Projection of Mortality Rates with Specific Reference to Immediate Annuitants and Life Office Pensioners

Terry Terieda Zvademoyo Sithole

A Thesis submitted for the degree of Doctor of Philosophy

Faculty of Actuarial Science and Statistics
Sir John Cass Business School, City of London
The City University
London, England

June 2003

Contents

Acknowledgements

I would like to express my gratitude to my supervisors, Prof. Steven Haberman and Prof. Richard Verrall for their advice, encouragement and patience during the course of this research.

I would also like to thank the CMI Bureau for providing the data, GE Frankona, UK and the National University of Science and Technology, Zimbabwe for financial assistance.

Some parts of Chapter 4 (§§ 4.5, 4.6) and Chapter 5 (§§ 5.2, 5.3, 5.4) have been reproduced from a paper published in the journal Insurance: Mathematics and Economics, vol 27 (2000).

During the period of this study, I have had the priviledge of gaining many friends whose encouragement and support has been invaluable and I am thankful. The work is dedicated to Rukudzo, Masimba, Farai and Rumbidzai, their mum Dinah, their grandmother and of course Evangeline Tanaka. The love and support I have received from my family and friends has made it possible for this work to be completed.

And thanks be to God.

Abstract

This thesis investigates the use of parametric models for projecting mortality rates. The basic framework used is that of generalised linear models and can be considered as an extension of the Gompertz-Makeham models (Forfar, McCutcheon and Wilkie, 1988) to include calendar period. The data considered are the CMI ultimate experience for immediate annuitants (male and female) over the period 1946 to 1994 and for life office pensioners (male and female) over the period 1983 to 1996.

The modelling structure suggested by Renshaw, Haberman and Hatzopoulos (1996) is used to investigate the data sets and to determine a range of suitable models, analysing the data by age and calendar period. The properties of these models are investigated and recommendations are made on which models are appropriate for use in projections.

Mortality improvement models are derived from the recommended models and the associated reduction factors are compared with CMI mortality reduction factors.

In addition, the female annuitants' ultimate experience is investigated using a method that combines parametric and time series models to generate forecasts. The procedure used by McNown and Rogers (1989) is used to project forces of mortality over time. The parametric models (Gompertz-Makeham type) are fitted in the framework of generalised linear models. Projected forces of mortality based on the combined parametric-time series model are compared with the projected forces of mortality recommended on the basis of the parametric models.

Chapter 1

Introduction

1.1. Objective and Outline

Forecasting mortality rates involves two different tasks, both of which are of interest to the actuary, (Tolley, Hickman and Lew, 1993). The first type of forecast involves the determination of the future mortality experience of groups of individuals, where the individuals of a group are assumed to be subject to mortality according to some known set of probabilities. For example, we might need to forecast the probability that an individual aged x exact, chosen at random from among the group of individuals, will live to age $x + t$. The principal tool for making such forecasts is the *life table*. The life table is one way of presenting the probability distribution of the random variable representing the future lifetime of an individual.

A second type of forecast involves assessment of future patterns of the mortality process. This approach involves forecasting the changes in the life table itself. As early as 1949, Jenkins and Lew observed that mortality rates were declining and would continue to follow a downward trend to lower levels. The unprecedented improvements in mortality over the last century have led actuaries, as well as demographers, economists and others to seriously deal with the problem of future changes in mortality. It is this second type of forecast that is the subject of this thesis.

In the UK, the Government Actuary's Department and the Continuous Mortality Investigation (CMI) Bureau of the Institute and Faculty of Actuaries undertake regular studies of mortality changes in the population of the UK in general and in the insured sub-population respectively. The Government Actuary's Department produces national population projections, which include a mortality projection basis (ONS publications). The most recently published reports describing trends in the mortality of assured lives in England and Wales are CMIR16 and CMIR17. A series of papers on UK mortality trends have appeared in the actuarial literature (e.g. Daykin, 1996). An international series of studies of mortality trends was started by Gwilt (1956) and continued by Anderson and Whitehead (1960), Giles and Wilkie (1973), Wilkie (1976) and Macdonald *et al* (1998).

For USA mortality experiences, the declining trend in mortality is well documented in mortality studies by the Social Security Administration (Actuarial Studies on United States life tables) and reports in The Transactions of the Society of Actuaries (e.g. Myers and Bayo, 1985, Society of Actuaries, 1981 and Wilkin, 1981).

Under conditions of improving mortality, projection of future mortality rates for annuitants is essential. A company selling life annuities at prices based on mortality levels current at the time the annuities were sold would be likely to make a loss since policyholders would live longer than anticipated (Benjamin and Pollard, 1993). If the standard mortality table used for calculation of annuity rates and reserves predicts higher mortality rates than actually experienced by the policyholders, the financial stability of the company selling the annuities would be in jeopardy since the policyholders would have been undercharged and reserves understated. It is therefore of considerable financial importance for insurance companies to be able to make accurate predictions of future rates of mortality for annuitants, thereby ensuring that they are able to determine correctly premium rates and reserves.

The aim of this study is to develop a model, or models, suitable for predicting future mortality rates for UK annuitants and pensioners. To this end, an analysis of recent trends in the mortality of immediate annuitants and life office pensioners is carried out, with the objective of developing models that take into account both the age variation in mortality, and the underlying time trends in the mortality rates. The main focus is on the modelling structure suggested by Renshaw, Haberman and Hatzopoulos (1996). The basic framework used for analysis is that of generalised linear models and can be considered as an extension of the Gompertz-Makeham models (Forfar, McCutcheon and Wilkie, 1988) to include calendar period. Renshaw *et al* (1996) developed a method of graduating mortality data that incorporates both the age variation in mortality and the underlying time trends. In this study, the focus has been on identifying the particular forms of the Renshaw *et al* (1996) modelling structure that are suitable for forecasting forces of mortality for annuitants and pensioners. The models are then compared with various forecasting methods, particularly the method currently used by the CMI Committee to project pensioners' and annuitants' mortality and the method based on time series analysis used by McNown and Rogers (1989).

The basic textbook on mortality referred to throughout this study is Benjamin and Pollard (1993). There are also frequent references to Elandt-Johnson and Johnson (1980), whose treatment of the subject is primarily focused on clinical data.

Chapter 2 covers measures of mortality (including the life table) and graduation methods. In Chapter 3 methods of projecting mortality are discussed. The results of modelling the immediate annuitants' experience are presented in Chapter 4 while the pensioners' experience is presented in Chapter 5. Chapter 6 is a brief discussion on modelling the experiences using parametric methods in combination with time series methods.

1.2. Notation

μ_x	the force of mortality at age x
μ_{xt}	the force of mortality at age x attained in calendar year t
(x)	a life aged x years exactly
$T(x)$	the future lifetime of (x)
${}_tq_x$	the probability that (x) will die within t years
${}_tp_x$	the probability that (x) will survive for at least t years
$\overset{o}{e}_x$	the complete expectation of life of (x)
e_x	the curtate expectation of life of (x)
R_x^c	the central exposed-to-risk at age x
R_{xt}^c	the central exposed-to-risk at age x in calendar year t

R_x	the initial exposed-to-risk at age x
R_{xt}	the initial exposed-to-risk at age x in calendar year t
a_x	observed number of deaths corresponding to R_x^c or R_x
a_{xt}	observed number of deaths corresponding to R_{xt}^c or R_{xt}
$RF(x,t)$	the mortality improvement factor for a life attaining age x at time t

Chapter 2

Measures of Mortality and Graduation

2.1 Introduction

Benjamin and Pollard (1993) state that '*the purpose of measuring mortality is to enable inferences to be drawn about the likelihood of death occurring within a specific population during a specific period of time*'.

The above statement implies that in assessing an individual's probability of dying (or surviving), we need to consider the number of deaths among a group of individuals under specified conditions. In addition to recording the number of deaths among the group of individuals under observation, we need to know for each individual, the period of *exposure to risk*, that is, the specific period of time during which the death of an individual will actually be recorded and contribute to the observed deaths (Elandt-Johnson and Johnson, 1979). Hence the basic measure of mortality is expressed as a proportion of the number of deaths to the period of exposure to the risk of death. Various age-specific probabilities involving mortality are described in Section 2.2.

Studies in mortality show that the risk of dying varies with a number of factors such as age, sex, and other factors that influence the environment of the people such as geographical location of residence, occupation and nutrition. The variation of mortality with age has always been of great interest to actuaries '*since many of their calculations of contingencies depend upon this variation*' (Benjamin and Pollard, 1993). The study of this variation has led to many attempts to fit analytic or mathematical expressions to the progression of mortality by age. The term '*law of mortality*' is used to describe a mathematical expression for the force of mortality, μ_x (or other measure of mortality) which can be explained from biological or other arguments (Scott, 1996). Such an expression has the advantage that calculations of mortality functions are greatly simplified and statistical inference is facilitated when only a few parameters need to be estimated (Gerber, 1995). A discussion of laws of mortality is covered in Section 2.3.

Most studies in mortality are based on mortality experiences that are samples from much larger experiences. Thus, the death rates derived at individual ages are subject to sampling error. The adjustment procedure that reduces the random errors is referred to as graduation and this procedure is described in Section 2.4. Section 2.5 covers the specific type of graduation involving generalized linear models, which is the method of graduation applied in this study.

2.2 Measures of mortality

In measuring mortality we are essentially concerned with probabilities of death (or equivalently, survival) of an individual at birth or at a given age x. There are various functions that are used as a measure of mortality and some of these functions are described in this section. Many books give detailed descriptions of survival distributions, among them Bowers *et al* (1986), Elandt-Johnson and Johnson (1980) and London (1988).

2.2.1 Probabilities of death and survival

Formulae for probabilities of death and survival have as their basis the distribution of failure time as developed in renewal theory (see for example Cox and Oakes, 1984).

Consider a person aged exactly x years, referred to as a life aged x or denoted by (x), and denote his future lifetime by T or $T(x)$. $T(x)$ is a random variable with a continuous distribution over the range $(0, \infty)$. The cumulative distribution function of $T(x)$

$$F(t) = P(T < t) \tag{2.1}$$

represents the probability that (x) will die between the ages x and $x + t$ years for any fixed value of t and is denoted by the symbol ${}_tq_x$ in International Actuarial Notation.

Complementary to $F(t)$ we have the *survivor function* $S(t)$, representing the probability that (x) survives to age $x + t$. In International Actuarial Notation, $S(t)$ is denoted by the symbol ${}_tp_x$. Thus

$${}_tp_x = S(t) = P(T \geq t) = 1 - F(t) \tag{2.2}$$

This implies that:

$$\left.\begin{array}{ll} {}_0q_x = F(0) = 0; & {}_\infty q_x = F(\infty) = 1 \\ {}_0p_x = S(0) = 1; & {}_\infty p_x = S(\infty) = 0 \end{array}\right\} \tag{2.3}$$

Thus, $F(t)$ is a non-decreasing function of t while $S(t)$ is a non-increasing function of t.

The probability that (x) dies between the ages x and $x + 1$ is denoted q_x while p_x is the probability that (x) survives for at least a year. q_x, the probability that (x) dies within the year is called the *initial rate of mortality* at age x.

The probability density function of the random variable T is

$$f(t) = F'(t) = -S'(t). \tag{2.4}$$

2.2.2 The force of mortality

The *force of mortality* or *hazard rate* of (x) at age $x + t$ is defined as

$$\mu_{x+t} = \lim_{dt \to 0^+} \frac{{}_{dt}q_{x+t}}{dt} \tag{2.5}$$

The force of mortality is a measure of mortality at the instant age $x + t$ is attained, expressed as an annual rate, and may therefore be defined as the instantaneous rate of mortality.

Noting that:

$$F'(t) = f(t) = \lim_{dt \to 0^+} \frac{P(t \le T < t + dt)}{dt}, \tag{2.6}$$

the force of mortality may be written as

$$\mu_{x+t} = \frac{F'(t)}{1 - F(t)}. \tag{2.7}$$

That is,

$$\mu_{x+t} = -\frac{d}{dt}\ln(1 - F(t)) \tag{2.8}$$

which is equivalent to

$$\mu_{x+t} = -\frac{d}{dt}\ln {}_tp_x. \tag{2.9}$$

Integrating (2.9) and using the condition that ${}_0p_x = 1$, we have

$${}_tp_x = \exp\left\{-\int_0^t \mu_{x+u}du\right\} \tag{2.10}$$

from which we obtain the formula:

$$ {}_tq_x = 1 - \exp\left\{ -\int_0^t \mu_{x+u} du \right\}. \tag{2.11} $$

Alternatively, using (2.2) and (2.4), expression (2.7) may be written as

$$ f(t) = {}_tp_x \mu_{x+t} \tag{2.12} $$

so that ${}_tq_x$ may be expressed as

$$ {}_tq_x = \int_0^t {}_up_x \mu_{x+u} du. \tag{2.13} $$

2.2.3 The central mortality rate

The *central mortality rate* m_x is defined as the average mortality rate over the age interval x to x+1, that is, the average risk to which the group of individuals under observation is exposed in the year of age x to x+1. The rate m_x may be expressed as

$$ m_x = \frac{\int_0^1 {}_tp_x \mu_{x+t} dt}{\int_0^1 {}_tp_x dt}. \tag{2.14} $$

By considering the number who survive to exact age x, m_x is seen to be the ratio of the expected deaths between exact ages x and x + 1, to the number of years expected to be lived between exact age x and exact age x + 1.

2.2.4 The expected future lifetime

The expected future lifetime of a life aged x, E[T] is written $\overset{o}{e}_x$ and is referred to as the *complete expectation of life*. The complete expectation of life of (x) is defined by

$$\overset{o}{e}_x = \int_0^\infty t\,{}_tp_x \mu_{x+t} dt \qquad (2.15)$$

On integrating by parts, formula (2.15) simplifies to

$$\overset{o}{e}_x = \int_0^\infty {}_tp_x dt \qquad (2.16)$$

The discrete random variable K or $K(x)$ representing the number of complete years to be lived by (x) in future is known as the *curtate future lifetime* of (x). The probability distribution of the random variable K is given by:

$$\begin{aligned} P(K = k) &= P(k \leq T < k+1) \\ &= S(k) - S(k+1) \\ &= {}_kp_x - {}_{k+1}p_x \,. \end{aligned} \qquad (2.17)$$

That is:

$$P(K = k) = {}_kp_x q_{x+k} \,. \qquad (2.18)$$

We may also deduce (2.18) by general reasoning since

$P(K = k) = P\{(x)$ survives to age $x + k$ and then dies between ages $x + k$ and $x + k + 1\}$

$= P\{(x)$ survives to age $x + k\}.P\{(x + k)$ dies within 1 year$\}$

The expectation of the random variable K is

$$E(K) = \sum_{k=0}^{\infty} k\,{}_kp_x q_{x+k} \qquad (2.19)$$

which simplifies to

$$E(K) = \sum_{k=0}^{\infty} {}_{k+1}p_x \qquad (2.20)$$

E(K) is denoted e_x and is called the *curtate expectation of life*.

An approximation to the relationship between the complete expectation of life and the curtate expectation of life is

$$\overset{o}{e}_x \approx e_x + \frac{1}{2} \qquad (2.21)$$

The approximation can be derived by letting *S* be the random variable representing the fraction of a year during which (*x*) is alive in the year of death, that is

$$T = K + S$$

Then,

$$\mathrm{E}[T] = \mathrm{E}[K + S] = \mathrm{E}[K] + \mathrm{E}[S]$$

S has a continuous distribution between 0 and 1, and approximating the expected value of *S* by ½, the result in (2.21) follows. The result is exact if deaths are uniformly distributed between ages *x* and *x* + 1.

2.2.5 The life table

The probability distribution of the future lifetime of a life aged *x* can be summarised by a *life table*, also called a *mortality table.* Gerber (1995) describes a life table as '*a table of one-year death probabilities* q_x, *which completely defines the distribution of K*'.

The concept of the life table can be traced back to John Graunt whose *Natural and Political Observations Mentioned in a following Index, and made upon the Bills of Mortality*, was published in 1662. Amongst the many observations he made in his essay, Graunt also made observations '*Of the Number of Inhabitants*' and produced a table of survivors from a group of lives followed through life and gradually reducing in numbers by deaths.

The first completed life table was constructed by Edmond Halley who in 1693, published *An Estimate of the Degrees of the Mortality of Mankind, Drawn from Various Tables of Births and Funerals at the City of Breslau.* This life table, called the Breslau Table, was based on the death records of Breslau for the period 1687 to 1691. From this beginning, the life table was developed. The textbooks by Benjamin and Pollard (1993) and Benjamin and Soliman (1993) give excellent historical accounts of the early development of the life table.

The basic function of the modern life table is a function denoted l_x. This is a positive non-increasing continuous function representing the number of lives expected to survive to age x, out of a *cohort* of l_α lives, where α is the youngest age for which l_x is tabulated. The value for l_α is usually taken as some convenient round number such as 1,000,000 or 100,000 and this value is called the *radix* of the life table.

The function l_x is chosen such that ${}_np_x$, the probability that a life aged x exact survives for at least n years, is equal to $\frac{l_{x+n}}{l_x}$. The expected number of deaths between ages x and $x+1$, from among the l_x lives attaining age x is denoted d_x, that is:

$$d_x = l_x - l_{x+1}. \tag{2.22}$$

Various probabilities involving mortality can be determined from the basic function l_x. For example, the probability that a life aged x exact will die within the year is given by:

$$q_x = 1 - p_x = 1 - \frac{l_{x+1}}{l_x} = \frac{l_x - l_{x+1}}{l_x} = \frac{d_x}{l_x},$$

while the central mortality rate is given by:

$$m_x = \frac{d_x}{\int_0^1 l_{x+t}\,dt}.$$

Mortality Rates at Non-Integer Ages

Although l_x is assumed to be continuous and hence defined for all values of x, it is usually tabulated for integral values of x and an approximate distribution of the future lifetime T, is found by interpolation. The approximate distribution of T is dependent on the assumptions made regarding the probability of death ${}_uq_x$, or the force of mortality μ_{x+u}, for x integer and $0 < u < 1$. Batten (1978) provides an extensive review of the various mortality assumptions made in determining rates at non-integral ages. The mortality assumptions are also discussed in other actuarial texts such as Gerber (1995), Bowers *et al* (1986) and Elandt-Johnson and Johnson (1980). Three such assumptions are summarised here.

Uniform distribution of deaths

Under this assumption ${}_uq_x$ is linear over the interval $0 < u < 1$, so that

$${}_uq_x = u.q_x, \text{ for } 0 < u < 1. \tag{2.23}$$

The assumption is based on Abraham de Moivre's hypothesis made in the eighteenth century, that the survivorship curve l_x, could be represented by a single straight line, (Batten, 1978). According to Neill (1977), the uniform distribution of deaths is the most commonly used assumption in interpolating fractional ages.

The Balducci Hypothesis

Under the Balducci hypothesis, it is assumed that the function ${}_{1-u}q_{x+u}$ is linear over the interval $0 < u < 1$, so that

$${}_{1-u}q_{x+u} = (1-u)q_x, \text{ for } 0 < u < 1. \tag{2.24}$$

Gaetano Balducci made extensive use of this hypothesis in his writings (see Batten, 1978, London, 1988).

Constant force of mortality

Under this assumption, the force of mortality μ_{x+u} is assumed to be constant over the unit age interval, that is:

$$\mu_{x+u} = \mu, \text{ for } 0 < u < 1. \tag{2.25}$$

Batten (1978) provides a comprehensive demonstration of the derivation of life table functions based on these three assumptions.

Life tables are useful in many fields of science. For example, demographers use life tables as tools in population projections while actuaries use life tables to build models for insurance systems. A published life table usually contains tabulations, by individual ages, of the basic functions q_x, l_x, d_x, and, possibly, additional derived functions. When the underlying mathematical formula is unknown, values of μ_x can be determined only approximately, (Jordan, 1967 and Neill, 1977). Neill (1977) goes on to discuss the formulae he considers as being the most useful for estimating μ_x. One of the formulae Neill describes is based on expression (2.10), that is

$$_tp_x = \exp\left\{-\int_0^t \mu_{x+u}du\right\}.$$

Taking logs of both sides and substituting $t = 1$ in the above expression, we have

$$-\log p_x = \int_0^1 \mu_{x+u}du. \tag{2.26}$$

The definite integral in expression (2.26) represents the mean value of μ between the ages x and $x + 1$. If this mean value is approximated to be $\mu_{x+1/2}$, then

$$\mu_{x+1/2} \approx -\log p_x \qquad (2.27)$$

Thus an estimate of $\mu_{x+1/2}$ is obtained from p_x (or equivalently, q_x).

Life tables are constructed from observed data and the construction involves estimation, graduation and forecasting techniques. The first two techniques are discussed in Sections 2.2.6 to 2.5 of this chapter and the third is the subject of Chapter 3.

Extensive discussions of life tables are made by Elandt-Johnson and Johnson (1980) and in textbooks on life contingencies such as Jordan (1967), Hooker and Longley-Cook (1953) and Neill (1977). Other books on actuarial science and demography include discussions of life tables. Examples are Pollard (1973), Keyfitz (1977, 1985) and Benjamin and Pollard (1993).

2.2.6 Deriving Crude Mortality Rates

In a mortality investigation, mortality rates have to be estimated from statistical data generated by a group of lives observed for a specified period, referred to as the *observation period*. The rates are derived by dividing the number of deaths of lives having a particular age definition, by the corresponding *exposed-to-risk*. The method of tabulating deaths will determine the ages at which mortality rates will be determined. The definition of age for the deaths is therefore of fundamental importance in estimating mortality rates. Batten (1978) and Benjamin and Pollard (1993) give detailed descriptions of methods of tabulating mortality data.

Exposed-to-risk

Consider a given population which is observed over a calendar year period from time $t = 0$ to $t = 1$ say, and denote by $P_x(t)$ the number of lives aged x at time t, where the age x is determined by reference to any one of the following points in time:

a) an individual's birthday (*life year*); or
b) a fixed point in the *calendar year*; or

c) the anniversary of the date the individual effected an insurance policy or joined a pension scheme, etc (*policy year*).

The period of time during which an individual's recorded age remains the same is called the *rate interval.*

The *central exposed-to-risk,* R_x^c at age x is the number of years of life lived by the $P_x(t)$ lives in the age range (x, x+1) during the year of observation. The corresponding number of deaths is denoted a_x. The central exposed-to-risk for a one-year observation period can be expressed as:

$$R_x^c = \int_0^1 P_x(t)dt\,. \tag{2.28}$$

For a mortality investigation covering a period of T years, R_x^c is:

$$R_x^c = \int_0^T P_x(t)dt\,. \tag{2.29}$$

If the period of exposure for each death is continued up to the end of the year (i.e. age $x + 1$ and time $t = 1$) and the total additional exposure is added to the central exposed-to-risk, the *initial exposed-to-risk* R_x, is obtained. Assuming that deaths occur uniformly over the year of age, the initial and central exposed-to-risk are related by:

$$R_x \approx R_x^c + \frac{a_x}{2} \tag{2.30}$$

Dividing a_x by R_x^c leads to an estimate of the force of mortality $\mu_{x+\frac{1}{2}}$ (or of the central mortality rate m_x) which may be denoted $\mu^*_{x+\frac{1}{2}}$ (or m^*_x) and division of a_x by R_x provides an estimate of the initial rate of mortality q_x which may be denoted q^*_x. The rates derived from the data in this manner are referred to as *age-specific crude rates*.

The central exposed-to-risk can be evaluated exactly using the *exact exposure method*, a method that involves counting the number of days of exposure for each life during the period of observation. However, in many practical situations, detailed data are available for deaths only and hence the exact exposure method cannot be adopted (Benjamin and Pollard, 1993). In this case, provided the values of $P_x(t)$ are available at specific points in time t (e.g. at $t = 0$ and $t = 1$ for a one-year observation period), then the values at these *census dates* can be used to obtain an estimate of R_x^c. Assuming that $P_x(t)$ varies linearly during the year of investigation, the trapezoidal rule of integration can be used to obtain an approximate integration over the period. This method of approximating the central exposed-to-risk is known as the *census method.*

The census formula approximations for the central exposed-to-risk depend on the type of rate interval being used. For a life-year or policy-year rate interval, assuming that birthdays or policy anniversaries and entries to and exits from the experience are uniformly spread over the year, the census formula is:

$$R_x^c \approx \tfrac{1}{2}\{P_x(0) + P_x(1)\}. \tag{2.31}$$

For a calendar-year rate interval, the census formula is:

$$R_x^c \approx \tfrac{1}{2}\{P_x(0) + P_{x+1}(1)\}. \tag{2.32}$$

The corresponding formulae for a more general investigation period, from time $t = 0$ to $t = T$ for a life-year or policy-year rate interval is:

$$R_x^c = \tfrac{1}{2}P_x(0) + \sum_{t=1}^{T-1} P_x(t) + \tfrac{1}{2}P_x(T) \tag{2.33}$$

and for a calendar-year rate interval the formula is:

$$R_x^c = \sum_{t=0}^{T-1} \tfrac{1}{2}\{P_x(t) + P_{x+1}(t+1)\}. \tag{2.34}$$

For any rate interval, the actual age to which the age label x applies is determined by the method of tabulating deaths. The derived crude rate is often not at exact age x (or $x + \frac{1}{2}$ for μ) but at some other age:

$$x + f \text{ (or } x + f + \tfrac{1}{2} \text{ for } \mu\text{)}, \tag{2.35}$$

which is the average actual age at the beginning of the rate interval.

Select Rates

The rates considered above depend only on age. Mortality functions might vary not only with age but also with the time that has elapsed since the individual entered into the class of lives under consideration. For example, the class of lives might be assured lives where medical evidence is obtained before acceptance; or annuitant lives where a person who is ill is unlikely to purchase an annuity. Based on this information it would be reasonable to assume that a life aged x at entry experiences lighter mortality than that of lives aged x in the group as a whole. Hence for lives aged x at entry, the form of mortality at age $[x] + t$, that is t years after entry into the specific class of lives, would be a function of both the attained age $[x] + t$ and the *duration t*. The square brackets are used to identify the variable representing the age at *selection*. Thus for example, $l_{[x]+t}$ denotes the number of lives expected to survive to age $x + t$ out of the $l_{[x]}$ lives who are select at age x, $d_{[x]+t}$ is the number expected to die between ages $x + t$ and $x + t + 1$ and $q_{[x]+t}$ is the one-year probability of death at age $x + t$. Rates according to age and duration are referred to as *select rates*.

The relationships between mortality functions in a non-select life table apply equally to the select table provided we are considering functions that apply to the same age at entry. Hence an equivalent expression for the number of lives expected to die between ages x and $x + 1$ is:

$$d_{[x]} = l_{[x]} - l_{[x]+1}. \tag{2.36}$$

The mortality advantage for select lives diminishes and becomes negligible for practical purposes after a few years. The period for which the duration effect is significant is called the *select period*. For durations equal to or greater than the select period, rates relating to age only would be required and these are referred to as *ultimate rates*.

Consider an individual whose exact age at entry (into assurance say) is x. With a select period of 2 years, the initial rates of mortality needed would be as follows:

$$q_{[x]},\ q_{[x]+1},\ q_{x+2},\ q_{x+3},\ \ldots \tag{2.37}$$

A life table that varies only with the attained age x is called an *aggregate life table*.

Select rates are not always lighter than non-select rates. For example, entry into 'select status' might mean retirement from a pension scheme due to ill health, in which case mortality is likely to be heavier. In this thesis, we are concerned with the mortality of annuitants and life office pensioners who retire at normal retirement age, and hence for these experiences, the effect of selection is a reduction in mortality, leading to the inequalities:

$$q_{[x]} < q_{[x-1]+1} < q_{[x-2]+2} < \cdots \tag{2.38}$$

A comprehensive discussion of *selection* in the context of mortality studies, including the different ways in which selection can arise, is provided by Benjamin and Pollard (1993). The type of selection described here is referred to as *temporary initial selection*, and is also called *self selection* for purchasers of annuities.

In the case of CMI data, each of the contributing offices carries out a census of the number of lives, or policies, or pound amounts in force at 1 January and 31 December of each year, subdivided by age nearest birthday, and for most investigations, by the number of complete years since the policy was effected (*curtate duration*). The number of deaths is recorded by lives, policies or amounts as appropriate, subdivided by age nearest birthday at death and where necessary, also by curtate duration at death. Hence for CMI data, we have a *life year rate interval* for age and a *policy year rate interval* for duration. Therefore the value of f in expression (2.35) is $-1/2$ so that a_x/R_x gives an estimate of $q_{x-1/2}$ while a_x/R_x^c provides an estimate of μ_x or $m_{x-1/2}$ (see for example, Forfar *et al*, 1988). The maximum select period available for CMI data is 5 years. In contrast, the continuous mortality investigations conducted by the Society of Actuaries in the USA have a select period of 15 years.

2.3 Laws of mortality

According to Jordan (1967), the earliest proposed law of mortality was that of *De Moivre* (1724). De Moivre postulated the existence of a maximum age w for human beings and assumed that the random variable T, the future lifetime of (x), was uniformly distributed over the range 0 to $w - x$. The formula for the force of mortality at age $x + t$ then becomes:

$$\mu_{x+t} = \frac{1}{w - x - t}, \quad \text{for } 0 < t < w - x \tag{2.39}$$

De Moivre himself noted that his formula was a rough estimation of the pattern of mortality and recommended that the assumption be used only for the age range 12 to 86 years (Jordan, 1967). A spin-off from De Moivre's formula that is in wide use today is the assumption of uniform distribution of deaths in estimating rates at non-integral ages (Batten, 1978). It should however be noted that this assumption (described by equation (2.23)), is applied on a year by year basis unlike (2.39).

Laws of mortality subsequently developed are reviewed in Sections 2.3.1 to 2.3.3.

2.3.1 Gompertz and Makeham laws

The most famous laws of mortality are those of Gompertz and Makeham. *Gompertz* (1825) postulated that the force of mortality would grow exponentially. That is:

$$\mu_x = Bc^x; \quad \text{where } B, c \text{ are constants and } B > 0, c > 1. \qquad (2.40)$$

Gompertz restricted the use of his formula to ages ranging from 10 to 55 or 15 to 60 (Jordan, 1967).

Applying the log transformation to (2.40) results in the log-linear relationship

$$\log \mu_x = \log B + x \log c \qquad (2.41)$$

which implies that if log μ_x is plotted against age x, the graph would be approximately linear. The constant B reflects the general level of mortality in the population under study while c reflects the rate at which the force of mortality increases with age.

Equivalently, the Gompertz function may be expressed as:

$$\mu_x = \exp(\beta_0 + \beta_1 x) \qquad (2.42)$$

where $B = e^{\beta_0}$, and $c = e^{\beta_1}$.

Makeham (1860) suggested an improvement of Gompertz law by adding a constant term to the formula. Thus Makeham's law assumes that

$$\mu_x = A + Bc^x, \qquad (2.43)$$

Makeham's law reflects the division of causes of death into those due to chance, reflected by the constant A and those due to deterioration (Benjamin, 1964).

As for Gompertz law, an equivalent form for Makeham's law is:

$$\mu_x = \alpha + \exp(\beta_0 + \beta_1 x) \qquad (2.44)$$

Jordan (1967) and Benjamin and Pollard (1993) observe that both Gompertz's and Makeham's laws possess properties of practical importance in the manipulation of functions involving more than one life so that both laws continue to be used today. Jordan (1967) and Neill (1977) demonstrate the evaluation of joint-life and contingent functions under these laws.

In general, the Gompertz law has been found to provide a good fit to mortality data at the adult ages. Recent studies include the study by Wetterstrand (1981) who used the Gompertz model to analyse US life insurance mortality data over the period 1948 to 1977 and showed that the Gompertz model was a good fit for ages 30 to 90. The Gompertz curve has also been found to provide a good fit at the higher ages for some UK experiences. For example, Humphrey (1970) derived mortality rates at ages 86 to 104 based on deaths in England and Wales in 1942-57 and found that the rates followed the Gompertz curve. Thatcher (1987), remarked in his study of mortality at the highest ages that '*mortality rates at high ages in England and Wales still look remarkably like Gompertz curves*'. However, Wilkin (1981) carried out a study of mortality among the aged based on US mortality experiences and concluded that above age 90 the observed patterns showed deviations from the Gompertz curve. Other studies where the validity of the Gompertz model has been questioned have been by individuals such as Myers and Bayo (1985) and Coale and Kisker (1990). Tuljapurkar and Boe (1998) observe that in spite of the problems associated with the model, the '*Gompertz picture of an exponentially rising force of mortality at old ages has strongly influenced much work on old-age mortality patterns*'.

After the early years of the 20th century, changes in the basic age-pattern of mortality meant that it was increasingly difficult to obtain satisfactory graduations over the whole lifetime using Makeham's formula, (Benjamin and Pollard, 1993). In an attempt to find a 'law' of mortality that would represent mortality experience over a wider range of ages more accurately, more complex mathematical formulae have been developed. Benjamin and Pollard (1993) provide an extensive review of the mathematical curves developed and a summary of some of these laws is presented in Section 2.3.2 below.

2.3.2 Other laws of mortality

Thiele (1872), postulated the relationship between the force of mortality μ and age x to be

$$\mu_x = a_1 e^{-b_1 x} + a_2 e^{\left[-\frac{1}{2} b_2 (x-c)^2\right]} + a_3 e^{b_3 x}, \tag{2.45}$$

where the first term is a decreasing Gompertz curve representing childhood mortality, the last term is a Gompertz curve representing old age mortality and the middle term is a normal curve representing mortality in adulthood.

Perks (1932) proposed a family of curves of the form

$$\mu_x = \frac{A + Bc^x}{1 + Dc^x} \tag{2.46}$$

and

$$\mu_x = \frac{A + Bc^x}{Kc^{-x} + 1 + Dc^x} \tag{2.47}$$

According to Benjamin and Pollard (1993), Perks' formulae represented the most promising attempt to fit a single curve to the whole range of ages at the time.

Weibull (1939) suggested the formula

$$\mu_{x+t} = k(x+t)^n \tag{2.48}$$

with fixed parameters $k > 0$ and $n > 0$.

In graduating *English Life Table Number* 11, based on the deaths in England and Wales in 1950 to 1952, and the population census of 1951, the formula adopted was:

$$m_x = a + \frac{b}{1 + e^{[-\alpha(x - x_1)]}} + ce^{\left[-\beta(x - x_2)^2\right]}. \tag{2.49}$$

The same form of curve was used to produce English Life Table Number 12, which was based on the deaths in England and Wales in the years 1960 to 1962 and the population census of 1961.

Formula (2.49) is seen to be a combination of a logistic curve with a symmetrical normal curve. Benjamin and Pollard (1993) observe that with 7 parameters to be estimated for each curve, the formula seems to provide a complicated method of graduating population data that generally requires little graduation.

The CMI Committee of the Institute and Faculty of Actuaries produced a new standard mortality table based on the pooled mortality experience of contributing life insurance offices for the years 1949 to 1952. Beard derived the following formula used in producing the table:

$$q_x = A + \frac{Bc^x}{Ec^{-2x} + 1 + Dc^x}. \tag{2.50}$$

Clearly (2.50) is related to Perks' family of curves.

Barnett suggested the following formula for graduating the United Kingdom assured lives' mortality experience for the years 1967 to 1970 (Joint Mortality Investigation Committee, 1974):

$$\frac{q_x}{p_x} = A - Hx + Bc^x \tag{2.51}$$

which is equivalent to

$$q_x = \frac{A - Hx + Bc^x}{1 + A - Hx + Bc^x}. \tag{2.52}$$

Experiments with the pensioners' experience for the same period indicated the following formula to be appropriate:

$$q_x = \frac{\exp\{P(x)\}}{1+\exp\{P(x)\}}, \tag{2.53}$$

where $P(x)$ is a polynomial in age x.

Formula (2.53) may be expressed as:

$$\ln\left(\frac{q_x}{p_x}\right) = P(x). \tag{2.54}$$

Graduations of the pensioners' and annuitants' mortality experiences for the period 1967-1970 were carried out using formula (2.54).

More recently, the following formula proposed by Heligman and Pollard (1980) has produced promising results over the whole life span:

$$\frac{q_x}{p_x} = A^{(x+B)^C} + D\exp\left\{-E(\ln x - \ln F)^2\right\} + GH^x. \tag{2.55}$$

The Heligman and Pollard law has a structure that is similar to (2.45), the law of mortality postulated by Thiele (1872). Whereas Thiele's law decomposes the force of mortality μ_x into three components, the Heligman and Pollard law decomposes q_x, the probability that a life aged x exact will die before age $x + 1$ into three components: an infant and child mortality curve; a hump representing mortality due to accidents at the younger adult ages; and a Gompertz curve representing mortality at older ages. Hence, when considering mortality at the older ages only, the first two terms can be neglected so that the Heligman and Pollard law is very similar to the Gompertz law at the older ages (see Thatcher 1990, Congdon 1993).

The formula has been used in studies by Forfar and Smith (1987) using data from English Life Tables, and McNown and Rogers (1989) on American mortality data.

2.4 Graduation

Graduation has been defined in various forms in the actuarial literature such as:
'*the process of securing, from an irregular series of observed values of a continuous variable, a smooth regular series of values consistent in a general way with the observed series of values*', Miller (1946);

'*an effort to represent a physical phenomenon by a systematic revision of some observations of that phenomenon*', Andrews and Nesbitt (1965);

'*the principles and methods by which a set of observed (or crude) probabilities is adjusted to provide a suitable basis for inferences to be made and further practical calculations to be made*', Haberman and Renshaw (1996).

Each of the above statements implies that graduation involves the revision of an initial set of data to produce a better representation of the underlying true values.

The most common application of graduation in actuarial science is the graduation of mortality rates with respect to age. We consider a set of age-specific crude probabilities of death q_x^*, or forces of mortality μ_x^* calculated from observed data. If the true rates of mortality were independent '*then the crude values would be our final estimates of the true underlying rates*', Haberman and Renshaw (1996). However, a prior opinion most frequently used in graduating mortality rates is that the true underlying rates at neighbouring ages are related and progress smoothly from age to age. This is because age is a continuous variable and any effect it has on mortality would be expected to change gradually, with a few exceptions at certain ages. Elphistone (1951) emphasized the relations between neighbouring mortality rates when he stated that: '*The theory of graduation is the theory of relations between neighbouring rates…*'.

If the sample sizes at each age could be increased to infinity, the crude rates would represent the true underlying rates, and would therefore progress smoothly with age under the hypothesis that mortality rates at neighbouring ages are related. In practice, there are severe limitations to sample sizes. The solution is to smooth the individual crude rates to obtain improved estimates of the unknown underlying $\{q_x\}$ or $\{\mu_x\}$ rates. '*This is done by systematically revising the crude values to remove the random fluctuations*', Haberman and Renshaw (1996). Benjamin and Pollard (1993) define graduation as: '*the adjustment procedure that reduces the random errors in the observed rates as well as smoothing them...*'.

Keyfitz (1982) and Bloomfield and Haberman (1987) list the following five distinct uses of graduation:

(a) *To smooth the data.* Graduation facilitates the processing of the data, makes it easier to handle and removes awkward irregularities and inconsistencies. A part of the reason under this heading might be called aesthetic, i.e. to make the set of rates look better. For insurance tables, if a reasonable degree of smoothness has not been achieved, complicated derived functions such as policy values might display worrying irregularities.

(b) *To make the results more precise*, on the assumption that the true experience underlying the observations follows a smooth curve.

(c) *To aid inferences from incomplete data.* In those populations for which complete registration of events, like births and deaths is not available, indirect methods of estimation based on graduation are important.

(d) *To facilitate comparisons of mortality.* One would like to be able to compare the mortality of two populations, or of two cohorts or of one population at two points in time, summarizing the difference in a set of parameters.

(e) *To aid forecasting and projection.* A clear progression over recent time in the values of a set of parameters enables extrapolation into the future to be used for forecasting of probabilities, rates and derived functions such as the life table.

Additionally, Keyfitz (1982) gives one other purpose of graduation which is *to construct life tables.*

The purpose of this study is to develop a model suitable for projecting mortality rates and hence we are ultimately concerned with (e). Since we are dealing with insurance data, objectives (a) and (b) are an essential part of the graduation process in this case.

Graduation methods are generally divided into two classes: parametric and non-parametric. In parametric graduation, an analytic expression, say $f(x)$, is used to represent all or part of the age pattern of mortality in terms of q_x or μ_x, or other mortality measure. London (1985) observes that our prior opinion regarding the smoothness and shape of the sequence of underlying mortality values is implicit in the form of the analytic expression chosen, so that $f(x)$ is expected to be a smoothly progressing continuous function of age x. The models of Gompertz and Makeham are examples of analytic expressions used in graduating mortality.

In non-parametric graduation, no analytic expression is specified although in some cases, a functional form is implicitly assumed for the underlying mortality rates. Examples of non-parametric graduation methods include moving-weighted-averages, Whittaker-Henderson graduation and Kernel smoothing.

There has been extensive written work on the subject of graduation in the actuarial literature, and statistical and applied mathematical periodicals. Benjamin and Pollard (1993) describe the more recent methods of graduation such as the graphic method; graduation by mathematical formula; graduation with reference to a standard mortality table and graduation using splines. Forfar *et al* (1988) give a full description of the graduation methodology currently used by the CMI Bureau to produce standard mortality tables for use by the UK life insurance industry.

London (1985) discusses both parametric and non-parametric graduation while Copas and Haberman (1983), Bloomfield and Haberman (1987), Gavin *et al* (1993, 1994, 1995) and Verrall (1996) discuss non-parametric methods.

Recent papers by Renshaw (1991, 1995), Renshaw and Haberman (1997) and Verrall (1996) demonstrate the applications of generalized linear models to graduation with respect to age. Renshaw *et al* (1996) and Renshaw and Hatzoupoulos (1996) develop the applications further by proposing a modelling structure in the framework of generalized linear models which incorporates both the age-variation in mortality and underlying time trends in the mortality rates.

2.4.1 Statistical Considerations

The two measures of mortality of primary interest in graduation, q and μ, are associated with two probability models: the binomial model and the Poisson model respectively. Batten (1978) and Elandt-Johnson and Johnson (1980) present thorough reviews of the estimation of q. Many other authors have discussed models for q and μ, among them are Forfar *et al* (1988) who describe in detail the estimation of both parameters while Benjamin and Pollard (1993) describe the model for q.

For brevity, in this section it is assumed that the mortality rates depend only on the age of the individual.

Estimation of q

Given the initial exposed-to-risk at age x, R_x, and the corresponding observed number of deaths a_x, the initial rate of mortality q_x is modelled using the binomial distribution. Under the binomial model, A_x, the random variable representing the number of deaths occurring in the year, has a binomial distribution with parameters q_x and R_x. It is assumed that the death or survival of each individual is independent of the death or survival of each of the other individuals and that the probability of death for each individual is the same. It is also assumed that the individuals are observed from exact age x to exact age $x + 1$ or until prior death.

The probability of a_x deaths is therefore given by:

$$P(A_x = a_x) = \binom{R_x}{a_x} q_x^{a_x} (1 - q_x)^{R_x - a_x} . \qquad (2.56)$$

Further, the *likelihood* of obtaining a_x deaths exactly is

$$L(\underline{q}) = \prod_x \binom{R_x}{a_x} q_x^{a_x} (1 - q_x)^{R_x - a_x} \qquad (2.57)$$

Ignoring the first term in expression (2.57), which does not depend on q_x, the likelihood becomes:

$$L(\underline{q}) \propto \prod_x \frac{q_x^{a_x}}{(1-q_x)^{a_x}}(1-q_x)^{R_x} \tag{2.58}$$

The natural logarithm of $L(\underline{q})$ is

$$\ell(\underline{q}) = \log L(\underline{q}) = \sum_x a_x \log\left(\frac{q_x}{1-q_x}\right) + R_x \log(1-q_x) \tag{2.59}$$

The value of $\underline{q}$ which maximises the likelihood $L(\underline{q})$ is the same as that which maximises the log-likelihood $\ell(\underline{q})$. By differentiating (2.59) and equating to zero, the *maximum likelihood estimator* of q_x, $\hat{q}_x$ is determined as:

$$\hat{q}_x = \frac{A_x}{R_x} \tag{2.60}$$

The expected value of $\hat{q}_x$ is q_x and the variance is $\frac{q_x(1-q_x)}{R_x}$.

The traditional actuarial approach described fully by Batten (1978) is to assume that the binomial model is applicable even when the '*total exposure is made up from a number of shorter periods of exposure*', Forfar *et al* (1988).

Estimation of μ

Given the central exposed-to-risk at age x, R_x^c and the corresponding observed deaths a_x, the appropriate probability model is the Poisson distribution with parameter $\mu_x R_x^c$. The choice of the Poisson distribution is based on the assumption that the force of mortality is constant in each age interval x to $x + 1$. Sverdrup (1965) discusses this point in detail with particular reference to multiple state models.

Under the Poisson model, the probability of a_x deaths occurring is:

$$P(A_x = a_x) = \frac{(\mu_x R_x^c)^{a_x} \exp(-\mu_x R_x^c)}{a_x!}. \tag{2.61}$$

The likelihood of obtaining a_x deaths is:

$$L(\underline{\mu}) = \prod_x \frac{(\mu_x R_x^c)^{a_x} \exp(-\mu_x R_x^c)}{a_x!} \tag{2.62}$$

and the log-likelihood is:

$$\ell(\mu) = \sum_x \left(a_x \log R_x^c + a_x \log \mu_x - \mu_x R_x^c\right) \tag{2.63}$$

ignoring the denominator which is not dependent on μ_x. By differentiating the log-likelihood (2.63) and equating to zero, the maximum likelihood estimator for μ_x is:

$$\hat{\mu}_x = \frac{A_x}{R_x^c}. \tag{2.64}$$

The expected value of $\hat{\mu}_x = \mu_x$ and the variance is $\frac{\mu_x}{R_x^c}$.

If the underlying mortality is assumed to follow a particular mathematical formula, μ_x or q_x will be a function of the unknown parameters in that formula. The maximum likelihood estimates of the unknown parameters are the values which maximise the log-likelihood (2.63) when modelling μ_x, or (2.59) when modelling q_x. In this thesis, the focus is on the force of mortality μ_x and hence we are concerned with parameter estimates that maximise the log-likelihood (2.63), which is equivalent to maximising the likelihood (2.62).

2.4.2 The current CMI graduation practice

A comprehensive presentation of the graduation methodology used by the CMI Committee to graduate the 1979-82 mortality experiences (CMIR 9, CMIR 10) and the 1991-94 mortality experiences (CMIR 16, CMIR 17), is given by Forfar *et al* (1988).

The Committee graduated μ_x or q_x using the "Gompertz-Makeham" class of functions defined as:

$$GM_x(r,s) = \sum_{i=0}^{r-1} \alpha_i x^i + \exp\left(\sum_{j=0}^{s-1} \beta_j x^j \right) \tag{2.65}$$

and the "Logit Gompertz-Makeham" formulae defined by:

$$LGM_x(r,s) = \frac{GM_x(r,s)}{1 + GM_x(r,s)}. \tag{2.66}$$

The Gompertz-Makeham formula $GM_x(r,s)$ is subject to the convention that when r=0, the first group of terms is absent, and when s = 0, the second group of terms is absent. The cases $\mu_x = GM_x(0,2)$, $\mu_x = GM_x(1,2)$ and $q_x = LGM_x(2,2)$ correspond respectively to the Gompertz formula (2.42), the Makeham formula (2.44) and the Barnett formula (2.51). Formula (2.53), used to graduate the pensioners' and annuitants' experiences for the years 1967 to 1970, is equivalent to $q_x = LGM_x(0,n)$ for some positive integer n.

Thus, for example,

$$GM_x(0,2) = \exp(\beta_0 + \beta_1 x) \tag{2.67}$$

$$GM_x(0,3) = \exp(\beta_0 + \beta_1 x + \beta_2 x^2) \tag{2.68}$$

$$GM_x(1,3) = \alpha_0 + \exp(\beta_0 + \beta_1 x + \beta_2 x^2) \tag{2.69}$$

For convenience, the Gompertz-Makeham and the Logit Gompertz-Makeham formulae were expressed in terms of orthogonal polynomials, with the age variable x transformed to $x' = (x - u)/v$ where u and v were chosen to be 70 and 50 respectively.

Orthogonal Polynomials

The particular set of orthogonal polynomials used is Chebycheff polynomials of the first type defined by:

$$C_k(\cos\theta) = \cos k\theta$$

so that

$$C_0(x) = 1, \qquad C_1(x) = x, \tag{2.70}$$

and by the addition formula for trigonometric functions:

$$C_{n+1}(x) = 2xC_n(x) - C_{n-1}(x), \qquad n \geq 1. \tag{2.71}$$

In general, the continuous form of orthogonal polynomials in relation to a given interval $[a, b]$ is obtained by constructing a sequence of polynomials $\{p_i(x),\ i = 0,1,2,\ldots\}$ such that

$$\int_a^b w(x)p_r(x)p_s(x)dx = \begin{cases} 0 & if \quad r \neq s \\ e_r & if \quad r = s \end{cases} \tag{2.72}$$

where $w(x)$ is a given positive weight function defined on the interval $[a, b]$ and e_r is some non-zero real number (Forfar *et al* 1988).

The Chebycheff polynomials are orthogonal on the interval [-1, 1] with $w(x) = \left[1 - x^2\right]^{-\frac{1}{2}}$. Over the same interval, when $w(x) \equiv 1$, we obtain the Legendre polynomials $\{L_i(x)\}$ defined by the initial equations:

$$L_0(x) = 1, L_1(x) = x, \tag{2.73}$$

and the recurrence relation:

$$(n+1)L_{n+1}(x) = (2n+1)xL_n(x) - nL_{n-1}(x) \quad \text{(integer } n \geq 1) \tag{2.74}$$

Numerical analysis textbooks such as Conte and de Boor (1980), give a detailed discussion of the derivation of orthogonal polynomials.

The use of orthogonal polynomials means that if $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \ldots, \alpha_n)$ defines the best-fitting polynomial of order n, and $\boldsymbol{\alpha'} = (\alpha'_1, \alpha'_2, \ldots, \alpha'_{n+1})$ defines the best-fitting polynomial of order $n + 1$, then

$$\alpha'_1 = \alpha_1, \alpha'_2 = \alpha_2, \ldots, \alpha'_n = \alpha_n \tag{2.75}$$

Thus by considering the value of the additional coefficient α'_{n+1}, it may be possible to explain the consequence of increasing the degree of best-fitting polynomial by one (Forfar *et al*, 1988).

The range of values μ_x and q_x can take is relevant to the choice of function type, that is *GM* or *LGM*. The possible range of values for μ_x is between 0 and infinity, making the *GM* class of formulae the more appropriate. On the other hand, q_x can take values between 0 and 1, so that the *LGM* formula is the more suitable for q_x.

In constructing the standard mortality tables based on the 1979 to 1982 experiences and the current standard mortality tables based on the 1991 to 1994 mortality experiences, the CMI Committee graduated the force of mortality μ_x, using central exposures, with maximum likelihood estimation of the parameters. To estimate the unknown parameters $\{\alpha_i\}$ and $\{\beta_j\}$, the actual (observed) number of deaths at age x, a_x, were modelled as independent realisations of Poisson random variables A_x, with mean and variance equal to $R_x^c \mu_x$.

Expressed in terms of the Chebycheff polynomials, the graduation formulae for μ_x are of the form:

$$\mu_x = GM_x(r,s) = \sum_{i=0}^{r-1} \alpha_i C_i(x') + \exp\left(\sum_{j=0}^{s-1} \beta_j C_j(x')\right) \tag{2.76}$$

which is equivalent to

$$\mu_x = GM_x(r,s) = \sum_{i=0}^{r-1} \alpha_i C_i\left(\frac{x-u}{v}\right) + \exp\left[\sum_{j=0}^{s-1} \beta_j C_j\left(\frac{x-u}{v}\right)\right]. \tag{2.77}$$

Thus, for example the $GM_x(0,2)$ and the $GM_x(1,3)$ formulae applied were:

$$\mu_x = GM(0,2) = \exp(\beta_0 + \beta_1 x') \tag{2.78}$$

and

$$\mu_x = GM(1,3) = \alpha_0 + \exp\{\beta_0 + \beta_1 x' + \beta_2 (2x'^2 - 1)\}, \tag{2.79}$$

with $x' = \dfrac{x-70}{50}$.

The graduation process involves choosing the lowest values of r and s, which produce a satisfactory fit for the experience.

In graduating the 1991-94 immediate annuitants' and life office pensioners' mortality experiences, the $GM_x(1,3)$ and the $GM_x(2,3)$ formulae were found to be the most satisfactory, (CMIR 16 and CMIR 17).

2.5 Graduation and generalized linear models

Based on the CMI graduation methodology detailed by Forfar *et al* (1988), Renshaw (1991) demonstrated that the models used for graduation could be formulated within the framework of generalized linear and non-linear models, (GLM and GNLM). He used the attendant statistical package GLIM to implement the graduations and advocated scrutiny of residual plots as an additional diagnostic check on any adopted graduations.

Renshaw *et al* (1996) extended the GLM framework proposed by Renshaw (1991) to include calendar time. The authors modelled the United Kingdom male assured lives' mortality trends over the calendar year period 1958 to 1990, by individual age x and individual calendar year t. Renshaw and Hatzopoulos (1996), experimented with amounts-based data and showed how the data could be graduated within the same framework of generalized linear models, also with respect to age and time. Haberman and Renshaw (1996) give a complete overview of the applications of generalized linear models to graduation and other actuarial problems.

Verrall (1996) presents graduation theory within the framework of statistical models in the computer software package S-PLUS. The framework of generalized linear models is extended to generalized additive models (GAM) to include non-parametric smoothing. Hence both parametric and non-parametric graduations are performed and compared within the same framework and using the same statistical package.

In this thesis, the modelling structure proposed by Renshaw *et al* (1996) is used to investigate annuitants' and pensioners' mortality experiences within the framework of statistical models in S-PLUS. The modelling structure has the distinct advantage that the models adopted can be used directly to project mortality rates into the future.

A brief introduction to generalized linear models is presented in Section 2.5.1. McCullagh and Nelder (1983, 1989) provide a thorough treatment of the subject while Dobson (1990) gives a shorter introduction.

2.5.1 Generalized linear models

Generalized linear models are an extension of classical linear models. The term generalized linear model is due to Nelder and Wedderburn (1972) who demonstrated how many statistical methods involving linear combinations of parameters could be unified. Many other papers on generalized linear models have since been written and international conferences on generalized linear models held (e.g. Gilchrist, 1982).

A generalized linear model is characterised by independent response variables, $\underline{Y} = \{Y_1, \cdots, Y_i, \cdots Y_n\}$ each of which is assumed to have the same distribution in the *exponential family*, taking the form:

$$f_Y(y;\theta,\phi) = \exp\left(\frac{y\theta - b(\theta)}{a(\phi)} + c(y,\phi)\right) \tag{2.80}$$

for some specific functions $a(\cdot)$, $b(\cdot)$ and $c(\cdot)$. If ϕ is known, the distribution of Y is a one-parameter exponential family with *canonical parameter* θ. An unknown ϕ may be regarded as a nuisance parameter and treated as though it were known.

The log-likelihood function considered as a function of θ and ϕ, with y given is:

$$\ell(\theta,\phi;y) = \log f_Y(y;\theta,\phi) = \frac{\{y\theta - b(\theta)\}}{a(\phi)} + c(y,\phi). \tag{2.81}$$

The mean and variance of Y can be derived from the log-likelihood function and the relations

$$E\left(\frac{\partial \ell}{\partial \theta}\right) = 0 \tag{2.82}$$

and

$$E\left(\frac{\partial^2 \ell}{\partial \theta^2}\right) + E\left(\frac{\partial \ell}{\partial \theta}\right)^2 = 0. \tag{2.83}$$

Denoting $m = E[Y]$, it is found that

$$m = b'(\theta), \tag{2.84}$$

and

$$\mathrm{Var}(Y) = a(\phi)b''(\theta). \tag{2.85}$$

The full derivation of the expectation and variance of Y is given in Appendix 1.

From (2.85), it is observed that the variance of Y is a product of two functions, $a(\phi)$, a function independent of θ, and $b''(\theta)$, a function of θ. The quantity $b''(\theta)$, which depends on the canonical parameter θ and hence on the mean m, is called the *variance function* and may be written as $V(m)$. The function $a(\phi)$ is usually of the form

$$a(\phi) = \phi/w \tag{2.86}$$

where ϕ, called the *dispersion parameter* or the *scale parameter*, is constant over observations, and w is a known prior weight that varies from observation to observation.

Well known distributions such as the Poisson, Normal, binomial and gamma distributions all belong to the exponential family. As an illustration, a Poisson random variable with parameter λ, has the likelihood:

$$f(y\,;\lambda) = \exp[y \log \lambda - \lambda - \log y!] \tag{2.87}$$

so that,

$$\theta = \log \lambda$$

$$b(\theta) = \lambda = \exp \theta$$

$$m = b'(\theta) = \exp \theta$$

$$V(m) = b''(\theta) = \exp \theta = m$$

$$\phi = 1$$

$$c(y;\phi) = -\log y!$$

Similarly, a binomial random variable with parameters n and p has the likelihood:

$$f(y;p) = \exp\left[y \log\left(\frac{p}{1-p}\right) + n \log(1-p) + \log\left(\frac{n!}{y!(n-y)!}\right)\right] \tag{2.88}$$

giving the following characteristics:

$$\theta = \log\left(\frac{p}{1-p}\right)$$

$$b(\theta) = \log\left[\left(1+e^{\theta}\right)^{n}\right]$$

$$m = b'(\theta) = \frac{ne^{\theta}}{1+e^{\theta}}$$

$$V(m) = b''(\theta) = \frac{ne^{\theta}}{\left(1+e^{\theta}\right)^{2}} = m\left(1-\frac{m}{n}\right)$$

$$\phi = 1\,.$$

In estimations involving the binomial distribution, we are often interested in the proportion having a given characteristic and hence we would consider the distribution of the random variable defined by Y/n.

The explanatory variables $\boldsymbol{x}_i$ influence the distribution of Y_i through a linear predictor

$$\eta_i = \sum_{j=1}^{p} x_{ij}\beta_j\,. \tag{2.89}$$

The linear predictor η_i is a function of m_i, the expected value of Y_i, and x_{ij} is the value of the jth covariate for Y_i. Thus

$$m_i = \mathrm{E}[Y_i], \tag{2.90}$$

and

$$\eta_i = g(m_i). \tag{2.91}$$

In matrix notation,

$$\eta_i = \underline{x}_i^T \boldsymbol{\beta} \tag{2.92}$$

where $\underline{x}_i$ is a $p \times 1$ vector of explanatory variables and $\boldsymbol{\beta}$ is a $p \times 1$ vector of unknown parameters, which have to be estimated from the data.

The function $g(\cdot)$, called the *link function*, is both monotonic and differentiable so that its inverse, $g^{-1}(\cdot)$ exists.

A generalized linear model is therefore specified by the following three components:

- The *random component*: independent response variables $Y = \{Y_1, \cdots, Y_i, \cdots, Y_n\}$ with the same distribution typically from the exponential family, with $E[Y] = m$.
- The *systematic component*: known covariates $\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_p$ which influence $\underline{Y}$ through a linear predictor $\boldsymbol{\eta} = \boldsymbol{x\beta}$, where $\boldsymbol{x}$ is a matrix of dimension $n \times p$.
- A monotone *link function* $g(\cdot)$, which specifies the relationship $g(m) = \eta$, between the random and systematic components.

The special link function where

$$\eta = g(m) = \theta$$

is called the canonical link function. From equations (2.87) and (2.88), it can be seen that the canonical link function for the Poisson distribution is the log-link, that is

$$\eta_i = \log(m_i) \text{ or } m_i = e^{\eta_i}; \tag{2.93}$$

while the canonical link for the binomial distribution is the log-odds or logit-link defined by:

$$\eta_i = \log\left(\frac{m_i}{1 - m_i}\right) \text{ or } m_i = \frac{e^{\eta_i}}{1 + e^{\eta_i}}. \tag{2.94}$$

Model fitting

The unknown parameters are estimated by maximising the quasi-log-likelihood function defined by Wedderburn (1974) as:

$$q(\underline{y};m)=\sum_i q_i(y_i;m_i)=\sum_i w_i \int_{y_i}^{m_i} \frac{y_i - s}{\phi V(s)} ds. \tag{2.95}$$

where m_i is the mean of the ith response Y_i, with the Y_i's independent, and $\phi V(m_i)/w_i$ is its variance. A comprehensive treatment of quasi-likelihood functions is given in Chapter 9 of McCullagh and Nelder (1989).

Wedderburn (1974) and McCullagh (1983) show that quasi-likelihood functions have properties analogous to those of likelihoods. In particular, the maximum quasi-likelihood estimator of β is asymptotically normal with mean β and the asymptotic covariances may be derived from the second derivative matrix of q. For members of the exponential family of distributions, quasi-log-likelihood estimators are identical to maximum likelihood estimators.

The quasi-log-likelihood function for a Poisson parameter m based on data y is given by:

$$q(y;m)= y\log m - m. \tag{2.96}$$

For the complete data $\underline{y}$, the quasi-log-likelihood is:

$$q(\underline{y};m)=\sum_i q_i(y_i;m_i)=\sum_i (y_i \log m_i - m_i), \tag{2.97}$$

which is equivalent to the log-likelihood.

The parameter estimates are obtained by differentiating (2.95) with respect to the unknown β_j's and solving the resulting system of linear equations:

$$\sum_i w_i \frac{y_i - m_i}{\phi V(m_i)} \frac{\partial m_i}{\partial \beta_j} = 0 \text{ for all } j. \tag{2.98}$$

Since the system of equations is non-linear in β, the maximum quasi-log-likelihood estimates are obtained by a numerical algorithm (see for example McCullagh and Nelder, 1989, Chambers and Hastie, 1993 and Venables and Ripley, 1997).

A special feature of quasi-likelihood functions is that they are specified entirely by the mean and variance functions rather than by a specific distribution or likelihood. Hence quasi-likelihood enables a fitting procedure to be defined for distributions with a given variance function, but without belonging to the class of distributions required for a generalized linear model proper.

Firth (1987) investigates the efficiency of quasi-likelihood estimation for models with constant variance, models with constant coefficient of variation and models with over-dispersion relative to some exponential family and concludes that "*quasi-likelihood estimation retains fairly high efficiency under 'moderate' departures from the corresponding natural exponential family*". Nelder and Pregibon (1987) propose an extended quasi-likelihood function to allow for comparison of variance functions as well as comparisons of linear predictors and link functions.

Measuring goodness of fit

Denote the resulting values of the parameter estimators, linear predictor and fitted values by $\hat{\beta}_j$, $\hat{\eta}_i$, and $\hat{m}_i$ respectively, where:

$$\hat{m}_i = g^{-1}(\hat{\eta}_i)$$

and

$$\hat{\eta}_i = \sum_{j=1}^{p} x_{ij}\hat{\beta}_j \ .$$

Given n observations, the *full* or *saturated model f*, has n parameters, one per observation, so that the fitted values match the observed data exactly, i.e. $\hat{m}_i = y_i$ for all i. The goodness-of-fit criterion or measure of discrepancy of the fit, can be written as:

$$-2\{q(\underline{y};\hat{m})-q(\underline{y};\underline{y})\}=\sum_{i=1}^{n}2w_i\int_{\hat{m}_i}^{y_i}\frac{y_i-s}{\phi V(s)}ds=-2q(\underline{y};\hat{m})=\frac{D(\underline{y};\hat{m})}{\phi} \qquad (2.99)$$

where $D(\underline{y};\hat{m})$ is known as the *deviance function* for the *current model* denoted *c*, and $D(\underline{y};\hat{m})/\phi$, which is minus twice the quasi-likelihood based on the current model, is known as the *scaled deviance*. The current model is defined to be the specific generalized linear model under investigation at any one time. The fitted values of the current and saturated model impact on the formula (2.99) through the lower and upper limits respectively.

