allows very accurate reproduction of the entire mortality patterns, gives explanations for the deviations of mortality from the exponential increase at young and very old ages and for the decrease in mortality over time. In Chapter 4 we put forward a few arguments to say that this model may reflect the real heterogeneous structure of a population and the differences between subpopulations can be due to genetic variation. Particularly, we show in Chapter 4 that evolutionary forces (i.e. natural selection) can explain the homogenisation of the Swedish population during the 20th century as predicted by the model.

However, the nature of those subpopulations is not entirely verified in this thesis and the exact factors that identify the mortality-related differences between real subgroups of individuals are not yet clear. Alongside genetic variations, the differences between subpopulations could eventually reflect environmental and socio-economic disparities, life-style conditions and many other factors. We aim to further investigate the nature of differences between those subpopulations in future work (see also other future plans in Chapter 5). Nevertheless, for the purposes of this study, the consideration of heterogeneity as a convenient constraint allows the development of an extremely accurate (in terms of the goodness-of-fit) model and the implementation of a comprehensive analysis of mortality dynamics explaining its main features related to age- and time- dependencies.

While the model includes parameters with demographic interpretations, we analyse the values of these parameters, as obtained by fitting the model to the observation data, and conclude with three main outcomes. The first outcome is related to the effect of heterogeneity on the dynamics of mortality across the lifespan which is examined in terms of the variation of model parameters (Chapter 2). The second outcome refers to the influence of heterogeneity on the evolution of mortality over time as described by the alteration of model parameters (Chapter 3). The third outcome is the justification of mortality-related homogenisation of the population during the 20th century and is derived by relating the existence of population heterogeneity with genetic variation between different subpopulations (Chapter 4). To perform this research and to derive the values

of parameters in the best fit model we have applied a specific fitting procedure and used a certain selection criterion as described in Sections 1.3.2 and 1.3.3 respectively.

1.3.2. Fitting procedure

In this research, we apply non-linear regression analysis to fit the model that we develop in Chapter 2, to mortality data. The Least Squares (LS) method is used to estimate the free (unknown) parameters that minimize the sum of squared residuals (a residual is the difference between the theoretical and the observed value). Linear and non-linear regressions have been shown to be less reliable than the Maximum Likelihood (ML) method in estimating the unknown parameters when simple mortality models are used (see (Wilson, 1993, Promislow et al., 1999, Yen et al., 2008)). Indeed the LS method overestimates mortality at early life and underestimates it at old ages when simple functions such as the Gompertz, Makeham or Weibull are used to fit the mortality rates. In contrast, when a complex nonlinear model fits very precisely an entire data set, the LS method is favoured over ML because it does not require the use of numerical methods for the partial differentiation of the function with respect to each unknown parameter.

In addition, both LS and ML methods have many similarities (Pletcher, 1999). When the errors (residuals) are independent and identically distributed, the linear and nonlinear regressions give estimations for the unknown parameters identical to those given by the maximum likelihood method (Burnham and Anderson, 2003). To see this link between ML and LS methods let us consider a sample of n variables, $(x_1, \dots, x_n)$, where each one follows a normal distribution with mean μ and variance σ^2. Then, the probability density of a variable x_j from the considered sample, is given by

$$f(x_j) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(x_j - \mu)^2 / 2\sigma^2}. \tag{1.11}$$

The likelihood function of the probability distribution is expressed as:

$$L(x) = \prod_{j=1}^{n} \frac{1}{\sqrt{2\pi}\sigma} e^{-(x_j - \mu)^2 / 2\sigma^2} = \left(\frac{1}{\sqrt{2\pi}\sigma}\right)^n e^{-\sum_{j=1}^{n}(x_j - \mu)^2 / 2\sigma^2}, \tag{1.12}$$

and the log-likelihood by

$$l(x) = \ln\big(L(x)\big) = n \ln \left(\frac{1}{\sqrt{2\pi}} \right) - n \ln(\sigma) - \frac{1}{2\sigma^2} \sum_{j=1}^{n} (x_j - \mu)^2. \qquad (1.13)$$

The maximum likelihood estimators are found by partial differentiation of the log-likelihood function with respect to each parameter. Consequently, the estimator of the normal distribution's mean is derived as:

$$0 = \frac{\partial l(x)}{\partial \mu} = \frac{1}{\sigma^2} \sum_{j=1}^{n} (x_j - \mu) \Rightarrow \hat{\mu} = \frac{1}{n} \sum_{j=1}^{n} x_j = \bar{x}. \qquad (1.14)$$

The estimation of mean is substituted into the log-likelihood function (equation (1.13)) and the partial differentiation respect to σ gives the estimator for the variance:

$$0 = \frac{\partial l(x)}{\partial \sigma} = -\frac{n}{\sigma} + \frac{1}{\sigma^3} \sum_{j=1}^{n} (x_j - \bar{x})^2 \Rightarrow \hat{\sigma}^2 = \frac{1}{n} \sum_{j=1}^{n} (x_j - \bar{x})^2. \qquad (1.15)$$

This shows that the estimators of the mean and the variance of the normal distribution as computed through ML are the same as computed by the LS method. In addition, the use of the normal distribution gives another important outcome. Using the above parameters' estimators, the maximum log-likelihood is expressed as

$$\ln(\hat{L}(x)) = n \ln \left(\frac{1}{\sqrt{2\pi}} \right) - n \ln(\hat{\sigma}) - \frac{1}{2\hat{\sigma}^2} \sum_{j=1}^{n} (x_j - \bar{x})^2 \qquad (1.16)$$

or

$$\ln\left(\hat{L}(x) \right) = -\frac{n}{2} \ln(\hat{\sigma}^2) - \frac{n}{2} \ln(2\pi) - \frac{n}{2}. \qquad (1.17)$$

The additive constants (terms which do not depend on the unknown parameters) are often removed from the log-likelihood as they do not contribute to its maximization. As a result, the maximized value of the log-likelihood function can be expressed as:

$$\ln\left(\hat{L}(x) \right) \approx -\frac{n}{2} \ln(\hat{\sigma}^2). \qquad (1.18)$$

This is a key point for the Bayesian Information Criterion (described in the next section) because it allows a simple mapping from the result of Least Squares analysis into the maximized value of the log-likelihood function.

Taking into account that the LS method is easier to implement, and that its estimations for the parameters are the same as for ML (in the case when the errors are independent and identically distributed), we use the LS as the core method of estimating the unknown parameters of our model. Fortunately the model that we develop provides a very accurate fit for the mortality data across all ages and the residuals in all data points have approximately constant mean with a value very close to zero and almost constant variance (i.e. the errors are homoscedastic). We also verify that the residuals are approximately normally distributed (see Q-Q plots in Chapter 5). This validates that the method used does not cause any significant biases to the estimates.

1.3.3. Model selection

Model selection process is used to determine the most optimal model, from a set of candidate models, i.e. the model which fits a considered set of data without over or under fit. Quality of fit can be represented by a certain quantity which estimates over/under fitting and depends on the number of model parameters and the size of the dataset. Generally speaking, an increase in sample size can compensate for the increase in the model parameters. However, if a model with a small number of parameters fits a large sample of data efficiently, then this model is preferred amongst other models with a larger number of parameters.

A number of different criteria have been introduced for the selection of the optimal model. The two most commonly used criteria that are closely related to each other, are the Akaike Information Criterion (AIC) (Akaike, 1973) and the Bayesian Information Criterion (BIC) which is attributed to Gideon Schwarz (Schwarz, 1978). Both criteria are based on the likelihood function but they differ in the penalty term used for the number of model parameters. The penalty term is important for fitting procedures, because the increase in the number of parameters usually increases the likelihood but also over-fits the data. The penalty term of BIC is $k \log(n)$ where k is the number of parameters and n is the number of observations, and the penalty term of AIC is $2k$. Therefore, for any sample size greater than eight, the penalty term in BIC is larger than in AIC. This implies that the BIC generally penalizes the number of unknown parameters more

strongly than the AIC does (and makes BIC favoured compared to AIC), as it depends on the size of n and the relative magnitude of n and k. Many studies have been performed to compare these information criteria (see for example (Koehler and Murphree, 1988, Kuha, 2004, Stone, 1979)). However, since both are very similar, it is usually difficult to choose one of them over the other. For this study, we use BIC for the selection of the optimal model for the only reason that it gives stronger penalties to the number of free parameters.

The BIC is represented by the formula:

$$\text{BIC} = -2\log(\hat{L}) + k\log(n) \tag{1.19}$$

where $\hat{L}$ is the maximized value of the likelihood function (Neath and Cavanaugh, 2012, Koehler and Murphree, 1988). Under the assumption that the residuals are independent and normally distributed, that is, the relationship expressed by the equation (1.18) is hold, the BIC can be rewritten as

$$\text{BIC} = n\log(\hat{\sigma}^2) + k\log(n). \tag{1.20}$$

This representation does not depend on the maximized likelihood and therefore can be used when the unknown parameters are estimated by the Least Squares method. Given a set of candidate models, the model with the lower value of BIC is the one to be preferred. Equation (1.20) indicates that the BIC is an increasing function of $\hat{\sigma}^2$ and an increasing function of k. Hence, using the LS method, lower BIC implies better fit as it gives a minimal number of parameters for the minimal sum of squared residuals ($\hat{\sigma}^2$ in equation (1.20) is the sum of squared residuals divided by the number of data).

Chapter 2.

A mathematical model of mortality dynamics across the lifespan combining heterogeneity and stochastic effects

2.1. Summary

The mortality patterns in human populations reflect biological, social and medical factors affecting our lives, and mathematical modelling is an important tool for the analysis of these patterns. It is known that the mortality rate in all human populations increases with age after sexual maturity. This increase is predominantly exponential and satisfies the Gompertz equation. Although the exponential growth of mortality rates is observed over a wide range of ages, it excludes early- and late-life intervals. In this Chapter we accept the fact that the mortality rate is an exponential function of age and analyse possible mechanisms underlying the deviations from the exponential law across the human lifespan. We consider the effect of heterogeneity as well as stochastic factors in altering the exponential law and compare our results to publicly available age-dependent mortality data for Swedish and US populations. In a model of heterogeneous populations we study how differences in parameters of the Gompertz equation describing different subpopulations account for mortality dynamics at different ages. Particularly, we show that the mortality data on Swedish populations can be reproduced fairly well by a model comprising four subpopulations. We then analyse the influence of

stochastic effects on the mortality dynamics to show that they play a role only at early and late ages, when only a few individuals contribute to mortality. We conclude that the deviations from exponential law at young ages can be explained by heterogeneity, namely by the presence of a subpopulation with high initial mortality rate presumably due to congenital defects, while those for old ages can be viewed as fluctuations and explained by stochastic effects.

2.2. Introduction

Analysis of the dynamics of human mortality over the life course is of great importance. Demographic comparisons between populations may reveal clues into differences in causes of mortality that may be related to intrinsic and extrinsic factors. Study of mortality dynamics over age has a long story. Many researchers following the early works of Lexis (Lexis, 1878) and Pearson (Pearson, 1897) have considered mortality at different age intervals to be affected by different factors, with ageing to be at play only after sexual maturation (Gompertz, 1825). Early considerations of age dependent mortality were rather philosophical: according to Lexis (Lexis, 1878) everyone should live the same length of time — the normal length of life — but some die earlier due to accidents. He distinguished "normal" deaths which occur at the normal age of death or are randomly distributed around that age, from premature deaths of adults and, a fortiori, deaths of children. Pearson's (Pearson, 1897) approach was far more scientific: he considered death as a random event and his statistical analysis of age distribution of death in England (1871-1880) revealed five different phases described by different probabilities of death for five age groups. The reasons why probability of death (or mortality rate) depends on age and should follow different dynamics at different age intervals are not well understood. Recent studies of mortality dynamics over the life course attempt to understand the mechanisms of age-related mortality based on the underlying physiological, molecular and genetic processes. It is not surprising that a number of studies have been conducted to analyse mortality data as a function of age (Vaupel, 2005, de Magalhaes et al., 2005, Gavrilov and Gavrilova, 2003).

Mortality rate m_i at age i is defined as number of deaths of individuals of age i (ΔN_i) divided by the number of person-years (PY_i) calculated for individuals of age i in the population (Preston et al., 2000). If an average person who died at age i have died a

years ($0 < a < 1$, i.e. a gives the fraction of a year) after his last birthday then $PY_i = N_i - (1 - a)\Delta N_i$ where N_i is the number of individuals who have reached age i. Thus:

$$m_i = \frac{\Delta N_i}{N_i - (1 - \alpha)\Delta N_i} \quad \text{or equally} \quad \Delta N_i = \frac{m_i N_i}{1 + (1 - \alpha)m_i}. \tag{2.1}$$

If an average person who died at age i had died 6 months after his birthday then $\alpha = 0.5$. Observations on mortality in human populations indicate that $\alpha = 0.5$ (with very high precision) for all ages except for the age zero, $i = 0$, for which parameter α is considerably smaller (typically $\alpha \approx 0.35$). Furthermore, the number of deaths of people aged i can be represented as:

$$\Delta N_i = N_i - N_{i+1}, \tag{2.2}$$

where N_{i+1} represents the number of people who has reached the age $i + 1$.

For most of the human lifespan the Gompertz equation (Gompertz, 1825) depicting the exponential increase in mortality with age fits the data well and has been widely used. Mathematically the exponential dynamics of mortality rate is represented as:

$$m_i = m_0 e^{\beta i}, \tag{2.3}$$

where m_0 is the initial mortality when $i = 0$ (can be derived from the mortality at the age when mortality rates begin to climb) and parameter β defines the rate of demographic ageing or how quickly the mortality rate is changing.

Combining equations (2.1-2.3) we have:

$$N_{i+1} = N_i - \Delta N_i = N_i - \frac{m_i N_i}{1 + 0.5 m_i} = \left(\frac{1 - 0.5 m_i}{1 + 0.5 m_i}\right)N_i = \left(\frac{1 - 0.5 m_0 e^{\beta i}}{1 + 0.5 m_0 e^{\beta i}}\right)N_i. \tag{2.4}$$

Equation (2.4) shows how the number of individuals of age $i + 1$ is defined by the number of individuals of age i. Using the derivation from equation (2.4) multiple times we can find the size N_i as a function of the initial size N_0, initial mortality, m_0, and parameter β:

$$N_i = \left(\frac{1 - 0.5m_0 e^{\beta(i-1)}}{1 + 0.5m_0 e^{\beta(i-1)}}\right) N_{i-1} = \cdots = N_0 \prod_{k=0}^{i-1} \left(\frac{1 - 0.5m_0 e^{\beta k}}{1 + 0.5m_0 e^{\beta k}}\right). \qquad (2.5)$$

Equations (2.1-2.5) represent a discrete counterpart of the continuous equations associated with the Gompertz law (Mueller et al., 1995).
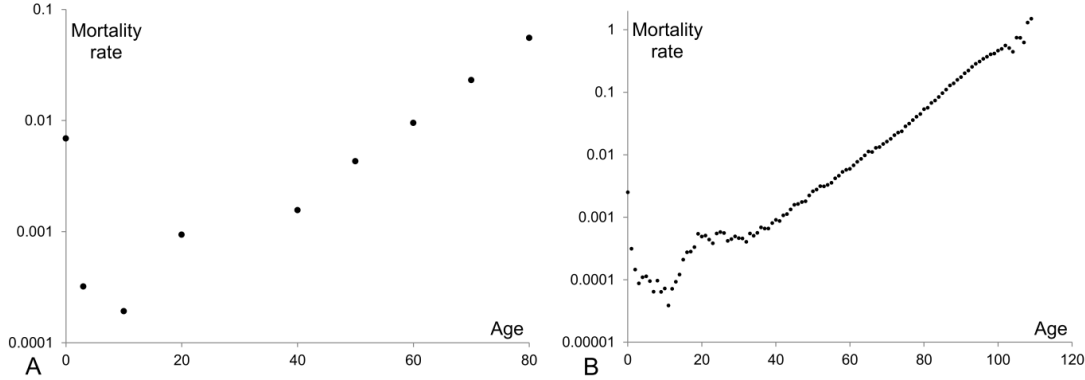

Figure 2.1: Mortality rate versus age for the United States population in the year 2002 (Panel A) and for the Swedish population in the year 2007 (Panel B). *The data presented in panel A is taken from the Centre for Disease Control and Prevention (http://www.cdc.gov/nchs/deaths.htm) and the data in panel B from the Human Mortality Database (http://www.mortality.org). Both panels represent the logarithm of mortality rate versus the age i. The data for the Swedish population is given for all ages while for the American population only selected ages are given. In both cases the data after the age of 25-30 fits into a straight line, i.e. indicates an exponential growth. The data for the Swedish population (which is more complete) shows the deviation from the exponential growth (which even includes drops in mortality rate) after the age of 100.*

Actual data on mortality in human populations can be found in different formats. Most commonly it is represented as a logarithm of mortality rate versus age (see Figure 2.1) which can be interpolated by a linear plot if it is given by the exponential function as in the case of equation (2.3). Plots in Figure 2.1 show that the mortality rate increases for most ages and this increase is approximately exponential. The deviations from the exponential law are observed in young (before 32) and old (after 102 in panel B) ages. The mortality plateau at late ages (Mueller and Rose, 1996, Pletcher and Curtsinger, 1998, Wachter, 1999) is one of particularly intriguing facets in human populations as well as in other non-human species.

A number of mathematical models have been developed and used to analyse observations on the mortality dynamics in human populations as well as in populations of other species (Yashin et al., 2000, Vaupel, 2010). Various explanations have been put

forward for the peculiarities of mortality dynamics at young and old ages as we have mentioned in Chapter 1. The deviations of mortality from the exponential law can be explained by heterogeneity (Vaupel and Yashin, 1985a, Vaupel et al., 1979), while heterogeneity can be explained and described using different models (Steinsaltz and Wachter, 2006, Lebreton, 1996). However, we feel that the systematic analysis of the mortality dynamics in heterogeneous populations is incomplete. Particularly, it would make an important exercise to construct a model of heterogeneous population with parameters fitting real observations as this could provide clues regarding biological, genetic and medical factors driving these mortality patterns. Another reason for deviations of mortality dynamics from Gompertz law can be associated with the random events affecting the longevity. The role of stochastic effects on the mortality dynamics as mediated by their impact on individual frailty have been addressed by many authors (Weitz and Fraser, 2001, Vaupel, 2010). And again the results of these studies have not been systematically compared with detailed observations available nowadays.

In this Chapter, we aim to model mortality across the whole lifespan presuming that the rate of mortality changes over age according to the Gompertz Law. Although many other models have been used to describe mortality dynamics over age (Pletcher, 1999) there is a genuine feeling that the fundamental processes underlying mortality should result in exponential law (Yashin et al., 2000). Therefore we analyse whether the deviations from Gompertz law can be explained by the heterogeneity of populations while the mortality in each subpopulation is still described by the Gompertz law. Also assuming that the mortality dynamics is described by Gompertz law we check whether the deviations from this law can be explained by stochastic effects. In order to perform this analysis we developed mathematical models comprising the heterogeneity of the population and/or stochastic effects. We also aim to derive likely Gompertz parameters for the model so that it would fit the observation data used in this study and infer clues about biological, social or biomedical processes at work. We first focus on the young ages and then on the old ages to model and analyse the irregularities in the mortality dynamics using data both from Sweden and the US. This allows us to explain these irregularities and to reproduce the observed data in the model.

2.3. Study of heterogeneous populations

2.3.1. Mathematical model of heterogeneous populations

Each human population can be seen as consisting of a number of subpopulations which differ genetically and/or by life style (for example associated with gender or occupation). The parameters defining the mortality dynamics (m_0, β in equation (2.3)) of each subpopulation can be different, reflecting the variations in the genotype and life style (see (Vaupel et al., 1998) and references therein). Therefore, we can model the whole heterogeneous population in the following way. We consider the population as consisting of n subpopulations and assume that the mortality rate in each subpopulation is defined by the Gompertz equation, although the equation parameters are different for different subpopulations (Vaupel, 2010). Let us use the notation $N_{j,0}$ for the initial size, $m_{j,0}$ for the initial mortality rate, and β_j for the rate of mortality dynamics of subpopulation j. According to the Gompertz law the mortality rate of subpopulation j is:

$$m_{j,i} = m_{j,0}e^{\beta_j i}. \tag{2.6}$$

If the entire population consists of n subpopulations then the equation (2.1) can be rewritten as:

$$
\begin{aligned}
m_i &= \frac{\Delta N_{1,i} + \Delta N_{2,i} + \cdots + \Delta N_{n,i}}{N_{1,i} + N_{2,i} + \cdots + N_{n,i} - 0.5\left(\Delta N_{1,i} + \Delta N_{2,i} + \cdots + \Delta N_{n,i}\right)} \\
&= \frac{\sum_{j=1}^{n} \Delta N_{j,i}}{\sum_{j=1}^{n} N_{j,i} - 0.5 \sum_{j=1}^{n} \Delta N_{j,i}}
\end{aligned}
\tag{2.7}
$$

where sub-index i denotes the age and sub-index j - the subpopulation. Taking into account equations (2.1) and (2.6), equation (2.7) can be rewritten as:

$$
m_i = \frac{\sum_{j=1}^{n} \dfrac{N_{j,i} m_{j,0} e^{\beta_j i}}{1 + 0.5 m_{j,0} e^{\beta_j i}}}{\sum_{j=1}^{n} N_{j,i} - 0.5 \sum_{j=1}^{n} \dfrac{N_{j,i} m_{j,0} e^{\beta_j i}}{1 + 0.5 m_{j,0} e^{\beta_j i}}}.
\tag{2.8}
$$

In this equation the actual sizes of subpopulations can be replaced by their fractions. That is, we can define $\rho_{j,i}$ as the fraction formed by subpopulation j over the total population at any age i:

$$\rho_{j,i} = \frac{N_{j,i}}{N_i} = \frac{N_{j,i}}{N_{1,i} + N_{2,i} + \cdots + N_{n,i}} \text{ with } \sum_{j=1}^{n} \rho_{j,i} = 1. \tag{2.9}$$

Then the equation (2.8) can be rewritten as:

$$m_i = \frac{\sum_{j=1}^{n} \dfrac{\rho_{j,i} m_{j,0} e^{\beta_j i}}{1 + 0.5 m_{j,0} e^{\beta_j i}}}{1 - 0.5 \sum_{j=1}^{n} \dfrac{\rho_{j,i} m_{j,0} e^{\beta_j i}}{1 + 0.5 m_{j,0} e^{\beta_j i}}}. \tag{2.10}$$

The fractions $\rho_{j,i}$ in the equation (2.10) are defined by the initial fractions $\rho_{j,0}$ by the equation similar to equation (2.5):

$$\rho_{j,i} = \frac{\rho_{j,0} \prod_{k=0}^{i-1} \left(\dfrac{1 - 0.5 m_{j,0} e^{\beta_j k}}{1 + 0.5 m_{j,0} e^{\beta_j k}} \right)}{\sum_{s=1}^{n} \left(\rho_{s,0} \prod_{k=0}^{i-1} \left(\dfrac{1 - 0.5 m_{s,0} e^{\beta_s k}}{1 + 0.5 m_{s,0} e^{\beta_s k}} \right) \right)}. \tag{2.11}$$

We will use equations (2.10-2.11) to define the mortality rate of the heterogeneous population as a function of age i and to examine the effect of model parameters on the dynamics of the mortality over age in the heterogeneous population. The continuous counterpart of equation (2.10) can be found in (Vaupel and Yashin, 1985a) (see also Chapter 5).

2.3.2. Mortality dynamics in the model of heterogeneous populations

We start our study by considering a heterogeneous population consisting of two subpopulations (Figure 2.2). The mortality of each subpopulation is described by equation (2.3) with parameters specific to the subpopulation. We can use equation (2.8) or equation (2.10) to analyse how the values of model parameters describing each subpopulation, namely initial sizes, $N_{1,0}$ and $N_{2,0}$ (or initial fractions $\rho_{1,0}$ and $\rho_{2,0}$), initial mortalities, $m_{1,0}$ and $m_{2,0}$, and parameters β_1 and β_2, defining ageing of subpopulations, affect the dynamics of the mortality rate of the entire population. Figure 2.2A shows the influence of the initial mortality rate of a subpopulation on the dynamics of the total mortality rate. Here we consider the case when the subpopulations have equal initial sizes and equal slopes of ageing, i.e., $\beta_1 = \beta_2$. We can see that the value of the mortality rate for the entire population is initially in between (exactly in the middle for age $i = 0$)

the mortality rates of the two subpopulations, but in the long run merges with the subpopulation that has the lower initial mortality. An increase in the difference in the initial mortality of subpopulations reduces the time needed for these to merge. The plot of mortality rate versus age has a single minimum that shifts down to smaller ages as the higher initial mortality is increased (compare red, green and blue solid lines in Figure 2.2A).
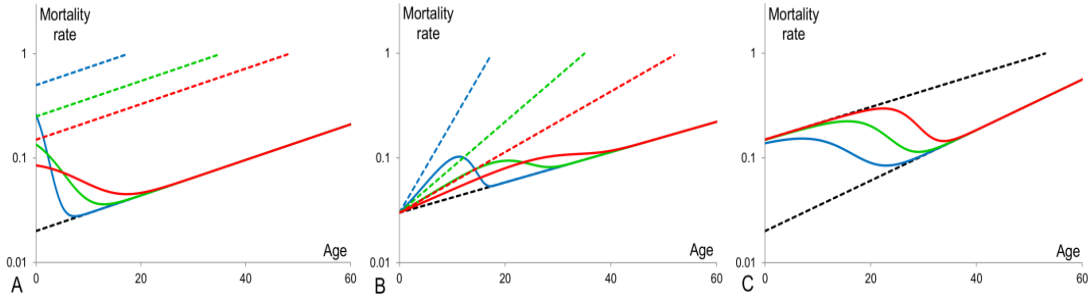

Figure 2.2: The effect of varying model parameters on the mortality dynamics of a heterogeneous population consisting of two subpopulations.
A: The effect of varying the initial mortality rate for one of the subpopulations. *Subpopulations have equal initial sizes ($\rho_{1,0} = \rho_{2,0} = 0.5$) and equal ageing slopes ($\beta_1 = \beta_2 = 0.039$). The initial mortality $m_{1,0}$ takes the values 0.5, 0.25 and 0.15 for the blue, green and red dashed lines, respectively, while the initial mortality $m_{2,0} = 0.02$ is constant. The total mortality of the entire population is represented by a solid line with the colour of the corresponding dashed line (indicating the value of $m_{1,0}$).*
B: The effect of varying the ageing slope. *Subpopulations have equal initial sizes $\rho_{1,0} = \rho_{2,0} = 0.5$ and equal initial mortality rates $m_{1,0} = m_{2,0} = 0.03$. The rate of ageing β_1 takes the values 0.2, 0.1 and 0.067 for the blue, green and red dashed lines, respectively, while $\beta_2 = 0.033$ is constant. The total mortality of the entire population is represented by a solid line with the colour of the corresponding dashed line (indicating the value of β_1).*
C: The effect of varying the initial size of the subpopulation. *Two subpopulations (dashed lines) with different ageing slopes ($\beta_1 = 0.036$, $\beta_2 = 0.056$) and different initial mortality rates ($m_{1,0} = 0.15$, $m_{2,0} = 0.02$) are considered. Blue, green and red lines show the total mortality of a whole population where the initial fraction $\rho_{1,0}$ is 0.9, 0.99 and 0.999 correspondingly.*

We have also checked how the difference in the ageing slopes, β_1 and β_2, of subpopulations influences the dynamics of the mortality rate of a heterogeneous population (Figure 2.2B). If the ageing slopes of subpopulations are different then the mortality rate of the entire population has a value in between the mortality rates of subpopulations. The total mortality increases at young ages, decreases for a short age interval and then increases again for old ages. In the long run the total mortality saturates to the level of the mortality rate of the subpopulation with the lower ageing coefficient, β. Generally, the graph of total mortality rate has a maximum and a minimum. They

32

both are shifting to old ages when the difference between ageing slopes of subpopulations decreases.

The effect of variation in the initial sizes of subpopulations on the mortality dynamics of a heterogeneous population is shown in Figure 2.2C. We have checked a general case of two subpopulations with different initial mortality rates ($m_{1,0} = 0.15$, $m_{2,0} = 0.02$) and different mortality coefficients ($\beta_1 = 0.0357$, $\beta_2 = 0.0556$) and varied the initial fractions $\rho_{1,0}$ and $\rho_{2,0}$ ($\rho_{2,0} = 1 - \rho_{1,0}$) of the subpopulations. The curves for the total mortality of the entire population (Figure 2.2C) confirm the conclusions made after the analysis of the first two cases (shown in Figure 2.2A and Figure 2.2B). Generally, there is a single minimum on the plot of the total mortality rate and this minimum shifts to old ages with the increase of the initial fraction of subpopulation 1 which displays both a higher initial mortality level and a slower mortality increase with age (i.e. lower β).

The analysis of simulations with varying parameters presented in Figure 2.2 can be used for reproduction of two sets of mortality data presented in Figure 2.1. Data on panel A of Figure 2.1 is represented by a few points while data on panel B is much more detailed. We have picked up these two sets of data to demonstrate that the technique we use for fitting model parameters to observation data works equally well for sparse and extensive data sets. Figure 2.1A gives the mortality data for the USA in the year 2002 taken from the Centres for Disease Control and Prevention http://www.cdc.gov/nchs/deaths.htm. Note, that this data could be interpolated either by a plot which has a single minimum (skipping the point at age 20) or by a plot with two minima. Our analysis (Figure 2.2) indicates that in order to reproduce a plot of mortality rate with a single extreme point we need to consider two subpopulations. Since each subpopulation is described by three model parameters (initial mortality, m_0, rate of ageing, β, and fraction, ρ) and $\rho_1 + \rho_2 = 1$ we need to find values of five parameters to fit the data in Figure 2.1A. In the general case of n subpopulations the number of free (unknown) parameters is $k = 3n - 1$.

To find values for the free parameters that could minimize the sum of squared residuals (residual is a difference between the theoretical prediction and the observation data) and therefore to fit the data we have used the Least Squares Method as described in Chapter 1. This method was implemented using nonlinear regression algorithm (provided by the command **DataFit** in Maple which is included in the **DirectSearch** package). Using the LS method we have fitted parameters of the models for heterogeneous populations

consisting of two (Figure 2.3A) and three (Figure 2.3B) subpopulations with the US data in Figure 2.1A. The next task is to find out which of these two models is a better fit. According to the BIC, the heterogeneous model with three subpopulations (BIC = -44.06) fits the US data better than the model with two subpopulations (BIC = -13.01). For the particular US data we cannot consider a population composed of four subpopulations because in that case we will need to calculate 11 free parameters while there are only 9 data points.
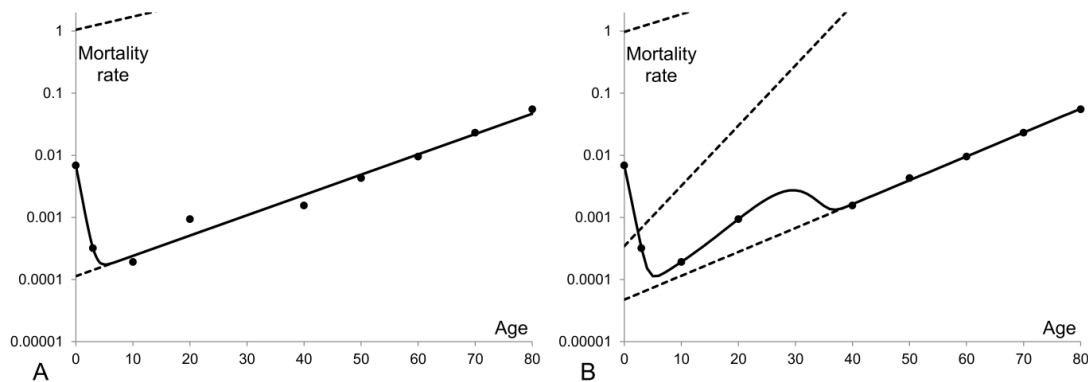

Figure 2.3: Fitting the heterogeneous model to the US mortality data using the Least Squares Method. *The data is denoted by the circle symbols, the mortality rates of modelled subpopulations are given by the black dashed lines and the mortality of the whole population by the solid black curve.*
A: The heterogeneous population composed by two subpopulations. *Model parameters: 1st subpopulation: $m_{1,0} = 1.052$, $\rho_{1,0} = 0.00978$, $\beta_1 = 0.0488$; 2nd subpopulation: $m_{2,0} = 0.00001133$, $\rho_{2,0} = 0.99022$, $\beta_2 = 0.0753$. Sum of squared residuals:* 0.625592.
B: The heterogeneous population composed by three subpopulations. *Model parameters: 1st subpopulation: $m_{1,0} = 0.9748$, $\rho_{1,0} = 0.01038$, $\beta_1 = 0.065$; 2nd subpopulation: $m_{2,0} = 0.000345$, $\rho_{2,0} = 0.02486$, $\beta_2 = 0.223$; 3rd subpopulation: $m_{3,0} = 0.0000475$, $\rho_{3,0} = 0.96476$, $\beta_3 = 0.0885$. Sum of squared residuals:* 0.009544.

Now let us consider the data on the death rates in Sweden for the year 2007 presented in Figure 2.1B (taken from the Human Mortality Database: http://www.mortality.org). This data is considerably more detailed compared to the US data presented in Figure 2.1A. From Figure 2.1B we see that the mortality rate is initially about 0.0025, and then declines to a minimum point at the age of 10 years, then increases until a local maximum value at the age of 25 years, drops slightly and advances exponentially (along a straight line on a logarithmic scale) from the age of 30 to about 100 years. At advanced ages, i.e. after approximately 100 years, the mortality data don not follow the monotonically

increasing line. This can be explained either by fluctuations in mortality data or by the fact that the mortality rate starts to decline with age.

