Statistical and Mathematical Modelling for Mortality Trends, and the Comparison of Mortality Experiences, through Generalised Linear Models and GLIM

Peter Hatzopoulos

A Thesis Submitted for the Degree of Doctor of Philosophy

The City University

Department of Actuarial Science and Statistics

1997

To my family

and to my teachers H. Karagiannis

V. Klonias

A. Renshaw

S. Haberman

Peter Hatzopoulos
Kritis *53*
12243 Egaleo
Athens - Greece
Tel : *5986960*

Contents

Acknowledgements

I am indebted to Professor Steven Haberman and Doctor Arthur Renshaw for their guidance and their helpful comments during the period which this work was undertaken.

I am very grateful to the Continuous Mortality Investigation Bureau for providing the data and to the *CMI* Bureau and *ABI* for financial support.

Declaration

Powers of discretion are granted to the University Librarian to allow this thesis to be copied in whole or in part without further reference to the author. This permission covers only single copies made for study purposes, subject to normal conditions of acknowledgement.

Abstract

The aim of the thesis is the statistical and mathematical modelling of trends over time in age - specific mortality rates based on lives, policies and amounts. The analysis is based on the theory of generalised linear models (*GLM's*).

Further, a method is advocated for the comparison of mortality experiences, as well as a method for the construction of a mortality table based on a standard mortality experience.

The results are based on the mortality experience of the *UK* life offices for whole life and endowment assurances, for the time period *1958 - 1990*, and for pensioners in pensions schemes, for the time period *1983 - 1990*, published by the *CMI* Bureau of the Institute and Faculty of Actuaries.

Introduction

The thesis consists of four parts.

The ***first part*** explains the need for the graduation process in the construction of life tables, and discusses the use of generalised linear modelling techniques.

Chapter I demonstrates how to construct a life table, explains the need for graduation, and reviews the history of graduation methods.

Chapter II discusses the history of the major mathematical formulae used to graduate mortality rates.

Chapter III outlines the theory of *GLM*'s.

Chapter IV describes the main statistical tests used for the justification of the model structure selected in the graduation process in relation to the theory of *GLM's*.

The ***second part*** considers various statistical distributions for modelling crude mortality rates.

Chapter V deals with the modelling of the central mortality rates

Chapter VI deals with the modelling of the initial mortality rates.

The ***third part*** deals with different approaches to the mathematical modelling of mortality rate trends.

Chapter VII considers the methodology advocated for the mathematical modelling of mortality rate trends.

In *Chapter VIII*, the Multiplicative model is applied for modelling

1. Male assured lives, duration *5+*, ages *24 - 89*, time period *1958 - 1990.*

2. Male assured lives, duration *5+*, ages *42 - 89*, time period *1958 - 1990.*

In *Chapter IX*, the Power model is applied for modelling

1. Male assured lives, duration *5+*, ages *24 - 89*, time period *1958 - 1990.*

2. Male assured lives, duration *5+*, ages *42 - 89*, time period *1958 - 1990.*

In *Chapter X*, the Additive model is applied for modelling

1. Male assured lives, duration *5+*, ages *24 - 89*, time period *1958 - 1990.*

In *Chapter XI*, the Complementary log-log model is applied for modelling

1. Pensioners, ages *60 - 95*, time period *1983 - 1990.*

In *Chapter XII*, the modelling of mortality data based on amounts for pensioners is analysed.

The ***fourth part*** describes two methods for the comparison of mortality tables.

In *Chapter XIII*, the first method deals with testing the hypothesis whether or not two mortality experiences can be modelled under the same mathematical structure using F - tests. As an illustration, durations *0, 1, 2, 3, 4* for male assured lives, for ages *23 - 62*, and time period *1958 - 1990*, are classified. The second method deals with the construction of mortality tables based on a standard mortality table with similar characteristics. This is illustrated by a number of examples. A pensioners' mortality table is constructed based on male assured lives' mortality experience, for the calendar year *1990*, and for the range of ages *64 - 89*. A mortality table for grouped durations *3 - 4* years is constructed based on male assured lives mortality experience for durations *5+*, for the time period *1958 - 1990*, and for the range of ages *23 - 62*. Further, mortality tables for durations *0, 1, 2* years are constructed based on grouped durations *3 - 4* years experience, for the time period *1958 - 1990*, and for the range of ages *23 - 62*.

Definitions

θ_x	the actual number of deaths for lives with age label x.
R_x^c	the central exposed to risk based on lives with age label x.
R_x^i	the initial exposed to risk based on lives with age label x.
P_x	the total number of policies giving rise to claims for lives with age label x.
${}^pR_x^c$	the central exposed to risk based on policy counts P_x for lives with age label x.
${}^pR_x^i$	the initial exposed to risk based on policy counts P_x for lives with age label x.
q_x	the probability that a life, attaining age label x, dies before attaining age label $x+1$.
μ_x	the force of mortality at exact age x in a life table.
$P_x(t)$	the population present at time t for lives with age label x, over the period of the mortality investigation.
l_x	the number of lives at exact age x in a life table.

Part 1

Graduation and Generalised Linear Models

CHAPTER I

Life table and graduation

1.1 Introduction

The *life table*, also referred to as the *mortality table*, is a statistical device for presenting and summarising the mortality experience of a population in a form that permits answering questions such as : What is the probability that a man aged x years will survive to age y, or what is the average number of years of life remaining for a person who has reached his x'th birthday?

Investigations connected with the construction of life tables began in the *17*th century. The Englishman John Graunt constructed in *1662* the first life table for the inhabitants of London. Later, the famous mathematician Wilhelm Leibniz presented, to the Royal Society in London, reliable statistical information for the city of Wroclaw. On the basis of this material, the English astronomer Edmund Halley constructed the first reliable life table in *1693*, using a method known subsequently as the Halley method. In *1760*, the Halley method was supplemented by the famous Swiss mathematician Leonhard Euler. Later modifications included the contributions of Per Wargentin (*1749*) and Richard Price (*1783*) and then, in *1812*, the French scientist Pierre Laplace proposed a direct method for the construction of life table from statistical data. (Gavrilov and Gavrilova, *1991,* Haberman and Sibbett, *1995*).

This initial historical stage can be described as the period of descriptive human mortality statistics rather than modelling in the modern sense. Besides their particular interest in human mortality, many scientists did not separate man and other living creatures in their investigations about mortality, which is justified by the recent tendency of integrating medico-biological and demographic research (Gavrilov and Gavrilova, *1991*).

The life table was used primarily in actuarial science for the analysis of life contingencies and for life insurance calculations such as the practical computation of premiums, and in demography to study population structure and change. Due to the work of health statisticians in medical follow-up studies in the early *1950's*, the life table began to attract the attention of biostatisticians. The advances in probability and statistical theory, and the life table's similarities with reliability theory and survival analysis have made it possible to address the life table from a purely stochastic point of view and to provide the subject with a rigorous theoretical foundation. Life table analysis has emerged as a rigorous and exact statistical method.

The life table method is applicable to the analysis of not only mortality but of many measurable censoring processes such as in the clinical studies of humans, or laboratory studies of animals. The applicability of the method can be generalised to non living things as for example to describe the life and death history of automobiles in a given year or to study the length of life of light bulbs and others. Consequently, the life table has become a valuable tool used by actuaries, biologists, physicists, demographers, manufacturers, public health workers and investigators in many other fields.

Two ways of categorising the life table are to consider the cohort (or generation) life table and the period life table. In the construction of the *cohort life table* one records the mortality experience of a group of individuals (all born at same period) from the birth of the first to the death of the last member. Besides the impracticality of long time delays in the construction of the cohort life table there are many other difficulties involved as many individuals may migrate or die unrecorded. However, cohort life tables have applications in the study of cause-specific mortality for humans, animal mortality and in the assessment of the durability of mechanical objects.

The *period life table* is the most effective means of analysing mortality and survival experience of a population. It is also a useful tool for comparing mortality experiences. The period life table is entirely dependent on the mortality rates prevailing in the time-period from which it is constructed. So, life expectancy based on a period life table means the expected number of years of life if the person were subjected throughout his life to the same mortality prevailing in the current year, which means that time is not taken into account as a factor influencing mortality. However, the construction of period life tables in successive time periods allows the factor time to affect mortality after calculating life expectancy on the basis of an 'artificial' cohort.

Cohort and current life tables may be either *complete* or *abridged*. In a complete life table the mortality rates are computed for each year of life; an abridged life table deals with grouped age intervals greater than one year.

1.2 The construction of a life table

In this section, we consider the construction of a life table from statistical data (crude rates) on deaths and lives under observation.

During the investigation period, we group the population concerned by noting age, sex and any other possible factors affecting mortality rates (such as social and cultural background, occupation, physical environment, standard of living, education and intelligence, mode of living or duration since initial selection). That is, we need a *homogeneous population* in which all individuals ~~have close~~ are exposed to a common force of mortality, in order to achieve accurate results. Of course, practical considerations require constraints to be placed on the degree of subdivision of the data so that there are adequate amounts of data available in each classified cell, in order to produce sound statistical results.

In the process of deriving mortality rates, it is not enough to count only the deaths occurring during the investigation period. We also need to know the amount of time that the lives under observation have been 'exposed' to the risk of death, so that we can estimate the *crude rate* of mortality. This quantity forms the divisor of the crude mortality rate and is known as the *exposed to risk* (Benjamin and Pollard, *1980*).

But before describing how the exposed to risk is calculated, it is necessary to define the period of time during which all the lives have the same 'label' which categorises the individuals according to those factors under consideration (for example, age). We need the definition of the label, according to which both the exposed to risk and deaths classify the individuals under observation, so that both deaths and exposed to risk correspond and are used to estimate the mortality rates correctly.

For example, we can count deaths and exposed to risk for individuals aged x last birthday, that is for individuals aged between exact x and exact $x+1$, during the investigation period. The only factor in this case is the age, and only the lives with age between x and $x+1$ are counted in the exposed to risk and the possible numbers of deaths.

This period of time is called the *rate interval* and is essential for the 'Principle of correspondence' according to which lives and deaths must be grouped under the same age label (Puzey, *1986*).

The same principles apply to the construction of crude mortality rates considering additional factors of mortality, as in the case of select rates after classifying exposed to risk and deaths according to the additional factor which is the duration since initial selection.

The type of crude mortality rate defined by the above procedure depends on the way the exposed to risk time for the lives under the investigation is calculated. There are two kinds of exposed to risk, both measured in units of years : central exposed to risk and initial exposed to risk.

1. central exposed to risk

One approach is to calculate the *exact time period of exposed to risk* for lives participating in the investigation. Using the same example as before, we consider a group of persons between ages x and $x+1$ observed during a time period. Assuming that the i^{th} person enters the investigation at age $x+a_i$ and leaves at age $x+b_i$, either by death or survival, the actual time he was under observation is $b_i - a_i$. The sum of all those individual exposures form the *central exposed to risk* symbolised by R_x^c. This computational procedure is called the *direct* method.

An alternative, more practical, approach is the *census* method. The central exposed to risk can be written as the integral of the population $P_x(t)$ present at time t for lives with age label x, over the whole period of investigation (Puzey, *1986*), that is

$$R_x^c = \int_0^T P_x(t)\, dt$$

This equation gives the total exposure time, in life years, during the investigation period $(0,T)$ of lives with age label x. Each individual contributes exposure time to the above integral only while he is alive with age label x.

For the calculation of the central exposed to risk by the census method, we can interpolate any mathematical formula which passes through the censuses, and integrate it explicitly as was indicated before. This mathematical expression will portray $P_x(t)$, the population present at time t for lives with age label x, over the period of the investigation. For instance, if there are available censuses only at the beginning and the end of the investigation and assuming that $P_x(t)$ varies linearly over the two successive censuses, the trapezium rule gives an approximation for the above integral. That is,

$$R_x^c \cong \frac{T}{2} \cdot (P_x(0)+P_x(T))$$

In either case (using the direct or the census method), dividing the number of deaths (θ_x) observed for lives with age label x by the corresponding central exposed to risk, R_x^c, the *crude central mortality rate* is obtained for every age label x in question.

Further, *assuming* that mortality is constant over the period with age label x, then the central rate over the period of age label x is identical with another mortality measure which is called the *crude force of mortality* and is symbolised by $\overset{0}{\mu}_x$. The assumption about constancy of the force of mortality over the period with the same age label x is utilised throughout the thesis. Thus, the crude central mortality rate over the period with age label x is

$$\overset{0}{\mu}_x = \theta_x / R_x^c \qquad (1.1)$$

The central mortality rate m_x is defined by the ratio

$$m_x = \frac{\int_0^1 l_{x+t} \cdot \mu_{x+t} \cdot dt}{\int_0^1 l_{x+t} \cdot dt}$$

and if $\mu_{x+t} = \lambda_x \quad \forall \quad t \in (0,1)$ then central mortality rates are identical with the force of mortality. Therefore, central mortality rates, in this thesis, are considered to be identical with the force of mortality, under the assumption of constancy of mortality during the interval for lives with age label x.

2. *initial exposed to risk*

If, in the event of death, the exposure time is continued up to the time where the individual would have normally left the investigation, and we add this extra time to the central exposed to risk, we form the initial exposed to risk R_x^i.

An approximation to the initial exposed to risk is given by the equation

$$R_x^i = R_x^c + \frac{\theta_x}{2}$$

on the assumption that the deaths are uniformly distributed over the rate interval x to $x+1$. Note that the above assumption may be inconsistent with the earlier assumption that the force of mortality μ_{x+t} is constant for $0 \leq t \leq 1$.

Dividing the number of deaths observed for lives with age label x by the initial exposed to risk the crude mortality rate is obtained for every age in question. This kind of mortality measure is called the *crude initial rate of mortality* and is symbolised by $\overset{o}{q}_x$. That is,

$$\overset{o}{q}_x = \theta_x / R_x^i \qquad \textbf{(1.2)}$$

The exact relationship between the force of mortality and the rate of mortality is obtained by

$$q_x = 1 - exp(-\int_x^{x+1} \mu_t \, dt)$$

Yet, assuming constancy for the force of mortality for each age interval $(x, x+1)$ the above formula becomes

$$q_x = 1 - exp(-\mu_{x+1/2})$$

Then, following the approach of Sverdrup (*1965*), the (maximum likelihood) estimator for the rate of mortality using the central exposed to risk is given by

$$\overset{o}{q}_x = 1 - exp(-\theta_x / R_x^c)$$

Quoting Sverdrup (*1965*), it is explained that "There is a real loss in information by disregarding the waiting time, such as in the case when $\overset{o}{q}_x = \theta_x / R_x^i$ is used in place of $\overset{o}{q}_x = 1 - exp\ (-\theta_x / R_x^c)$. When probabilities of death are small the frequencies give us the essential information needed, but as the probabilities become large the total waiting time R^c is of greater and greater importance, and when death is almost certain it is the waiting time that is pertinent". This point takes us beyond the restrictive assumption that deaths are uniformly distributed over the age interval $(x, x+1)$.

Moreover, " q_x has the weakness that it only reflects the total effect of mortality over a year, i.e. how many died by the end of a year, and is not affected by how these deaths are distributed over the year". (Puzey, *1986*)

"Central rates are very *efficient* (i.e. with little loss of information) and if the denominators are accurately computed, the main argument for their introduction was certainty of achieving *unbiasedness*" (Sverdrup, *1965*).

Furthermore, $se\{\theta_x / R_x^i\} \ / \ se\{1 - exp\ (-\theta_x / R_x^c)\} > 1,$ where se denotes the standard error (Sverdrup, *1965*) .

However, we should note that the census method used in the *CMI* Reports for computing R^c is only approximate, so that, in practice, it is often the case that exact information on R^c is *not* available.

Now, the force of mortality can be expressed as

$$\mu_x = \lim_{\delta_x \to 0^+} \frac{\Pr\ (\text{ death occurs between } x \text{ and } x + \delta_x \mid \text{survival to } x)}{\delta_x}$$

Therefore

$$\mu_x = \lim_{\delta_x \to 0^+} \frac{\delta_x q_x}{\delta_x}$$

where ${}_{\delta_x}q_x$ is the probability of death in the age interval x to $x + \delta_x$, conditional on the survival at age x.

In statistical terms μ_x , is identical to the *hazard rate function* $h(x)$. If T, the future lifetime, is considered as a random variable, for a homogeneous population of individuals for which failure is death, and each having a 'failure time' (lifetime) T, then

$$h(t) \cdot dt \cong P\,(\,t < T \leq t+dt \mid T > t\,), \quad \text{for small } dt$$

The *failure distribution* $F(t)$ is defined to be the probability of death before some time t, thus

$$F(t) = Pr\,(T < t\,)$$

The *survival function* is defined to be the probability of surviving to time t, thus

$$S(t) = Pr\,(T \geq t\,) = 1 - F(t) = exp\,[\,-H(t)\,]$$

and the *density function* or the *absolute instantaneous failure rate* $f(t)$ as

$$f(t) = h(t) \cdot S(t) = h(t) \cdot exp\,[\,-H(t)\,]$$

where

$$H(t) = \int_0^t h(x)\,dx$$

is called the *integrated hazard*. In the case of translating distributions, by introducing an additional parameter δ, everything can be converted into distributions on (δ,∞).

The construction of the life table is accomplished by computing the l_x values, which give the population present at the beginning of the interval for lives with age label x, from the following relationships

$$l_x = l_\alpha \cdot {}_{x-a}p_\alpha$$

for arbitrary l_α and computing

$$ {}_{x-\alpha}p_{\alpha} = \prod_{t=0}^{x-a-1} (1 - q_{\alpha+t}) $$

and

$$ 1 - q_{x+t} = \exp\left(-\int_0^1 \mu_{x+t+s}\, ds\right) = \exp\left(-\mu_{x+t+1/2}\right) $$

under the same assumption of constancy of the force of mortality over each year of age.

1.3 The nature of graduation

The above two mortality measures (the force of mortality and the initial rate of mortality) are subject to sampling errors giving an uneven progression from age to age. We *assume* initially that the irregularities are due only to the random variability inherent in the finite sample we observe (and we relax this assumption in a later paragraph on the following page). That is, increasing the size of the sample would lead to the irregularities being minimised and the crude rates would show an even progression through the ages. Thus, mortality rates are assumed to be a continuous and smooth function of age. *Graduation* is the practical means of compensating for the lack of availability of an infinite sample size with a practicable alternative of estimating the true mortality values as accurately as possible.

Copas and Haberman (*1983*) refer to the graduation problem and comment that, "the fundamental justification for the graduation of a set of observed probabilities like $\overset{0}{q}_x$ is the premise (*suggested by experience of nature*) that, if the number of individuals in the group on whose experience the data are based had been considerable larger, the set of observed probabilities would have displayed a much more regular progression with x".

Regarding the graduation problem, Puzey (*1986*) explains that, "the process of seeking to remove the random fluctuations is known as graduation".

Benjamin and Pollard (*1980*, page *240*), state that, "the art of smoothing the separate maximum likelihood values to obtain the best possible estimates of the underlying population values is called graduation".

In other words, graduation is the procedure of estimating the expected mortality rates, under the principle (*axiom*) that the resulting mortality values should show a smooth trend, or, that each set of neighbouring graduated values should satisfy the mathematical criteria of smoothness, differentiability and continuity.

Graduation should only remove random fluctuations. Crude rates can also include irregularities which are not due to sampling errors. In this case the true mortality rate in a particular range of ages inherits a specific feature which is not very smooth, and which has been called *intrinsic roughness* (Benjamin and Pollard, *1980*). A characteristic example of this phenomenon is the accident 'hump' occurring around age *18* among male lives in certain western European countries. Intrinsic roughness can be verified by previous experience, or by attaching a specific distribution to those crude rates and analysing the residual variations arising from a graduation, as will become clear in the following Chapters.

Removing only random sampling errors is different from the process in which the graduated rates have an unreasonably excellent goodness of fit to the crude rates (*undergraduation)*, and from the removal of intrinsic roughness or any other particular trend that the crude rates might include (*overgraduation)*. The graduation process intrinsically involves a trade off between smoothness and adherence to the crude rates. It could be stated, that the weights of this trade off depends on the fidelity to data and on tastes. More specifically, if the crude rates have been derived from a large population, like the English Life Tables (*ELT*) mortality investigations, adherence to the crude rates (undergraduation) is desirable.

We conclude that graduation is not achieved only by following an algorithm strictly, but is based as well on personal judgement and experience, and visual inspection should be an important part of the criteria for the acceptance of any particular graduation.

1.4 Methods of graduation

There are a number of methods for carrying out a graduation. These include the graphic method, graduation using splines, graduation by mathematical formula, non - parametric methods.

I. In the *graphic* method, a hand - drawn, curve is fitted to pass inside the corridor formed by the *95%* confidence intervals based on the crude mortality rates. This is a useful method for scanty data, where personal judgement is important, but there is the risk of bias being introduced. The graphic method in now mainly of historical interest.

II. Graduation by *mathematical formula* is the method when a mathematical model structure is applied to describe the mortality experience in question with the parameters involved being estimated by some optimisation criterion. Optimisation can be achieved

a) by *(weighted) least squares* method minimising the quantity

$$Q = \sum_x w_x \cdot (\overset{\circ}{z}_x - \hat{z}_x(\underline{\beta}))^2$$

in respect of $\underline{\beta}$, where $\underline{z} = \underline{q}$ or $\underline{\mu}$, and w_x are prior weights, or

b) by maximising the *(log) - likelihood* of the observed events, which is the sum of the (log) - likelihoods for each observation (under the independence assumption), after attaching an appropriate distribution to the observed rates, or

c) by minimising the X^2 value.

A special case of graduation by mathematical formula is the *reference to a standard table* method which is introduced where the mortality experience under analysis is believed to be related to a particular standard table. The method can be helpful again when the data are scanty. Various connections between the graduated rates and the mortality rates from the standard table have been suggested such as

$$q_x = a \cdot q_x^s + b \quad \text{or} \quad q_x = q_x^s \cdot (a + b \cdot x) \quad \text{or} \quad q_x = q_{x+h}^s + k \qquad \textit{(1.3)}$$

where q_x^s are the graduated values from the standard table, or models in terms of μ_x and μ_x^s.

In Benjamin and Pollard (*1980*) and Chadburn (*1991*), the reader can find a thorough analysis of the previous methods and a detailed consideration of the advantages or disadvantages of each method.

III. Graduation using *splines* has been used recently (*ELT*, No *14*) for mortality experiences which include different generations, where mathematical formulae commonly fail to produce an adequate fit for the whole range of ages under graduation. Spline functions can be considered to be an intermediate method between parametric and non - parametric methods.

According to their degree (*d*) they are defined by

$$f(x) = \sum_{j=0}^{d} \alpha_j \cdot x^j + \sum_{j=1}^{n} \beta_j \cdot (x - k_j)_+^d$$

where $(x-k)_+^d = (x-k)^d$ if $x > k$ and 0 otherwise, n the number of knots, and k the positions of the *knots* over the age range. Optimised estimation of the parameters (α_j, β_j, k_j) can be achieved as above.

IV. *Non - parametric methods* of graduation have long been developed, including Whittaker-Henderson graduation and moving weighted average methods (commonly used in the *U.S.A.*) and Kernel methods: further details are provided by London (*1985*), Copas and Haberman (*1983*) and Gavin et al (*1993*, *1994*). Verrall (*1992*) has shown how these methods can be put in a dynamic generalised linear modelling framework, with the estimated parameters being changed for each age in a 'time series' manner.

All the above approaches have their advantages and limitations. For example, non - parametric methods could be useful for graduations with a large range of ages such as English Life Tables, while the graphic method or reference to a standard table method could be very useful when a small sample participates in a mortality investigation.

The most relevant methods, for the kind of data being analysed in this thesis, seem to be the method of graduation by mathematical formula or the method of graduation using splines. This preference lies in the insight they provide when comparing different mortality experiences or when analysing any trends or even when attempting a forecast of future mortality rates. Moreover, the assured smoothness we automatically obtain using mathematical formulae, the rich gamut of them and the modern statistical computing packages all make these methods even more attractive.

The theory of Generalised Linear Models (*GLMs*) is used throughout this thesis and its connection with graduation is explained in Chapter *III*.

CHAPTER II

History of major Mathematical formulae fitted to Mortality data

2.1 Introduction

The first attempts to explain the *quantitative laws of life span* begin in the *19*th century, after the accumulation of reliable statistical data and the development of more sophisticated analytical methods.

But, before the description of the major mathematical laws it would be desirable to ask *what* we aim to achieve by modelling life span mathematically, and what conditions a mathematical model must satisfy.

According to Gavrilov and Gavrilova (*1991*), "the recognition that what we basically aim to achieve is nevertheless a *clarification* of the mechanism which determine the life span of organisms. Starting from this point, mathematical modelling is not a goal in itself, but only one of the means of achieving the intended goal. Therefore, we should pay particular attention not to cumbersome mathematical constructions which claim to be fundamental theories, but rather to comparatively simple heuristic working models which correspond to the known facts and predict new regularities".

"The best guarantee of success in applying the technique of mathematical modelling to biosystems is a dynamic change of models. A mathematical model should be investigated to see how its capabilities match the aims for which it was created, and once a model has been derived, it should be subjected to criticism and never made into a dogma for any length of time".

Thus, the construction of mathematical models is a method to describe the observed mortality experience. These mathematical models involve parameters and it would be desirable to use simple models that allow for change since time is an important factor for analysing mortality trends.