McCullagh and Nelder (1989) list forms of deviances for some common distributions including the binomial and Poisson distributions. For the Poisson distribution the deviance is:

$$2\sum_{i=1}^{n}\{y_i\log(y_i/\hat{m}_i)-(y_i-\hat{m}_i)\}. \qquad (2.100)$$

Writing $D(c,f)$ for $D(\underline{y};\underline{\hat{m}})$, then for a Gaussian modelling distribution with identity link, the scaled deviance has distribution:

$$S(c,f)=\frac{D(c,f)}{\phi}\sim\chi^2_{n-p}, \qquad (2.101)$$

where p is the number of parameters involved in the current model *c*. Hence an unbiased estimator of the scale parameter ϕ is:

$$\hat{\phi}=\frac{D(c,f)}{n-p}. \qquad (2.102)$$

Under other modelling distributions, the distribution for the scaled deviance (2.101) is approximate.

The differences in the values of scaled deviances in hierachical models can be used as a generalized measure of discrepancy for fixed modelling distributions with fixed link. Given model structures c_1 and c_2 with c_2 nested in c_1, and respective scaled deviances $S(c_1, f)$ and $S(c_2, f)$, the difference $S(c_2, f) - S(c_1, f)$ is asymptotically distributed as chi-square with $\nu_2 - \nu_1$ degrees of freedom, where ν_1 and ν_2 denote the degrees of freedom for c_1 and c_2 respectively. The distribution is exact only in the Gaussian family with identity link.

Further diagnostic checks on the adequacy of the fitted model can be made using residuals. Two types of residuals commonly used are given below.

i) The *deviance residuals* are defined by:

$$sign(y_i - \hat{m}_i)\sqrt{d_i} \tag{2.103}$$

where

$$d_i = 2w_i \int_{m_i}^{y_i} \frac{y_i - s}{V(s)} ds ,$$

is the ith component of the deviance $D(c, f)$ evaluated at the fitted value $\hat{m}_i$, i.e. the deviance residuals are the signed values of the square roots of the individual components contributing to the value of the deviance. Thus for example, the formula for the deviance residuals for the Poisson distribution is:

$$sign(y_i - \hat{m}_i)\sqrt{2\{y_i \log(y_i / \hat{m}_i) - (y_i - \hat{m}_i)\}} . \tag{2.104}$$

ii) The *Pearson residuals* are defined by:

$$\frac{y_i - \hat{m}_i}{\sqrt{V(\hat{m}_i)/w_i}} \tag{2.105}$$

that is, the signed values of the square roots of the individual components contributing to the value of the Pearson χ^2 goodness-of-fit statistic.

Dividing both types of residuals by the estimated value of $\sqrt{\phi}$ gives rise to *studentized residuals* of the particular type. Plots of residuals against some function of the fitted values provide informal visual checks on the various modelling assumptions.

The two sets of residuals are identical for the normal modelling distribution for which $V(m_i) = 1$ and $w_i = 1$ for all i. McCullagh and Nelder (1989) note that the deviance residual is generally preferred to the Pearson residual for model checking because it has distributional properties that are closer to the residuals arising in linear regression models. Pierce and Schafer (1986) provide an extensive examination of residuals in exponential family models.

2.5.2 Graduation with respect to age using GLMs

It is possible to reformulate and extend the graduation methodology used by the CMI by using generalized linear and non-linear models as described by Renshaw (1991) and Haberman and Renshaw (1996). Most of what follows in this section is a reproduction of the description of graduation with respect to age given by Haberman and Renshaw (1996).

As described in Section 2.4.1, the actual number of deaths A_x are modelled as Poisson random variables when targetting μ_x and as binomial random variables when targetting q_x. Hence for μ_x- graduations with responses $\{A_x\}$,

$$m_x = E(A_x) = R_x^c \mu_{x+\frac{1}{2}}, \qquad (2.106)$$

$$V(m_x) = m_x,$$

$$w_x = 1,$$

$$\phi = 1.$$

Equivalently, when graduating μ_x with responses A_x / R_x^c,

$$m_x = E\left[A_x / R_x^c\right] = \mu_{x+\frac{1}{2}}, \tag{2.107}$$

$$V(m_x) = m_x,$$

$$w_x = R_x^c,$$

$$\phi = 1.$$

For q_x-graduations with responses $\{A_x\}$,

$$m_x = E\left[A_x\right] = R_x q_x, \tag{2.108}$$

$$V(m_x) = m_x(1 - m_x / R_x),$$

$$w_x = 1,$$

$$\phi = 1.$$

The graduation formulae are presented as predictor-link relationships with age x as the sole covariate. Thus for μ_x-graduations, the focus is on either the log-link with responses $\{A_x\}$, so that from (2.106) we have,

$$\log m_x = \eta_x = \log R_x^c + \log \mu_{x+\frac{1}{2}} \tag{2.109}$$

with inverse

$$\mu_{x+\frac{1}{2}} = \exp\left(\eta_x - \log R_x^c\right), \tag{2.110}$$

or the parameterized power link with responses A_x / R_x^c, so that

$$\mu_{x+\frac{1}{2}}^{\gamma} = \eta_x \tag{2.111}$$

with inverse

$$\mu_{x+\frac{1}{2}} = \eta_x^{1/\gamma}. \tag{2.112}$$

The term $\log\left(R_x^c\right)$ in expression (2.109), which does not involve any unknown parameters, is known as the *offset*. An offset is a fixed known term in the linear predictor, which does not contain a parameter to be estimated. In (2.109), the offset $\log\left(R_x^c\right)$ automatically allows for the exposed-to-risk term in the log-likelihood.

For q_x-graduations with $\{A_x\}$ as responses, the focus is on the log-odds-link, so that

$$\log\left(\frac{m_x}{R_x^c - m_x}\right) = \log\left(\frac{q_x}{1-q_x}\right) = \eta_x \tag{2.113}$$

with inverse

$$q_x = \frac{e^{\eta_x}}{1+e^{\eta_x}} \tag{2.114}$$

or the complementary log-log-link,

$$\log\left\{-\log\left(1-\frac{m_x}{R_x}\right)\right\} = \log\{-\log(1-q_x)\} = \eta_x \tag{2.115}$$

with inverse

$$q_x = 1 - \exp\left(-e^{\eta_x}\right). \tag{2.116}$$

Also, for q_x-graduations, the parameterized family of link functions

$$\eta_x = \log\left\{\frac{(1-q_x)^{-\gamma} - 1}{\gamma}\right\} \tag{2.117}$$

with inverse

$$q_x = 1 - \left(1 + \gamma e^{\eta_x}\right)^{-1/\gamma} \tag{2.118}$$

reduces to the logit link when $\gamma = 1$ and the complementary log-log-link as $\gamma \to 0$.

The graduation formula

$$\mu_x = GM_x(r,s) = \sum_{i=0}^{r-1} \alpha_i C_i\left(\frac{x-u}{v}\right) + \exp\left[\sum_{j=0}^{s-1} \beta_j C_j\left(\frac{x-u}{v}\right)\right] \tag{2.119}$$

(equation 2.77) comprises the identity link ($\gamma = 1$) and the non-linear Gompertz-Makeham predictor $GM_x(r,s)$. When $r = 0$, the formula reduces to the log-link in combination with a polynomial predictor.

For q_x-graduations, A. D. Wilkie suggested the use of the log-odds link in combination with polynomial predictors, that is:

$$\log\left(\frac{q_x}{1-q_x}\right) = \sum_{j=0}^{r} \beta_j h_j\left(\frac{x-u}{v}\right) \tag{2.120}$$

with u and v suitably chosen and an orthogonal basis h_j, where h_j denote either Chebycheff polynomials of the first type or Legendre polynomials as defined in Section 2.4.2. The graduation formula (2.120) was used by the CMI Bureau to construct all q_x-graduations published by the Continuous Mortality Investigation Committee (1976).

The $LGM_x(r,s)$ formula:

$$LGM_x(r,s) = \frac{GM_x(r,s)}{1+GM_x(r,s)},$$

suggested by Forfar *et al* (1988), comprises the odds-link in combination with the Gompertz-Makeham predictor $\eta_x = GM_x(r,s)$. When $r = 0$, the formula reduces to the logit-link in combination with the polynomial predictor, and when $s = 0$ or 1 (with r > 0), the predictor is linear in combination with the odds-link; otherwise when $r > 0$ and $s > 1$, the predictor is non-linear.

Renshaw (1991) suggested the graduation formula:

$$q_x = 1 - \exp\left[-\exp\left\{\sum_{j=0}^{r} \beta_j h_j\left(\frac{x-u}{v}\right)\right\}\right] \tag{2.121}$$

comprising the complementary log-log-link in combination with a polynomial predictor and which includes the Gompertz formula as a special case when $r = 1$.

2.5.3 The distribution of deaths in the presence of duplicates

The effect of duplicate policies on the variance of the distribution of the number of deaths in a mortality experience based on policies has been extensively discussed in papers by, among others, Seal (1945), Beard and Perks (1949), Daw (1951), and more recently by Forfar *et al* (1988) and Renshaw (1992).

Beard and Perks (1949) showed that for a sample of N_x independent lives selected at random at age x, the variance of the number of claims A_x is given by

$$\operatorname{var}(A_x) = N_x q_x \sum_i i\pi_x^{(i)} \left\{ \frac{\sum_i i^2 \pi_x^{(i)}}{\sum_i i\pi_x^{(i)}} - q_x \sum_i i\pi_x^{(i)} \right\} \tag{2.122}$$

where $\pi_x^{(i)}$ is the probability that a policyholder aged x, holds i policies; $i = 1, 2, 3, \ldots$ and

$$\sum_i \pi_x^{(i)} = 1.$$

Expression (2.122) may be written as

$$\operatorname{var}(A_x) = \phi_x R_x q_x (1 - q_x) \tag{2.123}$$

with

$$\phi_x = \left\{ \frac{\sum_i i^2 \pi_x^{(i)}}{\sum_i i \pi_x^{(i)}} - q_x \sum_i i \pi_x^{(i)} \right\} (1 - q_x)^{-1} \tag{2.124}$$

and

$$R_x = N_x \sum_i i \pi_x^{(i)} \tag{2.125}$$

is the initial exposed-to-risk based on policies.

When there are no duplicates present, $\pi_x^{(1)} = 1$, $\pi_x^{(i)} = 0$ otherwise; so that $\phi_x = 1$ and A_x has a binomial distribution with parameters R_x and q_x as in Section 2.4.1.

Since q_x is small, equation (2.124) may be approximated by

$$\phi_x \approx \frac{\sum_i i^2 \pi_x^{(i)}}{\sum_i i \pi_x^{(i)}} > 1. \tag{2.126}$$

Use of the approximate form (2.126) is important in that in an empirical study of the distribution of duplicate policies, the parameter ϕ_x can be estimated without reference to q_x.

In an empirical study, the *variance ratios* ϕ_x are estimated by

$$r_x = \frac{\sum_i i^2 f_x^{(i)}}{\sum_i i f_x^{(i)}} \tag{2.127}$$

where $f_x^{(i)}$ is defined to be the proportion of policyholders in the study group aged x, who have i policies.

An equivalent formula for the variance of the number of claims A_x, when modelling μ_x in the presence of duplicates is:

$$\mathrm{var}(A_x) = \phi_x R_x^c \mu_{x+\frac{1}{2}} \tag{2.128}$$

with ϕ_x estimated as in (2.126) and (2.127) as before (Forfar *et al*, 1988).

Daw (1951) considered the practical aspects of the treatment of duplicates in a mortality study and suggested three courses of action that might be adopted. These are:

i) investigating the distribution of duplicates ;

ii) arranging the mortality data such that several independent estimates of each rate of mortality are obtained from which the variance can be estimated; and

iii) excluding duplicate policies and basing the investigation on lives.

In his discussion of the three suggested courses of action, Daw (1951) felt that the preferred method would be that of excluding duplicates since duplicates provide no further information about the mortality but simply have the effect of increasing the variance, provided mortality is independent of the number of policies held.

The CMI Committee carried out investigations into the distribution of duplicate policies held by assured lives in 1954, and in 1981 and 1982, (CMI Committee, 1957, 1986 and Forfar *et al*, 1988). The Committee observed that the number of lives with only one policy were highest at the youngest ages and tended to decrease as the lives became older. Consequently, the CMI Committee felt that any allowance for duplicates should be made on an age basis.

Forfar *et al* (1988) proposed a method based on the third course of action suggested by Daw (1951), whereby the number of deaths and the exposed-to-risk are each first divided by the estimated variance ratio r_x before modelling. The method relies on prior knowledge of the variance ratios and the authors proposed an investigation of the number of policies on each life among the deaths, to determine estimates of the variance ratios.

Forfar *et al* (1988) also demonstrated that although altering the variance, the existence of duplicates does not change the expected value of the number of deaths, so that

$E(A_x) = R_x q_x$, when modelling q_x;

and

$E(A_x) = R_x^c \mu_{x+\frac{1}{2}}$, when modelling μ_x.

Renshaw (1992) discusses the feasibility of using two-stage generalized linear (or non-linear) models to graduate mortality rates with allowance for over-dispersion attributed to duplicate policies. The first stage involves modelling the responses using generalized linear or non-linear models as appropriate; and the second stage involves modelling the dispersion parameters ϕ_x as secondary interrelating generalized linear models along the lines first proposed by Pregibon (1984) in his review of the 1st edition of the monograph by McCullagh and Nelder (1983), and subsequently developed by Nelder and Pregibon (1987). The method implies the joint modelling of the mean and the dispersion of the response variable, a subject covered in detail in Chapter 10 of McCullagh and Nelder (1989). The proposed model structure is as follows:

1. Model the A_x's as independent response variables with:

   $$\mathrm{E}(A_x) = m_x,\ \mathrm{var}(A_x) = \phi_x V(m_x), \qquad (2.129)$$

   and predictor-link $\eta_x = g(m_x) = \sum_j u_{xj}\beta_j$.

2. Model the unknown dispersion parameters ϕ_x using a dispersion parameter d_x with:

   $$\mathrm{E}(d_x) = \phi_x,\ \mathrm{var}(d_x) = \tau V_D(\phi_x), \qquad (2.130)$$

and predictor link $\xi_x = h(\phi_x) = \sum_j v_{xj}\delta_j$,

where $V_D(\cdot)$ is the variance function for the second stage generalized linear model and τ is a scale factor.

The u_{xj}'s and v_{xj}'s are known covariates, while the parameters β_j and δ_j need to be estimated. The optimisation procedure depends on the specific form of d_x.

Renshaw (1992) re-emphasises the point made by Cox (1983), that the modelling of excess variation in this context has little effect on the estimation of the parameters of primary concern (the β_j's), but that statistical tests and confidence intervals may be seriously in error unless the effect of the excess variation is taken into account.

2.5.4 Graduation with respect to age and time

Renshaw *et al* (1996) used generalized linear models to model the force of mortality μ_{xt}, at age x, in calendar year t, for the specific duration d = 5 years and over (5+), to give a trend analysis of United Kingdom male assured lives' mortality over the calendar year period 1958 to 1990, for ages 22 to 89 inclusive.

The actual number of deaths a_{xt}, are modelled as independent realisations of over-dispersed Poisson random variables A_{xt}, with mean and variance given by:

$$E(A_{xt}) = m_{xt} = R^c_{xt}\mu_{xt} \quad (2.131)$$

$$\text{Var}(A_{xt}) = \phi m_{xt} \quad (2.132)$$

The scale parameter ϕ is included to take account of the fact that the data are based on policy numbers rather than counts of lives, and as such there may be duplicate policies issued on the same lives, resulting in over-dispersion of the Poisson random variable. In the particular study by Renshaw *et al* (1996), it was felt that the effect of modelling the scale parameter ϕ as a function of age would be negligible and hence ϕ was modelled as a constant over all ages.

The function μ_{xt} is modelled using graduation formulae involving the log-link in combination with bivariate polynomial predictors with age and calendar year as the covariates, for the specific duration under consideration. The polynomial based-formulae are of the form:

$$\mu_{xt} = \exp\left\{\sum_{j=0}^{s}\beta_j L_j(x') + \sum_{i=1}^{r}\alpha_i t'^i\right\} \tag{2.133}$$

and

$$\mu_{xt} = \exp\left\{\sum_{j=0}^{s}\beta_j L_j(x') + \sum_{i=1}^{r}\alpha_i t'^i + \sum_{i=1}^{r}\sum_{j=1}^{s}\gamma_{ij} L_j(x')t'^i\right\} \tag{2.134}$$

subject to the convention that some of the γ_{ij} parameters may be pre-set to zero. x' and t' are transformed ages and calendar-years such that both the age and calendar-year ranges are mapped onto the interval [-1, 1], thus:

$$x' = (x - c_x) / w_x ; \; t' = (t - c_t) / w_t ;$$

c_x and c_t denote mid-points of the age and calendar-year ranges respectively, while w_x and w_t denote semi-ranges, that is

$$c_x = (x_{\min} + x_{\max})/2; \; w_x = (x_{\max} - x_{\min})/2 ;$$

with equivalent expressions for c_t and w_t. $L_j(x')$ denotes Legendre polynomials of degree j.

Equivalently, expressions (2.133) and (2.134) may be written as:

$$\log(\mu_{xt}) = \sum_{j=0}^{s}\beta_j L_j(x') + \sum_{i=1}^{r}\alpha_i t'^i , \tag{2.135}$$

and

$$\log(\mu_{xt}) = \sum_{j=0}^{s} \beta_j L_j(x') + \sum_{i=1}^{r} \alpha_i t'^i + \sum_{i=1}^{r}\sum_{j=1}^{s} \gamma_{ij} L_j(x') t'^i . \tag{2.136}$$

The predictor-link relationship is of the type

$$\log(m_{xt}) = \eta_{xt} = \log(R_{xt}^c) + \log(\mu_{xt}). \tag{2.137}$$

Hence,

$$\eta_{xt} = \log(m_{xt}) = \log(R_{xt}^c) + \beta_0 + \sum_{j=1}^{s} \beta_j L_j(x') + \sum_{i=1}^{r} \alpha_i t'^i \tag{2.138}$$

and

$$\eta_{xt} = \log(m_{xt}) = \log(R_{xt}^c) + \sum_{j=0}^{s} \beta_j L_j(x') + \sum_{i=1}^{r} \alpha_i t'^i + \sum_{i=1}^{r}\sum_{j=1}^{s} \gamma_{ij} L_j(x') t'^i . \tag{2.139}$$

The term, $\log(R_{xt}^c)$ offsets the general mean β_0, by a known amount, conditional on the values of age x and calendar time t.

The unknown parameters are estimated using the quasi-log-likelihood approach which involves maximising the expression:

$$\frac{1}{\phi}\sum_{x,t}\left(-m_{xt} + a_{xt}\log m_{xt}\right), \tag{2.140}$$

or, equivalently, minimising

$$\frac{1}{\phi}\sum_{x,t}\left(m_{xt} - a_{xt}\log m_{xt}\right). \tag{2.141}$$

Firstly, formulae of the form

$$\log \mu_{xt} = \beta_0 + \sum_{j=1}^{s}\{\beta_j L_j(x')\} + \sum_{i=1}^{r} \alpha_i t'^i ,$$

are fitted. As more terms are added to the structure of the linear predictor, the differences in scaled deviances are approximately distributed as χ^2, with degrees of freedom determined by the number of additional parameters added. Therefore, a test of whether the total improvement in the model as a result of the additional parameters is significant, can be carried out, and optimum values for r and s thereby determined.

The (unscaled) deviance for the current model is:

$$D(c,f) = 2\sum_{xt}\left\{a_{xt}\log\left(\frac{a_{xt}}{\hat{m}_{xt}}\right) - (a_{xt} - \hat{m}_{xt})\right\}, \tag{2.142}$$

where $\hat{m}_{xt} = R_{xt}^{c}\hat{\mu}_{xt}$, that is, $\hat{m}_{xt}$ is the number of deaths predicted by the model.

The scale parameter ϕ, can either be estimated from the unscaled deviance goodness-of-fit statistic, as

$$\hat{\phi} = \frac{D(c,f)}{v}, \tag{2.143}$$

or, from the Pearson goodness-of-fit statistic, as

$$\hat{\phi} = \frac{1}{v}\sum_{xt}\frac{(a_{xt} - \hat{m}_{xt})^2}{\hat{m}_{xt}}, \tag{2.144}$$

where ν is the number of degrees of freedom.

Starting with the pre-determined values of r and s, mixed polynomial terms in age and calendar-year effects are then introduced to the model. Thus, the model is now extended to:

$$\log(\mu_{xt}) = \sum_{j=0}^{s}\beta_j L_j(x') + \sum_{i=1}^{r}\alpha_i t'^i + \sum_{i=1}^{r}\sum_{j=1}^{s}\gamma_{ij} L_j(x')t'^i,$$

with some of the γ_{ij} parameters equal to 0.

It should be noted that although only polynomial effects of calendar time t and age x are allowed for in the models fitted, fractional polynomials could be used to improve the fit. In addition, semi-parametric procedures (such as Generalised Additive models) could be used to explore the functional link between mortality rates and (x, t).

Tests of Graduation

Noting that graduation can be regarded as a smoothing procedure of observed mortality rates, a satisfactory graduation should satisfy the following two criteria:

- **Overall goodness-of-fit or fidelity to data**: the graduated rates should be close enough to the crude mortality rates for them to be representative of the underlying mortality rates.
- **Smoothness**: Bizley (1958) defined smoothness as follows: '*a plane continuous curve is smooth at those points which are such that the absolute value of the rate of change of curvature with respect to distance measured along the curve is small*'. In the context of mortality, the requirement of a small rate of change is equivalent to a requirement that third order differences are small, which is consistent with the widely held view that low order polynomials are smooth.

Benjamin and Pollard (1993) observe that "*the two qualities 'smoothness' and 'goodness-of-fit' tend to conflict, in the sense that smoothness may not be improved beyond a certain point without some sacrifice of goodness-of-fit, and vice versa. Thus a graduation will often turn out to be a compromise between optimal fit and optimal smoothness*". Therefore the graduated rates should not follow the crude mortality rates too strictly at the expense of smoothness (*under-graduation*) or too loosely at the expense of goodness-of-fit (*over-graduation*).

As an initial check on the goodness-of-fit of the model, informal techniques involving a visual inspection of plots of the residuals are recommended (Renshaw, 1991, McCullagh and Nelder, 1989). A good model is taken as one that, among other things, "*leaves a patternless set of residuals*" (McCullagh and Nelder, 1989).

In graduating using GLMs, residuals can be used to explore the adequacy of the model with respect to choice of variance function, link function and terms in the linear predictor. They may also indicate the presence of anomalies requiring further investigation. Therefore the overall goodness-of-fit of a given model can be assessed using the residuals. Not surprisingly, the statistical tests that have been devised to test the adequacy of fit of a graduation are essentially applied to residuals.

Renshaw *et al* (1996) used two common types of residuals for model testing in graduating with respect to age and time using generalized linear models:

1) the deviance residuals defined by:

$$sign(a_{xt} - \hat{m}_{xt})\sqrt{2\left\{a_{xt}\log\left(\frac{a_{xt}}{\hat{m}_{xt}}\right) - (a_{xt} - \hat{m}_{xt})\right\}}\,; \tag{2.145}$$

and,

2) the Pearson residuals defined by:

$$\frac{a_{xt} - \hat{m}_{xt}}{\sqrt{\hat{m}_{xt}}}\,. \tag{2.146}$$

McCullagh and Nelder (1989) discuss a third type of residual proposed by Anscombe (1961). Noting that the Pearson residuals have the disadvantage that for non-normal distributions, the distribution is often markedly skewed and so may fail to have properties similar to those of a normal distribution, the residuals proposed by Anscombe are such that the distribution is as normal as possible. For the distribution of deaths (Poisson distribution), the Anscombe residuals take the form:

$$\frac{\frac{3}{2}\left(a_{xt}^{2/3} - \hat{m}_{xt}^{2/3}\right)}{\hat{m}_{xt}^{1/6}}\,. \tag{2.147}$$

However, in line with Renshaw *et al* (1996), in this study, statistical tests of graduation are applied on the studentized Pearson residuals and the studentized deviance residuals[1].

The null hypothesis tested is:
H_0: *the true underlying forces of mortality for the experience are the graduated rates.*

The statistical test of graduation applied to test for overall goodness-of-fit of the model is the **chi-square goodness-of-fit test**. The chi-square goodness-of-fit statistic is defined as:

$$\chi^2 = \sum_{x,t} z_{xt}^2 , \qquad (2.148)$$

where the z_{xt}^2 are the squared studentized Pearson residuals, that is

$$z_{xt} = \frac{a_{xt} - \hat{m}_{xt}}{\sqrt{\hat{m}_{xt}\hat{\phi}}} . \qquad (2.149)$$

Under the hypothesis that the studentized Pearson residuals (equivalent to the *relative deviations* referred to by Forfar *et al*, 1988) have a normal distribution, the statistic χ^2 is expected to be distributed as a chi-square random variable with *n-p* degrees of freedom, where *n* is the total number of ages or age groups and *p* is the number of parameters estimated. The model is rejected for high values of the χ^2 statistic.

The χ^2 distribution provides a good estimation provided the numbers in each cell are not too small. The approximation is generally considered to be unreliable if the number in a cell is small (less than 5 say). Therefore, where necessary, the data are grouped so that the expected number of deaths in each cell is at least 5.

---

[1] Studentized residuals are the residuals divided by the square root of the scale parameter ϕ.

The test has certain serious limitations when applied to mortality data and Benjamin and Pollard (1993) detail these limitations as follows:

a) it does not detect the existence of a number of excessively large deviations counterbalanced by a large number of small deviations;
b) it does not detect a large cumulative deviation over part or the whole of the age range;
c) it does not detect an excess of positive or negative deviations over part or the whole of the age range;
d) it does not detect excessive clumping of deviations of the same sign.

Because of these limitations, Benjamin and Pollard (1993) conclude that although the chi-square test provides a useful comparison in the form of a single statistic, further tests are necessary, and the authors further observe that some of the other tests may be of greater importance than the chi-square test. The tests applied in this study are described below.

- **Distribution of individual standardised deviations:-** test for overall goodness-of-fit and for detecting excessively large deviations. The individual standardised deviations test is used to test whether the individual numbers of deaths conform to the probability distribution assumed. Under the null hypothesis, the individual standardised deviations, that is the z_{xt}'s defined above, are expected to be distributed as $N(0,1)$. Most of the problems that might be present in a graduation can be detected by an examination of the distribution of the standardised deviations.

- **Signs of deviations test:-** test for overall bias in the graduation, i.e. whether the graduated rates are too high or too low. Under the null hypothesis, the deviations of the observed deaths from the expected are independent normal random variables. If there is no bias in the graduated rates, the signs of the individual deviations are independent and are equally likely to be positive or negative. Therefore, the distribution of the signs of the deviations would be expected to be binomial with parameters n and 0.5, where n is the number of ages or age groups (the number of data cells). An excessively large number of positive or negative deviations will indicate that the graduated rates are biased.

- **Cumulative deviations test:-** test for overall goodness-of-fit and for detecting large positive or negative relative deviations over a specified range. Under the null hypothesis, the number of deaths a_{xt} have an approximate normal distribution with mean $\hat{m}_{xt}$ and variance $\hat{m}_{xt}\hat{\phi}$. It follows that the total, $\sum_{x,t}(a_{xt} - \hat{m}_{xt})$, summed over a selected range of ages (chosen without reference to the data) in the appropriate calendar year (or years), will have an approximate normal distribution with mean 0 and variance $\sum_{xt}\hat{m}_{xt}\hat{\phi}$, so that the statistic

$$z = \frac{\sum_{xt}(a_{xt} - \hat{m}_{xt})}{\sqrt{\sum_{xt}\hat{m}_{xt}\hat{\phi}}}, \tag{2.150}$$

is expected to have a standardised normal distribution i.e. N(0,1). The null hypothesis is rejected if the absolute value of the test statistic is high.

- **Stevens' grouping of signs test:-** test for *over-graduation*. If the positive and negative standardised deviations are arranged in random order, R_+, the number of runs of positive deviations would be expected to have mean

$$\frac{n_+(n_- + 1)}{n_+ + n_-}, \tag{2.151}$$

and variance

$$\frac{(n_+ n_-)^2}{(n_+ + n_-)^3}, \tag{2.152}$$

where n_+ and n_- are the numbers of positive signs and negative signs respectively. If both n_+ and n_- are not small, the test statistic R_+ can be approximated by a normal distribution with mean and variance given by (2.151) and (2.152) respectively. The exact distribution of R_+ is the hypergeometric distribution.

The rates are over-graduated if there are too few runs. Daw (1954) pointed out that the test can lead to different conclusions depending on whether the test has been applied to positive or negative signs. However, the effect is of little consequence except in the case when the total number of signs is small.

The above summary of statistical tests of graduation is by no means exhaustive. Graduation tests are described in detail in textbooks on mortality such as Benjamin and Pollard (1993). In addition, the paper by Forfar *et al* (1988) includes an extensive discussion of statistical tests of graduation.

Chapter 3

Forecasting Mortality Rates

3.1 Background

The methods of forecasting or projecting mortality rates, and evaluating the forecast rates vary depending on the quality of the data available and the application of the projected rates. Regardless of the method adopted, the projection of age-specific mortality rates for actuarial applications is essentially a process of extrapolation of past mortality trends, as opposed to forecasts based on expected medical advances or the emergence of new diseases.

Keyfitz (1982) provides a review of some general aspects of forecasting mortality. He stresses the importance of curves fitted to mortality data (that is parametric graduation) as a tool for forecasting future mortality rates. His discussion focuses on minimising the parameter representation of the mortality function by aggregating ages. Implicit in the discussion is the trade-off between *parsimony* (using a smaller number of parameters) and the level of age detail to be used (more details require more parameters). He notes that '*the simpler the curve, the more realistic is likely to be its projection into the future*'.

Pollard (1987) discusses various methods of projecting age-specific mortality rates that have been used by actuaries and demographers: projection by extrapolation of mortality rates (or transformations of mortality rates) at selected ages; projection by reference to a 'law of mortality'; projection by reference to model life tables; projection by reference to another 'more advanced population'; projection by reference to an 'optimal' life table attainable under ideal conditions; and projection by cause of death. He assesses the potential strengths and weaknesses of each of the methods and gives examples of the use of these methods. The book by Benjamin and Soliman (1993) also provides a useful summary of methods of projecting age-specific mortality rates.

Tolley *et al* (1993) review some of the methods of forecasting mortality rates in the context of the private life insurance industry operating under U.S. insurance law. The authors observe that actuaries are not '*experts in stochastic models for forecasting changes in mortality*'. They go on to discuss various benefits of mortality forecasts to the insurance industry (and the actuarial profession) particularly in light of the unprecedented mortality improvement in the last century that has resulted in financial losses in the annuity business.

Tuljapurkar and Boe (1998) carry out a critical assessment of existing knowledge about mortality change and forecasting methods, based on forecasts of mortality change in Canada, Mexico and the U.S. The researchers provide a comprehensive discussion of broad mortality patterns in the U.S., Canada and Mexico and discuss analytical methods and theories used to study mortality and mortality change.

Macdonald (1997) documents mortality trends and methods of projecting mortality improvements in future for annuity and pension business adopted in some countries in Western Europe, notably Austria, France, Germany, Italy and the UK.

Renshaw *et al* (1996) develop a modelling structure based on the Gompertz-Makeham models described by Forfar *et al* (1988) to incorporate both age and time. The method is used to analyse United Kingdom ultimate mortality experience from the insured sub-population using generalized linear models.

In a paper presented at a meeting of the Staple Inn Actuarial Society in London, Willets (1999) provides a comprehensive analysis of mortality changes in the UK over the course of the 20th century and makes comparisons with mortality experience from other countries in Europe, Asia, Australia and the Americas. He discusses different methods of mortality projections and proposes an alternative method based on the *birth cohort* (that is the year of birth) for estimating future mortality improvements.

Felipe, Guillen and Nielsen (2001) propose a method for exploring the evolution of mortality rates based on kernel hazard estimation. The methodology, which involves analysing mortality as a two-dimensional multiplicative function of chronological time and age, is used to compare mortality experiences of Denmark and Spain.

In this chapter, various methods of projecting age-specific mortality rates are presented. However, the methods presented here are by no means exhaustive. In Sections 3.2 to 3.5, the methods described by Pollard (1987) are summarised. The Lee-Carter method of projecting mortality using statistical time series methods is described in Section 3.6 and methods used to project mortality for annuity business in Western Europe, including the Renshaw *et al* and the Willets methods, are described in Section 3.7.

3.2 Projection by Extrapolation of Mortality Rates

The simplest method of projecting age-specific mortality rates and the most widely used is that of projection by extrapolation of the mortality rates. The method has as its basis the assumption that the mortality rates are simple functions of time t, for both historic and future values of t. A particularly simple version of this assumption is that the proportinate reduction from one year to another in the age-specific mortality rates is relatively constant over a long period of time. This means that if the mortality rate at age x is plotted for successive years, the curve is close to being a straight line. The linearity may be improved by some transformation of the mortality rates such as the logarithmic transform, hence the method is sometimes referred to as the logarithmic method (see Benjamin and Soliman, 1993).

Extrapolation may be performed either graphically or by mathematical formula. The most commonly used formula for extrapolation is:

$$q_{x,t} = \beta_x \gamma_x^t , \tag{3.1}$$

where $q_{x,t}$ is the mortality rate at age x experienced in year (or time period) t ;

β_x is the level of mortality at age x at a particular point in time, that is, the initial level of mortality;

γ_x, $(0 < \gamma_x < 1)$ is the annual rate of improvement in mortality at age x.

The equivalent formula for the logarithmic transformation of $q_{x,t}$ is

$$\ln q_{x,t} = B_x + tC_x, \qquad (3.2)$$

where $B_x = \ln \beta_x$ and $C_x = \ln \gamma_x$.

The log-linear relationship holds for age-group rates as well as rates for single years (Benjamin and Soliman, 1993).

Formula 3.1 allows the mortality rate at age x to decrease indefinitely towards zero. An alternative formula preferred by those who postulate an ultimate level of mortality at age x, α_x say, is

$$q_{x,t} = \alpha_x + \beta_x \gamma_x^t. \qquad (3.3)$$

The function extrapolated need not be q_x. Other life table functions (such as μ_x) or transformations of life table functions may be used.

Extrapolation by mathematical formula involves estimating the improvement factor γ_x for various recent time periods using log-linear regression on the $q_{x,t}$ for time periods t when (3.1) holds, or using non-linear regression methods when (3.3) holds. Alternatively, when (3.1) is valid, γ_x may be found using the following formula valid between time periods s and t ($t > s$):

$$\gamma_x = \left(q_{x,t} / q_{x,s}\right)^{1/(t-s)} \qquad (3.4)$$

Projected mortality rates are then calculated using formula (3.1) or (3.3) as appropriate.

Under the graphical approach, mortality rates at each selected age x at recent time periods t are plotted against t. A smooth curve is drawn through the points and similar curves are drawn on the same graph for neighbouring ages x. The curves are all then extrapolated to yield projected values of $q_{x,t}$. Projected mortality rates at intervening ages are found by interpolation or by multiplying the base mortality rates (initial mortality rates) at those ages by the projected reduction factor at the nearest selected age.

The Institute and Faculty of Actuaries (UK) annuitant tables of 1924 (Anderson and Dow, 1964) based on life office annuitant mortality between 1900 and 1920 used mortality rates projected by formula (3.3). Subsequent annuitant tables have been derived using similar methodology. The current methodology used by the Institute and Faculty of Actuaries to project annuitant mortality is an extension of (3.3). The method is described in detail in Section 3.7. The Society of Actuaries (1981) used formula (3.1) to derive the '1983 Table *a*' for individual annuity valuation.

The method of extrapolation by mathematical formula has also been used in population projections. For example, two sets of population projections made in Canada in the 1950's by the Dominion Bureau of Statistics (1950, 1954) each projected future mortality rates on the basis of the log transformation given by formula (3.2). The Government Actuary's Department of the United Kingdom (1965) also used formula (3.1) in the projections of the population from 1965 to 2000. Golulapati, De Ravin and Trickett (1984) projected mortality rates of the Australian population from 1981 to 2020 by the same method, among others.

Pollard (1987) comments that the formula chosen for mathematical extrapolation should have few parameters and behave in a simple, appropriate and well-understood fashion as the time period t is varied. He further stresses the importance of the life table function chosen being sensitive to changes in mortality. An example given is that of the expectation of life $\mathring{e}_{x,t}$ which is affected by mortality at all ages above age x but is not particularly sensitive to any of the mortality rates; hence one can project the expectation of life for future time periods t and selected ages x, but mortality rates deduced from the extrapolation are unlikely to be reliable.

3.3 Projection through parameters

3.3.1 Projection by reference to a law of mortality

The curve representing a law of mortality (described in Chapter 2), is fitted to the observed age-specific mortality rates at each of several different time periods (e.g. by the method of maximum likelihood, least squares or minimum chi-square as described in Chapter 2) and the values of the parameters at each time period t are estimated. For example, in the case of Makeham's law (equation 2.43):

$$\mu_{x,t} = \alpha_t + \beta_t c_t^{\,x}; \tag{3.5}$$

the parameters α_t, β_t and c_t are estimated for various time periods t.

Trends in the parameters are extrapolated to provide estimates of the parameters at future time periods t. The extrapolation may be graphical or by mathematical formula. Projected age-specific mortality rates are then obtained by substituting the projected parameters and the various ages into the formula describing the law.

Cramer and Wold (1935) applied the method to Swedish mortality rates from 1801 to 1930 for lives aged between 30 and 90 using Makeham's formula. The parameter α_t was fitted as a straight line separately for males and females; β_t and c_t were both transformed by taking logarithms and then each was fitted to the logistic function:

$$y_t = \frac{A + B\exp[k(t+t_0)]}{1+\exp[k(t+t_0)]}, \tag{3.6}$$

where y_t is either $\log\beta_t$ or $\log c_t$. Cramer and Wold predicted mortality rates to 1980 using their model. The method however failed when it was applied in 1949 to the considerably shorter sequence of Australian mortality rates by Pollard (1949), because ‘*the small number of observations on the parameters provided no discernible trend*’, (Pollard, 1987).

The Heligman and Pollard eight-parameter curve has been used in studies by Forfar and Smith (1987) using the already graduated rates from English Life Tables 1-13, for ages from 0 to 85. Each of the parameter estimates is examined graphically for extrapolation purposes and, by adopting appropriate time periods, Forfar and Smith then used log-linear extrapolation to obtain parameter estimates for 1981.

McNown and Rogers (1989) also used the Heligman and Pollard curve in studies involving U.S. mortality data. The projected parameter values were obtained using Box and Jenkins univariate time series models.

Pollard (1987) observes that although the method of projection by reference to a law of mortality has a certain theoretical appeal, independent extrapolation of individual parameters may lead to projected mortality rates which are quite unreasonable, making it difficult to apply the method. Congdon (1993) however shows ‘*the ability of a parametric approach both to reproduce observed patterns with a high degree of accuracy and to facilitate comparisons across space and time*’.

Tolley *et al* (1993) noted that the method of projection by reference to a law of mortality is more common in epidemiology and demography than in actuarial science because ‘*such forecasting through parameters has been of little practical use to actuaries*’. Perhaps the basis of this statement is the authors’ assertion that actuaries do not view forecasting methods as scientific tools but rather as business tools.

3.3.2 Projection using Bayesian graduation

Hickman and Miller (1981) produced a graduation method that allows for forecasting by assuming that the forces of mortality $\mu_{x,t}$ are, a priori, distributed over time as correlated random variables. The researchers form the likelihood of the observed deaths and survivors of discrete age intervals in terms of $(\mu_{x,t})^{1/2}$. By assuming a normal likelihood for deaths given $(\mu_{x,t})^{1/2}$, and a normal prior distribution for $(\mu_{x,t})^{1/2}$ (the conjugate prior), posterior estimates of $\mu_{x,t}$ are obtained. The posterior distribution of $\mu_{x,t}$ for values of t beyond the data provides forecasts of future mortality rates.

More recently, Olivieri and Pitacco (2002) propose a Bayesian inferential model for future mortality changes based on the Weibull distribution. The probability distribution of the random lifetime for a fixed generation of lives is assumed to be represented by the Weibull model with two random parameters.

Tolley *et al* (1993) note the following points in favour of the Bayesian graduation method:

(a) The method provides a simple method of incorporating past data and any prior knowledge into the forecast.

(b) Because the method is a graduating technique, forecasts are smooth extensions of current experience.

(c) The method does not depend heavily on a parameterized relationship of the force of mortality over time.

3.4 Relational Models

3.4.1 The Brass two-parameter logit system

The basic idea of relational models for mortality projections, due to Brass (1971) is to construct a function of a life table function such as μ_x or l_x and relate the mortality under study to that in a reference population. Brass (1971) observed empirically that the logit transformation Λ_x, given by:

$$\Lambda_{x,t} = \frac{1}{2}\log\left(\frac{l_{0,t} - l_{x,t}}{l_{x,t}}\right), \tag{3.7}$$

can be expressed as a linear function of the logit Λ_x^s in a standard table, so that:

$$\Lambda_{x,t} = \alpha_t + \beta_t \Lambda_x^s, \tag{3.8}$$

where α_t and β_t are more or less independent of x.

For the purpose of projecting mortality, use of the Brass two-parameter logit system reduces the problem to the extrapolation of two times series, α_t and β_t. The parameter α_t reflects the level of mortality while β_t indicates the relationship between child and adult mortality relative to the standard (Benjamin and Soliman, 1993). The more β_t falls the lower is the predicted mortality at younger ages in relation to the standard and the higher is the predicted mortality at older ages compared with the standard (Congdon, 1993).

Brass (1971) used the system to obtain results for Sweden as well as England using generation data (that is, birth cohort data). He suggests that, for projection purposes, one of the life tables under study would be the best base to use. The system is one of the methods applied by Congdon (1993) in his study of mortality for the 32 London boroughs and Greater London.

3.4.2 Projection by Reference to Model Life Tables

Various systems of model life tables have been developed over the past quarter century (Brass, 1971; Coale and Demeny, 1966; Organisation for Economic Co-operation and Development, 1980; United Nations, 1955 and 1982). These model life tables are particularly useful for estimating complete life tables or abridged life tables from limited mortality data. They can also be used for projecting mortality.

The method of projection by reference to model life tables may be thought of as a special case of the method of projection by reference to a law of mortality (or vice versa). Furthermore, when certain systems of model life table are employed, the method can also be thought of as an example of projection by reference to more advanced populations (described in Section 3.4.3 below).

Firstly, a system of model life tables is chosen which, it is believed, represents and will continue to represent the mortality of the population of interest. The system may involve a single parameter or two or more parameters. In the single parameter case, the parameter of the system is measured in the population at each of several time periods. Any trend in the parameter is extrapolated graphically or by mathematical formula to provide estimates of the parameter at future time periods. Projected age-specific mortality rates are obtained by entering the model life table system for the various projected values of the parameter. To adjust for the fact that the observed base mortality rates in the population may not coincide with those in the model life table having the same parameter, the relative projected change on the model life table mortality rates is applied to the observed base mortality rates of the population.

Pollard (1987) observes that this method is one of several commonly used for projecting the mortality of less developed populations.

3.4.3 Projection by Reference to a More Advanced Population

According to Pollard (1987), the method of projection by reference to a more advanced population is one of the commonest methods of mortality projection adopted in respect of both developed and less developed countries.

The method may be summarized as follows:

- A more advanced population with adequate mortality statistics is chosen, having a mortality history which, it is hoped, the population under study will emulate.
- The mortality characteristics of the population under study are compared with those of the more advanced population and similarities are noted. For example, it may be that the mortality of the population under study is essentially the same as that of the more advanced population with a lag of, say, 20 years, which appears to be slowly shortening.
- Projections of mortality for the population under study are taken as those mortality rates already experienced by the more advanced population and (when necessary) projected for the more advanced population.

New Zealand mortality in the 1930s was the lightest in the world, and it was used in the period immediately after the Second World War as a model for projecting the mortality of other countries. A projection of Japanese mortality made in 1954 by Okazaki assumed a smooth decline in Japanese mortality of 1948, age by age, until the level of the 1934-1938 New Zealand mortality was attained in 1965 (Preston, 1974). No improvement in mortality was projected beyond 1965.

3.4.4 Projection by Reference to an 'Optimal' Life Table Attainable under Ideal Conditions

Several writers have addressed the question: 'What is the optimal life table one could expect in respect of a given population?' and a variety of approaches have been adopted in an attempt to answer the question.

For example, in 1947, Whelpton, Eldridge and Siegel studied the age-specific mortality rates in each of the states in the United States, and noted that the mortality rates in states with low mortality at any given time indicated the likely death rates for the nation as a whole some years later. Based on this information, they estimated expectations of life at birth of 68.4 and 71.8 for males and females, respectively, on the basis of individual state mortality rates in 1940. On the assumption that advances in public health and living standards would make it possible even to exceed these expectations, they concluded that the figures represented lower bounds for the year 2000. Using data from other nations rather than state data, and the same reasoning, they obtained expectations of life at birth of 68.6 and 70.9, for males and females respectively. (Cause-of-death trends were also used less formally and on a regional basis to estimate attainable reductions in mortality.)

Bourgeois-Pichat (1952) asked a similar question: 'Can mortality decline indefinitely or is there a limit, and if so, what is this limit?' He distinguished two categories of deaths: those that were exogenous (provoked by health conditions etc.) and those that were endogenous (coming from within). Using six broad groupings of cause of death and Norwegian data, Bourgeois-Pichat estimated ultimate expectations of life of 76.3 and 78.2 for males and females, respectively.

More recently, Benjamin (1982) has made some ‘extreme assumptions’ about improvements in mortality by cause in an attempt to come up with a life table under optimal conditions. Briefly, he assumed:

(a) Congenital/early infancy diseases reduced to one third;

(b) Smoking drastically reduced, eliminating 90 per cent of lung and bronchus cancer deaths and one third of pre-65 ischaemic heart disease deaths;

(c) Remaining heart disease deaths, cerebrovascular and other circulatory disease deaths deferred 10 years;

(d) Bronchitis, emphysema and asthma deaths prevented;

(e) Other cancer deaths eliminated;

(1) Accidental death unchanged;

(g) Small residual deaths from tuberculosis and diabetes;

(h) Unspecified causes of death deferred 10 years.

On this basis and using England and Wales data, he estimated an ultimate expectation of life at birth of 81.3 for males and 87.1 for females.

The essential steps in this method are the following:

- A suitable optimal life table attainable under ideal conditions is selected from those developed by other researchers or developed from the population's own cause-of-death data, taking account of optimal improvements for each cause along the lines suggested by Benjamin (1982).
- A decision is taken as to how the population will approach the optimal mortality schedule and how quickly it will do so. A formula like (3.3) will often be adopted, with α_x the optimal mortality rate at age x.

3.5 Projection by Cause of Death

Pollard (1949) studied various methods of projecting mortality using Australian data. Among them was the method of projecting total mortality from individual groups of diseases. He argued that a better projection of total mortality would be obtained by first separately projecting mortality from certain groups of disease and then adding these individual projections together. Pollard distinguished 13 cause groups: influenza, pulmonary tuberculosis, epilepsy, bronchitis and pneumonia, accidents, growths, intercranial lesions, diabetes, nephritis, appendicitis, diseases of the circulatory system, ulcers of the stomach and duodenum, and other causes. Mortality rates for these causes were calculated for selected ages for each of the years 1921-1938 and projected forward graphically towards 1970. In the event, the projected age-specific mortality rates obtained by combining the projected age-specific rates by cause were significantly higher than the rates obtained by any other method. Pollard (1987) argues that the high mortality rate predicted reflects the rapidly rising mortality rates from circulatory system diseases and, to a lesser extent, accidents.

Some of the more recent work in projecting mortality rates by cause of death has been done by actuaries in the United States Social Security Administration. Their method involves: a historical analysis of age-specific death rates by cause using 10 broad categories based on the International List of Diseases and Causes of Death code (see for example SSA Actuarial Study 112, Social Security Administration 1997); the use of expert opinion and judgement to assess future trends by cause; and, a mapping of history and expert opinion into projections.

The method of projection by cause of death may be summarized as follows:

- Cause-of-death statistics are used to calculate age-specific mortality rates by cause at each of several recent time periods for selected ages.
- The age-specific mortality rates by cause are projected separately for the selected ages, using one of the methods already outlined in the sections above.
- The projected age-specific mortality rates by cause are then combined to yield the projected mortality rates at the selected ages.
- Projected mortality rates at the intervening ages are found by interpolation or by an abridged life table technique.

The method of projection by cause of death is a complex approach not least because many different causes can result in the same disease. Wilmoth (1995) showed that when extrapolative methods are applied by cause of death and separately to aggregate mortality, the results are generally not the same over time. The cause of death that declines most slowly is the one that dominates in the end.

The method assumes that causes of death are independent. A more accurate method would involve allowing for dependency. For example, Carriere (1994) investigated the effect of removing heart and cerebrovascular diseases as a cause of death from the U.S. population. Assuming that these diseases were dependent on other causes, the dependence was modelled using the theory of *copula* functions. Loosely, a copula is defined as a multivariate cumulative distribution function that has uniform marginals. For a more precise definition refer to Carriere (1994) or Schweizer and Sklar (1983) for example.

Pollard (1987) also observed that the method requires reliable cause-of-death statistics at several recent time periods and that changes in the International Classification of Diseases, changes in medical diagnosis and medical 'fashions' can make comparisons over a number of years difficult.

3.6 Lee-Carter Method

In 1992, Lee and Carter published a model for forecasting the level and age pattern of mortality based on a combination of statistical time series methods and a simple approach to dealing with the age distribution of mortality.

They produced the following model, which is essentially a relational model:

$$\log m_{x,t} = \alpha_x + \beta_x k_t + \varepsilon_{x,t} \qquad (3.9)$$

where

$m_{x,t}$ is the central death rate for age x at time t;

α_x coefficients describe the average shape of the age profile over time;

β_x coefficients describe the pattern of deviations from the age profile when the parameter k_t varies;

k_t describes the variation in the rates of death with time t; and

$\varepsilon_{x,t}$ is an error term.

The model is over-parameterised in that if $\{\alpha_x, \beta_x, k_t\}$ are one solution, then for any constant c, $\{\alpha_x - \beta_x c, \beta_x, k_t + c\}$ and $\{\alpha_x, \beta_x c, k_t/c\}$ must also be solutions. Therefore k_t is determined only up to a linear transformation, β_x is determined only up to a multiplicative constant, and α_x is determined only up to an additive constant.

To distinguish a unique solution, two further conditions are imposed:

(i) $\sum_x \beta_x = 1$; and

(ii) $\sum_t k_t = 0$.

Under these conditions, the α_x coefficients are simply the average values over time of the $\log(m_{x,t})$ values for each age x, i.e.

$$\alpha_x = \frac{1}{n}\sum_{t=1}^{n}\log(m_{x,t}). \tag{3.10a}$$

Alternatively, the α_x coefficients can be viewed as the logarithm of the geometric mean of the central death rates $m_{x,t}$ averaged over time for each age x, that is,

$$\alpha_x = \log\left\{\prod_{t=1}^{n} m_{xt}^{1/n}\right\}. \tag{3.10b}$$

Lee and Carter (1992) used the single value decomposition (SVD) method to find a least-squares solution. The procedure is applied to the matrix of the logarithms of the death rates, after the averages over time of the (log) age-specific rates (3.10a or 3.10b) have been subtracted. The SVD procedure is available in many statistical packages including S-Plus.

The time factor k_t is modelled as a stochastic time series process using standard Box-Jenkins procedures. Lee (2000) observes that in most applications of the method so far, k_t is modelled by a random walk with drift, that is:

$$k_t = c + k_{t-1} + e_t \qquad (3.11)$$

where c is a constant average rate of change and e_t is a random term whose statistical properties are estimated from the data. In this case, the forecast of k_t changes linearly and each forecasted death rate changes at a constant (age-specific) exponential rate.

The log of each age-specific mortality rate is forecast s periods ahead from base period t_0 using the following equation:

$$\log\left(m_{x,t_0+s}\right) = \hat{\alpha}_x + \hat{\beta}_x \hat{k}_{t_0+s} \qquad (3.12)$$

where the ^ indicates estimates of the associated parameters. If the error term $\varepsilon_{x,t}$, in equation (3.9) is ignored, the variations in the $\log\left(m_{x,t}\right)$ values will be perfectly correlated with one another, because all are linear functions of the same time-varying parameter k_t. This means that probability bounds can be calculated on all (period) life table functions directly from the probability bounds on the forecasts of k_t.

Lee and Carter applied their method to the age-specific aggregate (sexes combined) US death rates from 1933 to 1987 and produced forecasts, with confidence intervals, for the period 1990 to 2065. A further study by Carter and Lee (1992) extended the basic method by implementing the model to male and female data separately. Wilmoth (1996) used the method to forecast Japanese mortality with variations to the model, in which the long-term decline of the time-factor is eventually forced to some specified level.

Lee (2000) discusses the forecasts to which the method has led and various extensions and applications of the method. One extension proposed involves estimating the a_x coefficients as the average of the $\log(m_{x,t})$ values of the most recent death rates, rather than the average of the $\log(m_{x,t})$ values over all t.

Brouhns, Denuit and Vermunt (2002) describe an extension to the Lee-Carter approach whereby the numbers of deaths are modelled as Poisson random variables $A_{x,t}$, that is:

$$A_{x,t} \sim \text{Poisson}\left(R^{c}_{x,t}\mu_{x,t}\right),$$

where $R^{c}_{x,t}$ is the exposure-to-risk and μ_{xt} is the force of mortality. The function $\mu_{x,t}$ is then modelled as:

$$\mu_{x,t} = \exp\left(\alpha_x + \beta_x k_t\right).$$

Hence the force of mortality is assumed to have the log-bilinear form:

$$\ln\left(\mu_{x,t}\right) = \alpha_x + \beta_x k_t .$$

The parameters α_x, β_x and k_t (defined as in the classical Lee-Carter model (3.9)) are estimated by maximising the log-likelihood. The time series part of the Lee-Carter methodology is not modified and is thus used to forecast k_t.

Renshaw and Haberman (2003) present a re-interpretation of the model underpinning the Lee-Carter methodology for forecasting mortality. A parallel methodology based on generalised linear modelling is introduced and the two methods are compared in terms of structure and assumptions.

3.7 Projections for Annuity Business

Under conditions of improving mortality, projection of future annuitant mortality is essential since any increased longevity will result in higher obligations in the future. Macdonald (1997) identifies two approaches to the setting of tariffs for annuity business in Western Europe. Firstly, he observes that if the tariff is regulated as in Sweden, and the interest basis is very strong, then technical mortality losses are likely to be of comparatively little significance, and the mortality basis can be chosen in a pragmatic way. On the other hand, where competition on the basis of price is allowed and offices often set their own tariffs using current interest rates, then the avoidance of technical losses is a critical matter, so that it is necessary to allow for future improvements in mortality, and possibly temporary initial selection as well. In the UK, both of these conditions apply, and consequently, the standard mortality tables produced by the Continuous Mortality Investigations (CMI) Committee of the Institute and Faculty of Actuaries for pensioners and annuitants allow for projected improvements in mortality.

Other countries in Western Europe making use of projected mortality for annuity business are Austria, France, Germany, Italy and The Netherlands (Macdonald *et al*, 1998 and Macdonald, 1997). The forecasting models adopted by these countries are described in Sections 3.7.1 and 3.7.2, using the study edited by Macdonald (1997) as the basis, with particular emphasis on the CMI model (described fully in CMI Report 10, 1990). In Section 3.7.3, the Renshaw *et al* (1996) GLM approach to modelling time trends in mortality, which is the main focus of this thesis, is described, while the Willets (1999) method is briefly described in Section 3.7.4.

3.7.1 Current CMI Practice

The current practice of the CMI is to graduate the force of mortality at age x, μ_x, by fitting the "Gompertz-Makeham" class of formulae described in Section 2.3.2 as:

$$\mu_x = GM_x(r,s) = \sum_{i=1}^{r-1} \alpha_i x^i + \exp\left[\sum_{j=0}^{s-1} \beta_j x^j\right], \tag{3.13}$$

with the convention that when $r = 0$, the polynomial term is absent, and when $s = 0$, the exponential term is absent (Forfar *et al*, 1988). Values of q_x are estimated from those of μ_x using Simpson's rule and Romberg integration or 'accelerated convergence'. A full description of the method of approximation is given by Waters and Wilkie (1987).

Tables resulting from the graduation, referred to as base tables, are then projected by applying time reduction factors, $RF(x,n)$, for an ultimate life attaining exact age x at time n, where n is measured in years from an appropriate origin (the *base year*), that is, $n = 0,1,2,\cdots$.

The projection formula used by the CMI Committee is of the form:

$$q_{x,n} = q_{x,0} \,.\, RF(x,n) \tag{3.14}$$

where:

$q_{x,n}$ is the rate of mortality for a life attaining exact age x during calendar year *base year* + n;

$q_{x,0}$ is the rate of mortality in the appropriate base table corresponding to lives attaining exact age x, in the base year;

and,

$$RF(x,n) = \frac{q_{x,n}}{q_{x,0}}. \tag{3.15}$$

Similarly for select lives (immediate annuitants), $RF(x,n)$ is the ratio of $q_{[x],n}$, the select rate of mortality for a life newly selected at age x in calendar year, *base year* + n, and $q_{[x],0}$, the select rate of mortality in the appropriate base table. The select period used for immediate annuitants is one year.

Having considered a variety of relevant factors and an assessment of likely changes in future improvements in mortality rates, the CMI Committee determined a time reduction factor of the form:

$$RF(x,n) = \alpha(x) + \{1 - \alpha(x)\}e^{-\beta_x n} \tag{3.16}$$

where, $0 < \alpha(x) \leq 1$ and $\beta_x > 0$ for all x; so that,

$$\lim_{n \to \infty} RF(x,n) = \alpha(x) \tag{3.17}$$

and

$$q_{x,\infty} = \lim_{n \to \infty} q_{x,n} = \alpha(x) q_{x,0} \,. \tag{3.18}$$

Thus the form of the model assumes that at each age x, the limiting rate of mortality is non-zero and that the rate of mortality decreases to its limiting value by exponential decay. The parameter β_x determines the speed with which the mortality rate decreases to its limiting value.

Denoting $f_s(x)$ to be the fraction of the total future reduction in q_x which will occur by some given future time s, that is,

$$q_{x,0} - q_{x,s} = f_s(x)(q_{x,0} - q_{x,\infty}), \tag{3.19}$$

it can be shown that:

$$e^{-\beta_x} = \{1 - f_s(x)\}^{1/s} \,. \tag{3.20}$$

By substituting expression (3.20) in equation (3.16), the time reduction factor becomes

$$RF(x,n) = \alpha(x) + \{1 - \alpha(x)\}\{1 - f_s(x)\}^{n/s} \,. \tag{3.21}$$

Assuming that relative to the limiting value, the speed of convergence does not depend on age, that is $\beta_x = \beta$ for all x, then expression (3.20) becomes:

$$e^{-\beta} = (1 - f_s)^{1/s} \tag{3.22}$$

and expression (3.16) becomes:

$$RF(x,n) = \alpha(x) + \{1 - \alpha(x)\}\{1 - f_s\}^{n/s} \,. \tag{3.23}$$

The more recent CMI mortality improvement model for pensioners and annuitants, is used with mortality tables based on mortality experience over the quadrennium 1991-94, (CMI Report 16, 1998; CMI Report 17, 1999). The graduated rates of mortality at age x apply on average to lives attaining exact age x in calendar year 1992 (i.e. halfway through 1992) and hence time is measured in years from 1992. The reduction factor is based on a study of the mortality experiences of the five quadrennia 1975-78 to 1991-94. The form of the model is that of expression (3.21) with $s = 20$, that is, it is assumed that a given percentage of the total future decrease in mortality, $f_s(x)$, will occur in the first 20 years, with the percentage varying by age.

The functions $\alpha(x)$ and $f_s(x)$ are both linear functions of age x for $60 \leq x \leq 110$ and are defined as:

$$\alpha(x) = \begin{cases} 0.13 & x < 60 \\ 1 + \dfrac{0.87(x-110)}{50} & 60 \leq x \leq 110 \\ 1 & x > 110 \end{cases} \qquad (3.24)$$

and,

$$f_s(x) = \begin{cases} 0.55 & x < 60 \\ \dfrac{(110-x)0.55 + (x-60)0.29}{50} & 60 \leq x \leq 110 \\ 0.29 & x > 110 \end{cases} \qquad (3.25)$$

The mortality improvement model is such that the rate of improvement in mortality is assumed to depend on both age and time for lives aged between 60 and 110 years only. At ages below 60 years, the rate of improvement is assumed to depend only on time while no improvement is assumed for lives aged 110 years and above. The same reduction factors apply for all pensioners' and annuitants' experiences, male and female, for data based on both lives and amounts.

The CMI mortality improvement model used with tables based on the 1979-82 mortality experiences (that is, the previously published mortality tables), was of the form of expression (3.23) with s also equal to 20. The exact format of the reduction formula was:

$$RF(x,n) = \alpha(x) + \{1 - \alpha(x)\}0.4^{n/20}, \tag{3.26}$$

where:

$$\alpha(x) = \begin{cases} 0.5 & x < 60 \\ \dfrac{(x-10)}{100} & 60 \le x \le 110 \\ 1 & x > 110 \end{cases} \tag{3.27}$$

As for the more recent mortality improvement model (expressions 3.21, 3.24 and 3.25), model (3.26) is such that the rate of improvement in mortality is assumed to depend on both age and time for lives aged between 60 and 110 years only. In the earlier model, it was further assumed that 60% of the total future decrease in mortality would occur in the first 20 years, for all values of x.

3.7.2 Other Methods of Projecting Annuity Business

As noted above, other European countries also make use of projected mortality for annuity business. The methods of forecasting are described in detail in the study edited by Macdonald (1997). In general cohort (generation) mortality tables based on population data form the basis of the projections. A summary of some of these methods is given below.

Austria

The generation mortality tables produced in 1986 took into account projected improvements in mortality, using information up to and including the 1980-82 experience. The forecasting model was:

$$q_x(t) = q_x(t_0)\exp\{-\lambda_x(t - t_0)\}, \tag{3.28}$$

where t_0 is an arbitrarily chosen calendar year, which acts as the origin for calendar time. Thus, for any given age x, changes in time from the base year t_0 are modelled by an exponential function, with a parameter λ_x depending on age.

The generation mortality rate in respect of a life aged x, born in calendar year τ is denoted q_x^τ, that is:

$$q_x^\tau = q_x(x+\tau). \tag{3.29}$$

The quantities $\ln q_x(t_0)$ and λ_x were graduated with respect to age x using polynomials of degree n and m respectively. Thus for a fixed age x, and for suitable parameters:

$$\left.\begin{aligned} \ln q_x(t_0) &= \alpha_0 + \alpha_1 x + \ldots + \alpha_n x^n \\ \lambda_x &= \beta_0 + \beta_1 x + \ldots + \beta_m x^m \end{aligned}\right\}. \tag{3.30}$$

The projection model therefore has the form:

$$\ln q_x(t) = \sum_{i=0}^{n} \alpha_i x^i - (t - t_0)\sum_{i=0}^{m} \beta_i x^i , \tag{3.31}$$

and the generation mortality rate is given by:

$$\ln q_x^\tau = \sum_{i=0}^{n} \alpha_i x^i - (x + \tau - \tau_0)\sum_{i=0}^{m} \beta_i x^i . \tag{3.32}$$

Two sets of data were used. The first database (I), which was considered most suitable for long-term projections, included population mortality for the periods 1865-75, 1870-80, 1901-05, 1906-10, 1930-33, 1949-51, 1959-61, 1970-72, 1980-82 and provisional data for 1990-92. The second database (II) included estimated mortality rates for each year 1947-92, and was considered to be most suitable for short-term projections.