Figure 2.4 shows three models of heterogeneous populations composed of three (panel A), four (panel B) and five (panel C) subpopulations fitting the data on the Swedish population presented in Figure 2.1B. Compared by eye the plots on panels B and C seem to be better fits than the plot on panel A. The BIC indicates (see Figure 2.4 legend) that the four-subpopulation model (panel B) fits to the data slightly better than the five-subpopulation model (panel C).
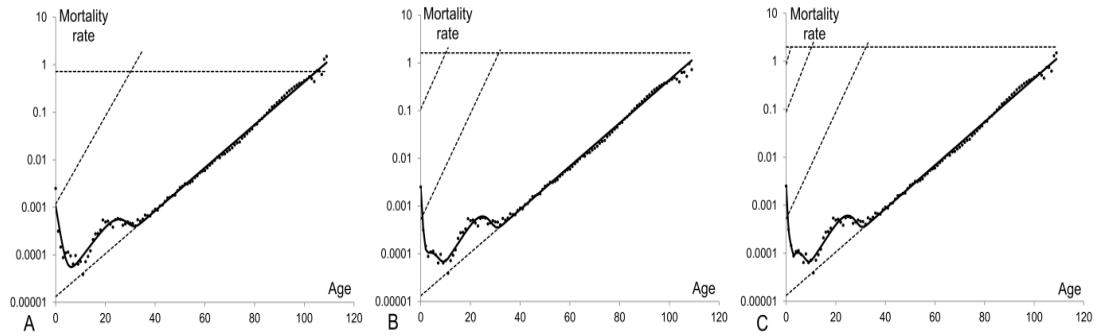

Figure 2.4: Fitting the heterogeneous model to the Swedish mortality data using the Least Squares Method. *The data is denoted by the circle symbols, the mortality rates of modelled subpopulations are given by the dashed lines and the mortality of the whole population by the solid curve.*

A: The heterogeneous population composed by three subpopulations. *Model parameters: 1st subpopulation: $m_{1,0} = 0.7211$, $\rho_{1,0} = 0.00198$, $\beta_1 = 0.67 \cdot 10^{-5}$. 2nd subpopulation: $m_{2,0} = 0.001169$, $\rho_{2,0} = 0.00483$, $\beta_2 = 0.2129$. 3rd subpopulation: $m_{3,0} = 0.00001317$, $\rho_{3,0} = 0.99319$, $\beta_3 = 0.1041$. Sum of squared residuals: 5.715275, $BIC = -287.7033$.*

B: The heterogeneous population composed by four subpopulations. *Model parameters: 1st subpopulation: $m_{1,0} = 1.6139$, $\rho_{1,0} = 0.00266$, $\beta_1 = 0.67 \cdot 10^{-5}$. 2nd subpopulation: $m_{2,0} = 0.108$, $\rho_{2,0} = 0.00057$, $\beta_2 = 0.2685$. 3rd subpopulation: $m_{3,0} = 0.00052$, $\rho_{3,0} = 0.00460$, $\beta_3 = 0.2558$. 4th subpopulation: $m_{4,0} = 0.000013146$, $\rho_{4,0} = 0.99217$, $\beta_4 = 0.1041$. Sum of squared residuals: 3.229884. $BIC = -336.3785$.*

C: The heterogeneous population composed by five subpopulations. *Model parameters: 1st subpopulation: $m_{1,0} = 1.986$, $\rho_{1,0} = 0.002$, $\beta_1 = 0.67 \cdot 10^{-5}$. 2nd subpopulation: $m_{2,0} = 0.859$, $\rho_{2,0} = 0.00074$, $\beta_2 = 0.4254$. 3rd subpopulation: $m_{3,0} = 0.088$, $\rho_{3,0} = 0.00052$, $\beta_3 = 0.3041$. 4th subpopulation: $m_{4,0} = 0.0005207$, $\rho_{4,0} = 0.00459$, $\beta_4 = 0.2558$. 5th subpopulation: $m_{5,0} = 0.00001316$, $\rho_{5,0} = 0.99215$, $\beta_5 = 0.1041$. Sum of squared residuals: 3.173179. $BIC = -324.2254$.*
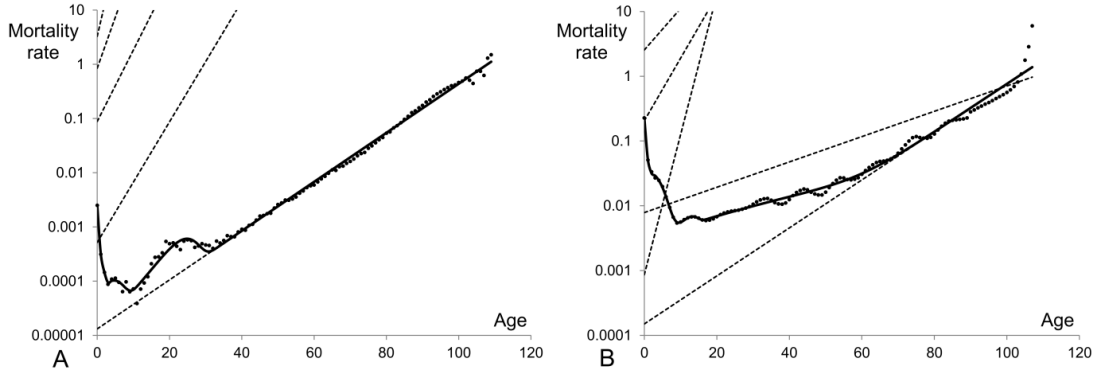
Figure 2.5: Extension of the four-subpopulation heterogeneous model with extra-subpopulation to fit parameter a (in equation (2.1)) in the Swedish mortality data for 2007 (panel A) and for 1751 (panel B). *Since $a = 0.5$ at all ages except for the age zero the fitting was done in 2 steps. 1. The four-subpopulation model is composed to fit all data except the very 1st point; 2. Extra (fifth) subpopulation was added to fit parameter a (see equation (2.1)) and number of deaths for the 1st point. The data is denoted by the circle symbols, the mortality rates of modelled subpopulations are given by the dashed lines and the mortality of the whole population by the solid curve.*
A: The heterogeneous population composed by five subpopulations. *Model parameters: 1st subpopulation: $m_{1,0} = 3.28$, $\rho_{1,0} = 0.00122$, $\beta_1 = 0.0066$, $a_{1,0} = 0.3$. 2nd subpopulation: $m_{2,0} = 0.8477$, $\rho_{2,0} = 0.00072$, $\beta_2 = 0.4324$; 3rd subpopulation: $m_{3,0} = 0.08824$, $\rho_{3,0} = 0.00052$, $\beta_3 = 0.3034$; 4th subpopulation: $m_{4,0} = 0.0005179$, $\rho_{4,0} = 0.0459$, $\beta_4 = 0.2561$; 5th subpopulation: $m_{5,0} = 0.000013163$, $\rho_{5,0} = 0.99295$, $\beta_5 = 0.1041$; Sum of squared residuals: 3.173157; BIC = −324.2261.*
B: The heterogeneous population composed by five subpopulations. *Model parameters: 1st subpopulation: $m_{1,0} = 2.544$, $\rho_{1,0} = 0.13045$, $\beta_1 = 0.1473$, $a_{1,0} = 0.32$. 2nd subpopulation: $m_{2,0} = 0.2054$, $\rho_{2,0} = 0.13239$, $\beta_2 = 0.2222$; 3rd subpopulation: $m_{3,0} = 0.007861$, $\rho_{3,0} = 0.28278$, $\beta_3 = 0.045$; 4th subpopulation: $m_{4,0} = 0.000862$, $\rho_{4,0} = 0.00587$, $\beta_4 = 0.4975$; 5th subpopulation: $m_{5,0} = 0.0001501$, $\rho_{5,0} = 0.44851$, $\beta_5 = 0.0853$; Sum of squared residuals: 4.814607; BIC = −270.3817.*

So far we assumed that $a = 0.5$ (see equation (2.1) for all subpopulations in the heterogeneous model. Observations indicate that this is indeed true for all ages except the very first year ($i = 0$) for which the value of parameter a is significantly smaller. We have taken this observation into account by introducing an extra subpopulation for which the parameter a was predefined for age 0 to fit its observed value. The combination of Gompertz law and small parameter a causes this extra subpopulation to die entirely within the first year (at age 0), and therefore this subpopulation affects only the first ($i = 0$) modelled point. This lets us use the following improved procedure for fitting the heterogeneous population model to the observation data. We remove the first data point ($i = 0$) from consideration and find the best four-subpopulation fit for the remaining data similar to what was done for Figure 2.4B. Then we add an extra subpopulation for

36

which parameter a is small (i.e., $a = 0.3$ which is close to the observed value of a for age 0). As this subpopulation entirely disappears during the 1st year, we just need to adjust its size to have an ideal fit to the first data point ($i = 0$) which was removed from consideration earlier. Figure 2.5A shows the outcome of the model, designed to fit the 2007 data for the Swedish population. It turns that this new model fits the data much better than models designed for Figure 2.4.

We have used the procedure described above to fit a model comprising five-subpopulations to Swedish 1751 data (Figure 2.5B). Although the 1751 data are considerably different from the 2007 data, the designed model also fits it remarkably well. A comparison of models for 1751 and 2007 data indicates that basically all parameters for all subpopulations did change over two-and-a-half century period. The most striking changes are the much higher initial mortality, $m_{1,0}$, in the main subpopulation in 1751 (>10 fold higher), and the fact that the main subpopulation, which comprises over 99% of the population in 2007, makes up less than half of the whole population in 1751. The latter observation also means that the other subpopulations (i.e. those with higher mortality) in 1751 were much larger. The detailed study of the evolution of model parameters could be done by making fits to the data for intermediate years (see in Chapter 3 the analysis on the evolution of mortality patterns).

2.4. Study of fluctuations in mortality dynamics

The important observation we can make is that all models designed so far fail to describe the noisy pattern of observed data in early and late ages. This section will be devoted to the analysis of noise in mortality dynamics.

2.4.1. Modelling the stochastic effects

Assume that the probability to die, q_i, within a year for any individual depends only on his age. Then the number of death of individuals of age i is $\Delta N_i = q_i N_i$ where N_i is the number of people who reached their i-th birthday. Combining this with equation (2.1) we get $q_i = m_i / (1 + 0.5 m_i)$ where m_i is presumed to follow the Gompertz Law (equation (2.3)). Furthermore, taking into account that if q_i is the probability to die then the probability to survive is $p_i = 1 - q_i$, we can find the probability that k out of N_i

individuals survive (while $N_i - k$ individuals die) within a one-year interval. To find this probability we consider the following binomial expansion:

$$1 = [q_i + (1 - q_i)]^{N_i} = \sum_{k=0}^{N_i} \frac{N_i!}{k!\,(N_i - k)!} q_i^{N_i - k} (1 - q_i)^k. \tag{2.12}$$

Here the right-hand-side contains $N_i + 1$ terms (corresponding to values of k from 0 to N_i), each giving the probability for k individuals to survive or, correspondingly, the probability that $\Delta N_i = N_i - k$. Therefore we can use formulas for the mean and the variance of the stochastic process described by the binomial distribution and conclude that the mean value of ΔN_i is $\langle \Delta N_i \rangle = N_i q_i$ and its variance, $\sigma^2 = N_i q_i (1 - q_i)$ (see (Ross, 2002, Morgan, 2000, Allen, 2010)). The mortality error, Δm, can be defined as the standard deviation of the number of deaths divided by the mean number of person years lived at age i:

$$\Delta m_i = \frac{\sigma}{\langle PY_i \rangle} = \frac{\sqrt{N_i q_i (1 - q_i)}}{N_i - 0.5 N_i q_i} = \sqrt{\frac{m_i (1 - 0.5 m_i)}{N_i}}. \tag{2.13}$$

The fluctuations in mortality are observable when the mortality error is high relative to the mean mortality or when the relative mortality error defined as the mortality error divided by the mean mortality is above some threshold, Th:

$$\frac{\Delta m_i}{m_i} = \sqrt{\frac{1 - 0.5 m_i}{N_i m_i}} > Th. \tag{2.14}$$

Equation (2.14) can be extended to the case of a heterogeneous population: the number N_i would represent the total size of the population (sum of all subpopulation sizes) and m_i represent the overall mortality of the heterogeneous population given by equation (2.10). The variables describing subpopulations in a heterogeneous population are independent and therefore the total variance, σ^2, for the whole population is a sum of the variances, σ_j^2, of the constituent subpopulations. Therefore, the mortality error for the entire heterogeneous population can be written in the form:

$$\frac{\Delta m_i}{m_i} = \frac{\left(\sum_{j=1}^n \sigma_{j,i}^2\right)^{1/2}}{\langle PY_i \rangle m_i} = \frac{\left(\sum_{j=1}^n N_{j,i} q_{j,i}(1 - q_{j,i})\right)^{1/2}}{m_i \sum_{j=1}^n (N_{j,i} - 0.5 N_{j,i} q_{j,i})} \tag{2.15}$$

where m_i is defined by equation (2.10).

The value of the relative error gives the amplitude of random fluctuations in mortality curve and can be used as an indicator when the fluctuations on mortality are observable. Similarly, since the binomial distribution can be approximated by normal distribution, we can use normal confidence intervals which show the variability of fluctuations for the same purpose. However, to make a judgement on what values of relative errors and/or confidence intervals are required to make fluctuations visible we need to perform simulations and to produce examples of mortality dynamics affected by fluctuations.

2.4.2. Numerical implementations of the stochastic model

The direct implementation of equation (2.12) for computer simulations requires the comparison of random computer-generated numbers with sums of terms on the RHS of equation (2.12) where binomial coefficients are found using the Pascal triangle. This algorithm would imply the use of inverse cumulative distribution method (Ross, 2007) in simulations of mortality dynamics and it works perfectly well if the size of the population is relatively small. However, for large populations, say when $N_i > 100$, we have to operate with extremely small numbers (like $q_i^{N_i}$ where $q_i < 1$) as well as extremely large numbers (like $N_i!$) hitting in both cases the computer limitations associated with handling real and integer numbers.

To overcome this difficulty we have used an alternative numerical algorithm which does not require equation (2.12). It is slow compared to the above algorithm but allows the consideration of populations of practically any size. If the number of individuals of age i is N_i then in order to find N_{i+1} we order the computer program to generate a set of N_i random numbers (each represented by a real number which is not less than zero and not more than one), i.e. one random number per each individual. Each of the generated random numbers is compared with the probability, q_i, $q_i = m_i/(1 + 0.5m_i)$ where m_i is defined by equation (2.3). The probability a random computer-generated number to be less than q_i is equal to q_i. Therefore, every time when the random number is less than q_i we conclude that one individual dies. Comparing N_i random numbers with q_i lets us make a decision (dies or stays alive) on each of N_i individuals and obtain the number of

individuals N_{i+1} who reach age $i + 1$. This procedure can also be extended to consider heterogeneous populations.

2.4.3. Fluctuations in the mortality dynamics in stochastic model

We have checked how the stochasticity affects the dynamics of mortality, which is presumed to follow the Gompertz law. Figure 2.6 indicates that the amplitude of fluctuations depends on the model parameters such as the size of the population and its initial mortality. We can also see that fluctuations can appear and disappear in different parts of the same mortality plot. The fluctuations are generally observed at early and advanced ages. The occurrence of the fluctuations in Figure 2.6 can be explained by equation (2.14), namely, we can claim that if the fluctuations in the mortality plot are due to stochastic effects, then they should become observable when the relative mortality error is high enough. Equation (2.14) indicates that this error is inversely proportional to the mortality rate and to the size of the population. This implies that deviations from the theoretical mortality data can be observed on two sides of the mortality patterns, at the initial ages where the mortality, m_i, is small and at advanced ages where the number of individuals, N_i, is small. In both cases the fluctuations become observable since the total number of deaths is relatively small.
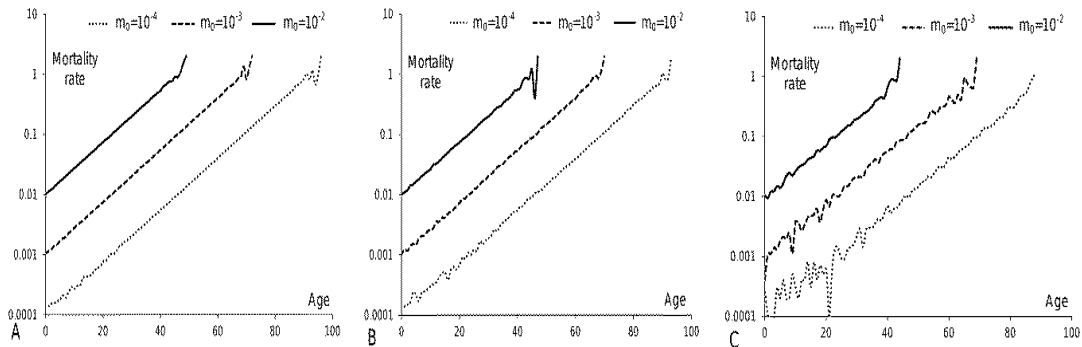
Figure 2.6: The mortality dynamics in a stochastic model of a homogeneous population. *Implementation of stochasticity: the mortality rate (calculated according to the Gompertz law) is converted to a probability, q, for each individual to die. Whether the actual death event takes place or not is decided according to a computer-generated random number. Each plot shows the mortality dynamics for three populations having different initial mortality rates ($m_0 = 0.0001$, $m_0 = 0.001$ and $m_0 = 0.01$ shown by the dotted, dashed and solid lines respectively). Each panel corresponds to the different initial size of population: 10^6 on panel A, 10^5 on panel B and 10^4 on panel C. Each population has mortality coefficient, $\beta = 0.1$.*

Equation (2.14) states that the relative mortality error should be above a certain threshold (Th) for the fluctuations in mortality of a population to become observable. Figure 2.7 shows the graphs of the relative mortality error (dashed lines) calculated for the mortality trajectories (solid lines) presented in Figure 2.6B. An analysis of these graphs indicates that the threshold value, Th, can be estimated by a number between 0.05 and 0.1 to reflect the transitions from observable to unobservable (and back) fluctuations on the mortality plots. As an example, the case of $Th = 0.06$ is shown in Figure 2.7 where a horizontal solid black line represents this threshold. Intersections of the graphs of relative mortality error with this horizontal line correspond to the transitions from fluctuating to non-fluctuating behaviour in the mortality plots.
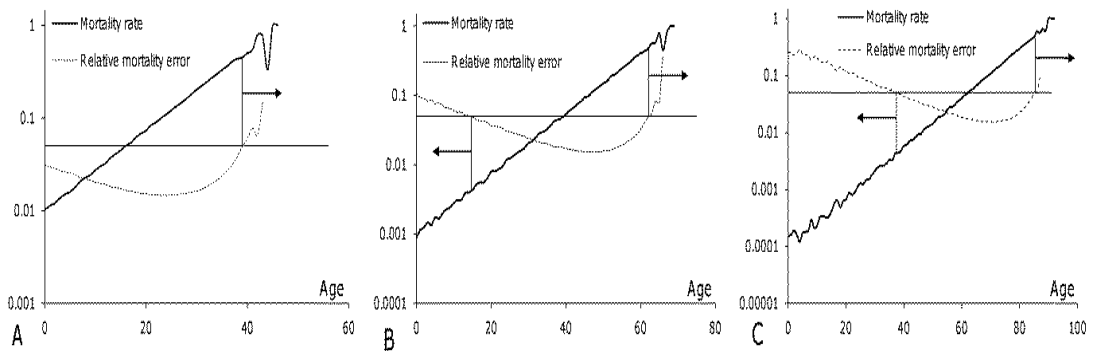

Figure 2.7: Correlation between the fluctuations in mortality dynamics and the relative mortality error. Three graphs of mortality trajectories from Figure 2.6B (solid lines) with the corresponding graphs of relative mortality error (dashed lines) are presented in three panels. When the relative mortality error is above the threshold (the dashed line is above the horizontal line) there are fluctuations in the mortality graphs, while when the error is below the threshold (dashed line below the threshold line), the fluctuations are not observed. Vertical lines indicate the ages when the error curves intersect the threshold line and the arrows show on what side of the vertical line the fluctuations are observed.

In practice, the term $1 - m_i$ in the numerator of the expression for the relative mortality error in equation (2.14) can be omitted (replaced by 1, as m_i is very small) at the transition point at young ages, and $Th = 0.06$ would represent the reciprocal of the square root of the product $N_i m_i$, $Th = 1/\sqrt{N_i m_i}$, or we could say that the young-ages transition takes place when $N_i m_i = 1/Th^2 \approx 300$. The transition point at advanced age takes place when the mortality m_i is not very small and therefore some estimate of the size of population at this transition point can be made. For example, assuming that the transition takes place when $m_i = 0.6$ we get for the size of the population $N_i \approx 180$ (this is a case for the plot shown in Figure 2.7A, when the fluctuations for the advanced ages

start at age $i = 41$ for which the mortality $m_i = 0.6$ and the number of individuals $N_i = 149$).

Figure 2.8 gives a few illustrations of the transition from non-fluctuating to fluctuating dynamics of mortality rate at advanced ages. The graph of mortality dynamics according to the Gompertz law is shown by the solid line while mortality in the stochastic model is shown by dashed line. Increase of the size of the population results in the disappearance of the fluctuations at early ages and further to the occurrence of fluctuations (the transition point) at progressively more advanced ages. Furthermore, despite the big differences in the initial sizes of the populations in all three shown simulations, the fluctuations start to become observable at the age represented by 150-180 individuals. The fluctuations affect the mortality dynamics in a random and unpredictable way which makes the shapes of mortality graphs after the transition points to be considerably different in all presented simulations.
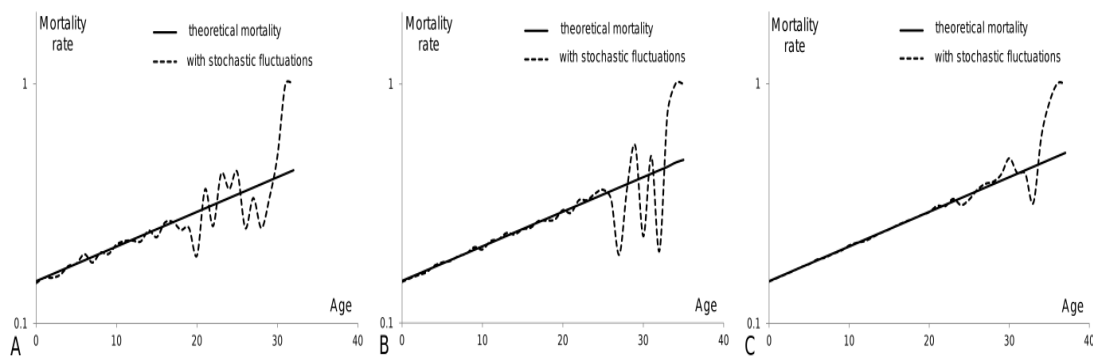
Figure 2.8: Variations in the mortality dynamics due to stochastic effects. *Plots of theoretical mortality (solid line) and actual mortality (dashed curve) for a population with initial mortality rate $m_0 = 0.15$ and mortality coefficient $\beta = 0.033$ are shown. The initial size of the population is 10^4 in panel **A**, 10^5 in panel **B** and 10^6 individuals in panel **C**.*

2.4.4. Fitting models to the observation data

Figure 2.4 demonstrates that the model of a heterogeneous population indeed reproduces the dynamics of observed mortality data very well. However the observation data has a feature which is not captured by the model, namely the fluctuations. The noisy background of mortality records is especially expressed at young and old ages. Many authors have treated the noise at old ages as a deviation from the exponential law (Vaupel et al., 1998) either by considering this deviation as a plateau on the graph or even as a decline in mortality (Partridge and Mangel, 1999). Figure 2.9A shows part of

42

the data from Figure 2.1B which is related to the elderly ages (solid black line). This data represent period data for the Swedish population for the year 2007 obtained from the Human Mortality Database (http://www.mortality.org). Dashed lines on the same panel in Figure 2.9 give parts of the period data for a few other years (1994, 2001, 2009) which are taken from the same database. The plots indicate that the mortality dynamics for all shown cases exhibit roughly the same growing pattern and in all cases the fluctuations appear after the age of 102.
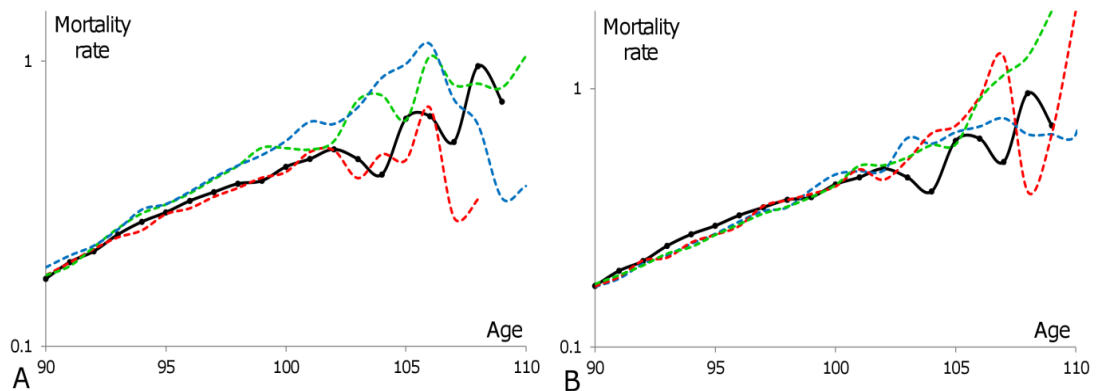

Figure 2.9: Plots of mortality dynamics for the advanced ages in period data for Swedish population (A) in the stochastic model (B). *The solid black line on both panels is the plot of the mortality rate for the advanced ages of the Swedish population in the year 2007 (part of the data from Figure 2.1B). The red, blue and green dashed curves on panel A correspond to the period data on mortality for the Swedish population in 1994, 2001 and 2009. The red, blue and green dashed curves on panel B correspond to three simulations of mortality dynamics in the stochastic model for the different sets of random computer-generated numbers. All plots (period data and simulations) show the deviation from exponential growth in a very similar way, i.e. deviations are represented by fluctuations which start at age 102. Model parameters: $N_{90} = 18660$, $m_{90} = 0.172479$, $\beta = 0.0926$.*

In order to check whether these fluctuations can be explained by the stochasticity in the dynamics of mortality we have performed computer simulations. Each time, we used the same model parameters (initial size of the population, and Gompertz parameters in equation (2.3)) but seeded different sets of random numbers to reproduce death events. The initial mortality, m_0, and initial size of the population in the model was chosen to fit the first data point (at age 90) for the Swedish data for year 2007 and the rate of ageing β to fit the slope of the data points. Figure 2.9B gives a few examples of computer simulated mortality dynamics (dashed lines) as compared with the actual data (black solid line). We can see that the simulations reproduce the data fairly well in a qualitative manner. However, the simulated mortality dynamics follows considerably different plots for different sets of random numbers generated by the computer. We can see that

fluctuations on all presented plots (for simulated data as well as for the data taken from the database) take place after the age of 102 (when about 150 survivors are left in the population) and these fluctuations are the main reason for the deviation from the Gompertz dynamics.

2.5. Model combining the heterogeneity of population with stochastic effects
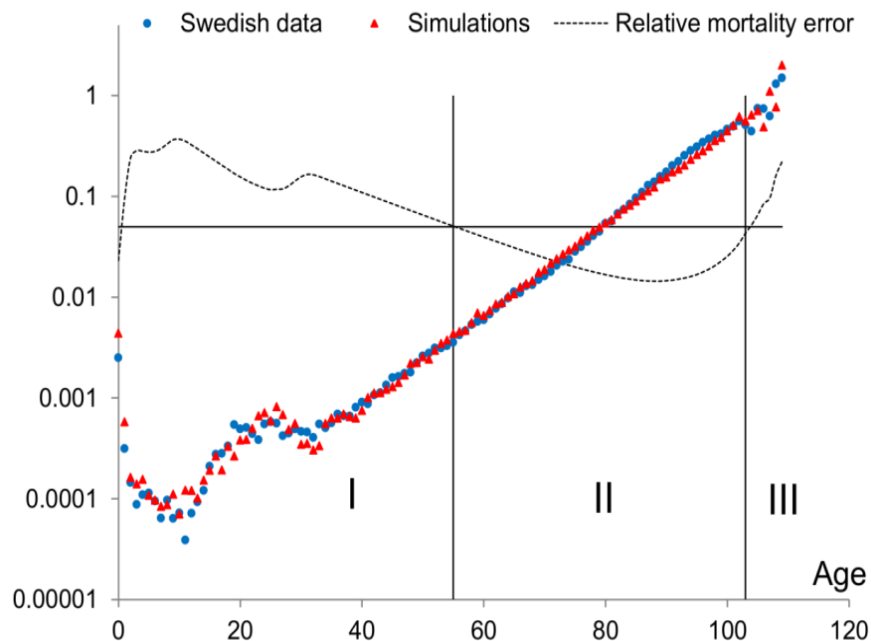
***Figure 2.10: Fitting the mortality data of the Swedish heterogeneous population with stochastic simulations.** The blue circles represent the mortality data for the Swedish population from Figure 2.1B while red triangles represent the simulation data. The model combining heterogeneity (the version used in Figure 2.4B) and stochasticity (implemented in the same way as in Figure 2.6-Figure 2.9) has been used. The graph of the relative mortality error for the simulated data is shown by the dashed line, the threshold level ($Th = 0.05$) is shown by the horizontal line. Vertical lines indicate ages at which the error graph crosses the threshold level. They divide the plot into 3 domains: domains I and III where the relative mortality error is above the threshold line (the fluctuations in both data sets are observed) and domain II where the relative mortality error is below the threshold line and both data sets are relatively free from the noise.*

Up to now we have considered either heterogeneous or stochastic models. Now we can combine these two models to reproduce the entire set of mortality data for the Swedish population presented in Figure 2.1B. We have already modelled this data assuming that the Swedish population is heterogeneous and comprised of four subpopulations (Figure

2.4B). Now we expand that model and introduce the stochastic effect to the mortality description of all four subpopulations. Figure 2.10 shows the results of this simulation. We see that the simulated data (red triangles) exhibit noise, which is very reminiscent of the noise in the real data (blue circles). The noise in both cases is enhanced for young and advanced ages. We have calculated the relative mortality error in simulations (dashed line) and identified its threshold value ($Th = 0.05$ shown by the horizontal solid line) such that the fluctuations (noise) in the mortality are visible if the error is above this threshold. This is observed in two domains indicated by I (up to the age of 53) and III (above the age of 102) in the figure. The relative mortality error in the domain II (between the ages of 53 and 102) is less than the threshold and correspondingly the noise in the mortality data (both actual and simulated) is negligible in this domain.
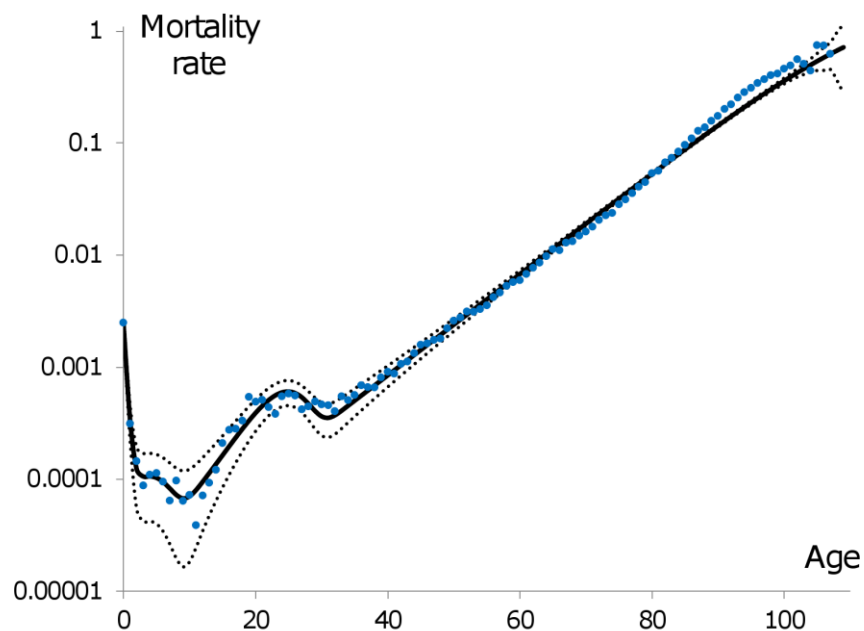
Figure 2.11: Average mortality rates with the confidence intervals in simulations as compared with the observation data. *The blue circles represent the mortality data for the Swedish population from Figure 2.1B while the solid line represent the average curve of simulation data derived by combining heterogeneity (the version used in Figure 2.4B) and stochasticity. On average the simulations at each age are equal to the probability of death predicted by the model. The dashed lines represent the 95% confidence interval around the average, showing the dispersion of fluctuations across the lifespan.*

The analysis of the variability of fluctuations in the observed and simulated mortality patterns presented in Figure 2.10 have been performed using the estimates of relative errors. This analysis can also be done using the estimates of normal confidence intervals. To justify this analysis we will have to assume that binomially distributed mortality rates

(this is definitely the case for simulated patterns) can be approximated by their normal distribution (and this is true when the size of the population is sufficiently large). Figure 2.11 shows the average of simulations (solid line) which is equal to the probability of death, q_i, (calculated by using the four-subpopulation model that reproduces the 2007 Swedish data) and its 95% confidence interval (dashed lines). The confidence intervals show that the dispersion of fluctuations is bigger at young (before age 50) and very old ages (after age 100) while their dispersion is almost zero at ages between 50 to 100 years. The sizes of confidence intervals in Figure 2.11 correlate with the values of relative errors in Figure 2.10 and lead to the similar conclusions concerning the age-related variations of fluctuations in mortality patterns.

2.6. Discussion

Modelling the dynamics of human mortality has long been a focus of research. It could help understand the human ageing process and causes of mortality, possibly providing insights that may help to improve human health and to extend lifetime. Not surprisingly, a number of studies have assessed the impact of heterogeneity on the dynamics of mortality, in particular at later ages (Rossolini and Piantanelli, 2001, Vaupel et al., 1979, Vaupel, 2010). In this Chapter we have developed a mathematical model which lets us reproduce and analyse the mortality dynamics across the entire human life span. Our model combines heterogeneity of population with stochastic effects and the model parameters can be easily tuned so that the simulated data fits well the actual data on mortality dynamics, as shown in Figure 2.4, Figure 2.9 and Figure 2.10.