The second question, which arises naturally, is *how* we achieve the mathematical modelling of life span, bearing in mind the above remarks.

Gavrilov and Gavrilova (*1991*) describes the following general '*Methodological principles for selecting the life span distribution law'.*

1. The principle of theoretical justification

According to this principle, only those equations should be used which have theoretical justifications. Then the recording of data using such an equation is simultaneously the first step to its interpretation. Starting from this principle, special attention should be devoted to formulae derived from theoretical hypotheses rather than to the empirical formulae.

2. The principle of universality

The aim of revealing general regularities which are valid for the widest possible range of natural phenomena is the very essence of the scientific world - view. In conformity with this principle, special value should be attached to general life span distribution laws which are valid for the greatest variety of organisms, including man.

3. The principle of the best approximation with the fewest parameters

A formula satisfying this principle allows data to be recorded in the most compact form, thereby permitting the distribution to be recovered with the minimal number of observations. This principle is a particular case of the idea that "entities are not to be multiplied beyond necessity", known as Occam's razor. As applied to the problem of life span, this principle points us not towards an absolutely exact description of the observed lifetime distributions using formulae with many parameters, but towards the use of models which reflect the most prominent characteristics of those distributions. In this connection, a promising approach might involve a factor analysis of mortality patterns, permitting a determination of the minimum number of parameters necessary to describe the salient features. Keyfitz (*1982*) provides a full review of the different approaches to the principle of a 'minimum parameter representation'.

4. The principle of local description

Since many systems have *critical periods* in their development when they qualitatively change their properties and behaviour, we should not try to describe the whole extent of the process in one go. The history of science demonstrates that the local description of a process is the most efficient path to take, with the subsequent 'dovetailing' of the various scientific approaches in the framework of a new, more general conception. Therefore, if a proposed life span distribution law is valid only for a restricted age interval, this is not in itself a basis for being critical towards it. The restricted applicability of the law does not demonstrate that it is incorrect, but merely that it is only a special case of another, more general and as yet unknown law.

Therefore, according to the third principle, "*simplicity,* represented by *parsimony* of parameters, is a desirable feature of a model. We do not include parameters that we do not need. Not only does a parsimonious model enable the analyst to think about his data, but one that is substantially correct gives better *predictions* than one that includes unnecessary extra parameters" (McCullagh and Nelder, *1983*, page *6*).

Moreover, if a model fits very closely to a particular set of data, it will not include changes or any measure for comparison that might be useful when another set of data relating to the same phenomenon is collected. Parsimony is related to *parameter invariance,* that is to parameter values that either do not change as some external condition changes or change in a predictable way (McCullagh and Nelder, *1983*).

Finally, quoting from McCullagh and Nelder (*1983*, page *6*), on the question of what constitutes a good model, we have that "Modelling in science remains, partly at least, an art. Some principles do exist, however, to guide the modeller. The first is that *all models are wrong*; some, though are better than others and we can search for the better ones. At the same time we must recognise that eternal truth is not within our grasp. The second principle (which applies also to artists!) is not to fall in love with one model, to the exclusion of alternatives. Data will often point with almost equal emphasis at several possible models and it is important that the analyst accepts this. A third principle involves checking thoroughly the fit of a model to the data, for example by using residuals and other quantities derived from the fit to look for outlying observations, and so on. Such procedures are not yet formalised (and perhaps never will be), so that imagination is required of the analyst here as well as in the original choice of models to fit".

2.2 History of major mathematical formulae

This section deals with the major mathematical formulae that have been used in the Actuarial literature. Fuller information can be obtained from the reviews written in Benjamin and Pollard (*1980*) and elsewhere.

I. De Moivre (*1725*)

A first attempt to describe a life table by a mathematical law was given by Abraham de Moivre (*1725*) in his hypothesis of equal decrements. In his book *Annuities upon Lives* he provided a thorough discussion of the valuation of annuities, but the underlying mortality hypothesis was defective as a representation of human mortality. His basic formula relates to l_x and is

$$l_x = k \cdot (\omega - x)$$

where ω is the 'maximum' age.

II. Gompertz's law (*1825*)

A major improvement in the mathematical analysis of law for life span dates from *1825*, when the English actuary Benjamin Gompertz gave a theoretical foundation that the force of mortality increases with age according to the geometric progression law, and he argued on physiological grounds that the intensity of mortality (in his terms, the average exhaustion of man's power to avoid or "resist" death) gained equal proportions in equal intervals of age. His law became the keystone of the biology of life span. Gompertz suggested that the rate at which the '*resistivity to death'* decreases is proportional to the resistivity itself. Since the force of mortality acts as a measure of the human susceptibility to death, Gompertz took its *reciprocal as a measure of resistivity*, thereby deriving the equation

$$\frac{d}{dx}\left(\frac{1}{\mu_x}\right) = -\alpha \cdot \left(\frac{1}{\mu_x}\right) \qquad \textbf{(2.1)}$$

where α is a non - negative constant.

After integration the equation (*2.1*) turns into

$$\mu_x = A \cdot \exp(\alpha \cdot x) \qquad \text{or} \qquad \mu_x = \mathrm{A} \cdot B^x$$

We notice that the log transformation of the above formula produces linearity, i.e.

$$\log(\mu_x) = \log(A) + \log(B) \cdot x = \beta + \alpha \cdot x$$

Thus, the graphical presentation of the force of mortality on the log scale should be linear over the ages. So, Gompertz's law is the same formula associated with the log link, used in the theory of *GLM*'s for the force of mortality, with the line (polynomial of first degree) as the linear predictor.

Formula *(2.1)* can be viewed as a linear differential equation with constant coefficients. More specifically it can be obtained from the following

$$f'(x) + \alpha \cdot f(x) = 0$$

where $f(x)$ denotes the resistivity to death. The solution to the above differential equation is Gompertz's law.

Gompertz's law is based on a theoretical justification with a parsimonious number of parameters and provides a description for the life span beyond about age *30*, where mortality is a monotonically increasing function of age. Another restriction imposed by this law is that on the log scale the force of mortality should be linear over the ages.

III. Gompertz - Makeham's Law (*1860*)

Gompertz noted that alongside this law of mortality there must exist an element of mortality which does not depend on age. He explains that "it is possible that death may be the consequence of two generally coexisting causes: the one chance, without previous disposition to death or deterioration, or increased inability to withstand destruction" (Gompertz, *1825*).

Gompertz's observation was taken into account in *1860* by the English actuary William Makeham, who stated the force of mortality as the sum of a constant (the Makeham term) and an exponential (the Gompertz function). The mathematical expression thus takes the following form

$$\mu_x = A + B \cdot C^x$$

Both laws gave satisfactory results for the late *19*th and early *20*th century.

Further modifications made to the Gompertz - Makeham law have included for example, the addition of a polynomial term to the Gompertz function, or of a component linearly dependent on age to the Makeham term. Another way to modify the Gompertz function is to divide the Gompertz function by a term ($1 + D \cdot C^x$), giving rise to a logistic formula, which was first suggested by Perks (see Perks' formulae, No *VI*).

IV. Oppermann (*1870*)

Oppermann suggested a formula in terms of the force of mortality suitable for infancy and childhood

$$\mu_x = \frac{a}{\sqrt{x}} + b + c \cdot \sqrt{x}$$

It has been shown (Hartmann, *1980*) that Oppermann's formula is an extremely flexible means for graduation of the first twenty years of life in any of the four regional families of the model life tables of Coale and Demeny (*1966*). However, it does not give a satisfactory graduation to the data for the middle and older ages.

V. Thiele and Steffensen (*1872*)

Modifications of Oppermann's formula were made by Thiele and Steffensen in their attempts towards finding graduation formulae valid for all ages. Thiele (*1872*) was of the opinion that such formulae should take into account the differences in mortality behaviour during the major epochs of life; childhood, adult and old ages.

Thus, he wanted to partition the force of mortality (and hence the survivorship curve) into three components

$$\mu_x = \mu_1(x) + \mu_2(x) + \mu_3(x)$$

where

$\mu_1(x) = a_1 \cdot \exp(-b_1 \cdot x)$ for childhood,

$\mu_2(x) = a_2 \cdot \exp(-b_2^2 \cdot (x-c)^2)$ for adult ages and

$\mu_3(x) = a_3 \cdot \exp(b_3 \cdot x)$ for old ages.

Thiele proposed this formula for the graduation of mortality throughout all ages and it was used for the graduation of Scandinavian mortality. It was widely acknowledged that the formula due to Thiele was too complicated for general use and his efforts became of historical importance only. This is discussed further in Steffensen (*1934*).

VI. Perks' formulae (*1932*)

$$\mu_x = \frac{A + B \cdot C^x}{1 + D \cdot C^x} \quad \& \quad \mu_x = \frac{A + B \cdot C^x}{K \cdot C^{-x} + 1 + D \cdot C^x}$$

The above formulae are the principal Perks' formulae, and constituted a successful attempt to fit a single curve to the whole range of ages.

"Perks found an analogy between the inability to withstand destruction' of Gompertz and the current physical concept of entropy change; the measure of the time progression of a statistical group from organisation to disorganisation." (Benjamin and Pollard *1980*, page *22*).

VII. Beard (*1951*)

Beard (*1951*) proposed a simplified version of the Perks' formula, i.e. with $A = 0$;

$$\mu_x = \frac{B \cdot C^x}{1 + D \cdot C^x}$$

VIII. Weibull distribution (*1951*)

Following Gavrilov and Gavrilova (*1991*), the Weibull distribution is widely used and is well known in reliability theory. It describes the variability of technical systems with respect to their 'lifetimes'. It was proposed by Weibull in *1951*, and is different in principle from the Gompertz distribution since the rate of failure (the analogue of the force of mortality) is described as a power function of age

$$\mu_x = B \cdot x^c \tag{2.2}$$

Recently, the Weibull distribution has also been applied in the description of the lifetime variability of organisms (Gavrilov and Gavrilova, *1991*).

By analogy with Gompertz's law, formula *(2.2)* can be viewed as a linear differential equation with variable coefficients. More specifically it can be obtained from the following form

$$f'(x) + \frac{c}{x} \cdot f(x) = 0$$

where $f(x)$ denotes the resistivity to death.

The Weibull distribution is valid for a wide (possible) range of natural phenomena. The restriction imposed by this law is that on the log scale the force of mortality must be in the following form

$$\log(B) + c \cdot \log(x)$$

Gavrilov and Gavrilova (*1991*) used a generalised form of the Weibull law

$$\mu_x = A + B \cdot x^c$$

and another law which then called the *generalised binomial law*

$$\mu_x = A + (b + c \cdot x)^n \tag{2.3}$$

IX. *ELT 11* (*1950 - 1952*) & *ELT 12* (*1960 - 1962*)

ELT 11 was based on the deaths in England and Wales in *1950 - 52* and the population census of *1951* and a mathematical formula was used for carrying out the graduation. This approach broke away from the traditional approach of dealing with population and deaths separately.

The mathematical formula advocated was a combination of a logistic curve with a symmetrical normal curve, involving seven parameters (Benjamin and Pollard, *1980*). The following expression shows the mathematical structure used for both sexes

$$m_x = a + \frac{b}{1 + e^{-\alpha \cdot (x - x_1)}} + c \cdot e^{-\beta \cdot (x - x_2)^2}$$

where m_x is the central death rate. The parameters were estimated by 'trial and error'. In the case of *ELT 12,* the formula was only applicable from the age of *27* upwards, so that the rates for the youngest ages still needed to be graduated by other methods.

X. Male assured lives mortality *(1949 / 1952)*

"In *1955*, the *CMI* Committee produced a new standard table of mortality based on the pooled experience of the contributing life offices for the years *1949 - 1952*. A two - year period of selection was adopted". For practical reasons, the Committee considered that the "construction of a smooth series of rates was more important than the achievement of a very good fit to the observed data. So this was not to be a graduation in the traditional sense. The Committee decided that the key features were to be (*i*) an almost flat level of q_x at young ages, (*ii*) a sharp upward turn between ages *40* and *55*, (*iii*) a flattening off in the curve at the oldest ages" (Benjamin and Pollard *1980*, page *306*). The mathematical formula used (due to Beard) is related to the Perks family of curves

$$q_x = \mathrm{A} + \frac{B \cdot C^x}{E \cdot C^{-2 \cdot x} + 1 + D \cdot C^x}$$

The parameters were estimated by 'trial and error' after many numerical experiments. This formula made no attempt to reproduce mortality rates decreasing with increasing age at the youngest ages (around the range of ages *22 - 30*), an effect that reflected the distribution of deaths from accidents (Benjamin and Pollard, *1980*).

XI. Male assured lives *(1967 / 1970)*

The data relate to male assured lives under whole life and endowment assurances issued in the United Kingdom and were collected by the *CMI* Bureau. The investigation was carried out in select form, the period of selection studied being five years. Computers were used for the first time in the graduation of such data-this allowed many separate graduations to be carried out, and tested, and the final graduations were made using the formula

$$\frac{q_x}{p_x} = A - H \cdot x + B \cdot C^x \qquad \text{(Barnett formula)}$$

with the parameters being estimated by maximum likelihood methods. This formula allowed mortality rates to decrease with increasing age at the youngest adult ages and produced a satisfactory graduation (Benjamin and Pollard, *1980*). The graduation was cut off below age *17* because the above formula gave inappropriate values for ages below *17*. Also, due to the errors in the exposed to risk for ages above *89*, which led to the exposed to risk being overstated, the data were ignored at these ages.

XII. Pensioners and annuitants *(1967 / 1970)*

Experiments showed that satisfactory results for the corresponding pensioners and annuitants experience could be obtained using the formula

$$q_x = \frac{e^{F(x)}}{1 + e^{F(x)}}$$

where $F(x)$ is a polynomial of x. This formula was used, at the suggestion of A. D. Wilkie, for all the graduations of the pensioners' and annuitants' experiences in the Second Report of the *CMI* Committee (*1976*). Two parameter polynomials gave satisfactory results for all graduations except female annuitants (*ult*) where a four parameter formula was more satisfactory.

In terms of a *GLM*, the above formula is identical with the log - odds or *logit link* function, when using a Binomial error.

XIII. Heligman and Pollard (*1980*)

The best known of all formulae which describe mortality over the entire age interval is the formula proposed by Heligman and Pollard

$$\frac{q_x}{p_x} = A^{(x+B)^c} + D \cdot e^{-E(\ln x - \ln F)^2} + G \cdot H^x$$

The curve reproduces three distinct features; "the mortality of a child adapting to its new environment, the mortality associated with the ageing of the body and the superimposed accident mortality; and the 'law' is applicable throughout the life span of more than a hundred ages" (Benjamin and Pollard *1980*, page *309*). Each parameter by Benjamin and Pollard (*1980*) is described as follows: "A is almost the same as q_1, c measures the rate of decline of mortality in early life (the rate at which a child adapts to his environment), B reflects the difference between q_0 and q_1, G indicates the level of senescent mortality, while H measures the rate of increase of that mortality, D represents the intensity of the accident hump, while F indicates the location of the hump and E its spread". Thus, in Heligman and Pollard's law each parameter has a significant explanatory contribution. Heligman and Pollard showed that the above formula graduates Australian mortality accurately (Heligman and Pollard, *1980*).

XIV. English Life Table (*14*) (*1980 - 1982*)

The data for English Life Table No *14* was graduated by J. J. McCutcheon, using a *cubic spline*, $s(x)$, defined on the interval [*2, 99*], with 'knots' at the points $x_1, x_2, \ldots, x_n$, a function which is piecewise - cubic on each of the subintervals $[2, x_1], [x_1, x_2], \ldots, [x_n, 99]$. $s(x)$ is twice - differentiable at each of the knots.

McCutcheon used $n = 10$ knots for males and $n = 11$ knots for females. His formula can be described as

$$\mu_x = \alpha + \beta \cdot x + \gamma \cdot x^2 + \delta \cdot x^3 + \sum_{i=1}^{n} \varepsilon_i \cdot (x - x_i)_+^3$$

where

$$(x-x_i)_+^3 = \begin{cases} 0 & if \quad x \le x_i \\ (x-x_i)^3 & if \quad x > x_i \end{cases}$$

For ages higher than *99* an extrapolation was carried out by a cubic polynomial, using the spline values at ages *90, 91* and *92* and the somewhat arbitrary value of *0.75* at age *105*.

He explains that "the method of cubic splines is in essence a refinement of the method of oscillatory interpolation devised by George King earlier this century", "in which (method) only one derivative exists at the knots" (Office of Population Censuses and Surveys, English Life Tables, No *14*).

XV. *UK* life - offices mortality experience *(1979 - 1982)*

A comprehensive description of the graduation of these data using so-called Gompertz - Makeham formula of the type

$$\mu_x \text{ or } \frac{q_x}{1-q_x} = GM_x(r,s) = \sum_{i=0}^{r-1} a_i \cdot x^i + exp\left(\sum_{j=0}^{s-1} b_j \cdot x^j\right)$$

in which the parameters are estimated by maximum likelihood methods, is given by Forfar et al (*1988*). Renshaw (*1991b*) has noted that their methodology can be reformulated and extended through the use of generalised linear and non - linear models. This methodology is extended to model trends in mortality in this thesis.

CHAPTER III

Generalised Linear Models (GLMs)

3.1 Introduction to GLMs

We introduce *GLM*s with two important quotations from McCullagh and Nelder (*1983* & *1989*).

"Classical linear models and least squares began with the work of Gauss and Legendre who applied the method to astronomical data" (McCullagh and Nelder, *1989*, page *1*).

"Generalised Linear Models permit us to study patterns of systematic variation in much the same way as ordinary linear models are used to study the joint effects of treatments and covariates" (McCullagh and Nelder, *1983*, page *6*).

In the theory of *GLM*s the data take the following form

$$(y_1, \underline{x}_1), (y_2, \underline{x}_2), \ldots, (y_n, \underline{x}_n) \qquad \text{for } n \text{ observations}$$

where $\{y_i\}$ is a vector of responses or dependent variables treated as a realisation of a vector of independent random variables $\{Y_i\}$. The vectors $\underline{x}_i = (x_{i1}, x_{i2}, \ldots, x_{ik}) \in R^k$ having a specific structure, $\forall\ i = 1, 2, \ldots, n$ are a set of qualitative covariates (*factors)*, or quantitative covariates (*explanatory variables*). We are interested in finding the underlying relationship between y_i and the $\underline{x}_i$ or in predicting y_i from the $\underline{x}_i$.

The modelling or *error* distribution imparted to the independent random variables Y_i's is specified by the first two moments

$$m_i = E(Y_i) \qquad \& \qquad Var(Y_i) = \frac{\varphi \cdot V(m_i)}{\omega_i} \tag{3.1}$$

where $\varphi > 0$ is the *scale parameter,* ω_i the *prior weights* and V( ) the *variance function.*

Another approach would be to consider an error distribution selected from the *exponential family* of distributions. The exponential family comprises a wide range of well - known and useful distributions such as the binomial, Poisson, Weibull, normal, inverse Gaussian, and gamma distributions. But, despite this wide range of error distributions, the first approach, given by the equations *(3.1)*, grants more freedom for the error distribution and this approach is advocated in this thesis.

Quoting McCullagh and Nelder (*1989*, page *23*), "With the introduction of *GLM*s, scaling problems are greatly reduced. Normality and constancy of variance are no longer required, although the way in which the variance depends on the mean must be known".

The covariate structure is specified through a *linear predictor* of the following form

$$\eta_i = \sum_{j=1}^{k} x_{ij} \cdot \beta_j$$

with known covariate structure (x_{ij}) and unknown parameters (β_j). This is linked to the mean response $m_i = E(Y_i)$ by the equation

$$g(m_i) = \eta_i \qquad \text{with inverse} \qquad g^{-1}(\eta_i) = m_i$$

Necessary restrictions imposed on the link function g are the existence of its inverse and its first derivative.

Thus, the response random variable of a *GLM* is considered to be decomposed into two parts : a systematic component linking the linear covariate structure to the mean, and a random component specified by the error distribution.

3.2 Model fitting

The unknown parameters β_j are estimated by *maximising the quasi log - likelihood* defined by the expression

$$Q(\underline{y};\underline{m}) = \sum_{i=1}^{n} Q(y_i, m_i) = \sum_{i=1}^{n} \omega_i \int_{y_i}^{m_i} \frac{y_i - u}{\varphi \cdot V(u)} \cdot du \qquad (3.2)$$

The β_j enter this expression through the upper limit m_i and the predictor - link expression

$$g(m_i) = \sum_{j=1}^{k} x_{ij} \cdot \beta_j$$

leading to the optimisation equations

$$\sum_{i=1}^{n} \omega_i \cdot \frac{y_i - m_i}{V(m_i)} \cdot \frac{\partial m_i}{\partial \beta_j} = 0 \qquad \forall j$$

These are called the *quasi - likelihood* estimating equations and may be solved by a numerical iterative weighting method (Newton - Raphson). The statistical package *GLIM* (Generalised Linear Interactive Modelling, Francis et al, *1993*) was specially written for fitting generalised linear models, and is used throughout to implement the graduations in this thesis.

The formula *(3.2)* behaves asymptotically like a log - likelihood, since it satisfies certain properties found in asymptotic theory connected with the log-likelihood. It also reduces to the log - likelihood for the specific distributions which are members of the exponential family of distributions.

If in the structure of the linear predictor certain additive terms are known in advance, then the sum of their contributions to the linear predictor is called an *offset*, so that

$$\eta_i = offset + \sum_j x_{ij} \cdot \beta_j$$

In fitting such a model, the offset term is first subtracted from the linear predictor, and the result can be regressed on the remaining covariates.

3.3 Goodness of fit and deviance

The *unscaled quasi - deviance* is defined by

$$D(c,f) = -2 \cdot \varphi \cdot Q(\underline{y};\underline{\hat{m}}) = \sum_{i=1}^{n} d_i = 2 \cdot \sum_{i=1}^{n} \omega_i \int_{\hat{m}_i}^{y_i} \frac{y_i - u}{V(u)} du \qquad (3.3)$$

where $\underline{\hat{m}}$ denote the *fitted values* for the *current* model c.

Then the scaled quasi - deviance is defined by

$$S(c,f) = D(c,f) / \varphi$$

For members of the exponential family of distributions this is identical to $-2 \cdot \log$ (likelihood ratio), that is

$$S(c,f) = -2 \cdot \log (l_c \;/\; l_f)$$

where l_c and l_f denote the values of the likelihood under the *current model* c, with fitted values $\hat{m}_i$, and under the *saturated model* f, with fitted values y_i, respectively.

The scale parameter φ (for the current model c) may be estimated by

$$\hat{\varphi} = \frac{D(c,f)}{p}$$

where p denotes the degrees of freedom of the current model c. It is defined to be $p = n - k$, where n is the number of observations and k is the dimension of the linear vector space generated by the linear predictor structure.

The scaled quasi - deviance or deviance of the current model is a measure of discrepancy between the responses and the fitted values. Comparison of different choices of nested predictor structures, that is *GLM*s with a fixed modelling distribution, fixed link and different predictor structures which are subsets of one another, can be based on the difference between the model deviances. In particular when the predictor structure of the model c_1 is nested in the predictor structure of the model c_2 the difference in the (scaled) deviances

$$S(c_1,f) - S(c_2,f)$$

is approximately distributed as the chi - square distribution with $p - q$ degrees of freedom, where p and q denote the degrees of freedom of c_1 and c_2 respectively. The addition of a greater number of nested parameters, for a fixed error and fixed link function, reduces the deviance and induces a pay - off situation.

Now, because the deviance is twice the difference between the maximum quasi - likelihood achieved by the *full model* and that achieved by the *current model*, the above statistic is the same as twice the difference between the maximum quasi - likelihoods achieved by the two nested current models, i.e. the following statistic

$$2 \cdot \log\{Q(\underline{y},\underline{m}_1,p) \,/\, Q(\underline{y},\underline{m}_2,q)\}$$

where $Q(\underline{y},\underline{m}_1,p)$ & $Q(\underline{y},\underline{m}_2,q)$ is the quasi - likelihood for the model with p & q degrees of freedom respectively.

Within this context the *Akaike Criterion* of best fit (Forfar et al, *1988*, page *49*) is given by

$$AC = Log\{\, Q(\underline{y},\underline{m}_1,p)\,\} - 2 \cdot (p - q) \qquad \textbf{(3.4)}$$

As an alternative to using the chi - square distribution as a rough means of assessing the relative merits of nested predictor structures, it is possible to use the approximate F - statistic

$$\frac{q \cdot \{S(c_1,f) - S(c_2,f)\}}{(p-q) \cdot S(c_2,f)} \cong F_{p-q,q}$$

The result is known to be exact under the normal distribution but not otherwise.

The implementation of the theory of *GLM*'s in the graduation of mortality rates (for the construction of a current life table), is achieved by modelling age as the explanatory variable and the crude mortality rates as realisations of the independent response variables.

3.4 Residuals

In classical linear regression theory where the independent responses $\{Y_i\}$ have the normal distribution $N(m_i, \sigma^2)$, the *standardised residuals,* defined by

$$\frac{y_i - \hat{m}_i}{\sqrt{\hat{\sigma}^2}}$$

where $\hat{m}_i$ denote the fitted values and $\hat{\sigma}^2$ estimates σ^2 are approximately distributed as $N(0,1)$.

The rationale behind the desirable normality property of such residuals arises in part from the simple visual judgement that can be made, as to the goodness of fit of the modelling distribution.