For each age x, the optimum projection period, denoted t_{switch}, before which the short-term data should be used and beyond which the long-term data should be used; and the earliest year in either set of data, denoted $t_{\min}^{(I)}$ and $t_{\min}^{(II)}$, which should be included in the fitting process were determined by considering confidence intervals of projected mortality rates. The values of t_{switch} and $t_{\min}$, which resulted in the smallest confidence intervals, were chosen. Because the results obtained were variable, averaged results were used as follows:

Males:

$$t_{\min}^{I} = \begin{cases} 1970 & x \in [67,82] \\ 1950 & otherwise \end{cases}$$

$$t_{\min}^{II} = \begin{cases} 1870 & x \in [10,85] \\ 1950 & otherwise \end{cases}$$

$$t_{switch} = 18 \quad x \in [1,100]$$

Females:

$$t_{\min}^{I} = 1950 \quad x \in [0,100]$$

$$t_{\min}^{II} = \begin{cases} 1903 & x \in [10,95] \\ 1950 & otherwise \end{cases}$$

$$t_{switch} = 16 \quad x \in [1,100].$$

This therefore resulted in four graduations for males and three for females, each resulting in a different projection model for $q_x(t)$. The separate models were then combined into a weighted average, taking into account the values of t_{switch} and $t_{\min}$, with the weights varying with age.

The published tables pertained to the 1950 generation, with age adjustments recommended for other generations.

Because the mortality rates obtained as above were based on population data, the pensioners' mortality tables incorporated further reduction factors to allow for the difference between population mortality and annuitants' mortality. At ages up to 85, reduction factors of 0.75 for males and 0.85 for females were used; at ages 96-99, a factor of 0.8 was used for both males and females; and at ages 86-95, the factors were interpolated.

The projections were reviewed in 1995 on the basis of the 1990-92 experience and some adjustments recommended. The age adjustments originally recommended for males in the 1986 study were all decreased by 1 year, while the age adjustments for females were left unchanged.

France

The published mortality table for annuity business sold after 1 July 1993, referred to as the TPRV 93, represents the mortality of female lives born in 1950, projected into the future, together with a table of age adjustments for other generations. The table is based on the mortality experience for the period 1961-87, in 5-year age groups. For fixed ages x, the crude rates were interpolated and extrapolated using functions of the form:

$$q_x(t) = \frac{\exp(f(t))}{1+\exp(f(t))}, \tag{3.33}$$

where $f(t)$ is a polynomial, that has the form:

$$f(t) = \sum_k c_k t^k .$$

(A linear function was actually used). For each fixed period t, mortality rates at individual ages were obtained by assuming an exponential progression of the probabilities of death between two 5-year age groups.

Germany

In projecting future mortality rates, an annual improvement rate, depending on age, based on the trend of mortality in the population is assumed. The improvement rate is applied to a projected table in the immediate future, using the most up-to-date information. The recommended mortality tables for annuity business are based on the population mortality projected to the year 2000 and adjusted as follows:

$$q'_x = f_x q_x - s_x^\alpha . \tag{3.34}$$

Here, f_x is a piecewise linear adjustment quantity ranging from 0.9 at ages between 0 and 20, to 0.75 at ages between 75 and 110, to allow for the lower mortality of annuitants. s_x^α is a safety loading for the statistical risk of variation calculated such that the actual number of deaths are kept above the expected number (on the basis of q'_x) with probability $1-\alpha$, where α is very small. Thus, s_x^α is such that:

$$P\left[T \geq \sum \left(q_x - s_x^\alpha\right) L_x\right] \geq 1-\alpha , \tag{3.35}$$

where,

$$T = \sum_x T_x$$

is the random number of deaths and L_x is the population aged x.

The resulting basic rates apply to the 1952-58 cohorts (male) and the 1952-57 cohorts (female) with age adjustments being used for other cohorts.

Italy

The first Italian annuity table was a select table based on adjustments to population mortality data, (see the article by Pietrobono in McCutcheon, 1986), with the adjustments only applying to mortality due to non-accidental causes. The more recent tables follow the same principle, adjusting the 1970-72 population data to allow for both future improvements and selection. Overall and accident–only (denoted by superscript a) mortality rates are projected using exponential factors as follows:

$$\left.\begin{aligned} q_x(1971+n) &= q_x(1971) r(x)^n \\ q_x^a(1971+n) &= q_x^a(1971) r^a(x)^n \end{aligned}\right\} \tag{3.36}$$

where,

$q_x(1971+n)$ is an estimate of the rate of mortality for a life attaining age x in calendar year $(1971 + n)$;

$q_x(1971)$ is the rate of mortality obtained from the base table (1971 mortality table);

$r(x)$ is the average annual rate of variation of mortality during the projection period, based on mortality variations with reference to the 1961, 1971 and 1978 mortality tables; and

n is the number of years of the projection period (20 years for males and 25 years for females).

A selection factor $g(x,t)$ depending on age and duration, is then applied to the rates of mortality excluding accidental deaths, giving the final projection formula:

$$q_{x+t}(1971+n) = q^a_{x+t}(1971+n) + g(x,t)\{q_{x+t}(1971+n) - q^a_{x+t}(1971+n)\}. \quad (3.37)$$

3.7.3 The GLM Approach

The modelling structure proposed by Renshaw *et al* (1996) detailed in Section 2.4.4 allows for mortality projections to be performed within the context of the model formula by evaluating μ_{xt} at future time periods t. The procedure involves fitting to historical data models of the form:

$$\mu_{xt} = \exp\left[\sum_{j=0}^{s}\beta_j L_j(x') + \sum_{i=1}^{r}\alpha_i t'^i + \sum_{i=1}^{r}\sum_{j=1}^{s}\gamma_{ij} L_j(x') t'^i\right], \quad (3.38)$$

where x' and t' are the transformed ages and calendar years respectively, such that both x' and t' are mapped onto the interval [-1, +1], and $L_j(x')$ are Legendre polynomials as defined in Section 2.3.2. In particular,

$x' = (x - c_x)/w_x$ and $t' = (t - c_t)/w_t$

where c_x and c_t denote mid-points of the range of x values and t values respectively, and w_x and w_t denote semi-ranges as described in Section 2.5.4.

Rewriting equation (3.38) as

$$\mu_{xt} = \exp\left\{\sum_{j=0}^{s} \beta_j L_j(x')\right\} \exp\left\{\sum_{i=1}^{r}\left(\alpha_i + \sum_{j=1}^{s} \gamma_{ij} L_j(x')\right) t'^{i}\right\}, \qquad (3.39)$$

the second multiplicative term in (3.39) may be interpreted as an age-specific trend adjustment term, provided at least one of the γ_{ij} terms is not pre-set to zero (Renshaw *et al*, 1996). In determining a suitable model, we are therefore interested in smoothness not only over age but also over time.

Assuming that a well-behaved model that adequately describes mortality trends in the experience under study can be determined, then it is possible to derive a mortality improvement model directly from the model formula. The same mortality improvement format as that used by the CMI is applied, except that the reduction factor is defined to be a ratio of forces of mortality rather than a ratio of mortality rates. Thus

$$\mu_{x,n} = \mu_{x,0} . RF(x,n), \qquad (3.40)$$

where,

$\mu_{x,0}$ is the force of mortality for a life attaining exact age x in the base calendar year (taken as year 0); that is, the base rate from the mortality table for the appropriate experience;

$\mu_{x,n}$ is the force of mortality for a life attaining exact age x in calendar year *base year* + n; and

$RF(x,n)$ is the reduction factor for an ultimate life attaining exact age x at time n where n is measured in years from the base calendar year, thus $n = 1,2,\cdots$.

Rewriting expression (3.40) as

$$RF(x,n) = \frac{\mu_{x,n}}{\mu_{x,0}}, \tag{3.41}$$

it is clear that

$$\log RF(x,n) = \log(\mu_{x,n}) - \log(\mu_{x,0}). \tag{3.42}$$

Denoting the base calendar year as t_0 (rather than 0), n may be expressed in terms of calendar year t and t_0 as

$$n = t - t_0 .$$

Hence taking logs of expression (3.38) or (3.39) and applying (3.42),

$$\log RF(x,n) = \log RF(x,t-t_0) = \sum_{i=1}^{r}\left(\alpha_i + \sum_{j=1}^{s}\gamma_{ij}L_j(x')\right)\left(\frac{(t-c_t)^i - (t_0-c_t)^i}{w_t^i}\right) \tag{3.43}$$

or

$$RF(x,n) = RF(x,t-t_0) = \exp\left\{\sum_{i=1}^{r}\left(\alpha_i + \sum_{j=1}^{s}\gamma_{ij}L_j(x')\right)\left(\frac{(t-c_t)^i - (t_0-c_t)^i}{w_t^i}\right)\right\}, \tag{3.44}$$

which is the trend adjustment term in the model formula (3.39), with adjustment for the transformed calendar time.

Therefore, in the GLM approach, rather than use ad-hoc methods to determine an appropriate mortality improvement factor, it is the form of the best-fitting graduation model that determines the form of the improvement factor, so that the improvement factor depends on the specific mortality experience. The general trend in mortality is represented by the α_i terms while the differences in mortality change with age are represented by the γ_{ij} terms.

In the specific case where the trend adjustment term is linear in time (that is $r = 1$), the mortality improvement factor simplifies to:

$$\log RF(x,n) = \left(\alpha_1 + \sum_{j=1}^{s} \gamma_{1j} L_j(x')\right) \frac{n}{w_t}. \quad (3.45)$$

The formula for the reduction factor (3.45) can be expressed as:

$$\log RF(x,n) = (a + b_x)n \quad (3.46)$$

where a is a constant and b_x is a polynomial in x.

The Renshaw *et al* method can be likened to the method suggested by Wetterstrand (1981). Wetterstrand used an extension of the Gompertz law to describe mortality trends of United States ultimate mortality experience from life insurance for the period 1948-77 at ages 30 to 90 years. He proposed a three-dimensional model with attained age x and experience year t (measured from an appropriate base year which was 1900 in the study) as independent variables and the force of mortality as the dependent variable. The model was of the form:

$$\mu_{x,t} = \exp(\alpha + \beta t + \gamma x). \quad (3.47)$$

Wetterstrand concluded that Gompertz's law described fairly accurately the ultimate mortality of insured lives over the 30-year period of study. He however did not use his model to forecast future forces of mortality.

The GLM approach to forecasting mortality rates is in fact a type of projection by extrapolation of mortality rates, with the extrapolation carried out by mathematical formula. Further studies on the methodology have been carried out by Renshaw and Haberman (2000, 2003).

Renshaw and Haberman (2000) suggest a method for modelling mortality reduction factors in the framework of generalized linear models. The methodology is both capable of assessing existing reduction factors, with the benefit of hindsight, and capable of projecting established data patterns in order to forecast future reduction factors. Three case studies based on the UK pensioner lives' mortality experience and the UK annuitants' mortality experience are presented to illustrate the different aspects of the methodology.

Renshaw and Haberman (2003) describe ways in which the Lee-Carter methodology of fitting and constructing mortality trends can be modified to forecast future behaviour of mortality reduction factors. A comparison is drawn with the GLM based regression methods described in Renshaw and Haberman (2000). By way of illustration, the UK male pensioner lives' experience and the UK male annuitants' experience are considered.

3.7.4 The Willets method

Willets (1999) developed a 'cohort' model for projecting mortality improvements over time and compared the model with the CMI projection model and the GLM-based model. In the Willets model, trends are projected by year of birth rather than attained age. The 'current' improvement rates for particular years of birth were derived from the England and Wales data for the period 1992 to 1997 and were chosen to be broadly consistent with values given in the 1996-based GAD National population projections and smoothed using a graphical method.

It was assumed that mortality improvement rates would remain at their current levels (by year of birth) for the period 1992 to 1997 and would move exponentially towards 'long-term' average rates, with half the change in rate occurring during the first 5 years and half of the remaining change occurring every 5 years. The 5-year time period was chosen to reflect the fact that smoking was considered to have begun to stabilise so that the part of improvements due to reduced smoking was assumed to fade away relatively rapidly. The long-term rates were based on average improvements from 1961 to 1997.

A final loading of 25% was applied to the improvement rates to convert from population improvements to improvement rates suitable for pensioners. The loading of 25% was broadly based on past differential between population and pensioner improvements. The alternative approach of using improvement rates by individual year of birth derived directly from pensioners' data was not done primarily because the appropriate data is not available. In addition, it was considered that the fluctuations in the pensioner experience would make it difficult to extract underlying trends, and that the population experience has a more stable exposed-to-risk.

Chapter 4

Modelling of Immediate Annuitants' Mortality Experience

4.1 Introduction

In this chapter, the GLM modelling structure proposed by Renshaw *et al* (1996), described in Chapter 2, Section 2.5, is used to investigate mortality trends for immediate annuitants' ultimate experiences over the period 1946 through to 1994, based on data provided by the CMI Bureau. The aim of the investigation is to identify the particular form of the models appropriate for forecasting ultimate mortality rates for annuitants.

The composition of the CMI immediate annuitants' investigation has changed to some extent during the period of study. Firstly, there was a well-documented change in the class of lives taking out annuity contracts as a result of the Finance Act 1956. Secondly, the experience has been declining rapidly over the years (see for example CMIR 16, 1998). The inclusion of annuities with a guarantee period from 1988 has gone some way to boost the experience (CMIR 14, 1995). However, this same inclusion might have the effect of changing the underlying mortality trends. Thirdly, in CMIR 16 (1998), the CMI Committee reported that '*a handful of substantial contributors*' were not able to contribute data for each year of the quadrennium 1991 to 1994. Again this might have an effect on the underlying mortality rates but any such effect has not been quantified. In this study, no attempt is made to investigate explicitly changes in mortality trends that might be due to changes in the composition of the experience.

The most recent standard mortality tables for immediate annuitants produced by the CMI Committee, were constructed with a one-year select period, providing rates for duration 0 and duration 1 year and over (1+). However, it was shown in CMIR 9 (1988), in the report on the graduation of the 1979-82 mortality experiences, that there was statistical justification for a five year select period, with durations 0 to 4 combined and duration 5 years and over (5+) separate. The Committee's decision to retain a one-year select period was based on the practical advantages, notably consistency with previous mortality tables for annuitants, the a(55) and a(90) tables, which were both produced with a one-year select period. In this study, the two definitions of ultimate experience were considered, but greater emphasis was placed on the experience at duration 1 year and over in order to facilitate comparisons with the results obtained by the CMI Committee.

Determining an appropriate trend model was done in stages. Firstly, a model that provides the best fit to the data was determined for each experience. Projections based on the model were then considered. As Forfar *et al* (1988) commented, '*if one graduation produces sensible extrapolations, and another (equally satisfactory in other respects) does not, then the former may be preferred, since it would then be possible to use the graduated rates without special adjustment*'. Therefore using the information obtained by fitting the model, and the projected mortality rates, the model was revised as necessary, noting that the preferred model is the model that produces sensible extrapolations.

An analysis of mortality trends over time in the annuitants' ultimate experience is presented in Section 4.2 separately for males and females. In Section 4.3, a method developed by Renshaw and Haberman (1997) for modelling select mortality rates relative to the corresponding ultimate rates is applied to female annuitants' data as an illustration of a complete GLM modelling procedure for mortality data that includes select rates.

4.2 Trend analysis of the immediate annuitants' mortality experience

The immediate annuitants' data analysed are for the calendar year period 1946 through to 1994, based on lives, excluding data for calendar years 1968, 1971, and 1975 which were not available. For each calendar year t, the CMI data are grouped by age x nearest birthday, and by curtate policy duration d, of 0, 1, 2, 3, 4, or 5 years and over (5+), tabulated separately for males and females. The data consist of the number of deaths occurring in calendar year t, and matching exposure to the risk of death tabulated as follows:

- **1946 - 1974**: numbers in force as at 1 January of each calendar year t for individual ages x ranging from 51 to 100 years, and classified into age groups for $x \leq 50$ and $x > 100$.

- **1976 - 1982**: initial exposed-to-risk for each calendar year t for individual ages 14 to 110 years.

- **1983 - 1994**: numbers in force at the start and end of each calendar year t for individual ages 0 to 110 years.

In addition, data pertaining to contracts issued prior to 1957 are tabulated separately for the calendar years 1969 to 1981.

Estimation of the central exposed-to-risk

In line with recent CMI graduation practice, (Forfar *et al* 1988, CMIR 9 1988, CMIR 10 1990, CMIR 16 1998 and CMIR 17 1999), the force of mortality is modelled using the central exposed-to-risk. It is therefore necessary to estimate the central exposed-to-risk, R^c_{xt}, at each age x in calendar year t for the annuitants' experience. In estimating R^c_{xt} it is assumed that birthdays, entries to and exits from the experience are spread uniformly over the calendar year.

For the actual number of deaths given, ${}^{d}a_{xt}$, at age x in calendar year t, with policy duration d, matching central exposures to the risk of death, ${}^{d}R_{xt}^{c}$, are estimated as follows:

- **1946 - 1974 and 1983 - 1994:**

  Average of the number in force at the beginning of the year, ${}^{d}P_{xt}(0)$, and the number in force at the end of the calendar year (or beginning of the following calendar year), ${}^{d}P_{xt}(1)$, where 0 and 1 denote the beginning and end of each calendar year t respectively. Thus:

$$ {}^{d}R_{xt}^{c} \approx \left({}^{d}P_{xt}(0) + {}^{d}P_{xt}(1)\right)/2 \quad (4.1) $$

- **1976 - 1982:**

  The difference between the initial exposed-to-risk, ${}^{d}R_{xt}$ and half the number of deaths, that is:

$$ {}^{d}R_{xt}^{c} \approx {}^{d}R_{xt} - {}^{d}a_{xt}/2 \quad (4.2) $$

4.2.1 Data Summary

An initial analysis of the immediate annuitants' mortality data revealed the following general features:

- The females' experience is by far the larger experience, with the females' exposed-to-risk constituting about 76% of the total exposed-to-risk for the period.
- For each calendar year t, the bulk of the experience is above age 65, with some occasional entries at younger ages.

A summary of the mortality experience is given in Table 4.1. For each experience (male or female) and for each duration d, the crude death rate is calculated as the total number of deaths divided by the corresponding central exposed-to-risk, ignoring the age variation. As can be seen from the table, the crude death rate is higher for males than for females at each of the durations. The crude death rates increase with policy duration, with the rates at policy duration 5 years and over being significantly higher than at either of durations 0, 1, 2, 3, or 4 years, evidently as a result of the self-selection known to be exercised by annuitants. The crude death rates pertaining to contracts issued prior to 1957 appear to be higher than for contracts issued after 1956, perhaps because of the combined effects of the Finance Act (1956) and the improvement in mortality over time. However, these crude rates are influenced by differences in the age structure for each experience, and as such can only serve as a broad indication of the underlying features of the mortality experience(s).

Although crude death rates for contracts issued prior to 1957 and contracts issued after 1956 would suggest a difference in the underlying mortality trends between the two periods, the data pertaining to the two periods were combined in the analyses, primarily because the pre-1957 investigation is only available for a limited period of 11 years, from 1969 to 1981 excluding 1971 and 1975. However, a separate analysis of mortality trends over the period 1958 to 1994 was also carried out.

Table 4.1 Immediate Annuitants' Mortality Experience: 1946-94

	Males			Females		
	Deaths	Central Exposed-To-risk	Crude Death Rate	Deaths	Central Exposed-To-risk	Crude Death Rate
Pre-1957 Duration 5+	1802	12608.75	0.1429	11603	97553.25	0.1189
Post 1956						
Duration 0	1784	48439.00	0.0368	2512	104775.25	0.0240
Duration 1	2133	47329.00	0.0451	3303	105251.50	0.0314
Duration 2	2247	43950.25	0.0511	3547	100133.50	0.0354
Duration 3	2220	40310.50	0.0551	3704	94920.75	0.0390
Duration 4	2236	37708.25	0.0593	3758	92022.50	0.0408
Duration 5+	37781	416783.50	0.0906	104536	1449620.25	0.0721
Post 1956 Total	48401	634520.50	0.0763	121360	1946723.75	0.0623
Grand Total	**50203**	**647129.25**	**0.0776**	**132963**	**2044277.00**	**0.0650**

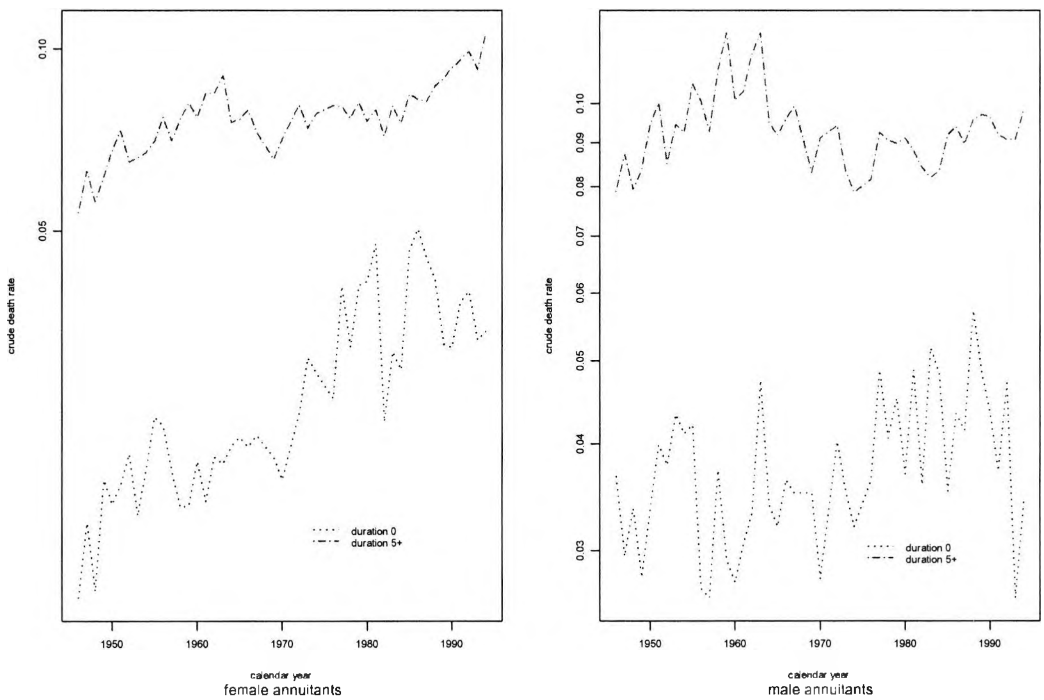


Figure 4.1 Immediate annuitants, durations 0 and 5+: crude death rates plotted on the log scale against calendar year

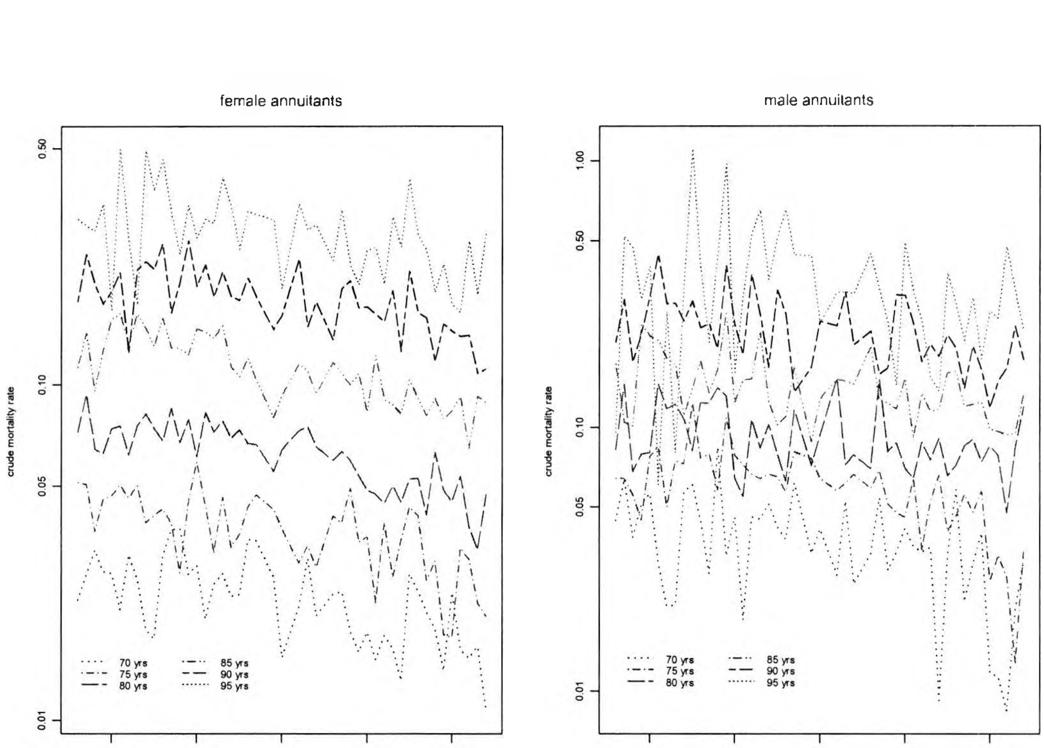


Figure 4.2 Immediate annuitants, duration 5+: crude mortality rates plotted on the log scale against calendar year

A further examination of the data reveals that the decline in numbers noted in Section 4.1 above is more pronounced in the females' experience; the females' exposed-to-risk for all durations combined, has declined from over 57000 in 1946 to just under 15000 in 1994, a decrease of about 74%. In contrast, the males' exposed-to-risk has decreased by approximately 37%, from around 13600 in 1946 to just over 8600 in 1994. As a consequence of the decline in numbers (and therefore a decline in new policyholders), the proportion of the population at risk at the higher ages is increasing, whilst the proportion of the population at risk at the younger ages is decreasing for both males and females, thus giving rise to an ageing experience. Consequently, the crude death rates shown in Figure 4.1 are increasing with time for the females' experience, as the population at risk is becoming older. For males, because the decline in numbers is relatively small, the decrease appears to have little effect on the overall trend.

The bulk of the data are at duration 5 years and over, and as such the general pattern in mortality trends may be discerned from a plot of crude mortality rates for the experience at this duration. Figure 4.2 shows crude mortality rates in 5-year age intervals from age 70 to 95 years, plotted on the log scale against calendar year, for male and female immediate annuitants' ultimate experiences at duration 5 years and over.

The crude mortality rate would be expected to increase steadily with age, and this is apparent from the plots. In addition, for the range of ages shown, there is clear evidence of a downward trend at each age, as would be expected with improvements in mortality over time. At the younger ages and at extreme old age, the data are sparse and as such the observed mortality rate is unreliable, and there is not much that can be deduced from a plot of the data. The pattern is clearer for the females' experience, reflecting the larger volume of the experience.

Since the aim is to make inferences about rates of mortality over time, the mortality experience is graduated over the range of ages for which trends over time are discernible, and this is at ages 65 to 95 for males and at ages 65 to 100 for females. Notwithstanding the observation made by Forfar *et al* (1988), that *'there is no justification for missing out ages where the data are just scanty'*, it is perceived that since consistent trends are virtually impossible to discern at ages where data are sparse, it would seem reasonable to exclude such ages when graduating with respect to age and time. It should be noted that methods such as those by Coale and Kisker (1990) and Coale and Guo (1989) could be used to analyse data at the extreme old ages excluded in this study (that is, ages above 95 for males and 100 for females).

A summary of the immediate annuitants' ultimate experience is given in Table 4.2. From the table, it can be seen that in terms of the exposed-to-risk, more than 90% of the observed experience is in the age range 65 to 95 for males and 65 to 100 for females. In other words, less than 10% of the available information is left unused when the experiences are modelled over these respective age ranges.

Table 4.2
Immediate Annuitants' Ultimate Experience, 1946-94

	Males				Females			
duration	Age range	Deaths	Central Exposed-to-risk	Percentage Of Total Exposure	Age range	Deaths	Central Exposed-to-risk	Percentage Of Total Exposure
d1+	all ages	48419	598690.3		all ages	130451	1939501.80	
	65 - 95	46393	545734.8	91%	65 - 100	127871	1754446.75	91%
d5+	all ages	39583	429392.3		all ages	116139	1547173.50	
	65 - 95	38068	401515.8	94%	65 - 100	114193	1440355.50	93%

4.2.2 Summary of Procedures and Models used

For each experience (male or female) and the specific ultimate duration d, the force of mortality, μ_{xt} at age x in calendar year t, is modelled using formulae of the type:

$$\mu_{xt} = \exp\left[\beta_0 + \sum_{j=1}^{s} \beta_j L_j(x') + \sum_{i=1}^{r} \alpha_i t'^i + \sum_{i=1}^{r}\sum_{j=1}^{s} \gamma_{ij} L_j(x') t'^i\right]. \tag{4.3}$$

x' and t' are the transformed ages x and calendar years t respectively, such that

$$x' = \frac{(x - c_x)}{w_x} \text{ and } t' = \frac{(t - c_t)}{w_t}, \tag{4.4}$$

where c_x, c_t denote the mid-points of the age and calendar year ranges respectively; w_x, w_t denote the semi-ranges; and $L_j(x')$ are Legendre polynomials defined in Chapter 2.

To estimate the unknown parameters α_i, β_j and γ_{ij}, the actual number of deaths a_{xt}, at age x in calendar year t, are modelled as independent realisations of Poisson response variables A_{xt}, of a generalised linear model with mean and variance given by:

$$E[A_{xt}] = m_{xt} = R^c_{xt}\mu_{xt}; \tag{4.5}$$

$$var(A_{xt}) = \phi m_{xt}. \tag{4.6}$$

R^c_{xt} is the central exposed-to-risk, and ϕ is a scale (dispersion) parameter to take account of the fact that, as the data are based on policy numbers rather than head counts, there may be duplicate policies issued on the same lives, resulting in over-dispersion of the Poisson random variable. The absence of over-dispersion would result in a value of ϕ equal to 1.

The unknown parameters are linked to the mean through the log function:

$$\eta_{xt} = \log m_{xt} = \log R^c_{xt} + \log \mu_{xt}, \tag{4.7}$$

so that,

$$\log m_{xt} = \log R^c_{xt} + \beta_0 + \sum_{j=1}^{s} \beta_j L_j(x') + \sum_{i=1}^{r} \alpha_i t'^i + \sum_{i=1}^{r}\sum_{j=1}^{s} \gamma_{ij} L_j(x') t'^i . \tag{4.8}$$

Estimation of the parameters is carried out by minimising the negative of the quasi-log-likelihood, that is, minimising the expression:

$$\frac{1}{\phi}\sum_{x,t}(m_{xt} - a_{xt}\log m_{xt}). \tag{4.9}$$

To determine the optimum values of r and s, the improvement in the scaled deviance for successive increases in the values of r and s, is compared with a χ^2 random variable with 1 degree of freedom as an approximation. The optimum values chosen are the minimum values of r and s beyond which improvement in the deviance is not statistically significant. It should be noted that a backward-type selection procedure could have been used to determine the optimum values of r and s.

Values of the χ^2 random variable with 1 degree of freedom, at selected probabilities, are given in the table below.

Table of χ_1^2 values at selected probabilities

probability	0.05	0.025	0.01
χ_1^2 value	3.841	5.024	6.635

From the table, it is seen that as an initial approximation, any improvement of less than about 4 in the deviance would not be considered significant.

The parameter estimates obtained are also checked for statistical significance at each stage.

In addition, when the choice between models on the basis of the analysis of deviance is not clear-cut, then based on the principle of *parsimony*, the model with the least number of parameters is chosen. According to Klugman, Panjer and Willmot (1998), the principle of parsimony states that '*unless there is a very good reason to do otherwise, the more parsimonious model should be used*'. The authors give the following definition of parsimony for modelling purposes:

'*One model is more parsimonious than another if it can be completely specified using a smaller number of objects. These objects are usually the number of parameters of the model*'.

They further note that models with a large number of parameters tend to perform poorly when used for prediction. Therefore in this study, considerable importance is attached to the principle of parsimony. Indeed Keyfitz's (1982) discussion of general aspects of forecasting mortality focuses on the principle of parsimony.

The (unscaled) deviance corresponding to the predicted forces of mortality, $\hat{\mu}_{xt}$ is:

$$D(c,f) = 2\sum_{x,t}\left\{a_{xt}\log\left(\frac{a_{xt}}{\hat{m}_{xt}}\right) - (a_{xt} - \hat{m}_{xt})\right\}, \tag{4.10}$$

where $\hat{m}_{xt} = R^c_{xt}\hat{\mu}_{xt}$, that is, $\hat{m}_{xt}$ is the number of deaths predicted by the model. The corresponding scaled deviance is defined as the deviance divided by the scale parameter ϕ, with ϕ estimated from dividing the deviance by the number of degrees of freedom v, determined from the most complex of the models being compared. Hence the number of degrees of freedom provides an estimate of the scaled deviance.

Diagnostic plots of studentized residuals and formal statistical tests of graduation are carried out on residuals computed from the data as a whole, and from each of the calendar years separately. The residuals used are:

- the deviance residuals, $sign(a_{xt} - \hat{m}_{xt})\sqrt{d_{xt}}$,

  where d_{xt} is the contribution of the particular observation in the domain $\{x, t\}$, to the deviance; and

- the Pearson residuals, $\frac{(a_{xt} - \hat{m}_{xt})}{\sqrt{\hat{m}_{xt}}}$,

  that is, the signed square root of the particular component of the Pearson goodness-of-fit statistic.

The statistical tests applied are the chi-square goodness-of-fit test; the distribution of individual standardised deviations test; the signs of deviations test; the cumulative deviations test and Stevens' grouping of signs test. Where necessary, before the residuals are computed, the data are grouped such that the expected number of deaths in each cell is at least equal to 5.

Implementation is done using the computer software package S-PLUS. The modelling procedure in S-PLUS is simplified by the availability of the function **glm** that computes the maximum likelihood estimates of the parameters $\boldsymbol{\beta}$, by solving the *score* equations

$$\frac{\partial \ell}{\partial \beta} = 0,$$

where ℓ is the (quasi) log-likelihood function (described in Chapter 2). Because the score equations are non-linear in $\boldsymbol{\beta}$, they are solved iteratively using an algorithm referred to as iteratively reweighted least-squares (IRLS). The procedure is described in detail by McCullagh and Nelder (1989) and Chambers and Hastie (1993).

Alternatively, the model can be fitted using the general minimisation function **ms** in S-PLUS.

Detailed descriptions of statistical modelling in S and S-PLUS can be found in Chambers and Hastie (1993) and Venables and Ripley (1994).

4.2.3 Analysis of the 1946-1994 female mortality experience at duration 1 year and over

The female immediate annuitants' experience at curtate policy duration 1 year and over has been analysed at individual ages x ranging from 65 to 100 years, over the calendar-year period 1946 to 1994, giving a total of 1656 data cells. As noted in Section 4.2.1, a separate analysis of mortality trends over the period 1958 to 1994 has also been carried out and the results are given in Section 4.2.5.

The fitting procedure adopted has been to determine an optimal value for s in the first instance (the Gompertz-Makeham term, $GM_x(0, s+1)$), and then to introduce the age independent trend adjustment term taking into account the value determined for s. In addition, under the assumption that the underlying shape for mortality graduation at adult ages is always of Gompertz (i.e. $GM_x(0, 2)$, (see for example Forfar *et al*, 1988) the focus is primarily on values of $s \geq 1$.

Table 4.3
Female Immediate Annuitants, Duration 1 year and over, 1946-1994 experience
Deviances for some polynomial predictors of degree *r* and *s*

	$r = 0$	$r = 1$	$r = 2$	$r = 3$	$r = 4$	$r = 5$	$r = 6$	$r = 7$
$s = 1$	5830.95	3580.84	3553.19	3529.93	3500.95	3491.80	3464.66	3459.38
$s = 2$	5686.45	3453.91	3428.96	3406.65	3381.55	3371.22	3345.09	3339.37
$s = 3$	5608.43	3367.18	3342.40	3321.41	3295.51	3284.65	3257.52	3251.88
$s = 4$	5606.40	3365.77	3340.99	3320.03	3294.18	3283.25	3256.12	3250.42

Table 4.3 shows the unscaled deviance profile resulting from fitting formulae composed of the Gompertz-Makeham term for age effects, together with the multiplicative age independent factor to adjust for calendar year effects. That is, a model of the form:

$$\mu_{xt} = \exp\left[\sum_{j=0}^{s}\beta_j L_j(x') + \sum_{i=1}^{r}\alpha_i t'^i\right], \tag{4.11}$$

with,

$$c_x = 82.5,\ w_x = 17.5, \text{ so that } x' = \frac{x - 82.5}{17.5}; \tag{4.12}$$

and

$$c_t = 1970,\ w_t = 24, \text{ giving } t' = \frac{t - 1970}{24}. \tag{4.13}$$

From the table, it can be seen that the improvement in deviance for successive increases in the value of s would suggest the optimum value of s to be 3, corresponding to a $GM_x(0, 4)$ graduation. When s is increased to 4, the improvement is not significant and β_4, the additional parameter introduced is not significantly different from zero.

Although there is a steady improvement in deviance for successive increases in the value of r, up to $r = 6$, when r is increased from 4 to 5, one of the 9 parameters in the model (assuming s is 3) is not statistically significant. However, subsequent increases in the value of r result not only in an improvement in deviance, but all the parameters introduced up to and including α_6 are statistically significant. The value of r can therefore be chosen to be either 4 or 6. Based on the principle of parsimony, the model with fewer parameters is preferred and hence the value of r is chosen to be 4. As observed by Klugman, Panjer and Wilmot (1998), the aim '*is to create simple models that adequately (but not necessarily perfectly) capture the essence of our data*'.

The complete model is therefore of the form:

$$\mu_{xt} = \exp\left[\sum_{j=0}^{3} \beta_j L_j(x') + \sum_{i=1}^{4} \alpha_i t'^i + \sum_{i=1}^{4}\sum_{j=1}^{3} \gamma_{ij} L_j(x') t'\right], \tag{4.14}$$

with $L_1(x') = x'$; $L_2(x') = \frac{3x'^2 - 1}{2}$; $L_3(x') = \frac{5x'^3 - 3x'}{2}$;

subject to the condition that some of the 12 γ_{ij} terms (i = 1, 2, 3, 4 and j = 1, 2, 3) may be set to zero.

The next stage involves introducing mixed product terms in age x and time t, that is, the γ_{ij} terms. The deviance profile shown in Table 4.4, indicates that there is significant improvement in deviance only when γ_{11} and γ_{12} are the additional parameters introduced. Replacing γ_{13} with γ_{21} in the model results in an improvement in deviance of 0.97, clearly not significant. Thus the model chosen is:

$$\mu_{xt} = \exp\left[\sum_{j=0}^{3} \beta_j L_j(x') + \sum_{i=1}^{4} \alpha_i t'^i + t'\sum_{j=1}^{2} \gamma_{1j} L_j(x')\right], \tag{4.15}$$

that is, a model consisting of a $GM_x(0,4)$ term in age effects and a trend adjustment term that is a polynomial of degree 4 in time t on the log scale. The interaction term in age and time is a linear function of t on the log scale, with the coefficient of t being a quadratic in age x.

Parameter estimates based on the model given by expression (4.15), their standard errors, and the associated t-statistics are given in Table 4.5. The absolute values of the t-statistics are all greater than 2, indicating statistical significance of each of the parameters. The estimated value of the scale parameter ϕ is 1.94.

Table 4.4
Female Immediate Annuitants, Duration 1 year and over, 1946-1994 experience
Deviance profile (terms added sequentially 1st to last)

Parameter	Deviance	Degrees of freedom	Difference in Deviance
β_0	77881.91	1655	
β_1	5830.95	1654	72050.96
β_2	5686.45	1653	144.49
β_3	5608.43	1652	78.03
α_1	3367.18	1651	2241.24
α_2	3342.40	1650	24.78
α_3	3321.41	1649	20.99
α_4	3295.51	1648	25.9
γ_{11}	3252.44	1647	43.07
γ_{12}	3214.41	1646	38.03
γ_{13}	3213.09	1645	1.32

Table 4.5
Female Immediate Annuitants, Duration 1 year and over, Model based on the 1946-94 experience

Parameter	Estimate	std. error	t-statistic
β_0	-2.565261	0.01	-313.68
β_1	1.810082	0.01	137.24
β_2	-0.086329	0.01	-6.35
β_3	-0.115944	0.01	-7.82
α_1	-0.264841	0.02	-14.48
α_2	0.108129	0.05	2.35
α_3	0.056555	0.03	2.17
α_4	-0.194091	0.05	-3.81
γ_{11}	0.069489	0.02	3.52
γ_{12}	0.107053	0.02	4.43

Table 4.6
Female Immediate Annuitants, Duration 1 year and over, analysis of the 1946-94 experience,
Distribution of individual studentized deviance residuals for the data as a whole

Range	(-∞,-3)	(-3,-2)	(-2,-1)	(-1,0)	(0,1)	(1,2)	(2,3)	(3,∞)
expected frequency	2.16	34.24	217.45	546.15	546.15	217.45	34.24	2.16
observed frequency	2.00	31.00	237.00	550.00	536.00	203.00	38.00	2.00

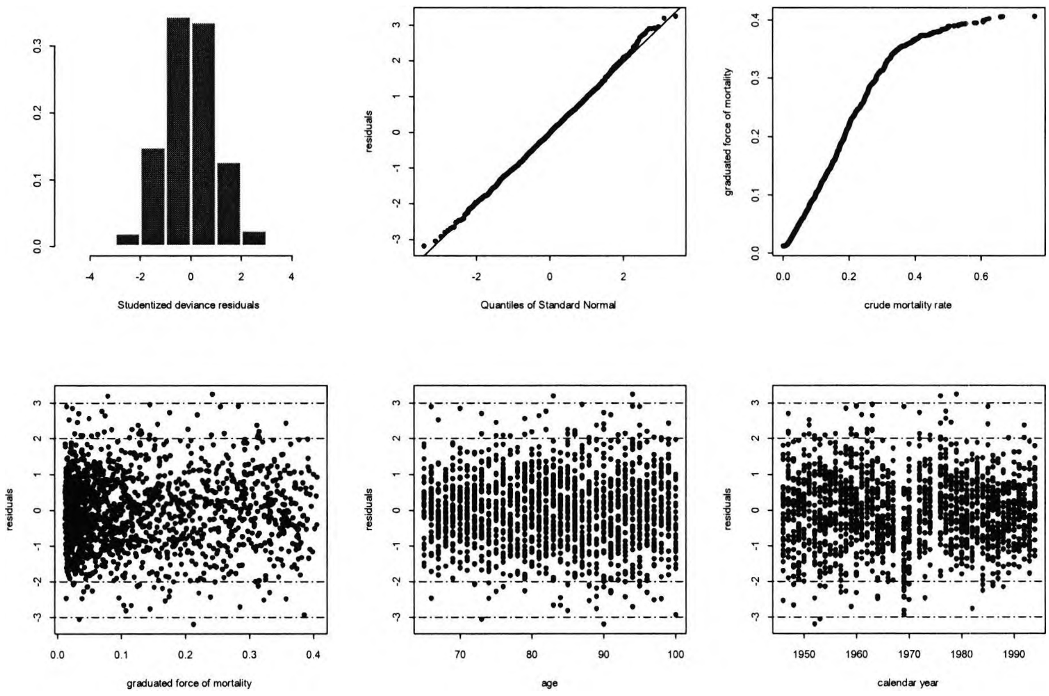


Figure 4.3 Female immediate annuitants, d1+ years, analysis of the 1946-94 experience, diagnostic plots: 10 parameter log-link model with $r = 4$, $s = 3$, γ_{11} and γ_{12}

Diagnostic plots of studentized deviance residuals for the data as a whole are shown in Figure 4.3. The plots, together with the distribution of studentized deviance residuals shown in Table 4.6, appear to indicate a normal distribution for the residuals, with less than 5% of the residuals exceeding 2 in absolute size. In addition, plots of the residuals against the graduated forces of mortality, age and calendar year do not show any obvious pattern. It would therefore appear that the model defined by equation (4.15) provides a good representation of the underlying forces of mortality over the period as a whole. The model passes all the statistical tests of graduation applied to the data as a whole, confirming the adequacy of the model and the distributional assumptions. The p-values for the various graduation tests are shown in Table 4.7.

Table 4.7
Female immediate annuitants, duration 1 year and over, analysis of the 1946-94 experience
***p*-values based on a 10-parameter model**

Chi square	0.50
Cumulative deviations	0.40
Individual standardised deviations	0.61
Grouping of signs	0.94
Signs of deviations	0.31

It is interesting to note that when the individual standardised deviations test is applied to the studentized Pearson residuals, the model fails this test (p-value 0.0178). As has been noted in chapter 2, Pierce and Schafer (1986) observed that the deviance residual is preferred to the Pearson residual for model checking procedures because its distributional properties are known and are closer to the residuals arising in linear regression models. The underlying distribution of deaths is non-normal and hence the Pearson residuals arising from model fitting might not have a normal distribution. Therefore, in all subsequent analyses in this study, the individual standardised deviations test is applied to the studentized deviance residuals.

When the graduation tests are applied to each of the calendar years separately, the graduation fails some of the tests applied in a significant number of years in the earlier part of the observation period, particularly in calendar years 1947, 1951, 1963, 1969, 1972, 1976 and 1977, when the graduation fails the chi-square goodness-of-fit test and the cumulative deviations test. In the later period from 1979, the statistical tests applied to each calendar year prove to be quite supportive of the model, with the graduation only failing the chi-square test in 1979; the cumulative deviations test and signs of deviations test in 1993; and the signs of deviations test in 1994. A summary of the p-values resulting from the statistical tests applied separately to each calendar year is given in Table 4.8, with the corresponding plots of deviance residuals against age x for each calendar year t, and plots of deviance residuals against calendar year t for each age x shown in Figures 4.4, 4.5, 4.6 and 4.7.

Although the overall shapes of the residual plots generally show a random pattern, there is however some evidence of patterns being exhibited in the residuals for some of the calendar years and ages. For example, the residuals for calendar years 1948 and 1962 appear to follow a cyclical pattern; for calendar year 1969, and at specific individual ages 67 and 100, there is an excess of negative deviations indicating that the graduated forces of mortality might be too high. Indeed, the residual plot pertaining to 1969 is notable for the non-null pattern exhibited.

Plots of graduated forces of mortality and crude mortality rates against age for each of the 46 calendar years, seem to confirm the biasedness of the graduated rates implied by the failure of the cumulative deviations test in a number of years, such as 1947, 1948, 1951, 1956, 1963 and 1969. These plots are shown in Figures 4.8 and 4.9. In the later part of the investigation period from 1979, the model appears to provide an adequate fit in each of the calendar years although the underlying forces of mortality in 1993 seem overestimated. This is very likely a problem with the data rather than the model, since some contributors were unable to provide data during this period.

Figure 4.10 shows crude mortality rates and graduated forces of mortality plotted against t at 5-year age intervals from age 65 to 100 years. For clarity, the plot has been split into two, with each plot showing the rates in 10-year age intervals, from 65 to 95 and from 70 to 100 years. A visual inspection of the plots indicates that the model provides a good fit at all ages shown with the exception of age 100, where the graduated forces of mortality tend to be higher than the observed crude rates in most years from 1964 onwards. From the plot of crude and graduated rates at each age x, for $x = 95$ to 100, shown in Figure 4.11, it appears that the graduation is satisfactory for all x less than 100 years. Observed mortality rates at old ages are generally considered unreliable and this seems to be confirmed by the absence of a smoothly progressing pattern at age 100.

Table 4.8
Female Immediate Annuitants, Duration 1 year and over, analysis of the 1946-94 experience
Tests of graduation applied to individual calendar years: *p*-values[1]

year	chi-square	cumdev	isd	runs	signs
1946	0.10	0.67	0.94	0.46	0.62
1947	0.00	0.00	0.00	0.90	0.41
1948	0.01	1.00	0.11	0.51	0.18
1949	0.17	0.80	0.40	0.19	1.00
1950	0.19	0.05	0.82	0.25	0.87
1951	0.03	0.00	0.02	0.81	0.00
1952	0.01	0.97	0.61	0.12	0.31
1953	0.00	0.89	0.08	0.17	0.09
1954	0.15	0.94	0.76	0.08	0.18
1955	0.58	0.71	0.31	0.20	1.00
1956	0.04	0.01	0.43	0.49	0.62
1957	0.11	0.97	0.12	0.07	0.62
1958	0.01	0.52	0.98	0.74	0.62
1959	0.03	0.00	0.17	0.36	0.31
1960	0.01	0.76	0.10	0.62	0.24
1961	0.14	0.08	0.72	0.25	0.87
1962	0.16	0.01	0.55	0.41	0.24
1963	0.00	0.00	0.00	0.88	0.01
1964	0.17	0.98	0.26	0.45	0.07
1965	0.41	0.98	0.25	0.45	0.07
1966	0.73	0.19	0.16	0.90	0.41
1967	0.31	1.00	0.04	0.87	0.03
1969	0.00	1.00	0.00	0.05	0.00
1970	0.03	0.97	0.27	0.19	0.07
1972	0.03	0.03	0.61	0.02	1.00
1973	0.46	0.59	0.82	0.09	1.00
1974	0.37	0.10	0.29	0.04	0.03
1976	0.00	0.00	0.02	0.32	0.01
1977	0.01	0.00	0.13	0.45	0.03
1978	0.21	0.15	0.47	0.18	0.41
1979	0.03	0.04	0.57	0.63	0.13
1980	0.60	0.85	0.40	0.08	0.18
1981	0.58	0.24	0.64	0.58	0.31
1982	0.43	0.98	0.48	0.71	0.09
1983	0.49	0.17	0.11	0.54	0.23
1984	0.18	0.95	0.30	0.06	0.23
1985	0.05	0.46	0.65	0.56	1.00
1986	0.28	0.44	0.35	0.29	0.73
1987	0.42	0.97	0.47	0.90	0.30
1988	0.52	0.78	0.50	0.11	0.22
1989	0.77	0.85	0.26	0.92	0.86
1990	0.57	0.15	0.78	0.90	0.60
1991	0.57	0.06	0.57	0.15	0.22
1992	0.39	0.22	0.07	0.89	1.00
1993	0.24	0.98	0.16	0.85	0.03
1994	0.47	0.35	0.09	0.07	0.01

[1] cumdev denotes cumulative deviations, isd denotes individual standardised deviations, runs denotes grouping of signs, signs denotes signs of deviations

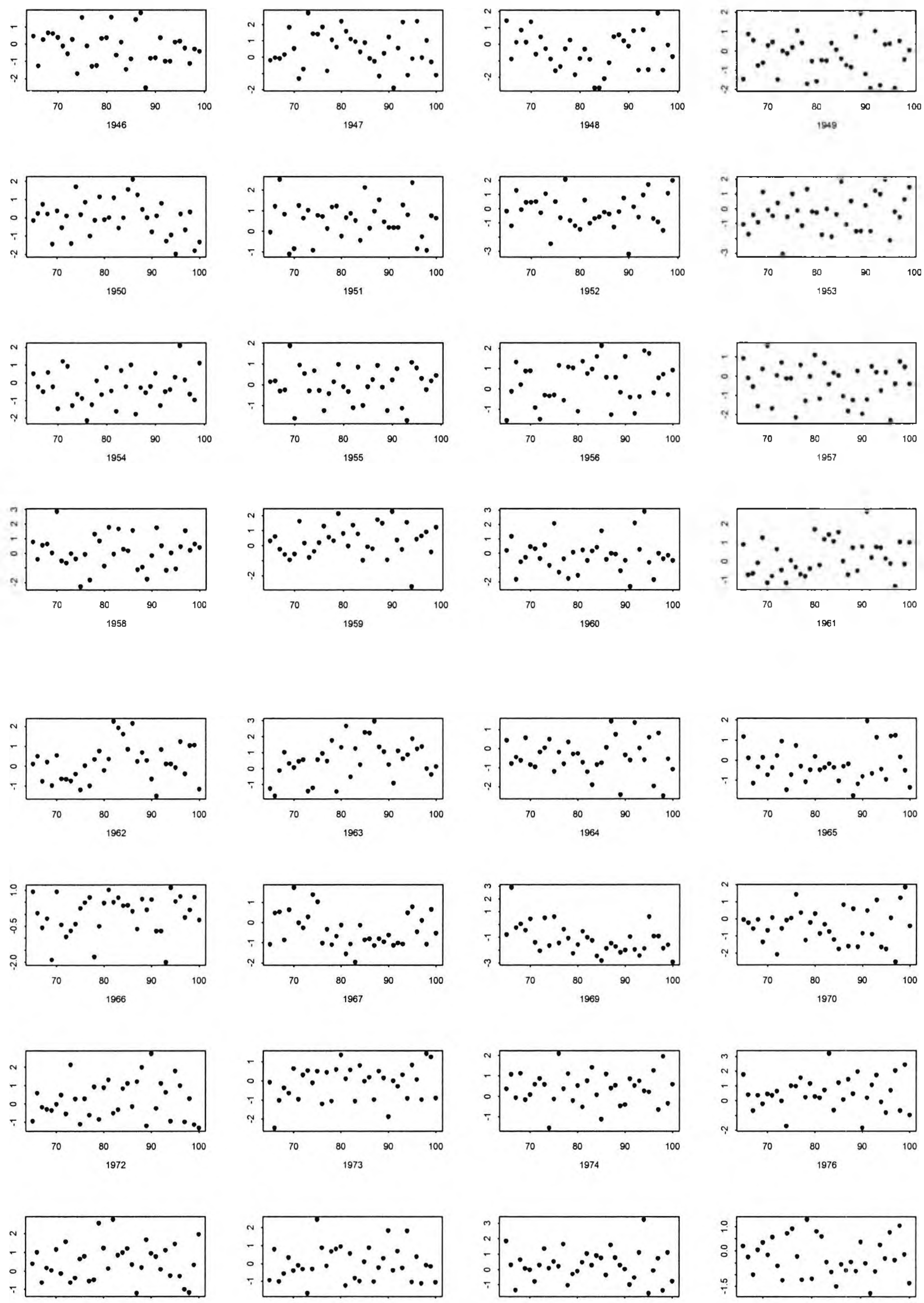


Figure 4.4 Female immediate annuitants, d1+ years, analysis of the 1946-94 experience, studentized deviance residuals plotted versus age on the log scale for individual calendar years 1946 to 1980

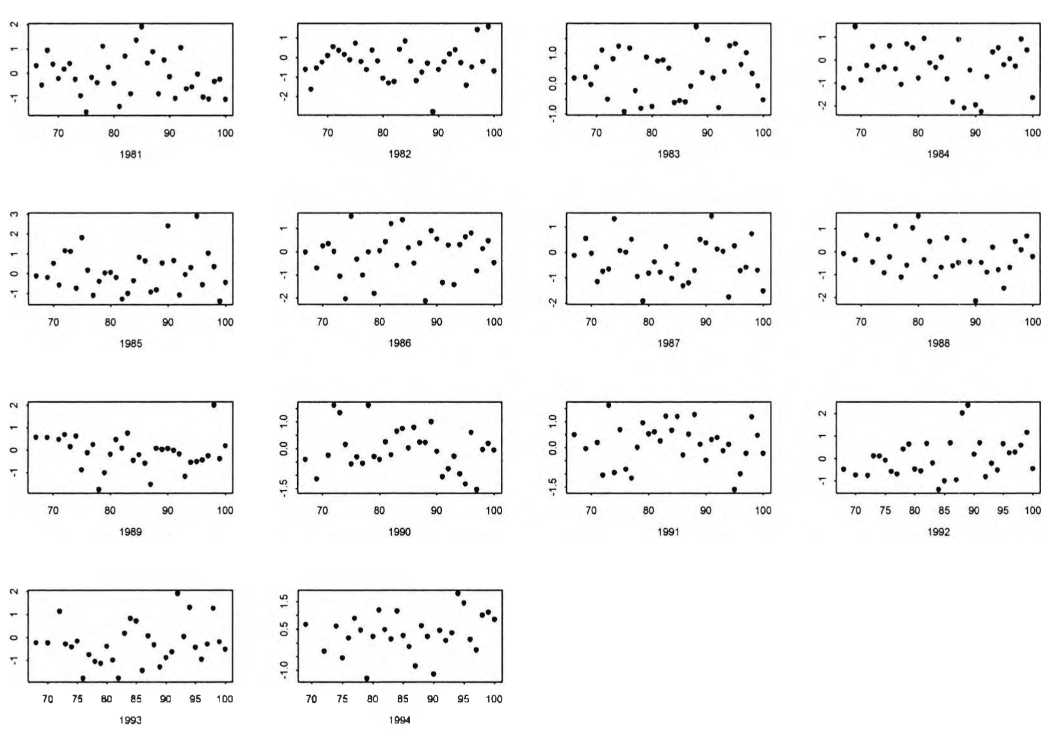


Figure 4.5 Female immediate annuitants, d1+ years, analysis of the 1946-94 experience, studentized deviance residuals plotted versus age on the log scale for individual calendar years 1981 to 1994

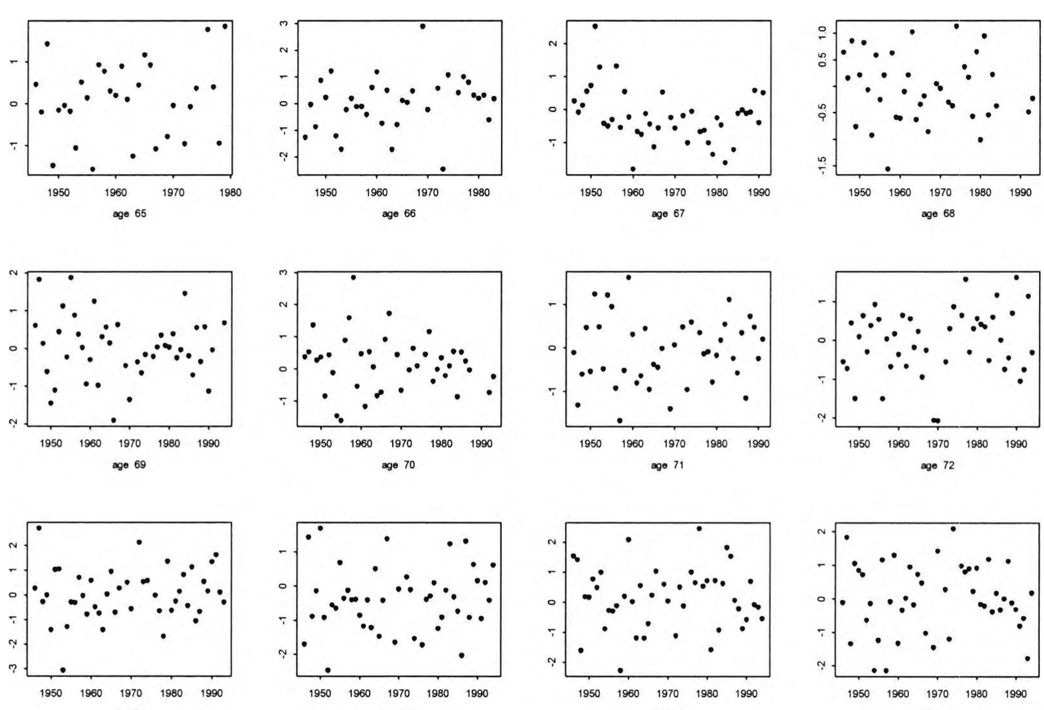


Figure 4.6 Female immediate annuitants, d1+ years, analysis of the 1946-94 experience, studentized deviance residuals plotted versus calendar year on the log scale for individual ages 65 to 76

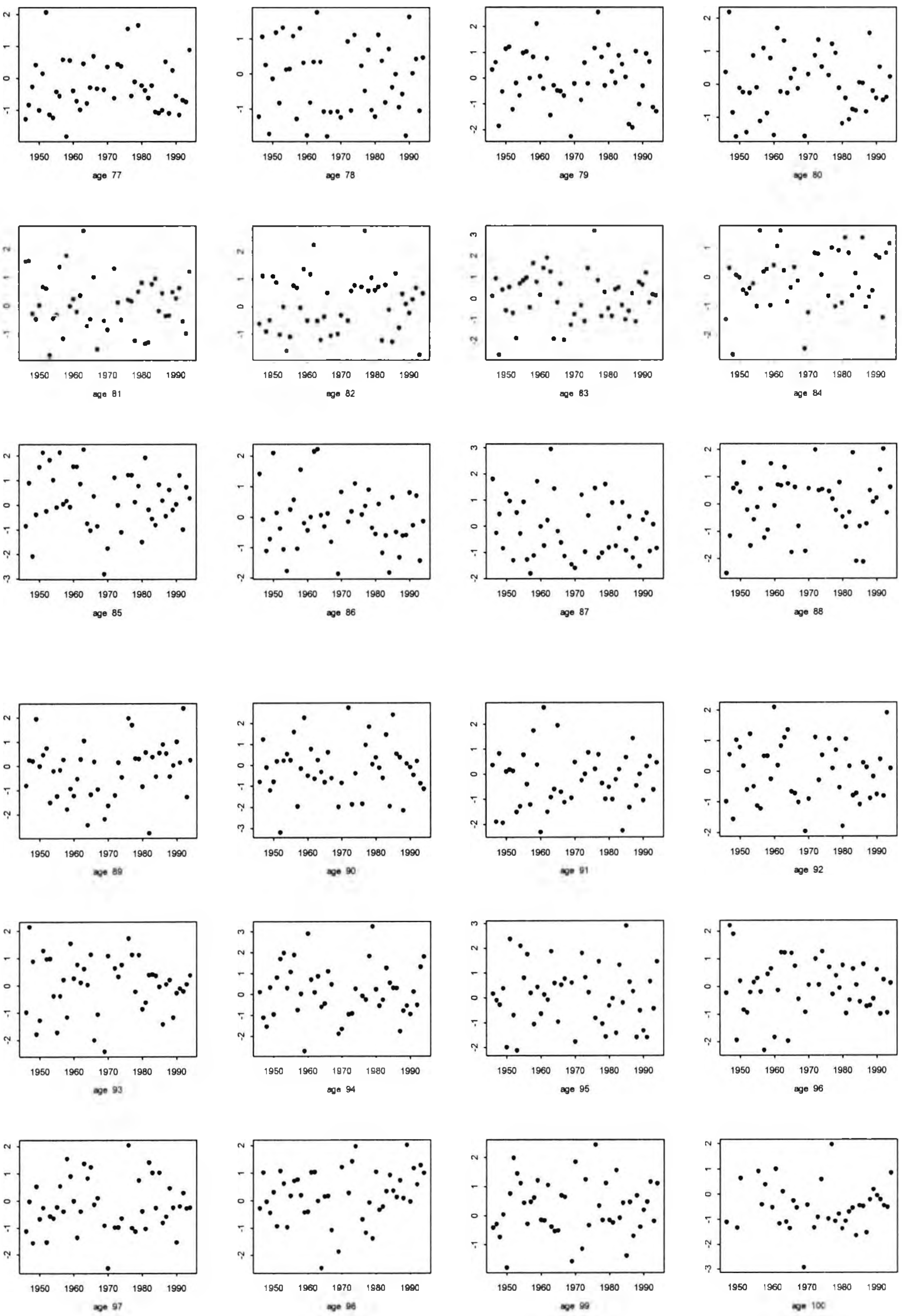


Figure 4.7 Female immediate annuitants, d1+ years, analysis of the 1946-94 experience, studentized deviance residuals plotted versus calendar year on the log scale for individual ages 77 to 100