We have shown that our model is capable of reproducing the actual data on human population mortality fairly well. We had to consider only five subpopulations to reproduce with sufficient accuracy the detailed period data for Swedish populations in 1751 and 2007. Even though this is an underestimate of actual heterogeneity of human populations, it shows how a simple mathematical model can represent actual human mortality well. One intriguing observation from the values of model parameters is that the main subpopulation makes up over 99% of the whole population (Figure 2.4), meaning that in modern populations heterogeneity is actually relatively low. A comparison of model parameters for 1751 and 2007 data (Figure 2.5) shows that model parameters such as initial mortality, m_0, and rate of ageing, β, have dramatically changed, which is not surprising since the conditions of life have also dramatically

changed. A more interesting point is that initial fractions of subpopulations, ρ, have also changed considerably. This, most likely, indicates that advances in medicine and hygiene over the last 250 years have caused fewer individuals to be susceptible or to be exposed to diseases, and infectious diseases in particular, essentially shifting individuals across the subpopulation. This argument can be rephrased in the following way. Each of the five considered subpopulations (say 1st level subpopulations) is also heterogeneous and composed of subpopulations (say 2nd level subpopulations). An improvement in life style over the last 250 years has caused for some of the 2nd level subpopulations to move across the 1st level subpopulations and contribute to the longer lasting fractions. This rearrangement has changed the fraction balance between the 1st level subpopulations.

Our simulations and analysis indicate that the contributions of heterogeneity and stochasticity are different at different ages. The effect of heterogeneity is profound when fractions formed by subpopulations are far from being zero or one. Our model suggests that at early ages a small subpopulation with high initial mortality explains the decline in mortality as this subpopulation gradually disappears. Generally, with an increased age the faster-ageing subpopulations are eliminated and the population starts to act more-and-more homogeneously as it would be composed by a single (having lowest mortality) subpopulation.

The leading causes of death in infants are congenital malformations, disorders related to short gestation and low birth weight, and sudden infant death syndrome (Kung et al., 2007). Therefore, a small subpopulation (the initial fraction comprised by the 1st subpopulation in the simulation shown in Figure 2.4B, $\rho_1 = 0.00266$ which is 0.27% of the total population) with high initial mortality is in line with epidemiological data. On the other hand, it is not clear what the phenomenological differences between other modelled subpopulations are. Our preliminary simulations indicate that fitting the model to describe the mortality dynamics for males or for females will also require four subpopulations (roughly the same as in Figure 2.4B) in both cases. Whether the modelled subpopulations are associated with different social groups has to be analysed in a follow-up study.

We also can identify a subpopulation (subpopulation 3 in Figure 2.4B) which lets us reproduce the mortality peak at about age 20. While this fits the data well, this subpopulation can be considered somewhat artificial because this mortality peak in the

teenage years is likely due to behaviour rather than intrinsic biological properties of a subset of individuals. It is possible, in fact, that the increase in mortality in teenage years is age-specific (goes up and down at a specific age range) and cannot be modelled by Gompertz law. Nonetheless, we speculate that risk-taking behaviour in a subset of individuals could make up such a hypothetical subpopulation.

In line with the results obtained earlier (Vaupel et al., 1979, Rossolini and Piantanelli, 2001, Vaupel, 2010) we conclude that the heterogeneity of a population is sufficient to explain the mortality plateaus observed at later ages. In addition, the size of the population declines with age and the effects of stochasticity become more pronounced. It seems that at later ages when the population is small the stochastic effects can explain the observed mortality plateaus as well as the high-amplitude fluctuations (high noise) in the mortality dynamics. Likewise, at earlier ages, when the number of death events is small, stochastic effects are also noticeable and cause the high-amplitude fluctuations in the mortality dynamics.

In conclusion, the assumption that the populations are heterogeneous and the mortality dynamics of each subpopulation follows the Gompertz equation with different parameters can account for observed deviations of the mortality dynamics (for the entire life course) from the Gompertz law. We also found that stochastic effects are important when relatively few individuals contribute to mortality. Our demographic modelling across the lifespan combining the effects of heterogeneity and stochasticity was successfully tested in simulations of human mortality data from populations in Sweden and the US.

Chapter 3.

Time-evolution of age-dependent mortality patterns in mathematical model of heterogeneous human population

3.1. Summary

The widely-known Gompertz law of mortality states the exponential increase of mortality with age in human populations. Such an exponential increase is observed at the adulthood span, roughly after the reproductive period, while mortality data at young and extremely old ages deviate from it. The heterogeneity of human populations, i.e. the existence of subpopulations with different mortality dynamics, is a useful consideration that can explain age-dependent mortality patterns across the whole life-course. A simple mathematical model combining the heterogeneity of populations with an assumption that the mortality in each subpopulation grows exponentially with age has been proven in Chapter 2 to be capable of reproducing the entire mortality pattern in a human population including the observed peculiarities at early- and late-life intervals. In this Chapter we fit this model to actual (Swedish) mortality data for consecutive periods and consequently describe the evolution of mortality dynamics in terms of the evolution of the model parameters over time. We find that the evolution of the model parameters validates the applicability of the compensation law of mortality to each subpopulation separately. Furthermore, this analysis indicates that the population structure changes so

that the population tends to become more homogeneous over time. Finally, our analysis of the decrease of the overall mortality in a population over time shows that this decrease is mainly due to a change in the population structure and to a lesser extent to a reduction of mortality in each of the subpopulations, the latter being represented by an alteration of the parameters that outline the exponential dynamics.

3.2. Introduction

Mathematical modelling of biological processes such as longevity, ageing and mortality is of interest for many scientists working on various subjects including demography, biology, statistics and actuarial sciences. The event of death and the forces that cause it have puzzled and inspired many philosophers and scientists from the 17th century onwards. Great works such as those by Joseph Addison (1672-1719), Karl Pearson (1857-1936) and Benjamin Gompertz (1779-1865) give us insights on the development of the concept of mortality over the past few centuries (see (Turner and Hanley, 2010) for a review). Addison in his allegorical essay "The vision of Mirza" (Addison, 1711) imagined the human life as a walkthrough over a bridge, "the bridge of human life", where hidden pitfalls open periodically and the people above them fall down and disappear, the forces causing death being then external. Almost two centuries after Addison, Pearson considered death as a random event and decomposed the entire mortality curve into five different phases, described by five different probability distributions (Pearson, 1897). Pearson's concept can be represented with humans crossing the bridge of life, where at each one of the five stages, a marksman attempts to kill them. From one stage to the next the precision of the marksman's weapon improves (five different precisions for the five different age groups) and consequently the chance of death increases. On the other hand, the work by Gompertz (Gompertz, 1825) is of greater importance as he was the first who considered death to be caused by internal forces in organisms and proposed a model for the force of mortality. According to Gompertz, the mortality force increases in a geometrical progression within a wide age-range of lifespan, that is from sexual maturity to considerably old ages.

Graphically the actual mortality data generates patterns which have certain common features (shown in Chapter 2) as well as some quantitative differences as compared to different cohorts and periods. A typical mortality pattern (Figure 3.1) originates from the initial mortality at age zero, falls down to a minimum point (approximately at the age of

50

10), increases to a local maximum (around the age of 25), then slightly decreases or remains constant and after the age of 35-40 advances exponentially satisfying the Gompertz law. At extreme old ages (above the age of 100) there is no common evidence on how the mortality curve behaves as the reported observations are controversial and provided with different explanations (Greenwood and Irwin, 1939, Olshansky, 1998, Gavrilov and Gavrilova, 2011). Various statements made about the mortality dynamics at old ages include the mortality levelling-off or so-called "late-life mortality plateau" (Economos, 1979, Mueller and Rose, 1996, Curtsinger et al., 2006), the late-life mortality deceleration (Depoid, 1973, Horiuchi and Wilmoth, 1998, Thatcher et al., 1998, Gavrilov and Gavrilova, 2001), the decline (Kannisto et al., 1994, Vaupel et al., 1998) or fluctuations at advanced ages (Avraam et al., 2013).
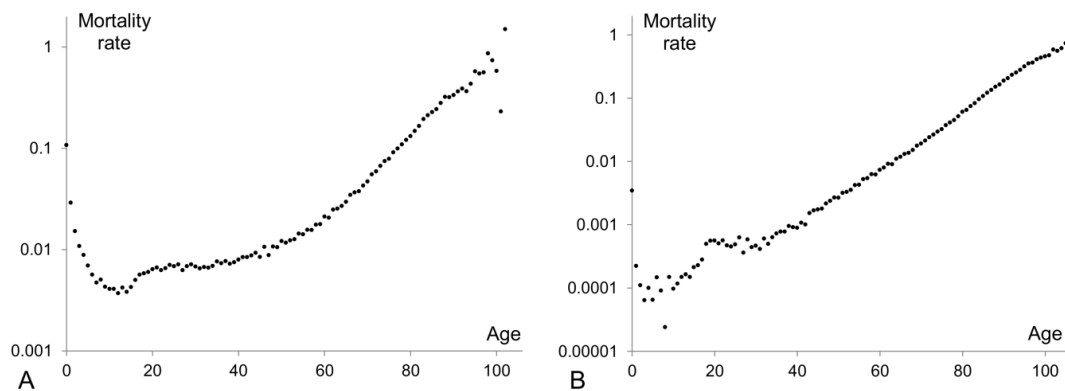
Figure 3.1: Mortality rates for the Swedish population in the period 1900 (panel A) and 2000 (panel B) presented in a semi-logarithmic scale. The data are taken from the Human Mortality Database, http://www.mortality.org.

The high initial level of mortality is due to the fact that new-borns are not particularly fit for the new environment they are born into and therefore, a relatively high proportion of them are not able to survive. As the forces of mortality due to environmental factors decrease, death rates decline. Mortality starts then to increase at the age of 10. One can state that mortality should increase exponentially from this age. However in actual mortality data the exponential increase of mortality is observable only after the ages of 35-40 (Figure 3.1) as between the ages 10 and 35 it overlaps with a local maximum on the mortality curve. This local maximum is apparent at the reproductive period of lifespan and is commonly called "the accidental hump" as it is related to the external causes of deaths (mainly accidents and maternal deaths) due to the risky behaviour of young adults.

Many studies have focused on the analysis of exponential increase of mortality in the range of ages 30 and above. By comparing parameters describing the exponential dynamics for data taken for different human societies it was found that in developed countries initial mortality, m_0, is lower while the mortality coefficient, β, is higher than these parameters describing data for less developed countries. This phenomenon, namely the inverse relationship between initial mortality and mortality coefficient appears to be fundamental (confirmed by all available data) and is the "compensation law" or "compensation effect" (Gavrilov and Gavrilova, 1979, Gavrilov and Gavrilova, 1991) which was described in Chapter 1.

A number of mathematical models have been proposed to analyse mortality dynamics and explain its deviations from the exponential law at early and late life intervals. Some models postulate that a few different processes take place in the population and affect its mortality dynamics (Heligman and Pollard, 1980, Thiele, 1872), while others analyse the impact of population heterogeneity on the dynamics of mortality (Vaupel et al., 1979, Vaupel and Yashin, 1985a). The model that we have developed in Chapter 2 based on the assumption that the mortality dynamics is indeed underlined by an exponential law and deviations from this law are due to the heterogeneity of human populations (Avraam et al., 2013). We have shown that the observed age-specific mortality patterns can be reproduced in a model of heterogeneous population consisting of a few (up to four) subpopulations each following the exponential law over all ages.

Time evolution of mortality dynamics in human populations is of great scientific interest and has practical implementations especially for actuaries, who use extrapolation methods to project mortality trends in order to estimate future life expectancy (Booth and Tickle, 2008, Pitacco, 2004), and to price several longevity products. An example of mathematical study of this evolution can be found in (Gaille, 2012), where the analysis of the evolution of the parameters of two conventional models (Heligman-Pollard and Lee-Carter) is used to forecast the Swiss mortality rates and to study the impact of longevity on Swiss pension funds. Mathematical analysis of the evolution of mortality dynamics could also be useful for demographers (to derive inferences on the population variance) and for biologists (to understand genetics underlying the evolutionary process of ageing).

In this Chapter we aim to describe the evolution of mortality dynamics as time evolution of the parameters in the model of a heterogeneous population so that we could gain insights in the processes governing mortality reductions over the past century. In Section 3.3 we present a simpler form of the model described in Chapter 2 and in Section 3.4 we present the used mortality data. In Section 3.5 we fit the model to various mortality data (cut at a certain age or including/excluding the extrinsic death factors) for consecutive periods and analyse the evolution of the model parameters. The results demonstrate that the population's structure is altered through time and a relative homogenization of the population occurs, explaining an important part of mortality reductions during the 20th century. The analysis also indicates that changes in the initial mortality and mortality coefficient of the exponential law for all subpopulations are in line with the compensation law. Discussion of presented results is provided in Section 3.6.

3.3. Mathematical model

In this Chapter we use the model proposed in Chapter 2, where a human population is considered as heterogeneous and composed of a number of subpopulations. The subpopulations are assumed to obey an exponential law but differ in their mortality parameters. The mortality of the entire population is modelled as a mixture of weighted exponential terms. The weights represent the relative sizes (fractions) of the subpopulations; they depend on age x and their sum is equal to unity at any age. Assuming that the entire population consists of n subpopulations, the total mortality rate is expressed as:

$$m_x = \sum_{j=1}^{n} \rho_{j,x} m_{j,x} = \sum_{j=1}^{n} \rho_{j,x} m_{j,0} e^{\beta_j x} \tag{3.1}$$

where $\rho_{j,x} = N_{j,x} / \sum_{j=1}^{n} N_{j,x}$ is the fraction or proportion representing the size, $N_{j,x}$, of the j-th subpopulation with respect to the whole population size, $\sum_{j=1}^{n} N_{j,x}$, at age x, and $m_{j,x}$ is the exponential function for the j-th subpopulation with initial mortality $m_{j,0}$ and mortality coefficient β_j. Equation (3.1) expresses the mortality rate of the entire heterogeneous population at age x within a cohort (with $N_{j,x}$ representing person-years) or time-period with the assumption that the population is stationary and its size and age-structure do not change over time.

3.4. Data

In this Chapter we used two series of datasets. The first series represents the Swedish mortality (combined for males and females) for a period-interval of one century 1900 – 2000, provided by the Human Mortality Database (http://www.mortality.org). These datasets provide single age mortality rates resulting from all causes of death. The second series of datasets is from the World Health Organization (WHO), which maintains a comprehensive cause-of-death mortality database (http://www.who.int). This database provides the sizes of mid-year populations and number of deaths by cause for various countries over the last 50 to 60 years. We obtained data for Sweden (males' population) from 1951 to 2010. The data are generally divided into five-year age-groups. Thus, our database is composed of nineteen groups, the first for infants less than one year old, a second for children aged one to four, thereafter in groups of five years, ending with the group aged 85 and above. The database of the WHO needs to be adjusted in order to analyse data consistently over time (proportional distribution of the number of deaths of unknown age; adjustments due to the changes of the International Classification of Diseases (ICD) over time (Table 3.1)). Details on these adjustments can be found in (Gaille and Sherris, 2011) and (Arnold (-Gaille) and Sherris, 2013). This database allows us to distinguish between extrinsic mortality and intrinsic mortality.

Causes of death	ICD 7 (1951-1968)	ICD 8 (1969-1986)	ICD 9 (1987-1996)	ICD 10 (1997-2010)
External causes	A138-A150	A138-A150	B47-B56	V00-Y89
Infectious and parasitic diseases	A001-A043	A001-A044	B01-B07	A00-B99

Table 3.1: International Classification of Diseases (ICD) codification. *The classification is regularly reviewed and updated and consequently it has evolved from ICD 7 in the 1950's to ICD 10 which is used nowadays.*

To justify our interest in the relative impacts of intrinsic and extrinsic factors to mortality dynamics we refer to the original Gompertz work where he mentioned two different mortality groups: a first mortality group related to *chance, without previous disposition to death or deterioration*; a second mortality group referring to *deterioration, or an increased inability to withstand destruction* (Gompertz, 1825). Today, the distinction is usually made between intrinsic and extrinsic causes of death, the intrinsic causes being

54

related to Gompertz's inability to oppose destruction. More specifically, the extrinsic causes of death represent external or environmental factors that produce death, while the intrinsic causes of death represent biological forces that lead to death, namely ageing or senescent (Makeham, 1867, Carnes and Olshansky, 1997, Shryock et al., 1975). Makeham (Makeham, 1867) suggested that the Gompertz law will fit much better with mortality due to biological causes (intrinsic causes). It is therefore interesting in our exponential modelling approach to apply our model to intrinsic causes of death. The extrinsic causes of death usually include the external causes of death (such as accidents, homicide and suicide) and the infectious and parasitic diseases, even if some studies recommend the inclusion of some other causes (see e.g. the classifications in (Carnes et al., 2006) and a review in (Carnes and Olshansky, 1997)). The WHO database allows us to analyse the mortality pattern excluding these extrinsic causes of death and then to focus on the mortality dynamics due to intrinsic mortality. Table 3.1 provides the ICD codes for the extrinsic causes we excluded.

3.5. Results

In this section we fit the model proposed in Section 3.3 to the data described in Section 3.4 in order to describe time-dependence of mortality patterns in terms of the evolution of the model parameters. In order to better capture the direction of the trend of each parameter over time, the evolution of each parameter is displayed with the most representative trendline within the examined time range. Two different monotonic functions (linear and exponential) are compared through the BIC (which, in the case of these two functions, works in the same way as R^2) statistics. These two functions were chosen due to their interpretability with respect to the compensation effect of mortality, as further explained in the following sections.

We start our analysis with data which can be fitted by a model comprising a small number of subpopulations and proceed with the analysis of more-and-more complete datasets. In Section 3.5.1 we consider the Swedish data for ages 40+, which is fitted by a one-subpopulation (homogeneous) model. In Section 3.5.2 we extend the considered data by including ages 20+ which are best fitted by a model comprising two subpopulations. In Section 3.5.3 we consider mortality data for all ages but excluding extrinsic death factors which are also best fitted by a model comprising two subpopulations. In Section 3.5.4 we consider the complete mortality dataset for Sweden

which has a best fit to a model comprising four subpopulations. Moving from a smaller to a larger number of subpopulations we make a comprehensive analysis of the evolution of the model parameters. This analysis reveals two effects that take place in the population through time that are compensation and homogenization, which are summarized in Sections 3.5.5 and 3.5.6 respectively.

3.5.1. Evolution of mortality parameters for ages 40+

As mentioned earlier the exponential increase of mortality with age is evident in the mortality pattern at adulthood span after the reproductive period. It follows that the mortality data from age 40 are best fitted by a model of homogeneous population. Our analysis shows that at the beginning of the century, the data for ages 40+ slightly diverge from a pure exponential growth. An example is the 1900 period dataset shown in Figure 3.2A. For those data, the BIC criterion indicates that a two-subpopulation model (BIC $= -184.73$) fits the data better than the homogeneous model (BIC $= -164.08$). On the other hand, more recent data indicate a stronger merge to the exponential growth and therefore the homogeneous model fits the data better than that involving two subpopulations. For example, the BIC numbers for homogeneous and two-subpopulation models fitting the data shown in Figure 3.2B are -319.88 and -319.07 respectively. It is therefore evident that the fit of the homogeneous model to the data for consecutive years gets better over time (one can read it by comparing panels A and B in Figure 3.2). This is mainly due to the extended impact that the accidental hump had on the mortality pattern at the beginning of the 20th century (Figure 3.2A). At that time the number of deaths caused by external factors, such as accidents, was considerably larger (as a result of poor education, inadequate transportation system, unsafe labour environment, lack of contraceptives, etc.) and therefore the accidental hump had a larger amplitude and impacted a wider age range. Over time the magnitude of the accidental hump decreased and at present it affects a smaller range of age groups. It is not surprising that some studies predict that the accidental hump will be less noticeable and even probably disappear in the coming years (Gaille, 2012). Despite this observation, we only consider the model of homogeneous population (described by the exponential law) to fit the Swedish data from age 40 for consecutive years, in order to be consistent in our first step and to study the evolution of mortality parameters.
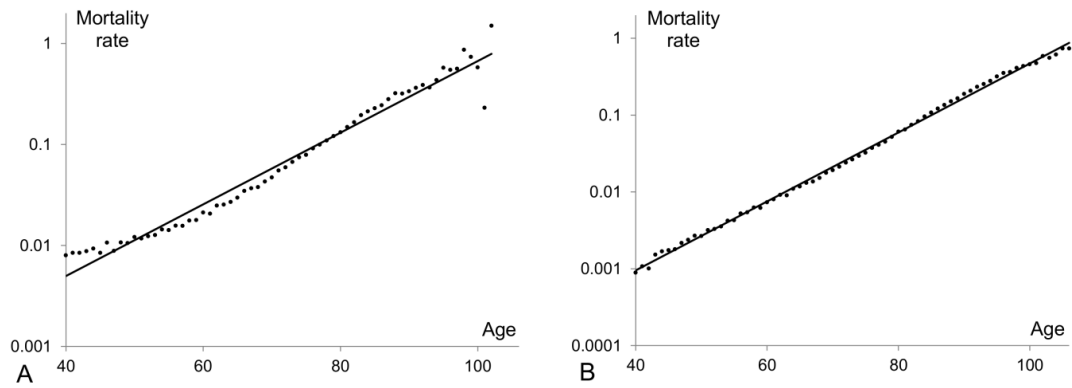
Figure 3.2: Homogeneous model (solid line) fitted to the 1900 (panel A) and 2000 (panel B) period Swedish data (dots) from age 40 and above. *The mortality parameters as estimated by the Least Squares Method are:* $m_0 = 0.000189$, $\beta = 0.0818$ *for panel A and* $m_0 = 0.000015$, $\beta = 0.1033$ *for panel B.*

As the mortality patterns evolve over time, the model parameters that provide the best fit to the actual data at different periods change. Figure 3.3 presents the evolution of the mortality parameters estimated by fitting the homogeneous model (as represented by the simple Gompertz function) to the Swedish mortality rates from age 40 and for every five years over the periods 1900 to 2000. The initial mortality shows an exponentially decreasing trend over time (Figure 3.3A) while the mortality coefficient a linearly increasing trend (Figure 3.3B). The decline of initial mortality is of no surprise, especially in developed countries where medicine and hygiene levels have improved rapidly over the last century. On the other hand the increase in the mortality coefficient is harder to explain. This observation is the phenomenon known as the compensation law of mortality which was described in Chapter 1.

The inverse relationship between the mortality parameters of the modelled homogeneous population is shown in panel A of Figure 3.4. This plot confirms that a high initial rate of mortality is associated with a low mortality coefficient. It is apparent that there is a linear relationship between the logarithm of m_0 and parameter β (this relationship is given by equation (1.4)). From equation (1.4) it follows that if the mortality coefficient is an increasing linear function of time (like the trendline in Figure 3.3B), then the initial mortality declines exponentially (like the trendline in Figure 3.3A). From equation (1.4) we can also estimate the target lifespan, X (the age at which the last survivor dies) and the target mortality rate, M (the mortality at which the last survivor dies).
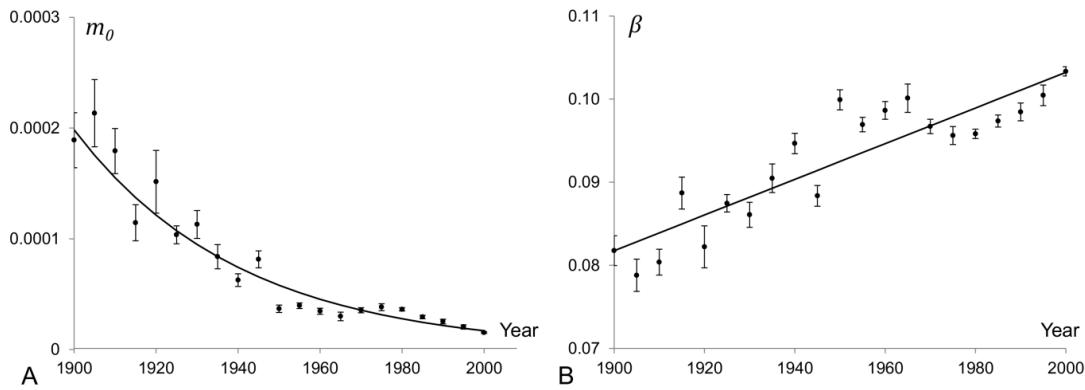
Figure 3.3: Evolution of the parameters of the homogeneous model fitted to the Swedish mortality rates from age 40, for the periods 1900 to 2000 with five-year intervals. Panel A shows an exponentially decreasing trend for the initial mortality and panel B - a linearly increasing trend for the mortality coefficient. In both plots the parameters are presented by dots and the trendlines by solid lines. Error bars represent the standard deviations of estimated parameters.

In panel B of Figure 3.4, the trajectories of the homogeneous model fitted to the Swedish data at four different years are plotted in the same semi-logarithmic plot. The mortality convergence is evident. The coordinates of the intersection point to which mortality trajectories are converging represent the target lifespan and target mortality rate. With a pure form of the compensation effect, all the exponential trajectories should cross that point. Even if the mortality trajectories do not cross strictly at one specific point within the area formed around the error bars of the point (X, M) the compensation effect is still valid in its weak form. The error bars of X and M represent the standard deviations of their values as estimated by the method described in (Reed, 1989). The mortality trajectories that fit the Swedish data above age 40 for different years should theoretically intersect at the point $(X = 103.6 \pm 6.1, M = 0.791 \pm 0.463)$ with a pure compensation effect. The coordinates (X and M) of this point define (as stated by equation (1.4)) the trendline (shown as a solid line) in panel A of Figure 3.4.

This first analysis showed that the compensation effect is evident in the Swedish population for ages above 40 when different periods are compared, although it is in its weak form. We are now interested in studying this effect for a wider age range and for heterogeneous populations, that is when a population is composed of a set of subpopulations.
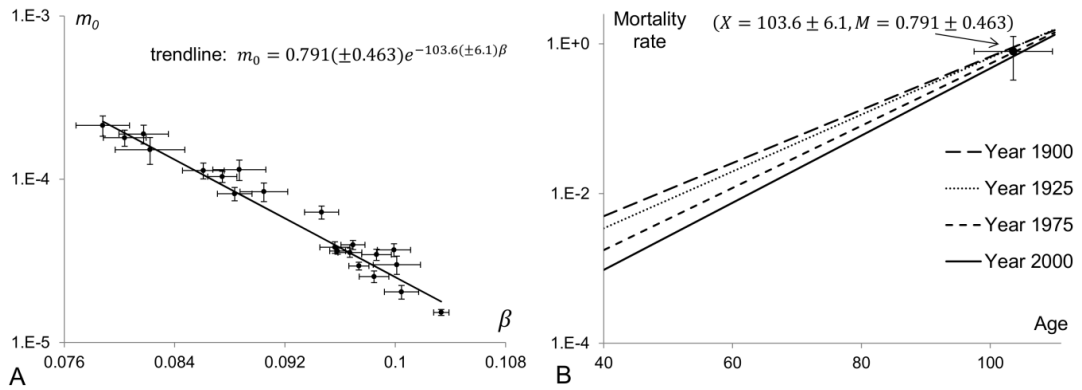
Figure 3.4: Compensation effect in 40+ mortality dynamics. *Panel A shows the inverse relationship between the parameters of the fitted homogeneous model to the Swedish mortality rates for ages 40+ for the period data in the interval 1900-2000. The relationship between the logarithm of the initial mortality and the mortality coefficient is shown to be linear. Panel B shows the convergence of the exponential functions fitted to the Swedish data for ages 40+ and for different periods. Both plots are set in a semi-logarithmic scale.*

3.5.2. Evolution of mortality parameters for ages 20+

It appears that a two-subpopulation model is the best fit for Swedish mortality data which are cut at age 20. The mortality pattern above age 20 includes a part of the accidental hump (usually a piece of its right tail) and the exponential rise of mortality at older ages. There is a local minimum in between these two parts and therefore two subpopulations should be involved to reproduce the mortality pattern observed above age 20. Figure 3.5 presents examples of mortality data and fitted models for 1900 and 2000. The mortality dynamics of the two subpopulations are shown by dashed lines while the mortality of the entire population as calculated by equation (3.1) is shown by the red solid curve. The first subpopulation having higher initial mortality is frailer than the other and thus is producing decline in mortality at the right tail of the accidental hump while the second subpopulation is responsible for the exponential growth of the entire population mortality after the reproductive period. For clarity and consistency purposes, in this thesis the first subpopulation represents the one with the highest initial mortality level ($j = 1$), the second is the one with the second highest initial mortality level ($j = 2$), etc.

An interesting observation from Figure 3.5A is that the heterogeneous model explains the decline of mortality at advanced ages observed in 1900 (local maximum at age ~98). This behaviour is observed when the mortality trajectories of the two subpopulations are crossing and a considerable number of individuals of the frailest subpopulation (the

subpopulation having the highest initial mortality) are still alive at that point. The mortality cross-section (or crossover as called in (Vaupel and Yashin, 1985a) occurs when one of the subpopulations has a lower mortality rate than the other subpopulation at younger ages, but higher at older ages. The two mortality trajectories of the subpopulations fitted to the Swedish 1900 period data (Figure 3.5A) cross each other at age 83, that is when 104 individuals of the first subpopulation (0.78% of the total population) are still alive at that age, while few of them remain alive at the local maximum age (age 98). For the 2000 period data (Figure 3.5B) the trajectories cross at age 84. The only individual of the frailest subpopulation alive at that age (0.002% of the total population at age 84) dies a couple of years after the point of intersection and therefore we don't observe the decline of mortality at advanced ages on this panel. Another observation from Figure 3.5 is about the constant mortality of the frailest subpopulation in panel B. For that particular period, the frailest subpopulation with zero-slope ($\beta_1 = 0$) appears to be optimal. This happens due to the constraint we set in the model, namely, that the mortality coefficients can't be negative. Without this constraint the best fitted models can involve negative mortality coefficients, which means that the models incorporate a process opposite to senescence. Thus here and later, some of our best fits will have a mortality coefficient equal to zero for the frailest subpopulation.

The time-evolution of mortality parameters in the model of a heterogeneous population fitted to the Swedish data for ages above 20 is shown in Figure 3.6. The initial mortality rate $m_{1,0}$, parameter β_1 and initial fraction $\rho_{1,0}$ of the first (frailest) subpopulation are shown in panels A, C and E respectively. Similarly the evolution of mortality parameters of the second (most robust) subpopulation is shown in panels B, D and F. Although the parameters are widely dispersed and presented results contain high standard errors, the shown trendlines reliably indicate the direction of trends.

Figure 3.6 indicates that the inverse relationship between the initial mortality rate and the mortality coefficient is observed for both subpopulations. We therefore conclude that the compensation law holds for each subpopulation, although in the first subpopulation it is reversed as compared to that in the second subpopulation. Indeed, in the most robust (second) subpopulation the initial mortality declines over time while its mortality coefficient increases (similar results as in Section 3.5.1). At the same time the initial mortality of the frailest (first) subpopulation increases while its mortality coefficient decreases.
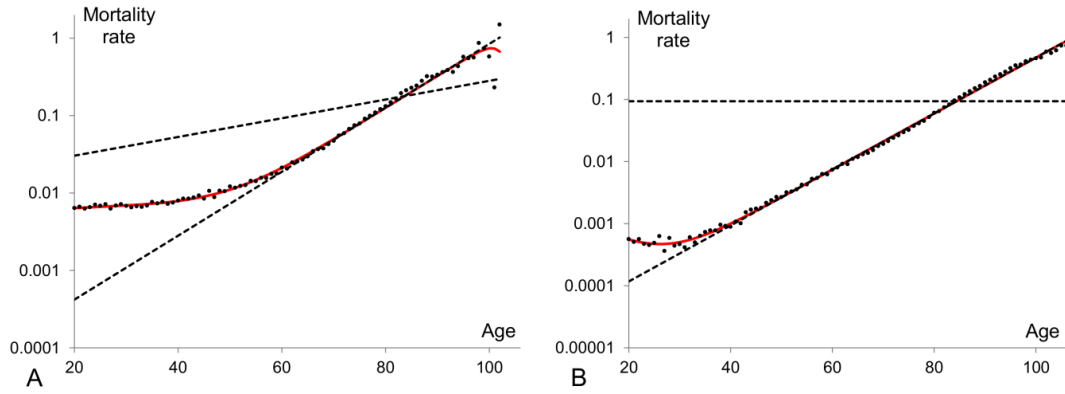
60

Figure 3.5: Heterogeneous model fitted to mortality data for ages 20+. *A two-subpopulation model is fitted to Swedish (males and females) data (dots) for the periods 1900 (panel A) and 2000 (panel B). Dashed lines represent the mortality rates of each subpopulation (exponential functions) while the solid (red) curve is the mortality of the whole population as described by the model of heterogeneous population. The model parameters as estimated by the Least Squares Method are for panel A: 1st subpopulation $m_{1,0} = 0.01742$, $\beta_1 = 0.0278$, $\rho_{1,0} = 0.28411$ and 2nd subpopulation $m_{2,0} = 0.000063$, $\beta_2 = 0.0951$, $\rho_{2,0} = 0.71589$ and for panel B: 1st subpopulation $m_{1,0} = 0.09361$, $\beta_1 = 0$, $\rho_{1,0} = 0.03114$ and 2nd subpopulation $m_{2,0} = 0.000015$, $\beta_2 = 0.10377$, $\rho_{2,0} = 0.96886$.*

It is also interesting to note on Figure 3.6 that the frailest subpopulation shows an important volatility especially for more recent years, when this subpopulation does not considerably affect the population mortality pattern above age 20. Specifically, the fraction of the frailest subpopulation becomes very small (close to zero) from around 1960. Thus, the mortality of the entire population is mainly reflecting the mortality of the most robust subpopulation. The parameter values of the frailest subpopulation do not have an important impact on the entire population mortality schedule, and thus, a wide range of parameter values provides a good fit.