It is necessary to extend this definition in the case of *GLM*s. Within the context of a *GLM* there are two types of residuals of interest.

1) The *Pearson residuals* defined by

$$\frac{y_i - \hat{m}_i}{\sqrt{\frac{V(\hat{m}_i)}{\omega_i}}}$$

where V denotes the variance function and ω_i denote the prior weights.

2) The *deviance residuals* defined by

$$r_i^D = sign(y_i - \hat{m}_i) \cdot \sqrt{d_i}$$

where d_i are the (unscaled) deviance components of equation *(3.3)*.

Both types of residuals may be *standardised* by dividing by $\sqrt{\hat{\varphi} \cdot (1 - h_i)}$ where $\hat{\varphi}$ is the estimate of the scale parameter and h_i a minor technical adjustment described in Francis et al (*1993,* pages *283 - 285*).

Residuals can detect the inadequacy of fit of a model in terms of inadequacies in the error distribution (as represented by the choice of the variance function) or inadequacies in the mathematical model (as represented by the link function and form of the linear predictor) (McCullagh and Nelder, *1989,* pages *391 - 400*).

CHAPTER IV

Statistical tests of a graduation

4.1 Introduction

Apart from assessing the goodness of fit, there are certain features which it is necessary to check in any graduation. According to Benjamin and Pollard (*1980*, page *226*) this include the checking of deviations for the possible existence of

a. a number of excessively large deviations (counter - balanced by a number of small deviations)

b. a large cumulative deviation over part (or the whole) of the age range

c. an excess of positive (or negative) deviations over the whole of the age range

d. an excessive clumping of deviations of the same sign over the whole of the age range.

Several *statistical criteria* have been devised to explore the adequacy of any proposed graduation model. We are mainly concerned with the *Chi - square* test, the *Individual Standardised Deviations* test, the *Sign* test and the *Runs* test. Each of the above tests examines certain desirable features of a graduation, and the failure of any of the tests may result in the reconsideration of the fitted model.

The statistics used for these tests in this thesis are the deviance

$$dev_x = y_x - \hat{m}_x$$

where y_x & $\hat{m}_x$ denote the observed responses and fitted values respectively of the *GLM* and

$$z_x^D = \frac{sign(dev_x) \cdot \sqrt{d_x}}{\sqrt{\hat{\varphi} \cdot (1 - h_x)}} \qquad \textbf{(4.1)}$$

the standardised deviance residuals of Section *3.4*.

Note that the corresponding studentised Pearson residuals of Section *3.4* are used in *CMI* graduations (Forfar et al, *1988* and Benjamin and Pollard, *1980*). Moreover, all of the graphical diagnostics in this thesis are based on deviance residuals, as we will see later in the next Sections.

4.2 The chi - square test

The *chi - square test* assesses the overall goodness of fit of a graduation. It involves the statistic X^2 defined as the sum of the squared residuals :

$$X^2 = \sum_{x=1}^{n} (z_x^D)^2$$

where z_x^D are the standardised deviance residuals of the associated *GLM*, defined by equation *(4.1)*, and which approximately follow the standard normal distribution.

The p - value of the test is the appropriate tail area, calculated using the chi-square distribution with $n - k$ degrees of freedom, based on n age cells (constructed by grouping adjacent ages where necessary) and a linear predictor involving k independent parameters.

If smoothness has been assured then we have an upper one - tailed test otherwise we have a two - tailed test allowing for the undesirable feature of undergraduation. Thus, one concludes that if the graduation has ~~being~~ been carried out by the use of mathematical formula, then the chi - square test becomes one - tailed. Thus the p - values quoted in this thesis for any test of any model structure are defined by

$$1 - F_{n-k}\left(\sum_{x=1}^{n} (z_x^D)^2\right)$$

where F_{n-k} is the cumulative distribution function for the chi - square distribution with $n - k$ degrees of freedom.

4.3 Other tests

As a standard practice with testing graduations, we have also used

I. The *individual standardised deviations* test which is designed to safeguard against the features described under *4.1a*. The test is based on an *upper one - tailed* p - value.

II. The *sign* test which is designed to safeguard against features described under *4.1c*. The test is based on an *two - sided* p - value.

III. The *runs* test which is designed to safeguard against features described under *4.1d*. The test is based on an *upper one - tailed* p - value.

IV. The *cumulative deviations* test , which is designed to safeguard against features described under *4.1b*. For reasons of simplicity we do not use this test since the results are usually satisfactory.

Full details are given in Benjamin and Pollard (*1980*, Chapter *11*).

4.4 Visual checks

As an additional check, the theory of *GLM*'s provides *visual tests* of the statistical analysis through *residual plots*. Standardised deviance residuals, as they are defined by equation *(4.1)*, are recommended in the text book by McCullagh and Nelder (*1989*, page *398)*, plotted either against the linear predictor, or against the fitted values transformed to the *constant information scale* (*CIS*) of the error distribution.

The *CIS* of the error distribution is defined by the formula

$$\int d\hat{\mu} \;/\; V^{1/2}(\hat{\mu})$$

Thus for Poisson errors we use $2 \cdot \sqrt{\hat{\mu}}$, for binomial errors we use $2 \cdot \sin^{-1}(\sqrt{\hat{\mu}})$ and for gamma errors $\log \hat{\mu}$.

Such a plot is capable of revealing isolated points with large residuals, or a general curvature, indicating unsatisfactory covariate scales or link function, or a trend in the spread with increasing fitted values, indicating an unsatisfactory variance function (McCullagh and Nelder *1983*, page *216)*.

If the model provides a satisfactory fit, residuals plot should show a 'corridor of values', or should not show any underlying pattern when plotted against the explanatory variables or against the fitted values.

Part 2

Statistical Modelling for Mortality Rates

CHAPTER V

Modelling central rates

5.1 Introduction

The Poisson process provides a useful theoretical background in the analysis of mortality rates, and the basic properties of this process are considered next.

The conditions which the stochastic point event counting process $\{X(t),\ t \geq 0\}$ must satisfy in order to form a ***Poisson process*** are given by the following four assumptions (Kakoulos *1990,* page *92*).

a. The number of point events in non - overlapping time intervals, (more generally, parametric sets) are independent events.

b. The probability that the number of point events, k, occurring in a given interval $[0,t]$, denoted by $a_k(t)$, is the same for all the intervals of the same length. This means that the process is *homogeneous (or stationary)* over time. So, for $k = 0, 1, \ldots$ we have

$$P\,[\,X(t+s) - X(s) = k\,] = a_k\,(t) \qquad \forall\ \ t \geq 0,\ s \geq 0$$

c. In the extremely short 'time' interval $(t\,,\,t+h)$ *one event at most* may occur. That is, there is a constant $\lambda > 0$ such that

$$a_1(h) = P\,[\,X(t+h) - X\,(t) = 1\,] = \lambda \cdot h + o(h)$$

$$a_0(h) = P\,[\,X(t+h) - X\,(t) = 0\,] = 1 - \lambda \cdot h + o(h) \qquad \textbf{(5.1)}$$

where $o(h)$ symbolises a function of h such that $o(h)\,/\,h$ tends to zero when $h \rightarrow 0.$

It follows from the above relationships that the probability that more than one event occurs in $(t,t+h)$ is

$$a_k(h) = o(h) \quad \forall \quad k > 1$$

and it follows from (5.1) that

$$\lim_{h \to 0^+} \frac{a_1(h)}{h} = \lambda$$

The parameter λ gives the rate with which the events occur, referred to as the *intensity* of the Poisson process, and is equal to the expected number of events in a unit time (or parametric) interval.

d. $X(0) = 0$, since we start to count the events at time 0.

So, $X(t)$ simply represents the number of point events occurring in the interval $(0,t)$ or, because of condition ***b***, in any interval $(s,\ s+t)$ with length t.

Any process satisfying the four conditions above is called a homogeneous or simple Poisson process having a *Poisson distribution* with mean $\lambda \cdot t$ (Kakoulos *1990,* page *93*). That is

$$P_k(t) = P[X(t) = k] = exp\ \{\ (-\lambda \cdot t)\} \cdot (\lambda \cdot t)^k / k!\ \cancel{\}}, \text{ for } \quad k = 0,\ 1,\ 2,\ \ldots$$

Next, the following three ***generalisations*** of the Poisson process are of interest in any mortality investigation (as will become clear later in context).

1. The parameter t usually represents time, so that $X(t)$ counts the number of point events occurring up to time t. But if t is a measure of length, area, volume, etc. we still have Poisson process but with parameter *space* instead of time.

2. We can allow the intensity of the process to depend on time t. Thus

$$P[\, X(t+h) - X(t) = 1 \,] = \lambda(t) \cdot h + o(h)$$

and $X(t)$ has again a Poisson distribution but with mean $\int_0^t \lambda(s)\, ds$. In this situation $\{X(t),\, t \geq 0\}$ is referred as a *non - homogeneous* or *time dependent Poisson process.*

3. If in a 'small' time interval more than one event may occur given that at least one event has occurred, we have the *generalised Poisson process* or the *compound Poisson process.* Further assuming that there is a probability function p_k such that for $k=1,2,...$ and $t \geq 0$

$$\lim_{h \to 0^+} P[\, X(t+h) - X(t) = k \;/\; X(t+h) - X(t) \geq 1 \,] = p_k$$

then it can be shown that $\{X(t),\ t \geq 0\}$ is a stochastic process with homogeneous and independent point-events and is a generalisation of the simple Poisson process for which $p_1 = 1$ and $p_k = 0$ for $k \neq 1$ (Kakoulos *1990*, page *100*) .

The compound Poisson process can be written in the form

$$X(t) = \sum_{n=1}^{N(t)} Y_n$$

where $\{N(t),\ t \geq 0\}$ is a simple Poisson process and $Y_n = 0,\ 1,\ 2,\ ...$ are independent random variables with the same distribution which are also independent of $N(t)$. Then,

$$E[X(t)] = E[N(t)] \cdot E[Y_n] = \lambda \cdot t \cdot E[Y_n] \qquad \&$$

$$V[X(t)] = E[N(t)] \cdot V(Y_n) + V[N(t)] \cdot E^2(Y_n) = \lambda \cdot t \cdot E[Y_n^{\,2}] \qquad \textbf{(5.2)}$$

Next, the following three basic ***properties*** of the Poisson process are of interest when modelling crude mortality rates.

1. The intermediate 'time' intervals between consecutive point events $i - 1$ and i say, denoted by T_i , are independent and identical distributed exponential random variables. Hence if W_v denotes the waiting time until the vth event, $T_i = W_i - W_{i-1}$ has density

$$f_{T_i}(t) = \lambda \cdot e^{-\lambda \cdot t}$$

So, the waiting time until the vth event $W_v = T_1 + T_2 + \ldots + T_v$, has the Erlang (gamma) distribution with parameters v and λ. That is

$$f_{W_v}(t) = \lambda \cdot e^{-\lambda \cdot t} \cdot \frac{(\lambda \cdot t)^{v-1}}{(v-1)!}$$

2. If $\{X(t),\ t \geq 0\}$ is a homogeneous Poisson process with intensity λ, then the distribution of the times $t_1 < t_2 < \ldots < t_v$ for the realisation of the v - events given that $X(t) = v$, is the same as the distribution generated by selecting a random sample of v observations from the uniform distribution on $[0, t]$ (Kakulos, *1990,* page *98*).

3. If we know the number of point events that occur in a given 'time' period, then the number of events which occur in any sub - interval depends only on the length of the sub - interval and follow the Bernoulli law. That is, if $\{X(t),\ t \geq 0\}$ is a Poisson process, then $\forall\ 0 < s \leq t$ and $k \leq v$, the distribution of $X(s)$ given $X(t) = v$ is Binomial $(v, s/t)$ (Kakulos, *1990,* page *99*).

5.2 ***Poisson process for deaths***

(using central exposed to risk)

Consider a group of lives all having the same age. Following Subsection *2.4* of Forfar et. al. (*1988*), if Θ denotes the number of deaths and R^c the central exposed to risk, with Θ (but not R^c) modelled as a random variable, then the number of deaths has a Poisson distribution with mean and variance both equal to $R^c \cdot \mu$, where μ denotes the force of mortality. That is,

$$\Theta \sim P(R^c \cdot \mu)$$

This may be likened to a Poisson process, in which the number of point events (deaths), in a fixed interval (the exposure to risk), has a Poisson distribution with intensity μ (the force of mortality).

The values θ and R^c are *minimal sufficient statistics* for μ. Hence it is natural to base all statistical inferences on these two quantities (Sverdrup *1965*). It is assumed that the force of mortality μ is piecewise constant within each age category and investigation period so that the ratio $\{\theta / R^c\}$ is the *maximum likelihood estimator* for μ.

Expressed as a *GLM* based on the independent response random variables $\{\Theta_x\}$ where x denotes age, we have, in comparison with equations (*3.1*)

$$E(\Theta_x) = m_x = R^c_x \cdot \mu_x \qquad \& \qquad Var(\Theta_x) = m_x$$

with scale parameter $\varphi = 1$, prior weights $\omega_x = 1$, and variance function $V(m_x) = m_x$.

For notational convenience, we shall use μ_x for the constant value of the force of mortality over the age interval under discussion, rather than $\mu_{x+\frac{1}{2}}$.

Evaluating the integral of expression (*3.2*), for this particular case, gives the expression for the deviance

$$S(c,f) \;=\; 2\cdot \sum_x \{ y_x \cdot \log(\frac{y_x}{\hat{m}_x}) \;-\; (y_x - \hat{m}_x) \}$$

where y_x denote the observed responses θ_x, and $\hat{m}_x$ denote the fitted values $R_x^c \cdot \hat{\mu}_x$ under the current model. Thus, the above expression for the deviance can be rewritten as

$$S(c,f) \;=\; 2\cdot \sum_x \{ \theta_x \cdot \log(\frac{\theta_x}{R_x^c \cdot \hat{\mu}_x}) \;-\; (\theta_x - R_x^c \cdot \hat{\mu}_x) \} \qquad \textbf{(5.3)}$$

Renshaw (*1991a*), describes the *implementation* of μ_x - graduations in *GLIM* based on these distributional assumptions coupled with the use of log link predictor formulae (the *canonical link* for the Poisson distribution) of the type

$$\eta_x \;=\; \log(m_x) \;=\; \log(\, R_x^c \cdot \mu_x) \;=\; \log(\, R_x^c) + \log(\mu_x) \;=\; \log(\, R_x^c) + \sum_j \beta_j \cdot x^j$$

in which the $\log R_x^c$ term is treated as an *offset* as described in Section *3.2*. Note that the graduation formula

$$\log(\mu_x) \;=\; \sum_j \beta_j \cdot x^j$$

implies that the force of mortality μ_x is modelled as an exponentiated polynomial in age x.

As an alternative to using offsets and / or in the implementation of other link based μ_x graduation formulae such as the power link, new responses $\{ Y_x \}$ based on the transformation

$$Y_x = \frac{\Theta_x}{R_x^c}$$

in which Θ_x is still the random variable are needed.

For these responses

$$E(Y_x) = m_x = \frac{1}{R_x^c} \cdot E(\Theta_x) = \mu_x \qquad \& \qquad Var(Y_x) = \frac{1}{(R_x^c)^2} \cdot Var(\Theta_x) = \frac{m_x}{R_x^c}$$

with scale parameter $\varphi = 1$, prior weights $\omega_x = R_x^c$, and variance function $V(m_x) = m_x$.

The expression *(5.3)* quoted for the deviance S (c, f) still applies (Renshaw, *1991a*).

5.3 Gamma distribution for the resistivity to death

(based on deaths)

As noted by Gerber (*1990*, page *113*), the expression for the log - likelihood under the assumption

$$\Theta \sim P(R^c \cdot \mu)$$

is identical to the expression for the log - likelihood under the assumption

$$R^c \sim G(\theta, \mu) \qquad \textbf{\textit{(5.4)}}$$

where $X \sim G(\alpha, \beta)$ means the gamma distribution with density

$$f(x) = \frac{\beta^\alpha}{\Gamma(\alpha)} \cdot x^{\alpha - 1} \cdot \exp(-\beta \cdot x)$$

Expressed as a *GLM* with the central exposures $\{ R_x^c \}$ as the independent response variables, it follows from the properties of the gamma distribution that

$$E(R_x^c) = m_x = \theta_x \cdot \frac{1}{\mu_x} \qquad \& \qquad Var(R_x^c) = \theta_x \cdot \frac{1}{\mu_x^2} = \frac{m_x^2}{\theta_x}$$

with scale parameter $\varphi = 1$, prior weights $\omega_x = \theta_x$, and variance function $V(m_x) = m_x^2$.

Evaluating the integral in expression (*3.3*), for this particular case, gives the following expression for the deviance

$$S(c,f) = -2 \cdot \sum_x \theta_x \cdot \{\log(\frac{R_x^c}{\hat{m}_x}) - \frac{R_x^c - \hat{m}_x}{\hat{m}_x}\}$$

where $\hat{m}_x = \dfrac{\theta_x}{\hat{\mu}_x}$ denote the fitted values under the current model. Note that the above formal expression for the deviance can trivially be shown to be identical to the expression for the deviance in the previous case, equation *(5.3)*.

Within this context, it is possible to target the resistivity of death μ_x^{-1}, so described by Gompertz (*1825*), by means of log link predictor formulae of the type

$$\eta_x = \log(m_x) = \log(\theta_x) + \log(1/\mu_x) = \log(\theta_x) + \sum_j \beta_j \cdot x^j$$

in which $\log(\theta_x)$ are treated as offsets. Note the graduation formula

$$\log(1/\mu_x) = \sum_j \beta_j \cdot x^j$$

which again implies that the force of mortality, μ_x, is modelled as an exponentiated polynomial in age *x*. Provided that the weights θ_x are all non - zero, the method produces identical graduations to the previous method : see Renshaw et al (*1996b*).

We note also, that assumption *(5.4)* implies

$$\frac{R^c}{\theta} \sim G(\theta, \theta \cdot \mu)$$

Expressed as a *GLM*, with the resistivity to death $Y_x = \dfrac{R_x^c}{\theta_x}$ as the independent response variables, it follows from the properties of the gamma distribution that

$$E(Y_x) = m_x = \frac{1}{\mu_x} \qquad \& \qquad Var(Y_x) = \frac{1}{(\theta_x)^2} \cdot Var(R_x^c) = \frac{m_x^2}{\theta_x}$$

with scale parameter $\varphi = 1$, prior weights $\omega_x = \theta_x$, and variance function $V(m_x) = m_x^2$.

The expression *(5.3)* quoted for the deviance S (*c, f*) still applies. Again, provided that the weights θ_x are all non zero the method produces identical graduations to the previous method.

5.4 Compound Poisson process for policies

(using central exposed to risk)

As in Section *5.2*, in this Section the number of deaths, Θ, is modelled as a Poisson random variable with $E(\Theta) = R^c \cdot \mu$. Again consider a group of lives all having the same age.

In a mortality investigation associated with assured lives the data available do not consist of the actual deaths and the exposures based on individual lives. Each policyholder may have more than one policy and any claim may subsequently give rise to more than one 'death'. The *actual data* available, for this kind of investigation, are the number of policies ceasing through death and the corresponding exposed to risk based on policies. Therefore, a simple Poisson process no longer describes the real process under which the assured lives data are generated.

Let P_i denote the number of *duplicate policies* giving rise to a claim from policyholder *i*. Let θ denote the actual number of deaths, and P the total number of policies giving rise to claims. Let R^c denote the central exposed to risk based on actual deaths. Then, we have the following relationship

$$P = \sum_{i=1}^{\theta} P_i$$

Then, assuming that the P_i's can be treated as independent and identically distributed random variables, it follows from the third generalisation of the Poisson process discussed earlier in Section *5.1* and equations *(5.2)* that

$$E(P) = \mu \cdot R^c \cdot E(P_i) \qquad \& \qquad Var(P) = \mu \cdot R^c \cdot E(P_i^2) \qquad \textbf{(5.5)}$$

under the assumption that there is no mortality selection among policyholders with different number of policies, such that $E(P_i)$ is an unbiased estimate for the average number of duplicate policies giving rise to claims for each policyholder.

Following Forfar et al (*1988*, page *30*), let ${}^{p}R^{c}$ denote the central exposed to risk based on policy counts, θ^{i} denote the number of policyholders who die (at age x) and have i policies and T^{i} denote the central exposure based on lives, arising from those cases for which the policyholders has i policies. Then, we have that

$$\Theta^{i} \sim P\,(T^{i} \cdot \mu) \qquad \& \qquad T^{i} \sim G\,(\theta^{i}, \mu) \tag{5.6}$$

and

$$P = \sum_{i} i \cdot \theta^{i} \qquad \& \qquad {}^{p}R^{c} = \sum_{i} i \cdot T^{i} \tag{5.7}$$

Further it has been proved that (Forfar et al, *1988*, page *31*),

$$E(P) = {}^{p}R^{c} \cdot \mu$$

Then, comparison with the first equation in the equation system *(5.5)* gives that

$${}^{p}R^{c} = R^{c} \cdot E(P_{i})$$

Now, equations *(5.5)* become

$$E(P) = \mu \cdot {}^{p}R^{c} \qquad \& \qquad \mathrm{Var}(P) = E(P) \cdot \{ E(P_{i}^{2}) / E(P_{i}) \} \tag{5.8}$$

In the context of a Poisson *GLM* this feature is described as *over - dispersion* because

$$\mathrm{Var}\,(P) > \mathrm{E}\,(P)$$

since $E\,(P_{i}^{2}) > E(P_{i})$ in practice. See for example Renshaw (*1992*).

Various techniques have been developed to facilitate the graduation process in the presence of over - dispersion. Forfar et al (*1988*), transform the data before modelling by dividing both policy counts and exposures by so-called *variance ratios*, defined as

$$r = \sum i^{2} \cdot f(i) / \sum i \cdot f(i)$$

where $f(i)$ denotes the proportion of policyholders holding i policies. A possible deficiency of this method is that the values of the variance ratios are not always readily available.

The over - dispersion parameter

$$\varphi = \frac{E(P_i^2)}{E(P_i)} = \frac{Var(P)}{E(P)}$$

defined by equation (5.8) is the ratio of the second moment of P_i divided by the first moment of P_i (under the same assumption about mortality selection), or the ratio of the second central moment of P to the first moment of P.

Renshaw (1992), describes a methodology of joint modelling of the mean and of the dispersion, using the over - dispersed Poisson model for policies, such that

$$E(P) = \mu \cdot {}^{p}R^{c} \qquad \& \qquad \mathrm{Var}(P) = \varphi \cdot E(P) \qquad \textbf{(5.9)}$$

where the over - dispersed parameter φ is independent of μ, and is the theoretical equivalent of the empirical variance ratio r discussed by Forfar et al (1988).

The method involves modelling the *unknown* dispersion parameter φ as a secondary inter - related *GLM* in order to model φ as a function of age. In this thesis, we will assume throughout that φ is independent of age since the effect on the graduation process is known to be small and we estimate φ as described in Section 3.3 (Renshaw, 1992).

Thus, expressed as a *GLM*, we model P_x, the total number of policies giving rise to claims at age x, as over - dispersed Poisson response variables where

$$E(P_x) = m_x = {}^{p}R_x^c \cdot \mu_x \qquad \& \qquad Var(P_x) = \varphi \cdot m_x$$

with scale parameter φ, prior weights $\omega_x = 1$, and variance function $V(m_x) = m_x$.

Evaluating the integral of expression (3.2), for this particular case, gives the expression for the deviance

$$S(c,f) = 2 \cdot \frac{1}{\varphi} \cdot \sum_{x} \{ \theta_x \cdot \log(\frac{\theta_x}{R_x^c \cdot \hat{\mu}_x}) - (\theta_x - R_x^c \cdot \hat{\mu}_x) \} \qquad \textbf{\textit{(5.10)}}$$

The same predictor link structures described in Section *5.2* also apply here.

As an alternative to using offsets and / or in the implementation of other link based μ_x graduation formulae such as the power link, new responses $\{ Y_x \}$ based on the transformation

$$Y_x = \frac{P_x}{{}^{p}R_x^c}$$

in which P_x is still the random variable are needed. For these responses

$$E(Y_x) = m_x = \frac{1}{{}^{p}R_x^c} \cdot E(P_x) = \mu_x \qquad \& \qquad Var(Y_x) = \frac{1}{({}^{p}R_x^c)^2} \cdot Var(P_x) = \varphi \cdot \frac{m_x}{{}^{p}R_x^c}$$

with scale parameter φ, prior weights $\omega_x = {}^{p}R_x^c$, and variance function $V(m_x) = m_x$.

The expression *(5.10)* quoted for the deviance *S(c , f)* still applies.

The estimates of the parameters are identical with the Poisson case (if the same mathematical formula is used). The only difference occurs in the standard errors of the parameter estimates and the p - values in the tests of a graduation, since the standardised deviance residuals include the over - dispersed parameter, φ.

To allow for over dispersion, the *Akaike Criterion* of best fit, expression *(3.4)*, is adjusted to

$$AC = Log \{ Q(\underline{y}, \underline{m}_1, p) \} - 2 \cdot \varphi \cdot k$$

5.5 *Gamma distribution for the resistivity to death*

(based on policies)

In this Section, the central exposed to risk based on policies ${}^{P}R^{c}$ is treated as a random variable conditional on the number of policies P ceasing due to deaths θ .