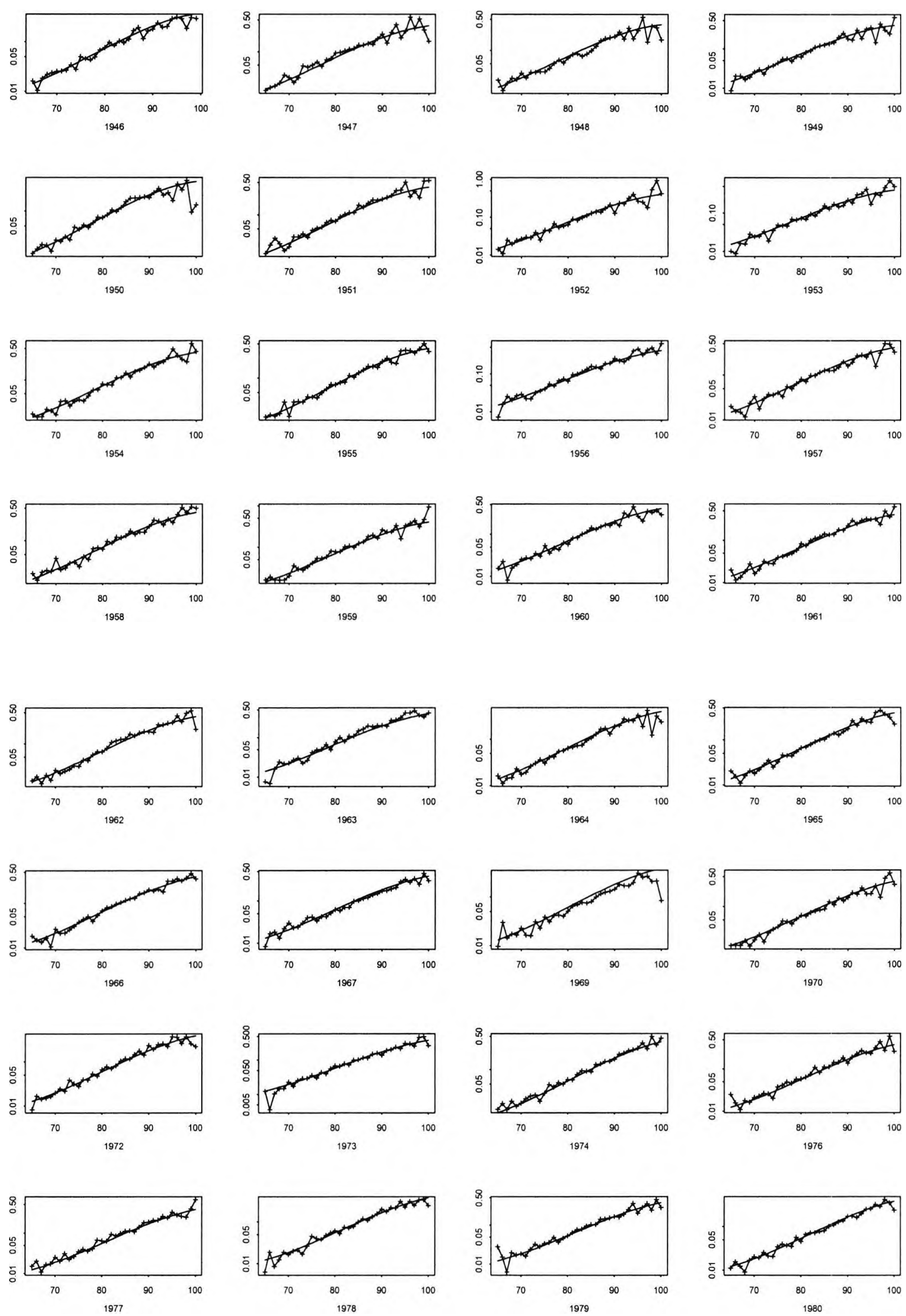


Figure 4.8 Female immediate annuitants, d1+ years, analysis of the 1946-94 experience, crude mortality rates and graduated forces of mortality plotted versus age on the log scale for individual calendar years 1946 to 1980

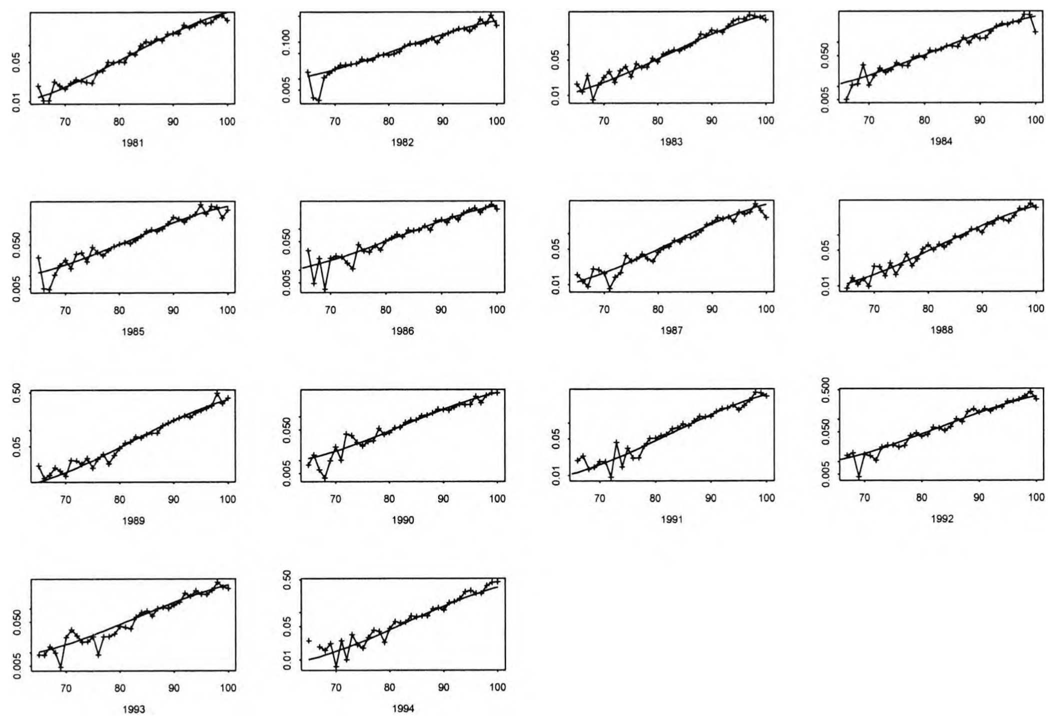


Figure 4.9 Female immediate annuitants, d1+ years, analysis of the 1946-94 experience, crude mortality rates and graduated forces of mortality plotted versus age on the log scale for individual calendar years 1981 to 1994

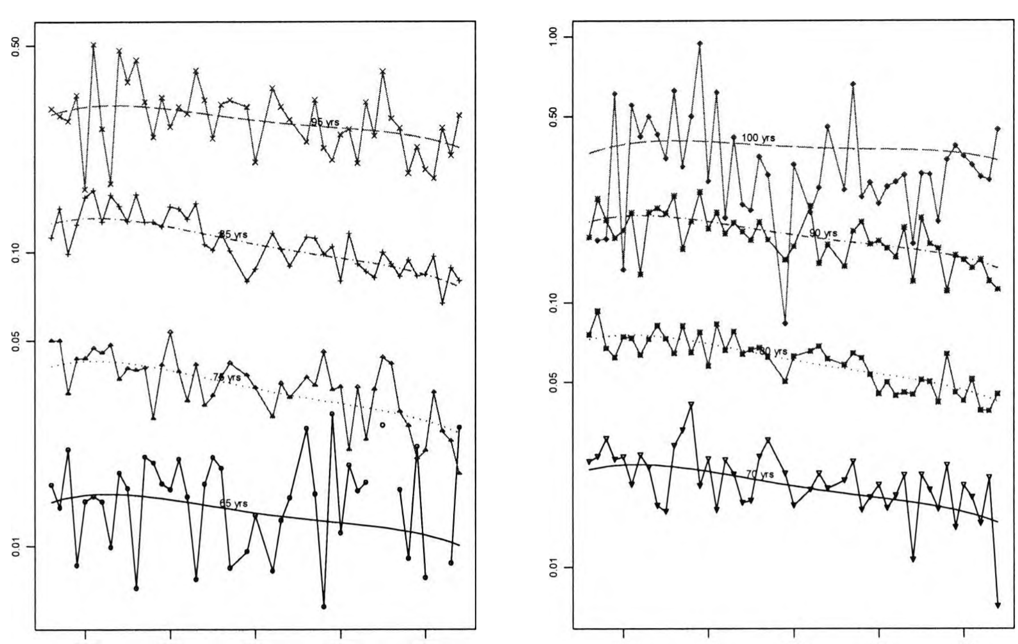


Figure 4.10 Female immediate annuitants, d1+ years, 1946-94 experience, ages 65-100: crude mortality rates and graduated forces of mortality plotted versus period on the log scale

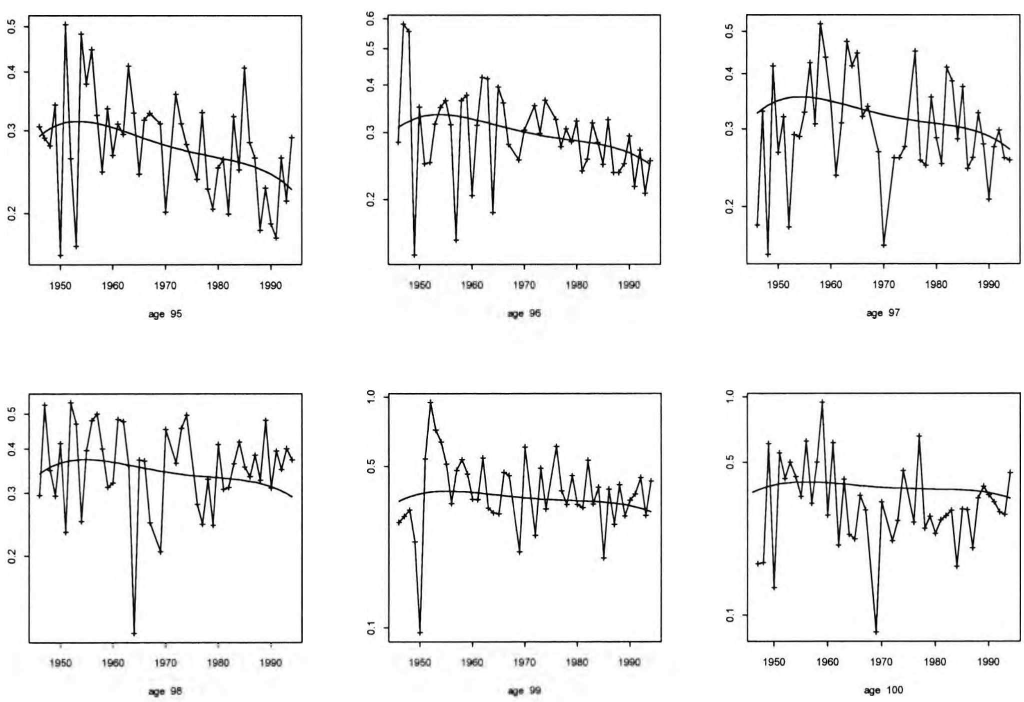


Figure 4.11 Female immediate annuitants, d1+ years, 1946-94 experience, crude mortality rates and graduated forces of mortality plotted versus period on the log scale for individual ages 95 to 100 years

4.2.4 Projection of female immediate annuitants' mortality at duration 1 year and over, based on the 1946-1994 mortality experience

By evaluating values of μ_{xt} at future time periods t, projections of the forces of mortality may be performed. Figure 4.12 shows predicted μ_{xt} values for female annuitants based on the 10-parameter model formula (4.15), for $t = 1946$ to 2014, at 5-year age intervals from age 55 to 110 years. The graduated rates are at ages 65 to 100 years in calendar years 1946 to 1994. The predicted forces of mortality are tabulated at quadrennial periods in Tables 4.9a, 4.9b and 4.9c.

It is clearly seen that the predicted forces of mortality for females do not have the desired shape. For the period considered, the μ_{xt} values increase with age x in a consistent manner when x lies in the range [59, 103] years. At ages below 59, the progression of the forces of mortality with age is consistent up to and including calendar year 1987. On the other hand, at ages above 103, the progression of μ_{xt} values is inconsistent in the earlier years where for example, the predicted force of mortality at age 110 in calendar year 1946 is lower than the corresponding rate at age 100.

The rapid rate of improvement in mortality over the projection period (1995 to 2014), apparent from the plot is unlikely to provide a satisfactory representation of the rate of improvement that might actually occur. Although it is important not to over estimate annuitant mortality, premium rates determined from a basis allowing for such a rapid improvement in mortality might prove to be too high and hence render the company using such rates uncompetitive. A limit lifetable controlling the asymptotic level of mortality could be introduced. However, in this study, an improved model has been identified by excluding the observed data in calendar years prior to 1958 from the analysis. The results are presented in Section 4.2.5.

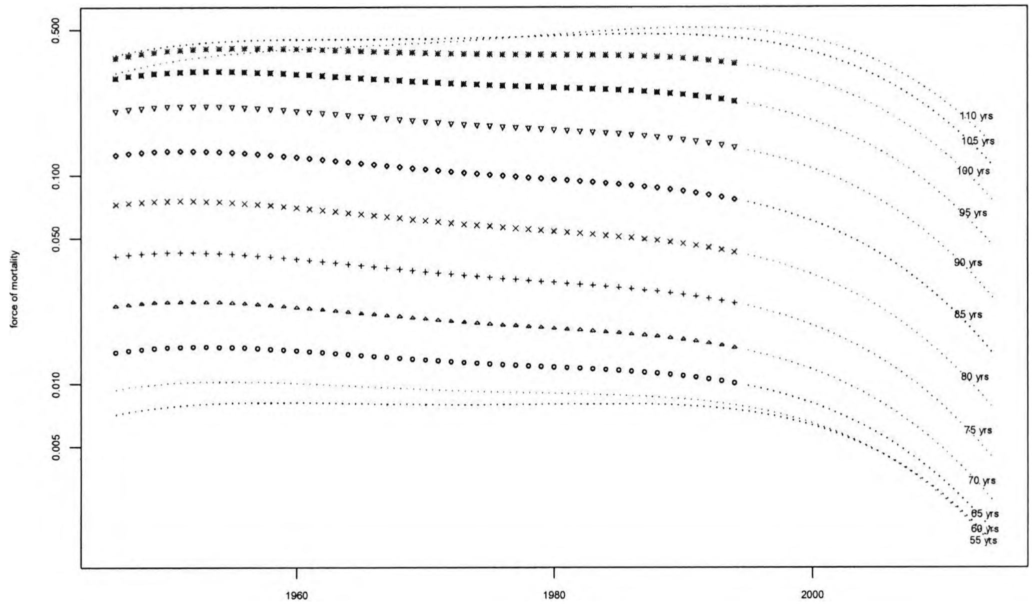


Figure 4.12 Female immediate annuitants, d1+ years, predicted forces of mortality based on a 10-parameter log-link model applied to the 1946-94 mortality experience

Table 4.9a Female Immediate Annuitants, Duration 1 year and over
Analysis of the 1946-1994 experience, predicted force of mortality at quadrennial periods 1946-66

age	1946	1950	1954	1958	1962	1966
55	0.007110	0.007749	0.008058	0.008136	0.008092	0.008010
56	0.007417	0.008051	0.008337	0.008385	0.008305	0.008188
57	0.007789	0.008422	0.008688	0.008703	0.008587	0.008434
58	0.008232	0.008868	0.009115	0.009097	0.008942	0.008750
59	0.008753	0.009397	0.009623	0.009571	0.009375	0.009141
60	0.009361	0.010015	0.010223	0.010133	0.009892	0.009613
61	0.010065	0.010735	0.010922	0.010791	0.010502	0.010173
62	0.010877	0.011566	0.011733	0.011558	0.011214	0.010830
63	0.011811	0.012524	0.012668	0.012444	0.012039	0.011594
64	0.012883	0.013623	0.013743	0.013464	0.012992	0.012478
65	0.014109	0.014883	0.014977	0.014636	0.014088	0.013497
66	0.015510	0.016323	0.016389	0.015979	0.015345	0.014668
67	0.017110	0.017968	0.018002	0.017514	0.016783	0.016009
68	0.018934	0.019845	0.019843	0.019268	0.018428	0.017543
69	0.021011	0.021983	0.021942	0.021268	0.020304	0.019294
70	0.023374	0.024416	0.024331	0.023546	0.022443	0.021293
71	0.026060	0.027182	0.027049	0.026138	0.024878	0.023569
72	0.029107	0.030322	0.030136	0.029085	0.027647	0.026160
73	0.032560	0.033883	0.033638	0.032429	0.030793	0.029105
74	0.036466	0.037913	0.037604	0.036220	0.034361	0.032448
75	0.040876	0.042466	0.042090	0.040511	0.038403	0.036238
76	0.045844	0.047601	0.047152	0.045358	0.042974	0.040528
77	0.051427	0.053377	0.052853	0.050822	0.048133	0.045376
78	0.057683	0.059857	0.059258	0.056968	0.053942	0.050842
79	0.064672	0.067107	0.066433	0.063864	0.060469	0.056991
80	0.072453	0.075192	0.074446	0.071577	0.067782	0.063892
81	0.081082	0.084174	0.083365	0.080177	0.075950	0.071613
82	0.090611	0.094112	0.093253	0.089731	0.085041	0.080225
83	0.101086	0.105060	0.104170	0.100302	0.095123	0.089795
84	0.112539	0.117062	0.116167	0.111948	0.106256	0.100388
85	0.124992	0.130147	0.129283	0.124713	0.118492	0.112062
86	0.138447	0.144329	0.143541	0.138631	0.131872	0.124864
87	0.152888	0.159599	0.158944	0.153717	0.146422	0.138829
88	0.168269	0.175926	0.175473	0.169962	0.162145	0.153974
89	0.184517	0.193244	0.193076	0.187333	0.179023	0.170292
90	0.201526	0.211455	0.211671	0.205762	0.197005	0.187751
91	0.219152	0.230424	0.231135	0.225146	0.216009	0.206286
92	0.237214	0.249972	0.251304	0.245341	0.235910	0.225796
93	0.255489	0.269880	0.271973	0.266159	0.256546	0.246140
94	0.273717	0.289883	0.292887	0.287368	0.277706	0.267131
95	0.291601	0.309677	0.313750	0.308688	0.299133	0.288537
96	0.308810	0.328916	0.334221	0.329796	0.320526	0.310081
97	0.324988	0.347226	0.353925	0.350326	0.341540	0.331438
98	0.339765	0.364207	0.372455	0.369880	0.361789	0.352244
99	0.352762	0.379449	0.389387	0.388034	0.380862	0.372097
100	0.363611	0.392543	0.404289	0.404352	0.398322	0.390572
101	0.371966	0.403095	0.416741	0.418395	0.413728	0.407226
102	0.377520	0.410746	0.426345	0.429746	0.426648	0.421619
103	0.380019	0.415187	0.432751	0.438020	0.436674	0.433325
104	0.379279	0.416178	0.435667	0.442886	0.443443	0.441952
105	0.375198	0.413559	0.434881	0.444084	0.446650	0.447160
106	0.367763	0.407266	0.430273	0.441441	0.446075	0.448679
107	0.357059	0.397338	0.421828	0.434883	0.441588	0.446328
108	0.343271	0.383921	0.409640	0.424448	0.433166	0.440025
109	0.326676	0.367269	0.393917	0.410288	0.420900	0.429796
110	0.307639	0.347731	0.374975	0.392665	0.404995	0.415786

Table 4.9b Female Immediate Annuitants, Duration 1 year and over
Analysis of 1946-1994 experience, predicted force of mortality at quadrennial periods 1970-1990

age	1970	1974	1978	1982	1986	1990
55	0.007948	0.007932	0.007959	0.008000	0.007995	0.007856
56	0.008092	0.008043	0.008038	0.008046	0.008008	0.007838
57	0.008302	0.008220	0.008183	0.008159	0.008090	0.007887
58	0.008582	0.008465	0.008396	0.008341	0.008239	0.008002
59	0.008934	0.008781	0.008678	0.008591	0.008456	0.008185
60	0.009363	0.009172	0.009035	0.008914	0.008745	0.008436
61	0.009877	0.009645	0.009470	0.009314	0.009108	0.008758
62	0.010483	0.010206	0.009991	0.009797	0.009551	0.009157
63	0.011192	0.010865	0.010606	0.010370	0.010081	0.009638
64	0.012013	0.011632	0.011324	0.011043	0.010707	0.010209
65	0.012962	0.012519	0.012158	0.011827	0.011439	0.010879
66	0.014053	0.013542	0.013121	0.012734	0.012288	0.011660
67	0.015306	0.014717	0.014229	0.013780	0.013269	0.012564
68	0.016739	0.016064	0.015501	0.014982	0.014398	0.013607
69	0.018378	0.017605	0.016958	0.016361	0.015695	0.014806
70	0.020249	0.019366	0.018624	0.017940	0.017182	0.016182
71	0.022381	0.021375	0.020526	0.019744	0.018883	0.017759
72	0.024810	0.023665	0.022697	0.021804	0.020826	0.019562
73	0.027573	0.026272	0.025169	0.024153	0.023046	0.021623
74	0.030712	0.029236	0.027984	0.026830	0.025576	0.023975
75	0.034275	0.032603	0.031184	0.029876	0.028458	0.026658
76	0.038311	0.036422	0.034817	0.033338	0.031738	0.029713
77	0.042877	0.040747	0.038936	0.037268	0.035466	0.033190
78	0.048031	0.045636	0.043599	0.041721	0.039696	0.037141
79	0.053839	0.051152	0.048867	0.046761	0.044489	0.041624
80	0.060366	0.057362	0.054806	0.052452	0.049910	0.046703
81	0.067683	0.064334	0.061487	0.058864	0.056030	0.052445
82	0.075859	0.072141	0.068982	0.066071	0.062920	0.058923
83	0.084964	0.080854	0.077364	0.074149	0.070660	0.066215
84	0.095067	0.090543	0.086708	0.083174	0.079326	0.074399
85	0.106229	0.101277	0.097085	0.093222	0.089000	0.083556
86	0.118506	0.113115	0.108562	0.104366	0.099757	0.093766
87	0.131939	0.126108	0.121198	0.116672	0.111671	0.105108
88	0.146557	0.140295	0.135040	0.130197	0.124808	0.117653
89	0.162366	0.155696	0.150120	0.144984	0.139221	0.131464
90	0.179351	0.172307	0.166449	0.161058	0.154948	0.146591
91	0.197463	0.190099	0.184015	0.178423	0.172007	0.163066
92	0.216622	0.209010	0.202774	0.197051	0.190391	0.180898
93	0.236709	0.228942	0.222647	0.216885	0.210060	0.200068
94	0.257560	0.249754	0.243515	0.237827	0.230939	0.220523
95	0.278968	0.271261	0.265216	0.259736	0.252911	0.242171
96	0.300678	0.293230	0.287538	0.282425	0.275811	0.264875
97	0.322389	0.315381	0.310222	0.305654	0.299426	0.288449
98	0.343753	0.337387	0.332959	0.329136	0.323490	0.312655
99	0.364386	0.358877	0.355394	0.352530	0.347683	0.337202
100	0.383870	0.379442	0.377127	0.375450	0.371635	0.361745
101	0.401766	0.398647	0.397726	0.397468	0.394931	0.385886
102	0.417625	0.416037	0.416733	0.418124	0.417113	0.409187
103	0.431008	0.431158	0.433679	0.436939	0.437698	0.431170
104	0.441497	0.443568	0.448098	0.453427	0.456187	0.451334
105	0.448718	0.452859	0.459550	0.467116	0.472082	0.469170
106	0.452356	0.458673	0.467634	0.477564	0.484906	0.484176
107	0.452176	0.460723	0.472011	0.484380	0.494222	0.495880
108	0.448039	0.458810	0.472422	0.487247	0.499654	0.503860
109	0.439908	0.452835	0.468704	0.485936	0.500910	0.507763
110	0.427865	0.442814	0.460805	0.480324	0.497797	0.507330

Table 4.9c Female Immediate Annuitants, Duration 1 year and over
Analysis of 1946-1994 experience, predicted force of mortality at quadrennial periods 1994-2014

age	1994	1998	2002	2006	2010	2014
55	0.007482	0.006781	0.005723	0.004384	0.002962	0.001708
56	0.007434	0.006711	0.005640	0.004303	0.002896	0.001663
57	0.007452	0.006700	0.005610	0.004264	0.002858	0.001635
58	0.007533	0.006748	0.005629	0.004262	0.002846	0.001622
59	0.007678	0.006854	0.005697	0.004298	0.002860	0.001625
60	0.007886	0.007016	0.005812	0.004371	0.002899	0.001641
61	0.008162	0.007238	0.005977	0.004481	0.002962	0.001671
62	0.008508	0.007522	0.006193	0.004629	0.003051	0.001716
63	0.008929	0.007873	0.006463	0.004817	0.003166	0.001776
64	0.009433	0.008295	0.006792	0.005048	0.003309	0.001851
65	0.010028	0.008796	0.007184	0.005326	0.003483	0.001944
66	0.010722	0.009384	0.007646	0.005656	0.003690	0.002055
67	0.011529	0.010068	0.008187	0.006043	0.003934	0.002186
68	0.012462	0.010861	0.008814	0.006493	0.004219	0.002339
69	0.013536	0.011776	0.009540	0.007015	0.004550	0.002518
70	0.014770	0.012829	0.010376	0.007618	0.004933	0.002726
71	0.016186	0.014039	0.011338	0.008313	0.005374	0.002966
72	0.017807	0.015426	0.012442	0.009111	0.005883	0.003243
73	0.019662	0.017014	0.013708	0.010027	0.006468	0.003561
74	0.021781	0.018830	0.015158	0.011077	0.007139	0.003927
75	0.024200	0.020906	0.016817	0.012280	0.007908	0.004347
76	0.026959	0.023276	0.018713	0.013657	0.008790	0.004829
77	0.030101	0.025980	0.020879	0.015232	0.009799	0.005381
78	0.033677	0.029060	0.023349	0.017030	0.010954	0.006014
79	0.037741	0.032566	0.026164	0.019083	0.012274	0.006739
80	0.042352	0.036549	0.029369	0.021423	0.013781	0.007567
81	0.047574	0.041068	0.033011	0.024087	0.015500	0.008514
82	0.053477	0.046186	0.037143	0.027116	0.017457	0.009593
83	0.060134	0.051971	0.041822	0.030552	0.019682	0.010823
84	0.067623	0.058492	0.047109	0.034443	0.022208	0.012222
85	0.076022	0.065824	0.053068	0.038839	0.025067	0.013810
86	0.085414	0.074043	0.059765	0.043793	0.028298	0.015608
87	0.095875	0.083225	0.067268	0.049358	0.031937	0.017640
88	0.107484	0.093445	0.075645	0.055589	0.036025	0.019928
89	0.120307	0.104772	0.084960	0.062541	0.040599	0.022497
90	0.134403	0.117269	0.095273	0.070265	0.045700	0.025371
91	0.149816	0.130987	0.106638	0.078809	0.051362	0.028573
92	0.166571	0.145962	0.119095	0.088212	0.057619	0.032126
93	0.184668	0.162210	0.132671	0.098505	0.064497	0.036048
94	0.204075	0.179721	0.147373	0.109705	0.072016	0.040354
95	0.224727	0.198456	0.163185	0.121811	0.080184	0.045055
96	0.246518	0.218338	0.180062	0.134803	0.088998	0.050154
97	0.269294	0.239253	0.197925	0.148638	0.098437	0.055646
98	0.292853	0.261040	0.216659	0.163241	0.108464	0.061516
99	0.316939	0.283489	0.236106	0.178510	0.119020	0.067737
100	0.341244	0.306339	0.256066	0.194306	0.130023	0.074269
101	0.365406	0.329282	0.276293	0.210454	0.141366	0.081056
102	0.389017	0.351957	0.296498	0.226745	0.152917	0.088029
103	0.411624	0.373962	0.316347	0.242933	0.164517	0.095101
104	0.432744	0.394857	0.335474	0.258740	0.175981	0.102170
105	0.451878	0.414177	0.353478	0.273857	0.187105	0.109119
106	0.468519	0.431446	0.369944	0.287959	0.197663	0.115817
107	0.482181	0.446188	0.384447	0.300705	0.207417	0.122124
108	0.492411	0.457953	0.396574	0.311755	0.216124	0.127892
109	0.498816	0.466331	0.405937	0.320782	0.223542	0.132973
110	0.501079	0.470974	0.412191	0.327481	0.229442	0.137218

The mortality improvement factor derived from the model (4.15), for an ultimate life attaining exact age x in calendar year t, is of the form of (3.44) with $r = 4$ and $s = 3$, that is:

$$RF(x,t-t_0)=\frac{\mu_{xt}}{\mu_{xt_0}}=\exp\left\{\sum_{i=1}^{4}\frac{1}{(w_t)^i}\left(\alpha_i+\sum_{j=1}^{3}\gamma_{ij}L_j(x')\right)\left((t-c_t)^i-(t_0-c_t)^i\right)\right\}; \tag{4.16}$$

where t_0 is the base calendar year. Denoting $t–t_0$ as n, $RF(x,n)=RF(x,t-t_0)$ is the reduction factor for an ultimate life attaining exact age x at time n, where n is measured in years from the base calendar year t_0.

Expression (4.16) is the same as:

$$RF(x,t-t_0)=\exp\left\{\sum_{i=1}^{4}\frac{\alpha_i}{(w_t)^i}\left[(t-c_t)^i-(t_0-c_t)^i\right]+\frac{(t-t_0)}{w_t}\sum_{j=1}^{2}\gamma_{1j}L_j(x')\right\}, \tag{4.17}$$

since only γ_{11} and γ_{12} are non-zero in equation (4.15). The parameters α_i (i = 1, 2, 3, 4) and γ_{1j} (j = 1, 2), are as given in Table 4.5. Noting that $L_1(x')=x'$ and $L_2(x')=\frac{5x'^2-1}{2}$, the improvement formula (4.17) may be written as:

$$\log RF(x,t-t_0)=\left\{\sum_{i=1}^{4}\frac{\alpha_i}{(w_t)^i}\left((t-c_t)^i-(t_0-c_t)^i\right)\right\}+\frac{(t-t_0)}{w_t}\left[\gamma_{11}x'+\gamma_{12}\left(\frac{5x'^2-1}{2}\right)\right] \tag{4.18}$$

From formulae (4.17) and (4.18), it is evident that the mortality improvement formula is essentially the trend adjustment term in the model formula, with adjustment for the transformation of the time variable t. The trend adjustment term (and hence the mortality improvement formula) involves higher order polynomials in both time t and age x, leading to undesirable features when extrapolating. Hence this 10-parameter model would not seem to be appropriate for mortality projections without major adjustments.

4.2.5 Analysis of the 1958-94 female annuitants' mortality experience at duration 1 year and over

In an effort to identify a model or models that would not only provide a good fit for the observed data but also provide reasonable projections of the forces of mortality at future time periods t, and for values of x outside the range of ages over which the model has been fitted, the observed mortality experience in calendar years prior to 1958 was excluded from the analysis of the female immediate annuitants' experience. The calendar year 1958 was chosen because of the apparent change in the class of lives taking out immediate annuity contracts as a result of the Finance Act 1956. The female annuitants' experience over the period 1958-94 was analysed over the same range of x values as for the 1946-94 experience, that is over the ages 65 to 100 years, giving a total of 1224 data cells.

An examination of the improvement in deviance as a result of increasing terms in r and s in the first instance, and then introducing mixed product terms, leads to the adoption of a 7-parameter model formula:

$$\mu_{xt} = \exp\left[\sum_{j=0}^{3}\beta_j L_j(x') + \left\{\alpha_1 + \sum_{j=1}^{2}\gamma_{1j}L_j(x')\right\}t'\right], \tag{4.19}$$

with,

$$x' = \frac{x-82.5}{17.5} \text{ as before;}$$

and

$$t' = \frac{t-1976}{18}, \text{ that is, } c_t = 1976 \text{ and } w_t = 18. \tag{4.20}$$

The unscaled deviance profile for successive increases in the values of r and s is shown in Table 4.10, while Table 4.11 shows the analysis of deviance for the model. It is clear from Table 4.10 that the optimum value of s is 3; increasing the value of s to 4 does not result in a significant improvement in deviance. As for the model based on the 1946-94 experience, the improvement in deviance for successive increases in the value of r shows some irregularities. When the value of r is increased from 1 to 2, the improvement is not significant, but subsequent increases of r up to and including r=6, result in improvements in deviance. However, the parameter α_2 remains non-significant. Therefore, on the basis of parsimony, and bearing in mind that higher order polynomials in the trend adjustment term tend to have features undesirable for projections, the value of r is chosen to be 1.

The fitted model (4.19) again consists of a $GM_x(0,4)$ term in age effects x, but the trend adjustment term is greatly simplified to a linear function of t on the log scale, while the coefficient of the trend adjustment term is itself a quadratic in x. Details of parameter estimates based on the 7-parameter model are given in Table 4.12.

Table 4.10
Female immediate annuitants, duration 1 year and over, analysis of the 1958-94 experience
Deviances for some polynomial predictors of degree *r* and *s*

	$r = 0$	$r = 1$	$r = 2$	$r = 3$
$s = 1$	3719.01	2459.17	2452.36	2425.50
$s = 2$	3630.68	2373.93	2367.48	2343.37
$s = 3$	3566.45	2308.54	2302.59	2276.96
$s = 4$	3565.18	2307.50	2301.58	2276.96

Table 4.11
Female Immediate Annuitants, Duration 1 year and over, Analysis of the 1958-1994 experience
Deviance profile (terms added sequentially 1st to last)

Parameter	Deviance	Degrees of freedom	Difference in Deviance
β_0	56046.00	1223	
β_1	3719.01	1222	52326.99
β_2	3630.68	1221	88.33
β_3	3566.45	1220	64.23
α_1	2308.54	1219	1257.92
γ_{11}	2292.80	1218	15.74
γ_{12}	2260.66	1217	32.14
γ_{13}	2259.75	1216	0.91

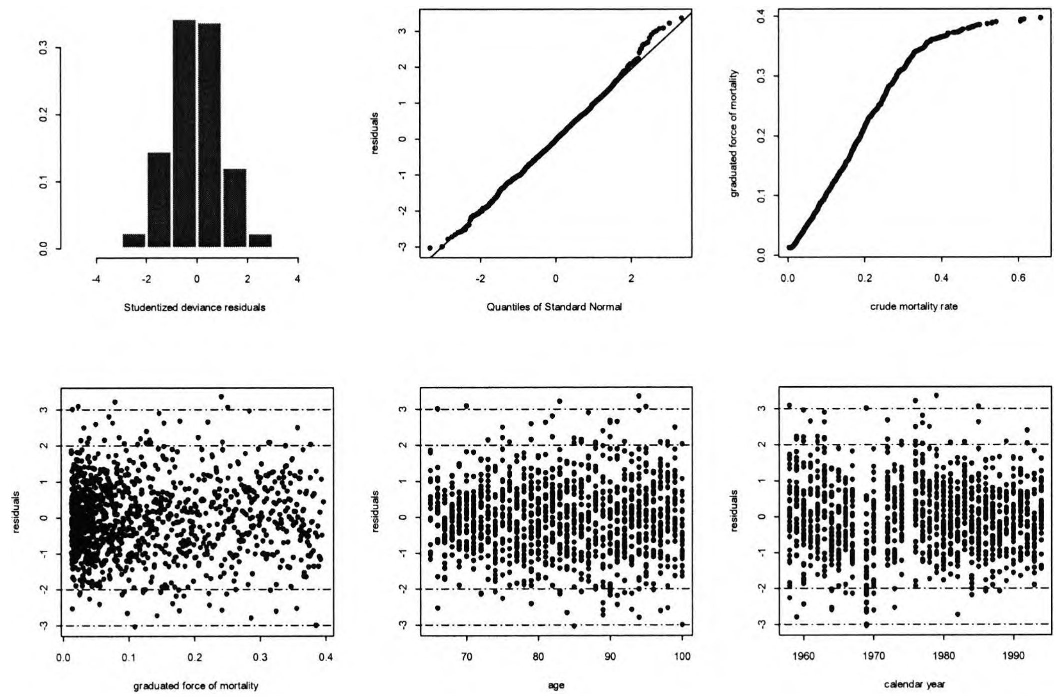


Figure 4.13 Female immediate annuitants, d1+ years, analysis of the 1958-94 experience, diagnostic plots: 7-parameter log-link model with $r = 1$, $s = 3$, γ_{11} and γ_{12}

Table 4.12
Female immediate annuitants, Duration 1 year and over, Analysis of the 1958-94 experience
7-parameter log-link model (ϕ = 1.8534)

parameter	Estimate	Standard error	t-value
β_0	-2.619459	0.0070	-376.8458
β_1	1.823799	0.0158	115.4447
β_2	-0.060129	0.0165	-3.6557
β_3	-0.117402	0.0171	-6.8481
α_1	-0.186520	0.0121	-15.4538
γ_{11}	0.022634	0.0230	0.9838
γ_{12}	0.116310	0.0279	4.1752

Although the data are supportive of the model as is evident from the diagnostic plots shown in Figure 4.13, some of the predicted forces of mortality outside the range of ages analysed are in fact increasing with time rather than decreasing. Figure 4.14 shows the predicted forces of mortality up to calendar year t = 2014, plotted against t on the log scale. At both ends of the age range, the predicted forces of mortality are such that there is a crossing over of forces of mortality. As an example, the predicted force of mortality for a life aged 55 in 2005 is higher than the predicted force of mortality for a life aged 66 in the same year. Hence it appears that this model would in general only be useful for predictions of future forces of mortality provided the predictions are made within the range of ages over which the model has been fitted.

From Table 4.12, it is observed that the 7-parameter model fitted includes one parameter that is not statistically significant (γ_{11}, with a t-value of 0.9838). The next step might then be to fit a model that excludes the parameter γ_{11}. However, on grounds of simplicity, the parameter γ_{12} that involves a term in x^2, was next excluded from the model formula. It turns out that when the parameter γ_{12} is excluded from the formula, all the remaining 6 parameters in the model (including γ_{11}) are statistically significant. Therefore the six-parameter model that excludes the quadratic coefficient in age effects from the trend adjustment term, was next fitted to the data. The revised model is:

$$\mu_{xt} = \exp\left[\beta_0 + \sum_{j=1}^{3} \beta_j L_j(x') + (\alpha_1 + \gamma_{11} L_1(x'))t'\right]. \tag{4.21}$$

Parameter estimates for the revised fit are given in Table 4.13.

The predicted forces of mortality based on each of the two models (4.19) and (4.21), together with the crude mortality rates, are plotted against time t on the log scale at 10-year age intervals, and shown in Figure 4.15. From the plot, it is clear that there is little difference between the predicted rates from the two models in the age range 75 to 95 years. However, outside this range of ages, the forces of mortality predicted from the two models follow differing patterns, with the 7-parameter model tending to follow the crude mortality rates more closely than the 6-parameter model.

Table 4.13
Female immediate annuitants, Duration 1 year and over,
6-parameter log-link model based on the 1958-94 experience (ϕ = 1.8846)

parameter	estimate	Standard error	t-value
β_0	-2.627386	0.0068	-384.9264
β_1	1.850637	0.0148	125.3246
β_2	-0.083299	0.0158	-5.2843
β_3	-0.099143	0.0167	-5.9264
α_1	-0.220583	0.0090	-24.4358
γ_{11}	0.061607	0.0213	2.8877

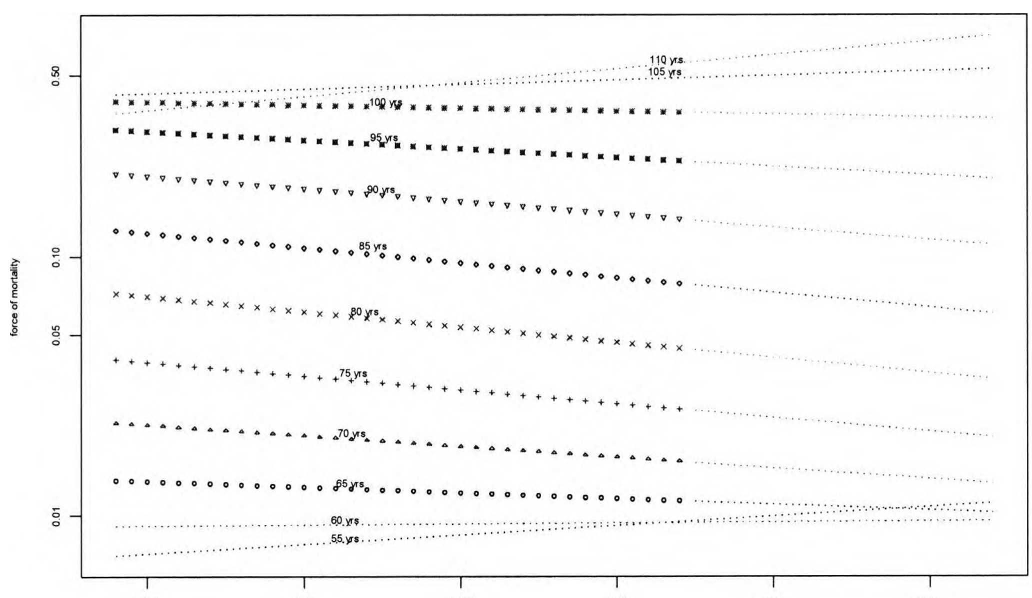


Figure 4.14 Female immediate annuitants, d1+ years, analysis of the 1958-94 experience, predicted forces of mortality based on a 7-parameter log-link model with $r = 1$, $s = 3$, γ_{11} and γ_{12}

It would be expected that the model involving more parameters would provide a better fit and this is obviously the case based on an analysis of deviance and a visual inspection of the plotted rates shown in Figure 4.15. However, from the chi-square goodness-of-fit test and other formal statistical tests of graduation (Tables 4.14 and 4.15), it is difficult to make any real distinction between the two models, with each model providing a satisfactory fit for the data overall. Therefore, all other factors being equal, the model with fewer parameters is preferred.

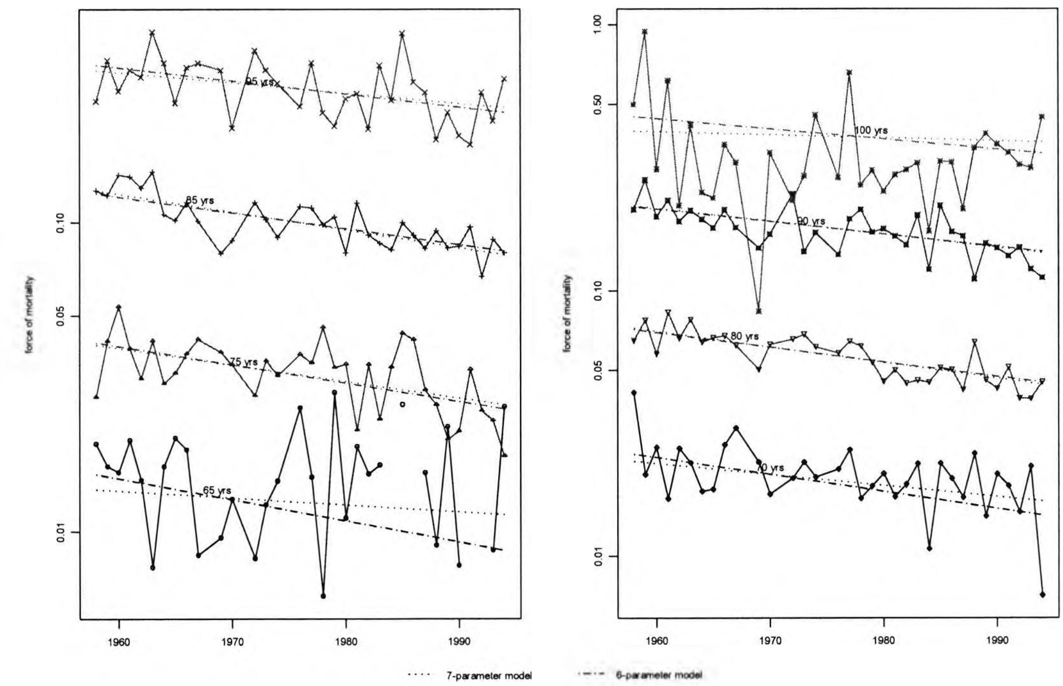


Figure 4.15 Female immediate annuitants, d1+ years, 1958-94 experience, crude and graduated forces of mortality plotted on the log scale; comparison of the 7-parameter and 6-parameter log-link models

Table 4.14
Female Immediate Annuitants, duration 1 year and over, Analysis of the 1958-94 experience
Comparison of p-values based on the 2 models fitted

Statistical test	6-parameter model	7-parameter model
Chi-square	0.4945	0.4945
Cumulative deviations	0.4290	0.4330
Individual standardised deviations	0.5521	0.5141
Grouping of signs of deviations	0.9785	0.9886
Signs of deviations	0.3965	0.3505

Table 4.15
Female Immediate Annuitants, d1 year and over, Analysis of the 1958-94 experience
Comparison of the distribution of individual studentized deviance residuals for the data as a whole

Range	**(-∞,-3)**	**(-3,-2)**	**(-2,-1)**	**(-1,0)**	**(0,1)**	**(1,2)**	**(2,3)**	**(3,∞)**
expected frequency	1.58	25.04	159.01	399.37	399.37	159.01	25.04	1.58
observed frequency								
6-parameter model	1.00	28.00	167.00	404.00	404.00	137.00	25.00	4.00
7-parameter model	2.00	28.00	171.00	403.00	397.00	142.00	27.00	5.00

Table 4.16
Female Immediate Annuitants, Duration 1 year and over, analysis of the 1958-94 experience
Tests of graduation applied to individual calendar years: *p*-values (6-parameter model)

year	chi-square	cumdev	isd	runs	signs
1958	0.03	0.44	0.99	0.91	0.87
1959	0.06	0.00	0.28	0.24	0.09
1960	0.03	0.71	0.08	0.41	0.41
1961	0.25	0.06	0.64	0.49	0.87
1962	0.25	0.00	0.13	0.75	0.87
1963	0.00	0.00	0.00	0.95	0.07
1964	0.23	0.98	0.03	0.20	0.01
1965	0.57	0.99	0.12	0.33	0.03
1966	0.86	0.23	0.05	0.91	0.62
1967	0.38	1.00	0.04	0.95	0.01
1969	0.00	1.00	0.00	0.05	0.00
1970	0.05	0.99	0.13	0.11	0.03
1972	0.10	0.06	0.63	0.02	1.00
1973	0.67	0.72	0.41	0.02	1.00
1974	0.60	0.16	0.29	0.18	0.03
1976	0.00	0.00	0.01	0.19	0.00
1977	0.03	0.00	0.13	0.45	0.03
1978	0.34	0.17	0.61	0.07	0.62
1979	0.06	0.04	0.36	0.42	0.04
1980	0.77	0.86	0.42	0.08	0.18
1981	0.76	0.21	0.25	0.37	0.50
1982	0.59	0.98	0.56	0.68	0.12
1983	0.40	0.13	0.06	0.47	0.12
1984	0.39	0.93	0.47	0.05	0.30
1985	0.05	0.38	0.70	0.29	0.73
1986	0.44	0.38	0.57	0.01	0.30
1987	0.60	0.97	0.61	0.90	0.60
1988	0.77	0.77	0.77	0.20	0.60
1989	0.83	0.87	0.23	0.92	1.00
1990	0.81	0.20	0.69	0.50	1.00
1991	0.80	0.12	0.57	0.39	0.22
1992	0.58	0.40	0.69	0.89	1.00
1993	0.28	1.00	0.02	0.30	0.00
1994	0.69	0.74	0.61	0.05	0.26

Based on the preferred 6-parameter model (4.21), various statistical tests of graduation are applied separately to each of the 34 calendar years. The p-values obtained are given in Table 4.16. Although in a number of years, the model fails some of the tests of graduation applied, the tests are generally quite supportive of the model. One of the problems of graduation using a mathematical formula is that no single curve can accurately fit a whole range of ages, more so when the covariates are both age and time period. Figure 4.16 is a graph of graduated forces of mortality and crude mortality rates, plotted on the log scale for the specific calendar years 1958 to 1992.

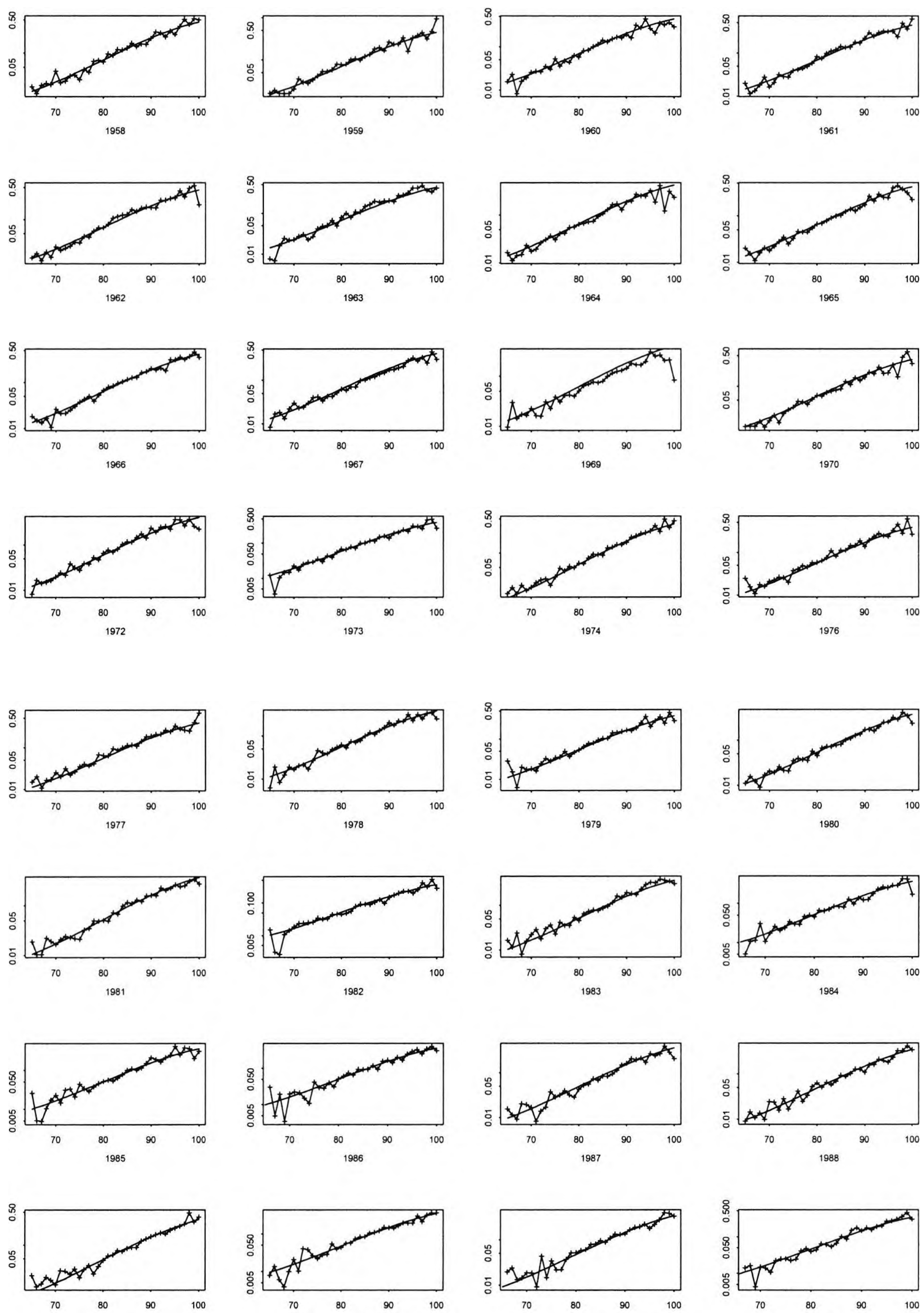


Figure 4.16 Female immediate annuitants, d1+ years, analysis of the 1958-94 experience, crude mortality rates and graduated forces of mortality plotted versus age on the log scale for individual calendar years 1958 to 1992: 6-parameter model

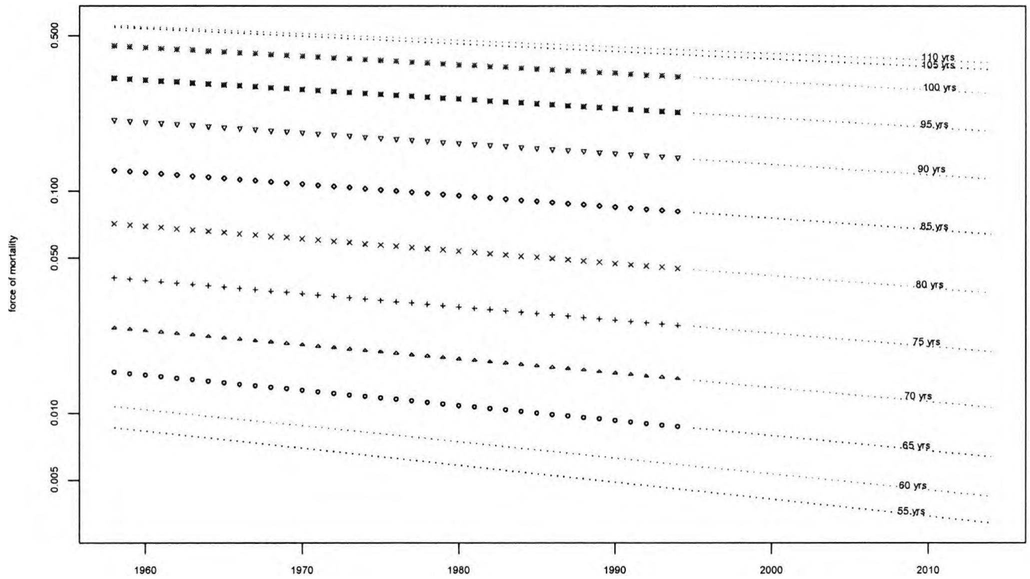


Figure 4.17 Female immediate annuitants, d1+ years, analysis of the 1958-94 experience, predicted forces of mortality plotted on the log scale; 6-parameter log-link model with $r = 1$, $s = 3$ and γ_{11}

Apart from the goodness-of-fit of the model, a further consideration in choosing an appropriate model is whether the shape of the predicted rates outside the main range of the data is sensible. Figure 4.17 is a plot of the predicted forces of mortality based on the 6-parameter model given by equation (4.21), plotted at 5-year age intervals from age 55 to 110 years. The predicted rates are also shown at quadrennial periods in Tables 4.17a, 4.17b and 4.17c. As for the projections based on an analysis of the 1946 to 1994 experience, the force of mortality was projected to calendar year 2014.

The revised model appears to provide reasonable predictions within and outside the range of ages over which the model was fitted. The predicted forces of mortality progress smoothly with respect to both age and time, and it can be seen that the model naturally predicts a reduction in the rate of improvement in mortality at the older ages. Therefore the 6-parameter model described by (4.21) would be preferred since this model seems to be appropriate for making predictions of future forces of mortality for female immediate annuitants, although the 7-parameter model defined by (4.19) provides a marginally better fit to the data.

Table 4.17a Female Immediate Annuitants, Duration 1 year and over
Analysis of 1958-94 experience, predicted force of mortality at quadrennial periods 1958-74

age	1958	1962	1966	1970	1974
55	0.008592	0.008007	0.007462	0.006953	0.006480
56	0.008868	0.008270	0.007713	0.007194	0.006709
57	0.009213	0.008599	0.008026	0.007491	0.006992
58	0.009631	0.008996	0.008403	0.007849	0.007332
59	0.010129	0.009468	0.008851	0.008274	0.007735
60	0.010713	0.010023	0.009377	0.008772	0.008207
61	0.011393	0.010667	0.009987	0.009351	0.008755
62	0.012178	0.011411	0.010692	0.010019	0.009388
63	0.013081	0.012267	0.011503	0.010787	0.010116
64	0.014116	0.013248	0.012433	0.011668	0.010950
65	0.015299	0.014369	0.013495	0.012675	0.011905
66	0.016647	0.015648	0.014708	0.013825	0.012995
67	0.018183	0.017105	0.016090	0.015136	0.014238
68	0.019931	0.018763	0.017664	0.016629	0.015655
69	0.021916	0.020648	0.019454	0.018329	0.017268
70	0.024169	0.022789	0.021488	0.020261	0.019104
71	0.026725	0.025219	0.023797	0.022456	0.021190
72	0.029622	0.027974	0.026418	0.024948	0.023561
73	0.032901	0.031095	0.029389	0.027776	0.026251
74	0.036610	0.034628	0.032753	0.030980	0.029302
75	0.040801	0.038622	0.036559	0.034607	0.032758
76	0.045528	0.043130	0.040859	0.038707	0.036668
77	0.050853	0.048213	0.045709	0.043336	0.041086
78	0.056842	0.053933	0.051172	0.048553	0.046068
79	0.063563	0.060357	0.057313	0.054422	0.051677
80	0.071090	0.067558	0.064200	0.061010	0.057978
81	0.079499	0.075607	0.071906	0.068386	0.065039
82	0.088867	0.084583	0.080505	0.076624	0.072930
83	0.099271	0.094560	0.090071	0.085796	0.081724
84	0.110788	0.105612	0.100678	0.095975	0.091491
85	0.123489	0.117812	0.112396	0.107228	0.102299
86	0.137438	0.131222	0.125287	0.119621	0.114211
87	0.152689	0.145898	0.139408	0.133208	0.127283
88	0.169283	0.161880	0.154801	0.148031	0.141558
89	0.187241	0.179193	0.171491	0.164120	0.157066
90	0.206563	0.197839	0.189483	0.181481	0.173816
91	0.227219	0.217793	0.208758	0.200098	0.191797
92	0.249148	0.238999	0.229264	0.219925	0.210967
93	0.272252	0.261367	0.250917	0.240884	0.231253
94	0.296392	0.284764	0.273592	0.262859	0.252546
95	0.321381	0.309015	0.297124	0.285691	0.274697
96	0.346988	0.333897	0.321300	0.309178	0.297514
97	0.372929	0.359141	0.345862	0.333074	0.320759
98	0.398875	0.384428	0.370504	0.357084	0.344150
99	0.424449	0.409395	0.394876	0.380871	0.367363
100	0.449232	0.433639	0.418586	0.404057	0.390032
101	0.472773	0.456720	0.441211	0.426230	0.411757
102	0.494596	0.478176	0.462301	0.446953	0.432114
103	0.514215	0.497533	0.481391	0.465773	0.450662
104	0.531146	0.514316	0.498020	0.482240	0.466960
105	0.544926	0.528073	0.511741	0.495913	0.480576
106	0.555131	0.538383	0.522140	0.506388	0.491110
107	0.561393	0.544882	0.528857	0.513303	0.498206
108	0.563419	0.547277	0.531597	0.516366	0.501572
109	0.561008	0.545361	0.530150	0.515364	0.500990
110	0.554062	0.539030	0.524406	0.510179	0.496338

Table 4.17b Female Immediate Annuitants, Duration 1 year and over
Analysis of the 1958-94 experience, predicted force of mortality at quadrennial periods 1978-94

age	1978	1982	1986	1990	1994
55	0.006039	0.005627	0.005244	0.004887	0.004554
56	0.006257	0.005835	0.005442	0.005076	0.004734
57	0.006526	0.006091	0.005685	0.005306	0.004952
58	0.006849	0.006397	0.005976	0.005582	0.005214
59	0.007231	0.006760	0.006319	0.005907	0.005522
60	0.007678	0.007183	0.006720	0.006287	0.005882
61	0.008197	0.007675	0.007186	0.006728	0.006299
62	0.008797	0.008243	0.007723	0.007237	0.006781
63	0.009486	0.008895	0.008342	0.007822	0.007335
64	0.010276	0.009644	0.009051	0.008494	0.007972
65	0.011181	0.010501	0.009863	0.009264	0.008701
66	0.012215	0.011481	0.010792	0.010144	0.009535
67	0.013394	0.012599	0.011852	0.011149	0.010488
68	0.014738	0.013875	0.013062	0.012297	0.011577
69	0.016270	0.015329	0.014442	0.013607	0.012820
70	0.018013	0.016984	0.016015	0.015100	0.014238
71	0.019996	0.018869	0.017805	0.016802	0.015855
72	0.022250	0.021012	0.019844	0.018740	0.017697
73	0.024810	0.023449	0.022162	0.020945	0.019796
74	0.027716	0.026215	0.024795	0.023453	0.022183
75	0.031009	0.029353	0.027785	0.026301	0.024897
76	0.034737	0.032908	0.031175	0.029533	0.027977
77	0.038953	0.036930	0.035012	0.033194	0.031471
78	0.043710	0.041473	0.039350	0.037336	0.035425
79	0.049070	0.046595	0.044245	0.042013	0.039894
80	0.055096	0.052358	0.049756	0.047283	0.044934
81	0.061855	0.058827	0.055947	0.053208	0.050603
82	0.069414	0.066068	0.062883	0.059851	0.056966
83	0.077845	0.074150	0.070631	0.067278	0.064085
84	0.087217	0.083142	0.079258	0.075555	0.072025
85	0.097596	0.093109	0.088829	0.084745	0.080849
86	0.109046	0.104114	0.099406	0.094910	0.090617
87	0.121621	0.116212	0.111043	0.106104	0.101385
88	0.135367	0.129448	0.123787	0.118374	0.113197
89	0.150315	0.143854	0.137671	0.131753	0.126090
90	0.166475	0.159445	0.152711	0.146261	0.140084
91	0.183841	0.176214	0.168904	0.161897	0.155181
92	0.202374	0.194130	0.186223	0.178637	0.171360
93	0.222007	0.213131	0.204609	0.196428	0.188574
94	0.242639	0.233120	0.223974	0.215187	0.206745
95	0.264127	0.253964	0.244191	0.234795	0.225760
96	0.286290	0.275489	0.265095	0.255094	0.245470
97	0.308899	0.297478	0.286479	0.275887	0.265686
98	0.331685	0.319671	0.308093	0.296934	0.286179
99	0.354334	0.341767	0.329646	0.317955	0.306678
100	0.376493	0.363425	0.350810	0.338633	0.326878
101	0.397775	0.384269	0.371221	0.358615	0.346438
102	0.417768	0.403898	0.390489	0.377525	0.364991
103	0.436041	0.421895	0.408207	0.394964	0.382150
104	0.452164	0.437836	0.423963	0.410530	0.397522
105	0.465713	0.451309	0.437351	0.423825	0.410717
106	0.476293	0.461924	0.447988	0.434472	0.421364
107	0.483554	0.469332	0.455529	0.442132	0.429128
108	0.487201	0.473242	0.459683	0.446513	0.433720
109	0.487017	0.473434	0.460229	0.447393	0.434915
110	0.482872	0.469772	0.457027	0.444628	0.432565

Table 4.17c Female Immediate Annuitants, Duration 1 year and over
Analysis of the 1958-94 experience, predicted force of mortality at quadrennial periods 1998-2014

age	1998	2002	2006	2010	2014
55	0.004244	0.003955	0.003686	0.003435	0.003201
56	0.004415	0.004117	0.003840	0.003581	0.003340
57	0.004622	0.004314	0.004027	0.003758	0.003508
58	0.004870	0.004549	0.004249	0.003969	0.003708
59	0.005162	0.004826	0.004511	0.004217	0.003942
60	0.005503	0.005148	0.004816	0.004506	0.004216
61	0.005898	0.005522	0.005170	0.004841	0.004532
62	0.006354	0.005954	0.005579	0.005227	0.004898
63	0.006879	0.006451	0.006049	0.005672	0.005319
64	0.007481	0.007021	0.006589	0.006184	0.005803
65	0.008172	0.007675	0.007208	0.006770	0.006359
66	0.008962	0.008424	0.007918	0.007443	0.006996
67	0.009866	0.009281	0.008730	0.008212	0.007725
68	0.010899	0.010260	0.009659	0.009093	0.008561
69	0.012078	0.011380	0.010721	0.010101	0.009517
70	0.013425	0.012658	0.011935	0.011254	0.010611
71	0.014961	0.014118	0.013322	0.012571	0.011863
72	0.016713	0.015783	0.014905	0.014076	0.013293
73	0.018709	0.017682	0.016712	0.015794	0.014928
74	0.020982	0.019846	0.018771	0.017755	0.016793
75	0.023567	0.022308	0.021117	0.019989	0.018922
76	0.026504	0.025108	0.023786	0.022533	0.021346
77	0.029837	0.028287	0.026819	0.025426	0.024106
78	0.033612	0.031892	0.030260	0.028711	0.027241
79	0.037882	0.035971	0.034157	0.032434	0.030798
80	0.042701	0.040578	0.038562	0.036645	0.034824
81	0.048126	0.045770	0.043530	0.041399	0.039372
82	0.054220	0.051606	0.049118	0.046750	0.044496
83	0.061043	0.058146	0.055386	0.052757	0.050253
84	0.068660	0.065453	0.062395	0.059480	0.056701
85	0.077132	0.073586	0.070204	0.066976	0.063897
86	0.086519	0.082606	0.078870	0.075303	0.071898
87	0.096875	0.092566	0.088449	0.084515	0.080756
88	0.108247	0.103513	0.098986	0.094658	0.090518
89	0.120670	0.115484	0.110520	0.105770	0.101223
90	0.134168	0.128502	0.123075	0.117877	0.112898
91	0.148744	0.142573	0.136659	0.130990	0.125556
92	0.164380	0.157685	0.151262	0.145100	0.139190
93	0.181035	0.173796	0.166848	0.160177	0.153772
94	0.198634	0.190841	0.183354	0.176161	0.169250
95	0.217073	0.208720	0.200689	0.192966	0.185541
96	0.236209	0.227298	0.218722	0.210471	0.202530
97	0.255863	0.246402	0.237292	0.228518	0.220069
98	0.275813	0.265823	0.256195	0.246915	0.237972
99	0.295802	0.285311	0.275192	0.265432	0.256018
100	0.315532	0.304580	0.294007	0.283802	0.273951
101	0.334675	0.323311	0.312333	0.301727	0.291482
102	0.352873	0.341158	0.329832	0.318881	0.308295
103	0.369752	0.357756	0.346149	0.334919	0.324053
104	0.384926	0.372729	0.360919	0.349483	0.338409
105	0.398014	0.385704	0.373775	0.362215	0.351012
106	0.408652	0.396323	0.384366	0.372770	0.361524
107	0.416507	0.404258	0.392368	0.380828	0.369628
108	0.421294	0.409223	0.397498	0.386110	0.375047
109	0.422785	0.410993	0.399530	0.388387	0.377555
110	0.420830	0.409413	0.398305	0.387499	0.376986

The trend adjustment term in the 6-parameter model (4.21) is a simple linear function of t on the log scale, with the coefficient of t being itself linear in x. Thus the formula for the reduction factor is simplified to:

$$RF(x,n) = \exp\left[\frac{n}{w_t}\{\alpha_1 + \gamma_{11}x'\}\right]; \qquad (4.22)$$

which is the same as

$$RF(x,n) = \exp\left[\frac{n}{w_t}\left\{\alpha_1 + \gamma_{11}\left(\frac{x - c_x}{w_x}\right)\right\}\right], \qquad (4.23)$$

where c_x=82.5, w_x=17.5, w_t=18 as defined above and n is measured in years from the base calendar year. In particular, based on the data-set for female immediate annuitants analysed and the parameter estimates given in Table 4.13, the formula for $RF(x,n)$ is:

$$RF(x,n) = \exp\left[\frac{n}{18}\left\{-0.220583 + 0.061607\left(\frac{x - 82.5}{17.5}\right)\right\}\right] \qquad (4.24)$$

or

$$RF(x,n) = \exp[(-0.028390 + 0.000196x)n] \qquad (4.25)$$

As for the CMI mortality improvement formula described in Chapter 3, the form of the mortality improvement model defined by expression (4.25) assumes that the rate of mortality decreases by exponential decay. However, unlike the CMI model, no limiting value is assumed for the improvement factor based on the log-link model and indeed there is no pre-determined maximum age assumed. The rate of improvement depends on both age and time for all ages x and the form of the model is such that mortality is assumed to improve at all ages x unless x is greater than the (unlikely) age of 145. By using this model, the problem of determining a precise maximum age for policyholders, which is the subject of ongoing debate among actuaries, does not arise (see for example Thatcher, 1999; Tuljapurkar and Boe, 1998; Wachter and Finch, 1997).