A new and interesting observation related to the structure of heterogeneous populations can be drawn from panels E and F in Figure 3.6. The fraction $\rho_{1,0}$ of the frailest subpopulation declines exponentially over time while the fraction $\rho_{2,0}$ of the most robust subpopulation increases accordingly. The frailest subpopulation represented 0.28% of the total population in 1900, while it only represented 0.03% in 2000. The population is then becoming more homogeneous over time. This important observation will be further discussed in the upcoming sections.
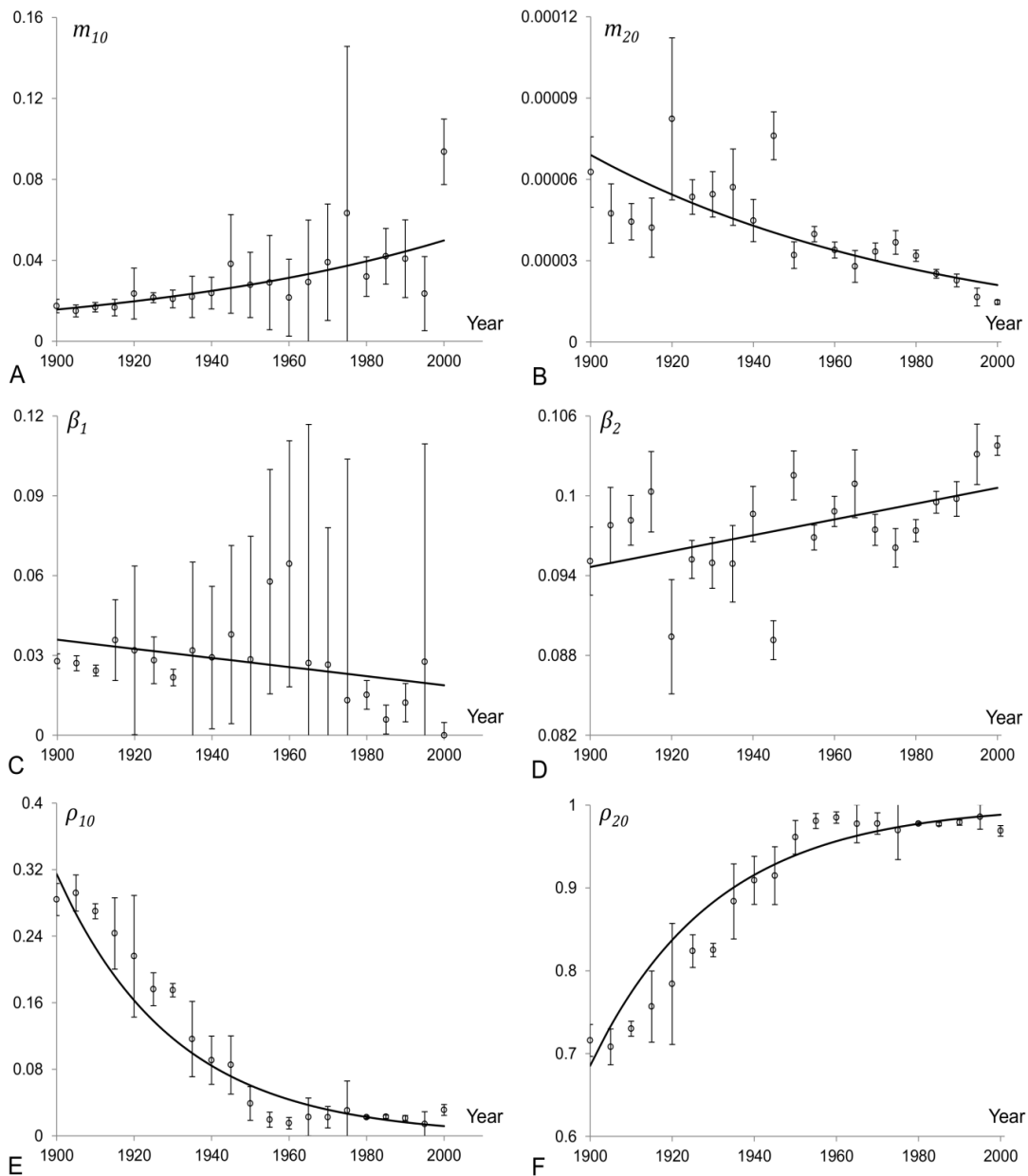
Figure 3.6: Evolution of the parameters of a two-subpopulation model fitted to the Swedish mortality rates for ages 20+. Panels A, C and E show evolution of parameters (initial mortality (m_0), mortality coefficient (β) and initial fraction (ρ_0)) associated with the first subpopulation while the panels B, D and F – with the second. The circle points with error bars correspond to the model parameters and their standard deviations as estimated by the Least Squares Method and the solid curves show trendlines for the evolution of the parameters over time.

3.5.3. Evolution of mortality parameters for all ages excluding the extrinsic causes of death

As mentioned in Section 3.4, the causes of death can be grouped into two categories: intrinsic and extrinsic. The intrinsic causes are related to the inability of biological

62

organisms to oppose destruction. Makeham (Makeham, 1867) suggested, among others, that the Gompertz law fits much better with mortality due to biological causes, this section focuses on the evolution of mortality due to intrinsic causes. Since the extrinsic causes of death include the external causes of death, the local maximum in the adulthood period (at approximately age 20 for the Swedish data in Figure 3.1), that is, the accidental hump is removed. Results in this section are based on mortality rates for Swedish males found on the WHO website, as mentioned in Section 3.4.
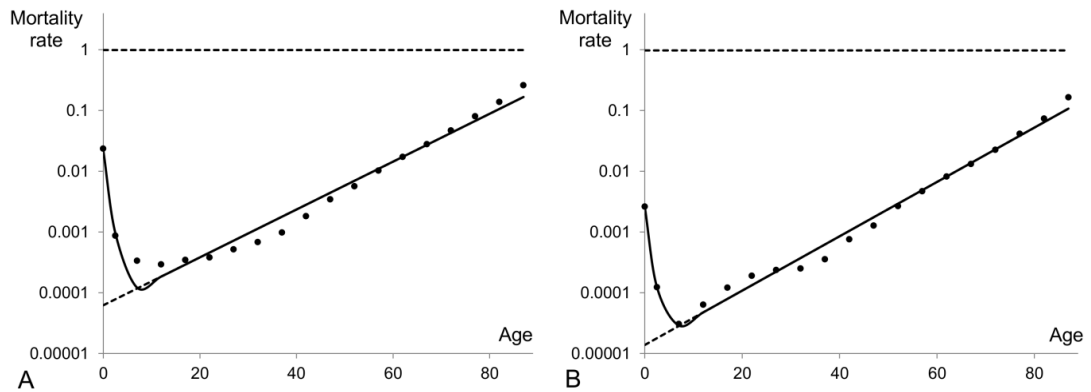
Figure 3.7: Heterogeneous model fitted to mortality data which exclude extrinsic causes of death. *A two-subpopulation model is fitted to Swedish male data (dots) for the periods 1951 (panel A) and 2010 (panel B). Dashed lines represent the mortality rates of each subpopulation (exponential functions) while the solid curve is the mortality of the whole population as described by the model of heterogeneous population. The model parameters as estimated by the Least Squares Method are for panel A: 1st subpopulation $m_{1,0} = 0.99013$, $\beta_1 = 0$, $\rho_{1,0} = 0.03324$ and 2nd subpopulation $m_{2,0} = 0.000062$, $\beta_2 = 0.0908$, $\rho_{2,0} = 0.96676$ and for panel B: 1st subpopulation $m_{1,0} = 0.97567$, $\beta_1 = 0$, $\rho_{1,0} = 0.00395$ and 2nd subpopulation $m_{2,0} = 0.000014$, $\beta_2 = 0.103$, $\rho_{2,0} = 0.99605$.*

Figure 3.7 presents observed Swedish male mortality (dots) excluding extrinsic causes of deaths for the periods 1951 (panel A) and 2010 (panel B). The high level of infant mortality reflects mainly the deaths due to severe birth defects, malformations, preterm births, the sudden infant death syndrome, etc. Therefore the mortality trajectory drops down to a minimum point and after the age of 10 increases exponentially. Thus, by excluding the accidental hump we observe that the exponential rise of mortality becomes apparent at the early stage of human life, just after age 10, and the pattern of mortality has a single minimum at that age. Therefore, a model of a heterogeneous population composed of two subpopulations should be sufficient to reproduce the actual data. Our studies show that a two-subpopulation model is indeed commonly a best fit for these data, although in some cases BIC values indicate that three-subpopulation models are

more accurate. In this section we will fit all the consecutive periods with the two-subpopulation model for consistency. In the two-subpopulation model (see Figure 3.7), the mortality dynamics of the first subpopulation explains the high infant mortality and the mortality decline at young ages, while the second subpopulation describes the exponential rise of mortality after the age of 10.

The second subpopulation represents a bigger proportion of the total population than the first subpopulation. The first subpopulation has usually a zero-slope coefficient and a high initial mortality, close to 1. Thus, the first-subpopulation individuals die at the very young ages due to the high level of initial mortality. By the age of 10, the first subpopulation is completely eliminated. Consequently, after the age of 10, the mortality rate of the entire population increases exponentially according to the mortality dynamics of the second subpopulation.

The time evolution of parameters of the two-subpopulation model is shown in Figure 3.8. The fitted model parameters are presented with circle points with error bars, while the solid lines represent trendlines. We found that while the model parameters $m_{1,0}$ and β_1 describing the first subpopulation do not show any trend, the evolution of the other model parameters follows a trend. Although the estimation of model parameters comes with considerably large error bars the trendlines can be reliably approximated by linear or exponential functions. The frailest subpopulation has a zero-constant mortality coefficient (Figure 3.8C) and an approximately constant (slightly decreasing) initial mortality (Figure 3.8A). The second subpopulation has an exponentially decreasing initial mortality (Figure 3.8B) and an exponentially increasing mortality coefficient (Figure 3.8D). The fraction of the first subpopulation declines exponentially over time (Figure 3.8E), while the fraction of the second subpopulation is increasing accordingly (Figure 3.8F). The first subpopulation can be considered as static, with a constant initial mortality and a constant mortality coefficient. Only its proportion to the entire population is changing over time. An analysis of the compensation effect requires then an analysis of the mortality dynamics for the second subpopulation only.

Similarly to the case of the 20+ data shown in Figure 3.6 we observe the decline of the initial mortality rate of the second (most robust) subpopulation and the increase of its parameter β and thus the inverse relationship between the two parameters is again observed, reflecting the compensation law of mortality.

64
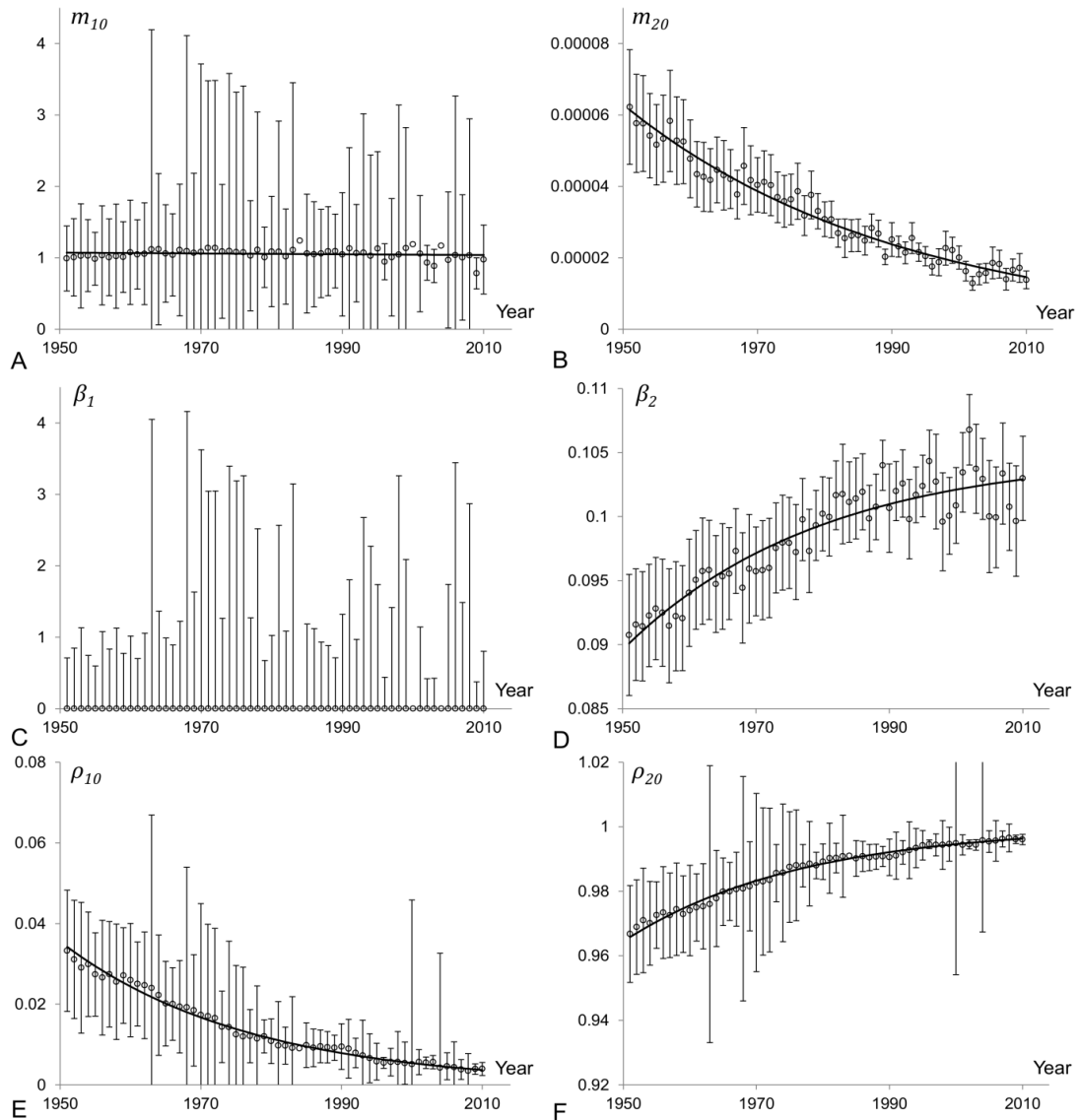
Figure 3.8: Evolution of the parameters of a two-subpopulation model fitted to the Swedish male mortality rates for all ages and excluding extrinsic causes of death. Panels A, C and E show evolution of parameters (initial mortality (m_0), mortality coefficient (β) and initial fraction (ρ_0)) associated with the first subpopulation while the panels B, D and F – with the second. The circle points (with error bars) give the model parameters as estimated by the Least Squares Method and the solid curves show trendlines for the evolution of the parameters over time.

As in the previous section, the heterogeneity in human populations is decreasing. The fraction $\rho_{1,0}$ declines over time and tends to zero while the fraction $\rho_{2,0}$ converges to 1. This could be interpreted as a reduction in the number of vulnerable new-borns who fail to survive in the new environment. This reduction can be viewed as a consequence of hygiene, medical and lifestyle improvements, etc. Since these individuals are no longer affected by fatal diseases at early ages, they are transferred to the most robust

65

subpopulation and thus add to the mortality dynamics in the same way as other individuals in the most robust subpopulation.

The decrease in human heterogeneity leads to a very interesting observation: the decline in mortality rates of the entire population is partly due to the decreasing proportion of the population related to the first subpopulation. Indeed, the change over time in the structure of the population, that is the change in the fractions of the first and second subpopulations, explains most of the mortality decline of the entire population. Most of the past mortality decline is thus not due to a decline in the mortality of each subpopulation (reflected with changes in the mortality parameters), but due to a change in the structure of the population. This result will be further discussed in the following sections.

3.5.4. Evolution of mortality parameters for all ages including all causes of death

In this section we will consider the model of heterogeneous populations as fitted to the mortality data on the entire age range including all causes of death. Swedish data for 1900-2000 have been used, as in Sections 3.5.1 and 3.5.2. According to BIC values a model consisting of four subpopulations is typically a best fit for these data as shown in Chapter 2. Figure 3.9 shows the fitted model for the first and the last periods under observation, that are 1900 (panel A) and 2000 (panel B).

The first subpopulation (with the highest initial mortality) describes the high infant mortality of the entire population and the deep mortality decline over the first few years. The second subpopulation has an impact on mortality in the age range from two (when first subpopulation is almost gone) to 10 when this subpopulation has also practically vanished. The third subpopulation describes the accidental hump which occurs due to the accidental mortality for young adult males and the accidental in addition to maternal mortality for young adult females. The last subpopulation (with the lowest initial mortality but the biggest initial size) explains the exponential mortality trajectory of the entire population after the reproductive period.

Figure 3.9 illustrates that the parameters of the four subpopulations are changing over time. Indeed, the points of intersections and the slopes of the dashed lines are different in panels A and B. Details are provided in Figure 3.10, where the time evolution of each

66

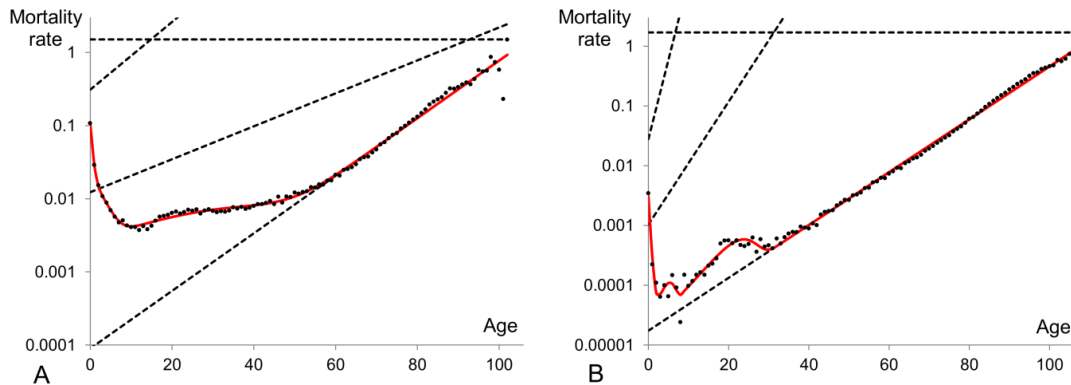parameter is shown. The plot has four columns, each column presenting the parameters of one subpopulation.

Figure 3.9: Heterogeneous model fitted to mortality data including extrinsic causes of death. *A four-subpopulation model is fitted to Swedish (males and females) mortality rates (dots) for the periods 1900 (panel A) and 2000 (panel B). Dashed lines represent the mortality rates of each subpopulation (exponential functions) while the red solid curve is the mortality of the whole population as described by the model of heterogeneous population. The model parameters as estimated by the Least Squares Method are for Panel A: 1st subpopulation $m_{1,0} = 1.50935$, $\beta_1 = 0$, $\rho_{1,0} = 0.09936$, 2nd subpopulation $m_{2,0} = 0.31023$, $\beta_2 = 0.10653$, $\rho_{2,0} = 0.0536$, 3rd subpopulation $m_{3,0} = 0.01229$, $\beta_3 = 0.05188$, $\rho_{3,0} = 0.17929$ and 4th subpopulation $m_{4,0} = 0.00009$, $\beta_4 = 0.09065$, $\rho_{4,0} = 0.66775$ and for Panel B: 1st subpopulation $m_{1,0} = 1.69434$, $\beta_1 = 0$, $\rho_{1,0} = 0.00331$, 2nd subpopulation $m_{2,0} = 0.02771$, $\beta_2 = 0.61132$, $\rho_{2,0} = 0.00028$, 3rd subpopulation $m_{3,0} = 0.00103$, $\beta_3 = 0.23672$, $\rho_{3,0} = 0.0045$ and 4th subpopulation $m_{4,0} = 0.000017$, $\beta_4 = 0.10179$, $\rho_{4,0} = 0.99191$.*

Figure 3.10 shows that the initial mortality of the first subpopulation increases while the initial mortalities of the other three subpopulations decrease over time (provided linear trends are not particularly accurate but confidently indicate the increasing/decreasing behaviour of data). The mortality coefficient of the first subpopulation (parameter β_1) is zero for almost all years. Since the parameter β_1 only affects the first few years of life, its value does not significantly influence the mortality pattern of the entire population (any value of the mortality coefficient combined with high level of initial mortality can reproduce the sharp initial decline of the mortality pattern). The mortality coefficients of the other three subpopulations (parameters β_2, β_3 and β_4) increase (approximately linearly) over time, which (taking into account the decrease in the initial mortality) confirms the validity of the compensation law of mortality for the last three subpopulations.
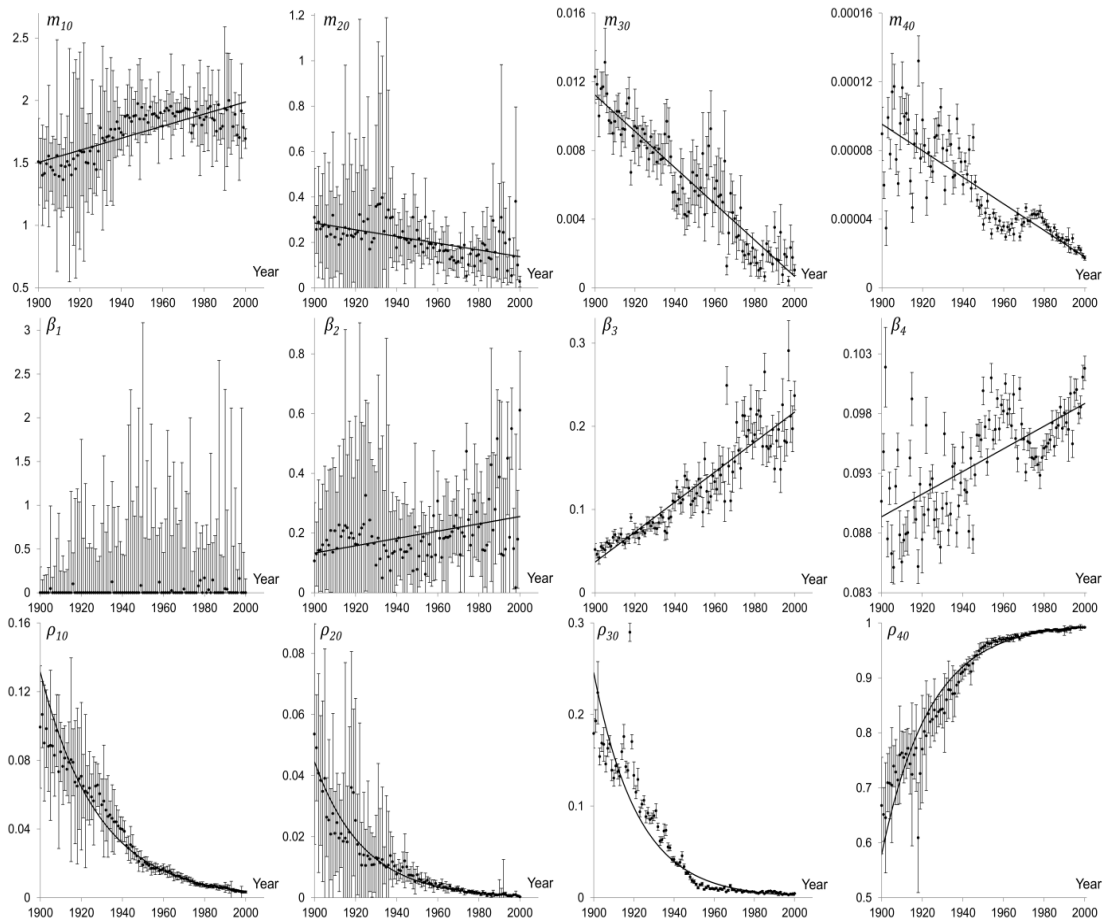
Figure 3.10: Evolution of the parameters of the four-subpopulation model fitted to Swedish mortality for each year over the periods 1900-2000. *Each of column shows evolution of parameters (initial mortality (m_0), mortality coefficient (β) and initial fraction (ρ_0)) associated with one of the subpopulations. The dot points (with error bars) give the model parameters as estimated by the Least Squares Method and the solid curves show trendlines for the evolution of the parameters over time.*

Figure 3.10 also reveals a very interesting observation, namely damped oscillations of initial mortality, $m_{4,0}$, and mortality coefficient, β_4, for the fourth subpopulation. These oscillations can reflect the effect of periodically changing external (e.g. climatic) factors on the mortality of the population, while their damping may indicate the evolution of the population's resistance to these factors. Finally, the homogenisation effect is also shown in Figure 3.10. As in previous sections, the most robust subpopulation (fourth subpopulation) is continuously growing and becoming a more important fraction of the total population, while the initial fractions of the first three subpopulations decline exponentially over time. At the year 1900, the proportion of the main subpopulation was 67% of the total population while this proportion increased to 99% in the year 2000.

3.5.5. Compensation effect
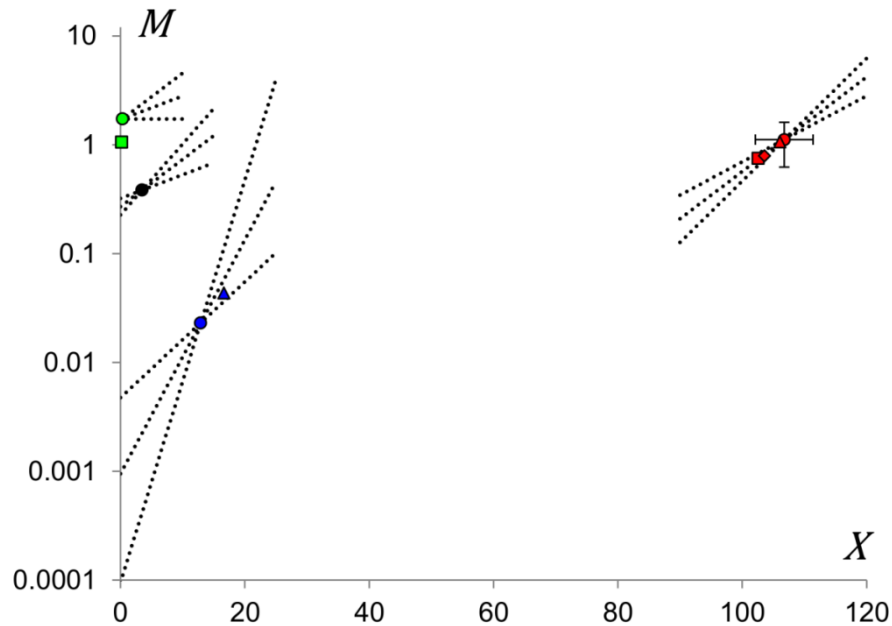
Figure 3.11: Compensation effect in heterogeneous models. The coordinates of coloured markers give the target lifespan and target mortality for each subpopulation of each model considered in Sections 3.5.1-3.5.4. The rhombus (red) marker corresponds to the homogeneous population fitted to the Swedish mortality from age 40 and above. The two triangle markers (red and blue) correspond to the subpopulations of the two-subpopulation model representing the Swedish mortality from age 20 and above. The two square markers (red and green) correspond to the subpopulations of the two-subpopulation model fitted to the Swedish male mortality excluding the extrinsic causes of death. The four circle markers correspond to the four subpopulations of the model fitted to the entire Swedish dataset. The colour used to draw a marker is the same for the subpopulations in different models having an impact in the same age interval: green is for infant mortality, black for child mortality, blue for the accidental hump and red for the exponential growth of mortality after the reproductive period. Error bars representing the standard deviations for the coordinates of the marker representing the most robust subpopulation in four-subpopulation model are shown. Error bars for the three other subpopulations of this model cannot be shown since they are smaller than the size of markers used in the figure. Sample mortality trajectories are presented in each intersection point of the four-subpopulation model to show the convergence of mortality at these points.

The inverse relationship between the time evolution of the initial mortality and the time evolution of the mortality coefficient was observed in previous sections for most subpopulations. A complementary phenomenon is the convergence of the mortality trajectories for each subpopulation. This convergence is manifested by the intersection of mortality trajectories (see Figure 3.4B) which, in the ideal case (a pure form of compensation effect), takes place in the same point for all trajectories, the coordinates of this point giving the target lifespan and target mortality, i.e. the age and the level of

69

mortality at which the last survivors in the subpopulation of interest die (Strehler and Mildvan, 1960, Strehler, 1978, Yashin et al., 2000, Gavrilov and Gavrilova, 1979, Gavrilov and Gavrilova, 1991, Gavrilov and Gavrilova, 2006). The target lifespans for the individuals of each subpopulation are found through the Strehler and Mildvan correlation (equation (1.4)). The target lifespans resulting from the preceding four sections are estimated as:

1. For homogeneous population fitting the Swedish data above age 40 (Section 3.5.1): 103.6 years;
2. Two subpopulations reproducing the Swedish mortality patterns for ages above 20 (Section 3.5.2): 16.7 and 106 years respectively;
3. Two subpopulations reproducing the Swedish male mortality patterns excluding the extrinsic mortality (Section 3.5.3): 0.2 and 102.6 years respectively;
4. Four subpopulations representing Swedish entire mortality schedule (Section 3.5.4): 0.3, 3.5, 12.9 and 106.8 years respectively.

Figure 3.11, showing the points of intersections for mortality trajectories of different subpopulations on the plane mortality/lifespan, summarizes these results. The intersection points are given by markers whose colours are the same for the "same" subpopulations in the four fitted models which can be distinguished by the shape of the markers: triangles for the two subpopulations representing the Swedish mortality at age 20+; squares for the two subpopulations reproducing the Swedish male mortality excluding the extrinsic death factors; circles for the four subpopulations reflecting the Swedish mortality for all ages and causes. For example, the first subpopulation of the model representing Swedish mortality at age 20+ explains the decline of mortality in the right tail of the accidental hump and thus, the intersection point for it has the same colour (blue) as the point for the third subpopulation of the four-subpopulation model which is also responsible for the accidental hump. Similarly the lifespan of the first subpopulation, reflecting the Swedish male mortality excluding the extrinsic causes of death, explains the sharp initial decline in the mortality pattern, and thus the intersection point for it has the same colour (green) as the point for the first subpopulation in the four-subpopulation model for Swedish mortality (green square and green circle points respectively in Figure 3.11). Finally the red colour represents the exponential rise of mortality after sexual maturity and thus the homogeneous population representing the Swedish mortality for age 40 and above is shown by the red rhombus, the second

subpopulation of the model representing the Swedish mortality at age 20+ is shown by the red triangle, while the second subpopulation reflecting the Swedish male mortality excluding the extrinsic causes of death is represented by the red square point, and the fourth subpopulation in the four-subpopulation model is represented by a red circle point. Samples of hypothetical mortality dynamics for each subpopulation of the four-subpopulation model are given by dashed lines.

The most interesting observation from Figure 3.11 is that the markers of the same colour are located close to each other. Thus, the four models studied in Sections 3.5.1-3.5.4, even if they were applied to different age ranges and datasets, provide similar results by indicating the existence of almost identical subpopulations. Therefore, the target lifespan for the total population in Sweden is reflected through the target lifespan for the most robust subpopulation (which appears in all four models) and thus lies between ages 102 and 107.

3.5.6. The role of homogenization in the evolution of mortality dynamics

Previous results lead to the following crucial conclusion: the reduction of mortality over time is not only affected by the change of mortality dynamics in each subpopulation but it is also a consequence of the change in the structure of the population. In Figure 3.12 the patterns formed by the model fitted to the 1900 (solid red curve) and 2000 (solid blue curve) Swedish data are shown in the same semi-logarithmic plot. The dashed blue curve in between them depicts an artificial pattern that is produced with the model composed by four subpopulations, using the initial mortalities and mortality coefficients obtained by fitting the 1900 period data and the initial fractions of the subpopulations reproducing the 2000 period data. It is then apparent that a reduction of mortality within one century is a result of

1. the alteration in population structure, that are changes in subpopulation's fractions, especially at young ages (difference between solid red and dashed blue curves in Figure 3.12) and

2. the alteration of the exponential dynamics of the subpopulations, that are changes in initial mortalities, $m_{j,0}$ and mortality coefficients, β_j (difference between dashed blue and solid blue curves in Figure 3.12).

Looking at Figure 3.12 we can also conclude that the decrease in mortality at younger ages is mostly due to the homogenization of the population while at older ages is entirely due to changes in the mortality parameters of the most robust subpopulation.
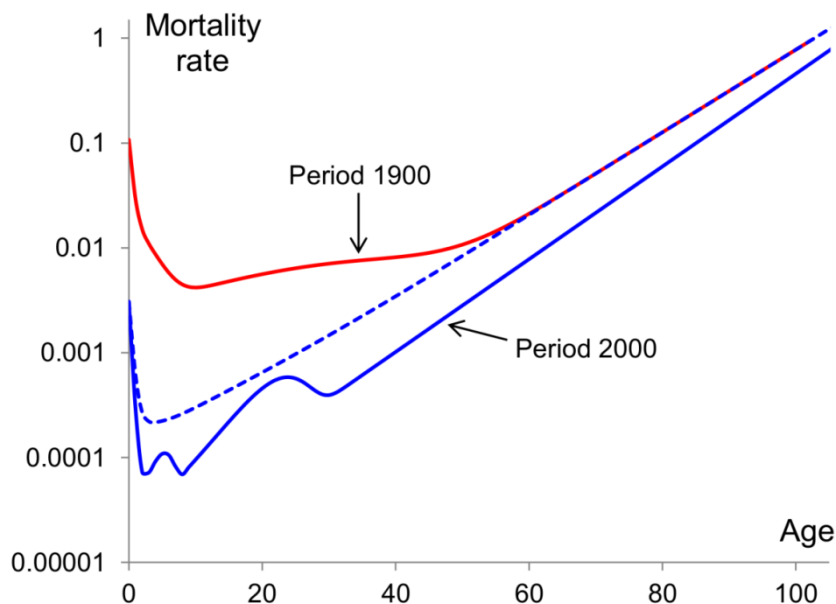
Figure 3.12: Reduction in the Swedish mortality rates within one century (from 1900 to 2000). *The patterns formed by the four-subpopulation model fitted to the 1900 (red solid curve) and 2000 (blue solid curve) Swedish mortality rates are plotted in a semi-logarithmic plot. The dashed blue curve indicates an artificial pattern formed by the four-subpopulation model with the initial mortalities and mortality coefficients estimated by fitting the 1900 period data and the fractions estimated by fitting the 2000 data.*

3.6. Discussion

Investigations of human mortality dynamics have practical implementations as they may help to find ways to increase our lifespan. These investigations also have a fundamental value as they help to understand biological processes and genetics underlying the process of ageing. One of the important ways to conduct such investigations is represented by mathematical modelling. Various assumptions have to be made to design a mathematical model, especially if one aims to model all relevant features observed in the mortality pattern of the entire lifespan. One of the commonly used assumptions is that a population is heterogeneous and composed of several subpopulations having different mortality dynamics (Rossolini and Piantanelli, 2001, Vaupel, 2010). The possible interpretation of the model parameters is extremely important to conduct deeper analyses and for forecasting purposes (Booth and Tickle, 2008). The time evolution of the parameters of a

model fitted to several observed periods provides then important insights in potential future evolutions. Such studies have been performed using various models of mortality dynamics (McNown and Rogers, 1989, McNown and Rogers, 1992, Lee and Carter, 1992, Bell, 1997, Tabeau et al., 2001, Felipe et al., 2002, Gaille and Sherris, 2011, Gaille, 2012, Njenga and Sherris, 2011).