From Section *5.3* the central exposed to risk, R^{c}, is the gamma random variable

$$R^{c} \sim G(\theta \ , \mu) \tag{5.11}$$

The expected number of *duplicate policies* $E(P_i)$ on an individual i is assumed to be the same for all policyholders, and it is assumed that there is no mortality selection among policyholders with different number of policies.

Following Forfar et al (*1988,* pages *30 - 32*), we have similarly, due to equations *(5.6) & (5.7)*, that

$$E({}^{P}R^{c}) = \sum_i i \cdot E(T^{i}) = \sum_i i \cdot \frac{\theta^{i}}{\mu} = \frac{P}{\mu}$$

and

$$Var({}^{P}R^{c}) = \sum_i i^2 \cdot Var(T^{i}) = \sum_i i^2 \cdot \frac{\theta^{i}}{\mu^2} = \frac{\sum_i i^2 \cdot \theta^{i}}{\sum_i i \cdot \theta^{i}} \cdot \frac{\sum_i i \cdot \theta^{i}}{\mu^2} = r \cdot \frac{P}{\mu^2} = \{E({}^{P}R^{c})\}^2 \cdot \frac{r}{P}$$

where

$$r = \frac{\sum_i i^2 \cdot \theta^{i}}{\sum_i i \cdot \theta^{i}}$$

the so - called variance ratios or the theoretical equivalent over - dispersed parameter, φ.

Expressed as a *GLM*, with the central exposed to risk based on policies $\{{}^{p}R^{c}_{x}\}$ acting as independent responses, comparison with equations *(3.1)* gives

$$E({}^{p}R^{c}_{x}) = m_x = \frac{P_x}{\mu_x} \qquad \& \qquad \mathrm{Var}({}^{p}R^{c}_{x}) = \frac{\varphi}{P_x}\cdot m_x^{2} \qquad \textbf{(5.12)}$$

with scale parameter $\varphi = r$, prior weights $\omega_x = P_x$, and squared variance function $\mathrm{V}(m_x) = m_x^{2}$.

The expression *(5.10)* quoted for the deviance *S(c , f)* still applies. The above *GLM* structure is suitable for use in combination with log link predictor formulae as described in Section *5.3*.

We note also, that equations *(5.12)* imply that

$$E\left(\frac{{}^{p}R^{c}_{x}}{P_x}\right) = m_x = \frac{1}{\mu_x} \qquad \& \qquad \mathrm{Var}\left(\frac{{}^{p}R^{c}_{x}}{P_x}\right) = \frac{\varphi}{P_x}\cdot m_x^{2}$$

Thus, as an alternative to using offsets and / or in the implementation of other link based μ_x graduation formulae such as the power link, new responses $\{Y_x\}$ based on the transformation

$$Y_x = \frac{{}^{p}R^{c}_{x}}{P_x}$$

in which ${}^{p}R^{c}_{x}$ is still the random variable are needed. Note that, this is now identical to the situation described in Section *5.3* subject to the introduction of a free standing scale parameter φ. Thus, expressed as a *GLM*, we get

$$\mathrm{E}(Y_x) = m_x = \frac{1}{\mu_x} \qquad \& \qquad \mathrm{Var}(Y_x) = m_x^{2}\cdot\frac{\varphi}{P_x}$$

with scale parameter φ, prior weights $\omega_x = P_x$, and squared variance function $\mathrm{V}(m_x) = m_x^{2}$.

The expression *(5.10)* quoted for the deviance *S(c, f)* still applies. The issues in this section are discussed in greater depth by Renshaw et al (*1996b*).

5.6 Normal distribution for the logarithm of the resistivity to death

Consider the gamma based *GLM* of the previous Section *5.5* with responses $\{Y_x\}$ such that

$$E(Y_x) = m_x = \frac{1}{\mu_x} \qquad \& \qquad Var(Y_x) = \frac{\rho}{P_x} \cdot m_x^2$$

where

$$Y_x = \frac{{}^{p}R_x^{c}}{P_x}$$

According to McCullagh and Nelder (*1989*, pages *285 - 286*), for small $\sqrt{\frac{\rho}{P_x}}$ (in association with the above error structure) we have that

$$E(\log Y_x) = \log \frac{1}{\mu_x} - \frac{1}{2} \cdot \sqrt{\frac{\rho}{P_x}} \qquad \& \qquad \text{Var}(\log Y_x) = \frac{\rho}{P_x}$$

Thus, the log transformation of the inverse of the force of mortality stabilises the variance. The removal of the skewness will be assumed under the normal approximation for the response variable, the natural logarithm of the empirical resistivity to death.

In this section we use this approximate result by modelling $\{\log Y_x\}$ as the responses where

$$E(\log Y_x) = m_x = \log \frac{1}{\mu_x} \qquad \& \qquad Var(\log Y_x) = \rho \cdot \frac{1}{P_x}$$

with scale parameter $\varphi = \rho$, prior weights $\omega_x = P_x$, and variance function $V(m_x) = 1$.

Thus, estimation of the predictor parameters is by the familiar weighted least squares method associated with the normal modelling distribution.

"If the analysis is exploratory or if graphical presentation is required, transformation of the data is convenient and indeed desirable. However, if the response variable Y is a variable with a physical dimension or if it is an extensive variable the method of analysis based on transforming to $\log Y$ seems unappealing on scientific grounds" (McCullagh and Nelder, *1983,* page *150*).

Therefore, the benefits of this approach depend on the results of the associated statistical tests after the normalisation of the response variable.

The following example (Section *5.7*) illustrates the various statistical approaches of this Chapter and illustrates the similarities and differences between them.

5.7 Example

The example is taken from the *CMI* assured males lives experience, for duration *5+*, in the time period *1987 - 1990* and for the age range *22* to *89*.

The graduation formula and hence linear predictor used involves *Legendre polynomials* defined either by

$$L_n(x) = \frac{1}{2^n \cdot n!} \cdot \frac{d^n}{dx^n} (x^2 - 1)^n, \qquad n = 0, 1, 2, ..$$

or by their recursive relationship

$$(n+1) \cdot L_{n+1}(x) - (2 \cdot n+1) \cdot x \cdot L_n(x) + n \cdot L_{n-1}(x) = 0, \quad n = 1, 2, 3, ..$$

$$\text{with} \quad L_0(x) = 1 \;\; \& \;\; L_1(x) = x.$$

The Legendre polynomials satisfy the following system of equations

$$\int_{-1}^{1} L_n(x) \cdot L_m(x)\, dx = 0 \qquad \forall \;\; n \neq m$$

$$\int_{-1}^{1} [L_n(x)]^2\, dx = \frac{2}{2 \cdot n + 1}$$

which implies orthogonality. To achieve this in practice we transform the x using

$$x' = \frac{x - (\frac{a+b}{2})}{(\frac{b-a}{2})}$$

where a and b are the youngest and oldest ages respectively, so that $x' \in [-1, 1]$.

The usefulness of employing orthogonal polynomials lies in the fact that the estimate of an additional coefficient in the predictor structure does not effect the estimates of the other coefficients and that this additional coefficient "may be capable of a relatively simple interpretation" (Forfar et al, *1988*, page *19*).

In this example, the *log link* is used, and the optimisation of the Akaike Criterion leads to the acceptance of *Legendre* polynomial of the fourth degree. Hence the graduation formula throughout is

$$\log \mu_x = \sum_{j=0}^{4} \alpha_j \cdot L_j(x')$$

Table *5.1* summarises the results obtained using the *GLIM* statistical package, for Poisson responses based on the death rate (Section *5.2*), for gamma responses based on the resistivity to death (Section *5.5*), and for normal responses based on the natural logarithm of the resistivity to death (Section *5.6*), as described in this Chapter. Note that the gamma responses, based on the resistivity to death (Section *5.5*), produces identical results with the compound Poisson responses (Section *5.5*), since there are no zero reported deaths in any of the age cells.

Table *5.1* : Results for Poisson, Gamma & Normal responses

	Poisson	***Gamma***	***Normal***
	Parameter estimates (Standard errors)	*Parameter (Standard errors)*	*Parameter (Standard errors)*
α_0	-5.104	-5.104	-5.100
	0.01213	0.0171	0.01698
α_1	3.182	3.182	3.176
	0.02585	0.03639	0.03613
α_2	0.3664	0.3664	0.3719
	0.02725	0.03842	0.03814
α_3	-0.4179	-0.4179	-0.4205
	0.02098	0.02956	0.02934
α_4	0.1468	0.1468	0.1476
	0.01644	0.02336	0.02319
p - values	***Poisson***	***Gamma***	***Normal***
p_{isd}	99	99	98
p_{sign}	40	40	60
p_{runs}	41	41	23
p_{chi}	49	55	48
	$\varphi = 1$	$\varphi = 2.028$	$\varphi = 1.998$

Note that the above *p* - values have been calculated using standardised deviance residuals. Their values, for each of the error distributions, reveal a satisfactory adherence of the graduated rates to the crude rates (chi - square value), with an excellent distribution of the graduated rates around the crude rates (*ISD* - value), a balanced fit (sign - value), and a satisfactory relationship of the resulting curve to the crude rates (runs - value).

The parameter estimates and deviance value *127.75* in association with *63* degrees of freedom are identical for the Poisson and gamma error models. This is to be expected certainly for the log-link model structure when there are no zero reported deaths in any of the age cells, as here. So, the graduated mortality rates are identical for these two cases. The corresponding parameter standard errors (gamma to Poisson) differ by a factor of approximately *1.41*, the square root of the estimated scale factor associated with the gamma model (the scale factor for the Poisson model being *1*). The deviance value *125.89* and the parameter estimates for the normal error model differ only slightly from the deviance and parameter estimates for the other two cases.

Figure *5.1* presents the crude rates (as dots) and the graduated force of mortality (as a continuous curve), on a log scale, for all three error distributions, plotted against age. Note that it is not possible to detect the small differences between the graduated values for the Poisson - gamma and normal error models on such a graph.

Plots of standardised deviance residuals against the appropriate constant information scale, for each error distribution, and against age for the gamma - compound Poisson error distribution, are presented in Figures *5.2 - 5.5*.

Figure 5.1 : Logarithm of $\mathring{\mu}_x$ and μ_x against age

Male assured lives *1987 - 1990*, duration *5+*

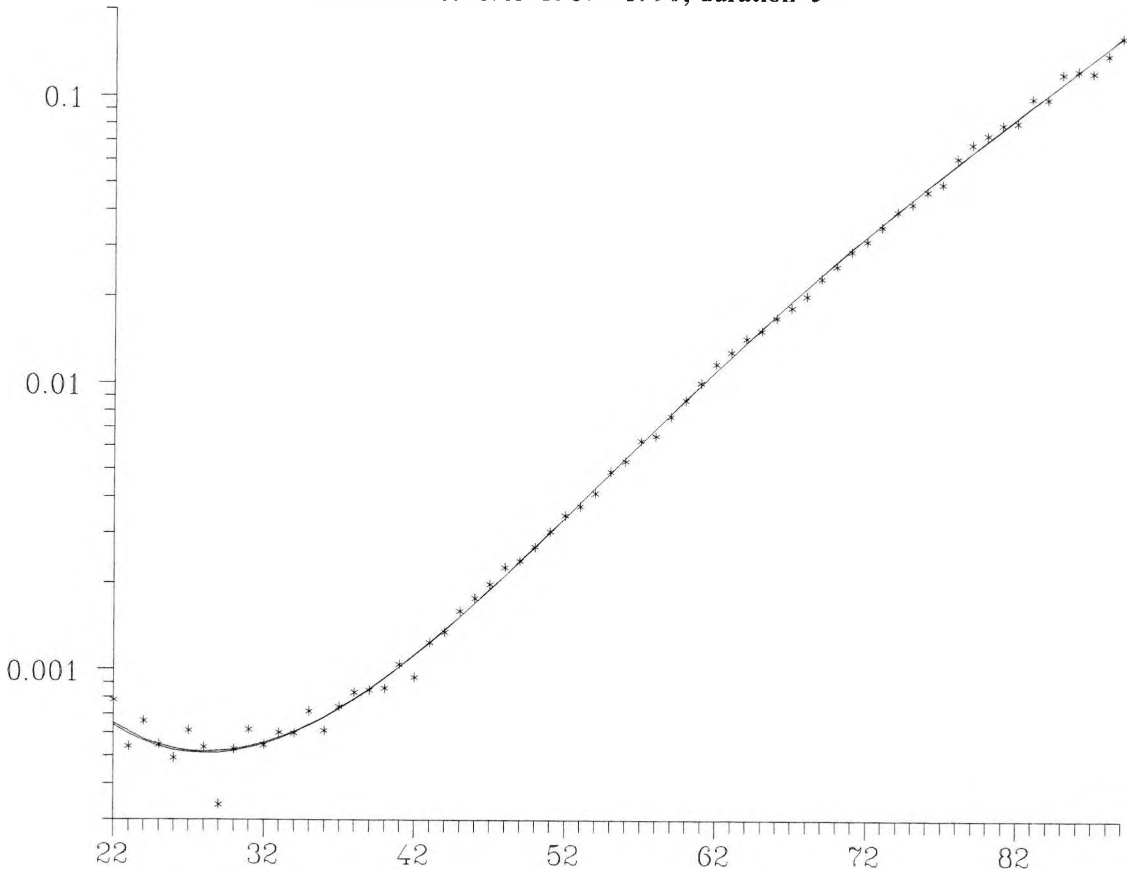


Figure 5.2 : Deviance residuals for *Poisson* error against $CIS = 2\cdot\sqrt{\hat{P}_x}$

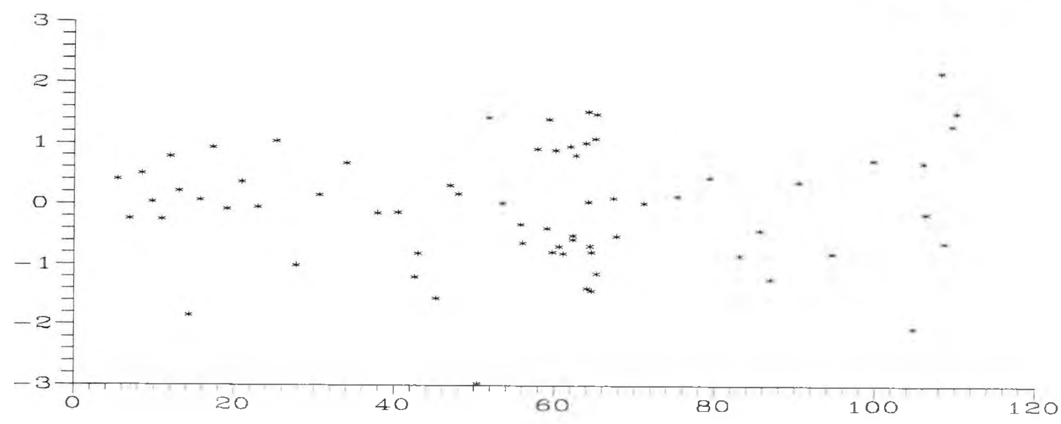


Figure 5.3 : Deviance residuals for *gamma - compound Poisson* error against

$$CIS = 2 \cdot \log(\frac{1}{\hat{\mu}_x})$$

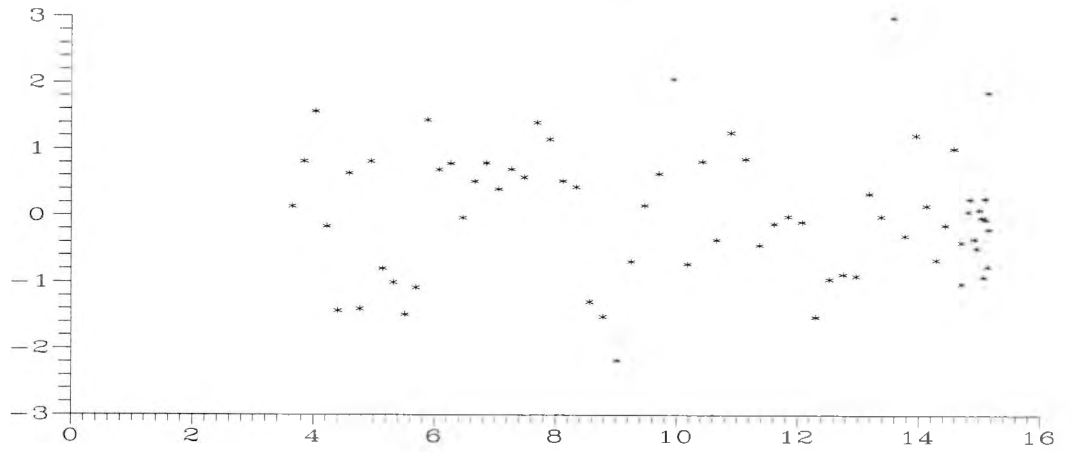


Figure 5.4 : Deviance residuals for *normal* error against $CIS = \log(\frac{1}{\hat{\mu}_x})$

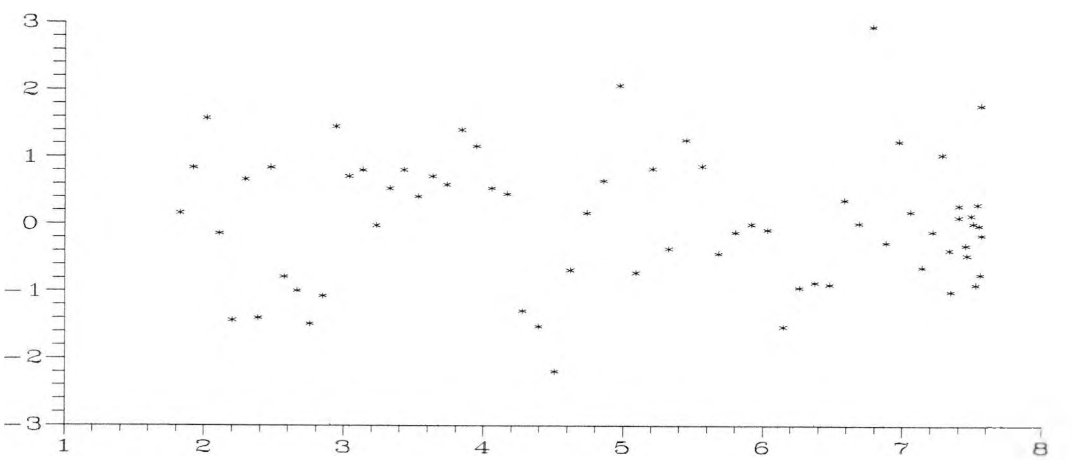
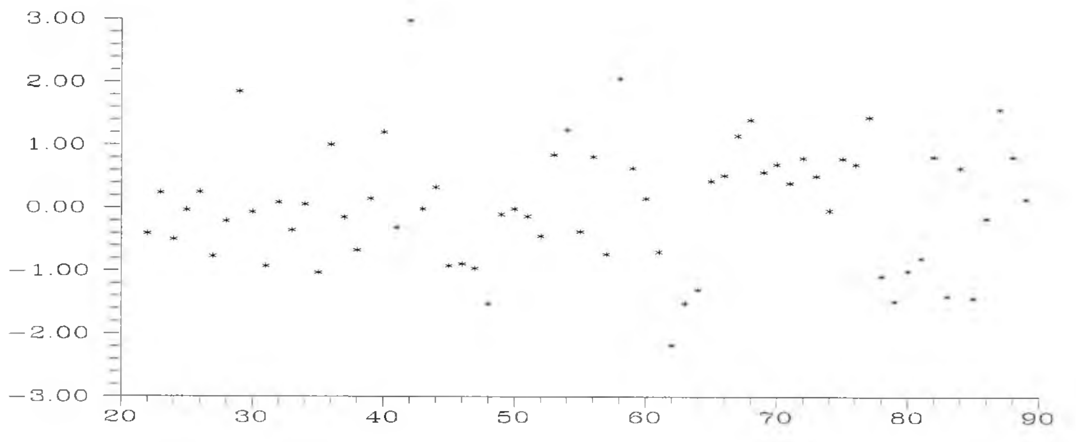


Figure 5.5 : Deviance residuals for *gamma - compound Poisson* error against age *x*



Each of the above figures is supportive of the particular error distributions concerned. The deviance residuals do not show any underlying pattern.

CHAPTER VI

Modelling initial rates

6.1 Binomial distribution for deaths

(Using initial exposed to risk)

Consider a group of lives all having the same age. As described in Chapter *I*, each life contributes a whole year to the exposed to risk on entry into investigation. For reasons of simplicity we assume that there are neither new entrants nor withdrawals, so that each life contributes a whole year to the initial exposed to risk.

Then, it is natural to assume that each life behaves as a *Bernoulli* trial, with a 'success' to denote death, and with the 'probability of a success' to denote a discrete measure of mortality, the *rate of mortality* q, as described in Chapter *I*.

The sum of all these 'successes' aggregates to give the number of deaths, Θ, which has the *binomial* distribution

$$\Theta \sim Bin\,(R^i, q)$$

where R^i denotes the initial exposed to risk. It is assumed that the death or survival of each life is independent of the death or survival of each of the others, for the particular age in question.

The crude rate of mortality, q, is estimated by the ratio of the number of deaths divided by the initial exposed to risk, as described in Chapter *I*, i.e. $\overset{o}{q} = \theta \,/ R^i$, which is the maximum likelihood estimator under the binomial distribution.

The rate of mortality, q_x, is the *conditional probability* of death in the rate interval associated with age x, given that an individual is alive at the beginning of the rate interval with age x.

Expressed as a GLM based on the independent responses $\{\Theta_x\}$ where x denotes age, comparison with equations (*3.1*) implies that

$$E(\Theta_x) = m_x = R^i_x \cdot q_x \qquad \& \qquad Var(\Theta_x) = m_x \cdot (1 - \frac{m_x}{R^i_x})$$

with scale parameter $\varphi = 1$, prior weights $\omega_x = 1$ and variance function

$$V(m_x) = m_x \cdot (1 - \frac{m_x}{R^i_x})$$

Evaluating the integral in expression (*3.3*), for this particular case, gives the expression for the deviance

$$S(c,f) = 2 \cdot \sum_x \left\{ \theta_x \cdot \log(\frac{\theta_x}{\hat{m}_x}) + (R^i_x - \theta_x) \cdot \log\left(\frac{R^i_x - \theta_x}{R^i_x - \hat{m}_x}\right) \right\} \qquad \textbf{\textit{(6.1)}}$$

where

$$\hat{m}_x = R^i_x \cdot \hat{q}_x$$

denote the fitted values under the current model.

Renshaw (*1991b*) describes the implementation of q_x graduations using $GLIM$ based on these distributional assumptions, coupled with the use of the following three (inverse) link functions in combination with polynomial predictors in age effects.

1. The *complementary log - log* link with inverse $\quad q_x = 1 - \exp(-\exp(\eta_x))$

2. The *logit* link with inverse $\quad q_x = \dfrac{\exp(\eta_x)}{1 + \exp(\eta_x)}$

3. The *probit* link with inverse $\quad q_x = \Phi(\eta_x)$

where Φ denotes the cumulative distribution of the standard normal distribution.

6.2 *Compound binomial distribution for policies*

(Using initial exposed to risk)

Consider a group of lives all having the same age. In the presence of duplicate policies, as in Section *5.4*, let P_j denote the number of *duplicate policies* giving rise to a claim from policyholder *j*. Let θ denote the actual number of deaths, and P the total number of policies giving rise to claims. Assume that the random variables P_j are independent and identically distributed for all *j*, and are independent of the number of deaths, θ. Let R^i denote the initial exposed to risk based on actual deaths. Then,

$$P = \sum_{j=1}^{\Theta} P_j$$

and it follows from the well - known relationships, for any compound process, that

$$E(P) = E(\Theta) \cdot E(P_j) \qquad \& \qquad Var(P) = E(\Theta) \cdot Var(P_j) + Var(\Theta) \cdot E^2(P_j) \qquad \textbf{(6.2)}$$

This assumes that there is no mortality selection among policyholders with different numbers of policies. So, $E(P_j)$ is an unbiased estimate for the average number of duplicate policies giving rise to claims for each policyholder.

Under the assumption $\Theta \sim Bin\ (R^i, q)$, so that $V(\Theta) = E(\Theta) \cdot (1\text{-}q)$, we can rewrite expression (*6.2*) for the variance as

$$Var(P) = E(\Theta)\ \{\ E(P_j^2) - E^2(P_j)\ \} + \{\ E(\Theta) \cdot (1\text{-}q)\ \} \cdot E^2(P_j)\ \}$$

This implies

$$Var(P) = E(\Theta) \cdot E(P_j^2)\ - E(\Theta) \cdot q \cdot E^2(P_j)$$

which reduces, on using the first of equations (*6.2*), to

$$Var(P) = E(P) \cdot \{\frac{E(P_j^{\ 2})}{E(P_j)} - q \cdot E(P_j)\}$$

Renshaw (*1992*) has shown that it is possible to rewrite this expression as the variance of an over - dispersed binomial variate, for which

$$Var(P) = \varphi \cdot E(P) \cdot (1 - q)$$

where

$$\varphi = \frac{E(P_j^{\,2})}{E(P_j)} \cdot \left(1 - \frac{E^2(P_j)}{E(P_j^{\,2})} \cdot q\right) \cdot (1-q)^{-1} \tag{6.3}$$

Further, expression (*6.3*) approximates to

$$\varphi \cong \frac{E(P_j^{\,2})}{E(P_j)} > 1$$

because of the relative smallness of q for all but the oldest ages, so that φ does not depend on the target q and may be interpreted as a dispersion parameter.