In determining a suitable mortality trend model for female immediate annuitants based on the 1958-94 experience, various other models involving higher order terms in time and/or higher order terms in age in the mixed product trend adjustment term were also fitted to the data. Although these models provided marginally better fits to the data, the predicted rates in each case were unsuitable. Including mixed product terms involving higher order terms in x resulted in mortality rates that did not progress smoothly at the extreme ages. On the other hand, introducing higher order terms in time t resulted in an unrealistically rapid improvement in mortality. It is apparent that in searching for a model that has a good shape for the purpose of making predictions, there has to be a trade-off between goodness-of-fit and predictive shape.

4.2.6 Analysis of the 1946-94 male annuitants' mortality experience at duration 1 year and over

The male immediate annuitants' mortality experience at curtate policy duration 1 year and over was analysed at individual ages x ranging from 65 to 95 years. An initial analysis of the experience at duration 5+ years revealed an inconsistency at age 94 in calendar year 1970, where there were 96 recorded deaths corresponding to a central exposed-to-risk of 52.5, resulting in a crude mortality rate of 1.8. This could be due to one individual (or several individuals) with a disproportionate number of policies, a factor that might not be adequately accounted for by the dispersion parameter, thereby resulting in a distortion of the results. Noting that the bulk of the experience with policy duration 1 year and over is comprised of the experience with policy duration 5 years and over, the data cell corresponding to this observation was excluded from the analyses. Hence the dataset analysed comprised of 1425 data cells with transformed age x' in the form:

$$x' = \frac{x-80}{15}, \text{ that is } c_x = 80 \text{ and } w_x = 15. \tag{4.26}$$

The transformed calendar year t' is given by (4.13), that is:

$c_t = 1970$, $w_t = 24$, giving $t' = \dfrac{t-1970}{24}$.

On the basis of the analyses of deviance shown in Tables 4.18 and 4.19, and the significance of the parameters introduced, the best fitting model was determined to be:

$$\mu_{xt} = \exp\left[\sum_{j=0}^{2}\beta_j L_j(x') + \sum_{i=1}^{2}\alpha_i t'^i + \gamma_{21} L_1(x') t'^2\right], \tag{4.27}$$

where $L_0(x') = 1$, $L_1(x') = x'$ and $L_2(x') = \dfrac{3x'^2 - 1}{2}$. The model adopted, which may be expressed as:

$$\mu_{xt} = \exp\left[\sum_{j=0}^{2}\beta_j L_j(x') + \alpha_1 t' + (\alpha_2 + \gamma_{21} x') t'^2\right], \tag{4.28}$$

consists of a $GM_x(0,3)$ term in age effects, and a trend adjustment term that is quadratic in time t on the log scale, with coefficients that are linear in age x.

The parameter estimates, their standard errors and the associated t-statistics are shown in Table 4.20. The values of the t-statistics associated with each parameter estimate are almost all greater than 2, indicating statistical significance of the parameters. The exception is the estimate of γ_{21}, the coefficient of the mixed product term in age and time, which has an associated t-statistic value just under 2. Statistical tests of graduation carried out on the data as a whole also indicate that the model fits the data reasonably well. The p-values for the various graduation tests applied are shown in Table 4.21.

Table 4.18
Male immediate annuitants, duration 1 year and over, analysis of the 1946-94 experience
Deviances for some polynomial predictors of degree *r* and *s*

	$r = 0$	$r = 1$	$r = 2$	$r = 3$
$s = 0$	20613.47	20498.29	20483.38	20411.93
$s = 1$	3530.96	2647.24	2642.49	2642.45
$s = 2$	3520.49	2639.41	2634.27	2634.25
$s = 3$	3520.23	2638.91	2633.76	2633.73

Table 4.19
Male Immediate Annuitants, Duration 1 year and over, Analysis of the 1946-1994 experience
Deviance profile (terms added sequentially 1st to last)

Parameter	Deviance	Degrees of freedom	Difference in Deviance
β_0	20613.47	1424	
β_1	3530.96	1423	17082.52
β_2	3520.49	1422	10.46
α_1	2639.41	1421	881.08
α_2	2634.27	1420	5.15
γ_{21}	2627.59	1419	6.67
γ_{22}	2624.26	1418	3.33

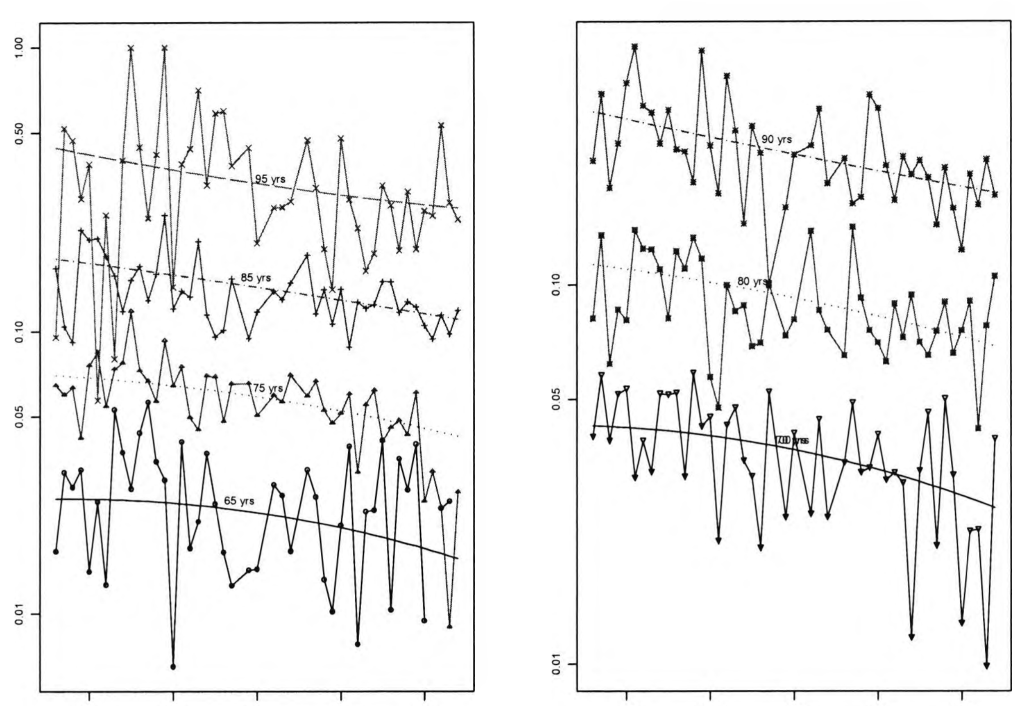


Figure 4.18 Male immediate annuitants, d1+ years, analysis of the 1946-94 mortality experience, crude mortality rates and graduated forces of mortality plotted on the log scale: 6-parameter model with $r = 2$, $s = 2$ and γ_{21}

Table 4.20
Male immediate annuitants, Duration 1 year and over,
log-link model based on an analysis of the 1946-94 experience (ϕ = 1.88)

parameter	estimate	Standard error	t-value
β_0	-2.4025	0.0102	-236.4201
β_1	1.3350	0.0204	65.3380
β_2	-0.0378	0.0177	-2.1305
α_1	-0.2470	0.0114	-21.6733
α_2	-0.0444	0.0219	-2.0304
γ_{21}	0.0883	0.0468	1.8866

Table 4.21
Male immediate annuitants, duration 1 year and over, analysis of the 1946-94 experience
***p*-values based on a 6-parameter model (4.27)**

Chi square	0.37
Cumulative deviations	0.50
Individual standardised deviations	0.07
Grouping of signs	0.36
Signs of deviations	0.04

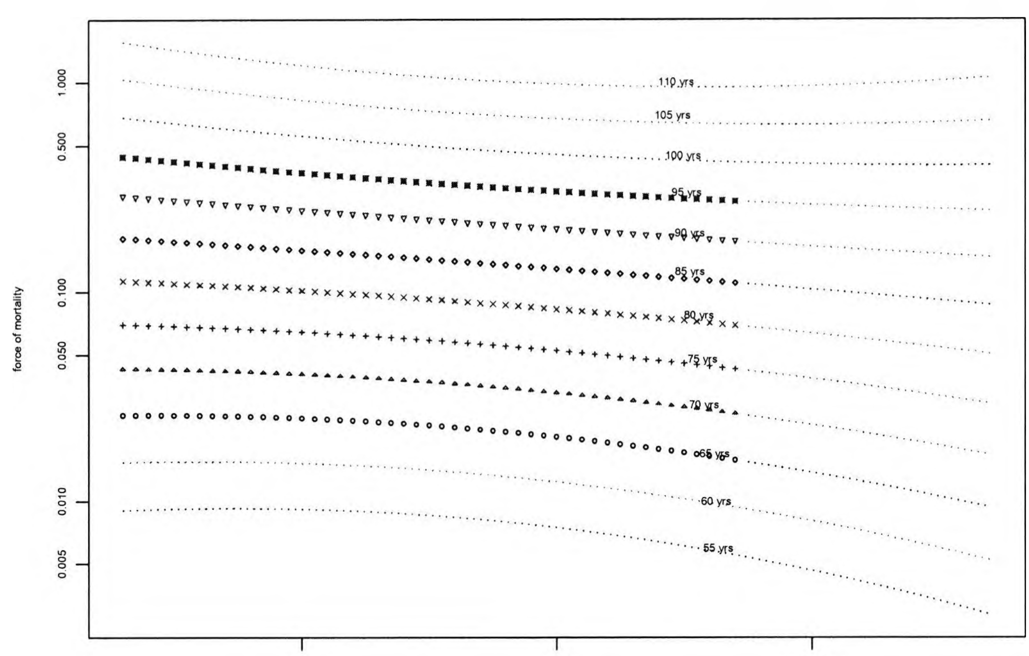


Figure 4.19 Male immediate annuitants, d1+ years, analysis of the 1946-94 mortality experience, predicted forces of mortality plotted on the log scale: based on a 6-parameter model with $r = 2$, $s = 2$ and γ_{21}

Graduated forces of mortality and crude mortality rates plotted on the log scale against calendar year are shown in Figure 4.18. From a visual inspection of this figure, the model seems to provide an adequate fit for the data.

Predicted forces of mortality for male immediate annuitants with policy duration 1+ years are shown in Figure 4.19. The forces of mortality are projected over the calendar-year period 1995 to 2014 and are shown at 5-year age intervals from age 55 to 110 years. From the plot, it is observed that the predicted rates show a consistent increase in mortality with age in each calendar year. However, although the predicted forces of mortality exhibit a downward trend at most ages, the rates at extreme old age (age 100 and above) tend to increase with time, a factor which is very likely a result of the rapid increase in the mixed product term involving t^2. In addition, the forces of mortality at these extreme ages seem exceptionally high, indicating that the recent improvements in male mortality might not be adequately represented by a model fitted to mortality experience from 1946, with each calendar year accorded the same weight in the modelling procedure.

As in the case of female annuitants, the male annuitants' mortality experience in calendar years prior to 1958 was then excluded from the analysis, in order to address the problem of recent changes in mortality trends.

4.2.7 Analysis of the 1958-94 male annuitants' mortality experience at duration 1 year and over

The male immediate annuitants' mortality experience over the period 1958-94, at curtate policy duration 1 year and over was analysed at individual ages x ranging from 65 to 95 years, excluding the one extreme observation at age 94 in calendar year 1970, giving a total of 1053 data cells. Thus x' is described by (4.26) while t' is described by (4.20), that is:

$$x' = \frac{x-80}{15}, \text{ that is, } c_x = 80 \text{ and } w_x = 15;$$

and

$$t' = \frac{t-1976}{18}, \text{ that is, } c_t = 1976 \text{ and } w_t = 18.$$

Based on an analysis of deviance (Tables 4.22 and 4.23) and the significance of the parameters introduced (Table 4.24), a simple 3-parameter model consisting of a $GM_x(0,2)$ term in age effects and an age-independent trend adjustment term linear in time t on the log scale, was found to fit the data adequately.

Thus the best-fitting model was:

$$\mu_{xt} = \exp[\beta_0 + \beta_1 L_1(x') + \alpha_1 t'] \quad (4.29)$$

or

$$\mu_{xt} = \exp[\beta_0 + \beta_1 x' + \alpha_1 t'] \quad (4.30)$$

since $L_1(x') = x'$.

The model defined by expression (4.30) is in fact equivalent to the model fitted by Wetterstrand (1981) to US mortality data from life insurance.

Table 4.22
Male Immediate Annuitants, Duration 1 year and over, Analysis of the 1958-94 experience
Deviances for polynomial predictors of degree *r* and *s*

	$r = 0$	$r = 1$	$r = 2$
$s = 0$	15677.61	15634.34	15516.68
$s = 1$	2319.39	1895.36	1895.25
$s = 2$	2318.50	1893.67	1893.60

Table 4.23
Male immediate annuitants, Duration 1 year and over, Analysis of the 1958-94 experience
Deviance profile (terms added sequentially 1st to last)

Parameter	Deviance	Degrees of freedom	Difference in Deviance
β_0	15677.61	1052	
β_1	2319.39	1051	13358.22
α_1	1895.36	1050	424.03
γ_{11}	1894.09	1049	1.27

Table 4.24
Male immediate annuitants, Duration 1 year and over,
3-parameter log-link model based on the 1958-94 experience (ϕ = 1.8223)

parameter	estimate	standard error	*t*-value
β_0	-2.467967	0.0075	-329.3263
β_1	1.352702	0.0158	85.7659
α_1	-0.194558	0.0127	-15.2822

The formula for the reduction factors derived from equation (4.30) becomes:

$$RF(x,n) = \exp\left[\alpha_1 \frac{n}{w_t}\right]. \tag{4.31}$$

For the particular data set modelled and the parameter estimates given in Table 4.24, the reduction factor is:

$$RF(x,n) = \exp\left\{-0.194558 \times \frac{n}{18}\right\} \tag{4.32}$$

or

$$RF(x,n) = \exp\{-0.010809n\} \text{ for all } x. \tag{4.33}$$

In applying this model to predict future mortality rates, there is an underlying assumption that mortality trends over time only depend on the time factor and not on the age of the individual. This assumption does not seem reasonable since the rate of improvement in mortality over time is expected to vary with age (see for example CMIR 14, 1995 and CMIR 16, 1998).

It is possible that because the male immediate annuitants' investigation is small, the underlying pattern is not fully captured by a model derived from analysing this experience over a period of 34 years. A longer investigation period could be considered. However, in view of the results of the analyses of the male and female annuitants' experiences over the period 1946-94, the observed changes over time in the composition of the annuitants' experience and the changes in mortality trends that have occurred in recent years, a longer period is likely to result in a model that does not represent the more recent trends adequately since each of the calendar years would be accorded the same weight in determining a suitable model under the methodology used. It was therefore considered appropriate to model the male annuitants' data using the 6-parameter model formula (4.21) adopted for the larger females' experience, but with the parameters estimated from the males' experience.

Table 4.25
Male immediate annuitants, Duration 1 year and over,
6-parameter log-link model based on the 1958-94 experience (deviance = 1894.09 on 1047 d.f., ϕ = 1.8314)

parameter	estimate	standard error	*t*-value
β_0	-2.473453	0.0086	-287.86
β_1	1.361608	0.0178	76.36
β_2	-0.024799	0.0210	-1.18
β_3	0.010189	0.0228	0.45
α_1	-0.199500	0.0136	-14.69
γ_{11}	0.029327	0.0292	1.00

Increasing the number of parameters can only improve the goodness-of-fit of the model, although the additional parameters introduced would not be expected to be statistically significant. The particular form of the six-parameter model results in rates that progress smoothly over both age and time, so that the smoothness criterion is also satisfied.

Table 4.25 gives the parameter estimates derived from again fitting a model of the form:

$$\mu_{xt} = \exp\left[\beta_0 + \sum_{j=1}^{3} \beta_j L_j(x') + (\alpha_1 + \gamma_{11} x')t'\right].$$

Equations (4.22) or (4.23) then give the form of the reduction factors. Based on the parameter estimates given in Table 4.25, the mortality improvement model for male annuitants becomes:

$$RF(x,n) = \exp\left[\frac{n}{18}\left\{-0.1995 + 0.029327\left(\frac{x-80}{15}\right)\right\}\right], \tag{4.34}$$

or

$$RF(x,n) = \exp[(-0.019773 + 0.000109x)n]. \tag{4.35}$$

The mortality improvement formula (4.35) is such that mortality is assumed to improve at all ages x up to age 180 approximately.

A third alternative is to fit a model consisting of the age effects term as determined from an analysis of the male annuitants' data, but with the form of the trend adjustment term constrained to be linear in both age and time on the log scale as in formula (4.21), that is, a 4-parameter model of the form:

$$\mu_{xt} = \exp[\beta_0 + \beta_1 x' + (\alpha_1 + \gamma_{11} x')t']. \tag{4.36}$$

Consequently, the mortality improvement model would again be of the form of equation (4.22) or (4.23). Based on the parameter estimates given in Table 4.26, the mortality improvement model is:

$$RF(x,n) = \exp\left[\frac{n}{18}\left\{-0.198336 + 0.023960\left(\frac{x-80}{15}\right)\right\}\right], \tag{4.37}$$

or

$$RF(x,n) = \exp[(-0.018118 + 0.0000887x)n]. \tag{4.38}$$

Table 4.26
Male immediate annuitants, Duration 1 year and over,
4-parameter log-link model based on the 1958-94 experience (deviance = 1891.48 on 1049 d.f., ϕ = 1.8240)

parameter	estimate	standard error	t-value
β_0	-2.468777	0.01	-326.42
β_1	1.353717	0.02	85.53
α_1	-0.198336	0.01	-14.67
γ_{11}	0.023960	0.03	0.84

Table 4.27
Male Immediate Annuitants, duration 1 year and over, Analysis of the 1958-94 experience
Comparison of p-values based on the 3 models fitted

Statistical test	6-parameter model	4-parameter model	3-parameter model
Chi-square	0.4941	0.4941	0.3723
Cumulative deviations	0.4626	0.3088	0.3280
Individual standardised deviations	0.0651	0.0172	0.0096
Grouping of signs of deviations	0.2405	0.3466	0.3572
Signs of deviations	0.2585	0.1246	0.1172

Table 4.28
Male Immediate Annuitants, duration 1 year and over, Analysis of the 1958-94 experience
Comparison of the distribution of individual studentized deviance residuals for the data as a whole

Range	(-∞,-3)	(-3,-2)	(-2,-1)	(-1,0)	(0,1)	(1,2)	(2,3)	(3,∞)
expected frequency	1.37	21.72	137.94	346.46	346.46	137.94	21.72	1.37
observed frequency								
6-parameter model	1.00	15.00	158.00	352.00	347.00	114.00	25.00	3.00
4-parameter model	2.00	13.00	161.00	358.00	348.00	109.00	24.00	3.00
3-parameter model	1.00	14.00	163.00	357.00	348.00	108.00	25.00	3.00

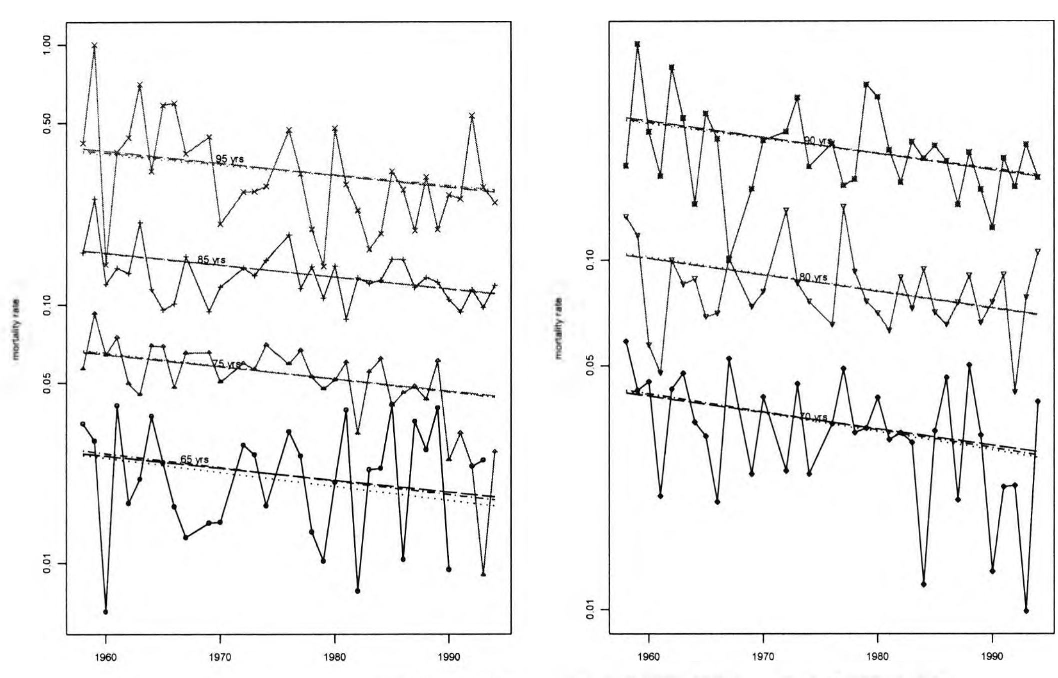


Figure 4.20 Male immediate annuitants, d1+ years, analysis of the 1958-94 experience, crude and graduated forces of mortality plotted on the log scale: comparison of 3 models

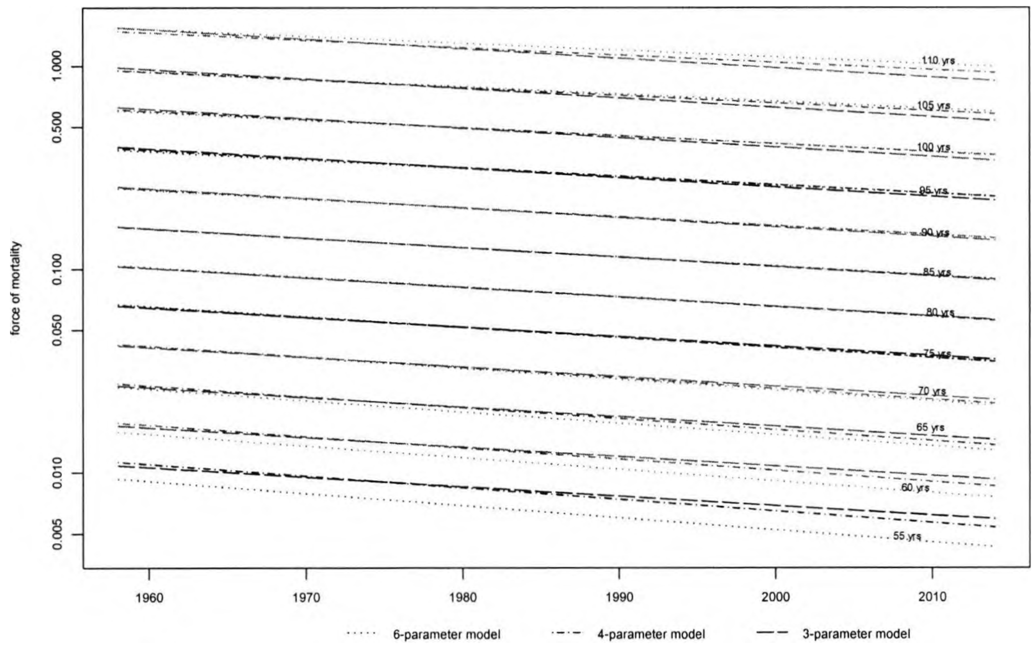


Figure 4.21 Male immediate annuitants, d1+ years, predicted forces of mortality plotted on the log scale: comparison of 3 models based on an analysis of the 1958-94 experience

The three models fitted are:

1. $\mu_{xt} = \exp[\beta_0 + \beta_1 x' + \alpha_1 t']$;
2. $\mu_{xt} = \exp[\beta_0 + \beta_1 x' + (\alpha_1 + \gamma_{11} x')t']$; and
3. $\mu_{xt} = \exp\left[\beta_0 + \beta_1 x' + \sum_{j=2}^{3}\{\beta_j L_j(x')\} + (\alpha_1 + \gamma_{11} x')t'\right]$.

From the plot of crude mortality rates and graduated forces of mortality shown in Figure 4.20, very little differences can be discerned between the three models, apart from the fact that the 6-parameter model results in the lowest forces of mortality at age 65. Each of the models provides an adequate fit to the data based on the Chi-square goodness-of-fit test, but the 4-parameter and the 3-parameter models fail the individual standardised deviations test (Tables 4.27 and 4.28).

Figure 4.21 is a comparative plot of the predicted forces of mortality at 5-year age intervals from age x = 55 to 110 years for the male annuitants' experience. The predicted forces of mortality are based on the three models given above, with the corresponding parameter estimates given in Tables 4.24, 4.25 and 4.26. As for the females' experience, predictions have been made to calendar year t = 2014. As already observed, there is little difference in the predicted rates from the three models for the period and age range over which the mortality experience has been analysed. The predicted forces of mortality based on each model progress smoothly with respect to both age and time. However, at the younger ages, predicted rates based on the 6-parameter model are noticeably lower than the predicted rates from the other two models, with the 3-parameter model providing the highest rates. At extreme old age, the reverse occurs with the 6-parameter model giving the highest forces of mortality, although the differences at these higher ages are relatively small.

The choice between the 3 models would be a matter of personal judgement although the 3-parameter model is the least favourable since using this model implies that changes over time do not depend on age. The principle of parsimony would suggest using the 4-parameter model while consideration of annuity pricing and reserves would suggest the 6-parameter model, the model that gives the lowest forces of mortality. In this study, the preferred model is the 6-parameter model, a model that has been identified from analysing a larger amount of data (the females' experience). The 6-parameter model exhibits the desired features for making predictions of future forces of mortality and the inclusion of the interaction term in age and time in the model formula means that changes over time in the forces of mortality vary with age, for all ages x. In addition, the model provides a satisfactory fit to the data based on each of the statistical tests of graduation applied, and, except for ages above 100 years, forces of mortality predicted on the basis of this model are generally the lowest over the projection period.

For all 3 models, the predicted forces of mortality at extreme old age exhibit a rapid increase with age. For example, in calendar year 1995, the force of mortality for a male aged 100 years is just under 0.5 while the corresponding predicted rate at age 110 is above 1. The predicted forces of mortality for males at these extreme ages still seem rather high. It would appear that the improvements in mortality, which have occurred in the male population in the more recent years, are not adequately reflected in the predicted forces of mortality at the older ages. In order to give more weight to the mortality experience in the more recent years, it might be necessary to determine the parameter estimates from an observation period, which excludes the earlier years, while still fitting the same model structure. This is considered in Section 4.2.8. An alternative approach, considered in Section 4.2.9, is to model the 1958-94 experience using weighted likelihood to estimate the parameters.

It should be borne in mind that the male annuitants' experience is small, so that at each individual age x in each calendar year t, the observed experience is even smaller and this can present problems in modelling the data.

4.2.8 Analysis of the 1974-94 annuitants' mortality experiences at duration 1 year and over

The 6-parameter log-link model structure

$$\mu_{xt} = \exp\left[\beta_0 + \sum_{j=1}^{3}\beta_j L_j(x') + (\alpha_1 + \gamma_{11}x')t'\right]$$

was also fitted to the male immediate annuitants' mortality experience at duration 1 year and over, for the calendar year period 1974 to 1994, so that the scaled calendar time t' is given by:

$$t' = \frac{t-1984}{10}\text{; that is, } c_t = 1984 \text{ and } w_t = 10. \qquad (4.39)$$

The period was chosen in order to put emphasis on the more recent experience. Although the choice of the specific calendar-year 1974 was arbitrary, it is worth noting that in deriving the current mortality improvement model for pensioners and annuitants, the CMI Committee considered mortality experiences over the quadrennia beginning with 1975-1978.

As for the 1958-94 experience, the 1974-94 male annuitants' mortality experience was analysed over the age range 65 to 95 years (620 data cells). Hence as before, the transformed age is defined as $x' = \dfrac{x-80}{15}$.

Parameter estimates for the model are given in Table 4.29. Here 2 parameters, β_3 and γ_{11} are not statistically significant (compared with 3 parameters in the 6-parameter model based on the 1958-94 experience). From the analysis of deviance (Tables 4.30 and 4.31), both β_3 and γ_{11} could well be excluded from the model. However the parameters were included because the aim was to determine parameter estimates that would result in reasonable projected forces of mortality, based on the pre-determined 6-parameter model structure given by (4.21).

In fact, the model provides a good fit to the data and the statistical tests of graduation carried out are supportive of the model. Some results of the tests of graduation on the data as a whole are shown in Tables 4.32 and 4.33.

Table 4.29
Male immediate annuitants, Duration 1 year and over,
6-parameter log-link model based on the 1974-94 experience (deviance = 924 on 614 d.f., ϕ = 1.5159)

parameter	estimate	standard error	t-value
β_0	-2.564089	0.0113	-226.6192
β_1	1.383162	0.0242	57.2702
β_2	-0.059597	0.0271	-2.1981
β_3	-0.028188	0.0283	-0.9948
α_1	-0.139243	0.0165	-8.4168
γ_{11}	0.026200	0.0342	0.7650

Table 4.30
Male Immediate Annuitants, Duration 1 year and over, Analysis of the 1974-94 experience
Deviances for polynomial predictors of degree *r* and *s*

	$r = 0$	$r = 1$	$r = 2$	$r = 3$
$s = 0$	8732.10	8721.45	8716.30	8716.29
$s = 1$	1051.07	937.68	931.72	931.06
$s = 2$	1044.20	926.53	921.30	920.50
$s = 3$	1043.48	924.96	919.39	918.57

Table 4.31
Male Immediate Annuitants, Duration 1 year and over, Analysis of the 1974-94 experience
Deviance profile (terms added sequentially 1st to last)

Parameter	Deviance	Degrees of freedom	Difference in deviance
β_0	8732.10	619.00	
β_1	1051.07	618.00	7681.03
β_2	1044.20	617.00	6.87
β_3	1043.48	616.00	0.72
α_1	924.96	615.00	118.52
γ_{11}	924.07	614.00	0.89

Table 4.32
Male Immediate Annuitants, Duration 1 year and over, Analysis of the 1974-94 experience
Distribution of individual studentized deviance residuals for the data as a whole: 6-parameter model

Range	(-∞,-3)	(-3,-2)	(-2,-1)	(-1,0)	(0,1)	(1,2)	(2,3)	(3,∞)
expected frequency	0.79	12.48	79.23	199.00	199.00	79.23	12.48	0.79
observed frequency	1.00	8.00	94.00	194.00	205.00	66.00	13.00	2.00

Table 4.33
Male Immediate Annuitants, Duration 1 year and over, Analysis of the 1974-94 experience
***p*-values based on a 6-parameter log-link model**

Statistical test	*p*-value
Chi-square	0.4922
Cumulative deviations	0.3933
Individual standardised deviations	0.2308
Grouping of signs of deviations	0.9736
Signs of deviations	0.6788

From the plot of predicted forces of mortality shown in Figure 4.22, this model would appear to result in forces of mortality that are more realistic, particularly at extreme old age. Whereas predicted forces of mortality based on parameter estimates determined from an analysis of the 1958-94 experience are as high as 1, the rates predicted from the model based on the 1974-94 experience have a maximum value just above 0.5. In general, predicted rates based on a trend analysis of the 1974-94 mortality experience are lower than predicted rates based on the 1958-94 mortality experience except for ages below 60 years. It would therefore appear that parameter estimates based on the 1974-94 mortality experience would give the preferred predicted forces of mortality for annuity pricing.

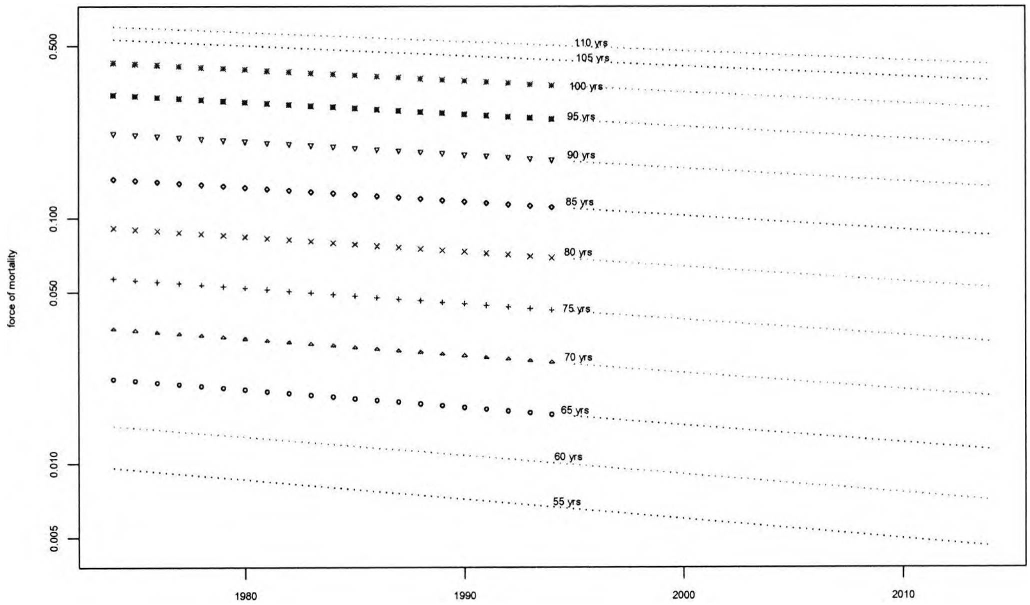


Figure 4.22 Male immediate annuitants, d1+ years, predicted forces of mortality plotted on the log scale; 6-parameter log-link model with $r = 1$, $s = 3$ and γ_{11}, based on an analysis of the 1974-94 mortality experience

The formula for the reduction factors based on the parameter estimates given in Table 4.29 is:

$$RF(x,n) = \exp\left[\frac{n}{10}\left\{-0.139243 + 0.026200\left(\frac{x-80}{15}\right)\right\}\right] \tag{4.40}$$

or

$$RF(x,n) = \exp[\{-0.0278976 + 0.0001747x\}n]. \tag{4.41}$$

For consistency, the same 6-parameter model structure was also fitted to the female immediate annuitants' mortality experience over the calendar-year period 1974-94 for age x = 65 to 95 years, giving a total of 620 data cells. The parameter estimates thus obtained, shown in Table 4.34, are all statistically significant with the exception of β_2. An analysis of deviance (Tables 4.35 and 4.36) shows that when the value of s is increased from 1 to 2, thus introducing β_2, the improvement in deviance is not significant. However, a subsequent increase in the value of s to 3 results in a significant improvement in deviance.

Based on the distribution of individual standardised deviations shown in Table 4.37 and the results of statistical tests of graduation shown in Table 4.38, the fitted values from the model provide a satisfactory representation of the underlying forces of mortality.

Table 4.34
Female immediate annuitants, Duration 1 year and over,
6-parameter log-link model based on the 1974-94 experience (deviance = 937 on 614 d.f., ϕ = 1.5264)

parameter	estimate	standard error	t-value
β_0	-2.958437	0.0113	-262.6085
β_1	1.622580	0.0236	68.8798
β_2	0.033533	0.0246	1.3612
β_3	-0.066741	0.0210	-3.1782
α_1	-0.177477	0.0151	-11.7715
γ_{11}	0.057479	0.0267	2.1519

Table 4.35
Female Immediate Annuitants, Duration 1 year and over, Analysis of the 1974-94 experience
Deviances for polynomial predictors of degree *r* and *s*

	$r = 0$	$r = 1$	$r = 2$	$r = 3$
$s = 0$	22302.26	22207.91	22185.40	22185.09
$s = 1$	1277.52	959.71	952.57	951.69
$s = 2$	1277.51	959.36	952.05	951.15
$s = 3$	1265.29	944.10	937.23	936.25
$s = 4$	1264.70	943.15	936.21	935.21

Table 4.36
Female Immediate Annuitants, Duration 1 year and over, Analysis of the 1974-94 experience
Deviance profile (terms added sequentially 1st to last)

Parameter	Deviance	Degrees of freedom	Difference in deviance
β_0	22302.26	619.00	
β_1	1277.52	618.00	21024.74
β_2	1277.51	617.00	0.01
β_3	1265.29	616.00	12.22
α_1	944.10	615.00	321.19
γ_{11}	937.01	614.00	7.09

Table 4.37
Female Immediate Annuitants, Duration 1 year and over, Analysis of the 1974-94 experience
Distribution of individual studentized deviance residuals for the data as a whole: 6-parameter model

Range	(-∞,-3)	(-3,-2)	(-2,-1)	(-1,0)	(0,1)	(1,2)	(2,3)	(3,∞)
expected frequency	0.77	12.20	77.47	194.57	194.57	77.47	12.20	0.77
observed frequency	1.00	14.00	82.00	189.00	210.00	61.00	10.00	3.00

Table 4.38
Female Immediate Annuitants, Duration 1 year and over, Analysis of the 1974-94 experience
p-values based on a 6-parameter log-link model

Statistical test	p-value
Chi-square	0.4921
Cumulative deviations	0.3490
Individual standardised deviations	0.3615
Grouping of signs of deviations	0.5990
Signs of deviations	0.9666

Figure 4.23 is a comparative plot of predicted forces of mortality for female immediate annuitants with policy duration 1 year and over, for the period 1992 to 2014, with model parameters estimated from the 1958-94 and the 1974-94 mortality experiences. A similar pattern as for the male annuitants emerges in that predicted forces of mortality based on the 1974-94 mortality experience are lower over the whole period for x greater than 65 years, and higher at ages below 60 years. For ages between 60 and 65 years, the rates based on the 1974-94 experience are generally higher during the earlier years of the projection period. Overall, it would appear that as for male annuitants with policy duration one year and over, the model based on the

1974-94 mortality experience gives lower predicted forces of mortality for female annuitants at duration 1 year and over.

The corresponding formula for the reduction factors based on parameter estimates given in Table 4.34 is:

$$RF(x,n)=\exp\left[\frac{n}{10}\left\{-0.177477+0.057479\left(\frac{x-80}{15}\right)\right\}\right] \quad (4.42)$$

or

$$RF(x,n)=\exp[\{-0.048403+0.000383x\}n]. \quad (4.43)$$

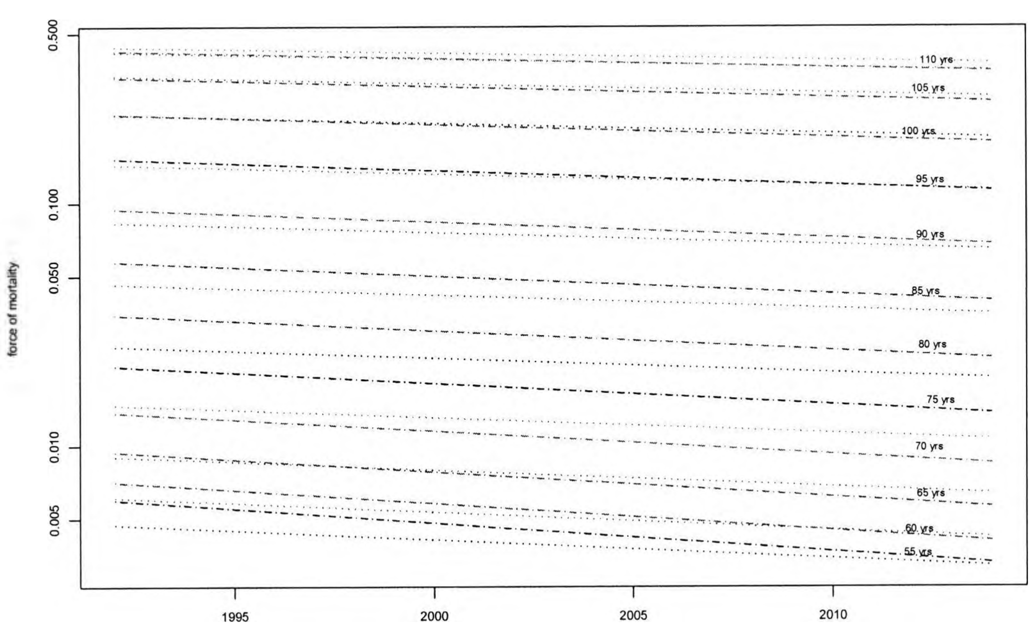


Figure 4.23 Female immediate annuitants, d1+ years, 1992-2014 predicted forces of mortality plotted on the log scale; 6-parameter log-link models with r = 1, s = 3 and γ_{11}, based on the 1958-94 and the 1974-94 mortality experiences

Predicted forces of mortality for male and female immediate annuitants, with projections made over a 20-year period, from 1995 to 2014 are shown in Figure 4.24. The predicted forces of mortality for male annuitants are consistently higher than predicted forces of mortality for female annuitants at all ages and in each of the 20 calendar years, for which the rates have been projected, with the highest differences occurring at the older ages. Tables 4.39a, 4.39b and Tables 4.40a, 4.40b give the predicted forces of mortality for males and females respectively, tabulated at quadrennial periods from 1974 to 2014.

As an illustration of confidence intervals for future forces of mortality for both male and female annuitants, Appendix 2 shows confidence intervals for predicted forces of mortality in the calendar-year 2014 based on modelling the 1974-94 mortality experiences.

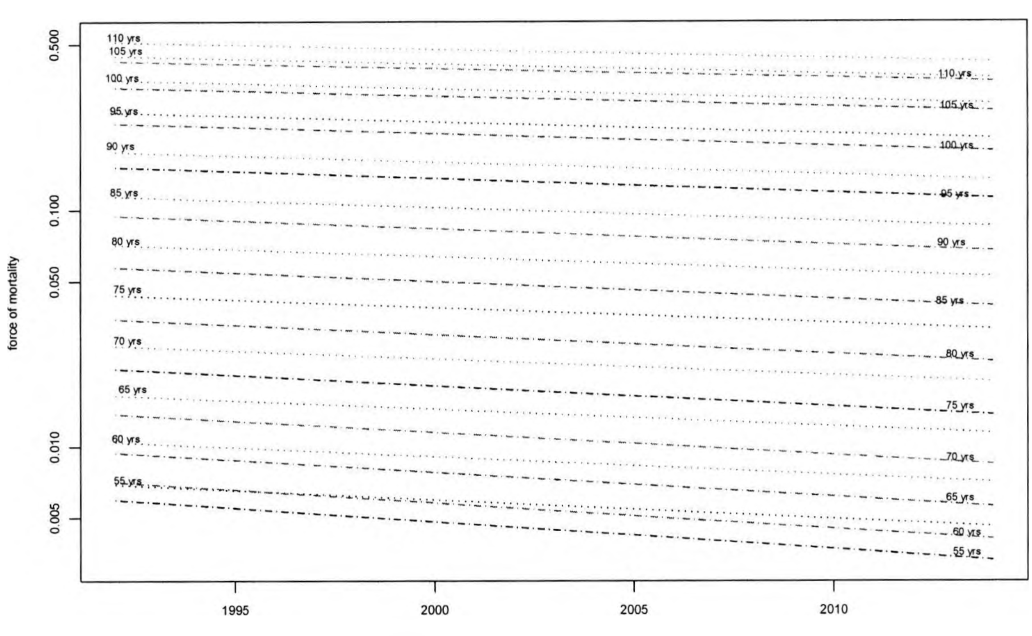


Figure 4.24 Immediate annuitants, d1+ years, 1992-2014 predicted forces of mortality based on the 6-parameter log-link models derived from the 1974-94 mortality experiences

Table 4.39a Male Immediate Annuitants, Duration 1 year and over
Analysis of the 1974-94 experience, predicted force of mortality at quadrennial periods 1974-1994

age	1974	1978	1982	1986	1990	1994
55	0.009569	0.008894	0.008267	0.007683	0.007141	0.006638
56	0.010316	0.009595	0.008924	0.008300	0.007720	0.007181
57	0.011146	0.010374	0.009656	0.008987	0.008365	0.007785
58	0.012067	0.011240	0.010469	0.009750	0.009082	0.008459
59	0.013091	0.012201	0.011372	0.010600	0.009879	0.009208
60	0.014227	0.013270	0.012377	0.011544	0.010767	0.010042
61	0.015489	0.014457	0.013493	0.012594	0.011754	0.010971
62	0.016889	0.015774	0.014733	0.013761	0.012853	0.012005
63	0.018443	0.017237	0.016111	0.015058	0.014074	0.013155
64	0.020166	0.018861	0.017641	0.016500	0.015433	0.014434
65	0.022077	0.020663	0.019340	0.018102	0.016943	0.015858
66	0.024196	0.022662	0.021226	0.019880	0.018620	0.017440
67	0.026543	0.024878	0.023317	0.021855	0.020484	0.019199
68	0.029142	0.027333	0.025637	0.024045	0.022553	0.021153
69	0.032018	0.030052	0.028206	0.026474	0.024848	0.023322
70	0.035199	0.033061	0.031052	0.029165	0.027393	0.025729
71	0.038714	0.036387	0.034200	0.032145	0.030213	0.028397
72	0.042594	0.040062	0.037680	0.035440	0.033333	0.031352
73	0.046872	0.044116	0.041523	0.039082	0.036784	0.034622
74	0.051584	0.048586	0.045761	0.043101	0.040596	0.038236
75	0.056768	0.053505	0.050430	0.047532	0.044800	0.042225
76	0.062462	0.058913	0.055566	0.052409	0.049431	0.046623
77	0.068707	0.064849	0.061207	0.057770	0.054526	0.051464
78	0.075544	0.071352	0.067392	0.063652	0.060120	0.056784
79	0.083017	0.078465	0.074162	0.070096	0.066252	0.062619
80	0.091168	0.086229	0.081558	0.077140	0.072961	0.069008
81	0.100040	0.094687	0.089620	0.084824	0.080285	0.075988
82	0.109675	0.103878	0.098388	0.093188	0.088263	0.083598
83	0.120111	0.113842	0.107901	0.102270	0.096932	0.091873
84	0.131387	0.124617	0.118196	0.112105	0.106329	0.100850
85	0.143535	0.136234	0.129305	0.122728	0.116485	0.110560
86	0.156584	0.148723	0.141257	0.134166	0.127430	0.121033
87	0.170555	0.162106	0.154076	0.146443	0.139189	0.132294
88	0.185463	0.176398	0.167777	0.159577	0.151778	0.144360
89	0.201311	0.191606	0.182369	0.173577	0.165209	0.157245
90	0.218094	0.207725	0.197849	0.188443	0.179483	0.170950
91	0.235793	0.224740	0.214204	0.204163	0.194592	0.185470
92	0.254375	0.242620	0.231408	0.220714	0.210515	0.200787
93	0.273792	0.261322	0.249420	0.238060	0.227218	0.216869
94	0.293977	0.280784	0.268183	0.256147	0.244652	0.233673
95	0.314847	0.300928	0.287624	0.274908	0.262754	0.251137
96	0.336299	0.321655	0.307650	0.294254	0.281441	0.269187
97	0.358207	0.342850	0.328150	0.314081	0.300615	0.287727
98	0.380429	0.364373	0.348995	0.334266	0.320158	0.306646
99	0.402800	0.386069	0.370034	0.354664	0.339933	0.325814
100	0.425134	0.407760	0.391097	0.375115	0.359785	0.345082
101	0.447228	0.429251	0.411997	0.395437	0.379542	0.364286
102	0.468861	0.450329	0.432530	0.415435	0.399015	0.383244
103	0.489798	0.470768	0.452477	0.434897	0.418000	0.401759
104	0.509793	0.490329	0.471607	0.453601	0.436281	0.419624
105	0.528591	0.508764	0.489681	0.471313	0.453635	0.436619
106	0.545935	0.525824	0.506455	0.487799	0.469830	0.452523
107	0.561567	0.541259	0.521685	0.502819	0.484636	0.467110
108	0.575239	0.554824	0.535134	0.516142	0.497824	0.480157
109	0.586715	0.566288	0.546572	0.527543	0.509176	0.491449
110	0.595775	0.575435	0.555789	0.536814	0.518486	0.500785

Table 4.39b Male Immediate Annuitants, Duration 1 year and over
Analysis of the 1974-94 experience, predicted force of mortality at quadrennial periods 1998-2014

age	1998	2002	2006	2010	2014
55	0.006169	0.005734	0.005329	0.004953	0.004604
56	0.006679	0.006212	0.005778	0.005374	0.004998
57	0.007246	0.006744	0.006277	0.005843	0.005438
58	0.007878	0.007338	0.006834	0.006366	0.005929
59	0.008582	0.007999	0.007456	0.006949	0.006477
60	0.009367	0.008736	0.008148	0.007600	0.007088
61	0.010240	0.009557	0.008920	0.008326	0.007771
62	0.011212	0.010472	0.009781	0.009136	0.008533
63	0.012295	0.011492	0.010741	0.010039	0.009383
64	0.013501	0.012627	0.011810	0.011046	0.010332
65	0.014842	0.013892	0.013002	0.012170	0.011390
66	0.016335	0.015299	0.014330	0.013422	0.012571
67	0.017995	0.016866	0.015808	0.014816	0.013887
68	0.019840	0.018608	0.017453	0.016370	0.015354
69	0.021890	0.020545	0.019284	0.018099	0.016988
70	0.024165	0.022697	0.021318	0.020023	0.018806
71	0.026690	0.025086	0.023578	0.022161	0.020829
72	0.029488	0.027735	0.026086	0.024536	0.023077
73	0.032586	0.030671	0.028868	0.027170	0.025573
74	0.036013	0.033920	0.031948	0.030091	0.028341
75	0.039798	0.037511	0.035355	0.033323	0.031408
76	0.043974	0.041476	0.039119	0.036897	0.034801
77	0.048574	0.045846	0.043272	0.040842	0.038548
78	0.053632	0.050656	0.047845	0.045190	0.042682
79	0.059185	0.055940	0.052872	0.049973	0.047233
80	0.065269	0.061733	0.058389	0.055226	0.052234
81	0.071922	0.068073	0.064430	0.060982	0.057719
82	0.079179	0.074995	0.071031	0.067277	0.063721
83	0.087078	0.082534	0.078226	0.074144	0.070274
84	0.095653	0.090724	0.086050	0.081616	0.077410
85	0.104936	0.099599	0.094533	0.089724	0.085161
86	0.114957	0.109186	0.103705	0.098499	0.093554
87	0.125740	0.119511	0.113591	0.107964	0.102616
88	0.137305	0.130594	0.124212	0.118141	0.112367
89	0.149664	0.142449	0.135582	0.129046	0.122824
90	0.162823	0.155082	0.147708	0.140686	0.133997
91	0.176776	0.168489	0.160591	0.153062	0.145887
92	0.191508	0.182658	0.174217	0.166166	0.158487
93	0.206992	0.197564	0.188566	0.179978	0.171781
94	0.223186	0.213170	0.203603	0.194466	0.185739
95	0.240035	0.229423	0.219280	0.209585	0.200319
96	0.257466	0.246255	0.235533	0.225277	0.215468
97	0.275391	0.263584	0.252283	0.241467	0.231114
98	0.293704	0.281308	0.269436	0.258064	0.247173
99	0.312281	0.299310	0.286878	0.274963	0.263542
100	0.330980	0.317455	0.304482	0.292039	0.280105
101	0.349644	0.335590	0.322101	0.309154	0.296727
102	0.368097	0.353548	0.339574	0.326153	0.313262
103	0.386150	0.371147	0.356726	0.342866	0.329545
104	0.403602	0.388192	0.373370	0.359114	0.345403
105	0.420242	0.404479	0.389307	0.374705	0.360650
106	0.435854	0.419798	0.404335	0.389440	0.375095
107	0.450217	0.433936	0.418243	0.403118	0.388540
108	0.463116	0.446680	0.430828	0.415538	0.400790
109	0.474339	0.457824	0.441885	0.426500	0.411652
110	0.483687	0.467174	0.451224	0.435819	0.420939

Table 4.40a Female Immediate Annuitants, Duration 1 year and over
Analysis of the 1974-94 experience, predicted force of mortality at quadrennial periods 1974-1994

age	1974	1978	1982	1986	1990	1994
55	0.009631	0.008652	0.007773	0.006984	0.006274	0.005637
56	0.009807	0.008823	0.007937	0.007140	0.006423	0.005778
57	0.010040	0.009044	0.008147	0.007338	0.006610	0.005954
58	0.010331	0.009318	0.008405	0.007581	0.006838	0.006167
59	0.010683	0.009648	0.008714	0.007870	0.007108	0.006419
60	0.011098	0.010037	0.009077	0.008208	0.007423	0.006713
61	0.011583	0.010489	0.009498	0.008600	0.007788	0.007052
62	0.012141	0.011008	0.009982	0.009050	0.008206	0.007441
63	0.012779	0.011602	0.010534	0.009564	0.008683	0.007883
64	0.013504	0.012277	0.011161	0.010146	0.009224	0.008386
65	0.014324	0.013040	0.011870	0.010805	0.009836	0.008954
66	0.015249	0.013900	0.012670	0.011548	0.010526	0.009595
67	0.016289	0.014867	0.013569	0.012385	0.011303	0.010317
68	0.017457	0.015954	0.014580	0.013324	0.012177	0.011129
69	0.018765	0.017171	0.015713	0.014379	0.013158	0.012041
70	0.020228	0.018535	0.016983	0.015562	0.014259	0.013066
71	0.021863	0.020060	0.018405	0.016887	0.015493	0.014215
72	0.023690	0.021764	0.019995	0.018369	0.016876	0.015504
73	0.025728	0.023667	0.021772	0.020028	0.018425	0.016949
74	0.028000	0.025791	0.023757	0.021883	0.020157	0.018567
75	0.030531	0.028160	0.025973	0.023956	0.022095	0.020379
76	0.033348	0.030799	0.028444	0.026270	0.024261	0.022407
77	0.036482	0.033737	0.031199	0.028852	0.026681	0.024674
78	0.039965	0.037006	0.034267	0.031731	0.029382	0.027207
79	0.043831	0.040640	0.037681	0.034938	0.032394	0.030036
80	0.048118	0.044673	0.041476	0.038507	0.035750	0.033191
81	0.052867	0.049147	0.045689	0.042474	0.039485	0.036707
82	0.058119	0.054101	0.050360	0.046878	0.043637	0.040620
83	0.063921	0.059579	0.055533	0.051761	0.048246	0.044969
84	0.070318	0.065628	0.061251	0.057166	0.053354	0.049796
85	0.077359	0.072295	0.067562	0.063139	0.059006	0.055143
86	0.085093	0.079627	0.074512	0.069726	0.065247	0.061056
87	0.093570	0.087675	0.082151	0.076975	0.072125	0.067581
88	0.102839	0.096487	0.090526	0.084934	0.079687	0.074765
89	0.112948	0.106110	0.099686	0.093651	0.087981	0.082655
90	0.123940	0.116589	0.109675	0.103171	0.097052	0.091296
91	0.135856	0.127967	0.120536	0.113536	0.106943	0.100733
92	0.148730	0.140277	0.132305	0.124786	0.117695	0.111006
93	0.162589	0.153550	0.145014	0.136953	0.129339	0.122149
94	0.177449	0.167805	0.158685	0.150060	0.141905	0.134192
95	0.193315	0.183049	0.173328	0.164123	0.155407	0.147154
96	0.210178	0.199278	0.188943	0.179144	0.169854	0.161045
97	0.228011	0.216471	0.205514	0.195112	0.185237	0.175861
98	0.246770	0.234588	0.223007	0.211998	0.201532	0.191583
99	0.266387	0.253569	0.241368	0.229754	0.218699	0.208176
100	0.286772	0.273332	0.260523	0.248313	0.236676	0.225584
101	0.307811	0.293771	0.280372	0.267583	0.255378	0.243730
102	0.329362	0.314753	0.300791	0.287449	0.274698	0.262513
103	0.351256	0.336117	0.321630	0.307767	0.294502	0.281808
104	0.373296	0.357677	0.342710	0.328370	0.314630	0.301465
105	0.395259	0.379218	0.363828	0.349063	0.334897	0.321305
106	0.416896	0.400503	0.384754	0.369625	0.355090	0.341127
107	0.437935	0.421267	0.405234	0.389811	0.374975	0.360704
108	0.458085	0.441230	0.424995	0.409358	0.394296	0.379788
109	0.477043	0.460094	0.443748	0.427982	0.412777	0.398111
110	0.494494	0.477553	0.461192	0.445391	0.430132	0.415395

Table 4.40b Female Immediate Annuitants, Duration 1 year and over
Analysis of the 1974-94 experience, predicted force of mortality at quadrennial periods 1998-2014

age	1998	2002	2006	2010	2014
55	0.005064	0.004550	0.004088	0.003672	0.003299
56	0.005198	0.004676	0.004207	0.003784	0.003404
57	0.005364	0.004831	0.004352	0.003920	0.003531
58	0.005563	0.005017	0.004525	0.004082	0.003682
59	0.005798	0.005236	0.004729	0.004271	0.003857
60	0.006071	0.005490	0.004965	0.004490	0.004060
61	0.006386	0.005783	0.005236	0.004742	0.004294
62	0.006747	0.006117	0.005547	0.005029	0.004560
63	0.007157	0.006498	0.005900	0.005357	0.004863
64	0.007623	0.006930	0.006300	0.005728	0.005207
65	0.008150	0.007419	0.006754	0.006148	0.005597
66	0.008745	0.007971	0.007266	0.006623	0.006037
67	0.009416	0.008594	0.007843	0.007159	0.006534
68	0.010170	0.009295	0.008494	0.007763	0.007094
69	0.011019	0.010083	0.009227	0.008444	0.007727
70	0.011972	0.010970	0.010052	0.009210	0.008440
71	0.013043	0.011967	0.010979	0.010074	0.009243
72	0.014244	0.013086	0.012022	0.011045	0.010147
73	0.015592	0.014343	0.013194	0.012138	0.011166
74	0.017103	0.015754	0.014511	0.013367	0.012313
75	0.018797	0.017337	0.015990	0.014748	0.013603
76	0.020694	0.019112	0.017651	0.016301	0.015055
77	0.022817	0.021101	0.019513	0.018045	0.016687
78	0.025193	0.023328	0.021602	0.020003	0.018522
79	0.027849	0.025822	0.023942	0.022199	0.020583
80	0.030815	0.028609	0.026561	0.024660	0.022895
81	0.034124	0.031723	0.029491	0.027416	0.025487
82	0.037811	0.035197	0.032763	0.030498	0.028390
83	0.041915	0.039068	0.036415	0.033942	0.031636
84	0.046475	0.043375	0.040482	0.037782	0.035263
85	0.051533	0.048159	0.045006	0.042060	0.039307
86	0.057134	0.053464	0.050029	0.046816	0.043808
87	0.063323	0.059333	0.055595	0.052092	0.048810
88	0.070146	0.065813	0.061747	0.057933	0.054354
89	0.077651	0.072950	0.068533	0.064384	0.060486
90	0.085882	0.080789	0.075997	0.071490	0.067250
91	0.094884	0.089374	0.084184	0.079296	0.074691
92	0.104697	0.098747	0.093135	0.087842	0.082850
93	0.115359	0.108946	0.102890	0.097170	0.091768
94	0.126899	0.120002	0.113480	0.107313	0.101480
95	0.139340	0.131940	0.124933	0.118298	0.112016
96	0.152693	0.144774	0.137266	0.130147	0.123398
97	0.166960	0.158509	0.150486	0.142869	0.135638
98	0.182125	0.173134	0.164587	0.156462	0.148738
99	0.198159	0.188624	0.179548	0.170909	0.162686
100	0.215012	0.204935	0.195331	0.186177	0.177451
101	0.232613	0.222003	0.211877	0.202212	0.192989
102	0.250869	0.239741	0.229106	0.218944	0.209232
103	0.269662	0.258039	0.246918	0.236275	0.226091
104	0.288851	0.276764	0.265183	0.254087	0.243455
105	0.308266	0.295755	0.283752	0.272237	0.261188
106	0.327713	0.314827	0.302447	0.290554	0.279129
107	0.346976	0.333770	0.321067	0.308848	0.297093
108	0.365814	0.352354	0.339389	0.326901	0.314873
109	0.383967	0.370325	0.357168	0.344479	0.332240
110	0.401163	0.387419	0.374146	0.361328	0.348949

4.2.9 Analysis of the 1958-94 male annuitants' mortality experience at duration 1+ years using weighted likelihood

From the preceding discussions, it can be inferred that the preferred model for projecting mortality for male immediate annuitants with policy duration 1 year and over would be a 6-parameter model based on an analysis of the 1974-94 experience. The model appears to provide a better representation of the likely future trends in male annuitants' mortality when compared to models based on the 1958-94 experience.