In this Chapter we have analysed the evolution of the parameters in the model of heterogeneity of human populations where each subpopulation follows an exponential law of mortality (Avraam et al., 2013). We fitted the model to four different datasets. First, the homogeneous exponential law was fitted to Swedish mortality rates for ages above 40 for the periods 1900 to 2000. Second, the model of heterogeneous population with two subpopulations was fitted to the Swedish mortality rates for ages 20 plus. Third, the two-subpopulation model was fitted to Swedish male mortality rates over the entire lifespan, excluding deaths due to extrinsic factors, for the periods 1951 to 2010. Fourth, the model of heterogeneous population with four subpopulations was fitted to the entire lifespan dataset of the Swedish population for the periods 1900-2000. Our model fitting approach results in four main findings.

The first remarkable observation concerns the model used, that is the best fit to the mortality dynamics for the most complete mortality data is given by the four-subpopulation model. The novel result associated with this model is given by the second subpopulation which distinguishes the impact of child mortality from infant mortality to the entire mortality pattern. Occurrence of the subpopulation which counts for the child mortality makes our model different from other notable models (such as Heligman-Pollard) where the mortality dynamics is commonly decomposed to only three stages corresponding to early childhood, accidental mortality and late-life adulthood.

Second, our model does not capture the "late-life mortality plateau" which was reported by many researches as described in the Introduction section. We have fitted the model to each year within the 20th century and only a few of the fits have captured the mortality deceleration at older ages. It turns out that the deviations in mortality dynamics from the exponential increase at older ages are not always significant enough to be captured by our model. As we use the BIC for evaluation of how the model fits the data, the best fit becomes very sensitive to the number of subpopulations in the model. For example five-subpopulation model is worse than four-subpopulation as it has more model parameters

although in terms of standard error it fits data better. To analyse the dynamics of mortality at old ages it is more appropriate to consider a shorter range of ages for example 80+. We have checked and confirm that the best fit for this range of ages is commonly given by two subpopulations and the model reproduces mortality deceleration around the age of 100 (see Chapter 4).

Third, the compensation law of mortality is confirmed at the subpopulation level. The inverse relationship (negative correlation) between the mortality parameters is shown with the reduction of the initial mortality and the increase of the mortality coefficient over time in almost all subpopulations. The only exception is regarding the frailest subpopulation, which has an approximately constant initial mortality and a constant mortality coefficient. The frailest subpopulation, namely the subpopulation with the highest initial mortality rate, represents a small proportion of the entire population and disappears after a couple of years (the individuals belonging to that subpopulation die a few years after their birth). Therefore, except for the very young ages, this subpopulation has a negligible impact on the mortality pattern. It is also interesting to note by comparing Figure 3.6 and Figure 3.8 that the evolution of the parameters of the two-subpopulation model for ages above 20 (Section 3.5.2) is similar in many ways to the evolution of the parameters of the two-subpopulation model excluding the extrinsic causes of death (Section 3.5.3). It is related to the fact that in both models, the most robust subpopulation explains the exponential rise of mortality over the adulthood period in the entire population and the frailest subpopulation describes a decline in the mortality pattern at young ages. Indeed, in the two-subpopulation model for ages 20+, the frailest subpopulation represents the decline forming the right tail of the accidental hump (Figure 3.5), while in the two-subpopulation model excluding extrinsic mortality, the frailest subpopulation is responsible for the initial decline of mortality at very young ages (Figure 3.7).

Our forth main finding is the homogenization of the population over time. Indeed, we have shown that the fractions of all subpopulations except for the most robust decline over time. We have found that the fraction of the most robust subpopulation gradually increases being equivalent to 67% of the total population in 1900 and 99% in 2000 for the Swedish data. The homogenisation we report here is related to the evolution of mortality in developed countries and does not directly reflect the variations in genotype in the population. It rather reflects the fact that in course of time, with improvements in

medical service and life conditions, the variations in genotype of people become less important in terms of their duration of life. Contemporary genetic studies (Cavalli-Sforza and Feldman, 2003) indicate that the genetic variability in human populations is increasing over time. This increase is rather explained by the fact that the mortality is gradually reducing and therefore not all of those who recently survive and give offspring (i.e. bring diversity into a gene pool) would be able to do so in the past.

In a view of the outlined above finding we can state that the mortality decrease over the last century can be decomposed into two components: first, a mortality decline due to changes in the structure of the population or its homogenization (decrease from the red solid curve to the blue dashed curve in Figure 3.12) and second, a decrease due to a mortality reduction in the subpopulations (decline from the blue dashed curve to the blue solid curve in Figure 3.12), which is reflected with a change in the mortality parameters of each subpopulation. The implications are remarkably important for potential future mortality improvements: once the homogenization process is over, that is the mortality of the entire population will only reflect the mortality of the healthiest subpopulation, the potential for future decrease in mortality will be relatively small compared to what we observed over the last century. New developments in mortality forecasting approaches should consider this aspect to avoid overestimation of future mortality improvements.

To conclude, the above findings can be used in a future work to further enhance our methodological approach. It would be of great interest to fit a mortality surface (rather than a line) over the plane given by two variables: age and time (see Chapter 4). Fitting a surface is a difficult task requiring further assumptions about its structure. This structure can now be postulated as based on the assumptions that the evolutions of initial mortality, m_0, is represented by a linear function while the evolution of mortality coefficient, β, by an exponent. This approach would allow us to obtain further results on the mortality structure of human populations, and thus could confirm, enhance and develop further the results presented in this Chapter.

Chapter 4.

On the heterogeneity of human populations as reflected by mortality dynamics

4.1. Summary

The heterogeneity of human populations is a common consideration in describing and validating their various age-related features. Heterogeneity, in particular, amongst other factors, is used to explain the variability of mortality rates across the lifespan and deviations from an exponential growth at young and very old ages. The mathematical model that combines the population heterogeneity with the assumption that the mortality of each constituent subpopulation increases exponentially with age, has been shown to successfully reproduce the entire mortality pattern across the lifespan as well as its evolution over time. Furthermore, the analysis of time-evolution of the mortality pattern, performed by fitting the model to actual data of consecutive periods, confirms the applicability of the compensation law of mortality to each subpopulation and concludes on the evolution of the population towards homogenisation.

In this Chapter we aim to show that the heterogeneity of human populations is not only a convenient consideration for fitting mortality data but is indeed the actual structure of the population as reflected by the dynamics of its mortality over age and time. In particular, we demonstrate that the model of heterogeneous populations fits mortality data better than most of the other models if the data are taken for the entire lifespan and better than

all other models if we consider only old ages. Also, we show that the model can reproduce seemingly contradicting observations in late-life mortality dynamics namely deceleration, levelling-off and mortality decline. Assuming that heterogeneity is reflected in genetic variations within the population, using Swedish mortality data for 20th century we show that the homogenisation of the population, observed in the model fits, can be associated with the evolution of allele frequencies.

4.2. Introduction

One of the central problems in ageing-related studies is the understanding of processes which underlie senescence and revealing the biological functions which deteriorate in organisms over the course of their lifespan. Ageing is associated with the acceleration of mortality expressed by the exponential increase of mortality rate over age, as described by the Gompertz law. The Gompertz law represents a fundamental mortality law and was verified by demographic observations across different countries, different time periods, and even different species (Gavrilov and Gavrilova, 1991). The analysis of available data on mortality rates for various diseases also indicates that for most diseases there is a considerably wide age range where the mortality rate also increases exponentially (Jones, 1956, Finch, 1994).

The exponential growth of mortality is not observed at young (before sexual maturity) and extremely old ages. Many researchers consider the exponential law of mortality to be "natural" while the deviations from it need to be explained. The observation that the exponential law of mortality does not apply at older ages was first made by Gompertz. In his 1825 paper, Gompertz stated that "*The near approximation in old age, according to some tables of mortality, leads to an observation, that if the law of mortality were accurately such that after a certain age the number of living corresponding to ages increasing in arithmetical progression, decreased in geometrical progression, it would follow that life annuities, for all ages beyond that period, were of equal value; for if the ratio of the number of persons living from one year to the other be constantly the same, the chance of a person at any proposed age living to a given number of years would be the same, whatever that age might be;*" recognising that the probability of surviving (and consequently, the mortality rate) levels-off at extremely old ages.

The divergence of mortality from its exponential increase at extremely old ages is generally accepted as valid by the majority of biologists and bio-demographers (Greenwood and Irwin, 1939, Olshansky, 1998, Gavrilov and Gavrilova, 2011) although some data do not support this observation (see for example (Gavrilova and Gavrilov, 2014)). Furthermore, existing data on mortality dynamics at advanced ages are so controversial (Olshansky, 1998) that performed studies have not given a definite answer on what mathematical function (logistic, quadratic, etc) can describe the data at those ages (Gavrilov and Gavrilova, 2001, Kannisto et al., 1994, Pham, 2011).

The absence of a definite explanation for the trend of mortality at very old ages renders the analysis of oldest-old mortality as an essential topic of demography. Raw mortality data are usually statistically noisy at extreme old ages as the number of survivors is small. The stochastic effects at very old ages are often seen as fluctuations in mortality dynamics (Avraam et al., 2013). On the other hand, different observations have been made for the trends of mortality behind these fluctuations. The late-life mortality slow-down phenomenon describes the divergence of mortality from the exponential law at very old ages. Late-life mortality slow-down could be associated with (a) an increase in late-life mortality at a slower rate than its exponential increase during the adulthood period, called deceleration (Depoid, 1973, Horiuchi and Wilmoth, 1998, Thatcher et al., 1998), (b) a levelling-off, commonly called the mortality plateau, which is the saturation of mortality trajectory on a constant horizontal rate asymptote to a certain limit (Economos, 1979, Mueller and Rose, 1996, Curtsinger et al., 2006) or (c) a decline of mortality with increasing age (Kannisto et al., 1994, Wilmoth, 1995). The late-life mortality slow-down was observed for human (Greenwood and Irwin, 1939, Bebbington et al., 2014) as well as non-human populations (Carey et al., 1992, Pletcher and Curtsinger, 1998, Economos, 1979, Economos, 1980).

The deviations in mortality from the exponential law, at young and older ages, segregate the analysis of mortality into two parts. The first part attempts to explain the biological processes underlying the exponential law, and the second part tries to explain the causes of deviations from the exponential law that are observed at young and extremely old ages. Mathematical verifications of the exponential law have been performed from different points of view ranging from a population genetic theory of ageing (Charlesworth, 1994) to the application of reliability theory to ageing and longevity (Gavrilov and Gavrilova, 2001).

78

The evolutionary theories that were proposed to explain ageing, and to answer the question of why organisms grow old and die, are mostly based on the assumption of a loss of selective significance of phenotypes developing during post-reproductive ages (Hamilton, 1966, Charlesworth, 2000, Rose et al., 2007). One such theory, first introduced by Medawar, is known as the mutation accumulation theory (Medawar, 1946, Medawar, 1952). Medawar's hypothesis states that gene alleles or mutations that are neutral at early life but deleterious at later life, escape natural selection and are transferred to the next generation before their deleterious effects become evident. Such mutations can therefore accumulate in the population by a genetic drift and reveal themselves via the diseases associated with the post-reproductive period. Another evolutionary theory, named the antagonistic pleiotropy, states that genes, with beneficial effects early in life but with deleterious actions later in life, could be favoured by selection and accumulate in the population (Williams, 1957). The theory of disposable soma, proposed by Kirkwood (Kirkwood, 1977, Kirkwood and Holliday, 1979), postulates that organisms have a limited amount of energy and that specific gene mutations save energy for reproductive aspects by reducing the amount of energy used for maintenance, leading to non-reproductive damages. Ageing is therefore a result of the accumulation of damages that are not repaired by the organism (Kirkwood and Austad, 2000). These hypotheses take for granted the fact that the length of the reproductive period itself may depend on a number of genetic determinants associated with environmental and population factors.

Other attempts (not based on evolutionary theories) to explain the Gompertz law include work by Strehler and Mildvan, who related mortality to inadequate responses of the organism to energy demands and showed that the exponential increase of mortality is associated with a linear decrease of vitality (where vitality was defined as the capacity of an individual to resist damage) (Strehler and Mildvan, 1960), work by Sacher and Trucco, who analysed the role of stochastic and homeostatic forces (Sacher and Trucco, 1962, Yashin et al., 2000), and a study by Shklovskii who modelled the exponentially rare escape of abnormal cells from immunological response (Shklovskii, 2005). Furthermore, Gavrilov and Gavrilova have applied the reliability theory to explain ageing and the Gompertz law by considering age-related failure kinetics of systems (machines) and their components (Gavrilov and Gavrilova, 2001). They have shown that the rate of machines' failure as a function of age can reproduce the known mortality laws

(Gompertz law, compensation effect, late-life deceleration) and therefore the reliability theory can be used to explain biological ageing.

Many mathematical models have been introduced to reproduce the observed mortality patterns (Gompertz, 1825, Makeham, 1860, Thiele, 1872, Siler, 1979, Heligman and Pollard, 1980). Some of them are designed to generate mortality patterns over the entire lifespan while others aim to reproduce a specific part of those patterns. For example, a function that outlines an inverse relationship between mortality and age, is used to generate the decline of mortality at very young ages (De Beer and Janssen, 2014) whereas the logistic-type and quadratic functions are used to create the late-life mortality plateau and the late-life mortality decline, respectively (Gavrilov and Gavrilova, 2001, Kannisto et al., 1994, Pham, 2011). A number of studies consider populations to be heterogeneous and model their mortality dynamics by assuming that the population is comprised of cohorts (subpopulations) such that the members of each cohort, at a given age, face the same probability of death (Vaupel and Yashin, 1985a, Vaupel et al., 1979, Manton et al., 1986). The concept of heterogeneity was used in Chapter 2, to model the dynamics of mortality and explain their deviations from the exponential growth at young and very old ages (see also (Wrigley-Field, 2014, Chen et al., 2013, Drapeau et al., 2000, Steinsaltz, 2005)).

Since the model of heterogeneous populations, introduced in Chapter 2, turned to be extremely useful for analysis of mortality data, there arises a question of whether this model reflects the real structure of the population and, if yes, what quality underlies the heterogeneity? This Chapter is devoted to the above question. We start by comparing the mortality model of heterogeneous populations with a wide-range of other commonly used parametric models and show that it is one of the best models when judgements are made on the basis of the quality of fit to observed data. We then show that the model can explain the apparent controversial observations for old-age mortality (deceleration, mortality plateau and decline) which are not in contradiction with one another, but reflect a similar and coherent process underlined by the heterogeneity of populations. Finally, we tackle the problem of the nature of a population's heterogeneity. Although the heterogeneity of populations can be conditioned by various factors such as disparities in life-style, environmental and other socio-economic conditions, the responses to environmental factors are largely shaped by an organism's genetic landscape. Particular gene polymorphisms may be more important in terms of increasing an organism's fitness

80

risks with age and would affect the dynamics of an organism's ageing and mortality. In this Chapter we check whether the population dynamics of putative gene variations can be aligned with the modelled dynamics of suggested distinct subpopulations required for representation of the global dynamics of mortality of integrated human population. We assume that individuals belonging to different subpopulations differ by genotype and have differential resistance to environmental perturbations. Changes in the environment would favour different subpopulations in different contexts and their resultant differential mortality may have an impact on the dynamics of the mortality characteristic for a population in a given time period. In Chapter 3 it was shown that the evolution of mortality dynamics in Sweden over the 20th century was for two reasons: changes in mortality dynamics of subpopulations, and changes in the structure of populations as represented by the fractions made by subpopulations (homogenisation of the population). While changes in the mortality dynamics of subpopulations are most likely driven by environmental changes (Black et al., 2010, Ahmad et al., 2000), the change in the population structure can be explained in terms of population dynamics. Based on the difference in mortality dynamics of subpopulations (and assuming that the difference is due to a single gene) we have calculated their relative fitnesses and confirmed that this allows for an explanation and accurate reproduction of the homogenisation process of populations.

This Chapter is structured as follows. In Section 4.3 we introduce different mathematical models for mortality and describe the technique used for an evaluation of allele frequency dynamics in population genomics. We then show that the model of a heterogeneous population is one of the best models for fitting actual mortality data over the entire lifespan (Section 4.4.1) or for ages over 80 (Section 4.4.2). We emphasise that it can explain controversial data on late-life mortality (Section 4.4.3). Finally, we show in Section 4.4.4 that the homogenisation of the heterogeneous population as revealed by the evolution of mortality data in Sweden over the 20th century can be explained by changes in allele frequency due to different fitnesses corresponding to different subpopulations. We conclude with a discussion of the obtained results and provide further arguments on genotypic differences between subpopulations having different mortality dynamics in Section 4.5.

4.3. Mathematical modelling

In this section we give a description of a few popular models (most of them were introduced in Chapter 1) which are commonly used for fitting mortality data (these models later will be compared with each other in terms of their fits to a given set of data). For the scope of this analysis, we use so-called "parametric" models that are models expressing mortality rates across the lifespan with fixed (i.e. time-independent) parameters. We exclude therefore the models that consider also the time dependency of mortality patterns such as the notable Lee-Carter model. The description of parametric models is followed by a description of the method of calculating the force of natural selection in a population of diploid organisms which is later used for the analysis of the evolution of heterogeneous populations.

4.3.1. Models of mortality

Exponential functions: Gompertz, Makeham and Weibull

The first developed parametric model and the one that remains the most notable in literature is the Gompertz model (Gompertz, 1825). The Gompertz model which became a universal law describes the exponential increase of mortality with age in a significant portion of lifespan (from sexual maturity to extremely old ages). According to the Gompertz law, the central death rate at age x, m_x, is given by

$$m_x = \alpha e^{\beta x}, \tag{4.1}$$

where α is the initial mortality rate (scale parameter) and β is the rate of change of mortality with age (shape parameter). It is remarkable that the Gompertz law does not only hold for human populations but also for many other biological species (Gavrilov and Gavrilova, 1991).

An extension of the Gompertz law is the Makeham model (Makeham, 1860), which represents the death rate as the sum of an age-dependent component (the Gompertz function) describing deaths due to age-related diseases or disorders, and an age-independent component (a constant γ) describing deaths due to external factors such as accidents or certain infectious diseases:

$$m_x = \gamma + \alpha e^{\beta x}. \qquad (4.2)$$

A third exponential parametric model is the Weibull model (Weibull, 1939) which expresses the mortality rate as a power function of age:

$$m_x = \alpha x^{\beta}. \qquad (4.3)$$

According to these three exponential models, mortality rates diverge to infinity as age tends to infinity. The difference in concavity or convexity of these functions, and the difference in their initial values when $x = 0$ ($m_0 = \alpha$ for Gompertz, $m_0 = \gamma + \alpha$ for Makeham and $m_0 = 0$ for Weibull), distinguish them in terms of their usage. The Gompertz and Makeham models are generally used to describe the mortality of biological species while the Weibull function is widely used to describe the ageing and failure rate of technical systems and devices (Weibull, 1951, Le Bras, 2008).

Logistic functions: Perks, Beard and Kannisto

The logistic-type functions which shape sigmoid curves are commonly used in the analysis of mortality dynamics at older ages. These curves saturate, reaching a horizontal asymptote, and can therefore produce the late-life mortality plateau (Perks, 1932, Kannisto, 1992). The general form of a logistic curve is expressed as a four-parameter function:

$$m_x = \gamma + \frac{\alpha e^{\beta x}}{1 + \delta e^{\beta x}}, \qquad (4.4)$$

which is known as the Perks model.

Different variations of logistic function can be used in order to reduce the number of parameters. A three-parameter logistic function is formed by setting $\delta = \alpha$ in equation (4.4) or the three-parameter function introduced by Beard (Beard, 1971) by setting $\gamma = 0$. Also, a simple two-parameter logistic function used by Kannisto (Kannisto, 1992) is formed by setting $\gamma = 0$ and $\delta = \alpha$ in equation (4.4).

The logistic function in equation (4.4) saturates asymptotically to $\gamma + \alpha/\delta$ as age increases while the Beard function tends to the constant α/δ. The Kannisto model has an asymptote equal to one and this model is used in a common procedure for the

construction of life tables in order to smooth the noisy death rates observed at ages 80 and above (Thatcher et al., 1998).

The Gompertz and Makeham models could be considered as special cases of equation (4.4). If $\delta = 0$, equation (4.4) is transformed into the Makeham model and if $\gamma = \delta = 0$ – into the Gompertz law. However, in both these models, the mortality rate tends to infinity as age increases which is in contrast to the logistic-type functions and due to the elimination of the denominator from the logistic form.

Michaelis-Menten kinetics

Michaelis-Menten kinetics are an outcome of a well-known model in biochemistry that describes the dynamics of catalysed reactions (Michaelis and Menten, 1913). The kinetics are represented by an equation which describes the saturation of a reaction rate when the substrate concentration is increasing. The Michaelis-Menten equation has also been used to model several other processes, for example, Monod who was working in the field of environmental engineering used this equation to model the growth rate of microorganisms as a function of the nutrient's concentration (Monod, 1949). In this Chapter, we suggest using the Michaelis-Menten equation (disregarding its parameters and variables terminology) to fit mortality data and to be compared with other asymptotic mortality functions that reproduce the mortality levelling-off at very old ages (i.e. the logistic-type functions). Following the form of the Michelis-Menten equation, the mortality at age x can be expressed as:

$$m_x = \alpha \exp\left(\frac{\beta x}{1 + \gamma x}\right). \tag{4.5}$$

Exponential-Quadratic function

An exponential-quadratic function (known also as the Coale-Kisker model) is usually used to fit mortality data and show the deceleration of mortality rate and its decline at very old ages (Coale and Kisker, 1990). The exponential-quadratic function is given by

$$\ln(m_x) = \alpha + \beta x + \gamma x^2, \tag{4.6}$$

where for a concave down parabola with a maximum point, γ should be less than zero.

Heligman-Pollard model

84

Heligman-Pollard model (Heligman and Pollard, 1980) is an eight-parameter function that can reproduce mortality patterns of the entire lifespan with sufficient accuracy. The model was originally formulated for the ratio of death and survival probabilities (q_x/p_x) and composed of three terms where the first term reflects the sharp decline of mortality at childhood, the second reflects the accidental hump that is observed during the reproductive period (ages 15-40), and the third term (which is a Gompertz function) reflects the exponential increase of mortality at post-reproductive ages:

$$\frac{q_x}{p_x} = A^{(x+B)^C} + De^{-E(\log(x)-\log(F))^2} + GH^x. \tag{4.7}$$

The last term of the Heligman-Pollard model is usually modified to the logistic form $GH^x/(1 + GH^x)$ to allow the saturation of mortality at extremely old ages. In this Chapter, we use the Heligman-Pollard model to fit actual mortality rates (which are best approximated by the central death rate, m_x) instead of the ratio q_x/p_x.

Model of heterogeneous population

Mathematically, the model of a heterogeneous population, which postulates the exponential mortality dynamics for constituent subpopulations, expresses the mortality rate m_x at age x, as a sum of weighted exponential terms:

$$m_x = \sum_{j=1}^{n} \rho_{j,x} m_{j,x} = \sum_{j=1}^{n} \rho_{j,x} \alpha_j e^{\beta_j x} = \sum_{j=1}^{n} \rho_{j,x} m_{j,0} e^{\beta_j x} \tag{4.8}$$

where the sub-index j indicates the *j-th* out of n subpopulation, $m_{j,x}$ is the central death rate at age x of subpopulation j, α_j is the initial mortality rate of the *j-th* subpopulation, and β_j is its mortality coefficient which gives the rate of change of mortality with age (Avraam et al., 2013, Avraam et al., 2014). The weights $\rho_{j,x}$ are fractions formed by each subpopulation j at age x in the entire population, and their sum is equal to unity at all ages. Finally, the mortality rate at age 0 of the subpopulation j is equal to α_j and thus we have the relation $m_{j,0} = \alpha_j$, which leads to the last term in equation (4.8).

4.3.2. Model of natural selection

Natural selection is an evolutionary process taking place within a population and states that individuals with certain heritable traits have the ability to survive and reproduce

offspring more often than individuals' deficient in those traits. Since these traits are heritable, the proportion of individuals carrying genotypes that express these traits is gradually increasing over time. Hence, natural selection similarly to the other primary evolutionary forces (mutation, migration and genetic drift) causes changes in allele frequencies in a population. The ability of any individual to pass genes to the next generation is determined by fitness. The more likely an individual is, to survive and live long enough to mate and reproduce, the higher their fitness is. A measure of fitness can be given by an average number of offspring that are born from parents of a given genotype (Futuyma, 2013). Selection is therefore conditioned by the variation of fitness between different genotypes. A simple model of natural selection that counts the frequencies of alleles (and subsequently the number of individuals with specific genotypes) over discrete generations is described in this section.

A diploid gene with alleles A and B splits the population into three groups of individuals having three distinct genotypes: AA, AB or BB. The notations p and q are used to denote the frequencies of alleles A and B respectively and the notations P, Q and R are used to define the frequencies of genotypes AA, AB and BB, where $p + q = 1$ and $P + Q + R = 1$. After a single step of random mating the frequencies of the three genotypes are $P = p^2$, $Q = 2pq$ and $R = q^2$ satisfying the Hardy-Weinberg equilibrium (Hartl and Clark, 2007). Each allele frequency can also be expressed in terms of genotype frequencies. In other words the frequency of an allele is equal to the frequency of homozygote genotype formed by two duplicates of that allele plus half of the frequency of the heterozygote genotype ($p = P + \frac{1}{2}Q$ and $q = R + \frac{1}{2}Q$).

The absolute fitness of each genotype (denoted as w_{AA}, w_{AB} and w_{BB} accordingly) is considered here by the average number of offspring produced by the individuals who carry this genotype. Relative fitness, i.e. the fitness of one genotype relative to that of another, is given by the ratio of their absolute fitnesses. Since this study deals with human populations, certain assumptions, i.e. organisms are diploid, reproduction is sexual and mating is random, are assured. It is also assumed that neither mutations or gene flows take place, and that stochastic effects due to genetic drift are negligible (the population size is large enough). Based on these assumptions the following formulas for the change of allele frequencies from generation i to generation $i + 1$ can be derived:

$$p^{i+1} = P^{i+1} + \frac{1}{2}Q^{i+1} = \frac{w_{AA}p^2 + w_{AB}pq}{\overline{w}}, \qquad (4.9)$$

$$q^{i+1} = R^{i+1} + \frac{1}{2}Q^{i+1} = \frac{w_{BB}q^2 + w_{AB}pq}{\overline{w}} \qquad (4.10)$$

where the denominator in both fractions is the normalised factor $\overline{w} = w_{AA}p^2 + 2w_{AB}pq + w_{BB}q^2$ (Hartl and Clark, 2007), representing the average number of children per individual in the population of interest. The changes in genotype frequencies between two subsequent generations are shown in Table 4.1.

Genotype	AA	AB	BB
Frequency of genotype at generation i	$P = p^2$	$Q = 2pq$	$R = q^2$
Absolute fitness	w_{AA}	w_{AB}	w_{BB}
Frequency of genotype at generation $i+1$	$\left(\dfrac{w_{AA}p^2 + w_{AB}pq}{\overline{w}}\right)^2$	$2\left(\dfrac{w_{AA}p^2 + w_{AB}pq}{\overline{w}}\right)\left(\dfrac{w_{BB}q^2 + w_{AB}pq}{\overline{w}}\right)$	$\left(\dfrac{w_{BB}q^2 + w_{AB}pq}{\overline{w}}\right)^2$

Table 4.1: Recurrence relation of genotype frequencies between two consecutive generations in a diploid genetics model with random mating.

4.4. Results

In this section, the mortality models described in Section 4.3.1 are fitted to actual mortality data of the entire lifespan in Section 4.4.1 and data from very old ages (above age 80) in Section 4.4.2, and comparisons between the fits of the models are performed. In Section 4.4.3 the study focuses on mortality at very old ages and shows that the model of a heterogeneous population can reproduce and explain various old-age mortality observations, namely deceleration, plateau and decline. In Section 4.4.4 the evolution of mortality dynamics in a heterogeneous population and specifically the homogenisation of this population over time is derived from the changes in genotype frequencies in successive generations through the process of natural selection.

4.4.1. Comparing mortality models by fitting data over the entire lifespan

The model of heterogeneous populations with different numbers of subpopulations as well as the Gompertz, Makeham, Perks and Heligman-Pollard models are fitted to period mortality data over the entire lifespan for the total (males and females) Swedish population. The Bayesian Information Criterion (BIC) values as calculated by fitting the models to the 2000-2010 period Swedish mortality data are shown in Figure 4.1. The model that gives the lowest BIC value provides the best fit to the data. The values in Figure 4.1 indicate that Gompertz, Makeham and Perks models are the weakest models in terms of data fitting. The model of heterogeneous populations gets better with an increase in the number of subpopulations from two to four but any further increase in the number of subpopulations does not result in significant improvements. One can see that the Heligman-Pollard model fits the data over the entire lifespan better than all other models including the model of heterogeneous populations (except for the 2010 period data for which both models provide fits with approximately the same accuracy). The actual fits of the four-subpopulation and Heligman-Pollard models to the 2004 period Swedish death rates are shown in Figure 4.2.
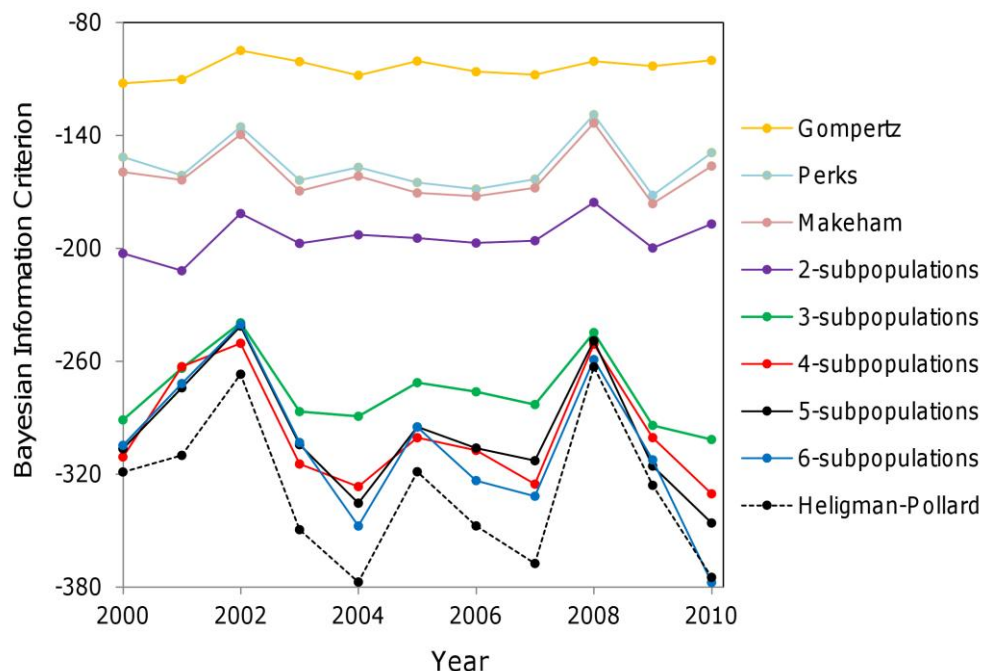

Figure 4.1: BIC values for different mortality models fitted to the Swedish mortality data for the period 2000-2010. *The fits by Gompertz, Makeham, Perks and Heligman-Pollard models and the fits by the model of heterogeneous populations consisting of two to six subpopulations are shown*

The Heligman-Pollard model fits the age-dependent mortality patterns very accurately since it imposes a pre-defined mortality pattern. Indeed the Heligman-Pollard model divides the mortality pattern into three distinct components observed over the past century, namely infant, accidental and adult mortality. On the other hand, the model of heterogeneous populations is more abstract, since it does not impose any pre-specified pattern: it assumes that the most basic feature of biological populations is their heterogeneity and all peculiarities of population mortality dynamics are conditioned by interplay between mortalities of subpopulations. Subpopulations in turn are homogeneous and their mortality dynamics simply follow the exponential law. The model can then be adapted to any dataset and can reproduce very different mortality curves. This flexibility allows (i) to fit mortality data very well for any part of the lifespan (see for example Section 4.4.2 on old ages), (ii) to reproduce different and potentially controversial observed mortality patterns (see Section 4.4.3 for an example related to old-age mortality) and (iii) to capture any new and thus unexpected mortality features (for example the reduction of external causes of death may result in the elimination of the accidental hump (Gaille, 2012)).
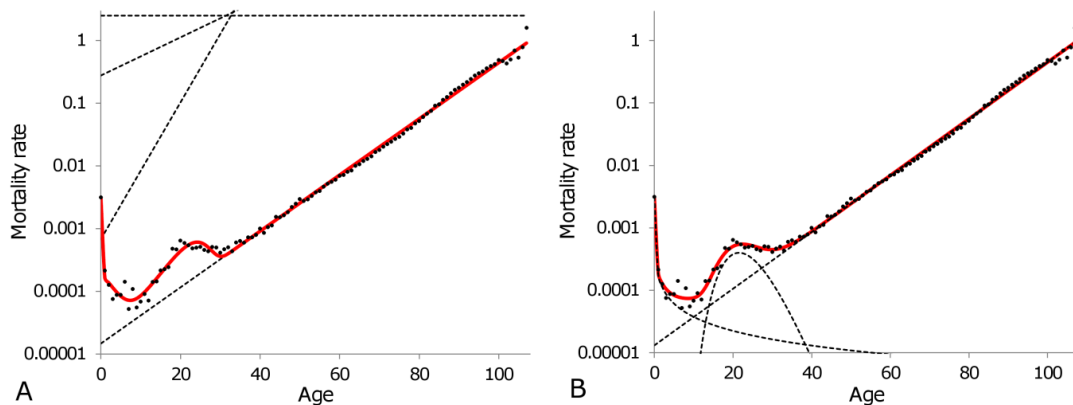
Figure 4.2: 2004 Swedish mortality data fitted by the model of a heterogeneous population composed of four subpopulations (panel A) and the Heligman-Pollard model (panel B). The dots represent the observed central death rates, while the dashed curves in panel A indicate the exponential mortality dynamics of each subpopulation in the model of a heterogeneous population and in panel B - the dynamics of the three components of the Heligman-Pollard model. Note that the plots are given in semi-logarithmic scale.