From the first of equations (*6.2*) and the assumption $\Theta \sim Bin(R^i, q)$ we obtain

$$E(P) = R^i \cdot E(P_j) \cdot q = {}^pR^i \cdot q$$

where we write

$${}^pR^i = R^i \cdot E(P_j)$$

to denote the exposed to risk based on policy rather than head counts. Recall $E(P_j)$ is the expected number of policies giving rise to a claim, per person j, and is the same for all individuals.

Expressed as a *GLM* therefore, the number of policies ceasing through death $\{P_x\}$, for individuals aged x, form the response variables with

$$E(P_x) = m_x = {}^pR_x^i \cdot q_x \qquad \& \qquad Var(P_x) = \varphi \cdot m_x \cdot \left(1 - \frac{m_x}{{}^pR_x^i}\right)$$

with scale parameter $\varphi > 1$, prior weights $\omega_x = 1$ and variance function

$$V(m_x) = m_x \cdot (1 - \frac{m_x}{{}^{p}R^{i}_{x}})$$

Evaluating the integral in expression (*3.3*), for this particular case, gives the expression for the deviance

$$S(c,f) = \frac{2}{\hat{\varphi}} \cdot \sum_{x} \left\{ P_x \cdot \log(\frac{P_x}{\hat{m}_x}) + ({}^{p}R^{i}_{x} - P_x) \cdot \log(\frac{{}^{p}R^{i}_{x} - P_x}{{}^{p}R^{i}_{x} - \hat{m}_x}) \right\} \qquad \textbf{\textit{(6.4)}}$$

where $\hat{\varphi}$ denote the estimated scale parameter, and

$$\hat{m}_x = {}^{p}R^{i}_{x} \cdot \hat{q}_x$$

denote the fitted values under the current model.

6.3 Example

The example is again taken from the *CMI* assured males lives experience, for duration *5+*, in the time period *1987 - 1990* and for the age range *22* to *89*.

The initial exposed to risk are approximated by the well known relationship $R_x^i = R_x^c + \dfrac{\theta_x}{2}$, since the data set is based on central exposures. Legendre polynomials are used as in the previous example.

The following table is a matrix of deviances, in which the columns correspond to the degree (k - 1) of the polynomial predictor and the rows correspond to the different link functions

Table 6.1 : Table of deviance for different link functions

	$k = 1$	$k = 2$	$k = 3$	$k = 4$	$k = 5$	$k = 6$
Log - log	106527.4	524.8	489.6	204.7	127.8	124.9
Logit	106527.4	476.1	464.5	206.7	127.3	124.9
Probit	106527.4	1260.2	343.7	162.8	**125.1**	124.9

For each of the different link function we choose the optimum deviance using the (modified) Akaike Criterion. Then, for the optimum choice for the whole table, we choose from the optimum deviances (using the Akaike Criterion) the deviance based on the least number of parameters, or the one with the minimum deviance value if all the optimum deviances are based on the same number of parameters.

In this example, each of the link functions attain their optimum deviance at $k = 5$, and the probit link function in combination with a quartic in age effects is chosen as the 'optimum' model. The details of the parameter estimates for this model are presented in Table *6.2*, where

$$\Phi^{-1}(q_x) = \sum_{j=0}^{4} \alpha_j \cdot L_j(x')$$

Table 6.2 : Parameter estimates and standard error for the probit link function in combination with a quartic in age effects.

	p.e.	*s.e.*
α_0	-2.419	0.00522
α_1	1.180	0.01098
α_2	0.2521	0.01225
α_3	-0.1012	0.00920
α_4	0.03895	0.00885
	$\hat{\varphi} = 1.986$	

Individual Standardised Deviations (or standardised Pearson residuals) are used as residuals, based on the normal approximation to the binomial distribution. Thus, if Θ has an over dispersed binomial distribution with parameters $(R\ ,\ q)$ then approximately

$$\Theta \approx N(R\cdot q,\ \varphi\cdot R\cdot q\cdot(1-q)) \qquad \& \qquad ISD = \frac{\Theta - R\cdot q}{\sqrt{\varphi\cdot R\cdot q\cdot(1-q)}} \approx N(0,1)$$

The p - values for the statistical tests are

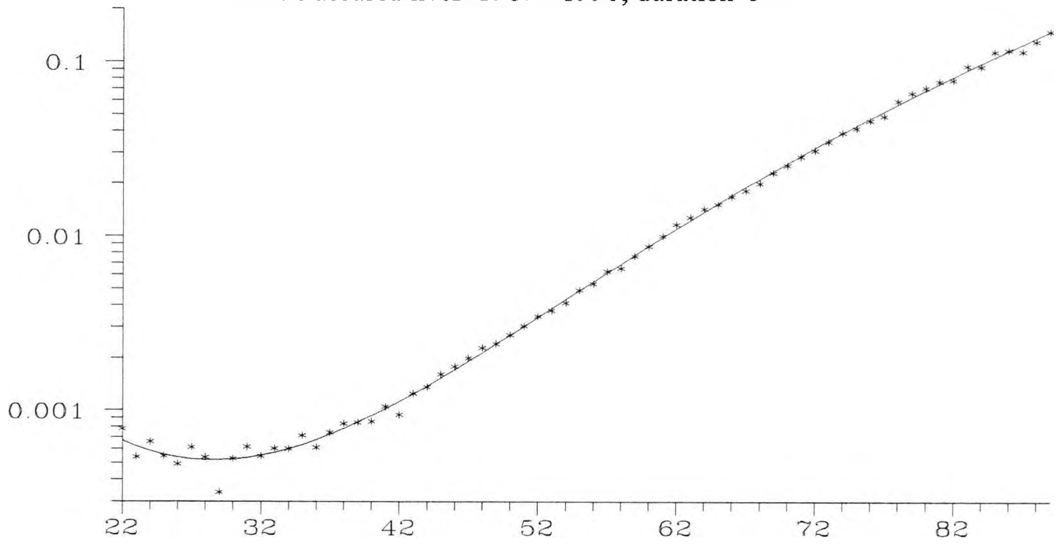
$$p_{ISD} = 0.88 \qquad p_{\mathrm{sign}} = 0.50 \qquad p_{\mathrm{run}} = 0.69 \qquad p_{\mathrm{chi}} = 0.56$$

The above p - values reveal a satisfactory adherence of the graduated rates to the crude rates (chi - square value), with an excellent distribution of the graduated rates around the crude rates (ISD - value), an excellent balanced fit (sign - value), and a satisfactory relationship between the resulting curve and the crude rates (runs - value).

Figure *6.1* displays the crude mortality rates (as dots) and the graduated mortality rates (as a continuous curve), on the log scale, plotted against age.

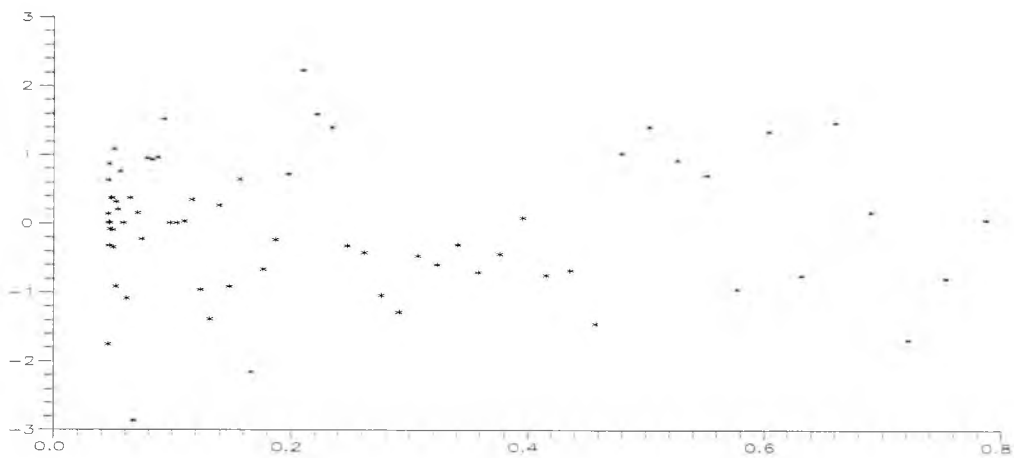
Figure 6.1 : Logarithm of $\overset{\circ}{q}_x$ and q_x against age

Male assured lives *1987 - 1990*, duration *5+*



The Individual Standardised Deviations (*ISD*) are plotted against the constant information scale, defined by $CIS = 2 \cdot \sin^{-1}(q_x)$, in Figure *6.2*.

Figure 6.2 : Individual standardised deviations against *CIS*



The lack of any underlying pattern is supportive of the model.

Part 3

Mathematical Modelling for Mortality Trends

CHAPTER VII

The methodology of modelling mortality trends

7.1 Introduction

The aim of this section is the construction of a mathematical relationship which describes the mortality trends in connection with age and time. The methodology employed can also be extended when further factors of mortality, other than age and time, are also included.

Moreover, using the constructed mathematical model, forecasting of future mortality rates can be considered. However, in order to make any hypothesis about future mortality values, we firstly need some strong remarks about the nature of the past experience and the degree to which this characterises the whole observed mortality experience. For these features, we will make the assumption that they will continue to apply for a sensible time span in the future.

Further, we should like to condense the information contained in the past experience into a set of critical parameters, which contain as much information as possible. This process will have the advantage of providing a better understanding of the evolution of the mortality through time and it will enable us to consider the forecasting of future mortality rates and to consider expanding the future forecasting period.

Forecasting of mortality rates depends strongly on the way that the mathematical modelling has been carried out. The following section describes the method advocated for the mathematical modelling of the mortality rates.

7.2 *Methodology*

Mathematical modelling, in this context, means the construction of a mathematical formula to describe the mortality trends through age and time. Therefore, we need a real function

$$f : R^2 \rightarrow R \qquad \text{such that} \qquad \mu_{x,t} = f(x, t, \underline{b})$$

where $\underline{b}$ is a vector of unknown parameters.

The methodology for the derivation of the function f will be based mainly on the construction of a mathematical formula capable of graduating the data in question for each year separately.

Given mortality data for a sequence of years $\{t\}$ and a sequence of ages $\{x\}$, we can define

$$\mu_{x,t} = h(x, \underline{b}_t)$$

to be the formula which graduates the data for each individual year t, where

$$\underline{b}_t = (\beta_{1t}, \beta_{2t}, \dots, \beta_{kt}) \in R^k$$

denotes a set of parameters for each year t.

Such structures are fitted using *GLIM* by declaring t as a factor. The resulting parameter estimates ($\hat{\beta}_{it}$) are examined for possible trends in time t, for each $i = 1, 2, \dots, k$. By this means, when trends are established, a drastic reduction in the number of parameters is possible by modelling t, as well as x, as an (independent) variable; whereby establishing an appropriate form for the parameterised function f. It is also possible to reverse the roles of x and t in the above process, which we shall do on occasions.

Using the mathematical formula f, we do not insist on 'perfect' tests of a graduation for each of the years concerned. The aim of this method is to derive a simple mathematical expression to describe the underlying pattern in mortality with age over time. The formula f will be extrapolated in time to investigate possible forecast mortality values.

7.3 General description of the mathematical modelling employed in Chapters VIII - XII

The following Chapters (*VIII - XI*) consider the methodology advocated in this Chapter for the mathematical modelling of age specific mortality trends through time. Various approaches are attested employing different mathematical models for the *UK* life offices for whole life and endowment assurances, for the time period *1958 - 1990*, and for pensioners in pensions schemes, for the time period *1983 - 1990*.

For male assured lives, duration *5+*, and ages *24 - 89*, the *log link* (Chapter *VIII*), the *power link* (Chapter *IX*) and the *additive* model structure (Chapter *X*) are analysed.

The log link function is deemed to be an adequate choice for the link for the central mortality rates, justified by the smooth progression imparted to the mortality trends when the log transformation is applied. It gives the minimum deviance when applying a polynomial predictor structure in age and time effects (Section *8.2.2*, model *8.4*). In association with a quadratic spline predictor structure in age effects and a fractional polynomial predictor structure (Royston & Altman, *1994*) in time effects, a flexible model is produced with a parsimonious number of parameters. The knots are located at the age points where the mortality curve changes curvature (distinctively for the multiplicative model, it seems that there exists a critical point in the neighbourhood of the age of *42*, where the mortality 'development' changes curvature, according to the principle of local description in Section *2.1*. This feature is imparted to the power model structures as well).

The power model structure gives the least number of parameters in association with the highest deviance when employing a polynomial predictor structure in age effects and a fractional polynomial predictor structure in time effects (Section *9.2.3*, model *9.5*). Also, employing the power model structure in association with a quadratic polynomial predictor structure, in age and time effects, we obtain a parsimonious number of parameters for each calendar year in question (Section *9.2.2*, model *9.2*).

The additive model produces sound results when it is associated with cubic spline functions in age effects and a fractional polynomial structure in time effects (Section *10.2.2*, model *10.4*).

Further, a different perspective of the above approaches is exercised, by discussing mortality trends through time, for each age in question as regards the multiplicative model structure (Section *8.2.4*), the power model structure (Section *9.2.4*), and the additive model structure (Section *10.2.2*).

Now, focusing on the range of ages [*42, 89*] we have derived some simple mathematical expressions in association with the multiplicative and power model structures.

For the multiplicative model, a simple model structure is presented (Section *8.3*), using a fractional polynomial structure in both age and time effects (model *8.20*).

For the power model, again a simple model structure is presented, using a fractional polynomial structure in time effects and a polynomial predictor structure in age effects (Section *9.3*, model *9.13*).

In Chapter *XI*, the Complementary log - log model is applied for modelling pensioners, ages *60 - 95*, time period *1983 - 1990*, using a polynomial structure in time effects and an inverse polynomial predictor structure in age effects (Section *11.2.2*, model *11.2*).

In Chapter *XII*, on the modelling of amounts, the approach developed for the graduation of 'amounts' provides some insight into the patterns of the claims amounts and of the modelling assumptions, using a polynomial structure in both time and age effects (Section *12.3*, model *12.7*). The methodology is strongly connected with the earlier work by Renshaw (*1992*) on duplicate policies where the effects on the graduation approach are modelled through over-dispersion.

CHAPTER VIII

Multiplicative models

8.1 Introduction

In this Chapter we focus on *log link* predictor relationships and define

$$\eta_{xt} = \log (m_{xt})$$

where η_{xt} denotes the parameterised linear predictor and m_{xt} the expected response. Offsets are declared where necessary.

As implied previously in various sections of Chapter *V*, the log link based parameterised mathematical formulae play a central role in modelling the force of mortality. The log link is the *canonical* or *natural link* when targeting the force of mortality, under the Poisson error distribution. The log link function is also applied in association with the gamma error distribution when targeting the resistivity to death. It is not, however, the canonical link when used in this context.

8.2 *UK male assured lives, duration 5+, period 1958 - 1990, ages 24 - 89*

8.2.1 *Description of the data*

The data, as supplied by the Continuous Mortality Investigation (*CMI*) Bureau, consist of the number of policies ceasing through death, and the *central* exposed to risk of death based on policies, for *UK male assured lives*, by individual calendar year from *1958* to *1990* inclusive and by individual ages.

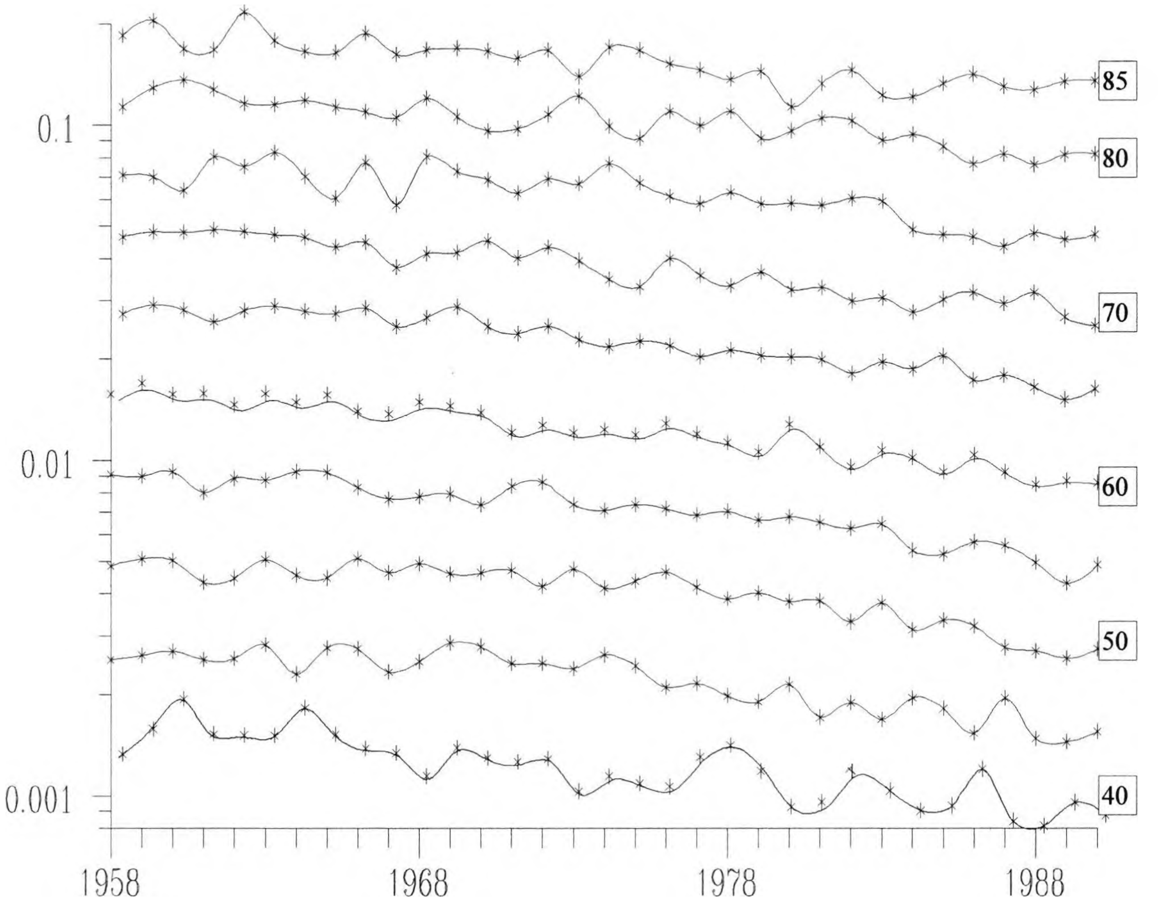
The data are further subdivided, for each calendar year, by duration of either *0, 1, 2, 3, 4* and *5+* years. Within this division the age range is defined by $x = 17+d, \ldots, 66+d$ years for duration $d = 0, 1, 2, 3, 4$ and $x = 24, 23, \ldots, 89$ years for duration *5+*. The data for duration *5+* are known to be suspect for ages in excess of *89* years.

We are mainly concerned here with the data for duration *5+* only, which comprise the bulk of the data. The data are presented in Appendix *A*, as published by the *CMI* Bureau of the Institute and Faculty of Actuaries.

The analysis of durations *0, 1, 2, 3, 4* is studied in Chapter *XIII*, where comparisons between mortality experiences are investigated.

By way of illustration the logarithms of the crude central mortality rates, $\overset{\circ}{\mu}_{xt}$, plotted against calendar years, at five yearly age intervals for duration *5+* years are reproduced in Figure *8.1*. The various curves, which are in descending order of age, starting with age *85* and reducing to age *40*, indicate a general improvement in mortality over the calendar period concerned.

Figure 8.1 : Logarithm of Crude Central Mortality rates from various ages against time



8.2.2 Modelling trends using polynomial predictor structures

We target the force of mortality in accordance with the distributional assumptions of Section 5.4. As an initial investigation *polynomial predictor structures* of degree k in age x were fitted separately for each year t. This heavily parametarised structure

$$\log\,(\mu_{x,t}) = \sum_{i=0}^{k} \beta_{i,t} \cdot L_i(x') \qquad \textbf{\textit{(8.1)}}$$

involves the declaration of time t as a factor with *33* levels (*1958 - 1990*).

The *optimum degree* k for each calendar year is of interest. This is investigated by applying F - tests as described in Section *3.3* (using the normal approximation for the logarithm of the resistivity to death as described in Section *5.6*), for the nested structures

$$H_0:\ \beta_{k,t} = 0 \qquad \text{vs} \qquad H_1:\ \beta_{k,t} \neq 0$$

$k = 4, 5, 6, 7$ and for each calendar year t. Table *8.1* lists the p - values of these F - tests.

Table 8.1 : p - values for hypothesis H_0: $k = 4, 5, 6, 7$' model *(8.1)*

Year	***k = 4***	***k = 5***	***k = 6***	***k = 7***
1958	**00.0**	**00.1**	52.9	05.0
1959	**00.0**	53.9	26.9	78.8
1960	**00.0**	06.6	08.3	95.0
1961	**00.0**	28.2	40.2	41.2
1962	**00.0**	07.5	57.6	**00.6**
1963	**00.0**	07.8	39.2	80.0
1964	**00.0**	45.4	12.3	37.8
1965	**00.0**	39.0	39.4	33.7
1966	**01.0**	64.7	98.6	66.8
1967	**00.0**	**00.0**	97.3	87.7
1968	**00.0**	**00.0**	79.9	05.9
1969	**00.0**	**03.4**	10.5	16.2
1970	06.8	**01.7**	**00.0**	47.3
1971	**00.0**	**00.0**	22.3	34.3
1972	**00.0**	**00.0**	**00.1**	**04.9**
1973	**00.0**	**00.0**	56.1	**00.5**
1974	**00.0**	**00.0**	**00.0**	29.2
1975	**00.0**	**00.0**	26.8	22.4
1976	**00.0**	**00.9**	97.2	37.7
1977	**00.0**	**00.0**	20.9	30.5
1978	**00.0**	**00.8**	06.3	09.8
1979	**00.0**	**00.0**	62.4	**00.7**
1980	**00.0**	**00.0**	21.5	83.4
1981	**00.0**	**00.0**	**00.4**	12.4
1982	**00.0**	**00.1**	17.3	80.1
1983	**00.0**	08.3	38.0	38.3
1984	**00.0**	**02.0**	97.6	34.9
1985	**00.0**	26.6	19.8	17.8
1986	**00.0**	06.1	05.6	18.3
1987	**00.0**	52.5	73.1	70.5
1988.	**00.0**	**03.7**	26.4	22.2
1989	**02.2**	21.2	06.8	91.2
1990	05.4	15.2	09.4	97.9

Significant p - values at the 5% level of significance are highlighted by bold. For $k = 6$ or 7 the null hypothesis $H_0 : \beta_{k,t} = 0$ gives consistently high p values, which means the null hypothesis is supported. For $k = 5$ or 4 the null hypothesis $H_0 : \beta_{k,t} = 0$ gives low significant p values, which means that the null hypothesis is rejected. Considering all these hypothesis tests, we conclude that the model structure with a *5th* degree polynomial is the most efficient parsimonious structure to carry out graduation for each calendar year.

Next, we fit an exponentiated polynomial graduation formula (in age effects) with a multiplicative age *independent* adjustment term for calendar year effects, given, on the log scale, by the following equation

$$\log(\mu_{x,t}) = \sum_{j=0}^{k} \beta_j \cdot L_j(x') + \sum_{i=1}^{r} \alpha_i \cdot t'^i \qquad \textbf{(8.2)}$$

where $L_j(x')$ denote Legendre polynomials of degree j, and x' and t' denote transformations of x and t respectively onto the interval $[-1,1]$ defined in Section *5.7*.

The optimum value of r is determined by monitoring the improvement in the model deviance as the value of r is increased (Recall that the optimum value of $k = 5$ was determined above). The resulting deviance profile is reproduced in Table *8.2*. The optimum value selected is $r = 2$, since there is no reduction of note in the deviance beyond this point..

Table 8.2 : Deviance profile for various additive polynomial predictors of degrees r and s

k	$r = 0$	$r = 1$	$r = 2$	$r = 3$	$r = 4$
5	*24759*	*5113*	***4374.5***	*4374*	*4373.8*

Finally, the structure of the linear predictor is further refined through the introduction of mixed polynomial terms in age and calendar year effects by switching to multiplicative age *dependent* adjustment term for calendar year effects, given, on the log scale, by the following equation

$$\log(\mu_{x,t}) = \sum_{j=0}^{k} \beta_j \cdot L_j(x') + \sum_{i=1}^{r} \alpha_i \cdot t'^i + \sum_{i=1}^{r} \sum_{j=1}^{k} \gamma_{ij} \cdot L_j(x') \cdot t'^i \qquad \textbf{(8.3)}$$

Starting with the predetermined values of $r = 2$ and $k = 5$, one possible sequence for introducing mixed product terms of increasing degree leads to the following extension of the deviance profile reported in Table *8.3*.