An alternative approach would be to analyse the 1958-94 experience using weighted likelihood to estimate the model parameters. The likelihood is weighted by calendar year, with the greatest weight given to the most recent calendar year in the investigation. This approach would mean that while trends are analysed over a longer period, greater contribution in the determination of the model parameters is provided by the more recent experience, which is considered to be of greater relevance to future mortality trends.

Denoting the weight factor applicable to calendar year t as ω_t, and the maximum weight as ω, with $0 < \omega \leq 1$ and $0 < \omega_t \leq \omega$, ω_t may be defined as

$$\omega_t = \frac{(t - t_{\min} + 1)}{(t_{\max} - t_{\min} + 1)} \times \omega , \tag{4.44}$$

for some suitably chosen value of ω. The unknown parameters are therefore estimated by maximising the expression:

$$\frac{1}{\phi} \sum_{x,t} \omega_t \{- m_{xt} + a_{xt} \log(m_{xt})\}. \tag{4.45}$$

Using this approach, the 6-parameter model

$$\mu_{xt} = \exp\left\{ \sum_{j=0}^{3} \beta_j L_j(x') + (\alpha_1 + \gamma_{11} x') t' \right\}$$

was refitted to the 1958-94 male annuitants' mortality experience with $\omega = 0.9$. Forces of mortality predicted on the basis of the revised parameter estimates, together with forces of mortality derived in Section 4.2.7 using unweighted likelihood, are shown in Figure 4.25. Clearly the rapid increase with age of forces of mortality at extreme old age has been reduced. In particular, the revised parameter estimates are such that in general, the predicted forces of mortality are lower at each age above 90 years and $\mu_{xt} < 1$ for all $x \leq 110$.

An examination of the parameter estimates involving terms in x and higher order terms in x, reveals that while the estimate of β_3 in the original model is positive, the revised estimate in the model based on weighted likelihood is negative, thereby having the general effect of reducing predicted forces of mortality particularly when x is large. Even though the parameter β_3 is not statistically significant, its assumed value will still have an effect on the values of the predicted forces of mortality at each age. In addition, the absolute value of β_2, the other negative parameter estimate involving terms in x, is larger in the revised model. The parameter estimates are shown in Table 4.41.

Table 4.41
Male immediate annuitants, d1+ years, 6-parameter log-link model based on the 1958-94 experience, parameter estimates based on weighted and unweighted likelihood

Parameter	Estimate based on unweighted likelihood	Estimate based on weighted likelihood[2]
β_0	-2.473453	-2.477385 (0.0091)
β_1	1.361608	1.357315 (0.0191)
β_2	-0.024799	-0.044192 (0.0208)
β_3	0.010189	-0.010001 (0.0220)
α_1	-0.199500	-0.202962 (0.0156)
γ_{11}	0.029327	0.050063 (0.0331)

[2] Standard errors are given in parenthesis.

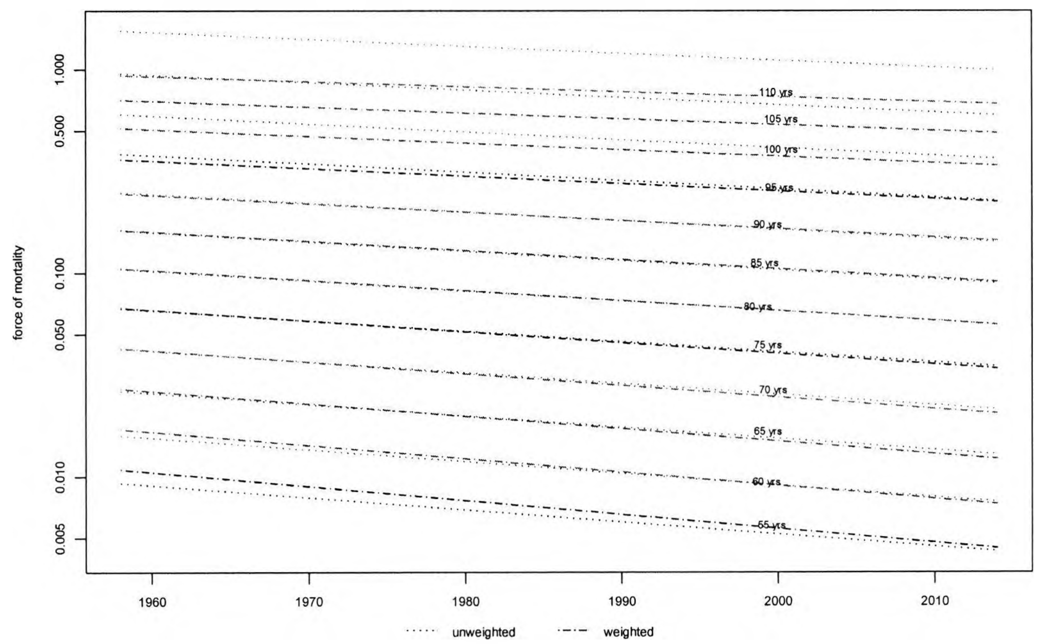


Figure 4.25 Male immediate annuitants, d1+ years, analysis of the 1958-94 mortality experience, 6-parameter log-link model with parameter estimates based on weighted and unweighted likelihood

The pattern of the revised predicted forces of mortality seems to confirm the assertion that the improvements in male annuitants' mortality that have occurred in the more recent years are not adequately reflected in a model derived from the experience from 1958, with equal weighting given to each calendar year. It is also worth noting at this point that parameter estimates derived from the 1974-94 experience still provide the lowest projected forces of mortality for male annuitants with policy duration 1 year and over.

For completeness, the predicted forces of mortality on the basis of the revised parameter estimates shown in Table 4.41, are given in Tables 4.42a, 4.42b and 4.42c.

Table 4.42a Predicted forces of mortality for male immediate annuitants, duration 1+ years, 1958-74
Based on an analysis of 1958-94 experience, model parameters estimated using weighted likelihood

age	1958	1962	1966	1970	1974
55	0.010840	0.010172	0.009545	0.008956	0.008404
56	0.011852	0.011130	0.010451	0.009814	0.009216
57	0.012965	0.012184	0.011449	0.010759	0.010111
58	0.014188	0.013343	0.012548	0.011801	0.011098
59	0.015533	0.014618	0.013758	0.012948	0.012185
60	0.017011	0.016021	0.015089	0.014211	0.013384
61	0.018635	0.017564	0.016554	0.015603	0.014706
62	0.020419	0.019260	0.018166	0.017135	0.016162
63	0.022379	0.021124	0.019940	0.018821	0.017766
64	0.024531	0.023173	0.021890	0.020677	0.019532
65	0.026894	0.025423	0.024033	0.022719	0.021477
66	0.029486	0.027894	0.026389	0.024964	0.023617
67	0.032329	0.030607	0.028976	0.027432	0.025971
68	0.035446	0.033582	0.031817	0.030144	0.028560
69	0.038861	0.036845	0.034934	0.033122	0.031404
70	0.042601	0.040421	0.038353	0.036391	0.034529
71	0.046694	0.044338	0.042100	0.039976	0.037959
72	0.051171	0.048625	0.046205	0.043906	0.041722
73	0.056063	0.053313	0.050698	0.048212	0.045847
74	0.061407	0.058438	0.055613	0.052924	0.050366
75	0.067238	0.064035	0.060984	0.058079	0.055312
76	0.073596	0.070142	0.066850	0.063712	0.060722
77	0.080522	0.076800	0.073249	0.069863	0.066634
78	0.088059	0.084051	0.080225	0.076573	0.073088
79	0.096254	0.091941	0.087821	0.083886	0.080127
80	0.105154	0.100517	0.096084	0.091847	0.087796
81	0.114810	0.109828	0.105062	0.100504	0.096143
82	0.125272	0.119926	0.114807	0.109907	0.105216
83	0.136596	0.130863	0.125370	0.120108	0.115067
84	0.148835	0.142694	0.136807	0.131162	0.125750
85	0.162047	0.155476	0.149172	0.143123	0.137320
86	0.176289	0.169266	0.162523	0.156048	0.149832
87	0.191618	0.184121	0.176917	0.169995	0.163344
88	0.208092	0.200099	0.192413	0.185022	0.177915
89	0.225770	0.217259	0.209068	0.201187	0.193602
90	0.244707	0.235657	0.226941	0.218547	0.210465
91	0.264958	0.255348	0.246087	0.237161	0.228559
92	0.286577	0.276387	0.266560	0.257083	0.247942
93	0.309611	0.298824	0.288413	0.278365	0.268667
94	0.334107	0.322706	0.311694	0.301058	0.290785
95	0.360104	0.348074	0.336446	0.325206	0.314342
96	0.387636	0.374965	0.362707	0.350850	0.339381
97	0.416732	0.403408	0.390511	0.378025	0.365939
98	0.447410	0.433426	0.419880	0.406757	0.394044
99	0.479679	0.465031	0.450832	0.437066	0.423720
100	0.513538	0.498226	0.483371	0.468959	0.454977
101	0.548975	0.533002	0.517494	0.502437	0.487818
102	0.585965	0.569338	0.553183	0.537486	0.522234
103	0.624468	0.607199	0.590407	0.574079	0.558203
104	0.664430	0.646534	0.629121	0.612177	0.595689
105	0.705779	0.687279	0.669265	0.651722	0.634640
106	0.748427	0.729350	0.710760	0.692644	0.674989
107	0.792269	0.772648	0.753512	0.734851	0.716652
108	0.837180	0.817052	0.797408	0.778237	0.759526
109	0.883016	0.862425	0.842315	0.822674	0.803491
110	0.929613	0.908610	0.888081	0.868017	0.848405

Table 4.42b Predicted forces of mortality for male immediate annuitants, duration 1+ years, 1978-94
Based on an analysis of 1958-94 experience, model parameters estimated using weighted likelihood

age	1978	1982	1986	1990	1994
55	0.007886	0.007400	0.006943	0.006515	0.006113
56	0.008654	0.008126	0.007631	0.007166	0.006729
57	0.009501	0.008929	0.008391	0.007885	0.007410
58	0.010437	0.009815	0.009230	0.008680	0.008163
59	0.011468	0.010793	0.010157	0.009559	0.008997
60	0.012606	0.011872	0.011182	0.010531	0.009919
61	0.013861	0.013064	0.012313	0.011605	0.010938
62	0.015244	0.014379	0.013562	0.012792	0.012066
63	0.016770	0.015829	0.014942	0.014104	0.013313
64	0.018451	0.017429	0.016464	0.015552	0.014691
65	0.020303	0.019193	0.018143	0.017151	0.016213
66	0.022342	0.021136	0.019995	0.018916	0.017895
67	0.024588	0.023278	0.022038	0.020864	0.019752
68	0.027058	0.025636	0.024288	0.023011	0.021801
69	0.029775	0.028231	0.026767	0.025378	0.024062
70	0.032762	0.031086	0.029495	0.027986	0.026554
71	0.036043	0.034225	0.032498	0.030858	0.029301
72	0.039646	0.037673	0.035799	0.034018	0.032325
73	0.043598	0.041459	0.039426	0.037492	0.035653
74	0.047931	0.045614	0.043408	0.041310	0.039313
75	0.052677	0.050168	0.047778	0.045502	0.043334
76	0.057872	0.055156	0.052568	0.050101	0.047749
77	0.063554	0.060616	0.057814	0.055142	0.052593
78	0.069761	0.066586	0.063555	0.060662	0.057901
79	0.076537	0.073107	0.069831	0.066702	0.063713
80	0.083924	0.080223	0.076685	0.073303	0.070071
81	0.091971	0.087980	0.084162	0.080511	0.077017
82	0.100725	0.096426	0.092310	0.088370	0.084599
83	0.110238	0.105611	0.101178	0.096932	0.092863
84	0.120562	0.115587	0.110818	0.106246	0.101862
85	0.131751	0.126409	0.121283	0.116365	0.111647
86	0.143863	0.138132	0.132629	0.127345	0.122272
87	0.156953	0.150812	0.144912	0.139242	0.133794
88	0.171081	0.164509	0.158190	0.152114	0.146271
89	0.186304	0.179280	0.172522	0.166018	0.159759
90	0.202681	0.195184	0.187966	0.181014	0.174319
91	0.220269	0.212280	0.204581	0.197160	0.190009
92	0.239126	0.230624	0.222424	0.214516	0.206889
93	0.259307	0.250272	0.241553	0.233137	0.225015
94	0.280862	0.271278	0.262021	0.253080	0.244444
95	0.303841	0.293690	0.283879	0.274395	0.265229
96	0.328287	0.317555	0.307174	0.297133	0.287420
97	0.354239	0.342913	0.331950	0.321336	0.311063
98	0.381729	0.369799	0.358241	0.347045	0.336198
99	0.410781	0.398238	0.386078	0.374289	0.362860
100	0.441411	0.428250	0.415482	0.403094	0.391075
101	0.473625	0.459844	0.446464	0.433474	0.420862
102	0.507416	0.493017	0.479028	0.465435	0.452228
103	0.542766	0.527756	0.513161	0.498970	0.485171
104	0.579645	0.564033	0.548842	0.534060	0.519676
105	0.618005	0.601806	0.586032	0.570671	0.555713
106	0.657784	0.641018	0.624679	0.608757	0.593240
107	0.698903	0.681594	0.664714	0.648252	0.632197
108	0.741265	0.723444	0.706050	0.689075	0.672508
109	0.784755	0.766456	0.748583	0.731128	0.714079
110	0.829237	0.810501	0.792189	0.774291	0.756797

Table 4.42c Predicted forces of mortality for male immediate annuitants, duration 1+ years, 1998-14
Based on an analysis of 1958-94 experience, model parameters estimated using weighted likelihood

age	1998	2002	2006	2010	2014
55	0.005736	0.005383	0.005051	0.004739	0.004447
56	0.006319	0.005933	0.005572	0.005232	0.004913
57	0.006963	0.006544	0.006149	0.005779	0.005430
58	0.007677	0.007220	0.006789	0.006385	0.006005
59	0.008467	0.007969	0.007499	0.007058	0.006642
60	0.009342	0.008798	0.008286	0.007804	0.007350
61	0.010310	0.009717	0.009159	0.008632	0.008136
62	0.011381	0.010735	0.010125	0.009550	0.009008
63	0.012566	0.011862	0.011196	0.010569	0.009976
64	0.013877	0.013109	0.012383	0.011697	0.011049
65	0.015327	0.014489	0.013697	0.012948	0.012240
66	0.016929	0.016016	0.015151	0.014333	0.013560
67	0.018700	0.017704	0.016761	0.015868	0.015022
68	0.020655	0.019569	0.018541	0.017566	0.016643
69	0.022814	0.021631	0.020509	0.019445	0.018437
70	0.025196	0.023907	0.022684	0.021523	0.020422
71	0.027822	0.026418	0.025085	0.023819	0.022618
72	0.030717	0.029188	0.027736	0.026356	0.025045
73	0.033904	0.032241	0.030660	0.029156	0.027726
74	0.037412	0.035603	0.033882	0.032244	0.030685
75	0.041270	0.039304	0.037431	0.035648	0.033950
76	0.045508	0.043372	0.041337	0.039397	0.037548
77	0.050162	0.047843	0.045631	0.043522	0.041510
78	0.055265	0.052750	0.050349	0.048057	0.045870
79	0.060858	0.058131	0.055526	0.053038	0.050662
80	0.066980	0.064027	0.061203	0.058504	0.055924
81	0.073675	0.070478	0.067420	0.064495	0.061696
82	0.080988	0.077531	0.074222	0.071054	0.068021
83	0.088966	0.085232	0.081654	0.078227	0.074944
84	0.097659	0.093629	0.089766	0.086062	0.082511
85	0.107119	0.102776	0.098608	0.094610	0.090773
86	0.117401	0.112724	0.108234	0.103922	0.099782
87	0.128560	0.123530	0.118697	0.114053	0.109590
88	0.140652	0.135249	0.130054	0.125058	0.120255
89	0.153736	0.147941	0.142364	0.136997	0.131832
90	0.167872	0.161663	0.155684	0.149926	0.144381
91	0.183118	0.176476	0.170075	0.163906	0.157962
92	0.199533	0.192438	0.185596	0.178997	0.172633
93	0.217176	0.209609	0.202307	0.195258	0.188456
94	0.236103	0.228046	0.220264	0.212748	0.205488
95	0.256368	0.247804	0.239525	0.231524	0.223789
96	0.278024	0.268936	0.260144	0.251640	0.243414
97	0.301117	0.291490	0.282170	0.273149	0.264416
98	0.325691	0.315511	0.305650	0.296098	0.286844
99	0.351780	0.341038	0.330625	0.320529	0.310742
100	0.379415	0.368102	0.357127	0.346479	0.336148
101	0.408616	0.396727	0.385184	0.373977	0.363096
102	0.439396	0.426928	0.414813	0.403043	0.391606
103	0.471754	0.458708	0.446022	0.433687	0.421694
104	0.505679	0.492059	0.478807	0.465911	0.453362
105	0.541147	0.526963	0.513151	0.499700	0.486602
106	0.578119	0.563384	0.549024	0.535030	0.521392
107	0.616541	0.601271	0.586380	0.571858	0.557696
108	0.656340	0.640560	0.625159	0.610129	0.595460
109	0.697428	0.681165	0.665282	0.649769	0.634617
110	0.739698	0.722986	0.706651	0.690685	0.675080

4.3 Modelling of immediate annuitants' select mortality experience

In general, immediate annuitants exercise a strong degree of self-selection, resulting in lower mortality rates during the initial years following selection. Therefore, in addition to modelling the ultimate experience, it is necessary to model the select experience for immediate annuitants. In this section, the method of modelling select mortality data using GLMs, applied by Haberman and Renshaw (1996), Renshaw and Haberman (1997), is applied to the mortality experience of female immediate annuitants as an illustration of a complete GLM modelling procedure for mortality data that includes select data.

Renshaw and Haberman (1997) suggested a two-stage modelling structure that firstly involves modelling the ultimate experience by any suitable method, and then modelling the log crude mortality ratios for the select experience relative to the ultimate experience. As modelling select experiences is not the focus of this study, only a brief description of the modelling procedure is given in Section 4.3.1.

4.3.1 Modelling select mortality: Methodology

For a given observation period, let ${}^{d}a_{x,t}$ be the observed number of deaths at age x, time t and duration d, and ${}^{d}R^{c}_{x,t}$ the matching central exposed-to-risk at age x, time t and duration d.

For an individual select duration d relative to the corresponding ultimate duration denoted d_+ at age x and time t, define the statistic

$$ {}^{d}z_{x,t} = \log\left\{\frac{{}^{d}a_{x,t}}{{}^{d}R^{c}_{x,t}}\right\} - \log\left\{\frac{{}^{d+}a_{x,t}}{{}^{d+}R^{c}_{x,t}}\right\}. \tag{4.46} $$

The statistic defined by expression (4.46) is simply the log of the ratio of the crude mortality rate at select duration d, to the crude mortality rate at ultimate duration d_+, at age x in time t and hence provides an estimate of the log of the ratio of the corresponding forces of mortality.

The underlying patterns are then modelled by targetting

$$E\left[{}^{d}Z_{x,t}\right] \approx \log\left\{\frac{{}^{d}\mu_{x+\frac{1}{2},t}}{{}^{d+}\mu_{x+\frac{1}{2},t}}\right\} = {}^{d}\eta_{x,t}; \tag{4.47}$$

with weights

$${}^{d}w_{x,t} = \frac{{}^{d}a_{x,t}\,{}^{d+}a_{x,t}}{{}^{d}a_{x,t} + {}^{d+}a_{x,t}}, \tag{4.48}$$

where ${}^{d}\mu_{x,t}$ and ${}^{d+}\mu_{x,t}$ denote the force of mortality at age x, time t, for select duration d and ultimate duration d_{+} respectively.

The resulting forces of mortality for the individual select durations are given by:

$${}^{d}\mu_{x+\frac{1}{2},t} = {}^{d+}\mu_{x+\frac{1}{2},t} \exp\left({}^{d}\eta_{x,t}\right). \tag{4.49}$$

For CMI data, $a_{x,t} / R^{c}_{x,t}$ provides an estimate of μ_x rather than of $\mu_{x+1/2}$, and hence the resulting forces of mortality are given by:

$${}^{d}\mu_{x,t} = {}^{d+}\mu_{x,t} \exp\left({}^{d}\eta_{x,t}\right). \tag{4.50}$$

The force of mortality at age x, in each time period t and select duration d can therefore be considered as a proportion of the corresponding force of mortality at age x and ultimate duration d_{+}, where the adjustment factor is given by $\exp\left({}^{d}\eta_{x,t}\right)$.

Assuming that the changes in mortality due to time are the same for select and ultimate lives, and that these are adequately represented in the modelling of the ultimate experience, the suffix t can be dropped from ${}^{d}\eta_{x,t}$ and the linear predictor denoted ${}^{d}\eta_x$. The responses ${}^{d}z_x$ are then modelled from the combined experience over the observation period, for each select duration d = 0, 1, . . . and the ultimate duration d_+. This is particularly useful when modelling the select experience for immediate annuitants where the data are scanty.

As an example of the application of the method, the select mortality experience for female immediate annuitants over the period 1974 to 1994 was modelled at select durations 0 and 1 to 4 combined, relative to the ultimate experience at duration d_+=5+ years. The results of analysing the ultimate experience and the select experience are given in Sections 4.3.2 and 4.3.3 respectively.

4.3.2 Modelling of the 1974-94 female immediate annuitants' mortality experience, duration 5+ years

The 6-parameter log-link model structure

$$\mu_{xt} = \exp\left[\beta_0 + \sum_{j=1}^{3} \beta_j L_j(x') + (\alpha_1 + \gamma_{11}x')t'\right]$$

applied to the immediate annuitants' mortality experience at duration 1 year and over, was fitted to the female immediate annuitants' mortality experience at duration 5 years and over, over the calendar-year period 1974 to 1994 for the age range 65-100 years, giving 720 data cells. The transformed calendar year t' and age x' are:

$$t' = \frac{t-1984}{10}, \text{ and } x' = \frac{x-82.5}{17.5}.$$

The results of graduating the ultimate mortality experience at duration 5 years and over, would generally be expected to be similar to the results of graduating ultimate mortality experience at duration 1 year and over, over the same age range. This is because the ultimate experience at duration 5 years and over is a subset of the ultimate experience at duration 1 year and over, with the bulk of the data being at duration 5 years and over. Thus the same features would be exhibited in the graduated forces of mortality for the two experiences. This is in fact borne out by the results obtained from modelling the female annuitants' mortality experience, even though the age range is not the same (ages 65-100 at d5+ years and 65-95 at d1+ years).

The parameter estimates obtained from the analysis are given in Table 4.43. As for the model of the female annuitants' experience at duration 1+ years, one parameter β_2, is not statistically significant.

The corresponding time reduction factor is:

$$RF(x,n) = \exp\left[\frac{n}{10}\left\{-0.170246 + .061594\left(\frac{x-82.5}{17.5}\right)\right\}\right] \tag{4.51}$$

or

$$RF(x,n) = \exp[(-0.046062 + 0.00035x)n] \tag{4.52}$$

Table 4.43
Female immediate annuitants, Duration 5 years and over,
6-parameter log-link model based on the 1974-94 experience (deviance = 1060.18 on 714 d.f., ϕ = 1.500)

parameter	estimate	standard error	t-value
β_0	-2.684160	0.0110	-243.8409
β_1	1.819016	0.0246	73.9200
β_2	-0.022700	0.0250	-0.9087
β_3	-0.103723	0.0240	-4.3168
α_1	-0.170246	0.0135	-12.5707
γ_{11}	0.061594	0.0293	2.1008

4.3.3 Modelling of the female immediate annuitants' select experience: 1974 to 1994

For each of the select durations $d = 0$ and 1-4 years combined, and ultimate duration $d_+ = 5+$ years, the data analysed were the 1974-94 combined mortality experiences, at individual ages 65-85 years. There was little data at ages above 85 years for the select durations, particularly at duration 0.

In line with Renshaw and Haberman (1997), the ${}^{d}z_x$ responses for the select durations are determined and the response plots against age x are shown in Figure 4.26. There is no clear pattern that can be discerned from the plots although models of the form:

$$ {}^{d}\eta_x = \theta_d \tag{4.53} $$

or

$$ {}^{d}\eta_x = \theta_d + \beta_d x \tag{4.54} $$

for each duration d would seem reasonable. The best fitting model for each of the select durations was of the form of expression (4.53). For each select duration d, the parameter β_d was not statistically significant. The model represents a horizontal straight-line structure and hence a constant adjustment factor at all ages x for each duration d.

Table 4.44 shows the linear predictors and adjustment factors for each select duration d.

Table 4.44
Female immediate annuitants, 1974-94 combined experience: Linear predictors and adjustment factors based on ultimate rates at duration 5 years and over

Duration d	Linear Predictor, θ_d	Adjustment factor: Exp(θ_d)
0	-0.3979885	0.67166975
1-4	-0.1605232	0.85169806

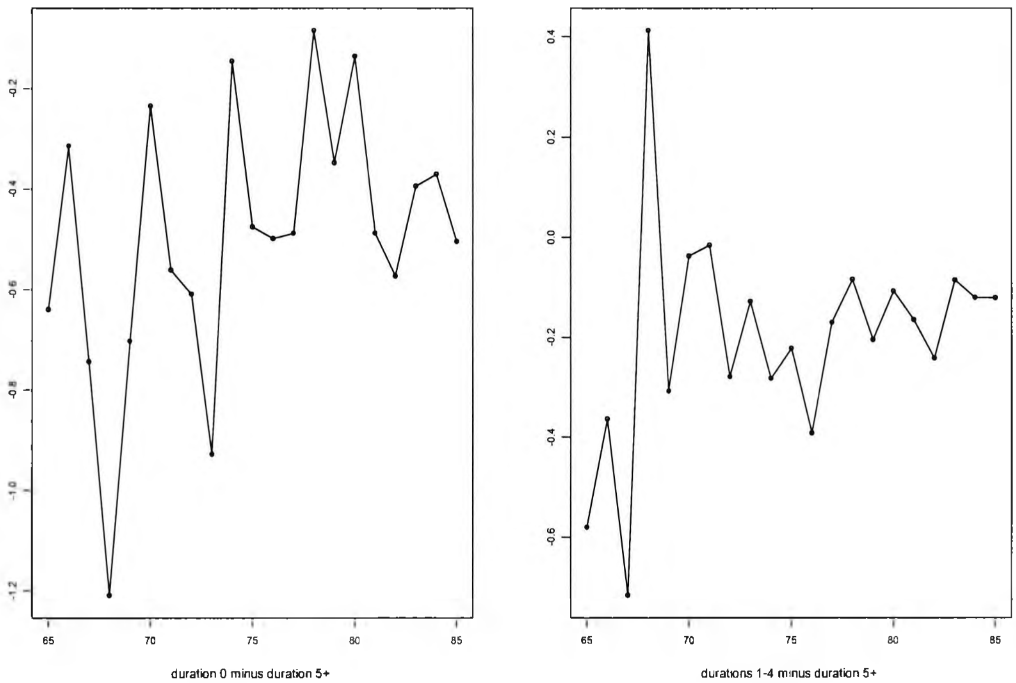


Figure 4.18 Female immediate annuitants, 1974-94 mortality experience: differences of log crude mortality rates plotted against age

The forces of mortality for each select duration d are obtained by applying the relevant adjustment factor to the predicted forces of mortality at ultimate duration 5+ years, based on modelling the 1974 to 1994 mortality experience with respect to both age and time as detailed in the preceding sections. For a given select duration d = 0 or 1-4 combined, the predicted force of mortality at age x in calendar year t is given by:

$$^{d}\mu_{x,t} = {}^{5+}\mu_{x,t} \exp(\theta_d). \tag{4.55}$$

The form of the model assumes that changes in mortality due to both age and time are the same for ultimate and select lives, and that these changes are adequately represented in the modelling of the ultimate experience at duration 5 years and over.

The constraints:

$^{0}\eta_x \leq {}^{d}\eta_x \leq 0$ for d = 1 to 4 combined, are satisfied, ensuring the ordering of forces of mortality with respect to duration d at each age x.

Chapter 5

Modelling Life Office Pensioners' Mortality Experience and Comparison of Mortality Improvement Models

5.1 Introduction

The Continuous Mortality Investigations Committee applies the same mortality improvement model for both annuitants and life office pensioners. The pensioners' mortality experience was therefore modelled using the same GLM procedures applied to immediate annuitants' mortality data, as detailed in Chapter 4, in order to make comparisons of improvement factors arising from the models derived.

The pensioners' data based on lives, were modelled for males and females separately, at individual calendar years t = 1983 to 1996. As with the immediate annuitants' experience, the CMI Bureau provided the data analysed and these pertain to pensioners who retired on or after normal retirement age. Data for the period prior to 1983 were not available for study. For each calendar year t, the data as provided by the CMIB are tabulated by age x nearest birthday and by curtate policy duration $d = 0,1,2,\cdots,9$ or 10 years and over, separately for males and females. The tabulated data consist of the number of deaths, a_{xt} occurring at each individual age x in calendar year t, and matching initial exposed-to-risk, R_{xt}.

A summary of the data provided is given in Table 5.1. The central exposed-to-risk, R^c_{xt} has been calculated as: $R^c_{xt} = R_{xt} - a_{xt}/2$. That is, the difference between the initial exposed-to-risk and half the number of deaths at age x in calendar year t.

Table 5.1
Life Office Pensioners' Mortality Experience: 1983-1996

	Deaths	Central Exposed-To-risk	Crude Death Rate
Male	253041	4009642	0.0631
Female	45706	1313868	0.0348
Total	**298747**	**5323510**	**0.0561**

In line with the current CMI practice, the pensioners' mortality experience was modelled for all durations combined and hence only the aggregate experience is presented in Table 5.1. In contrast to the immediate annuitants' mortality experience where the females' experience is the larger experience, here the male pensioners' experience is the larger experience, constituting just over 75% of the total pensioners' experience in terms of the exposed-to-risk.

For each experience (male or female), the crude death rate was calculated as the total number of deaths divided by the corresponding central exposed-to-risk. It is interesting to note that the crude death rate for pensioners is lower than the crude death rate for immediate annuitants at duration 5+ years, for both males and females (0.0763 and 0.0623 respectively for immediate annuitants). It is difficult to draw inferences from this observation since the crude death rate does not take into account the age of an individual. However, it is worth noting that the annuitants' experience is an ageing experience and that the pensioners' data is based only on the more recent experience, so that the heavier mortality of the earlier years is excluded (assuming there has been a decline in mortality over time).

As with the annuitants' experience, the force of mortality μ_{xt} at age x in calendar year t, was modelled using formulae of the form:

$$\mu_{xt} = \exp\left[\sum_{j=0}^{s} \beta_j L_j(x') + \sum_{i=1}^{r} \alpha_i t'^i + \sum_{i=1}^{r}\sum_{j=1}^{s} \gamma_{ij} L_j(x') t'^i\right], \tag{5.1}$$

with x', t' and $L_j(x')$ the transformed age, transformed calendar year and Legendre polynomials respectively, as detailed in Chapter 4.

5.2 Trend analysis of the male life office pensioners' mortality experience

The male life office pensioners' mortality experience was modelled over the age range 60 to 100 years, so that the scaled age x' is given by:

$$x' = \frac{x-80}{20};$$

and the scaled calendar year t' is

$$t' = \frac{t-1989.5}{6.5}.$$

The data analysed consisted of 574 data cells.

From an analysis of deviance shown in Tables 5.2 and 5.3 and the statistical significance of additional parameters introduced to the model formula, the best fitting model was determined to be a 7-parameter model of the form:

$$\mu_{xt} = \exp\left[\sum_{j=0}^{3}\{\beta_j L_j(x')\} + \{(\alpha_1 + \gamma_{11}x')t' + \alpha_2 t'^2\}\right]. \quad (5.2)$$

This is a $GM_x(0,4)$ term in combination with a trend adjustment term that is quadratic in time t on the log scale.

The estimated parameters are given in Table 5.4.

Table 5.2
Male life office pensioners, analysis of the 1983-96 experience
Deviances for polynomial predictors of degree *r* and *s*

	$r = 0$	$r = 1$	$r = 2$	$r = 3$
$s = 0$	105168.63	105158.32	105059.92	105052.67
$s = 1$	3472.58	1808.00	1788.47	1784.16
$s = 2$	2607.44	1071.36	1058.40	1053.26
$s = 3$	2607.00	1061.05	1048.20	1043.12
$s = 4$	2605.87	1060.85	1047.93	1042.85

Table 5.3
Male life office pensioners, analysis of the 1983-96 experience
Deviance profile (terms added sequentially 1st to last)

Parameter	Deviance	Degrees of freedom	Difference in deviance
β_0	105168.63	573	
β_1	3472.58	572	101696.05
α_1	1808.00	571	1664.58
β_2	1071.36	570	736.64
α_2	1058.40	569	12.96
β_3	1048.20	568	10.20
γ_{11}	937.14	567	111.06
γ_{12}	933.38	566	3.76

Table 5.4
Male life office pensioners, 7-parameter log-link model based on the 1983-96 experience (ϕ = 1.6600)

parameter	estimate	standard error	*t*-value
β_0	-2.524362	0.0056	-447.9598
β_1	1.822949	0.0140	129.9956
α_1	-0.127455	0.0044	-28.7098
β_2	-0.264690	0.0122	-21.6416
α_2	-0.031453	0.0078	-4.0422
β_3	-0.047932	0.0153	-3.1253
γ_{11}	0.105513	0.0129	8.1535

Renshaw and Hatzopoulos (1996) modelled the male life office pensioners' mortality experience over the period 1983-1990 and individual ages 60 to 95, by targetting the probability of death, q_{xt}. The number of pension policies was modelled as a binomial response variable of a generalised linear model with possible over-dispersion, in conjunction with the complementary log-log link function:

$$\log\{-\log(1-q_{xt})\} = \eta_{xt}. \quad (5.3)$$

Using the result that $\mu_{xt} = -\log p_{xt}$, assuming a constant force of mortality between ages x and $x + u$, where $0 \le u \le 1$, then modelling the complementary log-log link function (5.3) would be approximately the same as modelling $\log \mu_{xt}$, since $p_{xt} = 1 - q_{xt}$.

The formula adopted in that study was:

$$\log\{-\log(1-q_{xt})\} = \exp\left[\sum_{j=0}^{3}\{\beta_j L_j(x')\} + \{(\alpha_1 + \gamma_{11}x')t' + \alpha_2 t'^2 + \alpha_3 t'^3\}\right]. \quad (5.4)$$

The difference between equations (5.2) and (5.4) is the one additional term in equation (5.4) involving t'^3, ie the trend adjustment term in (5.4) is a cubic in time t on the log scale. In this study, no attempt was made to fit the same 8-parameter model since the analysis of the annuitants' data had shown that higher order terms in t result in unreasonable projections.

Based on the 7-parameter model described by equation (5.2), the corresponding mortality improvement model for a life attaining exact age x in calendar year t is a quadratic in time t on the log scale given by:

$$RF(x, t-t_0) = \frac{\mu_{xt}}{\mu_{xt_0}} = \exp\left\{\sum_{i=1}^{2}\alpha_i\left(t'^i - t_0'^i\right) + \gamma_{11}x'(t' - t_0')\right\}, \quad (5.5)$$

where t_0 is an appropriate origin (the base calendar year). The model (5.5) may be re-expressed as:

$$RF(x, t-t_0) = \frac{\mu_{xt}}{\mu_{xt_0}} = \exp\left\{\frac{(t-t_0)}{w_t}(\alpha_1 + \gamma_{11}x') + \frac{\alpha_2}{w_t^2}\left[(t-c_t)^2 - (t_0 - c_t)^2\right]\right\}, \quad (5.6)$$

with $c_t = 1989.5$, $w_t = 6.5$ and $x' = \dfrac{x-80}{20}$. As before, setting $n = t - t_0$, the improvement model (5.6) or (5.7) is the reduction factor for an ultimate life attaining exact age x at time n, where n is measured in years from the base calendar year t_0.

Figure 5.1 shows predicted forces of mortality at 5-year age intervals for x = 55 to 110 years and t = 1983 to 2016 together with mortality improvement factors for the period t = 1993 to 2016, based on the 7-parameter model. The predicted forces of mortality are plotted on the log scale. In deriving the mortality improvement factors, 1992 was taken as the base calendar year in order to be consistent with the current CMI improvement model.

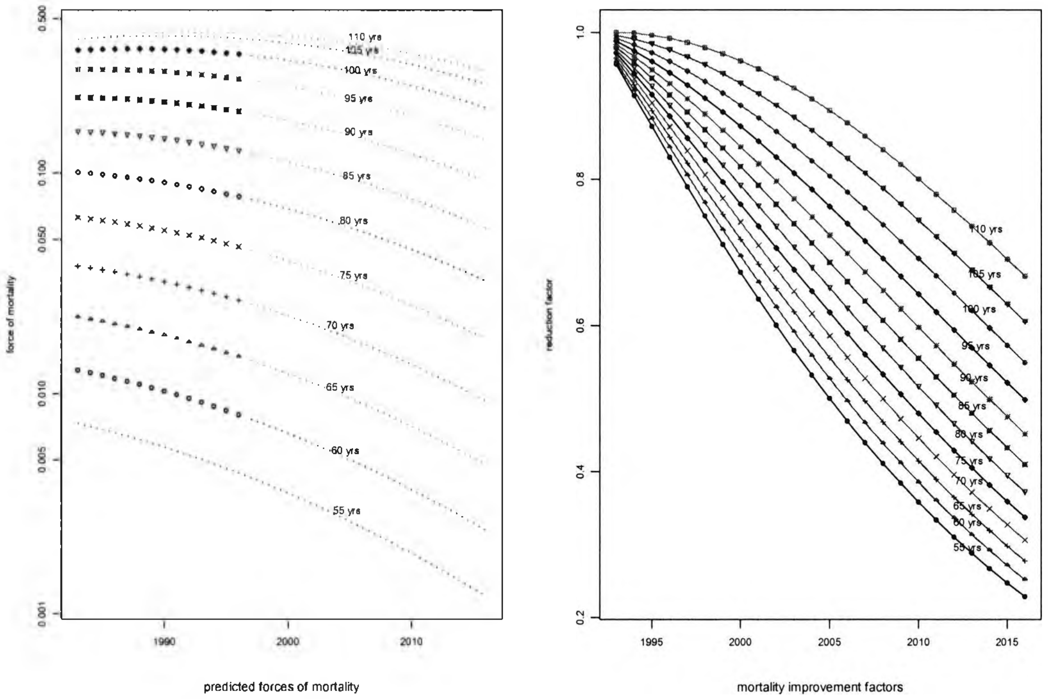


Figure 5.1 Male life office pensioners, predicted forces of mortality: 1983-2016, and mortality improvement factors: 1993-2016; based on a 7-parameter log-link model with $r = 2$, $s = 3$ and γ_{11}

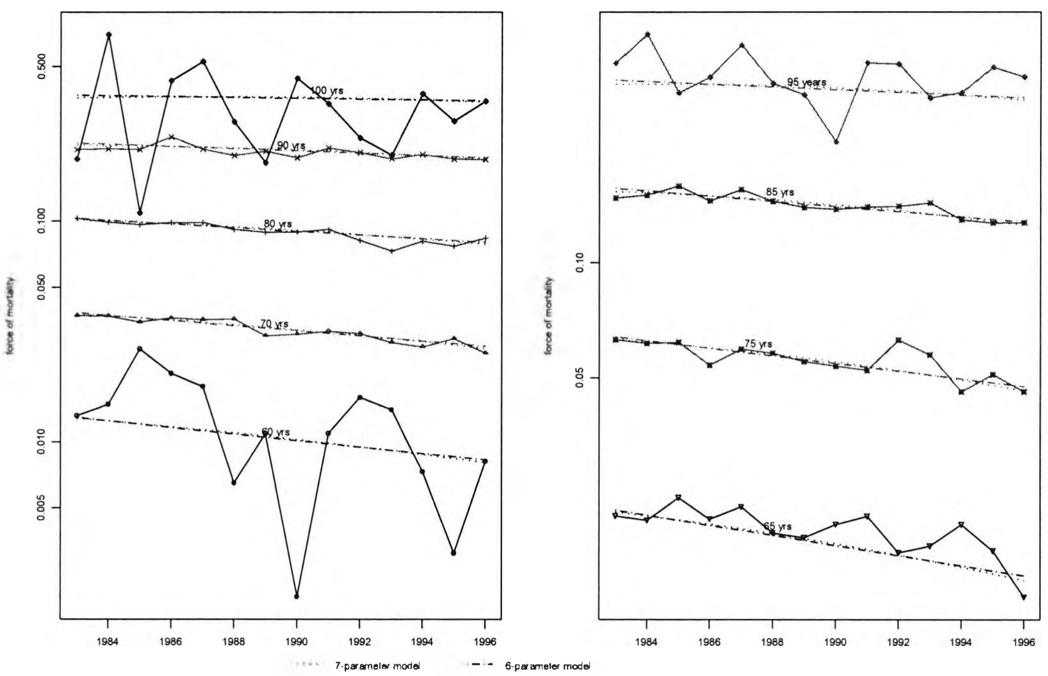


Figure 5.2 Analysis of male life office pensioners' mortality experience: 1983-96, crude mortality rates and predicted forces of mortality based on a 7-parameter and a 6-parameter log-link model

Although the predicted forces of mortality progress gradually with respect to age, the rates tend to progress rapidly with time, a feature that would not be desirable in determining prices and reserves. The rapid improvement in mortality over time is not surprising since the mortality improvement formula (5.5) or (5.6) includes the exponential of negative terms in t^2 thereby giving rise to a rapid decrease in $RF(x,n)$ as t increases.

In view of these results and the conclusions drawn in Chapter 4 in the analyses of immediate annuitants' experiences, a six-parameter model excluding the term involving t'^2 was then fitted to the data. Thus the form of the revised model was the same as the model formula adopted for immediate annuitants, that is:

$$\mu_{xt} = \exp\left[\beta_0 + \sum_{j=1}^{3}\{\beta_j L_j(x')\} + \{(\alpha_1 + \gamma_{11}x')t'\}\right]. \qquad (5.7)$$

Details of the parameter estimates for the 6-parameter model are given in Table 5.5.

Figure 5.2 shows crude mortality rates and graduated forces of mortality for male pensioners based on the 7-parameter model (5.2) and the 6-parameter model (5.7), plotted against calendar year t = 1983 to 1996 on the log scale. The rates are shown at 10-year age intervals from x = 60 to 100 years and from x = 65 to 95 years.

Although the 7-parameter model has a trend adjustment term that is quadratic in time t on the log scale, while the 6-parameter model is linear in t on the log scale, it is difficult to discern any differences between the graduated rates from the two models based on a visual inspection of the graph. A detailed examination of the predicted forces of mortality reveals that in general, the rates based on the 7-parameter model are lower than the rates based on the 6-parameter model. However, for the graduated rates the differences are small. Both models generally fit the data adequately as can be seen from the results of statistical tests of graduation shown in Tables 5.6 and 5.7.

Table 5.5
Male life office pensioners, 6-parameter log-link model based on the 1983-96 experience
(deviance = 964.33 on 568 d.f., ϕ = 1.7019)

parameter	estimate	standard error	t-value
β_0	-2.535627	0.0050	-510.5671
β_1	1.823223	0.0142	128.3969
β_2	-0.264853	0.0124	-21.3905
β_3	-0.047458	0.0155	-3.0559
α_1	-0.125691	0.0045	-28.2352
γ_{11}	0.096744	0.0129	7.5275

Table 5.6
Male life office pensioners, analysis of the 1983-96 mortality experience
Comparison of p-values based on 2 models

Statistical test	7-parameter model	6-parameter model
Chi-square	0.4921	0.4921
Cumulative deviations	0.5152	0.5048
Individual standardised deviations	0.9288	0.8653
Grouping of signs of deviations	0.9894	0.9757
Signs of deviations	0.7369	0.9331

Table 5.7
Male life office pensioners, analysis of the 1983-96 experience
Comparison of the distribution of individual studentized deviance residuals

Range	**(-∞,-3)**	**(-3,-2)**	**(-2,-1)**	**(-1,0)**	**(0,1)**	**(1,2)**	**(2,3)**	**(3,∞)**
expected frequency	0.77	12.13	77.06	193.54	193.54	77.06	12.13	0.77
Observed frequency								
7-parameter model	1.00	12.00	72.00	203.00	187.00	77.00	14.00	1.00
6-parameter model	2.00	14.00	71.00	198.00	194.00	73.00	14.00	1.00

Predicted forces of mortality and mortality improvement factors for male pensioners, based on the 6-parameter log-link model, are shown in Figure 5.3. The projections have been made over a 20 calendar year period from 1997 to 2016, and the projected forces of mortality are presented at quadrennial periods in Tables 5.8 and 5.9. From the graph, it can be observed that the model has a reasonable shape for projections at all ages other than extreme old age above 105 years, where the projected forces of mortality are increasing with time rather than decreasing. The resulting mortality improvement factors at ages less than 106 years exhibit a gradual progression with time. For example, based on the 7-parameter model, the force of mortality for a life age 55 in 2016 would be expected to be 23% of the force of mortality applicable in 1992 (a decrease of 77%), whereas on the basis of the 6-parameter model, the corresponding percentage reduction is 40%.

The mortality improvement model for male pensioners is simplified to the form:

$$RF(x,n) = \exp\left[\frac{n}{w_t}\{\alpha_1 + \gamma_{11}x'\}\right] \tag{5.8}$$

with w_t = 6.5. For the particular parameter estimates given in Table 5.5, the improvement formula is:

$$RF(x,n) = \exp[(-0.078872 + 0.000744x)n]. \tag{5.9}$$

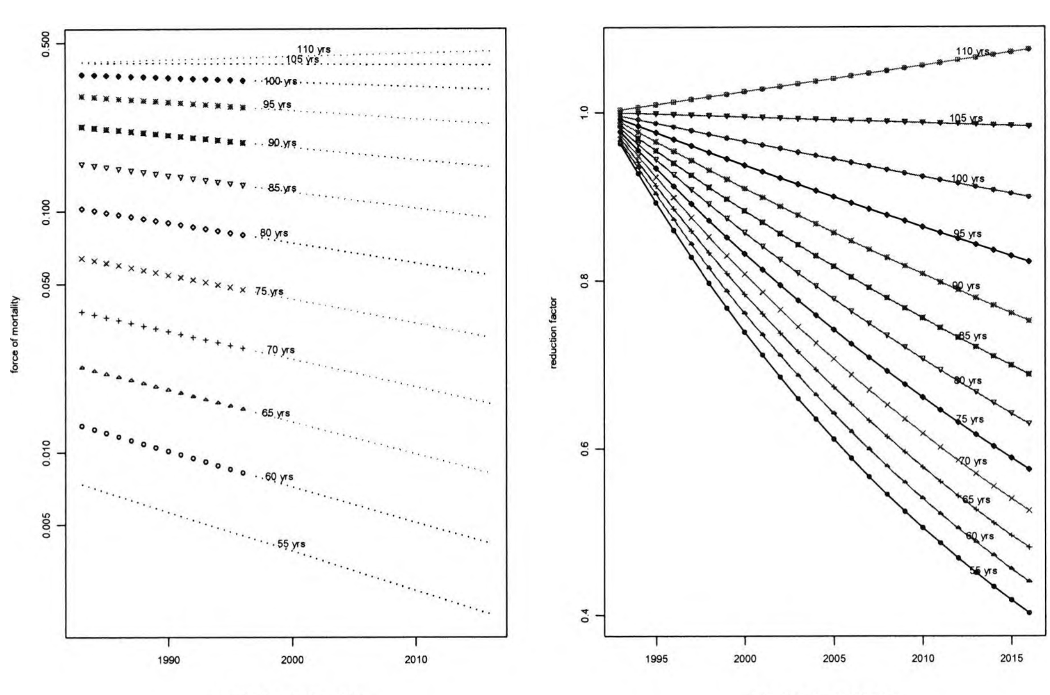


Figure 5.3 Male life office pensioners, predicted forces of mortality: 1983-2016, and mortality improvement factors: 1993-2016; based on a 6-parameter log-link model with $r = 1$, $s = 3$ and γ_{11}

Table 5.8
Male life office pensioners
Projected forces of mortality based on a 7-parameter log-link model

age	1996	2000	2004	2008	2012	2016
55	0.004349	0.003525	0.002789	0.002155	0.001626	0.001198
56	0.004912	0.003994	0.003171	0.002458	0.001861	0.001375
57	0.005549	0.004526	0.003605	0.002804	0.002129	0.001579
58	0.006268	0.005130	0.004099	0.003198	0.002437	0.001813
59	0.007081	0.005814	0.004661	0.003648	0.002789	0.002081
60	0.007999	0.006588	0.005299	0.004161	0.003191	0.002390
61	0.009033	0.007465	0.006023	0.004746	0.003651	0.002743
62	0.010199	0.008455	0.006845	0.005410	0.004176	0.003147
63	0.011510	0.009574	0.007775	0.006166	0.004775	0.003610
64	0.012985	0.010835	0.008828	0.007024	0.005457	0.004139
65	0.014640	0.012256	0.010019	0.007997	0.006233	0.004744
66	0.016496	0.013855	0.011362	0.009099	0.007115	0.005432
67	0.018574	0.015650	0.012877	0.010345	0.008115	0.006217
68	0.020896	0.017664	0.014581	0.011752	0.009249	0.007108
69	0.023487	0.019919	0.016496	0.013339	0.010532	0.008121
70	0.026374	0.022440	0.018644	0.015125	0.011981	0.009268
71	0.029583	0.025252	0.021048	0.017131	0.013615	0.010565
72	0.033143	0.028383	0.023735	0.019381	0.015453	0.012031
73	0.037085	0.031862	0.026731	0.021898	0.017516	0.013682
74	0.041439	0.035719	0.030064	0.024709	0.019829	0.015538
75	0.046238	0.039985	0.033764	0.027840	0.022414	0.017622
76	0.051513	0.044692	0.037861	0.031319	0.025298	0.019953
77	0.057297	0.049872	0.042387	0.035177	0.028506	0.022557
78	0.063621	0.055556	0.047371	0.039442	0.032066	0.025456
79	0.070516	0.061777	0.052847	0.044144	0.036006	0.028677
80	0.078009	0.068564	0.058844	0.049313	0.040353	0.032244
81	0.086128	0.075947	0.065392	0.054978	0.045135	0.036182
82	0.094895	0.083949	0.072517	0.061168	0.050380	0.040517
83	0.104329	0.092594	0.080245	0.067906	0.056112	0.045274
84	0.114441	0.101900	0.088597	0.075217	0.062355	0.050475
85	0.125240	0.111878	0.097589	0.083121	0.069131	0.056142
86	0.136725	0.122534	0.107232	0.091631	0.076457	0.062293
87	0.148886	0.133867	0.117530	0.100758	0.084346	0.068944
88	0.161706	0.145867	0.128482	0.110505	0.092805	0.076106
89	0.175155	0.158513	0.140074	0.120867	0.101838	0.083785
90	0.189194	0.171774	0.152287	0.131832	0.111438	0.091982
91	0.203769	0.185609	0.165087	0.143378	0.121592	0.100689
92	0.218816	0.199963	0.178433	0.155472	0.132277	0.109894
93	0.234255	0.214768	0.192266	0.168071	0.143461	0.119573
94	0.249993	0.229942	0.206521	0.181118	0.155101	0.129694
95	0.265925	0.245392	0.221113	0.194546	0.167142	0.140218
96	0.281931	0.261008	0.235949	0.208275	0.179519	0.151090
97	0.297879	0.276669	0.250920	0.222210	0.192153	0.162249
98	0.313626	0.292242	0.265905	0.236246	0.204955	0.173622
99	0.329017	0.307580	0.280772	0.250266	0.217823	0.185123
100	0.343891	0.322531	0.295377	0.264140	0.230647	0.196659
101	0.358081	0.336932	0.309568	0.277731	0.243303	0.208125
102	0.371415	0.350615	0.323187	0.290893	0.255661	0.219408
103	0.383722	0.363411	0.336072	0.303473	0.267585	0.230387
104	0.394834	0.375150	0.348056	0.315317	0.278933	0.240939
105	0.404589	0.385668	0.358978	0.326269	0.289560	0.250931
106	0.412833	0.394807	0.368680	0.336176	0.299323	0.260235
107	0.419429	0.402419	0.377010	0.344890	0.308080	0.268720
108	0.424255	0.408373	0.383832	0.352273	0.315697	0.276259
109	0.427208	0.412553	0.389021	0.358197	0.322050	0.282735
110	0.428212	0.414866	0.392475	0.362552	0.327026	0.288037

Table 5.9
Male life office pensioners
Projected forces of mortality based on a 6-parameter log-link model

age	1996	2000	2004	2008	2012	2016
55	0.004486	0.003854	0.003311	0.002845	0.002444	0.002100
56	0.005065	0.004365	0.003762	0.003242	0.002793	0.002407
57	0.005721	0.004944	0.004274	0.003694	0.003193	0.002759
58	0.006461	0.005601	0.004856	0.004209	0.003649	0.003163
59	0.007297	0.006345	0.005517	0.004797	0.004171	0.003626
60	0.008241	0.007186	0.006267	0.005465	0.004766	0.004156
61	0.009304	0.008138	0.007118	0.006226	0.005446	0.004763
62	0.010502	0.009213	0.008082	0.007090	0.006220	0.005457
63	0.011849	0.010426	0.009173	0.008072	0.007102	0.006249
64	0.013362	0.011792	0.010407	0.009184	0.008105	0.007153
65	0.015061	0.013331	0.011800	0.010445	0.009245	0.008183
66	0.016964	0.015060	0.013370	0.011870	0.010538	0.009356
67	0.019093	0.017002	0.015139	0.013480	0.012003	0.010688
68	0.021473	0.019177	0.017127	0.015296	0.013660	0.012200
69	0.024126	0.021611	0.019358	0.017340	0.015532	0.013913
70	0.027080	0.024329	0.021858	0.019638	0.017643	0.015851
71	0.030362	0.027360	0.024654	0.022216	0.020018	0.018039
72	0.034002	0.030731	0.027774	0.025102	0.022687	0.020504
73	0.038030	0.034474	0.031250	0.028327	0.025678	0.023277
74	0.042477	0.038620	0.035112	0.031923	0.029024	0.026388
75	0.047375	0.043201	0.039395	0.035924	0.032759	0.029873
76	0.052757	0.048252	0.044132	0.040364	0.036917	0.033765
77	0.058654	0.053805	0.049358	0.045278	0.041535	0.038102
78	0.065098	0.059895	0.055108	0.050703	0.046651	0.042922
79	0.072120	0.066553	0.061417	0.056676	0.052302	0.048265
80	0.079747	0.073812	0.068318	0.063233	0.058526	0.054170
81	0.088007	0.081699	0.075843	0.070408	0.065361	0.060677
82	0.096920	0.090242	0.084024	0.078234	0.072843	0.067824
83	0.106506	0.099463	0.092885	0.086743	0.081006	0.075649
84	0.116776	0.109378	0.102450	0.095960	0.089881	0.084188
85	0.127736	0.120002	0.112735	0.105909	0.099495	0.093471
86	0.139387	0.131337	0.123752	0.116605	0.109870	0.103525
87	0.151717	0.143381	0.135503	0.128058	0.121022	0.114372
88	0.164707	0.156121	0.147983	0.140269	0.132958	0.126027
89	0.178328	0.169536	0.161178	0.153232	0.145677	0.138495
90	0.192537	0.183591	0.175060	0.166926	0.159169	0.151773
91	0.207282	0.198240	0.189592	0.181321	0.173411	0.165846
92	0.222496	0.213424	0.204722	0.196375	0.188368	0.180688
93	0.238098	0.229071	0.220386	0.212030	0.203992	0.196258
94	0.253994	0.245093	0.236503	0.228215	0.220217	0.212500
95	0.270077	0.261389	0.252980	0.244842	0.236966	0.229343
96	0.286226	0.277844	0.269708	0.261810	0.254144	0.246702
97	0.302308	0.294331	0.286563	0.279001	0.271639	0.264470
98	0.318179	0.310706	0.303408	0.296282	0.289324	0.282528
99	0.333683	0.326817	0.320093	0.313507	0.307056	0.300738
100	0.348658	0.342503	0.336455	0.330515	0.324680	0.318947
101	0.362936	0.357591	0.352325	0.347136	0.342024	0.336987
102	0.376345	0.371908	0.367524	0.363190	0.358908	0.354677
103	0.388714	0.385276	0.381869	0.378491	0.375144	0.371826
104	0.399873	0.397518	0.395177	0.392849	0.390536	0.388236
105	0.409660	0.408461	0.407266	0.406075	0.404887	0.403702
106	0.417923	0.417942	0.417962	0.417982	0.418001	0.418021
107	0.424523	0.425809	0.427098	0.428391	0.429689	0.430990
108	0.429339	0.431923	0.434523	0.437138	0.439769	0.442415
109	0.432270	0.436168	0.440101	0.444070	0.448074	0.452114
110	0.433237	0.438447	0.443720	0.449056	0.454456	0.459921

5.3 Trend analysis of the female life office pensioners' mortality experience and comparison with male life office pensioners

The female life office pensioners' mortality experience was modelled over the age range x = 60 to 95 years, a total of 504 data cells. In modelling the experience, an analysis of the deviance suggested choosing s = 1 or 3, r = 2 and the age-specific trend adjustment term involving x and t, that is, γ_{11} (see Tables 5.10 and 5.11). Given the analyses of the immediate annuitants' experience in Chapter 4 and the male life office pensioners' experience in Section 5.2, the value of r was constrained to be 1. Assuming $r = 1$, $\gamma_{ij} = 0$ for $i, j \neq 1$, two models were then applied to the female pensioners' mortality data, a 4-parameter model with $s = 1$ and a 6-parameter model with $s = 3$. Hence the same 6-parameter model given by expression (5.6) was also applied to the female pensioners' experience. The 4-parameter model was of the form:

$$\mu_{xt} = \exp\left[\beta_0 + \beta_1 x' + \left(\alpha_1 + \gamma_{11} x'\right) t'\right]. \qquad (5.10)$$

Parameter estimates for the two models are presented in Tables 5.12 and 5.13. The 6-parameter model includes two parameters, β_2 and β_3, that could be excluded from the model on the basis of values of the associated t statistics. On the other hand all the parameters derived from the 4-parameter model defined by equation (5.10) are statistically significant. Statistical tests of graduation applied are all supportive of both models as can be seen in Tables 5.14 and 5.15.

Table 5.10
Female life office pensioners, analysis of the 1983-96 experience
Deviances for polynomial predictors of degree *r* and *s*

	r=0	r=1	r=2	r=3
s=0	33100.55	32904.21	32860.65	32851.49
s=1	882.79	778.84	755.53	750.75
s=2	877.81	776.01	752.97	748.18
s=3	871.15	770.98	748.24	743.41
s=4	870.99	770.68	748.06	743.24

Table 5.11
Female life office pensioners, analysis of the 1983-96 experience
Deviance profile (terms added sequentially 1st to last)

Parameter	Deviance	Degrees of freedom	Difference in deviance
β_0	33100.55	503	
β_1	882.79	502	32217.76
β_2	877.81	501	4.98
β_3	871.15	500	6.65
α_1	770.98	499	100.18
γ_{11}	760.76	498	10.22

Table 5.12
Female life office pensioners, 6-parameter log-link model based on the 1983-96 experience
(deviance = 760.76 on 498 d.f., ϕ = 1.5186)

parameter	estimate	standard error	t-value
β_0	-3.177138	0.0069	-462.4352
β_1	1.838014	0.0143	128.5235
β_2	-0.025876	0.0170	-1.5197
β_3	-0.036419	0.0189	-1.9252
α_1	-0.083041	0.0098	-8.4430
γ_{11}	0.055605	0.0215	2.5910

Table 5.13
Female life office pensioners, 4-parameter log-link model based on the 1983-96 experience
(deviance = 771.69 on 500 d.f., ϕ = 1.5248)

Parameter	estimate	standard error	t-value
β_0	-3.171249	0.0060	-526.4383
β_1	1.845645	0.0129	143.3259
α_1	-0.083762	0.0098	-8.5053
γ_{11}	0.045107	0.0208	2.1653

Table 5.14
Female life office pensioners, analysis of the 1983-96 mortality experience
Comparison of p-values based on 2 models

Statistical test	6-parameter model	4-parameter model
Chi-square	0.4916	0.4916
Cumulative deviations	0.5046	0.1244
Individual standardised deviations	0.6111	0.3546
Grouping of signs of deviations	0.9194	0.8117
Signs of deviations	0.5041	0.7552

Table 5.15
Female life office pensioners, analysis of the 1983-96 mortality experience
Comparison of the distribution of individual studentized deviance residuals from two log-link models

Range	(-∞,-3)	(-3,-2)	(-2,-1)	(-1,0)	(0,1)	(1,2)	(2,3)	(3,∞)
expected frequency	0.68	10.79	68.50	172.04	172.04	68.50	10.79	0.68
Observed frequency								
6-parameter model	1.00	7.00	77.00	175.00	165.00	71.00	8.00	0.00
4-parameter model	2.00	9.00	83.00	162.00	170.00	71.00	7.00	0.00

Figure 5.4 shows a comparative plot of forces of mortality predicted from the two models. Although there is little difference between the predicted rates within the range of ages over which the graduation was carried out (ages 60 to 95 years) there are marked differences outside this range of ages. The differences are more pronounced at the older ages where the 6-parameter model produces the lower forces of mortality. For example, the predicted force of mortality for a life age 110 in 1983 is approximately 0.68 based on the 6-parameter model and approximately 1.29 on the basis of the 4-parameter model. This characteristic is to be expected because the two additional parameters in the 6-parameter model are negative parameters both involving higher order terms in x, thereby having the general effect of producing lower forces of mortality at each age than those predicted on the basis of the 4-parameter model.