4.4.2. Comparing mortality models by fitting data of ages beyond 80

In this section we focus on mortality at very old ages (above age 80) and analyse the phenomenon of late-life mortality divergence from the exponential dynamics. For this

analysis we use the models designed for old-age mortality described in Section 4.3.1. All these models were fitted to period Swedish data for ages 80-110. The data were averaged for four-decade periods between the years 1970 and 2010. This adjustment allowed smoothing of the data by removing fluctuations. The BIC values calculated by fitting the models to the data are shown in Table 4.2.

Model	Equation	Number of parameters	BIC	$\lim_{x\to\infty} m_x$
Gompertz	$m_x = \alpha e^{\beta x}$	2	-151.731	∞
Makeham	$m_x = \gamma + \alpha e^{\beta x}$	3	-148.297	∞
Weibull	$m_x = \alpha x^{\beta}$	2	-161.819	∞
Heterogeneous 2-subpopulations	$m_x = \sum_{j=1}^{2} \rho_{j,x} m_{j,0} e^{\beta_j x}$	5	-167.380	∞
Heterogeneous 3-subpopulations	$m_x = \sum_{j=1}^{3} \rho_{j,x} m_{j,0} e^{\beta_j x}$	8	-154.080	∞
Perks	$m_x = \gamma + \dfrac{\alpha e^{\beta x}}{1 + \delta e^{\beta x}}$	4	-150.755	$\gamma + \dfrac{\alpha}{\delta}$
3-parameter Logistic	$m_x = \gamma + \dfrac{\alpha e^{\beta x}}{1 + \alpha e^{\beta x}}$	3	-127.547	$\gamma + 1$
Beard	$m_x = \dfrac{\alpha e^{\beta x}}{1 + \delta e^{\beta x}}$	3	-154.189	$\dfrac{\alpha}{\delta}$
Kannisto	$m_x = \dfrac{\alpha e^{\beta x}}{1 + \alpha e^{\beta x}}$	2	-122.069	1
Michaelis-Menten	$m_x = \alpha e^{\beta x/(1+\gamma x)}$	3	-158.714	$\alpha e^{\beta/\gamma}$
Exponential-Quadratic	$m_x = e^{\alpha + \beta x + \gamma x^2}$	3	-149.078	0 (for $\gamma < 0$)

Table 4.2: Comparison of BIC values for several parametric models fitted to the average 1970-2010 Swedish data for ages beyond 80. All model parameters are assumed to be greater than or equal to zero except for the parameters α and γ in the exponential-quadratic model where they are assumed to be negative.

The logistic-type models (Perks, 3-parameter logistic, Beard and Kannisto) and the Michaelis-Menten-type model show convergence to a certain limit as age increases and are therefore suitable to explain the late-life mortality deceleration and the existence of mortality plateaus. The exponential-quadratic model can generate a concave down parabola and therefore explains the decline of mortality at old ages. The exponential models by Gompertz, Makeham and Weibull fail to explain the late-life mortality slow down, because the death rates expressed by these functions tend to infinity as age increases. Even if the subpopulation mortality rates also diverge as age tends to infinity, the model of heterogeneous populations appears to be the only model which, due to interplay between subpopulations, can reproduce all observations (deceleration, plateau and decline) in mortality it later life.

From the results shown in Table 4.2, one can conclude that the model of a heterogeneous population composed of two subpopulations provides the best fit to the mortality data at old ages. For the averaged Swedish data, the mortality curve generated by the model of a heterogeneous population increases exponentially, asymptotically to the level of the dynamics of the frailest subpopulation between ages 80 and 90, then decelerates to reach the level of the dynamics of the most robust subpopulation and then keeps increasing exponentially at that level (Figure 4.3A).
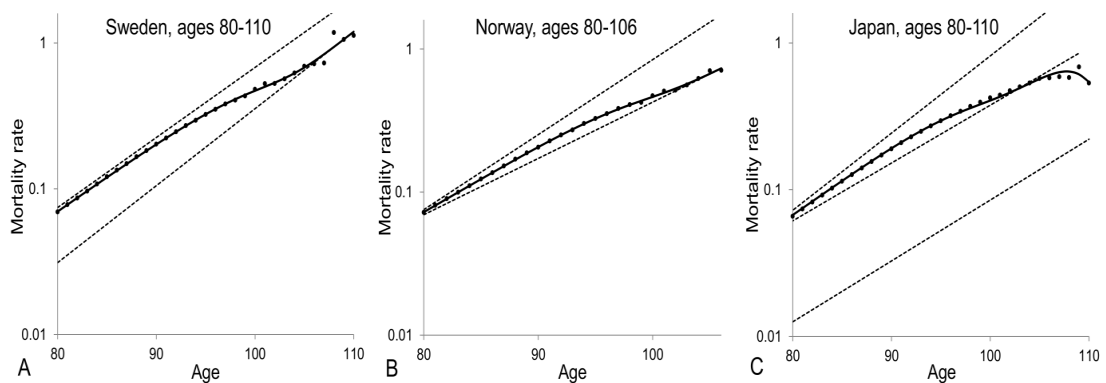
Figure 4.3: Model of heterogeneous populations fitted to average 1970-2010 death rates for ages over 80 for Swedish (A), Norwegian (B) and Japanese (C) populations. The dots represent the observed central death rates, while the exponential mortality dynamics of the subpopulations are shown by the dashed lines and the mortality dynamics of the entire population are shown by the black solid lines. Note that the plots are shown on a semi-logarithmic scale.

Similar results and conclusions have been derived by fitting the models presented in Table 4.2 to the death rates of ages 80+ for other developed countries, including Norway (Figure 4.3B) and Japan (Figure 4.3C). Interestingly, the Japanese data are better fitted

by the three-subpopulation model. The trajectory of mortality that fits the Japanese data increases exponentially along the level of the frailest subpopulation then decelerates for a couple of years, then re-accelerates and finally declines after the age of 108. On the basis of the analysis presented in Figure 4.3 we conclude that different observations on mortality dynamics at extremely old ages can be explained by the heterogeneity of populations, which is further developed in the following section.

4.4.3. Late-life mortality slow down due to population heterogeneity

Heterogeneity suggests that late-life mortality slow down is a result of the variation in robustness between sub-cohorts having a significant number of survivors at old ages. In addition, heterogeneity permits us to explicate the three different observations in late-life mortality, namely the deceleration, the saturation and the decline of mortality rates. Figure 4.4 shows that the simple model of a heterogeneous population composed of only two subpopulations can reproduce all these observations. In Figure 4.4A the frailest subpopulation (i.e. the subpopulation that dies out fastest) is the one that has the highest mortality rate at age 80, $m_{1,80} = 0.08$, and the highest mortality coefficient, $\beta_1 = 0.11$, as compared to the most robust subpopulation that has a mortality at age 80 of $m_{2,80} = 0.04$, and a mortality coefficient of $\beta_2 = 0.09$. The variation in the proportions of the two subpopulations determines the formation of the three different late-life phenomena. For example, if the fraction of the frailest subpopulation at age 80 in Figure 4.4A is $\rho_{1,80} = 0.5$, then the overall mortality of population shows a deceleration, if $\rho_{1,80} = 0.88$ - a plateau and if $\rho_{1,80} = 0.98$ - a decline.

A mortality cross-section, shown in Figure 4.4B, occurs when one of the subpopulations has a lower mortality rate than the other at younger ages, but higher at older ages (i.e. it is more robust initially but becomes frailer after a cross-section). In particular, the theoretical subpopulations presented in Figure 4.4B have mortality rates at age 80 and mortality coefficients $m_{1,80} = 0.09$, $\beta_1 = 0.07$ and $m_{2,80} = 0.04$, $\beta_2 = 0.15$ respectively. The fractions $\rho_{2,80} = 0.2, 0.5$ and 0.8 for the subpopulation with the lowest mortality rate at age 80 are used to reproduce deceleration, plateau and decline respectively.
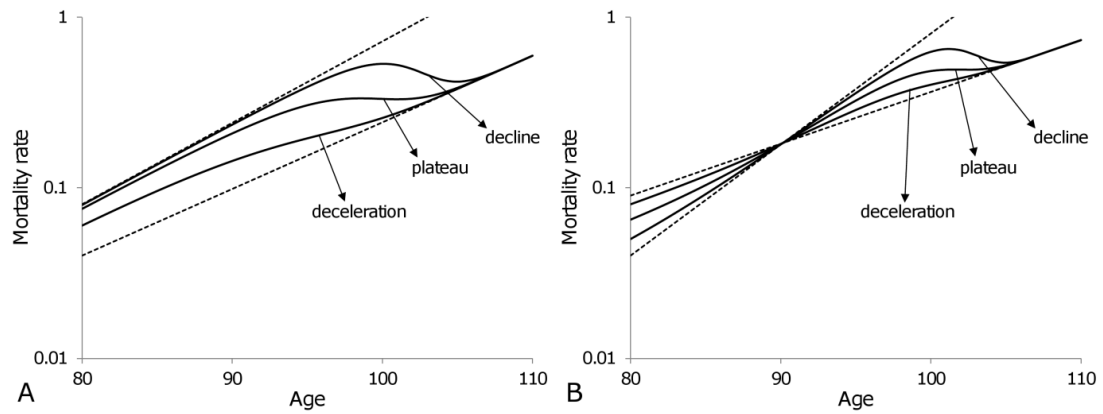
Figure 4.4: Theoretical trajectories (solid curves) of old-age (80-110) mortality dynamics for a heterogeneous population composed of two subpopulations. Variations in relative sizes of the subpopulations permit the reproduction of all three observations for late-life mortality: deceleration, plateau and decline. Once the individuals of the frailest subpopulation die out the mortality of the entire population follows the exponential dynamics of the most robust subpopulation. In panel A the frailest subpopulation remains the same one over ages, while in panel B the frailest subpopulation before age 90 becomes the most robust one from age 90. Note that the plots are shown on a semi-logarithmic scale.

The mortality trajectories presented in Figure 4.4 illustrate that apparently controversial observations in mortality dynamics for old ages are not necessarily in contradiction with each other and can be explained by the heterogeneity of populations.

4.4.4. Evolution of mortality dynamics: homogenisation and natural selection

In Chapter 3 the model of a heterogeneous population consisting of four subpopulations was used to study the evolution of Swedish death rates over the 20th century. In this model, the first subpopulation is used to reproduce the initial decline of mortality for infants, the second - the mortality at childhood, the third - the accidental mortality during reproductive period and the fourth - the exponential (Gompertz) growth of mortality at adult span (see Figure 4.2A). The analysis of mortality evolution, as examined by using this model, showed that the parameters which characterise the mortality dynamics of each subpopulation evolve through time displaying two remarkable features. The first is the confirmation of the compensation effect for each evolving subpopulation and the second is the homogenisation of the entire population manifested by the reduction in the initial fractions of the first three subpopulations (that are also the smallest subpopulations) and an increase in the initial fraction of the fourth subpopulation (from 67% at the beginning of the 20th century to 99% at its end).

An alternative way to examine the evolution of Swedish mortality dynamics over the 20th century is to modify the model of heterogeneous populations by making parameters time-dependent and fitting the model to the entire set of mortality data over age and time so that the fit will be represented on a three-dimensional surface. The death rates for ages 0 to 100 and for the one-century period (101 years from 1900 to 2000) compose a dataset of 10201 points. On the other hand, the four-subpopulation model has 12 parameters of which 11 are independent (the condition that the sum of the fractions $\rho_{j,x}$ at each age is equal to unity reduces by one the number of free parameters). Each of the 11 parameters is assumed to change linearly or exponentially over time according to the trendlines found in Chapter 3. Each linear or exponential trend is characterised by two parameters (a scale and a shape parameter) and therefore the modified time-dependent model has 22 free parameters. Thus, this approach requires the estimation of the values of only 22 parameters in order to fit the 10201 data points while the approach that was used in Chapter 3 requires estimation of the values of 11 unknown parameters to fit 101 data points for each period (or in other words, 1111 unknown parameters in total to fit the 10201 data points).

The 3-dimensional surface that is reproduced by fitting the modified model to age- and time-related Swedish data is shown in Figure 4.5D. The initial mortalities $m_{j,0}$ and the mortality coefficients β_j for each subpopulation j are assumed to change linearly over time as shown in Figure 4.5A and Figure 4.5B respectively. The negative correlation between the initial mortality and the mortality coefficient in each subpopulation indicates the validation of the compensation law of mortality. The initial fractions of the four subpopulations are assumed to change exponentially over time (Figure 4.5C). The phenomenon of homogenisation is evident as the initial fraction of the most robust subpopulation (red line in Figure 4.5C) increases over time and dominates at the end of the century, while the fractions of the other three subpopulations decrease and these subpopulations almost disappear by the end of the century. The most robust subpopulation has the smallest initial mortality rate, and more individuals belonging to this subpopulation survive to more advanced ages compared to the individuals from the other subpopulations.
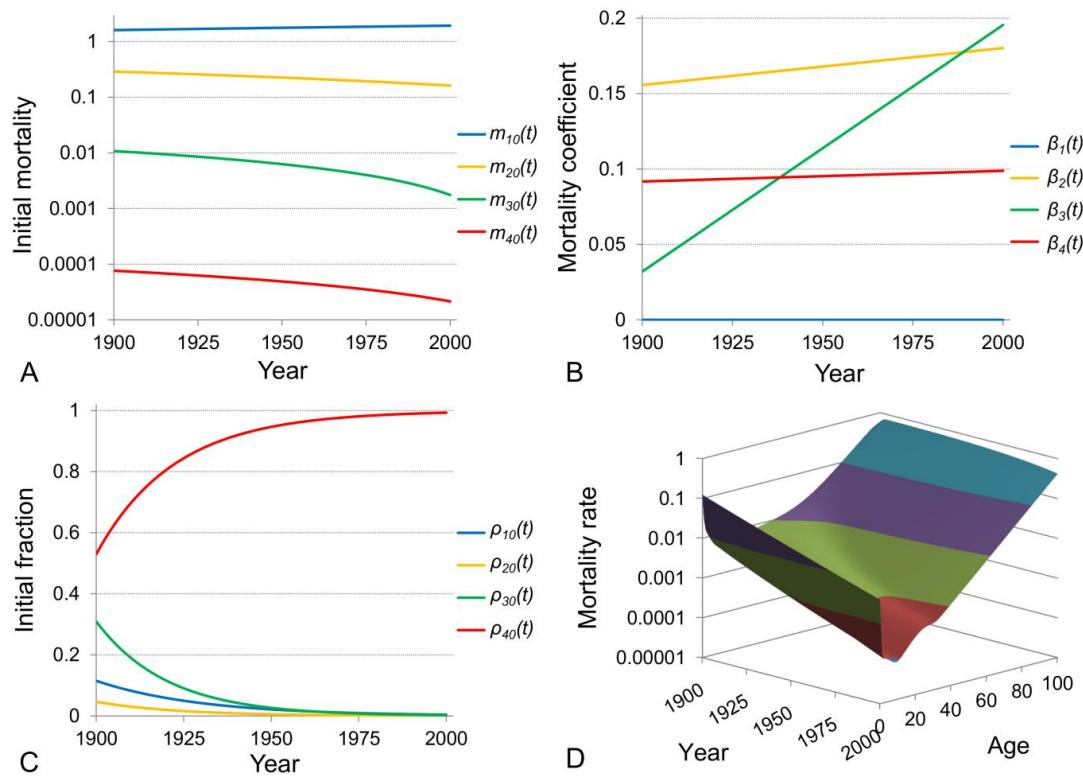
Figure 4.5: Time-evolution of mortality dynamics in the mathematical model of heterogeneous population. *The model of a heterogeneous population composed of four subpopulations is modified to contain time-dependent parameters and is used to fit Swedish death rates for ages 0 to 100 and for the entire 20th century period (1900-2000). The resulting fitted surface of the modified model to the age- and time-related Swedish data is shown in panel D. The initial mortalities and the mortality coefficients of subpopulations are assumed to change linearly over time (fits are shown in panels A and B respectively) while their initial fractions change exponentially (shown in panel C). Note that the plot in panel A is shown on a semi-logarithmic scale.*

Further examination of the results shown in Figure 4.5 indicates that all individuals from the first two subpopulations, reflecting infant and child mortality, die before sexual maturity and the reproductive period and therefore they do not leave offspring. The other two subpopulations have individuals that survive till reproductive age and consequently leave offspring who contribute to the next generation. However, the most robust subpopulation contributes relatively more and if we assume that these two subpopulations differ by genotype, the evolution of their initial fractions can be explained by natural selection. This problem is addressed in the following part of our study, namely we assume as a simplification, that the third and fourth subpopulations differ by a single gene (which has two alleles) and check whether the change in the

fractions of these subpopulations follows the changes in allele frequencies over generations under the force of natural selection.

We use the model for evolution of allele frequencies in diploid organisms as described in Section 4.3.2 and assume that alleles A and B indicate two distinct traits related to mortality dynamics. Choosing from two possibilities we pick up on an assumption that the allele A is dominant and therefore the heterozygotes AB have the same mortality-related phenotype as the homozygotes AA. Furthermore, the individuals carrying genotypes AA and AB are assumed to belong to the third subpopulation while individuals with BB genotype belong to the fourth. To calculate Darwinian fitness, we make the following simple assumptions concerning the reproductive behaviour of the individuals who make up the population: (1) reproductive behaviour does not depend on genotype (note that mortality depends on genotype and makes the fitness genotype specific); (2) reproductive age is set from the age of 20 to the age of 40; (3) within this age interval, reproduction takes place with the constant probability, ϕ, at any age (i.e. it is the same for both subpopulations and independent of age). We believe that by using these assumptions we can obtain a relatively good approximation of the spreading process of a favourite allele in the population due to its effect on mortality only, and thus the dynamics of the relative sizes of two subpopulations. For a more precise analysis, one can adjust the model assumptions by taking into account real fertility related data, the age dependence of reproduction probability, and by specifying the reproductive age-interval more accurately (which is different for males and females). However, here we prefer to keep the model as simple as possible and leave various extensions to the framework, which we are introducing here, for future studies.

Based on the above assumptions we can calculate the absolute fitnesses of individuals that belong to the third and the fourth subpopulations which will be denoted by w_3 ($= w_{AA} = w_{AB}$) and w_4 ($= w_{BB}$) respectively. Fitnesses can be evaluated based on the mortality dynamics and expressed as functions of the parameters that describe the exponential mortality dynamics of these two subpopulations as shown in Table 4.3. In this table, $N_{j,0}$ represents the number of individuals in subpopulation j at age 0.

The absolute and relative fitnesses of subpopulations are found using their initial mortalities and mortality coefficients which are obtained by fitting the heterogeneous model with four subpopulations to the Swedish period data. The estimated initial

fractions of two subpopulations (third and fourth), which are involved in reproduction, for the period 1900 (starting point of the examined time-interval) are normalised to have a sum equal to one (since we do not consider subpopulations 1 and 2 which are not involved in reproduction) and are then used to calculate the frequencies of alleles A and B (or values of p and q) in 1900.

Subpopulation	3rd	4th
Genotypes	$AA + AB$	BB
Initial fraction	$\rho_{3,0} = P + Q = p^2 + 2pq$	$\rho_{4,0} = R = q^2$
Absolute fitness	$w_3 = \phi \sum_{x=20}^{40} N_{3,0} \exp\left(\frac{m_{3,0}}{\beta_3}\left(1 - e^{\beta_3 x}\right)\right)$	$w_4 = \phi \sum_{x=20}^{40} N_{4,0} \exp\left(\frac{m_{4,0}}{\beta_4}\left(1 - e^{\beta_4 x}\right)\right)$
Relative fitness	w_3/w_4	1

Table 4.3: Genotype frequencies and fitnesses in terms of the model parameters. The absolute fitness of individuals belonging to each subpopulation is calculated by the sum of the number of survivors during the theoretical reproductive period (ages 20 to 40) multiplied by a probability to reproduce, ϕ, which is assumed to be age-independent and constant for all individuals. Note that since the initial mortality and mortality coefficient of each subpopulation change over time, the absolute and relative fitnesses also change over time.

Possessing all of the above considerations and the equations that describe the flow of alleles due to selection (equations (4.9) and (4.10)), the changes of genotype frequencies over generations are calculated presuming that each generation corresponds to 25 calendar years totalling four generations per century. Following this, changes in genotype frequencies are compared with the evolution of the initial fractions of the model of heterogeneous populations over the 20th century. The outcome of this analysis is shown in Figure 4.6. Thus, assuming that the difference in mortality dynamics of two subpopulations is conditioned by a difference in a single gene and taking an average value of relative fitness (black dashed line in Figure 4.6A), we calculate how the relative fractions of the subpopulations (corresponding to $AA + AB$ genotype frequency for subpopulation 3 and BB genotype frequency for subpopulation 4, black lines in Figure 4.6B) evolve due to natural selection. We can state that the obtained result is surprisingly close to the changes of the fractions (red triangles and blue dots in Figure 4.6B) in the best-fit model reported in Chapter 3. The significance of this result is that it serves as a self-consistency test for the model of heterogeneous populations.
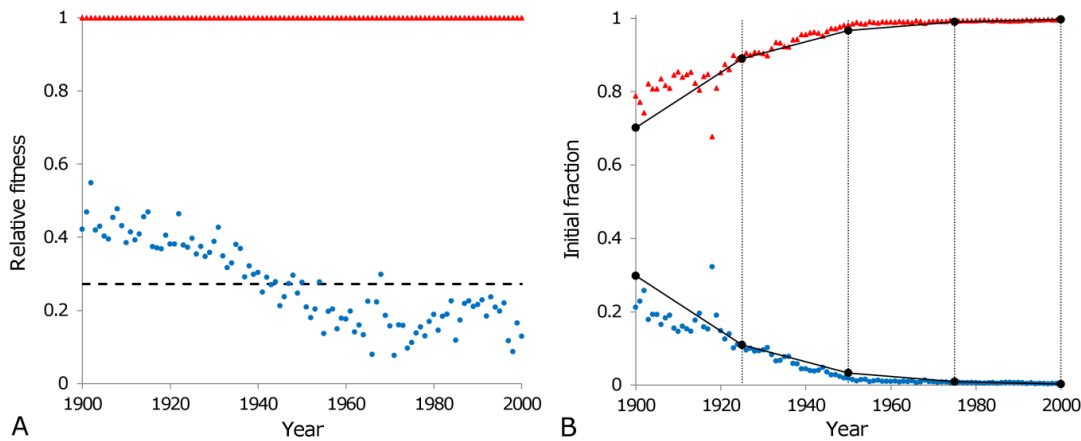
Figure 4.6: Population homogenisation as a consequence of natural selection. *The relative fitnesses of individuals belonging to the third (blue circles) and the fourth (red triangles) subpopulations, as calculated according to the formulas in Table 4.3, are shown in panel A. The relative fitness of the third subpopulation varies from year to year with an average value of 0.27 (black dashed line) over the entire century. This average value is used to calculate the changes in genotype frequencies due to natural selection as shown in panel B (black circles), over four generations (each lasting 25 calendar years as indicated by the vertical lines in panel B). Calculated genotype frequencies are interpolated linearly (solid lines connecting circles in panel B) within each generation to be comparable with the normalised initial fractions of the third (blue circles) and the fourth (red triangles) subpopulations modelled to fit the Swedish data for the period 1900 to 2000.*

Our further calculations have shown that dividing the century into three or five generations (i.e. the time interval between generations is 33 or 20 years respectively) does not significantly change the genotype frequencies observed in Figure 4.6B. (frequencies decrease/increase by about 15%). Averaging fitnesses of subpopulations over each generation rather than over the entire century, does not significantly change the obtained results either (note that in the model, where the evolving population is discretised into generations, consideration of fitnesses for each year or for any intervals shorter than the duration of the generation does not make sense).

4.5. Discussion

The aim of the study presented in this Chapter was to show the flexibility of the model of mortality in heterogeneous populations, its ability to model very different mortality patterns and its appealing interpretation of the peculiarities in mortality dynamics by the heterogeneous structure of human populations. This was done in a few steps. In the first step we compared this model with a number of other parametric models which were used to fit actual mortality data. We found (Sections 4.4.1 and 4.4.2) that the model of

98

heterogeneous populations provides the best fit to mortality data for old (over 80) ages and also provides the second best fit for the data over the entire lifespan. In the second step we demonstrated (Section 4.4.3) that contrary to other considered models, the model of heterogeneous population can reproduce and explain controversial observations in late-life mortality (deceleration, plateau and decline). In the third step we assumed that population heterogeneity reflects the genetic variation between subpopulations, and showed (Section 4.4.4) that the natural selection model based on differential mortality can explain and quantitatively reproduce the homogenisation of the Swedish population within a one-century period. Based on these results, we conclude that heterogeneity, beyond its convenient use in reproducing characteristics of age-structured populations, has a fundamentally inherent role in understanding the mortality dynamics across the lifespan and the evolution of these dynamics over time.

The model of Heligman-Pollard is shown to have an excellent fit for the mortality data over the entire lifespan. It is however important to note that our model of heterogeneous populations is very different in nature from the Heligman-Pollard model. While the Heligman-Pollard model imposes a pre-defined mortality pattern, the model of heterogeneous populations allows the mortality pattern to be adapted to the fitted data. Both models are thus extremely useful in different contexts. The Heligman-Pollard model is better for forecasting purposes, as it avoids projecting unrealistic patterns far into the future. However, as a wide variety of mortality patterns can be modelled using a different number of subpopulations, the model of heterogeneous populations allows us to capture new and unexpected patterns, providing a greater flexibility in data modelling. In addition to this, in the model of heterogeneous populations the model parameters do not lose their interpretation in demographic terms, even with an increase in the number of parameters in the model. This flexibility is important for data analysis as mortality patterns evolve through time due to several factors (medical improvements, changes in life-style conditions, biological evolution, etc) and this can be modelled as some subpopulations die out and some new ones become more pronounced in a quantitative sense. Furthermore, mortality dynamics of subpopulations change over time and these changes can be discovered via fitting procedures.

The introduction of subpopulations with different mortality characteristics can easily be justified on the basis of biological and medical observations. Certain diseases tend to follow others due to strong associations at genetic and cellular levels, and connections at

the cellular level get amplified at the population level when a number of diseases emerge as comorbid (Hidalgo et al., 2009, Barabasi et al., 2011, Chmiel et al., 2014). Susceptibility to a particular disease may stratify a part of population to having a particular dynamic of age-related accumulation of other associated diseases, and consequently a specific dynamic of mortality. To some extent this process is related to one described in the reliability theory (Gavrilov and Gavrilova, 2001), where age-related failure kinetics will be different for particular human physiological systems and their components, and will have a different impact on different human subpopulations. More studies are required on accumulation trends for particular diseases with an emphasis on mortality curves on one hand, and underlying genomic factors on the other. We believe that the complex structure of human populations in respect to evolving disease patterns in different age-groups will be revealed from these studies and more evidence will be available for refining the mathematical models. Having this knowledge we may be able to understand and better predict mortality dynamics in complex human populations by using a pallet of primary disease-associated genomic markers.

Our analysis of the evolution of allele frequencies (under the assumption that genomic differences are responsible for the difference in mortality rates between subpopulations) has indicated that the homogenisation of the Swedish population in the 20th century can be explained by the selection process in favour of a particular subpopulation better fitted to a changing environment during the studied period. The force of selection as calculated on the basis of mortality-related heterogeneity of the population is known as the force of mortality selection (Wrigley-Field, 2014). To provide an intuitive explanation of this force we note that the individuals belonging to frail subpopulations tend to die at younger ages (and more frequently before the reproductive ages) than ones from more robust subpopulations. Therefore frail subpopulations leave less offspring than more robust subpopulations. Consequently, the proportion of individuals belonging to more robust subpopulations increases through generations. Although we do not aim to propose a fully-specified and completely realistic evolutionary model, we show that using very simple assumptions we can relate the evolution of the heterogeneous structure of populations to genetics and natural selection. Thus this Chapter paves the way for many potential extensions regarding genetics and evolutionary theories.

There are many studies indicating that the currently observed increase in longevity is primarily associated with environmental changes (GBD 2013 Mortality and Causes of

Death Collaborators, 2015). In order to link this statement with our results on the evolution of allele frequencies, we would suggest considering the following hypothetical scenario. Consider the population carrying a gene with two alleles *A* and *B*. There are three different kinds of individuals: *AA*, *AB* and *BB* in this population for which we can assume the same pattern of mortality and reproduction (three identical subpopulations) so that the genetic structure of this population is in equilibrium. Now assume that due to some environmental change the mortality of individuals carrying *BB* is reduced. If this reduction hits the reproductive period, the frequency of *B* will tend to increase which will result in a change to the structure of the population and to gradual change in its mortality dynamics causing an increase in longevity. If the initial frequency of allele *A* is small then a jump in mortality patterns in cohort data should be observed with no further evolution. Contrary to this, if the initial frequency of allele *B* is small, then the jump will be replaced by a gradual evolution, associated with an increase of allele *B* frequency. Obviously the period data should show a gradual evolution of mortality patterns in both cases.

In this Chapter we focus on two subpopulations whose mortality dynamics evolve differently (in response to the same environmental changes) and this shows a change in the overall mortality pattern. Mortality patterns of both subpopulations change over the 20th century but for our analysis we have averaged the characteristics of the subpopulations by taking their average fitness. Thus we have reduced our analysis to the following idealised case: environmental change has happened on or before 1900 and this has changed the mortality patterns for subpopulations 3 and 4. The latter causes the changes in fitnesses of subpopulations, follow up gradual changes in the population structure and consequently lead to gradual increase in longevity. We do not address the question of why mortality of subpopulations changes in a certain way (in response to environmental changes), but taking these changes as granted we confirm that the change in the structure of populations (represented by fractions of subpopulations) correlates with the evolution of frequency of the hypothetical allele.

The surprising part of this result is the time scale of the process: the selection process causes significant changes to take place in the population within one century (over four generations). In the model we have assumed that there is no difference in the reproductive behaviour of individuals belonging to different subpopulations and thus the difference in fitnesses is only conditioned by the difference in mortality patterns of the

subpopulations, namely by their initial mortalities and mortality coefficients. As these parameters are time dependent, the relative fitnesses of subpopulations also change over time. In our study illustrated in Figure 4.6, we have ignored the fact that the fitnesses of the subpopulations evolve over time and used the average fitnesses over the entire century to calculate changes during four generations. This was done for two reasons: (1) to illustrate the process in a very simple case when the fitnesses do not change over time and (2) to properly account for variations in fitnesses from year to year, we would have to give up the idea of discrete generations and design a much more sophisticated model (i.e. design a virtual population).

An interesting question concerning the evolution of subpopulations analysed in Figure 4.6, concerns the relationship between them in the 19th century. Our preliminary study shows that subpopulation 3 (which is almost extinct by the end of the 20th century) had a higher relative fitness for most of the 19th century as its mortality rate was lower than that of the fourth subpopulation and as a consequence the fraction of subpopulation 3 was rather increasing in the 19th century (and then decreased in the 20th century).

In this Chapter we have presented a very simple, almost caricature, natural selection model to compare its outcome with the evolution of subpopulations in the fits of heterogeneous model to mortality data. We believe that our work will stimulate the development of more realistic models based on genetics and natural selection. However, surprisingly close correspondence between the time evolution of the subpopulation in the model of heterogeneous populations and the evolution of genome frequencies can already be highlighted on the basis of our simple model. This finding naturally paves the way for many interesting research questions and future research studies, such as the impacts of the environment on mortality changes. Indeed, the sharp reduction in overall mortality during the 20th century and especially the dramatic decline of premature (infant and child) mortality in almost all countries is mainly a result of environmental changes and improvements (Black et al., 2010, Ahmad et al., 2000) and to a lesser extent to biological evolution.

In our model framework, we do not explicitly account for environmental factors. Each subpopulation reacts in its own way to the environmental changes, and the mortality pattern of each subpopulation (here the scale and shape parameters of their exponential dynamics shown in Figure 4.5A and Figure 4.5B) evolves differently over time.

Therefore, we should further explore the effects of environmental changes on mortality dynamics of heterogeneous populations (that would be reflected in the evolution of the model parameters), on reproductive success, reproduction windows and duration of lifespan (interesting results can be found in (Stearns et al., 2010, Pettay et al., 2007)). Another potential extension of this work is associated with the consideration of the heritability of phenotypic mortality-related traits which are affected by genetic variations and environmental factors. Future research could also involve deeper consideration of the age-dependent fertility rate, male-female ratio, wider or narrower reproductive periods and changes over time including time-dependent fitnesses. More complex models of natural selection taking into account the effect of more than one gene polymorphism and naturally occurring splits in frequencies of different gene variants are to follow. In addition, the effects of in and out migration, mutations and genetic drift could be examined. Finally an extensive literature exists on biological ageing and its potential relation to some longevity genes. Linking this stream of research with the model proposed in this study could reveal new mortality modelling tools and improve our knowledge on mortality and longevity matters.

Chapter 5.

Discussion

5.1. Summary of the main findings

In this study we developed a mathematical model to analyse mortality data in human populations. The model was used for the analysis of mortality dynamics across the lifespan (Chapter 2) and for the analysis of the evolution of mortality patterns across the 20th century (Chapter 3). It has also been used to investigate whether heterogeneity is a convenient consideration in developing descriptive models that precisely reproduce mortality patterns, or indeed if it reflects the real structure of mortality dynamics in human populations (Chapter 4).

Chapter 1 provided an introduction to the subject of the thesis where we outlined the basic mortality-related observations and several mathematical approaches that have previously been performed to model the dynamics of mortality. We also stated our motivation which was the development of a mathematical model that considered two important characteristics: the heterogeneity of populations and the exponential dynamics of mortality. Following this introduction, we illustrated the development of the model and its features in Chapter 2. We demonstrated how the variations in model parameters affect mortality dynamics in heterogeneous populations and showed that a completed set of age-specific mortality data can be reproduced fairly well by a model comprising four-subpopulations. In Chapter 2, we also analysed the influence of stochastic effects on the dynamics of mortality and shown that this influence is significant at young and very old ages when only a few individuals contribute to mortality. The outcomes of Chapter 2 indicate that the deviations from the exponential law at young ages can be explained by

the heterogeneity of populations, while those for old ages can be viewed as fluctuations and explained by stochastic effects.

In Chapter 3 we fitted the model to Swedish data of consecutive periods over the 20th century in order to analyse the evolution of mortality dynamics in terms of the evolution of model parameters. This analysis displayed two interesting outcomes; (i) the applicability of the compensation effect for each subpopulation separately and (ii) the change in population structure towards homogenisation at the end of the 20th century. Based on these two outcomes we show that the decrease of overall population mortality over time is primarily a result of homogenisation of the population (change in the relative sizes of subpopulations) and secondly, this is a result of the reduction in mortality dynamics for each constituent subpopulation (change in scale and shape parameters of exponential components in the model).