Table 8.3 : Deviance profile for additional product terms

model	*deviance*	*d.f.*	*differences* deviance	d.f
$r = 2,\ k = 5$	*4374.5*	*217*		
			145.1	*1*
$+\gamma_{11}$	*4229.4*	*216*		
			38.2	*1*
$+\gamma_{21}$	*4191.2*	*216*		
			53.3	*1*
$+\gamma_{12}$	*4137.9*	*216*		
			24.6	*1*
$+\gamma_{22}$	*4113.3*	*216*		
			25	*1*
$+\gamma_{13}$	*4088.3*	*216*		
			45.5	*1*
$+\gamma_{23}$	*4042.8*	*216*		
			12.2	*1*
$+\gamma_{14}$	*4030.6*	*216*		
			12.9	*1*
$+\gamma_{24}$	*4017.7*	*216*		

Noting the reductions in the deviance as further model terms are added, coupled with the examination of the significance of the individual parameters, the final model adopted is

$$\log\,(\mu_{x,t}) = \sum_{j=0}^{5} \beta_j \cdot L_j(x') + \sum_{i=1}^{2} \alpha_i \cdot t'^{i} + \sum_{i=1}^{2} \sum_{j=1}^{3} \gamma_{ij} \cdot L_j(x') \cdot t'^{i} \qquad \textbf{(8.4)}$$

This later expression is quadratic in time, on the log scale, while the coefficients of the quadratic are themselves polynomials in age effects, x.

The quasi-likelihood parameter estimates and their standard errors are given in Table *8.4*.

Table 8.4 : Parameter estimates, standard error, and t - values for model *(8.4)*

	p.e.	s.e.	t - values
α_1	-0.2641	0.006214	-42.5
α_2	-0.05622	0.011548	-4.9
β_0	-4.7451	0.0049	-968.4
β_1	3.1899	0.010258	311
β_2	0.1457	0.010225	14.2
β_3	-0.3232	0.010467	-30.9
β_4	0.2139	0.007318	29.2
β_5	-0.0882	0.006679	-13.2
γ_{11}	0.0004535	0.01298	0.04
γ_{21}	-0.05589	0.02402	-2.3
γ_{12}	0.078137	0.011127	7.0
γ_{22}	0.121517	0.0206627	5.9
γ_{13}	-0.042916	0.010558	4.1
γ_{23}	-0.097016	0.019602	4.9

$\hat{\varphi} = 1.868$

With the exception of γ_{11} the t - statistic associated with each parameter estimate, calculated by dividing the parameter estimate by its standard error, has an absolute value in excess of *2*, indicating statistical significance. The scale parameter also quoted in Table *8.4*, is estimated by dividing the model deviance by the associated degrees of freedom. The magnitude of the scale parameter gives an indication of the degree of over - dispersion present.

By way of illustration the same predictor structure was refitted by targeting the resistivity to death in accordance with the distribution assumptions of Section *5.5.*

The values of the resulting parameter estimates are reproduced in Table *8.5*. It is of interest to note that the parameter estimates are identical in magnitute to those of Table *8.4* but opposite in sign due to the replacement of $\log(\mu_{xt})$ on the *LHS* of expression *(8.4)* with $\log(\mu_{xt}^{-1})$. This dual property of graduation under the assumptions of Section *5.4* and Section *5.5,* leading to basically identical graduations, is developed further in Renshaw et al (*1996b*).

Table 8.5 : Parameter estimates, standard error, and t - values for model *(8.4)* based on gamma responses

	p.e.	**s.e.**	***t* - values**
α_1	*0.2641*	*0.00622*	*42.5*
α_2	*0.05622*	*0.0116*	*4.8*
β_0	*4.7451*	*0.0049*	*968.4*
β_1	*-3.1899*	*0.01025*	*-311*
β_2	*-0.14575*	*0.01023*	*-14.2*
β_3	*0.3232*	*0.01049*	*30.9*
β_4	*-0.2139*	*0.00739*	*-28.9*
β_5	*0.0882*	*0.00672*	*13.2*
γ_{11}	*-0.0004326*	*0.013*	*-0.04*
γ_{21}	*0.05589*	*0.02417*	*2.3*
γ_{12}	*-0.07815*	*0.01115*	*-7.0*
γ_{22}	*-0.12155*	*0.02080*	*-5.8*
γ_{13}	*0.04293*	*0.010612*	*4.0*
γ_{23}	*0.09705*	*0.01972*	*4.9*
	$\hat{\varphi} = 1.868$		

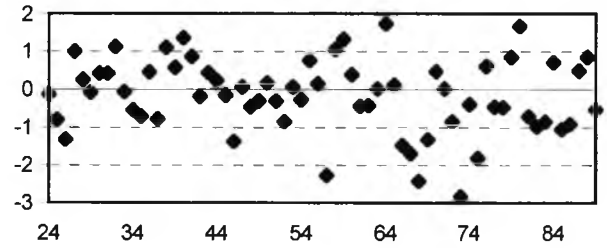
Next, a summary of some of the formal statistical tests of a graduation, applied to all *33* separate calendar years, is presented in Table *8.6*. These involve an analysis of the standardised deviance residuals, for each calendar year, t, based on the tests described in Chapter *IV*. The few significant entries, implying failure of the test concerned, all at the *5%* level of significance, are highlighted by bold.

Table 8.6 : p - values, formal graduation tests for each calendar year separately for model *(8.4)*

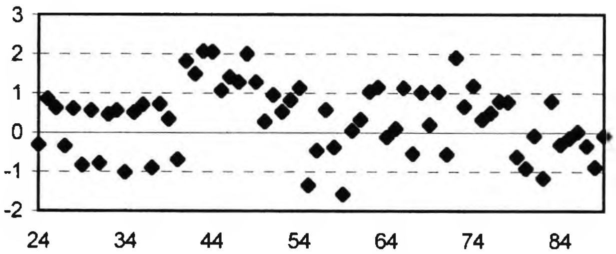
Year	***ISD***	***Sign***	***Runs***	***Chi***
1958	82	31	41	45
1959	94	68	23	48
1960	61	31	41	42
1961	92	40	84	46
1962	77	42	**1**	44
1963	16	**98**	69	57
1964	86	40	6	48
1965	60	93	**4**	54
1966	80	68	78	48
1967	**0**	**0**	72	39
1968	60	4	**2**	43
1969	**2**	93	15	61
1970	15	**99**	29	55
1971	31	10	**1**	43
1972	41	93	15	52
1973	80	83	8	49
1974	27	68	89	48
1975	60	4	18	41
1976	74	40	40	48
1977	32	**2**	59	46
1978	82	7	**2**	47
1979	74	7	31	44
1980	65	93	**4**	47
1981	43	16	44	46
1982	99	23	42	49
1983	31	83	87	51
1984	43	10	97	44
1985	78	40	69	51
1986	68	68	61	47
1987	26	23	97	40
1988	72	4	74	41
1989	53	4	8	50
1990	44	**98**	**2**	48

Figure *8.2* displays just a few of the standardised deviance residual plots against age examined, for each calendar year.

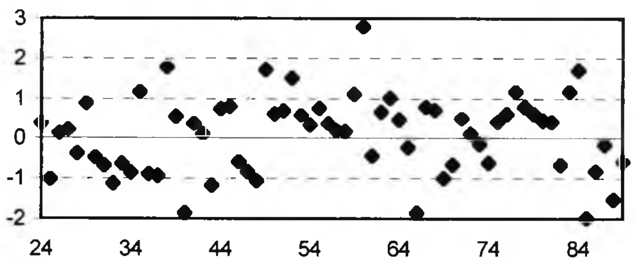
Figure 8.2 : Standardised deviance residuals vs. age, various calendar years, model *(8.4)*



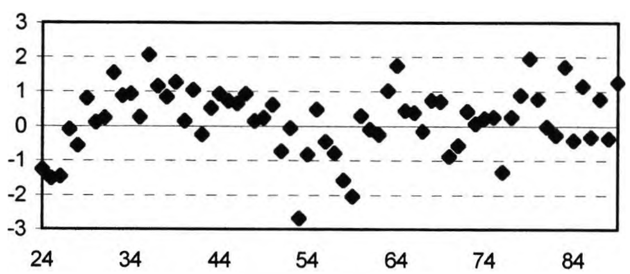
calendar year *1960*



calendar year *1970*



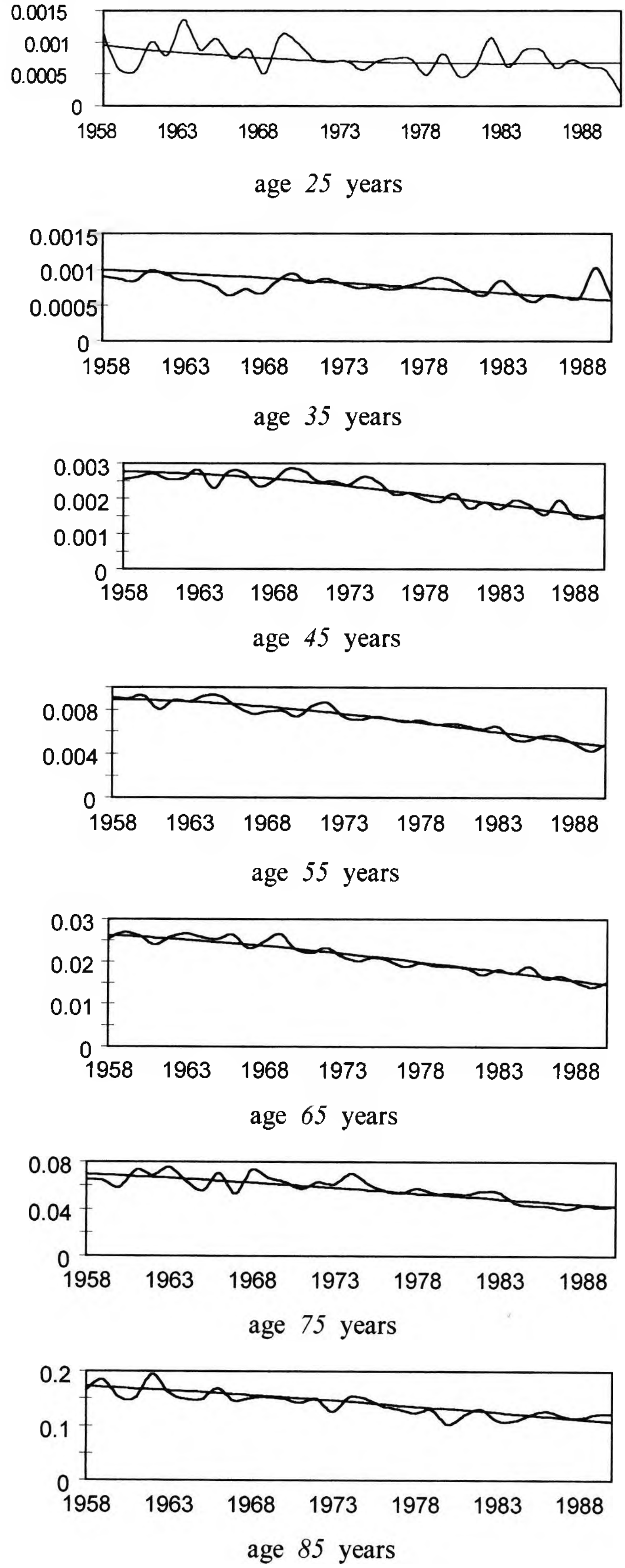
calendar year *1980*



calendar year *1990*

To illustrate the impact of the age specific trend adjustments on mortality, we plot the predicted force of mortality against calendar year at ten yearly age intervals in Figure *8.3*.

Figure 8.3 : Crude and predicted force of mortality vs. calendar year, various ages, model *(8.4)*



Here we have superimposed the crude mortality curves on the corresponding predicted mortality rates. This acts as a further visual check on the predictive qualities of the model. At each age, the graduated values are given by an exponentiated quadratic in calendar time with age specific polynomial coefficients (Renshaw et al, *1996a*).

Finally, for this model, the predicted values of the force of mortality, μ_{xt}, in the age range x = *24* to *89* years, over the calendar period t = *1960* to *1990* at *5* yearly intervals, are presented for completeness in Table *8.7*.

Table 8.7 : Predicted force of mortality, quinquennial periods - ages, model *(8.4)*

	1960	1965	1970	1975	1980	1985	1990
25	0.00093	0.00082	0.00074	0.00070	0.00068	0.00069	0.00072
30	0.00078	0.00071	0.00066	0.00061	0.00057	0.00054	0.00051
35	0.00098	0.00092	0.00086	0.00079	0.00072	0.00065	0.00058
40	0.00155	0.00148	0.00139	0.00127	0.00113	0.00099	0.00085
45	0.00274	0.00264	0.00248	0.00226	0.00200	0.00172	0.00144
50	0.00496	0.00480	0.00451	0.00411	0.00363	0.00312	0.00260
55	0.00887	0.00856	0.00803	0.00733	0.00651	0.00561	0.00471
60	0.01535	0.01474	0.01381	0.01264	0.01130	0.00986	0.00841
65	0.02573	0.02454	0.02297	0.02111	0.01905	0.01687	0.01467
70	0.04216	0.03998	0.03741	0.03454	0.03146	0.02828	0.02508
75	0.06817	0.06447	0.06038	0.05601	0.05145	0.04680	0.04216
80	0.10889	0.10320	0.09695	0.09031	0.08339	0.07634	0.06928
85	0.16903	0.16158	0.15272	0.14270	0.13184	0.12042	0.10875

From the above Table it is deduced that the predicted force of mortality for the age *25* is raised in the last years, even if the observed values, for that age, show a general decline over the years in question. This feature could be granted to the high level of 'noise' in observed values.

8.2.3 Modelling trends using quadratic spline predictor structures in age effects and fractional polynomial predictor structures in time effects

For the data under consideration, it is suspected that the empirical central rate of mortality changes curvature with age, on the logarithmic scale, in the region of age 42 years, for each calendar year. It is observed that the force of mortality for ages less than approximately 42 years has a convex shape, changing to a concave shape for ages greater than approximately 42 years. This characteristic can be modelled by using polynomial predictors of degree greater than one in which the second derivative changes sign at the critical age of approximately 42 years. Further, it will be shown that quadratic predictors can be used to graduate the two age ranges in a very satisfactory way.

An alternative way to describe this feature mathematically in the case of a *quadratic predictor* is for the coefficient of the quadratic term to be positive in the age range less than the critical age, changing to negative in the age range greater than the critical age. Thus

$$\log(\mu_x) = \begin{cases} \alpha_1 + \beta_1 \cdot x + \gamma_1 \cdot x^2 & \text{if } x < k \quad (\gamma_1 > 0) \\ \alpha_2 + \beta_2 \cdot x + \gamma_2 \cdot x^2 & \text{if } x \geq k \quad (\gamma_2 < 0) \end{cases} \tag{8.5}$$

for a specific t, where k denotes the critical age.

The use of such quadratic expressions, in association with the log transformed force of mortality, can be justified on the basis of the *theoretical hypothesis* that the rate at which the *resistivity to death* decreases with age, divided by the resistivity itself, is a linear function of age, that is

$$\frac{d}{dx}\left(\frac{1}{\mu_x}\right) = -(\beta + \gamma \cdot x) \cdot \left(\frac{1}{\mu_x}\right)$$

This may be viewed as a generalisation of Gompertz' law, for which $\gamma = 0$ (as discussed in Section 2.2). The above relationship can be viewed as a linear differential equation of the first degree with variable coefficients, namely

$$f'(x) + (\beta + 2 \cdot \gamma \cdot x) \cdot f(x) = 0, \qquad f(x) = \frac{1}{\mu_x}$$

and solved to give expressions of the type

$$\log(\mu_x) = \alpha + \beta \cdot x + \gamma \cdot x^2$$

In order to ensure continuity at the critical age k, expressions *(8.5)* are combined into a *quadratic spline function* with a single knot located at the critical age. Thus,

$$\log(\mu_x) = \alpha + \beta \cdot x + \gamma \cdot x^2 + \delta \cdot (x-k)_+^2 \qquad \textbf{(8.6)}$$

where
$$(x-k)_+^2 = \begin{cases} 0 & if \quad x \le k \\ (x-k)^2 & if \quad x > k \end{cases}$$

Hence, expression *(8.6)* can be rewritten in the form of expression *(8.5)* as

$$\log(\mu_x) = \begin{cases} \alpha + \beta \cdot x + \gamma \cdot x^2 & \text{if} \quad x < k \\ (\alpha + \delta \cdot k^2) + (\beta - 2 \cdot k \cdot \delta) \cdot x + (\gamma + \delta) \cdot x^2 & \text{if} \quad x \ge k \end{cases} \qquad \textbf{(8.7)}$$

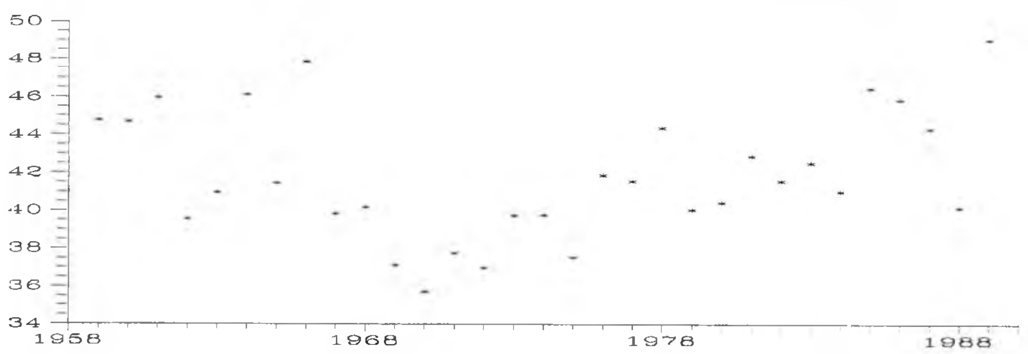
All predictor link structures in this subsection were fitted by targeting the resistivity to death in accordance with the distributional assumptions of Section *5.5*.

To start with, structure *(8.6)* was fitted for each calendar year separately, so that

$$\log(\mu_{x,t}) = \alpha_t + \beta_t \cdot x + \gamma_t \cdot x^2 + \delta_t \cdot (x - k_t)_+^2 \qquad \textbf{(8.8)}$$

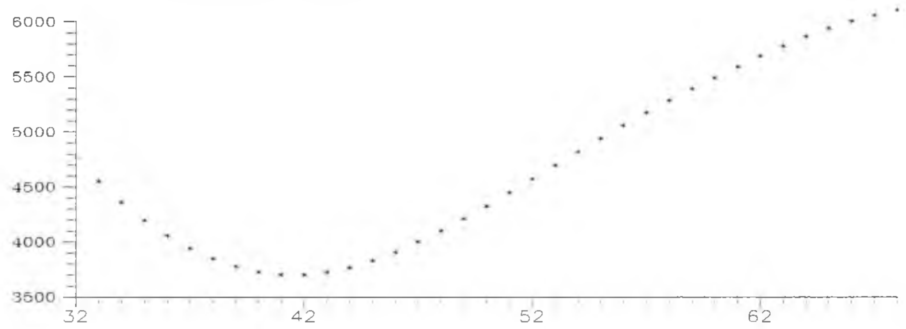
The optimum knot positions, determined by minimising the deviance, for each calendar year in question are shown in Figure *8.4*.

Figure 8.4 : Optimum knot position against calendar year, model (8.8)



From Figure *8.4,* it is reasonable to assume a constant critical age $k_t = k = 42$ years. Further supportive evidence for the actual positioning of a single constant knot $k_t = k = 42$ in equation *(8.8)*, is to be found in the deviance profile for this structure, reproduced in Figure *8.5.*

Figure 8.5 : Profile of deviance against knot position k, model *(8.8)* with $k_t = k$



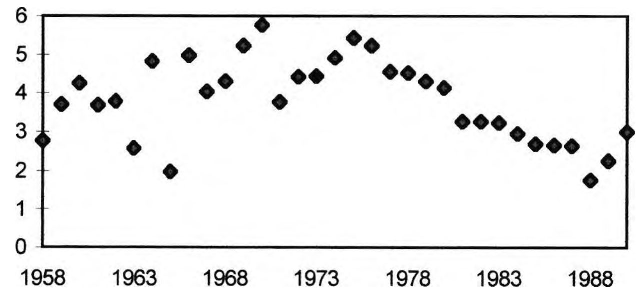
It is easily concluded from this Figure that the minimum value of the deviance is obtained when the knot position approximates the age *42* years, where the deviance is *3703* on *2046* degrees of freedom. Experiments involving the introduction of a second knot were tried and rejected on the basis of deviance profiles.

The trend in the parameter estimates for the heavily parameterised model structure

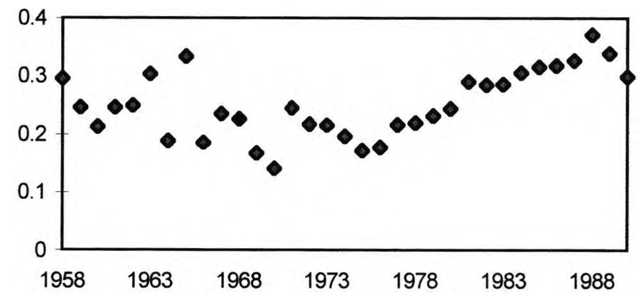
$$\log(\mu_{x,t}) = \alpha_t + \beta_t \cdot x + \gamma_t \cdot x^2 + \delta_t \cdot (x - 42)_+^2 \qquad \textbf{(8.9)}$$

are displayed in Figure *8.6,* while the choice of model has also been further justified on the basis of the statistical tests of a graduation, described in Chapter *IV*, but not reproduced here.

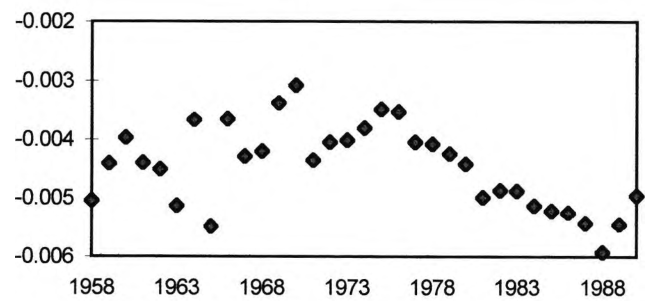
Figure 8.6 : Trend in parameters estimates through time, model *(8.9)*



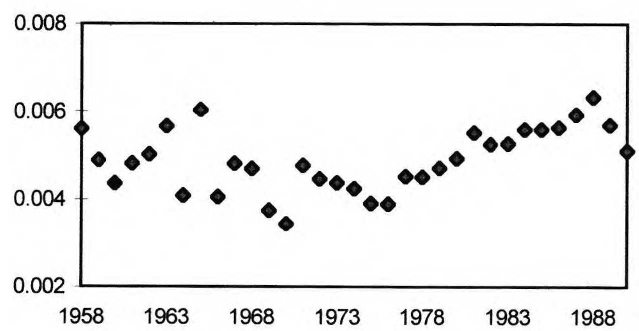
α_t - trend



β_t - trend



γ_t - trend



δ_t - trend

In an attempt to simplify the heavily parameterised structure *(8.9)* and produce a model with smoother parametric trends, the values of the deviance for various nested structures, determined by setting certain of the parameters equal to a constant, are presented in Table *8.8*.

Table 8.8 : Deviances for various simplifications of model *(8.9)*

constant	***deviance***	***d.f.***
δ	3774	2078
γ	3777	2078
β	3777	2078
α	**3766**	2078
γ, δ	3927	2110
β, δ	3897	2110
α, δ	**3862**	2110
β, γ	3865	2110
α, γ	3967	2110
α, β	4227	2110

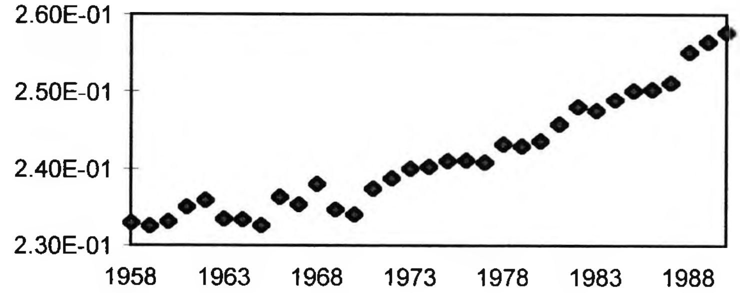
From Table *8.8* it is revealed that the model structure with the minimum deviance is attained when the parametric vector α is kept constant (from among the models with one constant parameter), and when the parametric vectors α, δ are kept constant (from among the models with two constant parameters). Noting that the difference in the unscaled deviance between model *(8.9)* and the nested model with α, δ treated as constants is *159* on *64* degrees of freedom. The effective (approximate) p - value is *6*%, allowing for a scale parameter with value *2.04*. Alternatively the value of the F - statistic is *1.37* on $(64,\infty)$ degrees of freedom, with an approximate p - value of *8*%.

It is also desirable to investigate the model structure in which the parametric vectors α, δ are kept constant (since it leaves only two time dependent parameters, a desirable property according to Anson, *1988*). The trends in the two sets of parameter estimates, for the time dependent parameters in the model structure

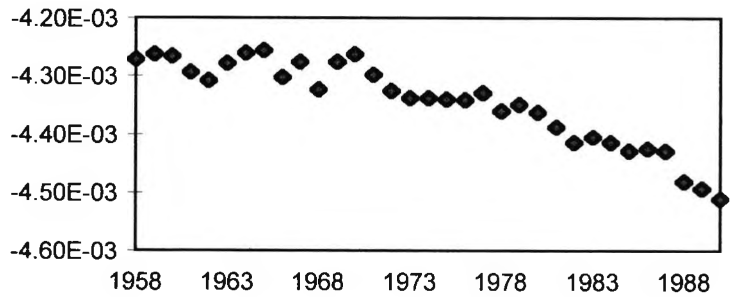
$$\log(\mu_{x,t}) = \alpha + \beta_t \cdot x + \gamma_t \cdot x^2 + \delta \cdot (x-k)_+^2 \qquad \textbf{(8.10)}$$

are displayed in Figure *8.7*, and the adequacy of the model, on the basis of the statistical tests of a graduation, has been justified. Note that the trends in the estimated parameters are quite smooth.