The preferred model is the model that produces lower forces of mortality and this again is the 6-parameter model defined by expression (5.7).

The mortality improvement model based on either model is given by:

$$RF(x,n) = \exp\left[\frac{n}{6.5}\left\{\alpha_1 + \gamma_{11}\left(\frac{x-77.5}{17.5}\right)\right\}\right] \tag{5.11}$$

that is, $x' = \dfrac{x-77.5}{17.5}$ and w_t = 6.5. Based on the parameter estimates for the 6-parameter model given in Table 5.12, the formula for the reduction factors is

$$RF(x,n) = \exp[(-0.050660 + 0.000489x)n]. \tag{5.12}$$

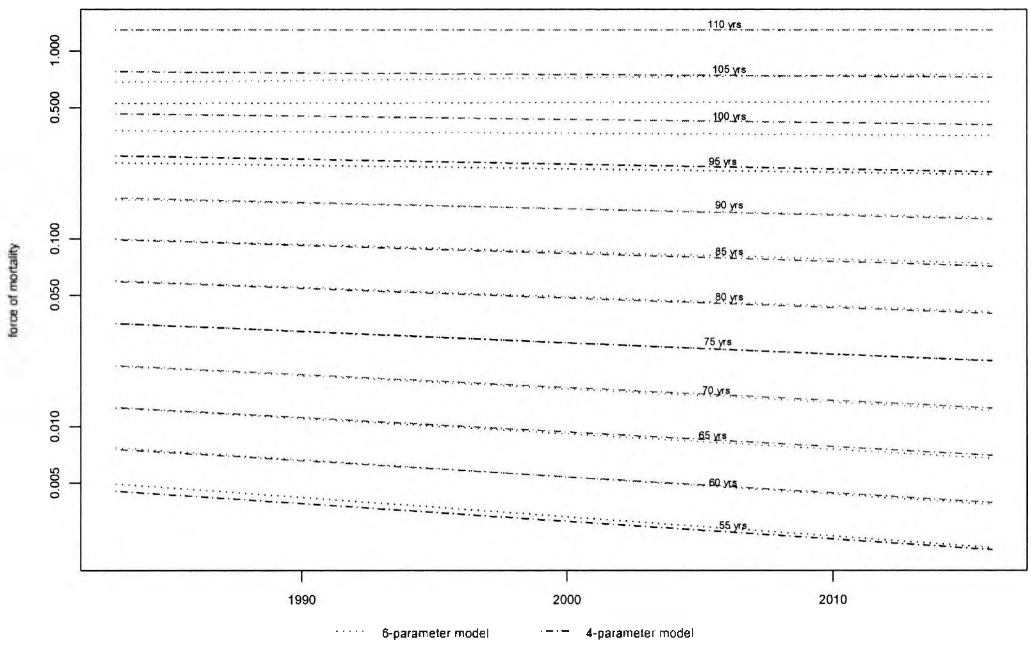


Figure 5.4 Female life office pensioners, comparison of predicted forces of mortality based on 6-parameter and 4-parameter log-link models: 1983-2016

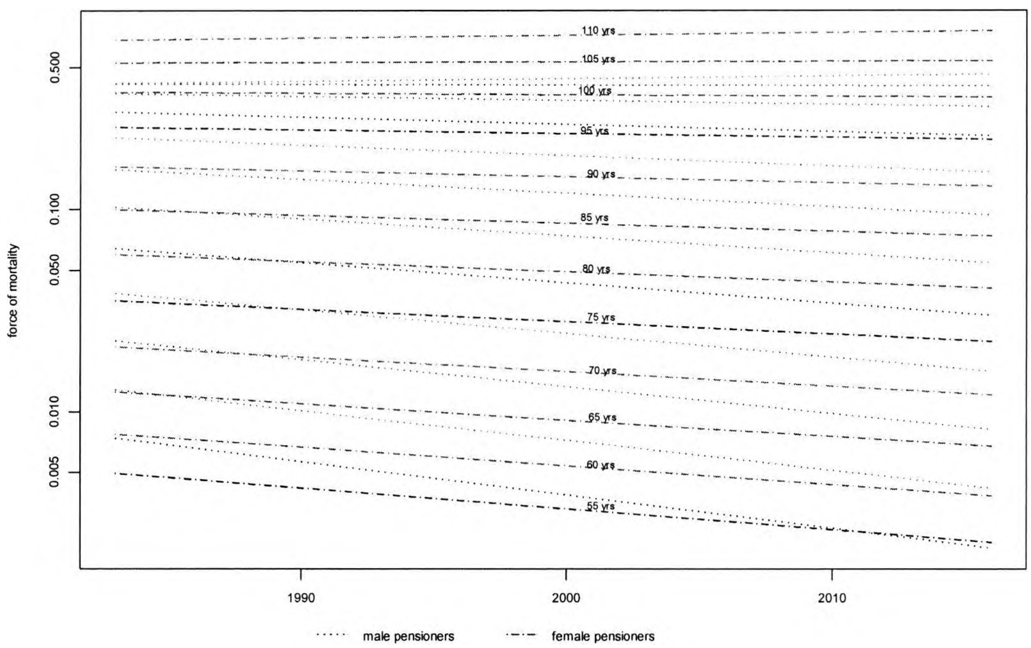


Figure 5.5 Comparison of predicted forces of mortality for male and female life office pensioners: 1983-2016; predictions based on 6-parameter log-link models with $r = 1$, $s = 3$ and γ_{11}

A comparative plot of male and female life office pensioners' predicted forces of mortality up to calendar year 2016 is shown in Figure 5.5. The predicted forces of mortality show a declining trend at all ages below 103 years for females and 106 years for males. At all ages the predicted rates for males appear to be improving at a faster rate than the females' predicted rates. For ages above 95 years, the predicted rates for male pensioners are lower than the predicted rates for females. Although this could be a reflection of the form of the underlying forces of mortality, it is also possible that this is the result of problems arising from extrapolating the rates.

It is worth noting that although the life office pensioners' mortality experience has been modelled over a period of 14 years only, the model derived still predicts reasonable future forces of mortality at most ages and for the 20 calendar-year period over which forces of mortality have been projected.

5.4 Comparison of CMI mortality improvement factors with log-link based mortality improvement factors for pensioners and annuitants

The most recent CMI mortality improvement model for pensioners and annuitants, used with mortality tables based on graduation of the mortality experience over the quadrennium 1991-94, was summarized in Chapter 3 and described in detail in CMI Report 17 (1999). The graduated rates of mortality at age x apply on average to lives attaining exact age x in calendar year 1992 (i.e. halfway through 1992) and hence time is measured in years from 1992. The same reduction factors apply for all pensioners' and annuitants' experiences, male and female, for data based on both lives and amounts.

The CMI improvement model is such that a given percentage of the total future decrease in mortality is assumed to occur in the first 20 years. The model is defined by:

$$RF(x,n) = \alpha(x) + \{1 - \alpha(x)\}\{1 - f_{20}(x)\}^{n/20} \tag{5.13}$$

where n is measured in years from the base calendar year (1992 in this case), $\alpha(x)$ is the limiting value of the reduction factor for a life age x in the base year, i.e.

$$\alpha(x) = \lim_{n \to \infty} RF(x,n) \ , \tag{5.14}$$

and $f_{20}(x)$ is the percentage of the total future decrease in mortality for a life age x, that will occur in the first 20 years.

The functions $\alpha(x)$ and $f_{20}(x)$ are defined as:

$$\alpha(x) = \begin{cases} 0.13 & x < 60 \\ 1 + \dfrac{0.87(x-110)}{50} & 60 \le x \le 110 \\ 1 & x > 110 \end{cases} \tag{5.15}$$

and,

$$f_{20}(x) = \begin{cases} 0.55 & x < 60 \\ \dfrac{(110-x)0.55 + (x-60)0.29}{50} & 60 \le x \le 110 \\ 0.29 & x > 110 \end{cases} \tag{5.16}$$

Although the CMI mortality improvement model is applied to initial rates of mortality, while the log-link models are based on forces of mortality, the reduction factors can still be compared directly since the general trend in mortality rates and forces of mortality would be the same.

Figure 5.6 is a comparison of CMI mortality improvement factors with improvement factors derived from the log-link models for male and female pensioners. The factors are shown for the period 1993 to 2016 at 5-year age intervals from age 55 to 100 years. The highest age for which the factors have been plotted is 100 because at ages above 100, there are inconsistencies in the patterns exhibited by the log-link based reduction factors. For consistency with the CMI mortality improvement model, 1992 has been taken as the base year for the log-link models.

In general, the log-link based reduction factors for pensioners exhibit a similar pattern to CMI reduction factors, although the reduction factors based on the log-link models would predict lower forces of mortality for male pensioners and higher forces of mortality for female pensioners for a given base table of forces of mortality (or mortality rates).

A comparative plot of the CMI and log-link based reduction factors for immediate annuitants is shown in Figure 5.7, for ages 55 to 110 years over the calendar year period 1993 to 2016. The log-link model reduction factors are based on the analysis of the 1974-94 mortality experiences for male and female immediate annuitants, that is, the reduction factors described by expression (4.36) or (4.37) for males and expression (4.38) or (4.39) for females.

Given a base table of forces of mortality for immediate annuitants, the CMI reduction factors predict lower forces of mortality at the younger ages (up to age 70) and higher forces of mortality at the older ages than those predicted by the log-link based reduction factors. In addition, reduction factors based on the log-link models are relatively close to the CMI reduction factors at the younger ages, with large differences occurring at ages above 90 years. By age 110, the CMI mortality improvement model assumes that there is no longer improvement in mortality over time.

The log-link based mortality improvement models assumed for both pensioners and immediate annuitants are reproduced in Table 5.16.

Table 5.16
Mortality improvement models for life office pensioners and immediate annuitants

Experience	$RF(x,n)$
Male annuitants, d1+ (basis: 1974-94 experience)	$\exp[(-0.027898+0.000175x)n]$
Female annuitants, d1+ (basis: 1974-94 experience)	$\exp[(-0.048403+0.000383x)n]$
Male pensioners (basis: 1983-96 experience)	$\exp[(-0.078872+0.000744x)n]$
Female pensioners (basis: 1983-96 experience)	$\exp[(-0.050660+0.000489x)n]$

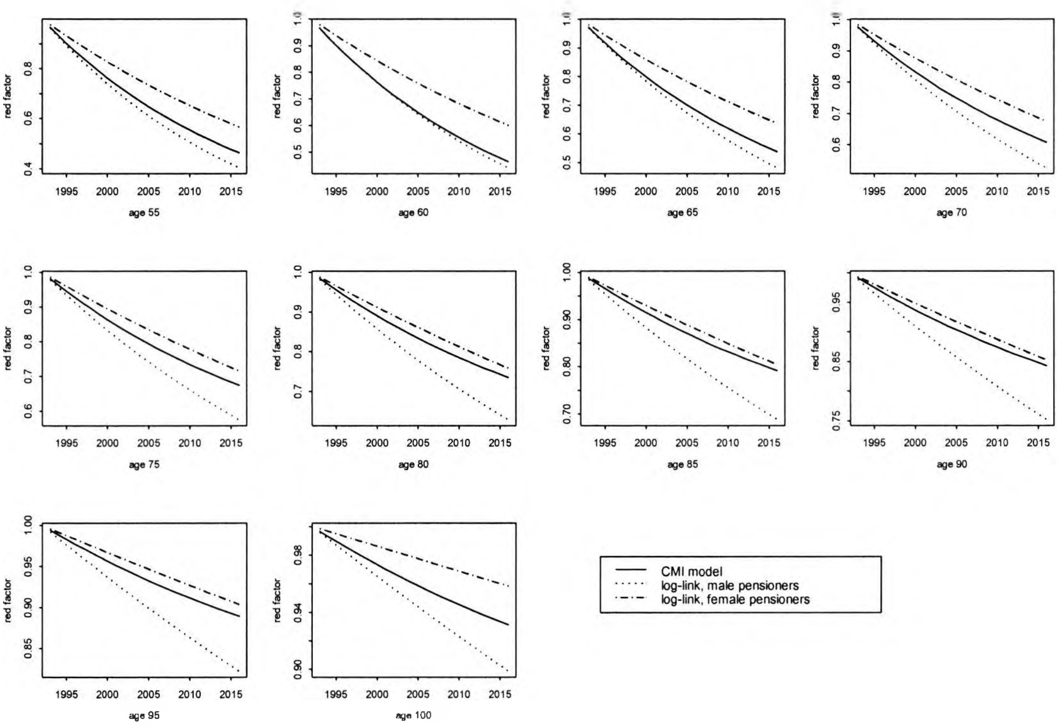


Figure 5.6 Comparison of time reduction factors, 1993 to 2016: CMI mortality improvement model and log-link models for life office pensioners based on analyses of the 1983-96 mortality experiences

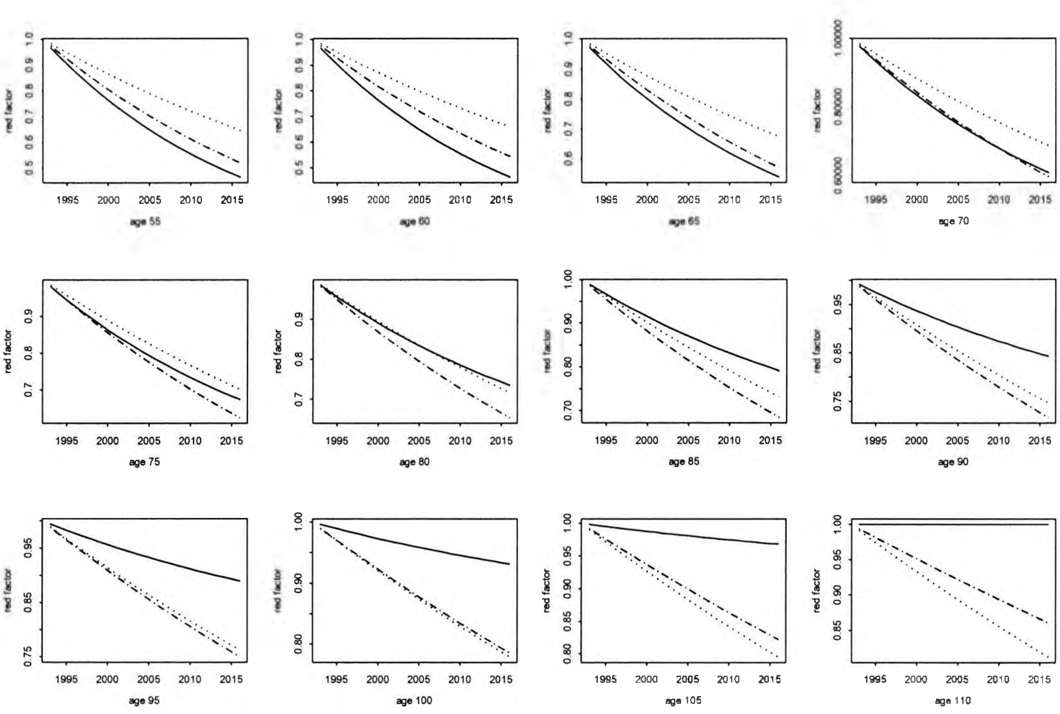


Figure 5.7 Comparison of time reduction factors, 1993 to 2016: CMI mortality improvement model and log-link models for immediate annuitants based on analyses of the 1974-94 mortality experiences

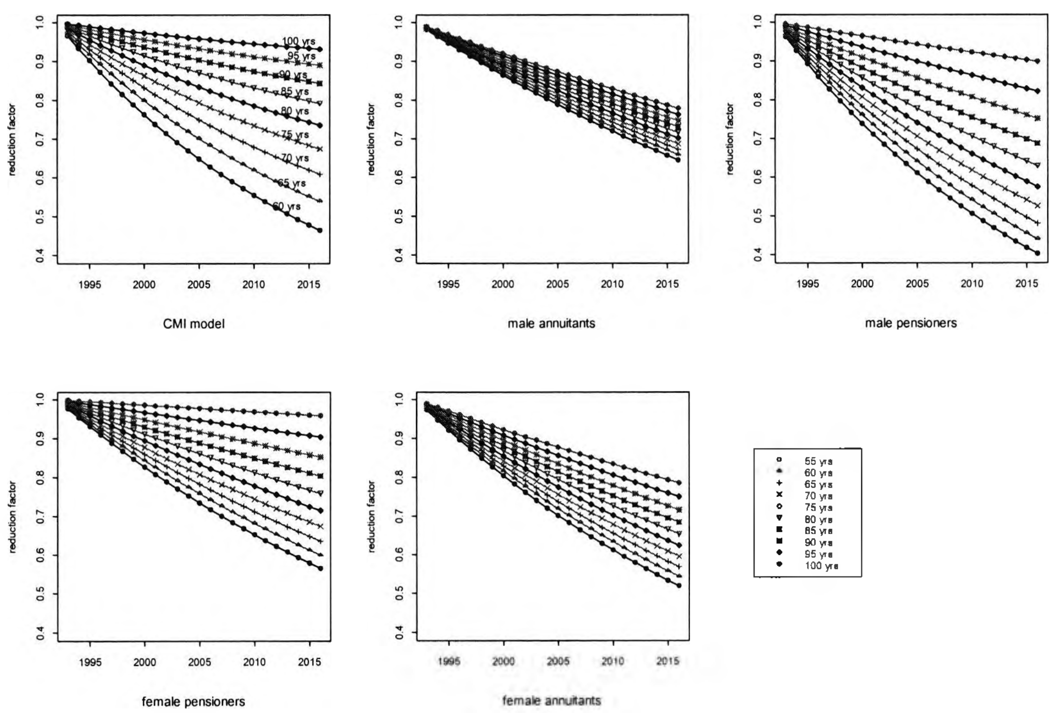


Figure 5.8 Comparison of log-link based reduction factors and CMI reduction factors for immediate annuitants, dl+ and life office pensioners

The mortality improvement factors based on the CMI model and on each of the log-link-based models, i.e. male and female pensioners, and male and female annuitants are shown separately in Figure 5.8. The factors are shown at 5-year age intervals from age 55 to 100 years. It should be noted that the CMI reduction factors at ages below 60 are given by the reduction factors at age 60 years.

From the plots it can be seen that while the CMI reduction factors and the log-link based reduction factors for pensioners have a similar shape, the log-link based reduction factors for immediate annuitants have a very different shape. It would appear that the CMI reduction factors have a shape with respect to age and time that is different from that which emerges from a model-based analysis of the immediate annuitants' mortality experience, raising doubts about the suitability of the CMI reduction factors for making predictions of future forces of mortality (or mortality rates) for immediate annuitants.

The reduction factors that would apply in 2014 based on the log-link models assumed and the CMI model are shown in Table 5.17. In practice, where the model-based reduction factor exceeds one, the value would be set to one.

Table 5.17
Reduction factors for pensioners and annuitants[3] applicable in calendar year 2014

Attained Age	male annuitants d1+	female annuitants d1+	male pensioners	female pensioners	CMI
55	0.6687	0.5482	0.4340	0.5927	0.4915
60	0.6817	0.5717	0.4710	0.6255	0.4915
65	0.6949	0.5964	0.5112	0.6600	0.5630
70	0.7084	0.6220	0.5548	0.6965	0.6301
75	0.7221	0.6488	0.6021	0.7350	0.6927
80	0.7361	0.6768	0.6535	0.7756	0.7506
85	0.7504	0.7059	0.7092	0.8184	0.8039
90	0.7650	0.7363	0.7697	0.8636	0.8526
95	0.7798	0.7680	0.8354	0.9113	0.8966
100	0.7949	0.8010	0.9067	0.9617	0.9358
105	0.8104	0.8355	0.9840	1.0148	0.9703
110	0.8261	0.8715	1.0680	1.0709	1.0000

5.5 Revised mortality improvement models for life office pensioners

The models derived for projecting life office pensioners' mortality exhibit some anomalies at extreme old age. Firstly, for both the male and female models, the projected forces of mortality above specific ages increase with time contrary to expectations. Secondly, for all ages above 95 years, the forces of mortality for males are lower than the forces of mortality for females, again contrary to expectations. These inconsistencies would require further investigation. In this section the discussion is confined to the problem of forces of mortality increasing with time.

[3] Reduction factors for immediate annuitants are based on analyses of the 1974-94 experiences

For brevity, the parameters α_1 and γ_{11} are denoted α and γ respectively. For each experience, the mortality improvement model derived can then be expressed as:

$$RF(x,n) = \exp\left[\frac{n}{w_t}\left\{\alpha + \gamma\left(\frac{x-c_x}{w_x}\right)\right\}\right],$$

from which it can be deduced that for a given value of x, the force of mortality decreases with time if $\alpha + \gamma\left(\frac{x-c_x}{w_x}\right) < 0$. Assuming that $\alpha < 0$, there are four scenarios to consider:

Case 1: $\gamma < 0$ and $x > c_x$

The condition $\alpha + \gamma\left(\frac{x-c_x}{w_x}\right) < 0$ is satisfied without imposing further conditions on the values of α and γ.

Case 2: $\gamma < 0$ and $x < c_x$

The condition $\alpha + \gamma\left(\frac{x-c_x}{w_x}\right) < 0$ is satisfied if $\gamma\left(\frac{x-c_x}{w_x}\right) < |\alpha|$.

Case 3: $\gamma > 0$ and $x > c_x$

As in case 2, the condition $\alpha + \gamma\left(\frac{x-c_x}{w_x}\right) < 0$ is satisfied if $\gamma\left(\frac{x-c_x}{w_x}\right) < |\alpha|$.

Case 4: $\gamma > 0$ and $x < c_x$

As in case 1 the condition $\alpha + \gamma\left(\frac{x-c_x}{w_x}\right) < 0$ is satisfied without imposing further conditions on the values of α and γ.

Therefore for each experience, if parameter values α^* and γ^* can be determined such that the condition $\gamma^*\left(\frac{x-c_x}{w_x}\right) < \left|\alpha^*\right|$, is true for all $x \le {}^{max}x$, where ${}^{max}x$ is a suitably large value of x which should be at least as large as the highest age assumed in the (base) mortality table, then the mortality improvement models derived would give forces of mortality decreasing with time for all $x \le {}^{max}x$.

The GLM modelling procedure used to derive mortality trend models is such that the associated standard errors of the parameter estimates are provided directly. Denoting the parameter estimates as $\boldsymbol{\beta}$, the estimates could be reported as $\boldsymbol{\beta} \pm$ *standard error*. The true value of each parameter will probably lie within 2 standard errors of the estimated value with an approximate probability of 0.95.

Here, the problem is essentially that of reducing the current values of the reduction factors at extreme old age, and hence finding revised parameter estimates that are lower than the 'current' estimates obtained from the model would have the desired effect. It therefore seems reasonable to take sets of parameter values determined from $\boldsymbol{\beta}$ – (*standard error*) or $\boldsymbol{\beta}$ – 2(*standard error*) to produce revised sets of mortality improvement factors. Clearly, being optimistic in the determination of the improvement factors leads to a less risky option for pricing and reserving annuities.

The mortality improvement models for male and female pensioners thus revised are shown in Table 5.18 and the reduction factors determined from the revised models are given in Table 5.19.

Table 5.18
Revised mortality improvement models for life office pensioners and immediate annuitants

Experience	***RF(x,n)***	
	β - (*standard error*)	β - 2(*standard error*)
Male pensioners	$\exp[(-0.071647+0.000645x)n]$	$\exp[(-0.064423+0.000546x)n]$
Female pensioners	$\exp[(-0.037552+0.000300x)n]$	$\exp[(-0.024443+0.000112x)n]$

Table 5.19
Revised reduction factors for life office pensioners applicable in calendar year 2014

Attained age	Life office pensioners' revised log-link models				CMI model
	Basis: β - (*standard error*)		Basis: β - 2(*standard error*)		
	Male pensioners ($^{max}x = 111$)	Female pensioners ($^{max}x = 125$)	Male pensioners ($^{max}x = 118$)	Female pensioners ($^{max}x = 219$)	
55	0.4514	0.6294	0.4695	0.6684	0.4915
60	0.4846	0.6506	0.4986	0.6767	0.4915
65	0.5203	0.6724	0.5295	0.6850	0.5630
70	0.5585	0.6950	0.5623	0.6935	0.6301
75	0.5996	0.7183	0.5971	0.7020	0.6927
80	0.6437	0.7424	0.6341	0.7107	0.7506
85	0.6911	0.7673	0.6734	0.7195	0.8039
90	0.7419	0.7931	0.7151	0.7283	0.8526
95	0.7965	0.8197	0.7594	0.7373	0.8966
100	0.8551	0.8472	0.8065	0.7464	0.9358
105	0.9180	0.8757	0.8564	0.7556	0.9703
110	0.9855	0.9051	0.9095	0.7650	1.0000

From the reduction factors given in Table 5.17, it can be seen that if we assume the maximum age in a given life table to be 110 years, then the objective is achieved. The reduction factors applicable at all ages up to age 110 are all less than 1 and these reduction factors decrease with time for all $x \leq {}^{max}x$. Modifying the mortality improvement models in this way has a greater effect on reduction factors applicable to female pensioners. This is because the parameter estimates for α_{11} and γ_{11} in the female pensioners' model have relatively large standard errors compared with parameter estimates for the male pensioners' model. Although the forces of mortality for male pensioners are still predicted to improve at a faster rate than for female pensioners at most ages, higher improvements in mortality would be predicted for female pensioners at extreme old age.

It would seem that the reduction factors determined from β – (*standard error*) would be the most appropriate for forecasting forces of mortality for life office pensioners. The factors vary steadily with age and while an improvement in mortality is predicted at all ages, the improvement at extreme old age is small.

5.6 A brief note on log-link based mortality improvement models for pensioners and annuitants

The mortality improvement model $RF(x,n)$ pertains to attained age x in calendar year $t=t_0+n$ where t_0 is the base calendar year. The forecasting model for the force of mortality for a life attaining exact age x in calendar year t is:

$$\mu_{xt} = \mu_{xt_0} \cdot RF(x,n). \tag{5.17}$$

Denoting as μ_x^τ the force of mortality for a life age x born in calendar year τ, a forecasting model pertaining to cohort (year of birth) experience can be expressed as

$$\mu_x^\tau = \mu_{xt_0} \cdot RF(x, x+\tau-t_0). \tag{5.18}$$

Assuming that the force of mortality for a life age x in calendar year t is given by:

$$\mu_{xt} = \exp\left\{\sum_{j=0}^{3}\beta_j L_j(x') + (\alpha+\gamma x')t'\right\}, \tag{5.19}$$

with

$$RF(x,n) = \exp\left\{\frac{n}{w_t}(\alpha+\gamma x')\right\}, \tag{5.20}$$

and noting that

$$\mu_x^\tau = \mu_{x,x+\tau}, \tag{5.21}$$

μ_x^τ can be expressed as:

$$\mu_x^\tau = \exp\left\{\sum_{j=0}^{3}\beta_j L_j(x') + (\alpha+\gamma x')(x+\tau)'\right\}. \tag{5.22}$$

Further, it can be deduced that given forces of mortality in a base calendar year t_0, the mortality improvement model for a life age x born in calendar year τ is:

$$RF(x, x+\tau-t_0) = \exp\left\{\frac{(x+\tau-t_0)}{w_t}(\alpha+\gamma x')\right\} \tag{5.23}$$

with $x' = \dfrac{x-c_x}{w_x}$ and $t' = \dfrac{t-c_t}{w_t}$ $\left(note: (x+\tau)' = \dfrac{x+\tau-c_t}{w_t}\right)$.

A parallel model to the GLM-based cohort-forecasting model described here is the Austrian mortality-forecasting model described in Chapter 3 (pages 90-92):

$$q_x^\tau = q_x(x+\tau) \tag{5.24}$$

where:

$$q_x(t) = q_x(t_0)\exp\{-\lambda_x(t-t_0)\}. \tag{5.25}$$

The difference is that the Austrian improvement model is defined in terms of the initial rate of mortality, as is the CMI model. The equivalent form of the function λ_x in the Austrian model (5.25) is described by a linear function of x in the GLM-based models (expression (5.23)).

In effect, the year of birth cohort mortality improvement model (5.23) and the period mortality improvement model (5.20) are equivalent since $t = x + \tau$. The predicted forces of mortality derived in this study are largely supportive of this form particularly in the case of male life office pensioners. Inconsistencies are apparent at a few ages at both extremes of the age range over which forces of mortality have been predicted. These inconsistencies are probably a result of the observed data at these and neighbouring ages being scanty or perhaps a result of the assumed model being inappropriate for these specific ages. A detailed investigation of this problem would be required but such an investigation has not been undertaken in this study.

Alternatively, a GLM-based mortality trend analysis by year of birth can be carried out by replacing age x with year of birth $z = t - x$ in fitting the model structure (5.1) as was done by Renshaw *et al* (1996) in analysing mortality trends in UK male assured lives.

It is worth noting that a satisfactory analysis of mortality data by cohort generally requires a long series of consistent data, which is not always available.

Chapter 6

Comparison of GLM-based Mortality Forecasts and a Time Series-based Approach: Female Annuitants

6.1 Introduction

In this chapter, an alternative approach to modelling forces of mortality over time is considered. The forecasting procedure applied is that used by McNown and Rogers (1989) to develop forecasts of US mortality, a method that combines parametric models and time series methods to generate forecasts. The approach is essentially a three-stage process. The first stage involves fitting to mortality data for each calendar year t the same age-dependent parametric model. McNown and Rogers (1989) used the Heligman-Pollard model described by expression (6.1) to model US mortality over the period 1900 to 1985:

$$q_x = A^{(x+B)^C} + D\exp\left\{-E(\ln x - \ln F)^2\right\} + GH^x/\left(1 + GH^x\right) \tag{6.1}$$

The parameter estimates obtained represent a set of observations over time. The second stage is then to model the trends in the observations using time series methods to develop forecasting models for the mortality measures. McNown and Rogers (1989) applied univariate *autoregressive integrated moving average* (ARIMA) models to observations over time on each of the 8 parameters in the Heligman-Pollard model (6.1) to describe mortality trends in the US population. The time-series model for each of the parameters was identified according to the methods of Box and Jenkins

(1976), using 35 observations from 1941 to 1975. (Trends in the observations for the period prior to 1941 were considered to be unsuitable for forecasting purposes.)

The final stage is to use the forecasting models developed to make predictions of the mortality measures.

Felipe, Guillen and Artis (1998) have applied the NcNown and Rogers' (1989) methodology to model the evolution of Spanish mortality patterns using Heligman-Pollard law number two. Thus for each calendar year t, the probability of death is modelled using:

$$q_{x,t} = A_t^{(x+B_t)^{C_t}} + D_t \exp\left\{-E_t\left(\ln x - \ln F_t\right)^2\right\} + G_t H_t^x \Big/ \left(1 + K_t G_t H_t^x\right).$$

As an illustration of the McNown and Rogers' method, only one of the four data sets modelled in Chapters 4 and 5, that is, the female immediate annuitants' mortality experience, is considered. For each calendar year t, the force of mortality μ_{xt} at age x is modelled using Gompertz-type parametric models, fitted in the framework of GLMs. Time series methods are then used to project the forces of mortality and comparisons made with forces of mortality projected by applying the GLM approach of Chapter 4.

A brief description of time series models is given in Section 6.2. Detailed descriptions can be found in textbooks such as Box and Jenkins (1976), Chatfield (1996), Hamilton (1994) and many other books on time series analysis. The procedures applied in analysing the mortality of female immediate annuitants using time series, and the results obtained are presented in Section 6.3.

6.2 Univariate time series models

Suppose that the series of observations over time $\{y_t : t = 1,2,\cdots,n\}$ is a sample of size n of some random variable $\{Y_t : t = 1,2,\cdots,n\}$. The sequence $\{Y_t\}$ is referred to as a time series process. Each random variable Y_t has a distribution, a mean and a variance.

Probability models for time series are collectively called *stochastic processes*. Chatfield (1996) defines a stochastic process as '*a collection of random variables which are ordered in time and defined at a set of time points which may be continuous or discrete*'.

The expectation of the tth observation of a time series process $\{Y_t\}$ is the mean function m_t:

$$m_t = E[Y_t]. \tag{6.2}$$

The second moments are specified by the variance function and the *autocovariance function*. The usual definition of the variance applies to the variance function σ_t^2, thus:

$$\sigma_t^2 = E\left[(Y_t - m_t)^2\right]. \tag{6.3}$$

The jth autocovariance function γ_{jt}, is the covariance of Y_t and Y_{t+j}, and is defined by:

$$\gamma_{jt} = E\{(Y_t - m_t)(Y_{t+j} - m_{t+j})\}. \tag{6.4}$$

The difference in time between Y_t and Y_{t+j} is called the *lag* and hence γ_{jt} is the autocovariance at lag j.

A key assumption in using Box-Jenkins time series procedures is that of *stationarity*. Broadly '*a time series is said to be stationary if there is no systematic change in the mean, if there is no systematic change in the variance, and if strictly periodic variations have been removed*' (Chatfield, 1996).

If neither the mean m_t nor the autocovariances γ_{jt} depend on t, then the sequence $\{Y_t\}$ is said to be *covariance-stationary* or *weakly stationary*:

$$E[Y_t] = m \qquad \text{for all } t$$

and,

$$E\{(Y_t - m_t)(Y_{t+j} - m_{t+j})\} = \gamma_j \quad \text{for all } t \text{ and any } j.$$

Thus the process $\{Y_t\}$ is weakly stationary if its mean is constant and its autocovariance function depends only on the lag j.

In contrast a time series process is said to be *strictly stationary* if the joint distribution of $(Y_t, Y_{t+j_1}, Y_{t+j_2}, \cdots, Y_{t+j_n})$ depends only on the intervals between $j_1, j_2, \cdots, j_n$ and not on the time t.

The *autocorrelation function* at lag j, ρ_j of a stationary time series process is defined by

$$\rho_j = \frac{\gamma_j}{\gamma_0}, \tag{6.5}$$

where γ_0 is the autocovariance at lag 0, i.e. γ_0 is the variance function of the process. The autocorrelation function ρ_j measures the correlation between Y_t and Y_{t+j} and hence satisfies the usual property of a correlation coefficient, that is

$$|\rho_j| \leq 1.$$

For any covariance-stationary process, the autocovariance is an even function:

$$\gamma_j = \gamma_{-j} \quad \text{for } j = 1,2,\cdots$$

Some classes of stochastic processes useful for time-series modelling are described in Section 6.2.1 below.

6.2.1 Probability models for time-series

White noise process

The simplest example of a time-series model is a *white noise process*. Suppose the random variables $\{\varepsilon_t : t = 1,2,3,\cdots,n\}$ are independent and identically distributed (i.i.d.) with mean 0 and variance σ^2, the process $\{\varepsilon_t\}$ is called a white noise process. The white noise process is the basic building block for time series models considered in this thesis.

Moving average process

The time series process $\{Y_t : t = 1,2,3,\cdots,n\}$ is a moving average process of order q (denoted MA(q) process) if

$$Y_t = m + \varepsilon_t + \theta_1\varepsilon_{t-1} + \theta_2\varepsilon_{t-2} + \cdots + \theta_q\varepsilon_{t-q} \tag{6.6}$$

where m and $\{\theta_i : i = 1,2,\cdots,q\}$ are constants and $\{\varepsilon_t : t = 1,2,\cdots,n\}$ is a white noise process with mean 0 and variance σ^2.

A moving average process is covariance-stationary for any values of the $\{\theta_i\}$ constants. The expectation of a MA(q) process is:

$$E(Y_t) = E(m + \varepsilon_t + \theta_1\varepsilon_{t-1} + \cdots + \theta_q\varepsilon_{t-q}) = m \tag{6.7}$$

and the variance function is:

$$\gamma_0 = \mathrm{var}(Y_t) = \mathrm{var}(m + \varepsilon_t + \theta_1\varepsilon_{t-1} + \cdots + \theta_q\varepsilon_{t-q}) = \sigma^2(1 + \theta_1^2 + \cdots + \theta_q^2) \tag{6.8}$$

since the $\{\varepsilon_t\}$ are independent and identically distributed with mean 0 and variance σ^2. By considering

$$\gamma_j = E[(\varepsilon_t + \theta_1\varepsilon_{t-1} + \cdots + \theta_q\varepsilon_{t-q}) \times (\varepsilon_{t-j} + \theta_1\varepsilon_{t-j-1} + \cdots + \theta_q\varepsilon_{t-j-q})], \tag{6.9}$$

and noting that $E[\varepsilon_i\varepsilon_k = 0]$ for $i \neq k$, it is seen that for $j = 1,2,\cdots,q$, the autocovariance function γ_j is

$$\gamma_j = \sigma^2(\theta_j + \theta_{j+1}\theta_1 + \theta_{j+2}\theta_2 + \cdots + \theta_q\theta_{q-j}), \tag{6.10}$$

and is 0 for $j > q$.

Autoregressive process

The process $\{Y_t\}$ is said to be an autoregressive process of order p (denoted AR(p) process) if

$$Y_t = \zeta + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \cdots + \phi_p Y_{t-p} + \varepsilon_t \tag{6.11}$$

where ζ and $\{\phi_k : k = 1,2,\cdots,p\}$ are constants and $\{\varepsilon_t\}$ is again a white noise process with mean 0 and variance σ^2.

Let B denote the *backward shift operator* such that

$$B^j Y_t = Y_{t-j} \quad \text{for all } j. \tag{6.12}$$

By applying the backward shift operator to equation (6.11), the equation may be expressed as:

$$Y_t(1 - \phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p) = \zeta + \varepsilon_t \tag{6.13}$$

An AR(p) process will only be stationary if the parameters $\{\phi_k\}$ lie within a certain range. In particular, an AR(p) process will only be stationary if the roots of the equation

$$\phi(B) = (1 - \phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p) = 0 \tag{6.14}$$

lie outside the unit circle. For example, in the case of an AR(1) process:

$$Y_t = \zeta + \phi Y_{t-1} + \varepsilon_t \tag{6.15}$$

the process is covariance-stationary if $|\phi| < 1$.

Assuming that the stationarity conditions are satisfied, the mean function of an AR(p) process is given by

$$m = \frac{\zeta}{1 - \sum_{k=1}^{p} \phi_k}, \tag{6.16}$$

and the second moments satisfy the equations:

$$\gamma_0 = \phi_1\gamma_1 + \phi_2\gamma_2 + \cdots + \phi_p\gamma_p + \sigma^2 \tag{6.17}$$

and

$$\gamma_j = \phi_1\gamma_{j-1} + \phi_2\gamma_{j-2} + \cdots + \phi_p\gamma_{j-p} \quad \text{for all } j > 0 \tag{6.18}$$

where, as before, $\sigma^2 = \text{var}(\varepsilon_t)$.

Using the fact that $\gamma_j = \gamma_{-j}$, autocovariances are found by solving the system of equations (6.17) and (6.18) for $j = 0,1,2,\cdots,p$ as functions of $\sigma^2,\phi_1,\phi_2,\cdots\phi_p$. For example, for an AR(2) process:

$$Y_t = \zeta + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \varepsilon_t \tag{6.19}$$

$$\gamma_0 = \phi_1\gamma_1 + \phi_2\gamma_2 + \sigma^2 \tag{6.20}$$

and

$$\gamma_j = \phi_1\gamma_{j-1} + \phi_2\gamma_{j-2} \quad \text{for } j > 0\ . \tag{6.21}$$

Solving equation (6.21) recursively for γ_1 and γ_2 yields

$$\gamma_1 = \frac{\phi_1}{1-\phi_2}\gamma_0 \tag{6.22}$$

and

$$\gamma_2 = \left(\frac{\phi_2 - \phi_2^2 + \phi_1^2}{1-\phi_2}\right)\gamma_0 . \tag{6.23}$$

Substituting the expressions for γ_1 (6.22) and γ_2 (6.23) in expression (6.20) gives an equation for γ_0 as a function of σ^2, ϕ_1 and ϕ_2, thus

$$\gamma_0 = \frac{(1-\phi_2)\sigma^2}{\left((1-\phi_2)^2 - \phi_1^2\right)(1+\phi_2)} . \tag{6.24}$$

Mixed autoregressive moving average processes

A mixed autoregressive moving average process of order *p*,*q* (denoted ARMA(*p*,*q*) process) has *p* autoregressive terms and *q* moving average terms:

$$Y_t = \zeta + \phi_1 Y_{t-1} + \cdots + \phi_p Y_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \cdots + \theta_q \varepsilon_{t-q} . \tag{6.25}$$

Stationarity of an ARMA(*p*,*q*) process depends entirely on the autoregressive parameters $\{\phi_k\}$.

The mean function of an ARMA(*p*,*q*) process is simply the mean of the autoregressive terms, that is

$$m = E(Y_t) = \frac{\zeta}{1-\phi_1-\phi_2-\cdots-\phi_p}$$

as given by equation (6.16). The equation for the mean function may be written in the form:

$$\zeta = m(1 - \phi_1 - \phi_2 - \cdots - \phi_p)\,. \tag{6.26}$$

Using the result (6.26), the process (6.25) may be expressed as

$$Y_t - m = \phi_1(Y_{t-1} - m) + \cdots + \phi_p(Y_{t-p} - m) + \varepsilon_t + \theta_1\varepsilon_{t-1} + \cdots + \theta_q\varepsilon_{t-q}\,. \tag{6.27}$$

By multiplying (6.27) by $(Y_t - m)$ and taking expectations, the autocovariance function of an ARMA(p,q) process can be obtained. Indeed the same method can be used to find the autocovariance function of an AR(p) process.

For an ARMA(1,1) process

$$Y_t = \zeta + \phi Y_{t-1} + \varepsilon_t + \theta\varepsilon_{t-1}\,, \tag{6.28}$$

it can be shown that

$$\gamma_0 = \frac{(1 + 2\phi\theta + \theta^2)}{1 - \phi^2}\sigma^2\,, \tag{6.29}$$

$$\gamma_1 = \frac{(1 + \phi\theta)(\phi + \theta)}{1 - \phi^2}\sigma^2 \tag{6.30}$$

and

$$\gamma_j = \phi\gamma_{j-1} \text{ for } j \geq 2\,. \tag{6.31}$$

The full derivation of the autocovariance functions described by expressions (6.29), (6.30) and (6.31) is shown in Appendix 2.

Random walk

A process $\{Y_t\}$ is said to be a random walk if

$$Y_t = Y_{t-1} + \varepsilon_t . \tag{6.32}$$

where $\{\varepsilon_t\}$ is a purely random process with mean 0 and variance σ^2 (i.e. the random variables $\{\varepsilon_t\}$ are independent and identically distributed). The process has the same form as an AR(1) process (equation 6.15), except that ϕ is one for a random walk process.

Repeatedly substituting for past values of Y_t gives

$$Y_t = Y_0 + \sum_{i=1}^{t} \varepsilon_i . \tag{6.33}$$

If the initial value Y_0 is fixed, the mean and variance functions are $m_t = y_0$ and $\sigma_t^2 = t\sigma^2$ respectively. A random walk is therefore a non-stationary process.

The first differences of the process $\{Y_t\}$ are defined by:

$$\nabla Y_t = Y_t - Y_{t-1} = \varepsilon_t \tag{6.34}$$

The sequence $\{\nabla Y_t\}$ is the same as the purely random process $\{\varepsilon_t\}$. Hence a stationary process can be derived from a random walk by taking the first differences of the random walk process.

A random walk with drift δ is defined as the process:

$$Y_t = \delta + Y_{t-1} + \varepsilon_t \tag{6.35}$$

where δ is a constant.

Autoregressive integrated moving average processes

Suppose $\{Y_t\}$ is a time series process and $\{W_t\}$ is a sequence formed by taking the d-th differences of the process $\{Y_t\}$, that is

$$W_t = \nabla^d Y_t . \tag{6.36}$$

For example,

$$\nabla^2 Y_t = \nabla Y_t - \nabla Y_{t-1} = Y_t - 2Y_{t-1} + Y_{t-2} . \tag{6.37}$$

Writing $BY_t = Y_{t-1}$, then $\nabla Y_t = Y_t - Y_{t-1} = (1-B)Y_t$. Expression (6.36) may thus be written as:

$$W_t = (1-B)^d Y_t . \tag{6.38}$$

$\{Y_t\}$ is an autoregressive integrated moving average process (ARIMA(p,d,q) process) if:

$$W_t = \varphi_1 W_{t-1} + \varphi_2 W_{t-2} + \cdots + \varphi_p W_{t-p} + \varepsilon_t + \psi_1 \varepsilon_{t-1} + \psi_2 \varepsilon_{t-2} + \cdots + \psi_q \varepsilon_{t-q} \tag{6.39}$$

is an ARMA(p,q) process. Thus an ARIMA(p,d,q) process is a non-stationary process whose d-th difference produces a stationary ARMA(p,q) process.

The Box-Jenkins approach to time series analysis is a method for forecasting non-stationary time series based on ARIMA models. Hamilton (1994) breaks down the Box-Jenkins approach into four steps:

(1) Transform the data, if necessary, so that the assumption of covariance-stationarity is a reasonable one.

(2) Make an initial guess of small values of p and q for an ARMA(p,q) model that might describe the (transformed) series.

(3) Estimate the unknown autoregressive and moving average parameters.

(4) Perform diagnostic analysis to confirm that the model is indeed consistent with the observed features of the data.

6.2.2 Model Selection

An initial indication of a suitable model can be provided by an examination of the sample autocorrelation function. The autocovariance function γ_j at lag j is usually estimated by:

$$c_j = \frac{1}{n}\left\{\sum_{t=1}^{n-j}(y_t - \bar{y})(y_{t+j} - \bar{y})\right\} \tag{6.40}$$

where

$$\bar{y} = \frac{1}{n}\sum_{t=1}^{n} y_t$$

is the sample mean of the process and n is the number of observations. The estimator c_j is a biased estimator of γ_j (with bias of order $1/n$). However, it is asymptotically unbiased, that is

$$\lim_{n\to\infty} E(c_j) = \gamma_j \tag{6.41}$$

Having obtained an estimate of the autocovarince function, the autocorrelation function ρ_j is then estimated by:

$$r_j = \frac{c_j}{c_0} \,. \tag{6.42}$$

For a MA(q) process, γ_j is zero for $|j| > q$ so that a plot of the sample autocorrelation function against lag (called a *correlogram*) would be a useful tool in identifying a MA(q) process.

In contrast, if the process is an AR(p) process, the autocovariances are generally all non-zero and the autocorrelation coefficients ρ_j would tend towards zero exponentially. An AR(p) model can however be identified from a plot of the *partial autocorrelation function*. The pth partial autocorrelation coefficient is defined as the last coefficient ϕ_p after regression on $Y_t, Y_{t+1}, \cdots, Y_{t+p-1}$ when fitting an AR(p) model. It '*measures the excess correlation at lag p which is not accounted for by an* AR($p-1$) *model*' (Chatfield, 1996). For an AR(p) process, the partial autocorrelation between Y_t and Y_{t+j} is zero for $j > p$.

The coefficients of an AR(p) process satisfy the Yule-Walker equations:

$$\rho_j = \phi_1 \rho_{j-1} + \phi_2 \rho_{j-2} + \cdots + \phi_p \rho_{j-p} \quad \text{for all } j > 0. \tag{6.43}$$

By solving the Yule-Walker equations (6.43) with ρ replaced by r, the partial autocorrelation function and the unknown $\{\phi_j\}$ parameters can be estimated.

Model identification among ARMA processes can also be done using Akaike's information criterion (AIC) which is minus twice the log-maximised likelihood plus twice the number of parameters estimated. The model chosen is that which minimises AIC. Hamilton (1994) provides a detailed discussion of maximum likelihood estimation of time series processes.

A common approximation to the likelihood function for an ARMA(p,q) process is conditioned on initial values of the $\{Y_t\}$ and $\{\varepsilon_t\}$. Box and Jenkins (1976) recommended setting the $\{\varepsilon_t\}$ to 0 and the $\{Y_t\}$ to their observed values. Thus assuming the $\{\varepsilon_t\}$ are i.i.d. random variables with a normal distribution, mean 0 and variance σ^2, the log-likelihood function for an ARMA(p,q) process conditioned on the first p values of the series can be shown to be:

$$\log f\left(Y_n, \cdots, Y_{p+1} \middle| Y_p, \cdots, Y_1\right) = -\frac{n-p}{2} \log\left(2\pi\sigma^2\right) - \sum_{t=p+1}^{n} \frac{\varepsilon_t^2}{2\sigma^2}. \tag{6.44}$$

When the observations are normally distributed, the likelihood function can be expressed in terms of the residuals, $\{e_t\}$ (the differences between the observed values and the predicted values). This is known as *prediction error decomposition* (Harvey, 1993). The variance of e_t is the same as the conditional variance of Y_t, i.e.

$$\operatorname{var}(e_t) = \operatorname{var}(Y_t | Y_{t-1}, Y_{t-2}, \cdots, Y_1). \tag{6.45}$$

Writing var(e_t) as $\sigma^2 f_t$, the prediction error decomposition form of the conditional log likelihood is:

$$\log f(Y_n, \cdots, Y_{p+1} | Y_p, \cdots, Y_1) = -\frac{(n-p)}{2}\log(2\pi\sigma^2) - \frac{1}{2}\sum_{t=p+1}^{n}\log f_t - \frac{1}{2\sigma^2}\sum_{t=p+1}^{n}\frac{e_t^2}{f_t} \tag{6.46}$$

The procedure for fitting ARIMA models available in S-Plus uses this latter form (equation 6.46) of the conditional log-likelihood.

The optimal forecast of Y_{t+1} is the conditional expectation $E(Y_{t+1} | Y_t, \cdots, Y_1)$.

6.2.3 Model Diagnostics

As in the case of fitting GLM models discussed in Chapter 3, the adequacy of fit of an ARIMA model can be explored through an examination of the residuals. The analysis of residuals for time-series models is discussed in detail in textbooks on time series such as Box and Jenkins (1976) and Harvey (1993) for example.

If the correct ARIMA model has been fitted, a plot of the standardised residuals $\left(e_t / \sqrt{f_t}\right)$ against time should behave like a standardised normal random variable (a normal random variable with mean 0 and variance 1) approximately. The plot against time will reveal any outliers and any obvious non-randomness in time.

In addition, Chatfield (1996), notes that the autocorrelations of the true errors should be uncorrelated and have an approximate normal distribution with mean 0 and variance $1/n$ for reasonably large values of n. Although the correlogram of the residuals has somewhat different properties, it turns out that $1/\sqrt{n}$ forms an upper bound for the standard error of the residuals and hence the presence of autocorrelations which lie outside the range $\pm 2/\sqrt{n}$ would give evidence of the inadequacy of the model at the 5% significance level (see for example Box and Pierce, 1970).

Another useful diagnostic described by Box and Jenkins (1976) is based on the sample autocorrelations taken as a whole. The *portmanteau test statistic* Q is defined by:

$$Q = n_1 \sum_{k=1}^{K} r_k^2 \qquad (6.47)$$

where r_k is the estimated autocorrelation function defined in (6.42), K is a fixed maximum number of lags and n_1 is the number of observations used to compute the (log) likelihood. If the appropriate ARIMA model is fitted and the data have a normal distribution, then Q has an approximate χ^2 distribution with $K\text{–}p\text{–}q$ degrees of freedom, where p and q are the number of autoregressive and moving average terms respectively, in the model.

6.3 Time series analysis of female annuitants' mortality experience

The mortality experience of female immediate annuitants with policy duration 1 year and over (1+ years) was analysed over the calendar year period 1958 to 1994 (excluding the experience for the years 1968, 1971 and 1975), giving 34 observations for each of the parameter estimates. No attempt was made to include the pre-1958 mortality experience in the time-series analysis since the inclusion of this experience in the GLM-based modelling procedures of Chapter 4 resulted in models unsuitable for forecasting. For each calendar year t, the experience was analysed over the age range $x = 65, 66, \cdots, 100$ years.

6.3.1 Age-dependent parametric models

The first stage of the time-series analysis of mortality trends of female annuitants involves modelling for each calendar year t, the force of mortality μ_{xt} at age x using Gompertz-type parametric models of the form:

$$\mu_{xt} = \exp\left\{ \sum_{j=0}^{s} \beta_{jt} L_j(x') \right\}. \tag{6.48}$$

$L_j(x)$ are the Legendre polynomials described in Chapter 2 and x' is age x scaled such that the maximum value of x' is 1 and the minimum value is -1. For the female immediate annuitants' experience modelled, $x' = \dfrac{x - 82.5}{17.5}$ for each age x.

The GLM modelling procedures used in Chapters 4 and 5 were also used to estimate the unknown β_{jt} parameters for each calendar year t. Thus the actual number of deaths a_{xt} at age x in calendar year t are modelled as independent realisations of Poisson response variables A_{xt} of a generalised linear model with mean and variance given by:

$$E[A_{xt}] = m_{xt} = R^c_{xt} \mu_{xt}; \tag{6.49}$$

$$\operatorname{var}(A_{xt}) = \kappa_t m_{xt}. \tag{6.50}$$

R^c_{xt} is the central exposed-to-risk and κ_t is a scale parameter to take account of possible over-dispersion of the Poisson random variable due to duplicate policies on the same lives. The unknown parameters are linked to the mean through the log function:

$$\eta_{xt} = \log m_{xt} = \log R^c_{xt} + \log \mu_{xt}, \tag{6.51}$$

that is,

$$\eta_{xt} = \log R^c_{xt} + \sum_{j=0}^{s} \beta_{jt} L_j(x'). \tag{6.52}$$

Estimation of the parameters is carried out by minimising the negative of the quasi-log-likelihood, that is by minimising:

$$\frac{1}{\kappa_t} \sum_x (m_{xt} - a_{xt} \log m_{xt}) \tag{6.53}$$

The optimum value of s chosen is the minimum value of s beyond which improvement in the scaled deviance for successive increases in the value of s is not statistically significant, assuming a χ^2 distribution for the improvement. A more detailed discussion of model fitting using GLM procedures is given in Chapters 3 and 4 of this study while McCullagh and Nelder (1989) give an authoritative discussion of GLMs.

The parametric model adopted for each calendar year t is in the form of a $GM_x(0,3)$ model defined by:

$$\mu_{xt} = \exp\left\{ \sum_{j=0}^{2} \beta_{jt} L_j(x') \right\}, \tag{6.54}$$

which is equivalent to:

$$\mu_{xt} = \exp\left\{ \beta_{0t} + \beta_{1t} x' + \beta_{2t} \left[\frac{3x'^2 - 1}{2} \right] \right\} \tag{6.55}$$

since $L_0(x') = 1$, $L_1(x') = x'$ and $L_2(x') = \frac{3x'^2 - 1}{2}$. For each calendar year t, the parameter β_{0t} represents the general level of mortality while β_{1t} and β_{2t} represent the mortality level pertaining to age x.

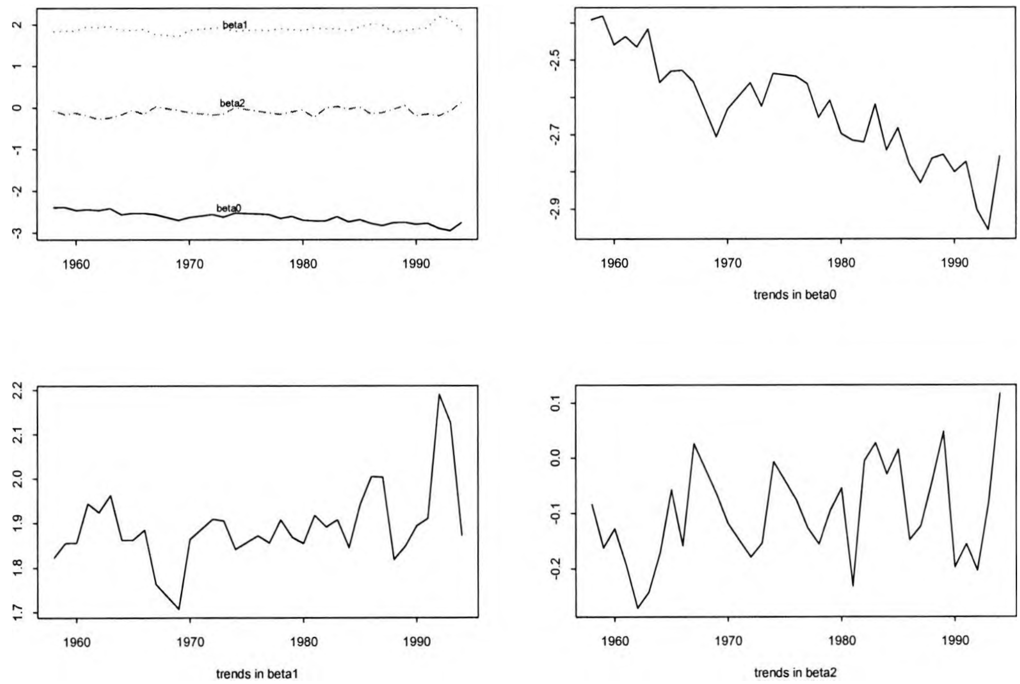


Figure 6.1 Female immediate annuitants, d1+ years: plots of observed parameters over time; estimates based on $GM_x(0,3)$ graduation models fitted to mortality experience for each calendar year from 1958 to 1994 (excluding 1968, 1971 and 1975)

Table 6.1 shows the 34 observations of each of the 3-parameter estimates, together with the associated standard errors and t-values. Although some of the observations on the β_{2t} parameter have an absolute t-value less than 2, indicating statistical non-significance, (and hence suggesting possible over-parameterisation), the 3-parameter age-dependent model was still retained because of the better forecasting performance over a 2-parameter model, for example.

From the plots of 'observed' parameters over time shown in Figure 6.1, it is evident that there has been a steady decline in the general level of mortality as indicated by the steady decrease in the estimates of the parameter β_{0t} over t. The observed values of β_{1t} appear to fluctuate about some value in the range 1.7 to 2 while the observed values of β_{2t} tend to fluctuate about some value between –0.3 and 0.

Table 6.1
Female immediate annuitants, d1+ years: parameter estimates based on $GM_x(0,3)$ graduation models for each calendar year 1958 to 1994 (excluding 1968, 1971 and 1975)

	β_0			β_1			β_2		
Year (*t*)	estimate	Std error	*t*-value[1]	estimate	Std error	*t*-value[1]	estimate	Std error	*t*-value[1]
1958	-2.391899	0.0391	-61.22	1.823163	0.0735	24.82	-0.083837	0.0986	-0.85
1959	-2.382317	0.0338	-70.57	1.855283	0.0630	29.44	-0.162014	0.0850	-1.91
1960	-2.459115	0.0414	-59.34	1.856512	0.0781	23.76	-0.127725	0.1034	-1.24
1961	-2.437456	0.0325	-74.91	1.943654	0.0614	31.66	-0.193643	0.0802	-2.41
1962	-2.465134	0.0310	-79.59	1.924045	0.0582	33.03	-0.271126	0.0763	-3.55
1963	-2.416211	0.0341	-70.96	1.962524	0.0643	30.54	-0.242308	0.0836	-2.90
1964	-2.561308	0.0350	-73.21	1.862250	0.0667	27.94	-0.170808	0.0861	-1.98
1965	-2.529934	0.0294	-86.18	1.862973	0.0565	32.96	-0.057560	0.0717	-0.80
1966	-2.527692	0.0270	-93.60	1.885102	0.0516	36.55	-0.158785	0.0662	-2.40
1967	-2.559028	0.0235	-109.04	1.763904	0.0452	38.99	0.025532	0.0580	0.44
1969	-2.706943	0.0362	-74.86	1.707638	0.0697	24.50	-0.065999	0.0885	-0.75
1970	-2.632070	0.0332	-79.22	1.864585	0.0638	29.22	-0.118763	0.0812	-1.46
1972	-2.561615	0.0336	-76.33	1.909122	0.0643	29.68	-0.179197	0.0816	-2.19
1973	-2.624880	0.0265	-99.12	1.905449	0.0508	37.50	-0.154161	0.0644	-2.39
1974	-2.536236	0.0238	-106.57	1.841314	0.0458	40.18	-0.006697	0.0582	-0.12
1976	-2.545036	0.0358	-71.13	1.872427	0.0687	27.24	-0.076539	0.0866	-0.88
1977	-2.564476	0.0325	-78.86	1.855735	0.0623	29.77	-0.126989	0.0786	-1.62
1978	-2.654422	0.0353	-75.28	1.906954	0.0680	28.06	-0.155942	0.0837	-1.86
1979	-2.608433	0.0386	-67.54	1.868590	0.0750	24.92	-0.095427	0.0907	-1.05
1980	-2.697848	0.0295	-91.52	1.854805	0.0575	32.25	-0.055267	0.0682	-0.81
1981	-2.715723	0.0367	-74.01	1.917469	0.0711	26.95	-0.231196	0.0834	-2.77
1982	-2.721537	0.0372	-73.21	1.891316	0.0722	26.19	-0.006064	0.0854	-0.07
1983	-2.619136	0.0383	-68.31	1.907554	0.0744	25.63	0.026394	0.0870	0.30
1984	-2.742584	0.0455	-60.34	1.845451	0.0875	21.09	-0.029333	0.1027	-0.29
1985	-2.683218	0.0561	-47.83	1.942125	0.1088	17.86	0.014987	0.1226	0.12
1986	-2.779569	0.0599	-46.39	2.004058	0.1157	17.32	-0.148679	0.1274	-1.17
1987	-2.830635	0.0550	-51.43	2.002781	0.1073	18.67	-0.123694	0.1129	-1.10
1988	-2.765239	0.0455	-60.74	1.818199	0.0890	20.43	-0.040018	0.0939	-0.43
1989	-2.755176	0.0415	-66.42	1.848914	0.0811	22.81	0.047771	0.0840	0.57
1990	-2.801633	0.0491	-57.03	1.893994	0.0949	19.96	-0.197790	0.0977	-2.03
1991	-2.774284	0.0502	-55.24	1.910234	0.0974	19.61	-0.155815	0.0984	-1.58
1992	-2.903917	0.0695	-41.80	2.190260	0.1357	16.14	-0.203313	0.1259	-1.61
1993	-2.957471	0.0796	-37.17	2.125543	0.1558	13.65	-0.081956	0.1423	-0.58
1994	-2.759032	0.0557	-49.50	1.871659	0.1092	17.13	0.116824	0.1005	1.16

[1] The *t*-value is calculated as the estimate divided by the standard error

Having determined the observed parameter estimates using GLM procedures, the second stage in the modelling process is to identify appropriate time series models that describe the trends in the β_{jt} parameters (j = 0, 1, 2). To this end univariate time series models for each of the 3 parameters were determined according to the methods of Box and Jenkins (1976).

6.3.2 Univariate ARIMA models of the β_{jt} parameters

Time series modelling of each of the $\{\beta_{jt} : j = 0,1,2\}$ parameters was implemented in the S-Plus environment. To facilitate the analysis using methods available in S-Plus, the missing observations for calendar years 1968, 1971 and 1975 were approximated from the available observations to produce regularly spaced time-series processes.