In Chapter 4 we compared the fits to a given set of data obtained by using different parametric mortality models including the model of heterogeneous populations. We show that the model of heterogeneous populations fits mortality data better than most of the other models if the data are taken for the entire lifespan and better than all other models if we consider only old ages. We also demonstrated that the model can explain controversial observations in late-life mortality, namely deceleration, levelling-off and mortality decline. Furthermore in Chapter 4, we considered that population heterogeneity reflects the genetic variation between subpopulations, and show that the homogenisation of the Swedish population within the 20th century can be quantitatively reproduced by a model of population genetics that describes the changes in allele frequencies over generations. Based on these results we concluded that heterogeneity, beyond its convenient mathematical usage in reproducing and explaining several characteristics of age-structured populations, has a fundamental, inherent role on mortality processes, and on the evolution of mortality over time.

The strengths (including advantages and further applications) and limitations of the model are discussed in Section 5.2. In Section 5.3 we propose some ideas for future work and we conclude the thesis in Section 5.4.

5.2. Analysis of the model

The model that we have introduced and used for this research has several advantages compared to other models of mortality. One of them is that its parameters have simple demographic interpretations which makes the modelling results easy to analyse and useful for understanding age and time specific observations related to mortality. We have demonstrated that this model can be used as an efficient mathematical tool for the analysis of mortality data and the exploration of biological or demographical processes that underlie ageing and mortality.

A simple version of the model, for a population composed of only four subpopulations, was shown to precisely reproduce a set of mortality data of the entire lifespan, explaining the deviations of mortality from exponential growth, at young and very old ages (as shown in Chapter 2). The model can also generate different (and mutually contradicting) observations of mortality (late-life deceleration, levelling off and mortality decline) at advanced ages and therefore population heterogeneity can explain and verify the existence of these phenomena (as shown in Chapter 4). By fitting the model to consecutive period data we analysed the changes in model parameters over time and made interesting observations concerning mortality evolution. The two main outcomes of this analysis are the validation of the compensation effect for each constituent subpopulation and the homogenisation of the (Swedish) population over the 20th century (as shown in Chapter 3).

The formulation of our model allows for the consideration of any number of subpopulations required to reproduce a dataset, without the model parameters losing their demographic meanings and avoiding any over or under parameterisation of the model. This flexibility is important as the mortality patterns evolve over time and any gained or lost feature of a pattern can be modelled not only by changes in characteristics of subpopulations but also by change in the number of subpopulations, i.e. by the inclusion or exclusion of subpopulations.

The statistical method used in our study (see the description of fitting procedure in Section 1.3.2) was appropriate for the specific formulation of the model we used and does not lead to any faulty conclusions. One important point is that the log transformation of mortality rates reduces the variability of errors (residuals) while another is that the highly precise "fits" obtained from this model make the variability

almost constant at all ages, indicating that the errors are approximately homoscedastic. Also, the estimated residuals are approximately normally distributed and therefore the method used does not cause any significant biases to the estimates. The Quantile-Quantile (Q-Q) plots in Figure 5.1 show linearity of the residuals estimated by fitting the four-subpopulation model to Swedish mortality data (Figure 5.1B) suggesting that the errors are normally distributed while this is not the case when the data are fitted by Gompertz function representing a homogeneous population (shown in Figure 5.1A). This confirms that the Least Squares method used for the estimation of parameters of the heterogeneous population model gives results as accurate as those that can be derived from the Maximum Likelihood method.
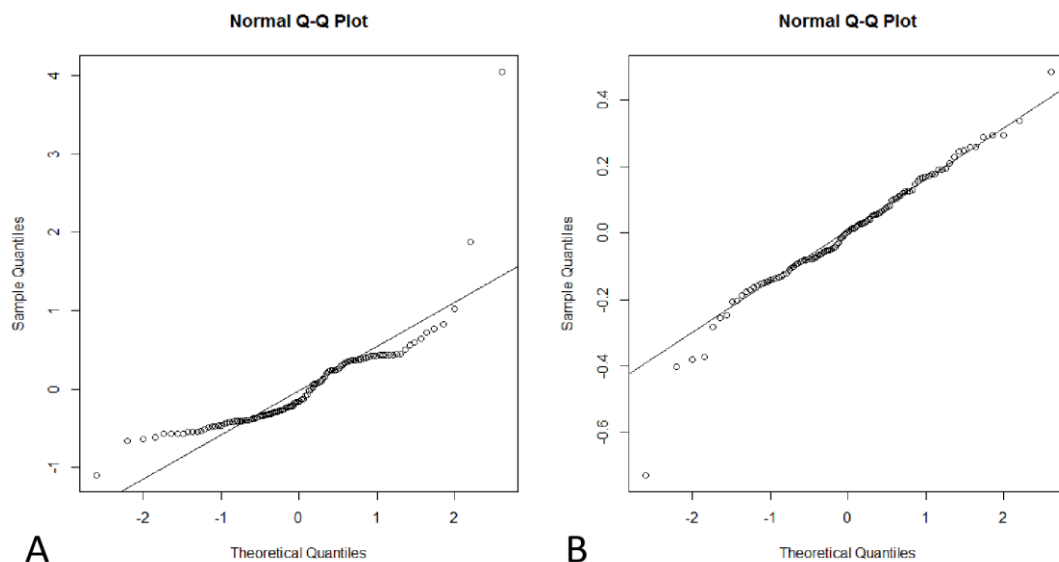
Figure 5.1: Normal Quantile-Quantile plots. *The Q-Q plot for the residuals estimated by fitting a homogeneous Gompertz model (panel A) and a four-subpopulation model (panel B) to 2007 Swedish mortality rates.*

While applying our fitting procedure for estimating unknown parameters, we set a few constraints to the ranges of the parameters' values (as mentioned in Chapter 3) in order to avoid any misleading results. For example, a negative mortality coefficient (i.e. negative value for a parameter β_j) could result in a better fit to a dataset (according to goodness-of-fit measurements) but would represent an unrealistic process opposite to senescence and therefore is set to take only positive values. We consider these constraints as a limitation of the model and we therefore suggest an alternative approach for future work that includes the development of a new mortality model that does not consider population heterogeneity but uses the formulation of the superposition of

exponential terms (see this idea and other plans for future work in Section 5.3). Despite this limitation which is further discussed in Section 5.2.3, the model of heterogeneous populations can be used to express the rate of mortality at continuous ages (Section 5.2.1) and can also be applied to other mortality-related measurements such as probability density and survival function (Section 5.2.2).

5.2.1. Continuous model of mortality in a heterogeneous population

In the continuous model, age is defined by a real number x (continuous age) rather than by the integer number i. For continuous age x, the force of mortality, $\mu(x)$, of a homogeneous population is defined by equation (1.2). Substituting the Gompertz law in the LHS of equation (1.2) and solving the differential equation, results in

$$N(x) = Ae^{-(\mu_0/\beta)e^{\beta x}}, \tag{5.1}$$

where the constant of integration A is equal to $N_0 = e^{\mu_0/\beta}$ as estimated by the initial condition $N(x = 0) = N_0$. This means that the expression for the population size N at age x depends on the initial mortality, μ_0 and the mortality coefficient β:

$$N(x) = N_0 e^{(\mu_0/\beta)(1-e^{\beta x})}. \tag{5.2}$$

In a heterogeneous population, formula (5.2) is used to describe the size of each subpopulation at age x. Therefore, the subscript j is added in each parameter. As a result, the mortality of the entire population in continuous age is expressed by

$$\mu(x) = \frac{\sum_{j=1}^{n} \mu_j(x)N_j(x)}{\sum_{j=1}^{n} N_j(x)} = \frac{\sum_{j=1}^{n} \mu_{j,0} e^{\beta_j x} N_{j,0} e^{(\mu_{j,0}/\beta_j)(1-e^{\beta_j x})}}{\sum_{j=1}^{n} N_{j,0} e^{(\mu_{j,0}/\beta_j)(1-e^{\beta_j x})}}. \tag{5.3}$$

By solving equation (5.3) at integer values of age $(x = i)$, equation (2.8) is found, providing a link between the dynamics of mortality for both the continuous and discrete models.

5.2.2. Application of the model to probability density and survival function

The consideration of heterogeneity in human populations can be used for the derivation of models for other mortality-related variables that exist in human life tables. Such
108

variables are the number of survivors and the number of deaths at age x. In this section, the models of probability density and survival function for heterogeneous populations are developed in continuous time.

In a homogeneous population, $S(x + \Delta x)$ denotes the probability of an individual surviving at age $x + \Delta x$ (usually called survival function) and is calculated as the difference between the probability of surviving at age x and the probability of dying between ages x and $x + \Delta x$:

$$S(x + \Delta x) = S(x) - S(x)\mu(x)\Delta x \qquad (5.4)$$

$$\Rightarrow \frac{S(x + \Delta x) - S(x)}{\Delta x} = -\mu(x)S(x). \qquad (5.5)$$

The limit of LHS of equation (5.5) when Δx tends to 0, is the derivative of $S(x)$ with respect to x:

$$\lim_{\Delta x \to 0} \frac{S(x + \Delta x) - S(x)}{\Delta x} = \frac{dS(x)}{dx}, \qquad (5.6)$$

and therefore equation (5.5) can be rewritten as the differential equation

$$\frac{dS(x)}{dx} = -\mu(x)S(x). \qquad (5.7)$$

The solution of the differential equation (5.7), when the force of mortality $\mu(x)$ follows the Gompertz law, is

$$S(x) = Ae^{-(\mu_0/\beta)e^{\beta x}}, \qquad (5.8)$$

where the constant of integration is $A = e^{\mu_0/\beta}$ (given by initial condition $S(x = 0) = 1$).

Multiplying the survival function (equation (5.8)) with the initial size of population N_0 we get the number of surviving individuals at age x (which is the same expression as equation (5.2)). As a result, the number of survivors of a heterogeneous population at age x is given by:

$$N(x) = N_0 S(x) = N_0 \sum_{j=1}^{n} \rho_{j,0} \exp\left(\frac{\mu_{j,0}}{\beta_j}\left(1 - e^{\beta_j x}\right)\right). \tag{5.9}$$

The probability density function $f(x)$ of a heterogeneous population is derived similarly. The probability $q(x)$ of an individual dying by age x is the complement of the probability of surviving at the same age (i.e $q(x) = 1 - S(x)$) and therefore the probability density function is obtained by differentiating the cumulative distribution function $q(x)$ with respect to x:

$$f(x) = q'(x) = \mu_0 \exp\left(\beta x - \frac{\mu_0}{\beta}\left(e^{\beta x} - 1\right)\right). \tag{5.10}$$

By multiplying the probability density function with the initial size of the population, we get the theoretical distribution of deaths across the lifespan, $\Delta N(x) = N_0 f(x)$.

The distribution of deaths across lifespan, in a heterogeneous population composed of n subpopulations, is given by the sum of the number of deaths of individuals from each subpopulation:

$$\Delta N(x) = \sum_{j=1}^{n} N_{j,0} f_j(x) = N_0 \sum_{j=1}^{n} \rho_{j,0} \mu_{j,0} \exp\left(\beta_j x - \frac{\mu_{j,0}}{\beta_j}\left(e^{\beta_j x} - 1\right)\right). \tag{5.11}$$

The models described above are fitted to datasets of different mortality-related measurements of the 2010 period Swedish population, taken from the Human Mortality Database. The fitting procedure described in Chapter 1 is applied to death rates, number of deaths and number of survivors by using equations (5.3), (5.11) and (5.9) respectively. The BIC values indicate that the best fit to the observed number of deaths and survivors is obtained with a model composed of four subpopulations.

Consequently, the analysis shows that the assumption of population heterogeneity provides mathematical models that fit the mortality-related data better than a model of a homogeneous population. On the other hand, the three attempts to fit mortality-related data of the same population do not give the same values for the model parameters. The model of mortality in heterogeneous populations with optimal parameter values is able to reproduce the mortality pattern for the entire lifespan, since using the logarithm of mortality rates during the fitting procedure we increase the weight of young ages. The

other two models (equations (5.9) and (5.11)) provide parameters that minimize the residuals mainly across adulthood, since the differences between theoretical values and observations at young and extremely old ages are negligible. Besides, the theoretical relationships between equations (2.8), (5.3), (5.9) and (5.11) are only valid for cohort data with no migration where the number N_{i+1} of surviving individuals at age $i + 1$ in year t is equivalent to the number of surviving individuals at age i in year $t - 1$ minus the number of individuals who died at age i in year $t - 1$, $N_i - \Delta N_i$. However, since we do not fit cohort data but period data, and since the Swedish population is subject to migration flows, this relationship does not hold, partly explaining the observed differences between the values of the parameters of the three fitted models.
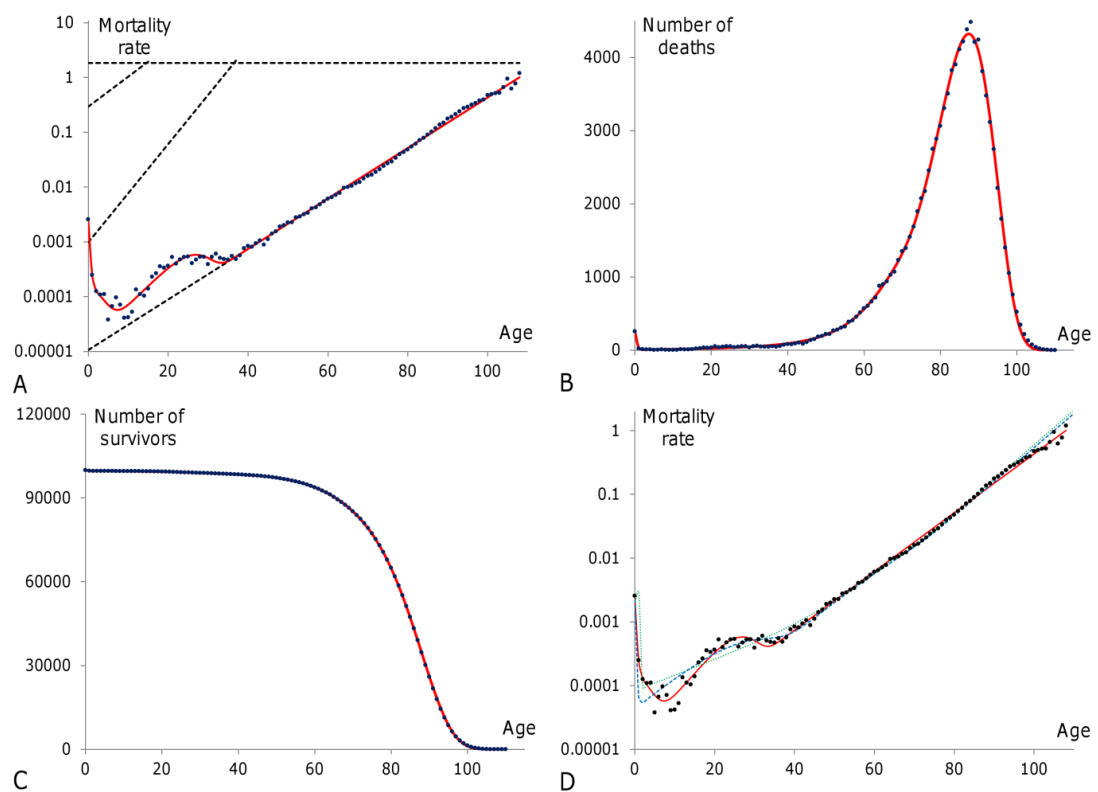
Figure 5.2: The model of a heterogeneous population fitted to the 2010 Swedish mortality data. *A: The mortality model of a heterogeneous population composed of four subpopulations is fitted to observed mortality rates. B: The density function of a heterogeneous population composed of four subpopulations is fitted to actual numbers of deaths and C: The survival function of a heterogeneous population composed of four subpopulations is fitted to actual numbers of survivors. D: Different fits of the four-subpopulation model to mortality rates: the solid (red) curve represents the mortality pattern resulting from the model fitted to the mortality rates (same pattern as in panel A) while the dotted (green) and dashed (blue) curves show the mortality pattern resulting from the model fitted to the numbers of deaths and the numbers of survivors, respectively.*

The mortality rates of the entire population resulting from the model applied to the three different sets of Swedish data are shown in Figure 5.2D. The dotted and dashed curves indicate that the parameters obtained by fitting the number of deaths and survivors, fail to accurately model the peculiarities of mortality patterns at early and extremely old ages. However both curves create a smooth dip at around age 75 and thus better capture the mortality pattern at adult ages compared to the curve of mortality obtained by fitting mortality rates (solid curve in Figure 5.2D).

5.2.3. Limitations of the model

We have shown that the model can explain the main characteristics of mortality dynamics across the lifespan and can provide interesting observations on the time-evolution of population mortality. However, a limitation in the fitting procedure should be mentioned as this could otherwise lead to unrealistic outcomes. As described in Chapter 2, each subpopulation is characterised by three parameters, the initial fraction $\rho_{j,0}$, the initial mortality $m_{j,0}$ and the mortality coefficient β_j. The sum of the fractions $\rho_{j,i}$ at each age i is equal to unity. Then, for a heterogeneous population consisting of n subpopulations, the fitting procedure estimates the values of $3n - 1$ unknown parameters. These estimations are obtained by setting some constraints to the parameters, in order to avoid values with unrealistic meanings (for example to avoid negative values for initial mortality or negative values for initial fractions). These constraints are: (a) the mortality coefficients are non-negative, (b) the initial mortalities are positive (c) the initial fractions can take any value from 0.00001 to $1 - 0.00001 \cdot (n - 1)$ in order to have at least one individual in each subpopulation initially (within a theoretical initial entire population of 100000 individuals) and (d) the sum of the initial fractions is equal to one.

Using these constraints we observed that in some cases the best fit to the data is given by a four-subpopulation model where the frailest subpopulation has a mortality coefficient equal to zero (see for example Figure 3.7 and Figure 3.9). This implies a non-ageing process however all the individuals of that subpopulation die within the first few years of life due to the high level of initial mortality. To avoid the zero value of one of the mortality coefficients, the fitting procedure can be repeated by ignoring the constraint for non-negative β_j. Removing this constraint, we get a slightly improved fit to the data (according to the BIC) with the frailest subpopulation having a negative mortality

coefficient (negative slope). In that case, the fraction of the frailest subpopulation tends to a constant value with increasing age and saturates as age tends to infinity. This saturation means that some individuals of that subpopulation are "immortal" (they never die) which is unrealistic. Two fits of the model of heterogeneous populations (without using restrictions for the Gompertz slopes) to mortality data of the Swedish population and to the intrinsic mortality of the French male population, are shown in Figure 5.3A and B respectively. The fit to the French male data results in the conclusion that the frailest subpopulation has a number of "immortal people". For an entire set of mortality data, the occurrence of "immortal people" in the frailest subpopulation is observed in cases when mortality at advanced ages significantly declines (a local maximum at those ages exists instead of fluctuations). In that case, the negative slope of the frailest subpopulation is sharper than the negative slope shown in Figure 5.3A, and is responsible for creating a maximum at older ages.
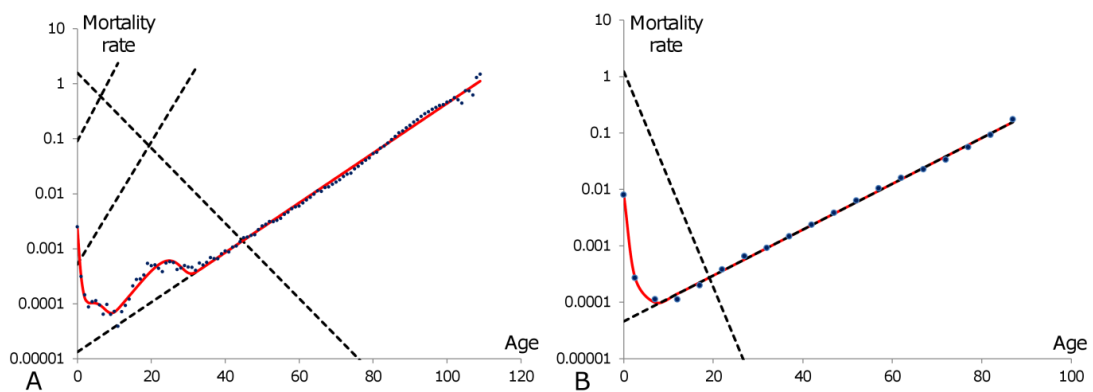

Figure 5.3: Fitting the model of heterogeneous populations to actual mortality data using Least Squares Method without constraints in the parameters. A: The model of heterogeneous population consisting of four subpopulations was used to fit the Swedish 2007 Period data ($BIC = -336.99$). B: The model of heterogeneous population consisting of two subpopulations was used to fit the French Male 1990 data which excluded external causes of deaths ($BIC = -74.50$). In both plots the actual data are denoted by circle points, the mortality of the total population by the red-solid curve and the mortality of the consisting subpopulations by black-dashed lines.

The reference to "immortal people" is unrealistic for biological populations and therefore these conclusions indicate that the concept of heterogeneous populations fails to model mortality when restrictions for the parameters are not taken into account. However, mathematically the mixture of exponential terms with no restrictions on the parameters accurately fits an entire set of mortality data. For that reason, a new model of mortality, with a similar structure as the model of heterogeneous populations, could be developed

without considering the heterogeneity but taking into account other mechanisms underlying ageing and mortality. The new model would express mortality rate as a superposition of exponential terms with negative and positive slopes and therefore will be able to accurately reproduce patterns of mortality. This idea and other plans for future work are discussed in the next section.

5.3. Directions for future research

We are planning to extend the work described in this thesis in two main directions. Firstly, we want to extend the approach used in Chapters 2-4 to analyse mortality data for different death causes associated with different diseases. This would provide further insights concerning the nature of mortality dynamics in human populations that can be beneficial for medicine and health service practical use and regulations. Secondly, we want to launch a series of studies to investigate the nature of exponential law of mortality using different model approaches. The extension of our research will be conducted in four different approaches described in the following four sections.

5.3.1. Analysis of mortality data for different causes of death

The techniques and procedures used in this thesis could be repeated for the analysis of mortality data (available from http://www.who.int/ and http://www.cdc.gov/) for different (intrinsic) causes of death in developed countries. This would enable us to study the structure of mortality patterns as related to specific causes of death and its evolution over time. In this way, the numbers and weights of subpopulations that would reproduce mortality patterns from illnesses related to different organ systems (i.e. circulatory, respiratory, digestive, etc.) could be examined. The evolution of mortality patterns related to each kind of disease over the last century could be also analysed. Comparing any upcoming results with the results presented in this thesis, we would be able to confirm how the heterogeneous structure of the population (as related to its total mortality dynamics) is related to its structures associated with different types of diseases for different period data over the 20th century. The study of population heterogeneity could also be extended by a functional analysis of the ageing-defining genes predicted in Genome-Wide Association Studies (GWAS) (Walter et al., 2011).

114

5.3.2. Modelling interactions between different causes of death

The analysis of mortality data for different types of diseases could be extended by the study of correlations between them. In particular, we could investigate whether interactions between diseases can explain correlations between observed shapes of mortality curves for considered diseases. A main goal for this research would be to determine to what extent particular shapes of mortality curves for considered diseases are affected by the interactions between them and whether these interactions can explain why overall mortality follows the Gompertz law. For this study, a discrete type model based on consideration of a set of events (corresponding to illness from particular diseases) which take place over the lifespan could be developed. We would introduce probabilities of these events taking place and probabilities of each event causing death. Using this model would enable us to check how different hypotheses related to the interactions between diseases affect the mortality dynamics over age. The proposed interactions would reflect the fact that an individual already affected by an illness, is more vulnerable and more susceptible (has larger probability) to get another disease than a "healthy" person. We would consider stochastic (e.g. Poisson) processes to model these interactive events with a probability that the occurrence of every next event is affected by past events. In particular, we would check the model assumptions against the conventional laws of mortality (Gompertz law, compensation effect and late-life mortality deceleration).

Our preliminary results indicate that relatively simple rules of interactions between "theoretical" diseases can result in a mortality curve represented by modifications of the Gompertz law (i.e. exponent multiplied by age or exponent divided by age) which in some cases fits mortality data even better than the exponent. This could be observed if we consider that the probability of an individual having n diseases at age x is given by the Poisson distribution

$$P(n, x) = \frac{1}{n!} (\lambda x)^n e^{-\lambda x}, \qquad (5.12)$$

where λ is the rate of one disease per unit of age. An average of δ diseases cause the death of the individual and therefore the force of mortality is given by

$$\mu(x) = \sum_{n=0}^{\infty} P(n, x)\,(\delta n) = \delta \sum_{n=0}^{\infty} \frac{n}{n!}(\lambda x)^n e^{-\lambda x}. \tag{5.13}$$

By using the Taylor's expansion $\sum_{n=0}^{\infty} \frac{k^n}{n!} = e^k$ we can show that for constant λ the force of mortality is a linear function of age (i.e. $\mu(x) = \delta \lambda x$). If λ is an age-dependent function, $\lambda = f(x)$, then the mortality is given by $\mu(x) = \delta f(x)x$. This can be further extend to more complicated functions if λ depends on age, x, and on number of diseases, n.

The "diseases" here may correspond to events of physiological decay of different tissues and organs that may have different dynamics of ageing, but that do interact and affect the integrated organism's function and the probability of death. We would align our modelling predictions with an analysis of tissue specificity with an expression of the known ageing-related genes in different organism's tissues and an analysis of the known population polymorphisms in these genes. Analysis of the existing GWAS data on genetic associations of rates of ageing in human and in dogs would help to approximate a level of heterogeneity of the populations in respect to ageing-related mutations. Analysis of a functional crosstalk between the defined genes would be performed to identify groups of mutations with potentially similar phenotypical outcomes and to re-evaluate approximate levels of real population heterogeneity in respect to ageing dynamics. The examples of one-mutation dependence of ageing dynamics come from the analysis of ageing mutants in *Caenorhabditis elegans* and we would expect a decreasing homogeneity of effect of one mutation in respect to ageing phenotypes from dogs to human, and potentially decreasing weights of sub-populations with the similar ageing phenotype. GWAS studies can be also extremely useful for disclosing and explaining interplays between different causes of death (Tsai et al., 2012).

5.3.3. Modelling the impact of reproductive period on mortality dynamics

A disease or a deleterious mutation that is manifested early in the lifespan (before the reproductive period) has a low probability of passing to the next generation while those manifested after the reproductive period can accumulate in a population. These considerations are central for evolutionary theories of ageing which, despite a certain progress (Kirkwood, 1977, Williams, 1957, Medawar, 1946), have struggled to answer

116

the key questions related to fundamental mortality laws. Here we want to develop a computational model which will describe the evolution of mortality dynamics in a virtual population as affected by the reproductive period. In the model we will consider a number (say, one hundred thousand) of entities representing living organisms. Each entity is characterised by its age and number of heterozygous mutations it carries. At each time step the entity gets older and with some probability (depending on its genotype) dies. When it is in the reproductive age it mates (with some probability) and leaves an offspring. The offspring carries new mutations, for example, with a probability of mutations proportional to the age of the parents if we assume that the mutations in germ cells occur with a constant rate. We will only consider genes for which at least one allele causes disease and affects mortality. We will run the program for a number of generations to see how mortality dynamics evolves in the population and how they are affected by the parameters defining the reproductive age, the effect of mutations on the mortality rate and an age-dependence of the probability of new mutations. In the basic version of this model we will assume that all one-mutation dependent outcomes are independent and there are no interactions between them. We anticipate that in the case of no interactions between diseases, mortality will increase slower than it is described by the Gompertz law. We will further extend the model with various hypotheses (see Section 5.3.2) related to interactions between mutations/diseases in order to find conditions when the Gompertz law will hold. In this way, we can try to identify mechanisms underlying universal mortality laws. The computational model could also verify and extent the analysis performed in Chapter 4.

The outcomes of this study would be compared with the outcomes of other evolutionary theories of mutation accumulation and antagonistic pleiotropy (Charlesworth, 2000, Rose et al., 2007, Medawar, 1946, Williams, 1957). This study would also be extended by the examination of a model of homozygous species. The results obtained using this computational model would be verified on mortality data for pure breeds of domestic animals such as dogs (Kraus et al., 2013) or age-/stress-/mutation-related mortality data on different strains of *C. elegans* (Stroustrup et al., 2013).

5.3.4. Development of mechanistic models for mortality dynamics

An explanation of the Gompertz law could be based on the observation that the exponential function (describing the Gompertz law) occurs naturally as a mathematical

solution of the equation $d\mu/dx = \beta\mu$. This equation is based on an assumption that the rate of change in the force of mortality is proportional to the force of mortality. However, there is not any universal biological justification so far to explain that proportionality. Ideally, the development of a mathematical model based on assumptions about the dynamics of physiological and biological processes that affect mortality rates and is able to biologically verify these processes, and accurately reproduce the observed mortality patterns, could address this problem.

However, in this particular case we would act the other way around. We know that mortality data can be represented as a mixture of exponents. Solutions of systems of linear differential equations are also often represented by a superposition of exponents. Therefore, the mortality data for intrinsic causes of death (superposition of two exponents) can appear as a solution of a coupled system of two differential equations. The two variables in the model should be associated with the physiological states (i.e. vulnerability to diseases and ability to recover) of each individual in a population. Such model can easily be fit to mortality data for intrinsic causes of death and even extended to reproduce the total mortality dynamics (see Figure 5.4). The advantage of this model is that it allows negative exponentiation in the solution (which is forbidden in the heterogeneous model since it can result in immortality).

In our preliminary work we developed the model by denoting u the process of "illness" and v the process of "recovery". We consider that the rate of change of illness increases with probability a over age x, because if an individual is ill then there is a probability that they will get more ill. The rate decreases by probability b as a sick individual could recover, and finally the rate increases by probability c, which is the probability that the individual was healthy and becomes ill. Mathematically, this change is expressed by the differential equation:

$$\frac{du}{dx} = au - bu + cv. \tag{5.14}$$

Similarly, the rate of change of recovery with age increases with probability b if the individual was ill and recovers, and decreases by probability c if the individual was healthy and gets ill:

118

$$\frac{dv}{dx} = bu - cv. \tag{5.15}$$

The system is then converted to a single second order differential equation of variable u, by elimination. The reduction to one equation is achieved by substituting equation (5.15) into the first derivative of equation (5.14) with respect to x. The solution of the 2$^{\text{nd}}$ order ODE is the superposition of two exponents:

$$u(x) = c_1 e^{\lambda_1 x} + c_2 e^{\lambda_2 x}, \tag{5.16}$$

where c_1 and c_2 are constants and λ_1 and λ_2 are the eigenvalues having opposite signs (as the probabilities a, b and c are always positive). A proportion δ of diseases leads the individual to death and therefore the force of mortality is $\delta u(x)$. Hence,

$$\mu(x) = \delta u(x) = C_1 e^{\lambda_1 x} + C_2 e^{\lambda_2 x}, \tag{5.17}$$

where $C_1 = \delta c_1$ and $C_2 = \delta c_2$. The constants can be calculated by using the initial conditions $\mu(0) = \int_0^1 C_1 e^{\lambda_1 x} + C_2 e^{\lambda_2 x}\, dx$ and $\mu'(10) = 0$ (as the minimum point of mortality pattern exist at age 10).

The accidental hump can be modelled by including an extra cause of death, the accidents, by using a Poisson process. In other words, the probability of dying at age x after k accidental events is given by

$$P(k, x) = \frac{(\lambda x)^k}{k!} e^{-\lambda x}, \tag{5.18}$$

where parameter λ is the expected number of events that occur per unit of age. Then, the mortality is expressed as the sum of intrinsic (equation (5.17)) and extrinsic (equation (5.18)) death factors:

$$\mu(x) = C_1 e^{\lambda_1 x} + C_2 e^{\lambda_2 x} + C_3 \frac{(\lambda x)^k}{k!} e^{-\lambda x} \tag{5.19}$$

where C_3 is the fraction of individuals in the population that are dying due to accidents. The fit of equation (5.19) to actual mortality data is shown in Figure 5.4.

The presented model requires a proper biological justification which in turn would provide appropriate interpretations for the model coefficients and parameters. Various

nonlinear interactions between considered variables can be analysed by modifications of the basic model. The model can also be extended (by considering a separate set of variables to describe the dynamics of each disease) for investigation of the interplay between the main death factors (circulatory system, cancer, respiratory system, external, infectious and parasitic diseases and other causes) to the overall dynamics of mortality. By applying this model to different period data we can draw conclusions about the evolution of the interactions between different death factors over time.
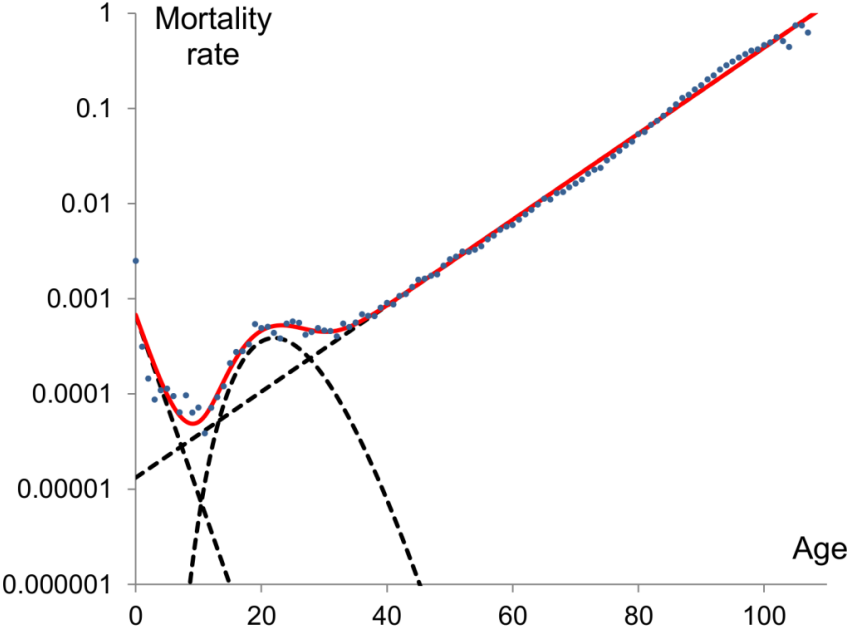

Figure 5.4: Fitting the model of illness-recovery to the 2007 Swedish mortality rates. *The force of mortality as given by equation (5.19) is shown by the red-solid curve while the straight dashed lines indicate the superposition of two inversely directed exponents and the curved dashed line the Poisson distribution term.*

5.4. Conclusion

This thesis is devoted to the analysis of mortality dynamics over age and time by means of mathematical modelling. The model that we have developed and used is based on two main assumptions: the heterogeneity of human populations and the exponential mortality dynamics in each of the constituent subpopulations. On the basis of our analysis we make a few broad conclusions about the mortality-related structure of heterogeneous populations across the lifespan and their evolution over time and on the effect of selective processes on the population mortality dynamics. The model that we have developed links the heterogeneity of populations and the exponential age-specific growth

of mortality, as described in Chapters 2 and 3, and is proven to be an efficient tool for analysing the dynamics of mortality across the lifespan and its alterations over time. Also, as we have shown in Chapter 4, our model not only allows accurate fits to actual mortality data but explains fundamental features of population mortality.