Figure 8.7 : Trend in parameter estimates through time, model *(8.10)*



β_t - trend



γ_t - trend

Finally, in order to produce a parsimonious model, the number of parameters is reduced further using *fractional polynomials* (Royston and Altman, *1994*) of the type

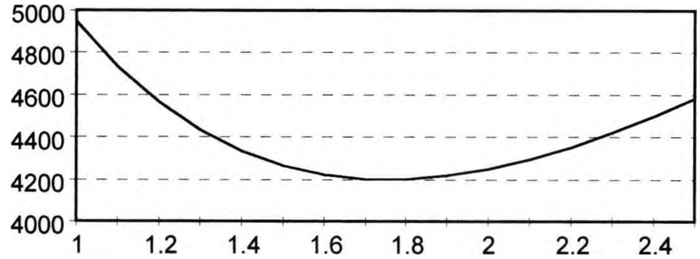
$$a + b \cdot t^k$$

to represent the variation in *both* β_t & γ_t (the empirical coefficient of correlation, between the parametric vectors β & γ, takes the value $\hat{\rho}_{\beta,\gamma} = -0.995$), where k is a fixed index. The value $k = 1.8$ is based on the deviance profile of Figure *8.8*, constructed by fitting the model structure

$$\log(\mu_{x,t}) = \alpha + (\beta_1 + \beta_2 \cdot t^k) \cdot x + (\gamma_1 + \gamma_2 \cdot t^k) \cdot x^2 + \delta \cdot (x - 42)_+^2 \qquad \textbf{(8.11)}$$

for different values of k (in steps of 0.1), where $t =$ calendar year - 1957.

Figure 8.8 : Deviance profile for different values of k, model *(8.11)*



The parameter estimates, standard errors, and t - values for the model structure

$$\log(\mu_{x,t}) = \alpha + (\beta_1 + \beta_2 \cdot t^{1.8}) \cdot x + (\gamma_1 + \gamma_2 \cdot t^{1.8}) \cdot x^2 + \delta \cdot (x - 42)_+^2 \qquad \textbf{(8.12)}$$

are as shown in Table *8.9.*

Table 8.9 : Parameters estimates, standard errors, and t - values, model *(8.12)*

	p.e.	*s.e.*	*t - values*
α	-4.003	0.1737	-23.04
β_1	-0.2328	0.00867	-26.85
β_2	-0.00004336	0.00000093	-46.62
γ_1	0.004263	0.000107	39.84
γ_2	0.0000004075	0.0000000138	29.52
δ	-0.00477	0.0001126	-42.36
	$\hat{\varphi} = 1.934$		

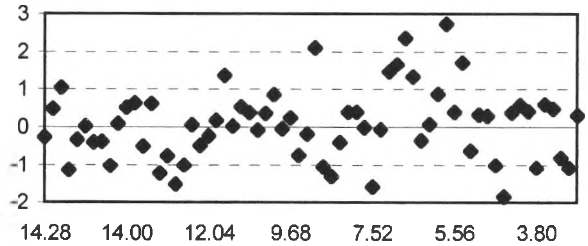
The deviance for the model structure is *4200.5* on *2172* degrees of freedom.

The p - values for the statistical tests of a graduation are presented in Table *8.10,* and just some of the many standardised deviance residual plots examined (on the constant information scale $CIS = 2 \cdot \log(1/\mu_{xt})$), for various calendar years, presented in Figure *8.9.*

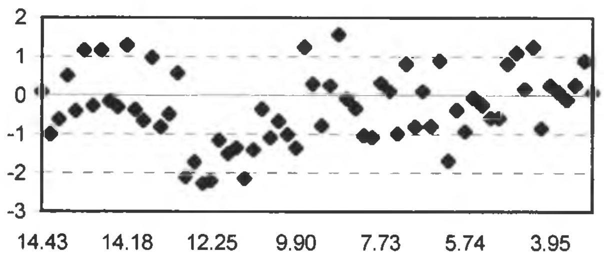
Table 8.10 : *p* - values, formal statistical tests for each calendar year separately, model *(8.12)*

Year	ISD	Sign	Runs	Chi
1958	69	23	42	44
1959	74	23	11	41
1960	30	76	33	58
1961	73	59	50	50
1962	83	76	17	58
1963	**1**	**0**	46	31
1964	77	76	7	49
1965	57	4	7	40
1966	43	68	84	53
1967	**0**	**0**	33	67
1968	26	**99**	29	63
1969	16	31	16	27
1970	15	**2**	39	37
1971	28	93	**1**	62
1972	57	10	6	44
1973	88	40	16	50
1974	91	59	77	55
1975	44	**98**	14	65
1976	98	59	7	51
1977	**4**	**99**	71	53
1978	6	**99**	7	52
1979	62	**98**	9	58
1980	54	10	**3**	52
1981	39	95	8	53
1982	96	76	62	47
1983	63	10	75	43
1984	77	89	83	55
1985	63	68	78	41
1986	68	16	18	51
1987	88	50	89	64
1988	94	68	41	60
1989	68	89	6	48
1990	15	**1**	8	44

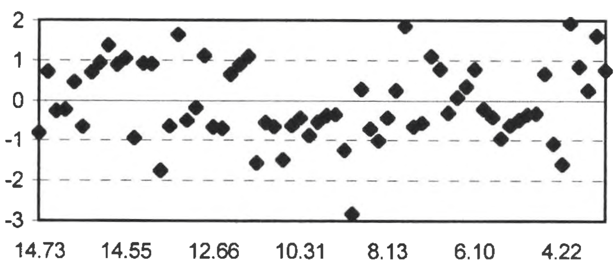
Figure 8.9 : Standardised deviance residuals vs. *CIS*, various calendar years, model *(8.12)*



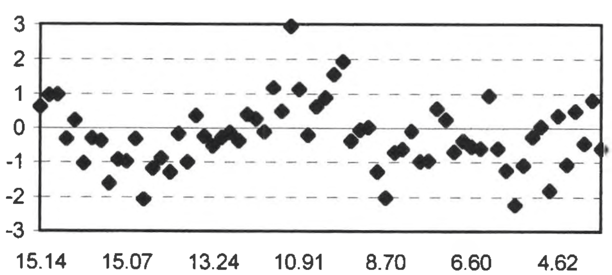
calendar year *1960*



calendar year *1970*



calendar year *1980*



calendar year *1990*

Such diagnostics are supportive of the structure, but before reporting further findings, we take a different perspective of the structure.

8.2.4 *Analysis of age specific mortality trends*

It is desirable to take a different perspective of the above approach by discussing mortality trends through time, for each age in question. In particular it is possible to rearrange equation *(8.12)* in the following way

$$\log\ (\mu_{t,x}) = A(x) + B(x) \cdot t^{1.8} \qquad \textbf{(8.13)}$$

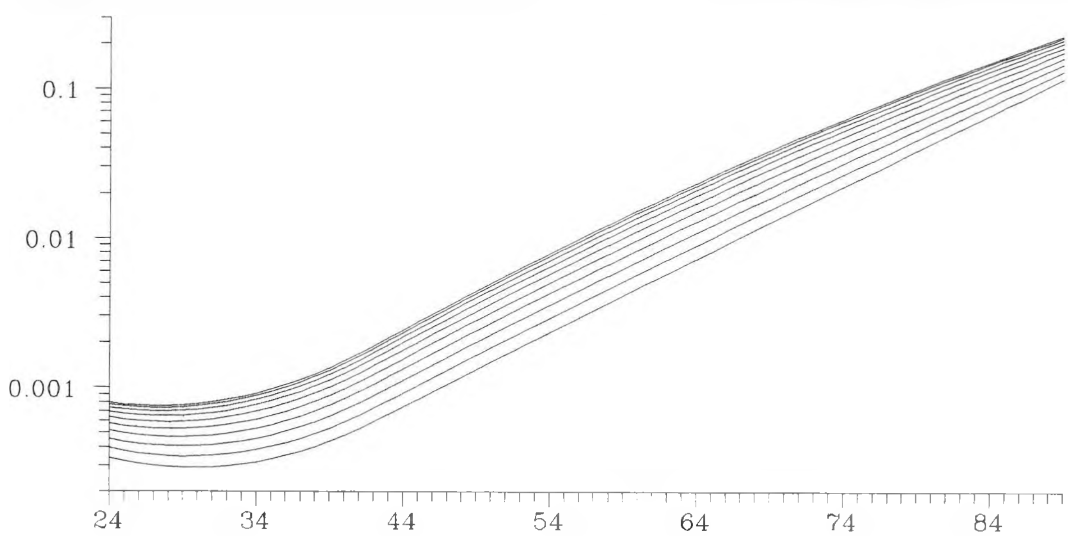
where

$$A(x) = \alpha + \beta_1 \cdot x + \gamma_1 \cdot x^2 + \delta \cdot (x-42)_+^2 \qquad \& \qquad B(x) = \beta_2 \cdot x + \gamma_2 \cdot x^2$$

and the values of the parameters are as quoted in Table *8.9*. Thus, the log of the force of mortality is represented by a fractional polynomial in time effects with age dependent coefficients.

By way of illustration, the graphs in Figure *8.10* illustrate the force of mortality plotted against age, as predicted by the model structure *(8.13),* at five yearly time periods, starting with *1960* through to *2005.*

Figure *8.10* : Log - mortality against age, various periods, based on model structure *(8.13)*



The highest mortality curve corresponds to the calendar year *1960*, moving downwards through progress calendar years, to the year *2005.* This represents a fairly uniform overall improvement in mortality across all ages.

As a further check on the model structure *(8.13)* it was decided to fit the following model structure

$$\log(1 / \mu_{t,x}) = A_x + B_x \cdot t^{1.8} \qquad \textbf{(8.14)}$$

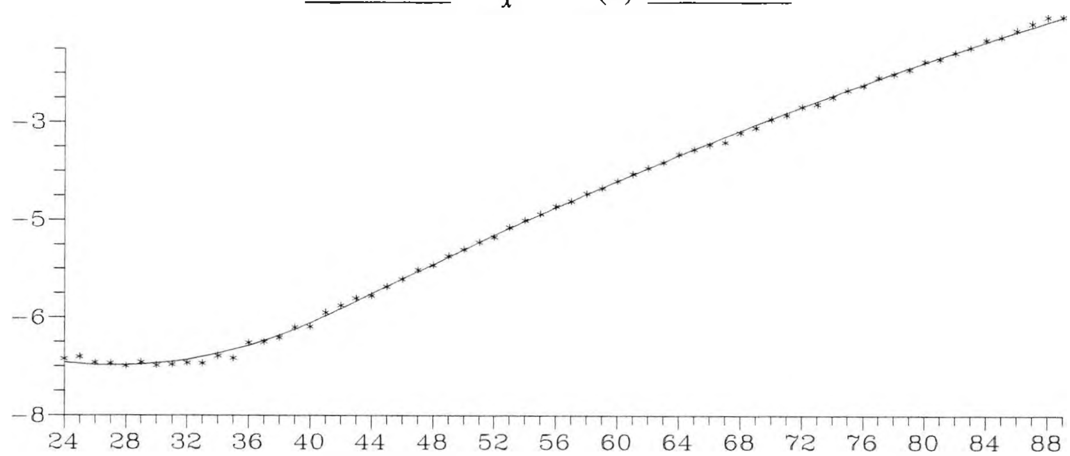
treating age as a factor. Not suprisingly this heavily parameterised structure was found to fit the data well. As evidence of this p - values for the dual statistical tests based on each age (rather than period) are presented in Table *8.11*.

Table 8.11 : p - values, formal statistical tests for each age separately, model *(8.14)*

Age	*ISD*	*Sign*	*Runs*	*Chi*
24	16	97	70	41
25	99	57	81	65
26	50	43	19	74
27	99	30	72	59
28	100	43	30	48
29	18	30	72	75
30	95	57	6	43
31	99	30	20	52
32	96	70	58	37
33	100	57	43	46
34	84	19	98	59
35	96	57	**2**	43
36	58	70	44	55
37	94	30	82	50
38	99	43	97	58
39	83	30	31	51
40	94	57	**2**	43
41	96	70	58	39
42	50	43	57	44
43	97	57	29	47
44	28	19	92	51
45	99	57	30	45
46	85	57	**3**	37
47	97	70	**3**	50
48	100	57	**0**	45
49	68	30	45	50
50	64	43	89	49
51	12	70	11	52
52	74	57	70	51
53	78	43	30	48
54	78	81	13	45
55	80	30	45	49
56	99	43	30	47
57	34	89	68	44
58	95	57	**2**	47
59	96	30	59	49
60	78	19	22	46
61	99	43	11	50
62	93	70	59	44
63	99	43	**0**	47
64	93	43	**3**	48
65	85	70	11	44
66	43	89	53	41
67	89	81	62	47
68	96	30	72	52
69	77	70	95	49
70	100	57	19	50
71	72	57	94	46
72	77	30	44	44
73	84	57	70	50
74	87	19	6	51
75	95	30	31	50
76	87	19	96	52
77	47	19	22	53
78	97	57	70	48
79	44	57	70	52
80	37	89	67	47
81	91	70	72	45
82	97	70	19	49
83	93	70	44	48
84	82	70	**1**	49
85	84	70	20	47
86	84	70	72	44
87	94	57	89	49
88	36	19	75	50
89	55	89	68	40

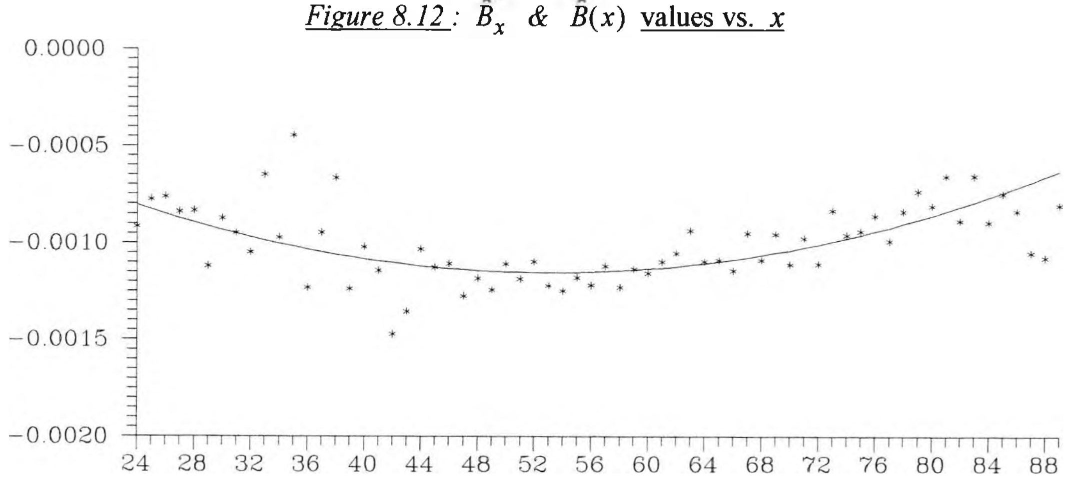
It is informative to plot the two sets of parameter estimates $\hat{A}_x$ & $\hat{B}_x$ for this model *(8.14)* against the respective curves $\hat{A}(x)$ & $\hat{B}(x)$ defined above for model *(8.13)*. This is done in Figures *8.11* & *8.12* respectively. Both Figures are supportive of the choice of model *(8.13)*. Note the different scale used for Figures *8.11* & *8.12*.

Figure 8.11 : $\hat{A}_x$ & $\hat{A}(x)$ values vs. x



We also note that $\hat{A}_x$ & $\hat{A}(x)$ are similar in shape to 'crude' and 'graduated' mortality curves respectively, on the log scale, at time $t = 1$ *(1958)*.

Figure 8.12 : $\hat{B}_x$ & $\hat{B}(x)$ values vs. x



The values $\hat{B}_x$ represent the pace of mortality improvement in time, on the log scale, for each age x. Lower values denote faster improvement. So Figure 8.12 indicates that mortality improvement in the middle ages is higher than that for the youngest and the oldest ages, which have about the same pace of mortality improvement, on the log scale.

Additional diagnostic evidence for model *(8.13)* is provided by the plots of standardised deviance residuals against the constant information scale $(CIS = 2 \cdot \log(1/\mu_{tx}))$ at ten yearly age intervals in Figure *8.13*. The predicted force of mortality (for the time period *1990* to *2010*), against calendar year at ten yearly age intervals is shown in Figure *8.14*.

Figure 8.13 : Standardised deviance residuals vs. *CIS*, various ages, model *(8.13)*

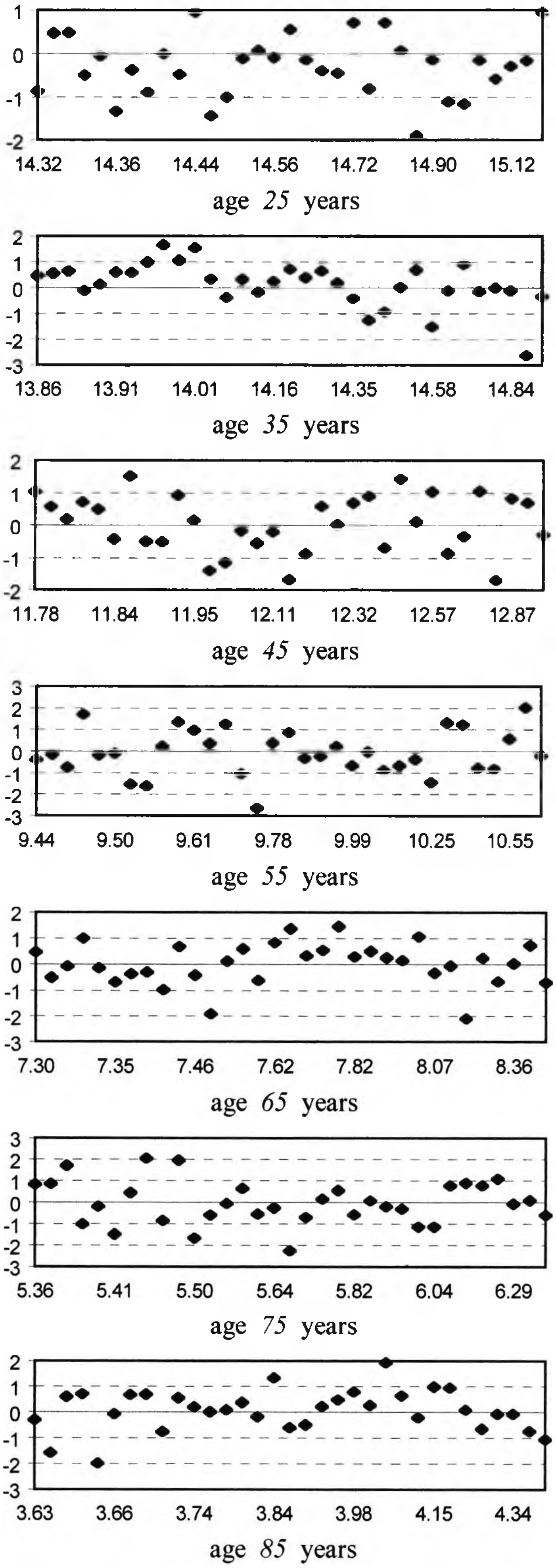
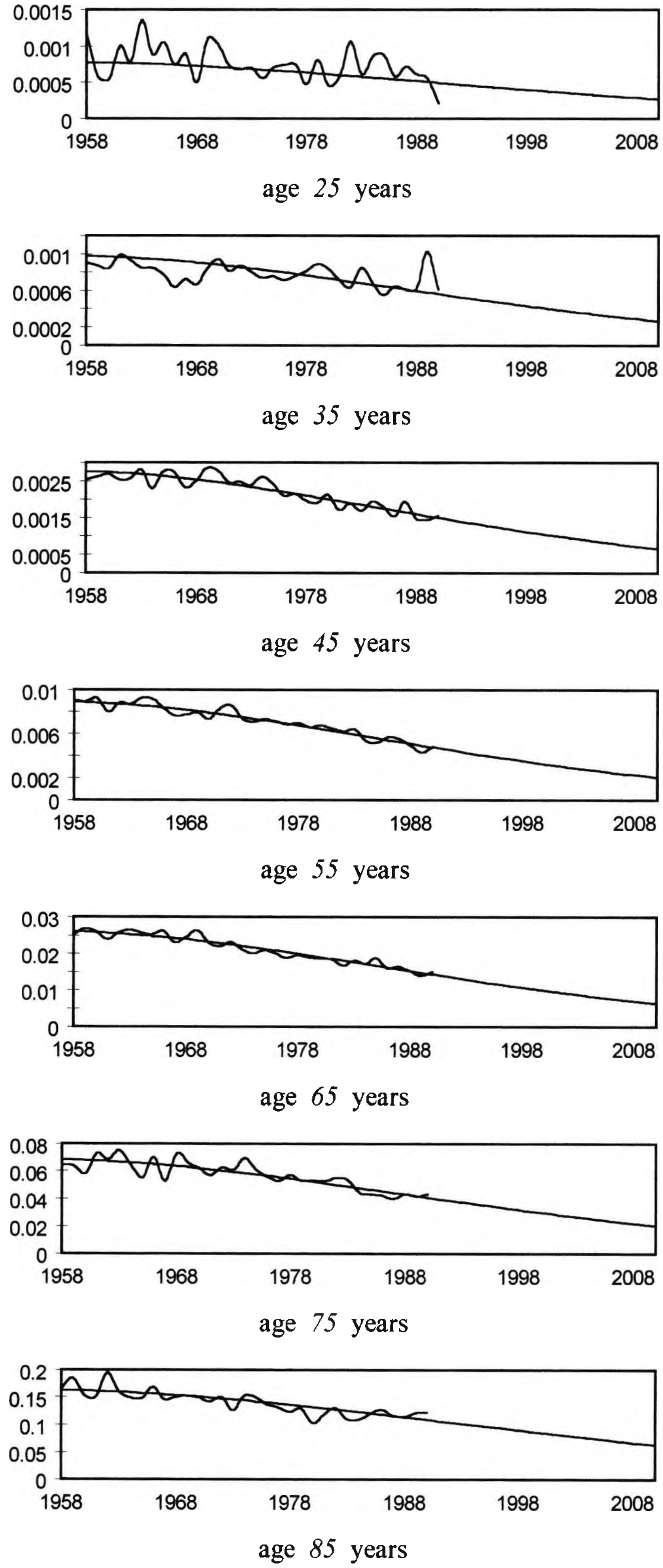


Figure 8.14 : Crude and predicted force of mortality vs. calendar year, various ages, model *(8.14)*



Here we have superimposed the estimated mortality curves on the corresponding crude mortality rates. This acts as a further visual check on the predictive qualities of the model. At each age, the graduated values are given by an exponentiated fractional polynomial of the type $a + b \cdot t^{1.8}$ in calendar time, with age specific quadratic polynomial coefficients.

Finally for this model, the predicted values of the force of mortality, in the age range $x = 24$ to 89 years, over the calendar period $t = 1960$ to 1990 at 10 yearly intervals, and the forecast values of the force of mortality over the calendar period $t = 2000$ to 2010 at 10 yearly intervals, are presented for completeness in the following Table 8.12.

Table 8.12 : Predicted and forecasting force of mortality, 10 years period, quinquennial ages, model (8.13)

	1960	**1970**	**1980**	**1990**	**2000**	**2010**
25	*0.00077*	*0.00072*	*0.00062*	*0.00050*	*0.00038*	*0.00027*
30	*0.00078*	*0.00071*	*0.00060*	*0.00047*	*0.00035*	*0.00024*
35	*0.00097*	*0.00088*	*0.00073*	*0.00056*	*0.00040*	*0.00027*
40	*0.00150*	*0.00136*	*0.00111*	*0.00084*	*0.00059*	*0.00038*
45	*0.00275*	*0.00247*	*0.00201*	*0.00151*	*0.00104*	*0.00066*
50	*0.00500*	*0.00448*	*0.00364*	*0.00270*	*0.00185*	*0.00117*
55	*0.00886*	*0.00795*	*0.00645*	*0.00479*	*0.00327*	*0.00207*
60	*0.01533*	*0.01378*	*0.01122*	*0.00837*	*0.00575*	*0.00366*
65	*0.02586*	*0.02333*	*0.01912*	*0.01440*	*0.01002*	*0.00648*
70	*0.04253*	*0.03858*	*0.03196*	*0.02443*	*0.01734*	*0.01146*
75	*0.06822*	*0.06233*	*0.05237*	*0.04086*	*0.02976*	*0.02031*
80	*0.10668*	*0.09839*	*0.08417*	*0.06737*	*0.05070*	*0.03599*
85	*0.16269*	*0.15174*	*0.13265*	*0.10951*	*0.08572*	*0.06381*

It is of interest to investigate some properties, over time, of the model structure (8.13). We first note that, under this model structure, the force of mortality does not increase with time, since $B(x)$ is always negative (Figure 8.12).