For each of the series $\{\beta_{jt} : j = 0,1,2\}$, the models fitted were assumed to have a Gaussian distribution i.e. the $\{\varepsilon_{j,t}\}$ were assumed to be independent and identically distributed normal random variables with mean 0 and variance σ_j^2. Model identification was performed using Akaike information criterion (AIC).

The series $\{\beta_{0t} : t = 1,2,\cdots,36\}$ was modelled by a random walk with drift, δ:

$$\beta_{0,t} = \beta_{0,t-1} + \delta + \varepsilon_t \, . \tag{6.56}$$

The initial value $\beta_{0,0}$ was assumed to be the observed value in calendar year 1958. The k-steps ahead forecast function is therefore:

$$\tilde{\beta}_{0,n+k} = \beta_{0,n} + k\delta \quad , \tag{6.57}$$

where $\beta_{0,n}$ is the observed value of β_0 in calendar-year 1994 and k is the number of years after 1994.

The estimated value of the parameter δ is –0.010763 with standard error 0.012490. Thus the fitted model is:

$$\hat{\beta}_{0,t} = \hat{\beta}_{0,t-1} - 0.010763\,, \tag{6.58}$$

with σ_0^2 estimated as 0.005466.

The series $\{\beta_{1t} : t = 1,2,\cdots,37\}$ was modelled by an ARIMA(2,0,0) model (effectively an AR(2) model):

$$\beta_{1,t} - \hat{m} = \phi_1\left(\beta_{1,t-1} - \hat{m}\right) + \phi_2\left(\beta_{1,t-2} - \hat{m}\right) + \varepsilon_{1,t}\,, \tag{6.59}$$

where m is the process mean. The model fitted is:

$$\hat{\beta}_{1,t} - \hat{m} = 0.686683\left(\hat{\beta}_{1,t-1} - \hat{m}\right) - 0.363251\left(\hat{\beta}_{1,t-2} - \hat{m}\right), \tag{6.60}$$

with an estimate of σ_1^2 given by 0.005527 and standard error of the parameter estimates $\hat{\phi}_1$ and $\hat{\phi}_2$ given by 0.159784. The mean of the process is estimated as 1.891219 giving an estimate of ζ of 1.279537 (equation (6.26)).

The ARIMA(1,0,0) model fitted to the series $\{\beta_{2t} : t = 1,2,\cdots,37\}$ is:

$$\hat{\beta}_{2,t} - \hat{m} = 0.324959\left(\hat{\beta}_{2,t-1} - \hat{m}\right), \tag{6.61}$$

with σ_2^2 estimated as 0.007552, and standard error of the derived coefficient given by 0.162191. The process mean is estimated to be –0.099999 with ζ estimated as –0.0675035.

The models fitted are summarised in Table 6.2 and the predicted parameter values for the 20-year period from 1995 to 2014 are given in Table 6.3. As would be expected, predicted values of β_0 are declining steadily with time while the predicted values of β_1 and β_2 converge to their respective means. Figure 6.2 shows the predicted parameter values together with the associated approximate 95% confidence intervals.

Table 6.2
Time-series models fitted to observations of the parameters β_0, β_1 and β_2 over the period t = 1958-94

Parameter	Model	Parameter estimates (with standard errors)	Mean $\hat{m}$	$\hat{\sigma}^2$
β_{0t}	Random Walk with drift	$\hat{\delta} = -0.010763\,(0.012490)$	-2.63739	0.005466
β_{1t}	ARIMA(2,0,0)	$\hat{\phi}_1 = 0.686683\,(0.159784)$ $\hat{\phi}_2 = -0.363251\,(0.159784)$	1.891219	0.005527
β_{2t}	ARIMA(1,0,0)	$\hat{\phi} = 0.324959\,(0.162191)$	-0.099999	0.007552

Table 6.3
Female immediate annuitants, d1+ years: predicted parameter values for the period 1995-2014, based on time-series models fitted to observations over the period 1958-1994

Year(t)	β_0	β_1	β_2
1995	-2.769796	1.792669	-0.029541
1996	-2.780559	1.830651	-0.077103
1997	-2.791322	1.885426	-0.092559
1998	-2.802086	1.909242	-0.097581
1999	-2.812849	1.905699	-0.099214
2000	-2.823612	1.894615	-0.099744
2001	-2.834376	1.888291	-0.099916
2002	-2.845139	1.887974	-0.099972
2003	-2.855902	1.890054	-0.099990
2004	-2.866665	1.891598	-0.099996
2005	-2.877429	1.891902	-0.099998
2006	-2.888192	1.891550	-0.099999
2007	-2.898955	1.891198	-0.099999
2008	-2.909719	1.891084	-0.099999
2009	-2.920482	1.891134	-0.099999
2010	-2.931245	1.891209	-0.099999
2011	-2.942009	1.891243	-0.099999
2012	-2.952772	1.891239	-0.099999
2013	-2.963535	1.891224	-0.099999
2014	-2.974298	1.891215	-0.099999

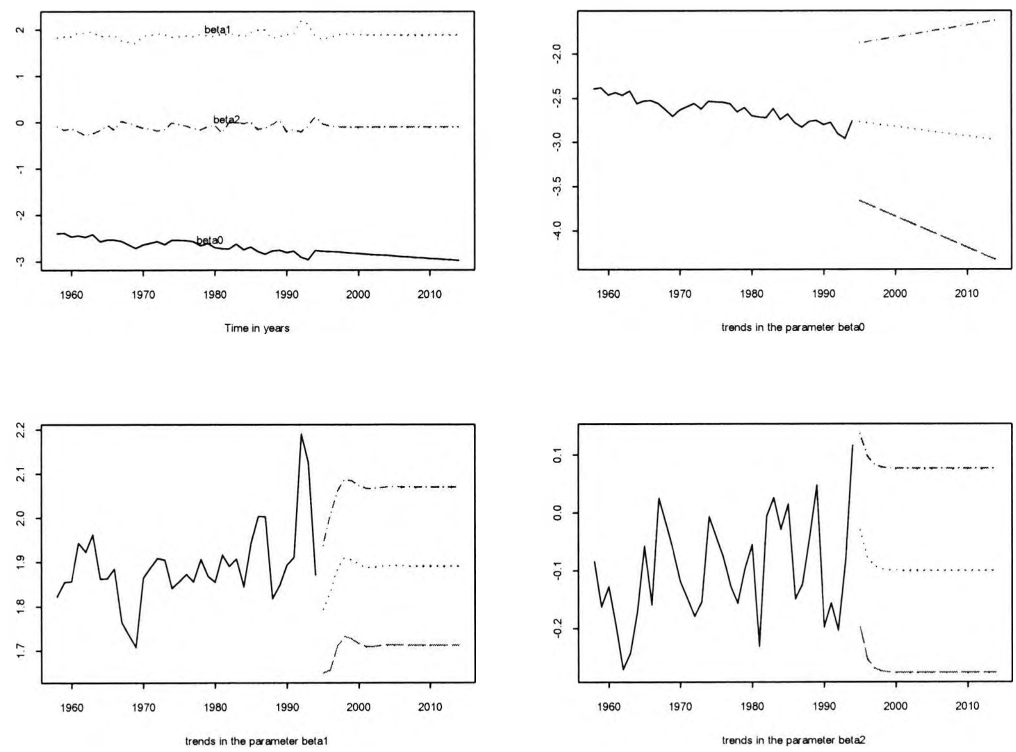


Figure 6.2 Female immediate annuitants, d1+ years: plots of observed (1958-1994) and predicted (1995-2014) parameters β_0, β_1 and β_2 over time t, together with approximate 95% confidence intervals for the predicted values

The final stage of the time-series modelling procedure applied is to use the parametric model (equation (6.54)) to predict future forces of mortality for female immediate annuitants with policy duration 1 year and over, assuming the time-series based forecast parameter values for each year of prediction. Figure 6.3 is a comparative plot of the forces of mortality thus predicted and the forces of mortality predicted from the 6-parameter log-link model derived from modelling the 1958-94 mortality experience for female annuitants. The forces of mortality have been predicted over a 20 calendar-year period from 1995 to 2014 for ages 55 to 110 and are shown at 5-year age intervals in the graph. The predicted forces of mortality applying in calendar years 1998, 2002, 2006, 2010 and 2014 are also shown in Table 6.4.

Table 6.4

Female immediate annuitants, d1+ years: predicted forces of mortality based on a 6-parameter log-link model for modelling forces of mortality and ARIMA models for forecasting parameters

age	1998	2002	2006	2010	2014
55	0.002210	0.002172	0.002068	0.001982	0.001899
56	0.002529	0.002484	0.002366	0.002268	0.002172
57	0.002891	0.002838	0.002704	0.002592	0.002482
58	0.003302	0.003240	0.003088	0.002959	0.002834
59	0.003768	0.003695	0.003522	0.003375	0.003233
60	0.004296	0.004209	0.004013	0.003846	0.003684
61	0.004893	0.004791	0.004569	0.004378	0.004194
62	0.005568	0.005448	0.005196	0.004979	0.004769
63	0.006330	0.006188	0.005904	0.005657	0.005419
64	0.007189	0.007023	0.006701	0.006421	0.006151
65	0.008157	0.007962	0.007599	0.007281	0.006974
66	0.009246	0.009018	0.008609	0.008248	0.007901
67	0.010471	0.010204	0.009743	0.009335	0.008942
68	0.011846	0.011534	0.011015	0.010554	0.010109
69	0.013390	0.013026	0.012442	0.011921	0.011419
70	0.015120	0.014695	0.014040	0.013452	0.012885
71	0.017058	0.016563	0.015828	0.015164	0.014525
72	0.019225	0.018650	0.017826	0.017078	0.016358
73	0.021647	0.020979	0.020056	0.019214	0.018404
74	0.024351	0.023576	0.022543	0.021597	0.020686
75	0.027367	0.026468	0.025314	0.024251	0.023229
76	0.030726	0.029686	0.028397	0.027204	0.026058
77	0.034465	0.033262	0.031825	0.030487	0.029203
78	0.038622	0.037233	0.035632	0.034134	0.032695
79	0.043240	0.041638	0.039855	0.038178	0.036569
80	0.048362	0.046518	0.044535	0.042661	0.040863
81	0.054041	0.051918	0.049716	0.047622	0.045616
82	0.060328	0.057890	0.055445	0.053109	0.050871
83	0.067282	0.064484	0.061774	0.059170	0.056677
84	0.074966	0.071760	0.068758	0.065859	0.063083
85	0.083448	0.079778	0.076457	0.073231	0.070145
86	0.092801	0.088606	0.084934	0.081349	0.077921
87	0.103103	0.098314	0.094259	0.090279	0.086475
88	0.114440	0.108979	0.104505	0.100090	0.095873
89	0.126902	0.120682	0.115751	0.110860	0.106188
90	0.140586	0.133512	0.128083	0.122667	0.117499
91	0.155598	0.147561	0.141589	0.135600	0.129886
92	0.172048	0.162928	0.156366	0.149749	0.143439
93	0.190055	0.179720	0.172517	0.165213	0.158252
94	0.209746	0.198049	0.190149	0.182095	0.174422
95	0.231256	0.218033	0.209378	0.200506	0.192057
96	0.254728	0.239799	0.230326	0.220562	0.211268
97	0.280315	0.263479	0.253121	0.242386	0.232173
98	0.308177	0.289214	0.277901	0.266110	0.254897
99	0.338485	0.317152	0.304807	0.291869	0.279571
100	0.371419	0.347449	0.333991	0.319808	0.306333
101	0.407167	0.380268	0.365611	0.350078	0.335328
102	0.445930	0.415778	0.399833	0.382839	0.366708
103	0.487916	0.454160	0.436830	0.418255	0.400632
104	0.533345	0.495600	0.476783	0.456500	0.437266
105	0.582447	0.540291	0.519881	0.497755	0.476783
106	0.635462	0.588436	0.566319	0.542206	0.519361
107	0.692640	0.640243	0.616301	0.590048	0.565188
108	0.754241	0.695930	0.670038	0.641484	0.614457
109	0.820536	0.755720	0.727748	0.696720	0.667366
110	0.891806	0.819844	0.789654	0.755972	0.724122

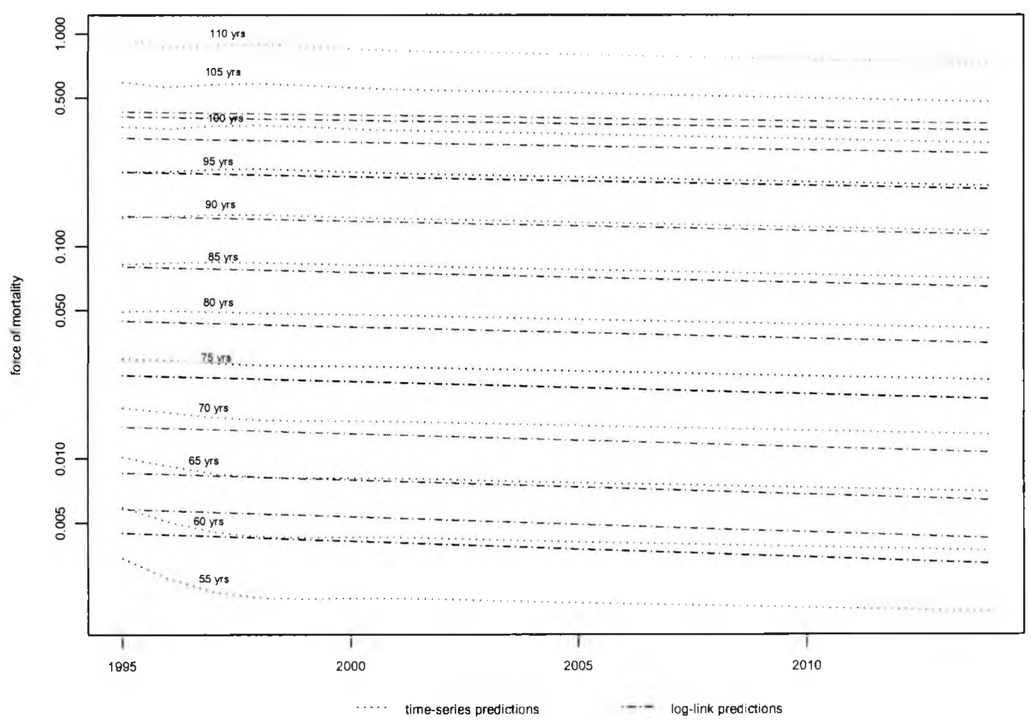


Figure 6.3 Female immediate annuitants, d1+ years: a comparison of predicted forces of mortality derived from a parametric model and predicted forces of mortality derived from a combination of parametric and univariate ARIMA models

From the plot in Figure 6.3, it can be observed that in general, forces of mortality projected on the basis of the parametric model are lower than the time-series based projected forces of mortality, except at ages below 65 years. The greatest differences occur at extreme old age, where the highest projected forces of mortality are just under 0.5 for the parametric model and just under 1 for the time-series model. However both modelling approaches would appear to produce forces of mortality that would be suitable for premium calculation and determining of reserves. It is interesting to note that based on the forces of mortality projected over the calendar-year period from 1995 to 2014, there is little difference between the forces of mortality projected on the basis of the two approaches at ages between 65 and 100 years (that is, the range of ages over which the data has been modelled).

The GLM modelling structure that has age x and time t as covariates might be preferred to the time series approach applied in this thesis. The GLM-approach is an integrated procedure that provides a single model to describe past trends in mortality and provide future predicted mortality rates. On the other hand the time series approach involves some three stages requiring a comparatively large number of computations: modelling the historical data for each calendar year t to determine an appropriate age-dependent model; determining forecast models for each of the parameters in the age-dependent model and deriving the forecasts; and using the forecast parameters to derive future mortality rates. It therefore seems that the integrated GLM procedure would be easier to implement. In addition, in general, the parametric model assumed here produces the lower forces of mortality, suggesting that this would be the preferred model for premium calculation and determining of reserves.

The univariate time series procedure applied in this section is based on the assumption that the series $\beta_{jt}, j = 1, 2, 3$ are independent. Empirical covariances computed seem to support this assumption. However, the parameter forecasts might be improved by modelling the trends using multivariate time series methods, thereby incorporating (any) interaction effects between the model parameters. This has however not been applied in this study. The computed covariances are presented in Appendix 4.

An alternative time series approach to determine mortality forecast models is the Lee-Carter method (Lee and Carter, 1992; Lee, 2000). The method, described in Chapter 3, has been applied to UK immediate annuitants' and pensioners' mortality experiences and compared with the GLM-based method by Renshaw and Haberman (2003b).

Chapter 7

Conclusion

7.1 Summary

The aim of the thesis is to investigate trends in the mortality of UK immediate annuitants and life office pensioners, and to develop a model or models suitable for predicting future patterns in mortality. The measure of mortality considered is the force of mortality. The salient points presented in the thesis are summarised in this section.

Chapter 2

Measures of mortality are discussed in general and methods of deriving crude mortality rates to estimate from observed mortality data, the force of mortality μ_x, and the initial rate of mortality q_x at age x, are also discussed. The concept of selection is briefly described and its relevance to annuitants' mortality experience noted. Various laws of mortality, i.e. mathematical formulae that have been used to describe levels of mortality, and methods of graduating mortality data are presented. Two graduation methods are described: the Continuous Mortality Investigation (CMI) Committee method of graduation and the extension to incorporate generalised linear models (GLMs) proposed by Renshaw (1991). The Renshaw *et al* (1996) GLM-based trend analysis methodology for the force of mortality is discussed in detail with a view to applying the same procedure to immediate annuitants' and pensioners' mortality data. The tests of graduation applied in the study are described.

Chapter 3

Methods of projecting mortality rates discussed by Pollard (1987) are described. The methods described are: projection by extrapolation of mortality rates; projection through parameters by reference to a law of mortality or using Bayesian graduation; projection using relational models, by relating the mortality under study to either a reference population, model life tables, a more advanced population, or an 'optimal' life table attainable under ideal conditions; and projection by cause of death. The time-series based Lee-Carter method of projecting central death rates is also described. Finally, methods that have been used in some countries of Western Europe to project mortality rates for annuity business are presented, with emphasis on the CMI model. The Renshaw *et al* (1996) GLM model is presented as an alternative to the CMI model. It is shown that a mortality improvement formula of the same basic format as the CMI mortality improvement model can be derived directly from the GLM-based graduation model that incorporates age and time. The mortality improvement formula is of the form:

$$\mu_{x,n} = \mu_{x,0} \cdot RF(x,n)$$

where,

$\mu_{x,0}$ is the force of mortality for a life attaining exact age x in the base calendar year (taken as year 0); that is, the base rate from the mortality table for the appropriate experience;

$\mu_{x,n}$ is the force of mortality for a life attaining exact age x in calendar year *base year* + n; and

$RF(x,n)$ is the reduction factor for an ultimate life attaining exact age x at time n, where n is measured in years from the base calendar year, thus $n = 1,2,\cdots$.

Chapter 4

The Renshaw *et al* (1996) modelling structure is used to investigate trends in the mortality experience of immediate annuitants' (ultimate). The ultimate experience analysed pertains to policyholders with policy duration 1 year and over and the data covers the calendar year period from 1946 to 1994.

It is shown that a trend analysis of the female experience over the period 1946-94 results in a 10-parameter model formula that is unsuitable for projections although the model provides a good fit to the observed experience. The trend adjustment term in this 10-parameter model is a polynomial of degree 4 in time t on the log scale. Predicted forces of mortality at future time periods are seen to exhibit a rapid improvement in mortality that is considered unrealistic. It is therefore concluded that the model is unsuitable for projections of forces of mortality at future time periods.

Taking note of the changes in the class of lives taking out immediate annuity contracts as a result of the Finance Act 1956, data prior to 1958 is then excluded and GLM models refitted to the experience. Two models are identified as providing the best fit to the observed experience. The first model consists of a Gompertz-Makeham term $GM_x(0,4)$ in combination with a trend adjustment term that is a linear function of t on the log scale, with a coefficient that is a quadratic function of age x (7 parameters). The second model is a 6-parameter model consisting of a $GM_x(0,4)$ term plus a 2-parameter trend adjustment term linear in time t on the log scale, with a coefficient that is linear in age x. It is shown that forces of mortality projected on the basis of the 7-parameter model do not progress smoothly at the extreme ages, particularly at ages outside the range of ages over which the model is fitted. On the other hand the 6-parameter model produces forces of mortality that progress smoothly with respect to both age and time and would therefore be a suitable model for mortality projections for annuity business. In addition, it is shown that a simple mortality improvement model can be readily derived from the trend model formula. The 6-parameter model is of the form:

$$\mu_{xt} = \exp\left[\beta_0 + \sum_{j=1}^{3}\beta_j L_j\left(\frac{x-c_x}{w_x}\right) + \left\{\alpha + \gamma\left(\frac{x-c_x}{w_x}\right)\right\}\left(\frac{t-c_t}{w_t}\right)\right],$$

and the corresponding mortality improvement model is of the form:

$$RF(x,n) = \exp\left[\frac{n}{w_t}\left\{\alpha + \gamma\left(\frac{x - c_x}{w_x}\right)\right\}\right].$$

$L_j(x)$ are Legendre polynomials; c_x, w_x, c_t and w_t are chosen such that age x and calendar-year t respectively, are each mapped onto the interval [-1,+1] in the model structure; and n is time measured in years from an appropriately chosen base calendar-year.

The best-fitting model for the male immediate annuitants' ultimate mortality experience over the period 1946 to 1994 is shown to be a 6-parameter model consisting of a $GM_x(0,3)$ term in age effects and a trend adjustment term that is quadratic in time t on the log scale, with coefficients that are linear in age x. Projected forces of mortality on the basis of this model are found to have some undesirable features at extreme old age: the rates tend to increase with time and also exhibit a rapid increase with age at ages above 95 years.

As for the females' experience, data prior to 1958 is then excluded from the analysis of the male annuitants' mortality experience. The simplest model providing a good fit to the male immediate annuitants' mortality experience over the period 1958 to 1994 is found to be a 3-parameter model consisting of a $GM_x(0,2)$ term in age effects and an age-independent trend adjustment term linear in time t on the log scale. The projected forces of mortality for male annuitants based on this 3-parameter model are compared with projected forces of mortality based on the same model structure as for the corresponding female experience, and a third model consisting of a $GM_x(0,2)$ term in age effects and a trend adjustment term of the same format as for the female annuitants' 6-parameter model. The 6-parameter model structure is shown to provide the lowest projected forces of mortality over a 20-year period. It is however suggested that the choice between the three models fitted would be a matter of personal judgement since all three models exhibit patterns suitable for projections, although the 3-parameter model would be an unlikely candidate because of the age-independent trend adjustment term. It can therefore be inferred that the GLM-based modelling structure can be used to predict future forces of mortality for immediate annuitants when the trend adjustment term is linear in time t on the log scale and the coefficient of this term is itself linear in age effects.

The forces of mortality for male annuitants predicted on the basis of a trend analysis of the 1958-94 experience are however seen to exhibit a rapid increase with age at extreme old age (over 100 years). It is suggested that a projection model for male annuitants based on the experience from 1958, with equal weighting given to each calendar year, does not adequately reflect the improvements in mortality in the male population in the more recent years. In order to give more weight to the mortality experience in the more recent years, two approaches are presented. One approach taken is to fit the 'adopted' 6-parameter model structure to the male annuitants' ultimate experience over a period that excludes the earlier years. In this case the period excluded is the calendar years prior to 1974. A second approach is to estimate the model parameters on the basis of the 1958-94 experience using likelihood weighted by calendar year, with the greatest weight given to the most recent calendar year of the investigation period. In both cases, it is shown that the rapid increase of forces of mortality with age at extreme old age is greatly reduced, producing rates that are preferable for annuity pricing.

As an illustration of a complete GLM modelling procedure for mortality data that includes select data, the method proposed by Renshaw and Haberman (1997) is applied to female annuitants' data. In this specific case the ultimate experience is taken to be policies with duration 5 years and over.

Chapter 5

The results of applying the GLM modelling structure to life office pensioners' mortality experience over the 14-year period from 1983 to 1996 are presented. The best fitting model for male pensioners is shown to be a 7-parameter model consisting of a $GM_x(0,4)$ term in combination with a trend adjustment term that is quadratic in time t on the log scale. Projections based on this 7-parameter model are shown to exhibit unrealistically rapid improvements in mortality with time. Fitting the 6-parameter model structure fitted to the annuitants' experiences is shown to result in reasonable projections for both male and female pensioners at most ages, particularly at ages below 100 years.

It is therefore concluded that the GLM-based modelling structure proposed by Renshaw *et al* (1996) can be used to predict future forces of mortality for immediate annuitants and pensioners provided the trend adjustment term is linear in time t on the log scale and the coefficient of this term is itself linear in age effects. Including interaction terms involving higher order terms in x results in forces of mortality that do not progress smoothly at the extreme ages. On the other hand, introducing higher order terms in time t results in unrealistically rapid improvements in mortality.

It is suggested that the future mortality pattern for pensioners could be better projected by a model based on a longer period of observation (which is not available), or by using the lower confidence limits derived for the parameters in the model formula.

Despite the short observation period over which the pensioners' projected forces of mortality are based, the mortality improvement factors derived for life office pensioners are shown to be consistent with CMI mortality improvement factors, which are based on the pensioners' mortality experience. It is also shown that the improvement factors derived from modelling the immediate annuitants' experience do not have a similar pattern. The underlying mortality trends in the two experiences are different. For example, because the annuitants exercise some degree of self-selection, there is less improvement in mortality over time and less variation in mortality changes over age than there is for pensioners. The selection processes for immediate annuitants and for pensioners are different: pensioners' population effectively corresponds to compulsory purchase of annuity, whereas annuitant population effectively corresponds to voluntary purchase of annuity. Finkelstein and Poterba (2002) have explored adverse selection in the voluntary and compulsory individual annuity markets in the UK. They observe that the mortality differences between annuitants and non-annuitants are more pronounced in the voluntary than the compulsory annuity market and estimate that the amount of adverse selection in the compulsory market is about one half of that in the voluntary sector. Thus, mortality improvement factors appropriate for the pensioners' experience are not necessarily appropriate for the immediate annuitants' experience.

It is also shown that to some extent, the period mortality improvement formula derived can be considered equivalent to a mortality improvement model based on year of birth (cohort) experience.

Chapter 6

Consideration of an alternative approach to modelling time trends in annuitants' mortality using time series methods is made for comparative purposes. The time series approach applied is that used by McNown and Rogers (1989), a method that combines parametric models and time-series models to generate forecasts. Univariate time series procedures are used to forecast each of the parameters in the age-dependent model for the force of mortality μ_x and then the forecast parameters are applied in the age dependent model to project forces of mortality at future time periods, t.

The data set considered is the female annuitants' experience over the period 1958 to 1994. Projected forces of mortality based on the GLM modelling procedure and the time-series based modelling procedure are found to be quite similar, particularly within the range of ages over which the data is modelled. It is however suggested that the 6-parameter log-link model (with a linear trend adjustment term in time t on the log scale), derived in Chapter 4 is the better model since the forces of mortality projected on the basis of the univariate time-series models are generally higher than the forces of mortality projected on the basis of the parametric model. It is further suggested that the GLM modelling procedure is more attractive to apply as it is an integrated procedure providing a single model to describe past trends in mortality and to produce projected forces of mortality.

The Renshaw *et al* (1996) GLM modelling procedure is such that the model determined from analysing the historical data incorporates an appropriate mortality improvement model for predicting future mortality rates. Hence a mortality improvement model is determined statistically rather than by using *ad hoc* methods, as is the case with the CMI model. Notwithstanding this, it is recognised that, in choosing appropriate mortality reduction factors, a degree of personal judgement will always be exercised bearing in mind that overestimating mortality will endanger the company's financial position; while an underestimate will undermine the company's competitive position.

7.2 Areas of Further Research

Projection using Bayesian methods

The GLM-based projected forces of mortality are based on a deterministic approach for future forces of mortality. An area of further research would be the derivation of stochastic models for $\mu_{x,t}$ using Bayesian graduation methods (e.g. Carlin, 1992, Dellaportas, Smith and Stavropoulos, 2001)

Year of birth cohort mortality improvement model

In Chapter 5, Section 5.6, it is suggested that the period (year of attaining age x) mortality improvement model is equivalent to a year of birth cohort mortality improvement model. A detailed investigation of the annuitants' and pensioners' mortality experiences by year of birth should provide empirical evidence in support (or otherwise) of this assertion.

Time-series modelling of parameters in the age-dependent model

The time series approach applied in this thesis is to use univariate time series procedures to forecast each of the parameters in the age-dependent model, ignoring any interaction effects between the parameters. It would be worthwhile to consider multivariate time series methods that would then incorporate interaction effects in the modelling procedure.

Investigation of links between GLM-based mortality forecasting methods and Lee-Carter framework

Links between the GLM approach to forecasting mortality data and the Lee-Carter time-series approach have not been investigated in this thesis. A detailed investigation of the links between the methods would be desirable. Renshaw and Haberman (2003b) have suggested ways in which the Lee-Carter method can be modified to forecast mortality reduction factors and have noted similarities in the results with results obtained from applying GLM forecasting methods.

Investigation of links between GLM-based mortality forecasting methods and the log-bilinear approach

The 6-parameter log-link model proposed in this thesis is very close to the log-bilinear approach briefly described in Chapter 3, Section 3.6. There is need for a careful comparison of the two approaches in order to identify their respective drawbacks.

Appendix 1

Likelihood functions for generalized linear models

A vector of observations **y** having n components is assumed to be a realisation of a random variable $\boldsymbol{Y}$ whose components are independently distributed with means $\boldsymbol{m}$. Each component of $\boldsymbol{Y}$ is assumed to have a distribution in the exponential family, taking the form:

$$f_Y(y;\theta,\phi) = \exp\left(\frac{y\theta - b(\theta)}{a(\phi)} + c(y,\phi)\right). \tag{A1.1}$$

for some specific functions $a(.)$, $b(.)$ and $c(.)$. If ϕ is known, the distribution of $\boldsymbol{Y}$ is a one-parameter exponential family with canonical parameter θ. If ϕ is unknown the distribution may or may not be a two-parameter exponential family.

The mean and variance of $\boldsymbol{Y}$ can be derived from the log-likelihood function:

$$\ell = \log f_Y(y;\boldsymbol{\theta},\boldsymbol{\phi}) = \frac{\{y\boldsymbol{\theta} - b(\boldsymbol{\theta})\}}{a(\boldsymbol{\phi})} + c(y,\boldsymbol{\phi}) \tag{A1.2}$$

and the relations:

$$E\left(\frac{\partial \ell}{\partial \boldsymbol{\theta}}\right) = 0 \tag{A1.3}$$

and

$$E\left(\frac{\partial^2 \ell}{\partial \boldsymbol{\theta}^2}\right) + E\left(\frac{\partial \ell}{\partial \boldsymbol{\theta}}\right)^2 = 0. \tag{A1.4}$$

From (A1.2) we have

$$\frac{\partial \ell}{\partial \boldsymbol{\theta}} = \frac{\{y - b'(\boldsymbol{\theta})\}}{a(\boldsymbol{\phi})} \tag{A1.5}$$

and

$$\frac{\partial^2 \ell}{\partial \theta^2} = -\frac{b''(\theta)}{a(\phi)}. \qquad \text{(A1.6)}$$

From (A1.3) and (A1.5), we have

$$0 = E\left(\frac{\partial \ell}{\partial \theta}\right) = \frac{\{m - b'(\theta)\}}{a(\phi)},$$

so that

$$E(Y) = m = b'(\theta).$$

Similarly, from (A1.4), (A1.5) and (A1.6) we have

$$0 = -\frac{b''(\theta)}{a(\phi)} + \frac{\operatorname{var}(Y)}{a^2(\phi)},$$

so that

$$\operatorname{var}(Y) = b''(\theta)a(\phi).$$

Appendix 2

Confidence intervals for predicted forces of mortality in 2014 based on the 6-parameter log-link models derived from the 1974-94 annuitants' mortality experiences

Figure A2.1

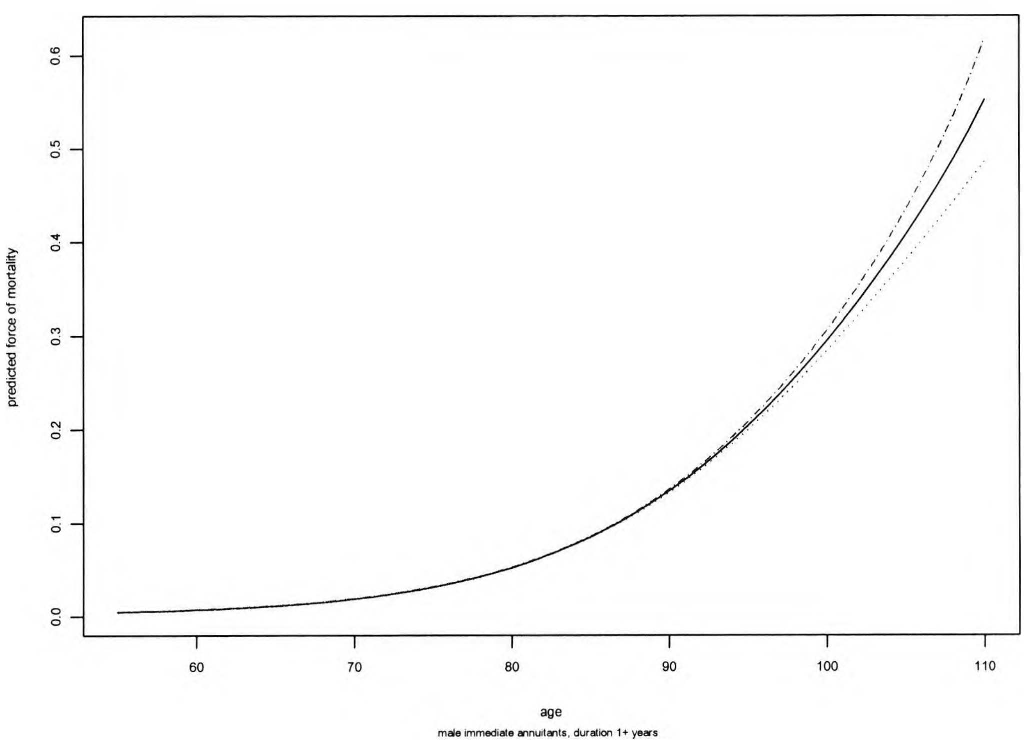
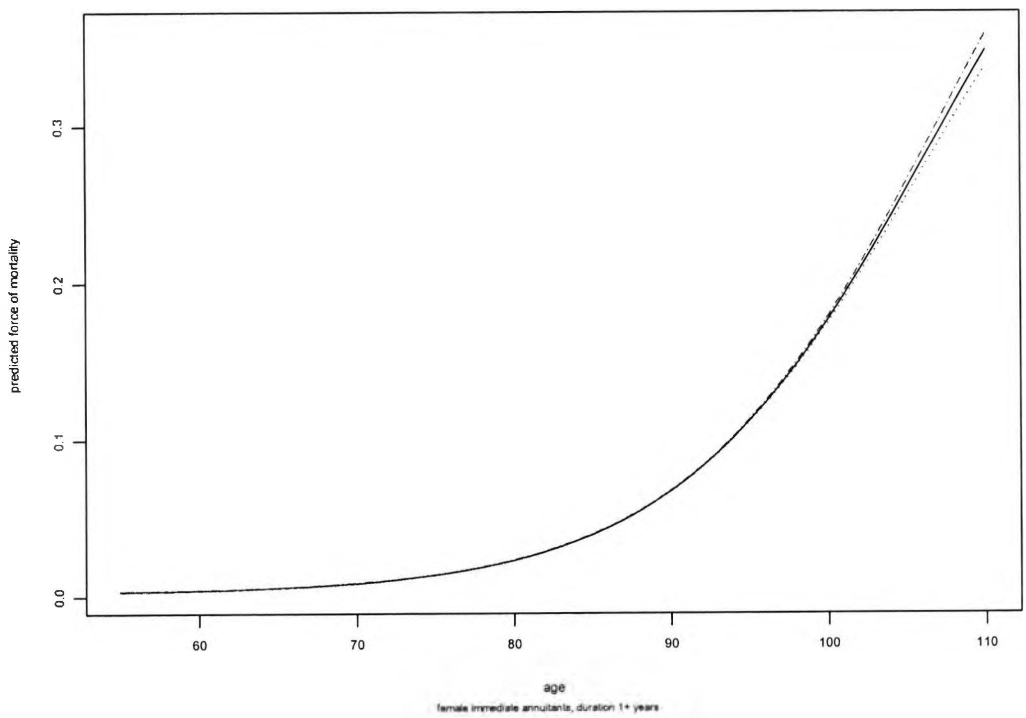


Figure A2.2


predicted force of mortality
0.0
0.1
0.2
0.3
60
70
80
90
100
110
age
female immediate annuitants, duration 1+ years

Appendix 3

Derivation of variance and covariance functions for ARMA(1,1) process

An ARMA(1,1) process is defined by:

$$Y_t = \zeta + \phi Y_{t-1} + \varepsilon_t + \theta\varepsilon_{t-1}. \qquad \text{(A3.1)}$$

where the $\{\varepsilon_t\}$ is a white noise process with mean 0 and variance σ^2. Taking expectations of (A3.1) and noting that $E(Y_t) = m$ for all t, we have

$$m = \frac{\zeta}{1-\phi}, \qquad \text{(A3.2)}$$

or

$$\zeta = m(1-\phi). \qquad \text{(A3.3)}$$

Using the result (A3.3), (A3.1) may be expressed as:

$$Y_t = m(1-\phi) + \phi Y_{t-1} + \varepsilon_t + \theta\varepsilon_{t-1},$$

that is:

$$Y_t - m = \phi(Y_{t-1} - m) + \varepsilon_t + \theta\varepsilon_{t-1}. \qquad \text{(A3.4)}$$

The covariances of Y_t and the white noise terms ε_t and ε_{t-1} are derived as follows:

$$\text{cov}(Y_t, \varepsilon_t) = \text{cov}(\zeta + \phi Y_{t-1} + \varepsilon_t + \theta\varepsilon_{t-1}, \varepsilon_t) = \sigma^2 \qquad \text{(A3.5)}$$

$$\text{cov}(Y_t, \varepsilon_{t-1}) = \text{cov}(\zeta + \phi Y_{t-1} + \varepsilon_t + \theta\varepsilon_{t-1}, \varepsilon_{t-1}) = \sigma^2(\phi + \theta). \qquad \text{(A3.6)}$$

The variance function γ_0, is given by:

$$\gamma_0 = \text{cov}(Y_t, Y_t) = \text{cov}(\zeta + \phi Y_{t-1} + \varepsilon_t + \theta\varepsilon_{t-1}, Y_t) = \phi\gamma_1 + \sigma^2 + \theta\sigma^2(\phi + \theta). \qquad \text{(A3.7)}$$

Using similar relationships, it is found that:

$$\gamma_1 = \text{cov}(Y_t, Y_{t-1}) = \phi\gamma_0 + \theta\sigma^2, \tag{A3.8}$$

$$\gamma_2 = \text{cov}(Y_t, Y_{t-2}) = \phi\gamma_1 \tag{A3.9}$$

and the recurrence relation

$$\gamma_j = \phi\gamma_{j-1} \text{ for all } j \geq 2. \tag{A3.10}$$

Expressing γ_1 (A3.8) in terms of γ_0 in the expression for γ_0 (A3.7), the results

$$\gamma_0 = \sigma^2 \frac{(1 + 2\phi\theta + \theta^2)}{1 - \phi^2}, \tag{A3.11}$$

and

$$\gamma_1 = \frac{\sigma^2(\phi + \theta)(1 + \phi\theta)}{1 - \phi^2} \tag{A3.12}$$

Appendix 4

Female immediate annuitants, d1+ years: covariances of parameter estimates based on $GM_x(0,3)$ graduation models for each calendar year 1958 to 1994 (excluding 1968, 1971 and 1975)

year	cov(β_0,β_1)	cov(β_0,β_2)	cov(β_1,β_2)
1958	-0.0003	0.0026	-0.0001
1959	-0.0003	0.0019	-0.0003
1960	-0.0006	0.0028	-0.0008
1961	-0.0005	0.0017	-0.0008
1962	-0.0005	0.0016	-0.0008
1963	-0.0007	0.0019	-0.0010
1964	-0.0006	0.0019	-0.0008
1965	-0.0004	0.0013	-0.0006
1966	-0.0004	0.0011	-0.0005
1967	-0.0002	0.0008	-0.0002
1969	-0.0006	0.0019	-0.0008
1970	-0.0006	0.0016	-0.0009
1972	-0.0008	0.0017	-0.0012
1973	-0.0005	0.0010	-0.0008
1974	-0.0004	0.0008	-0.0007
1976	-0.0011	0.0019	-0.0020
1977	-0.0010	0.0016	-0.0018
1978	-0.0013	0.0019	-0.0024
1979	-0.0016	0.0023	-0.0030
1980	-0.0010	0.0013	-0.0019
1981	-0.0017	0.0022	-0.0031
1982	-0.0016	0.0022	-0.0030
1983	-0.0018	0.0023	-0.0033
1984	-0.0026	0.0034	-0.0047
1985	-0.0043	0.0050	-0.0077
1986	-0.0052	0.0058	-0.0092
1987	-0.0047	0.0047	-0.0080
1988	-0.0032	0.0032	-0.0054
1989	-0.0027	0.0026	-0.0045
1990	-0.0038	0.0036	-0.0062
1991	-0.0041	0.0037	-0.0065
1992	-0.0083	0.0067	-0.0127
1993	-0.0109	0.0084	-0.0163
1994	-0.0053	0.0040	-0.0079

Bibliography and References

Anderson, J. L. & Dow, J. B. (1964). *Actuarial Statistics*, vol. 2. Cambridge University Press, Cambridge.

Anderson, J. L. & Whitehead, D. (1960). *The Trend of Mortality Rates in Seventeen Countries in the Last Twenty Years*. Transactions of the 16th International Congress of Actuaries, **2**, 379-400.

Andrews, G. H. & Nesbitt, C. J. (1965). *Periodograms of Graduation Operators*. T.S.A., **XVII**, I, 1-27.

Anscombe, F. J. (1961). *Examination of Residuals.* Proc. Fourth Berkeley Symposium, **1**, 1-36.

Batten, R. W. (1978). *Mortality Table Construction.* Prentice-Hall, Inc., Englewood Cliffs, New Jersey.

Beard, R. E. & Perks, W. (1949). *The Relation between the Distribution of Sickness and the Effect of Duplicates on the Distribution of Deaths*. J.I.A., **75**, 75-86.

Benjamin, B. (1964). *Demographic and Actuarial Aspects of Ageing, with Special Reference to England and Wales*. J.I.A., **90, III**, 211-253.

Benjamin, B. (1982). *The Span of Life*. J.I.A., **109**, 319-340.

Benjamin, B. & Pollard, J. H. (1993). *The Analysis of Mortality and other Actuarial Statistics,* 3rd Edition. Institute of Actuaries and Faculty of Actuaries.

Benjamin, B. & Soliman, A. S. (1993). *Mortality on the Move*. Actuarial Education Service.

Bloomfield, D. S. F. & Haberman, S. (1987). *Graduation: Some Experiments with Kernel Methods*, J.I.A., **114**, 339-369.

Bourgeois-Pichat, J. (1952). *Essai sur la mortalité 'biologique' de l'homme*. Population, **3**, 381-394.

Bowers JR, N. L., Gerber, H. U., Hickman, J. C., Jones, D. A., Nesbitt, C. J. (1986). *Actuarial Mathematics*. The Society of Actuaries.

Box, G. E. P., & Jenkins, G. M. (1976). *Time Series Analysis, Forecasting and Control.* Holden-Day, San Francisco.

Box, G. E. P., Jenkins, G. M. & Reinsel, G. C. (1994). *Time Series Analysis, Forecasting and Control*. Prentice Hall, New Jersey.

Box, G. E. P., & Pierce, D. A. (1970). *Distribution of Residual Auto-correlations in Autoregressive-integrated Moving Average Time-series Models.* Journal of the American Statistical Association, **65**, 1509-1526.

Brass, W. (1971). *On the Scale of Mortality* in Brass, W., ed., *Biological Aspects of Demography.* Taylor and Francis, London.

Brouhns, N., Denuit, M. & Vermunt, J. K. (2002). *A Poisson Log-bilinear Regression Approach to the Construction of Projected Lifetables.* Insurance: Mathematics and Economics, **31,** 373-393.

Carlin, B. P. (1992). *A Simple Monte Carlo Approach to Bayesian Graduation.* Transactions of the Society of Actuaries, **44**, 55-76.

Carriere, J. F. (1994). *Dependendent Decrement Theory*, Transactions of the Society of Actuaries, **46**, 45-74.

Carter, L. & Lee, R. D. (1992). *Modelling and Forecasting U. S. Mortality: Differentials in Life Expectancy by Sex.* International Journal of Forecasting, **8**, no. 3 (November), 393-412.

Chambers, J. M. & Hastie T. J. (1993). *Statistical Models in S.* Chapman and Hall, London.

Chatfield, C. (1996). *The Analysis of Time Series.* Fifth Edition, Chapman and Hall/CRC.

CMI Committee (1957). *Continuous Investigation into the Mortality of Assured Lives: Memorandum on a Special Inquiry into the Distribution of Duplicate Policies.* J.I.A., **83, I**, 34-36 and T.F.A., **24**, 94.

CMI Committee (1976). *The Graduation of Pensioners' and Annuitants' Mortality Experience* 1967-70. CMIR, **2**, 57.

CMI Committee (1986). *An Investigation into the Distribution of Policies per Life Assured in the Cause of Death Investigation Data.* CMIR, **8,** 49-.

CMI Committee (1988). *The Graduation of the 1979-1982 Mortality Experiences.* CMIR, **9,** 1-102.

CMI Committee (1990). *Standard Tables of Mortality Based on the 1979-1982 Experiences.* CMIR, **10**.

CMI Committee (1995). *The Mortality of Assured Lives, Pensioners and Annuitants 1987-90.* CMIR, **14**, 1-77.

CMI Committee (1998). *The Mortality of Immediate Annuitants, Holders of Retirement Annuity Policies, and Holders of Personal Pension Plans 1991-1994.* CMIR, **16**, 45-64.

CMI Committee (1998). *The Mortality of Pensioners in Insured Group Pension Schemes 1991-1994.* CMIR, **16**, 65-82.

CMI Committee (1998). *Proposed New Tables for Life Office Pensioners, Normal, Male and Female, based on the 1991-1994 Experiences.* CMIR, **16**, 45-51.

CMI Committee (1999). *Standard Tables of Mortality based on the 1991-94 Experiences*. CMIR, **17**.

Coale, A. J. & Demeny, P. (1966). *Regional Model Life Tables and Stable Populations*. Princeton University Press, Princeton.

Coale, A. J. & Guo, G. (1989). *Revised Regional Model Life Tables at Very Low Levels of Mortality.* Population Index **55**: 613-643.

Coale, A. J. & Kisker, E. E. (1990). *Defects in Data on Old-Age Mortality in the United States: New Procedures for Calulating Mortality Schedules and Life Tables at the Highest Ages.* Asian Pacific Population Forum 4, **1**, 1-31.

Congdon, P. (1993). *Statistical Graduation in Local Demographic Analysis and Projection*, J. R. Statistic. Soc. A, **156**, 237-270.

Conte, S. D. & de Boor, C. (1980), *Elementary Numerical Analysis,* Third Edition. McGraw-Hill, Kogashuka, Tokyo.

Copas, J. & Haberman, S (1983). *Non-parametric Graduation using Kernel Methods*, JIA, **110**, 135-156.

Cox, D. R. (1983). *Some Remarks on Overdispersion.* Biometrika, **70**, 1, 269-274.

Cox, D. R. & Oakes, D. (1984). *Analysis of Survival Data.* Chapman and Hall, London.

Cramér, H. & Wold, H. (1935). *Mortality Variations in Sweden: A study in Graduation and Forecasting.* Scandinavian Actuarial Journal, **11**, 161-241.

Cryer, J.D. (1986). *Time Series Analysis.* Duxbury Press, Boston.

Daw, R. H. (1946). *On the Validity of Statistical Tests of the Graduation of a Mortality Table.* J.I.A., **72,** 174-202.

Daw, R. H. (1951). *Duplicate Policies in Mortality Data.* J.I.A., **77**, 261-267.

Dellaportas, P., Smith, A. F. M. & Stavropoulos, P. (2001). *Bayesian Analysis of Mortality Data.* J. R. Statist. Soc. A, **164**, 275-291.

Dobson, A. J. (1990). *An Introduction to Generalized Linear Models.* Chapman and Hall, London.

Dominion Bureau of Statistics (1950). *Memorandum on the Projection of Population Statistics*. (84-D-69), Ottawa.

Dominion Bureau of Statistics (1954). *Memorandum on the Projection of Population Statistics*. (84-D-69A), Ottawa.

Elandt-Johnson, R. C. & Johnson, N. L. (1980). *Survival Models and Data Analysis*. John Wiley and Sons, New York.

Elphinstone, M. D. W. (1951). *Summation and Some Other Methods of Graduation – The Foundations of Theory*. TFA, **XX**, 15-57.

Felipe, A., Guillen, M. & Artis, M. (1998). *Recent Mortality Patterns in the Spanish Population*. Transactions of the 26th International Congress of Actuaries, **9,** 55-74.

Felipe, A., Guillen, M. & Nielsen, J. P. (2001). *Longevity Studies based on Kernel Hazard Estimation*. Insurance: Mathematics and Economics, **28**, 191-204.

Finkelstein, A. & Poterba, J. (2002). *Selection Effects in the United Kingdom Individual Annuities Market*. The Economic Journal, **112**, 28-50.

Firth, D. (1987). *On the Efficiency of Quasi-likelihood Estimation*. Biometrika, **74**, 2, 233-245.

Forfar, D. O., McCutcheon, M. A. & Wilkie, A. D. (1988). *On Graduation By Mathematical Formula*. J.I.A., **115,** 1-135 and T.F.A., **41**, 97.

Forfar, D. & Smith, D. (1987). *The Changing Face of English Life Tables*. T.F.A., **40**, 98.

Gavin, J. B., Haberman, S. & Verrall, R. J. (1993). *Moving Weighted Average Graduation using Kernel Estimation*. Insurance: Mathematics and Economics, **12**, 113-126.

Gavin, J. B., Haberman, S. & Verrall, R. J. (1994). *On the Choice of Bandwidth for Kernel Graduation*. J.I.A., **121**, 119-134.

Gavin, J. B., Haberman, S. & Verrall, R. J. (1995). *Graduation by Kernel and Adaptive Kernel Methods with a Boundary Correction*. Transactions of the Society of Actuaries, **47**, 1-38.

Gerber, H. U. (1995). *Life Insurance Mathematics*, Second Edition. Springer-Verlag, Berlin, Heidelberg.

Gilchrist, R. (1982). *GLIM*82: *Proceedings of the International Conference on Generalized Linear Models*. Springer-Verlag, New York.

Giles, P. & Wilkie, A. D. (1973). *Recent Mortality Trends*: *Some International Comparisons*. T.F.A., **33**, 375-514.

Golulapati, R., De Ravin, J. W. & Trickett, P. J. (1984). *Projections of Australian Mortality Rates*. Occasional Paper No 1983/2, Australian Bureau of Statistics, Canberra.

Gompertz, B. (1825). *On the Nature of the Function of the Law of Human Mortality and on a New Mode of Determining the Value of Life Contingencies*. Phil. Trans. of Royal Society, **115**, 513-585.

Graunt, J. (1662). *Natural and Political Observations Mentioned in a following Index, and made upon the Bills of Mortality*. Reprinted by the Institute of Actuaries (1964), J.I.A. **90**, 1-61.

Gwilt, R. L. (1956). *Mortality in the Past Hundred Years*. T.F.A. **24**, 40-176.

Haberman, S. & Renshaw, A. E. (1996). *A Different Perspective on U.K. Assured Lives Select Mortality*. Actuarial Research Paper No. **92**, Department of Actuarial Science and Statistics, City University, London.

Haberman, S. & Renshaw, A. E. (1996). *Generalized Linear Models and Actuarial Science*. The Statistician, **45**, 1-31.

Harvey, A. C., (1990). *The Econometric Analysis of Time Series*. Philip Allan, Hemel Hempstead, Hertfordshire.

Harvey, A. C., (1993). *Time Series Models*. Harvester Wheatsheaf, Hemel Hempstead, Hertfordshire.

Hamilton, J. D. (1994). *Time Series Analysis*. Princeton University Press, Princeton, New Jersey.

Heligman, L. & Pollard, J. H. (1980). *The Age Pattern of Mortality*. J.I.A. **107**, 49-80

Hickman, J. C. & Miller, R. (1981). *Bayesian Bivariate Graduation and Forecasting*. Scandinavian Actuarial Journal, **57**, 129-150.

Hooker, P. F. & Longley-Cook, L. H. (1953). *Life and Other Contingencies*. Cambridge University Press.

Humphrey, G. T. (1970). *Mortality at the Oldest Ages*. J.I.A. **96**, 105.

Hustead, E. C. (1989). 100 *Years of Mortality*. Society of Actuaries.

Jenkins, W. A. & Lew, E. A. (1949). *A New Mortality Basis for Annuities*. Transactions of The Society of Actuaries, **I**, 369-498.

Joint Mortality Investigation Committee (1974). *Consideration Affecting the Preparation of Standard Tables of Mortality*. J.I.A., **101**, II, 133-216.

Jordan, C. W. (1967). *Life Contingencies*. Second Edition, Society of Actuaries, Chicago.

Keyfitz, N. (1968). *Introduction to the Mathematics of Population with revisions.* Addison-Wesley, Reading, Massachusetts.

Keyfitz, N. (1977, 1985). *Applied Mathematical Demography.* Second Edition, Springer-Verlag, New York.

Keyfitz, N. (1982). *Choice of Function for Mortality Analysis: Effective Forecasting Depends on a Minimum Parameter Representation.* Theoretical Population Biology, **21**, 329-352.

Klugman, S. A., Panjer, H. H. & Willmot, G. E. (1998). *Loss Models From Data to Decisions.* John Wiley & Sons, Inc., New York.

Lee, R. (2000). *The Lee-Carter Method for Forecasting Mortality, with Various Extensions and Applications.* North American Actuarial Journal, **4**, 1, 80-93.

Lee, R. D. & Carter, L. R. (1992). *Modelling and Forecasting U. S. Mortality.* Journal of the American Statistical Association, **87**, 419, 659-675.

London, D. (1985). *Graduation: The Revision of Estimates.* ACTEX, Winsted.

London, D. (1988). *Survival Models and their Estimation.* 3rd Edition, ACTEX, Winsted.

Macdonald, A. S. (ed) (1997). *The Second Actuarial Study of Mortality in Europe.* Groupe Consulatatif des Associations d'Actuaires des Pays des Communautés Européennes.

Macdonald, A. S., Cairns, A. J. G., Gwilt, P. L. & Miller, K. A. (1998). *An International Comparison of Recent Trends in Population Mortality.* B.A.J., **4**, I, 3-141.

Makeham, W. M. (1860). *On the Law of Mortality.* J.I.A., **13**, 325-358.

McCullagh, P. (1983). *Quasi-likelihood functions.* The Annals of Statistics, **7**, 1-26.

McCullagh, P. & Nelder, J. A. (1989). *Generalized Linear Models.* 2nd Edition, Chapman and Hall, London.

McCutcheon, J. J. (ed.) (1986). *An Actuarial Study of Mortality in the Countries of the European Communities.* Groupe Consulatatif des Associations d'Actuaires des Pays des Communautés Européennes.

McNown, R. & Rogers, A. (1989). *Forecasting Mortality: A Parameterized Time Series Approach.* Demography, **26**, 4, 645-660.

Myers, R. J. & Bayo, F. R. (1985). *United States Life Tables for 1979-1981.* Transactions of the Society of Actuaries, **XXXVII**, Part I, 303-350.

Nelder, J. A. & Pregibon, D. (1987). *An Extended Quasi-Likelihood Function.* Biometrika, **74**, 2, 221-232.

Nelder, J. A. & Wedderburn, R. W. M. (1972). *Generalized Linear Models.* J. Royal Statist. Soc. A, **135**, 370-384.

Neill, A. (1977). *Life Contingencies.* Heinemann, London.

Olivieri, A. & Pitacco, E. (2002). *Inference about Mortality Improvements in Life Annuity Portfolios.* Transactions of the 27th International Congress of Actuaries, Cancun, Mexico.

Okazaki, A. (1954). *The Present and Future of Japan's Population.* Paper Prepared for the 12th Conference of the Japanese Institute of Pacific Relations, Kyoto.

Organisation for Economic Co-operation and Development (1980). *Mortality in Developing Countries*, vol **III**. *New Model Life Tables for Use in Developing Countries.* Paris.

Perks, W. (1932). *On Some Experiments in Graduation of Mortality Statistics.* J.I.A., **63**, 12-40.

Pierce, D. A. & Schafer, D. W. (1986). *Residuals in Generalized Linear Models.* J. Am. Statist. Assoc. **81**, 977-986.

Pollard, A. H. (1949). *Methods of Forecasting Mortality Using Australian Data.* J.I.A., **75**, 151-170.

Pollard, J. H. (1973). *Mathematical Models for the Growth of Human Populations.* Cambridge University Press.

Pollard, J. H. (1987). *Projection of Age-Specific Mortality Rates.* Population Bulletin of the United Nations **21,** 22, 55-69.

Pregibon, D. (1984). *Book Review: Generalized Linear Models.* The Annals of Statistics, **12**, 4, 1589-1596.

Preston, S. (1974). *An evaluation of Postwar Mortality Projections in Australia, Canada, Japan, New Zealand and the United States.* World Health Statistics Report, Special Subject II. World Health Organization, Geneva.

Renshaw, A. E. (1991). *Actuarial Graduation Practice and Generalized Linear and Non-linear Models.* J.I.A., **118**, II, 295-312.

Renshaw, A. E. (1992). *Joint Modelling for Actuarial Graduation and Duplicate Policies.* J.I.A., **119**, **I**, 69-85.

Renshaw, A. E. (1995). *Graduation and Generalized Linear Models: An Overview.* Actuarial Research Paper No. **73**, Department of Actuarial Science and Statistics, City University, London.

Renshaw, A. E. & Haberman, S. (1997). *Dual Modelling and Select Mortality.* Insurance: Mathematics and Economics **19** (**2**), 105-126.

Renshaw, A. E. & Haberman, S. (2000). *Modelling for Mortality Reduction Factors.* Actuarial Research Paper No. **127**, City University, London.

Renshaw, A. E. & Haberman, S. (2003a). *Lee-Carter Mortality Forecasting, a Parallel GLM Approach, England & Wales Mortality Projections.* J. R. Statist. Soc., C (Applied Statistics), **52**, 119-137.

Renshaw, A. E. & Haberman, S. (2003b). *On the Forecasting of Mortality Reduction Factors.* To be published in Insurance: Mathematics and Economics.

Renshaw, A. E., Haberman, S., & Hatzoupoulos, P. (1996). *The Modelling of Recent Mortality Trends in United Kingdom Male Assured Lives.* B.A.J., **2**, II, 449-477.

Renshaw, A. E. & Hatzoupoulos, P. (1996). *On the Graduation of 'Amounts'.* B.A.J., **2**, I, 185-205.

Schweizer, B., & Sklar, A. (1983). *Probabilistic Metric Spaces.* North Holland, New York.

Scott, W. F. (1996). *Life Contingencies.* Heriot-Watt University, Edinburgh.

Seal, H. (1945). *Tests of a Mortality Table Graduation.* J.I.A., **71**, 3.

Society of Actuaries (1981). *Report of the Committee to Recommend a New Mortality Basis for Individual Annuity Valuation* (*Derivation of the* 1983 *Table a*). Transactions of the Society of Actuaries, **33**, 675-735.

Sverdrup, E. (1965). *Estimates and Test Procedures in Connection with Stochastic Models for Deaths, Recoveries and Transfers between different States of Health.* Skandinavisk Aktuarietidskrift, **48**, 184.

Thatcher, A. R. (1987). *Mortality at the Highest Ages.* J.I.A., **114**, 327-338.

Thatcher, A. R. (1990). *Some Results on the Gompertz and Heligman Laws of Mortality.* J.I.A., **117**, 135-151.

Thatcher, A. R. (1999). *The Long-term Pattern of Adult Mortality and the Highest Attained Age.* J. R. Statist. Soc. A, **162,** 5-43.

Thiele, P. (1872). *On a Mathematical Formula to Express the Rate of Mortality Throughout the Whole of Life.*

Tolley, H. D., Hickman, J. C. & Lew, E. A. (1993). *Actuarial and Demographic Forecasting Methods* in Manton, K., Singer, B. & Suzman, R. eds., *Forecasting the Health of Elderly Populations*. Springer.

Tuljapurkar, S. & Boe, C. (1998). *Mortality Change and Forecasting: How Much and How Little Do We Know*? North American Actuarial Journal, **2,** 4, 13-46.

United Kingdom, Government Actuary's Department (1965). *Projecting the Population of the United Kingdom.* Economic Trends, iii. H. M. Stationery Office, London.

United Nations (1955). *Age and Sex Patterns of Mortality: Model Life Tables for Underdeveloped Countries.* United Nations Publication, Sales No. 1955.XIII.9.

United Nations (1982). *Model Life Tables for Developing Countries.* United Nations Publication, Sales No. 81.XIII.7.

United States Department of Commerce, Bureau of the Census (1977). *Projections of the Population of the United States*: 1977 to 2050. Population Estimates and Projections, Series P-25, No 704. Government Printing House, Washington, D. C.

United States Social Security Administration (1997). *Social Security Area Population Projections*: 1996. Actuarial Study No. 112. Baltimore, MD.

Venables, W. N. & Ripley, B. D. (1994, 1997). *Modern Applied Statistics with S-PLUS.* Springer-Verlag, New York.

Verrall, R. J. (1996). *A Unified Framework for Graduation.* Actuarial Research Paper No. **91,** Department of Actuarial Science and Statistics, City University, London.

Wachter, R. & Finch, C., (ed. 1997). *Biodemography of Aging.* National Academy Press, Washington, D.C.

Waters, H. R. & Wilkie, A. D. (1987). *A Short Note on the Construction of Life Tables & Multiple Decrement Tables.* J.I.A., **114**, 569-580.

Wedderburn, R. W. M. (1974). *Quasi-likelihood Functions, Generalized Linear Models and the Gauss-Newton Method.* Biometrika, **61**, 439-447.

Wetterstrand, W. H. (1981). *Parametric Models for Mortality Data: Gompertz's Law over Time.* Transactions of the Society of Actuaries, **33,** 159-176.

Whelpton, P. K., Eldridge, H. T. & Siegel, J. S. (1947). *Forecasts of the Population of the United States of 1945-75.* Bureau of the Census, United States Department of Commerce. Washington, D.C.

Wilkie, A. D. (1976). *An International Comparison of Recent Trends in Population Mortality.* Transactions of the 20th Congress of Actuaries, Tokyo, **2**, 761-781.

Wilkin, J. C. (1981). *Recent Trends in the Mortality of the Aged.* Transactions of the Society of Actuaries, **XXXIII**, 11-62.

Willets, R. (1999). *Mortality in the Next Millenium.* Staple Inn Actuarial Society, London.

Wilmoth, J. R., (1995). *Are Mortality Rates Always More Pessimistic When Disaggregated by Cause of Death*? Mathematical Population Studies **5,** 4, 293-319.

Wilmoth, J. R., (1996). *Mortality Projections for Japan: A Comparison of Four Methods* in Caselli, G. & Lopez, A. eds., *Health and Mortality among Elderly Populations*. Oxford: Oxford University Press, 266-87.