Using mathematical modelling techniques to analyse the dynamics of mortality we have reached the premise of the thesis and we have presented a significant work that matches the scope and the motivation stated in introduction. The major conclusions of this research are: (a) the heterogeneity of human populations can explain the deviations of mortality from its exponential increase at young and very old ages, (b) deviations of mortality from the exponential increase at very old ages can be seen as fluctuations and explained by stochastic effects, (c) heterogeneity can explain controversial observations of late-life mortality (the deceleration, the levelling-off and the decline), (d) the evolution of mortality described by the evolution of parameters of the model of heterogeneous populations confirms the validation of compensation effect in each subpopulation, (e) the mortality-related structure of a heterogeneous population changes over time and for the Swedish population those changes caused its homogenisation within the 20th century, (f) the decrease of the overall mortality within the 20th century is primarily a result of homogenisation of the population and secondly of the reduction of mortality dynamics in constituent subpopulations, (g) if heterogeneity reflects the genetic variation, then common models of population genetics can explain the changes in the mortality-related structure of populations that occur over time.

The outcomes of this research could be beneficial for researchers representing a number of disciplines within academia and may also be potentially beneficial to public and private sectors. On one hand, mathematical modelling of mortality rates is of great interest in order to analyse the dynamics of mortality and the processes underlying ageing in biological populations. On the other hand, it has many practical implementations in actuarial sciences (pension funds, life-insurance companies, government's financial and economic policies) where extrapolation methods are used for the projection of mortality trends in order to estimate future life expectancy and to price several longevity products. In particular, the analysis of evolution of mortality dynamics over time is commonly used for forecasting future mortality patterns and studying the impact of longevity on pension funds.

An extension of this study should explore the biological and genetic factors of the processes underlying ageing and the fundamental role of heterogeneity on the dynamics of mortality. Outcomes of mortality and longevity related studies should result in the development of solid tools to support the goal of building a general theory of mortality and ageing. Such a theory would affect many medical implementations and health service regulations. This indeed will be achieved if scientists understand what fails in the organisms with age in order to potentially contribute to the prevention, delay and treatment of those changes (for example, to delay some inevitable age-related diseases or to prevent some interactions between them). We anticipate that the presented work, and any potential results derived with the implementation of work described in our future plans, will be beneficial to the development (in the long-run) of a theory of ageing and will be of interest to a wide range of researchers and practitioners working in different academic fields as well as in the public and commercial private sector.

References

ADDISON, J. 1711. The vision of Mirza. *Spectator,* 159.

AHMAD, O. B., LOPEZ, A. D. & INOUE, M. 2000. The decline in child mortality: a reappraisal. *Bulletin of the World Health Organization,* 78**,** 1175-1191.

AKAIKE, H. Information theory and an extension of the maximum likelihood principle. *In:* PETROV, B. N. & CSAKI, F., eds. Second International Symposium on Information Theory, 1973. Akadémiai Kiado, 267-281.

ALLEN, L. J. S. 2010. *An introduction to stochastic processes with applications to biology,* Boca Raton, FL, Chapman & Hall/CRC.

ARNOLD (-GAILLE), S. & SHERRIS, M. 2013. Forecasting mortality trends allowing for cause-of-death mortality dependence. *North American Actuarial Journal,* 17**,** 273-282.

AVRAAM, D., ARNOLD, S., JONES, D. & VASIEV, B. 2014. Time-evolution of age-dependent mortality patterns in mathematical model of heterogeneous human population. *Experimental Gerontology,* 60**,** 18-30.

AVRAAM, D., DE MAGALHAES, J. P. & VASIEV, B. 2013. A mathematical model of mortality dynamics across the lifespan combining heterogeneity and stochastic effects. *Experimental Gerontology,* 48**,** 801-811.

BARABASI, A.-L., GULBAHCE, N. & LOSCALZO, J. 2011. Network medicine: a network-based approach to human disease. *Nature Reviews Genetics,* 12**,** 56-68.

BEARD, R. E. 1959. Note on Some Mathematical Mortality Models. *In:* WOLSTENHOLME, G. E. W., O'CONNOR, M. (ed.) *The Lifespan of Animals.* John Wiley & Sons, Ltd.

BEARD, R. E. 1971. Some aspects of theories of mortality, cause of death analysis, forecasting and stochastic processes. *In:* BRASS, W. (ed.) *Biological Aspects of Demography.* London: Taylor and Francis.

BEBBINGTON, M., GREEN, R., LAI, C.-D. & ZITIKIS, R. 2014. Beyond the Gompertz law: exploring the late-life mortality deceleration phenomenon. *Scandinavian Actuarial Journal,* 2014**,** 189-207.

BELL, W. R. 1997. Comparing and assessing time series methods for forecasting age-specific fertility and mortality rates. *Journal of Official Statistics,* 13**,** 279-303.

BLACK, R. E., COUSENS, S., JOHNSON, H. L., LAWN, J. E., RUDAN, I., BASSANI, D. G., JHA, P., CAMPBELL, H., WALKER, C. F., CIBULSKIS, R., EISELE, T., LIU, L. & MATHERS, C. 2010. Global, regional, and national causes of child mortality in 2008: a systematic analysis. *The Lancet,* 375**,** 1969-1987.

BOOTH, H. & TICKLE, L. 2008. Mortality modelling and forecasting: a review of methods. *Annals of Actuarial Science,* 3**,** 3-43.

BRONIKOWSKI, A. M. & FLATT, T. 2010. Aging and its demographic measurement. *Nature Education Knowledge,* 3**,** 3.

BURNHAM, K. P. & ANDERSON, D. R. 2003. *Model selection and multimodel inference: a practical information-theoretic approach*, Springer New York.

CAIRNS, A. J. G., BLAKE, D. & DOWD, K. 2006. A two-factor model for stochastic mortality with parameter uncertainty: theory and calibration. *Journal of Risk and Insurance,* 73**,** 687-718.

CAIRNS, A. J. G., BLAKE, D., DOWD, K., COUGHLAN, G. D., EPSTEIN, D. & KHALAF-ALLAH, M. 2011. Mortality density forecasts: an analysis of six stochastic mortality models. *Insurance: Mathematics and Economics,* 48**,** 355-367.

CAIRNS, A. J. G., DAVID, B., DOWD, K., COUGHLAN, G. D., EPSTEIN, D., ONG, A. & BALEVICH, I. 2009. A quantitative comparison of stochastic mortality models using data from England and wales and the United States. *North American Actuarial Journal,* 13**,** 1-35.

CANUDAS-ROMO, V. 2008. The modal age at death and the shifting mortality hypothesis.

Demographic Research, 19**,** 1179-1204.

CAREY, J. R., LIEDO, P., OROZCO, D. & VAUPEL, J. W. 1992. Slowing of mortality rates at older ages in large medfly cohorts. *Science,* 258**,** 457-461.

CARNES, B. A., HOLDEN, L. R., OLSHANSKY, S. J., WITTEN, M. T. & SIEGEL, J. S. 2006. Mortality partitions and their relevance to research on senescence. *Biogerontology,* 7**,** 183-198.

CARNES, B. A. & OLSHANSKY, S. J. 1997. A biologically motivated partitioning of mortality. *Experimental Gerontology,* 32**,** 615-631.

CARSTENSEN, B. 2007. Age-period-cohort models for the Lexis diagram. *Statistics in Medicine,* 26**,** 3018-3045.

CAVALLI-SFORZA, L. L. & FELDMAN, M. W. 2003. The application of molecular genetic approaches to the study of human evolution. *Nature Genetics,* 33**,** 266-275.

CHARLESWORTH, B. 1994. *Evolution in age-structured populations*, Cambridge University Press.

CHARLESWORTH, B. 2000. Fisher, Medawar, Hamilton and the evolution of aging. *Genetics,* 156**,** 927-931.

CHEN, H. Y., ZAJITSCHEK, F. & MAKLAKOV, A. A. 2013. Why ageing stops: heterogeneity explains late-life mortality deceleration in nematodes. *Biology letters,* 9**,** 20130217.

CHEUNG, S. L., ROBINE, J. M., TU, E. J. & CASELLI, G. 2005. Three dimensions of the survival curve: horizontalization, verticalization, and longevity extension. *Demography,* 42**,** 243-258.

CHMIEL, A., KLIMEK, P. & THURNER, S. 2014. Spreading of diseases through comorbidity networks across life and gender. *New Journal of Physics,* 16**,** 115013.

COALE, A. J. & KISKER, E. E. 1990. Defects in data on old-age mortality in the United States: new procedures for calculating mortality schedules and life tables at the highest ages. *Asian and Pacific Population Forum,* 4**,** 1-31.

CURTSINGER, J. W., FUKUI, H. H., TOWNSEND, D. R. & VAUPEL, J. W. 1992. Demography of genotypes: failure of the limited life-span paradigm in Drosophila melanogaster. *Science,* 258**,** 461-463.

CURTSINGER, J. W., GAVRILOVA, N. S. & GAVRILOV, L. A. 2006. Biodemography of aging and age-specific mortality in Drosophila melanogaster. *Handbook of the biology of aging*. Amsterdam: Elsevier Academic Press.

DE BEER, J. & JANSSEN, F. 2014. The NIDI mortality model. A new parametric model to describe the age pattern of mortality. Netherlands Interdisciplinary Demographic Institute (NIDI).

DE MAGALHAES, J. P., CABRAL, J. A. & MAGALHAES, D. 2005. The influence of genes on the aging process of mice: a statistical assessment of the genetics of aging. *Genetics,* 169**,** 265-274.

DEPOID, F. 1973. Mortality of old people over 85. *Population,* 28**,** 755-792.

DRAPEAU, M. D., GASS, E. K., SIMISON, M. D., MUELLER, L. D. & ROSE, M. R. 2000. Testing the heterogeneity theory of late-life mortality plateaus by using cohorts of Drosophila melanogaster. *Experimental Gerontology,* 35**,** 71-84.

ECONOMOS, A. C. 1979. A non-Gompertzian paradigm for mortality kinetics of metazoan animals and failure kinetics of manufactured products. *Age,* 2**,** 74-76.

ECONOMOS, A. C. 1980. Kinetics of metazoan mortality. *Journal of Social and Biological Structures,* 3**,** 317-329.

FELIPE, A., GUILLEN, M. & PEREZ-MARIN, A. M. 2002. Recent mortality trends in the Spanish population. *British Actuarial Journal,* 8**,** 757-786.

FINCH, C. E. 1994. *Longevity, senescence, and the genome*, University of Chicago Press.

FRIES, J. F. 1980. Aging, natural death, and the compression of morbidity. *The New England Journal of Medicine,* 303**,** 130-135.

FUTUYMA, D. J. 2013. *Evolution*, Sinauer Associates, Inc.

GAILLE, S. 2012. Forecasting mortality: when academia meets practice. *European*

Actuarial Journal, 2, 49-76.

GAILLE, S. & SHERRIS, M. 2011. Modelling mortality with common stochastic long-run trends. *The Geneva Papers on Risk and Insurance-Issues and Practice, 36*, 595-621.

GAVRILOV, L. A. 1984. Does a limit of the life span really exist? *Biofizika, 29*, 908-911.

GAVRILOV, L. A. & GAVRILOVA, N. S. 1979. Determination of species length of life. *Doklady Akademii Nauk SSSR: Biological Sciences, 246*, 905-908.

GAVRILOV, L. A. & GAVRILOVA, N. S. 1991. *The biology of life span: a quantitative approach,* New York, Harwood Academic Publisher.

GAVRILOV, L. A. & GAVRILOVA, N. S. 2001. The reliability theory of aging and longevity. *Journal of Theoretical Biology, 213*, 527-545.

GAVRILOV, L. A. & GAVRILOVA, N. S. 2003. The quest for a general theory of aging and longevity. *Science of Aging Knowledge Environment, 2003*, 1-10.

GAVRILOV, L. A. & GAVRILOVA, N. S. 2006. Reliability theory of aging and longevity. *In:* MASORO, E. J., AUSTAD, S.N. (ed.) *Handbook of the Biology of Aging.* 6th ed. San Diego, CA, USA: Elsevier Academic Press.

GAVRILOV, L. A. & GAVRILOVA, N. S. 2011. Mortality measurement at advanced ages: a study of the Social Security Administration Death Master File. *North American Actuarial Journal* 15, 432-447.

GAVRILOV, L. A., GAVRILOVA, N. S. & IAGUZHINSKII, L. S. 1978. Basic patterns of aging and death in animals from the standpoint of reliability theory. *Zhurnal Obshchej Biologii, 39*, 734-742.

GAVRILOVA, N. S. & GAVRILOV, L. A. 2014. Mortality trajectories at extreme old ages: a comparative study of different data sources on U.S. old-age mortality. *Living to 100 monograph,* 2014.

GAVRILOVA, N. S. & GAVRILOV, L. A. 2015. Biodemography of old-age mortality in humans and rodents. *Journals of Gerontology Series A-Biological Sciences and Medical Sciences, 70*, 1-9.

GBD 2013 MORTALITY AND CAUSES OF DEATH COLLABORATORS 2015. Global, regional, and national age–sex specific all-cause and cause-specific mortality for 240 causes of death, 1990–2013: a systematic analysis for the Global Burden of Disease Study 2013. *The Lancet, 385*, 117-171.

GOMPERTZ, B. 1825. On the nature of the function expressive of the law of human mortality, and on a new mode of determining the value of life contingencies. *Philosophical Transactions of the Royal Society of London, 115*, 513-583.

GREENWOOD, M. & IRWIN, J. O. 1939. The biostatistics of senility. *Human Biology, 11*, 1-23.

HABERMAN, S. & SIBBETT, T. A. 1995. *History of actuarial science,* London, Pickering & Chatto.

HAMILTON, W. D. 1966. The moulding of senescence by natural selection. *Journal of theoretical biology, 12*, 12-45.

HARTL, D. L. & CLARK, A. G. 2007. *Principles of population genetics*, Sinauer Associates.

HAYFLICK, L. 1965. The limited in vitro lifetime of human diploid cell strains. *Experimental Cell Research, 37*, 614-636.

HAYFLICK, L. & MOORHEAD, P. S. 1961. The serial cultivation of human diploid cell strains. *Experimental Cell Research, 25*, 585-621.

HELIGMAN, L. & POLLARD, J. H. 1980. The age pattern of mortality. *Journal of the Institute of Actuaries, 107*, 49-80.

HIDALGO, C. A., BLUMM, N., BARABÁSI, A.-L. & CHRISTAKIS, N. A. 2009. A dynamic network approach for the study of human phenotypes. *PLoS Computational Biology, 5*, e1000353.

HOBCRAFT, J., MENKEN, J. & PRESTON, S. 1982. Age, period and cohort effects in demography: a review. *Population Index, 48*, 4-43.

HORIUCHI, S. & WILMOTH, J. R. 1998. Deceleration in the age pattern of mortality at

older ages. *Demography,* 35**,** 391-412.

HUNT, A. & BLAKE, D. 2014. A general procedure for constructing mortality models. *North American Actuarial Journal,* 18**,** 116-138.

JONES, H. B. 1956. A special consideration of the aging process, disease, and life expectancy. *Advances in Biological and Medical Physics,* 4**,** 281-337.

KANNISTO, V. 1992. Presentation at a workshop on old age mortality at Odense University.

KANNISTO, V. 2001. Mode and dispersion of the length of life. *Population,* 13**,** 159-171.

KANNISTO, V., LAURITSEN, J., THATCHER, A. R. & VAUPEL, J. W. 1994. Reductions in mortality at advanced ages: several decades of evidence from 27 countries. *Population and Development Review,* 20**,** 793-810.

KEYFITZ, N. 1982. *Choice of function for mortality analysis: effective forecasting depends on a minimum parameter representation*, Academic Press.

KEYFITZ, N. & LITTMAN, G. 1979. Mortality in a heterogeneous population. *Population Studies,* 33**,** 333-342.

KIRKWOOD, T. B. L. 1977. Evolution of ageing. *Nature,* 270**,** 301-304.

KIRKWOOD, T. B. L. & AUSTAD, S. N. 2000. Why do we age? *Nature,* 408**,** 233-238.

KIRKWOOD, T. B. L. & HOLLIDAY, R. 1979. The evolution of ageing and longevity. *Proceedings of the Royal Society of London. Series B.*

KOEHLER, A. B. & MURPHREE, E. S. 1988. A comparison of the Akaike and Schwarz criteria for selecting model order. *Journal of the Royal Statistical Society. Series C (Applied Statistics),* 37**,** 187-195.

KRAUS, C., PAVARD, S. & PROMISLOW, D. E. 2013. The size-life span trade-off decomposed: why large dogs die young. *The American Naturalist,* 181**,** 492-505.

KUHA, J. 2004. AIC and BIC: comparisons of assumptions and performance. *Sociological Methods & Research,* 33**,** 188-229.

KUNG, H., XU, J. & MURPHY, S. 2007. *E-stat deaths: preliminary data for 2005 health E-stats* [Online]. Available: http://www.cdc.gov/nchs/products/pubs/pubd/hestats/prelimdeaths05/prelimdeaths05.htm.

LE BRAS, H. 2008. *The nature of demography,* Princeton, N.J. ; Oxford, Princeton University Press.

LEBRETON, J. D. 1996. Demographic models for subdivided populations: the renewal equation approach. *Theor Popul Biol,* 49**,** 291-313.

LEE, R. D. & CARTER, L. R. 1992. Modeling and forecasting U. S. mortality. *Journal of the American Statistical Association,* 87**,** 659-671.

LEXIS, W. 1875. *Einleitung in die theorie der Bevölkerungs-Statistik,* Strasburg, Trübner, K. J. .

LEXIS, W. 1878. Sur la durée normale de la vie humaine et sur la théorie de la stabilité des rapports statistiques. *Annales de Démographie Internationale,* 2**,** 447-460.

LINDOP, P. J. 1961. Growth rate, lifespan and causes of death in SAS/4 mice. *Gerontologia,* 5**,** 193-208.

MAKEHAM, W. M. 1860. On the law of mortality and the construction of annuity tables. *The Assurance Magazine, and Journal of the Institute of Actuaries,* 8**,** 301-310.

MAKEHAM, W. M. 1867. On the law of mortality. *Journal of the Institute of Actuaries,* 13**,** 325-358.

MANTON, K. G., AKUSHEVICH, I. & KULMINSKI, A. 2008. Human mortality at extreme ages: data from the NLTCS and linked medicare records. *Mathematical Population Studies,* 15**,** 137-159.

MANTON, K. G., STALLARD, E. & VAUPEL, J. W. 1981. Methods for comparing the mortality experience of heterogeneous populations. *Demography,* 18**,** 389-410.

MANTON, K. G., STALLARD, E. & VAUPEL, J. W. 1986. Alternative models for the heterogeneity of mortality risks among the aged. *Journal of the American Statistical Association,* 81**,** 635-644.

MANTON, K. G. & TOLLEY, H. D. 1991. Rectangularization of the survival curve:

implications of an ill-posed question. *Journal of Aging and Health,* 3**,** 172-193.

MCNOWN, R. & ROGERS, A. 1989. Forecasting mortality: a parameterized time series approach. *Demography,* 26**,** 645-660.

MCNOWN, R. & ROGERS, A. 1992. Forecasting cause-specific mortality using time series methods. *International Journal of Forecasting,* 8**,** 413-432.

MEDAWAR, P. B. 1946. Old age and natural death. *Modern Q,* 1**,** 30-56.

MEDAWAR, P. B. 1952. *An unsolved problem of biology,* London, H.K. Lewis and Company.

MICHAELIS, L. & MENTEN, M. L. 1913. Die kinetik der invertinwirkung. *Biochemische Zeitschrift,* 49.

MONOD, J. 1949. The growth of bacterial cultures. *Annual Review of Microbiology,* 3**,** 371-394.

MORGAN, B. J. T. 2000. *Applied stochastic modelling*, Arnold Publishers.

MUELLER, L. D., NUSBAUM, T. J. & ROSE, M. R. 1995. The Gompertz equation as a predictive tool in demography. *Experimental Gerontology,* 30**,** 553-569.

MUELLER, L. D. & ROSE, M. R. 1996. Evolutionary theory predicts late-life mortality plateaus. *Proceedings of the National Academy of Sciences of the United States of America,* 93**,** 15249-15253.

NEATH, A. A. & CAVANAUGH, J. E. 2012. The Bayesian information criterion: background, derivation, and applications. *Wiley Interdisciplinary Reviews: Computational Statistics,* 4**,** 199-203.

NJENGA, C. N. & SHERRIS, M. 2011. Longevity risk and the econometric analysis of mortality trends and volatility. *Asia-Pacific Journal of Risk and Insurance,* 5**,** 22-73.

OLSHANSKY, S. J. 1998. On the biodemography of aging: a review essay. *Population and Development Review,* 24**,** 381-393.

OLSHANSKY, S. J. & AULT, A. B. 1986. The fourth stage of the epidemiologic transition: the age of delayed degenerative diseases. *Milbank Quarterly,* 64**,** 355-391.

OLSHANSKY, S. J., CARNES, B. A. & CASSEL, C. 1990. In search of Methuselah: estimating the upper limits to human longevity. *Science,* 250**,** 634-640.

OMRAN, A. R. 2005. The epidemiologic transition: a theory of the epidemiology of population change. *Milbank Quarterly,* 83**,** 731-757.

PARTRIDGE, L. & MANGEL, M. 1999. Messages from mortality: the evolution of death rates in the old. *Trends in Ecology & Evolution,* 14**,** 438-442.

PEARSON, K. 1897. *The chances of death and other studies in evolution,* London, Edward Arnold.

PERKS, W. 1932. On some experiments in the graduation of mortality statistics. *Journal of the Institute of Actuaries (1886-1994),* 63**,** 12-57.

PETTAY, J. E., HELLE, S., JOKELA, J. & LUMMAA, V. 2007. Natural selection on female life-history traits in relation to socio-economic class in pre-industrial human populations. *PLoS One,* 2**,** e606.

PHAM, H. 2011. Modeling U.S. mortality and risk-cost optimization on life expectancy. *IEEE Transactions on Reliability,* 60**,** 125-133.

PITACCO, E. 2004. From Halley to "frailty": a review of survival models for actuarial calculations. *Giornale dell'Istituto Italiano degli Attuari,* 67**,** 17-47.

PLETCHER, S. D. 1999. Model fitting and hypothesis testing for age-specific mortality data. *Journal of Evolutionary Biology,* 12**,** 430-439.

PLETCHER, S. D. & CURTSINGER, J. W. 1998. Mortality plateaus and the evolution of senescence: why are old-age mortality rates so low? *Evolution,* 52**,** 454-464.

PRESTON, S. H., HEUVELINE, P. & GUILLOT, M. 2000. *Demography : measuring and modeling population processes,* Oxford, Blackwell.

PROMISLOW, D. E. L., TATAR, M., PLETCHER, S. D. & CAREY, J. R. 1999. Below-threshold mortality: implications for studies in evolution, ecology and demography. *Journal of Evolutionary Biology,* 12**,** 314-328.

REED, B. C. 1989. Linear least-squares fits with errors in both coordinates. *American*

Journal of Physics, 57**,** 642-646.

RENSHAW, A. E. & HABERMAN, S. 2003. Lee–Carter mortality forecasting with age-specific enhancement. *Insurance: Mathematics and Economics,* 33**,** 255-272.

ROBINE, J. M. 2001. Redefining the stages of the epidemiological transition by a study of the dispersion of life spans: the case of France. *Population,* 13**,** 173-193.

ROGERS, R. G. & HACKENBERG, R. 1987. Extending epidemiologic transition theory: a new stage. *Social Biology,* 34**,** 234-243.

ROSE, M. R. 1991. *The evolutionary biology of aging,* New York, Oxford University Press.

ROSE, M. R., RAUSER, C. L., BENFORD, G., MATOS, M. & MUELLER, L. D. 2007. Hamilton's forces of natural selection after forty years. *Evolution,* 61**,** 1265-1276.

ROSS, S. M. 2002. *Simulation, statistical modeling and decision science*, Academic Press.

ROSS, S. M. 2007. *Simulation*, Elsevier Science.

ROSSOLINI, G. & PIANTANELLI, L. 2001. Mathematical modeling of the aging processes and the mechanisms of mortality: paramount role of heterogeneity. *Experimental Gerontology,* 36**,** 1277-1288.

SACHER, G. A. 1966. The Gompertz transformation in the study of the injury-mortality relationship: application to late radiation effects and ageing. *In:* LINDOP, P. J. & SACHER, G. A. (eds.) *Radiation and Ageing.* London: Taylor and Francis.

SACHER, G. A. & TRUCCO, E. 1962. The stochastic theory of mortality. *Annals of the New York Academy of Sciences,* 96**,** 985-1007.

SCHWARZ, G. 1978. Estimating the dimension of a model. *Annals of Statistics,* 6**,** 461-464.

SHEPARD, D. S. & ZECKHAUSER, R. J. 1980. Long-term effects of interventions to improve survival in mixed populations. *Journal of Chronic Diseases,* 33**,** 413-433.

SHKLOVSKII, B. I. 2005. A simple derivation of the Gompertz law for human mortality. *Theory in Biosciences,* 123**,** 431-433.

SHRYOCK, H. S., SIEGEL, J. S. & LARMON, E. A. 1975. *The methods and materials of demography,* Washington, DC, U.S. Dept of Commerce, Bureau of the Census, U.S. Govt. Printing Office.

SILER, W. 1979. A competing-risk model for animal mortality. *Ecology,* 60**,** 750-757.

STEARNS, S. C., BYARS, S. G., GOVINDARAJU, D. R. & EWBANK, D. 2010. Measuring selection in contemporary human populations. *Nature Reviews Genetics,* 11**,** 611-622.

STEINSALTZ, D. 2005. Re-evaluating a test of the heterogeneity explanation for mortality plateaus. *Experimental Gerontology,* 40**,** 101-113.

STEINSALTZ, D. R. & WACHTER, K. W. 2006. Understanding mortality rate deceleration and heterogeneity. *Mathematical Population Studies,* 13**,** 19-37.

STONE, M. 1979. Comments on model selection criteria of Akaike and Schwarz. *Journal of the Royal Statistical Society. Series B (Methodological),* 41**,** 276-278.

STREHLER, B. L. 1978. *Time, cells, and aging,* New York, Academic Press.

STREHLER, B. L. & MILDVAN, A. S. 1960. General theory of mortality and aging (a stochastic model relates observations on aging, physiologic decline, mortality and radiation). *Science,* 132**,** 14-21.

STROUSTRUP, N., ULMSCHNEIDER, B. E., NASH, Z. M., LOPEZ-MOYADO, I. F., APFELD, J. & FONTANA, W. 2013. The Caenorhabditis elegans lifespan machine. *Nature Methods,* 10**,** 665-670.

SZILARD, L. 1959. On the nature of the aging process. *Proceedings of the National Academy of Sciences of the United States of America,* 45**,** 30-45.

TABEAU, E., VAN DEN BERGH JETHS, A. & HEATHCOTE, C. 2001. *Forecasting mortality in developed countries. Insights from a statistical, demographic and epidemiological perspective,* Dordrecht, Kluwer Academic Press.

TATAR, M., CAREY, J. R. & VAUPEL, J. W. 1993. Long-term cost of reproduction with and without accelerated senescence in Callosobruchus maculatus: analysis of age-specific mortality. *Evolution,* 47**,** 1302-1312.

THATCHER, A. R., KANNISTO, V. & VAUPEL, J. W. 1998. *The force of mortality at ages*

80 to 120, Odense, Odense University Press.

THIELE, P. N. 1872. On a mathematical formula to express the rate of mortality throughout the whole life. *Journal of the Institute of Actuaries,* 16**,** 313-329.

TSAI, K. L., NOORAI, R. E., STARR-MOSS, A. N., QUIGNON, P., RINZ, C. J., OSTRANDER, E. A., STEINER, J. M., MURPHY, K. E. & CLARK, L. A. 2012. Genome-wide association studies for multiple diseases of the German Shepherd Dog. *Mammalian Genome,* 23**,** 203-211.

TURNER, E. L. & HANLEY, J. A. 2010. Cultural imagery and statistical models of the force of mortality: Addison, Gompertz and Pearson. *Journal of the Royal Statistical Society, Series A,* 173**,** 483-499.

VAUPEL, J. W. 2005. Lifesaving, lifetimes and lifetables. *Demographic Research,* 13**,** 597-614.

VAUPEL, J. W. 2010. Biodemography of human ageing. *Nature,* 464**,** 536-542.

VAUPEL, J. W., CAREY, J. R., CHRISTENSEN, K., JOHNSON, T. E., YASHIN, A. I., HOLM, N. V., IACHINE, I. A., KANNISTO, V., KHAZAELI, A. A., LIEDO, P., LONGO, V. D., ZENG, Y., MANTON, K. G. & CURTSINGER, J. W. 1998. Biodemographic trajectories of longevity. *Science,* 280**,** 855-860.

VAUPEL, J. W., MANTON, K. G. & STALLARD, E. 1979. The impact of heterogeneity in individual frailty on the dynamics of mortality. *Demography,* 16**,** 439-454.

VAUPEL, J. W. & YASHIN, A. I. 1985a. The deviant dynamics of death in heterogeneous populations. *Sociological Methodology***,** 179-211.

VAUPEL, J. W. & YASHIN, A. I. 1985b. Heterogeneity's ruses: some surprising effects of selection on population dynamics. *American Statistician,* 39**,** 176-185.

WACHTER, K. W. 1999. Evolutionary demographic models for mortality plateaus. *Proceedings of the National Academy of Sciences of the United States of America,* 96**,** 10544-10547.

WALTER, S., ATZMON, G., DEMERATH, E. W., GARCIA, M. E., KAPLAN, R. C., KUMARI, M., LUNETTA, K. L., MILANESCHI, Y., TANAKA, T., TRANAH, G. J., VOLKER, U., YU, L., ARNOLD, A., BENJAMIN, E. J., BIFFAR, R., BUCHMAN, A. S., BOERWINKLE, E., COUPER, D., DE JAGER, P. L., EVANS, D. A., HARRIS, T. B., HOFFMANN, W., HOFMAN, A., KARASIK, D., KIEL, D. P., KOCHER, T., KUNINGAS, M., LAUNER, L. J., LOHMAN, K. K., LUTSEY, P. L., MACKENBACH, J., MARCIANTE, K., PSATY, B. M., REIMAN, E. M., ROTTER, J. I., SESHADRI, S., SHARDELL, M. D., SMITH, A. V., VAN DUIJN, C., WALSTON, J., ZILLIKENS, M. C., BANDINELLI, S., BAUMEISTER, S. E., BENNETT, D. A., FERRUCCI, L., GUDNASON, V., KIVIMAKI, M., LIU, Y., MURABITO, J. M., NEWMAN, A. B., TIEMEIER, H. & FRANCESCHINI, N. 2011. A genome-wide association study of aging. *Neurobiology of Aging,* 32**,** 2109.e15-2109.e28.

WEIBULL, W. 1939. *A statistical theory of the strength of materials*, Generalstabens litografiska anstalts förlag.

WEIBULL, W. 1951. A statistical distribution function of wide applicability. *Journal of Applied Mechanics,* 18**,** 293-297.

WEITZ, J. S. & FRASER, H. B. 2001. Explaining mortality rate plateaus. *Proceedings of the National Academy of Sciences of the United States of America,* 98**,** 15383-15386.

WEON, B. M. & JE, J. H. 2009. Theoretical estimation of maximum human lifespan. *Biogerontology,* 10**,** 65-71.

WEON, B. M. & JE, J. H. 2011. Plasticity and rectangularity in survival curves. *Scientific Reports,* 1.

WILLIAMS, G. C. 1957. Pleiotropy, natural selection, and the evolution of senescence. *Evolution,* 11**,** 398-411.

WILMOTH, J. R. 1995. Are mortality rates falling at extremely high ages? An investigation based on a model proposed by Coale and Kisker. *Population Studies,* 49**,** 281-295.

WILMOTH, J. R., DEEGAN, L. J., LUNDSTRÖM, H. & HORIUCHI, S. 2000. Increase of

maximum life-span in Sweden, 1861-1999. *Science,* 289**,** 2366-2368.

WILMOTH, J. R. & LUNDSTROM, H. 1996. Extreme longevity in five countries: presentation of trends with special attention to issues of data quality. *European Journal of Population,* 12**,** 63-93.

WILSON, D. L. 1993. A comparison of methods for estimating mortality parameters from survival data. *Mechanisms of Ageing and Development,* 66**,** 269-281.

WOODBURY, M. A. & MANTON, K. G. 1977. A random-walk model of human mortality and aging. *Theoretical Population Biology,* 11**,** 37-48.

WRIGLEY-FIELD, E. 2014. Mortality deceleration and mortality selection: three unexpected implications of a simple model. *Demography,* 51**,** 51-71.

YANG, S. S., YUE, J. C. & HUANG, H.-C. 2010. Modeling longevity risks using a principal component approach: a comparison with existing stochastic mortality models. *Insurance: Mathematics and Economics,* 46**,** 254-270.

YASHIN, A. I., IACHINE, I. A. & BEGUN, A. S. 2000. Mortality modeling: a review. *Mathematical Population Studies,* 8**,** 305-332.

YASHIN, A. I., VAUPEL, J. W. & IACHINE, I. A. 1994. A duality in aging: the equivalence of mortality models based on radically different concepts. *Mechanisms of Ageing and Development,* 74**,** 1-14.

YEN, K., STEINSALTZ, D. & MOBBS, C. V. 2008. Validated analysis of mortality rates demonstrates distinct genetic mechanisms that influence lifespan. *Experimental Gerontology,* 43**,** 1044-1051.