Further, it can be seen Figure 8.14 that the predicted mortality curves change their curvature during the time period involved. This feature indicates that the rate of the mortality decline through time reaches a maximum, in that time period, and afterwards diminishes. In mathematical terms, this turning point can be viewed as the time point where the second derivative, with respect to time, equals zero. That is, when

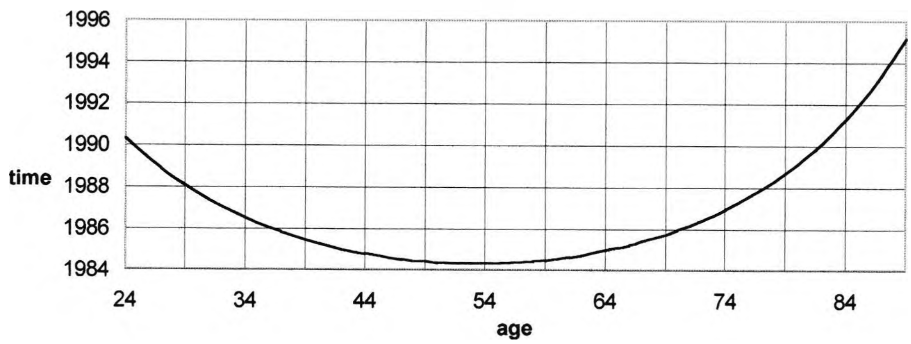
$$\frac{\partial^2}{\partial t^2} \mu_{xt} = 0$$

which leads, after some algebrical manipulations of formula *(8.13)*, to the point

$$t = \left(-0.444 \cdot \frac{1}{B(x)}\right)^{\frac{1}{1.8}}$$

The 'points of inflection' t_x are plotted against x in Figure *8.15*.

Figure 8.15 : Time - points where the second derivative equals zero, with respect to time, for model structure *(8.13)*



It is possible to conclude from Figure *8.15* that the rate of the mortality decrease reaches its maximum during the *1980*'s decade for ages in the neighbourhood of *53*.

8.3 UK male assured lives, duration 5+, period 1958 - 1990, ages 42 - 89

As noted in the previous section, the force of mortality viewed as a function of age, changes curvature in the neighbourhood of age *42* years. This enables one to investigate some simpler trend structures in the restricted age range *42* to *89* years.

In this section we begin by investigating model structure of the type

$$\mu_{x,t} = A_t \cdot B_t^{\sqrt{x}} \tag{8.15}$$

This structure was arrived at by first trying models of the type

$$\mu_x = A \cdot B^{x^k}$$

for each period t and different predetermined values of k.

By analogy with the Gompertz type differential equation defined by the relationship *(2.1)*, we obtain the following linear differential equation (of degree one) with variable coefficients

$$f'(x) - a \cdot x^{k-1} \cdot f(x) = 0 \tag{8.16}$$

where *f(x)* denotes the resistivity to death at age x. This generalisation includes Gompertz's law as an obvious special case when $k = 1$. The only difference with Gompertz's law lies in the fact that we are to use $\sqrt{x}$ instead of x. This reflects the fact that the force of mortality, on the log scale, is no longer linear in age but is linear in the square root of the age.

There is evidence in the literature to suggest that mortality rates increase less rapidly from age to age at the oldest ages compared to the younger adult ages (see Perks (*1932*), Redington (*1969*) for example). This suggests that $k < 1$ would be an expected choice given that we are here focusing in the age range *42* to *89*.

To implement equation *(8.15)* we note that it is equivalent to the equation

$$\log(\mu_{x,t}) = \alpha_t + \beta_t \cdot \sqrt{x} \qquad \textbf{(8.17)}$$

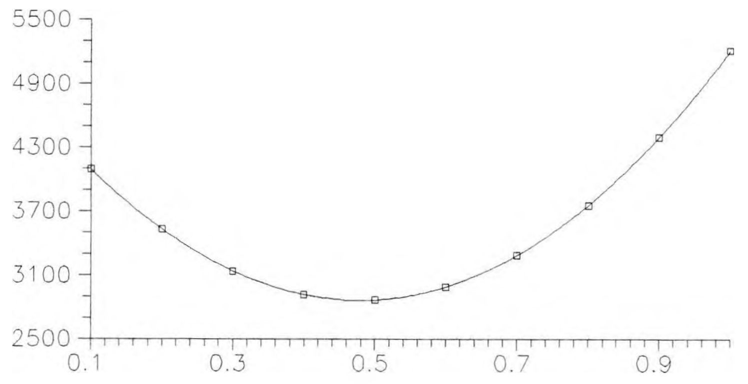
where we choose to target the resistivity to death in accordance with the distributional assumptions of Section *5.5* for consistency. This gives identical results to the targeting of the force of mortality based on over - dispersed Poisson responses, since all data cells contain non-zero numbers of deaths.

To justify the choice of $k = 0.5$ the deviance profile for various values of k (in steps of 0.1) under the model structure

$$\mu_{xt} = A_t \cdot B_t^{x^k} \qquad \textbf{(8.18)}$$

in which t is treated as a factor, is reproduced in Figure *8.16*.

Figure 8.16 : Deviance profile for various values of k for the model structure (8.18)



The p - values for the statistical tests based on model *(8.17)* were then obtained using standardised deviance residuals, and the results indicate high acceptance for the model used, except for a few (randomly) scattered years, where the runs test fails.

The trends in the two sets of parameter estimates under the model structure *(8.17)* are presented in Figures *8.17* & *8.18*.

Figure 8.17 : $\hat{\alpha}_t$ against t, model (8.17)

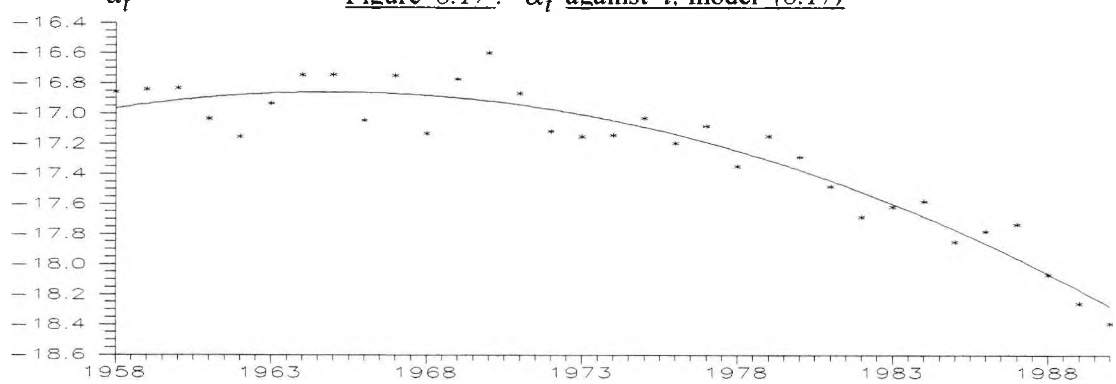
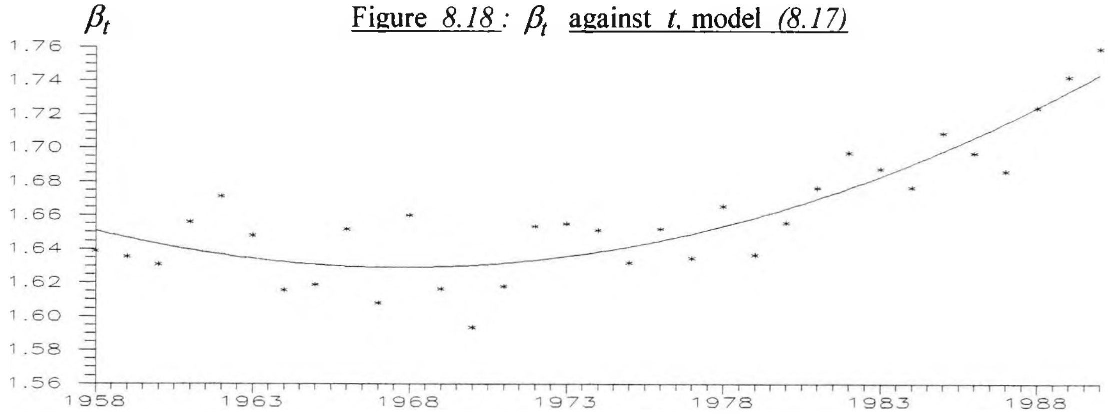


Figure 8.18 : $\hat{\beta}_t$ against t, model (8.17)



For the reasons described in Section 8.2.3 it was decided to replace both sets of parameters α_t & β_t by fractional polynomial of the type

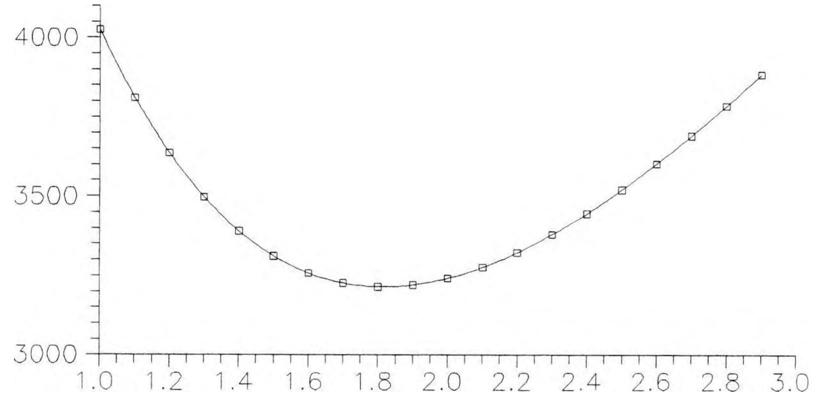
$$\alpha_1 + \alpha_2 \cdot t^k$$

Thus, the model structure

$$\log(\mu_{x,t}) = a + b \cdot t^k + c \cdot \sqrt{x} + d \cdot \sqrt{x} \cdot t^k \qquad \textbf{(8.19)}$$

was considered next. The deviance profile for various values of k (in steps of 0.1) is reproduced in Figure 8.15 which implies an optimum value of $k = 1.8$.

Figure 8.19 : Deviance profile for various values of k, model *(8.19)*



leading to the adoption of the model structure

$$\log(\mu_{x,t}) = a + b \cdot t^{1.8} + c \cdot \sqrt{x} + d \cdot \sqrt{x} \cdot t^{1.8} \qquad \textbf{(8.20)}$$

It is of interest to compare this structure with that of model *(8.13)*, which, for $x > 42$, can be written as

$$\log(\mu_{xt}) = A(x) + B(x) \cdot t^{1.8}$$

where both $A(x)$ and $B(x)$ are quadratic in x. Here, model *(8.20)* can be expressed in exactly the same general form, but where both $A(x)$ and $B(x)$ are linear in $\sqrt{x}$.

The associated parameter estimates, standard errors, and t - values are as shown in Table *8.13*.

Table 8.13 : Parameters estimates, standard errors, and t - values, model *(8.20)*

	p.e.	***s.e.***	***t - values***
a	-16.76	0.0311	-538.9
b	-0.002464	0.000125	-19.7
c	1.624	0.003931	413.1
d	0.000177	0.0000157	11.2
	$\hat{\varphi} = 2.034$		

The deviance for the model structure *(8.20)* is *3214.1* on *1580* degrees of freedom.

The p - values for the statistical test of a graduation using standardised deviance residuals are as shown on Table *8.14*.

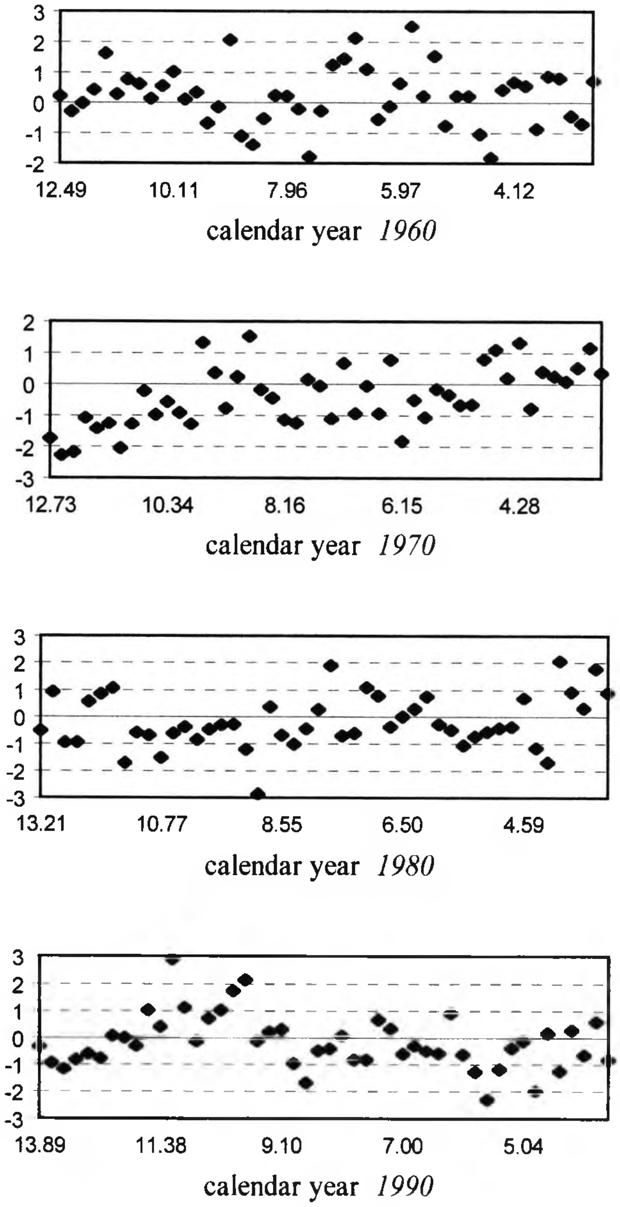
Table 8.14 : p - values, formal statistical tests for each calendar year separately, model (8.20)

Year	ISD	Sign	Run	Chi
1958	92	19	14	36
1959	88	50	7	40
1960	55	95	8	51
1961	96	50	50	45
1962	68	80	1	41
1963	29	7	62	28
1964	24	80	14	44
1965	94	61	19	38
1966	77	50	61	40
1967	0	100	92	57
1968	42	95	21	42
1969	13	28	8	25
1970	14	2	0	27
1971	13	95	8	55
1972	37	4	32	33
1973	92	19	0	38
1974	77	19	75	38
1975	67	87	34	47
1976	68	80	2	42
1977	46	95	68	42
1978	19	98	17	41
1979	59	92	6	46
1980	43	4	1	41
1981	60	71	29	43
1982	54	28	40	42
1983	33	7	93	39
1984	67	87	91	43
1985	68	80	14	35
1986	86	61	61	44
1987	74	50	99	45
1988	96	71	20	45
1989	22	99	4	48
1990	28	4	21	40

Considering the simplicity of the model used, the above table gives very satisfactory results.

The standardised deviance residuals plotted against the constant information scale ($CIS = 2 \cdot \log(1/\mu_{xt})$) for some of the calendar years are presented in Figure *8.20*.

Figure 8.20: Standardised deviance residuals vs. *CIS*, various calendar years, model *(8.20)*



Finally for this model, the predicted values of the force of mortality, μ_{xt}, in the age range $x =$ *42* to *89* years, over the calendar period $t = 1960$ to *1990* at *10* yearly intervals, and the forecast values of the force of mortality for the years *2000* and *2010* are presented for completeness in the Table *8.15*.

Table 8.15 : Predicted and forecasting force of mortality, *10* yearly intervals, quinquennial ages, model *(8.20)*

	1960	1970	1980	1990	2000	2010
45	0.002814	0.002496	0.00198	0.001423	0.000934	0.000562
50	0.005075	0.004529	0.003635	0.002656	0.00178	0.001098
55	0.008893	0.007982	0.006477	0.004809	0.003288	0.002079
60	0.015200	0.013716	0.011249	0.008479	0.005910	0.003824
65	0.025415	0.023055	0.019102	0.014608	0.010371	0.006862
70	0.041679	0.038001	0.031794	0.024656	0.017819	0.012045
75	0.067172	0.061545	0.051979	0.040855	0.030037	0.020730
80	0.106578	0.098111	0.083621	0.066584	0.049773	0.035045
85	0.166713	0.154173	0.132570	0.106899	0.081205	0.058296

Furthermore, the simplicity of the model $\mu_{x,t} = \mathrm{A}_t \cdot \mathrm{B}_t^{\sqrt{x}}$ used to graduate the available data for each calendar year separately, means that we are able to interpret the trend in mortality through an examination of the values of α_t and β_t. The parameter α_t indicates the *level* of mortality for year *t*, and the parameter β_t indicates the *growth*, or the rate of increase of mortality with age, for the year *t*. Figure *8.18* shows that the growth of mortality decreases between *1958* and *1970*, and subsequently is projected to increase in a quadratic manner. We conclude from Figures *8.17* & *8.18* that for the period *1958 - 1970*, the downward mortality trend favours the oldest ages. But for the years *1970 - 1990*, the nearly quadratic decrease in the level parameter and the nearly quadratic increase in the growth parameter shift the graduated curves downwards and bend the curves in favour of the middle ages.

CHAPTER IX

Power link models

9.1 Introduction

In this Chapter we focus on power link predictor relationships of the type

$$\eta_{xt} = m_{xt}^{p}$$

for some predetermined power $p \neq 0$, where η_{xt} denotes the parameterised linear predictor, and m_{xt} the expected response.

It includes the *identity* link when $p = 1$, and the *log* link when $p \to 0$.

The optimum value of p for a specific linear predictor structure is determined by repeatedly fitting the structure over a range of values of p and monitoring the resulting deviance profile.

Given a close approximation p_0 to the optimum value of p, determined by the above process, it is possible to determine a closer approximation to the optimum value of p using the method proposed by Pregibon (*1980*) (McCullagh and Nelder, *1989*, pages *375 - 376*).

The optimisation of the p - value or equivalently the minimisation of the deviance for different p - values is achieved through the approximation of the expansion of the link function in a Taylor series about a fixed value p_0. This approximation is achieved by keeping only the linear term. Thus, for the power family we have

$$g(m;p) = m^{p} \cong g(m;p_0) + (p - p_0) \cdot g'_p(m;p_0)$$

so that

$$g(m;p) \cong m^{p_0} + (p - p_0) \cdot m^{p_0} \cdot \log(m)$$

Thus we can approximate the correct link function $\eta = m^{p}$ by

$$\eta_0 = m^{p_0} = m^{p} - (p - p_0) \cdot m^{p_0} \cdot \log(m) = \sum \beta_i \cdot x_i - (p - p_0) \cdot m^{p_0} \cdot \log(m)$$

Given a first estimate p_0 of p we can fit the model by including an extra covariate $-m^{p_0} \cdot \log(m)$ in the linear predictor, whose parameter estimate measures p - p_0, the first order adjustment to p_0. To obtain the optimum value for p we have to repeat the above process forming a new adjusted value for p at each stage. Convergence is not guaranteed however and requires that the starting value p_0, is sufficiently close to p for the process to converge.

Using the power link function, we choose to target the resistivity to death for consistency throughout this Chapter, unless otherwise stated, in accordance with the distributional assumptions of Section *5.5*.

9.2 UK male assured lives, duration 5+, period 1958 - 1990, ages 24 - 89

9.2.1 Description of the data

The methods of Section *9.1* are applied to the *UK* male assured lives data set, for duration *5+*, period *1958* to *1990* and ages *24* to *89* years, both inclusive, as described in Section *8.2.1*. The data are presented in Appendix *A*, as published by the *CMI* Bureau of the Institute and Faculty of Actuaries.

Since these data have at least one reported death in each cell, as with log-link formulae, the targeting of the resistivity of death is identical to the targeting of the force of nortality, (subject to a change in sign in the estimated power index under the two approaches), in the case of power-link formulae (Renshaw et al, *1996b*).

9.2.2 *Modelling trends using polynomial predictor structures*

The number of terms needed in the polynomial predictor is determined by the shape of the crude mortality curve. For the data in question, the crude mortality rates are not in monotonic order over the whole of the age range (*24* to *89* years), so that a polynomial of degree higher than one is needed when formulating the linear predictor. A quadratic predictor in age effects has been found to be sufficient, for each calendar year in question, so that we start the analysis with the following model structure.

$$\mu_{xt}^{-p_t} = \alpha_t + \beta_t \cdot x + \gamma_t \cdot x^2 \qquad \textbf{(9.1)}$$

The fitting of this structure is equivalent of using a power link function and quadratic linear predictor in age effects to graduate the mortality experience for each calendar year separately.

The sum of deviances over all years, based on the fitting model structure *(9.1)*, is *3884.8* on *2046* degrees of freedom.

The results of the tests of a graduation based on standardised deviance residuals, for each calendar year, are reported in Table *9.1*.

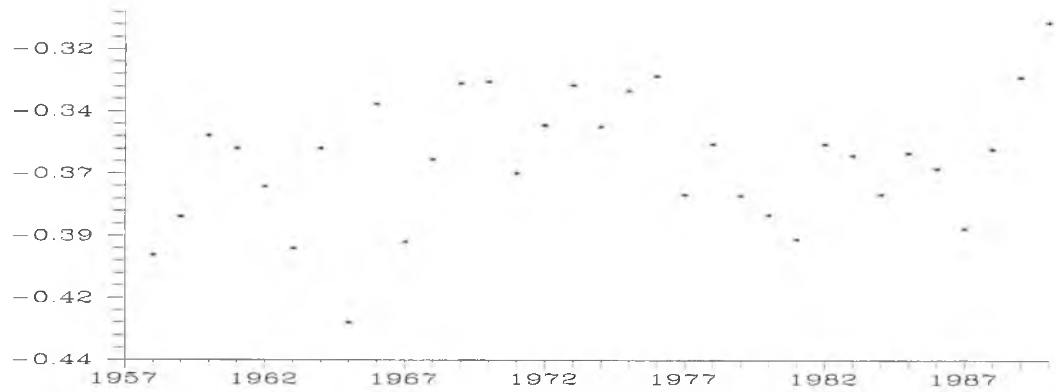
Table 9.1 : p - values, formal statistical tests for each calendar year separately for model (9.1)

Year	***ISD***	***Sign***	***Runs***	***Chi***
1958	65	89	47	34
1959	94	69	11	40
1960	97	69	16	9
1961	61	50	69	50
1962	96	69	51	18
1963	15	93	22	52
1964	25	40	16	82
1965	99	40	**1**	64
1966	22	69	96	22
1967	88	50	59	87
1968	69	77	33	50
1969	62	96	**1**	47
1970	22	31	**4**	33
1971	24	11	9	4
1972	76	50	**0**	94
1973	96	60	50	21
1974	8	16	35	50
1975	84	69	**4**	31
1976	86	69	16	53
1977	30	84	64	77
1978	63	89	**0**	17
1979	92	50	11	38
1980	95	50	7	34
1981	69	77	17	24
1982	95	50	30	19
1983	85	60	84	13
1984	90	77	79	28
1985	89	69	31	47
1986	63	60	84	42
1987	59	23	98	67
1988	89	60	23	55
1989	83	69	23	40
1990	74	50	50	91

The resulting p - values indicate the acceptance of the model used to carry out graduation for the data in question. Moreover, all tests do not show any trend through time, giving no preference, for the choice of the formula used, to any specific time period.

The optimum values of the power link parameter p_t have been obtained for each calendar year as described in Section *9.1*. These values are displayed in Figure 9.1.

Figure 9.1 : Estimated p_t against *t*, model *(9.1)*



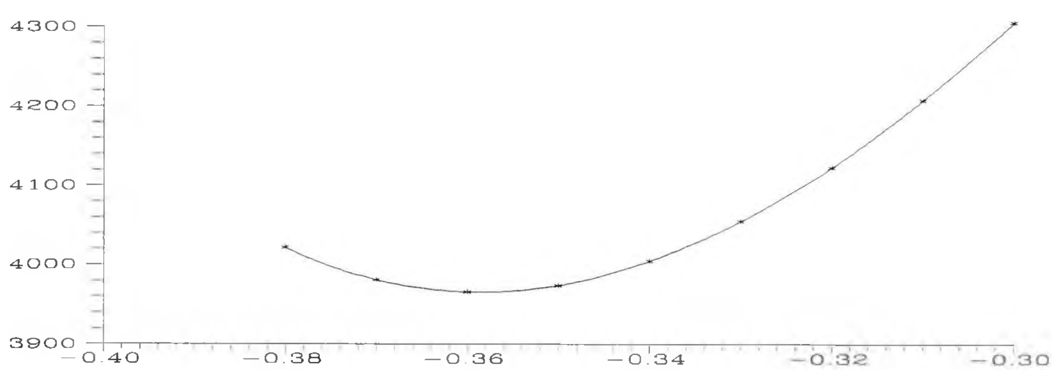
As implied, by Figure *9.1*, the estimated values of p_t are banded about the value *-0.36*, with no clear trend. As a consequence of this, and also because of the potential difficulty of modelling the power parameter as a function of *t*, it was decided to model p_t as a constant *p*.

The deviance profile, produced by fitting the model structure

$$\mu_{xt}^{-p} = \alpha_t + \beta_t \cdot x + \gamma_t \cdot x^2 \tag{9.2}$$

for various values of *p*, is reproduced in Figure *9.2*. This has an optimum value at $p = -0.36$.

Figure 9.2 : Deviance profile against *p*, model *(9.2)*

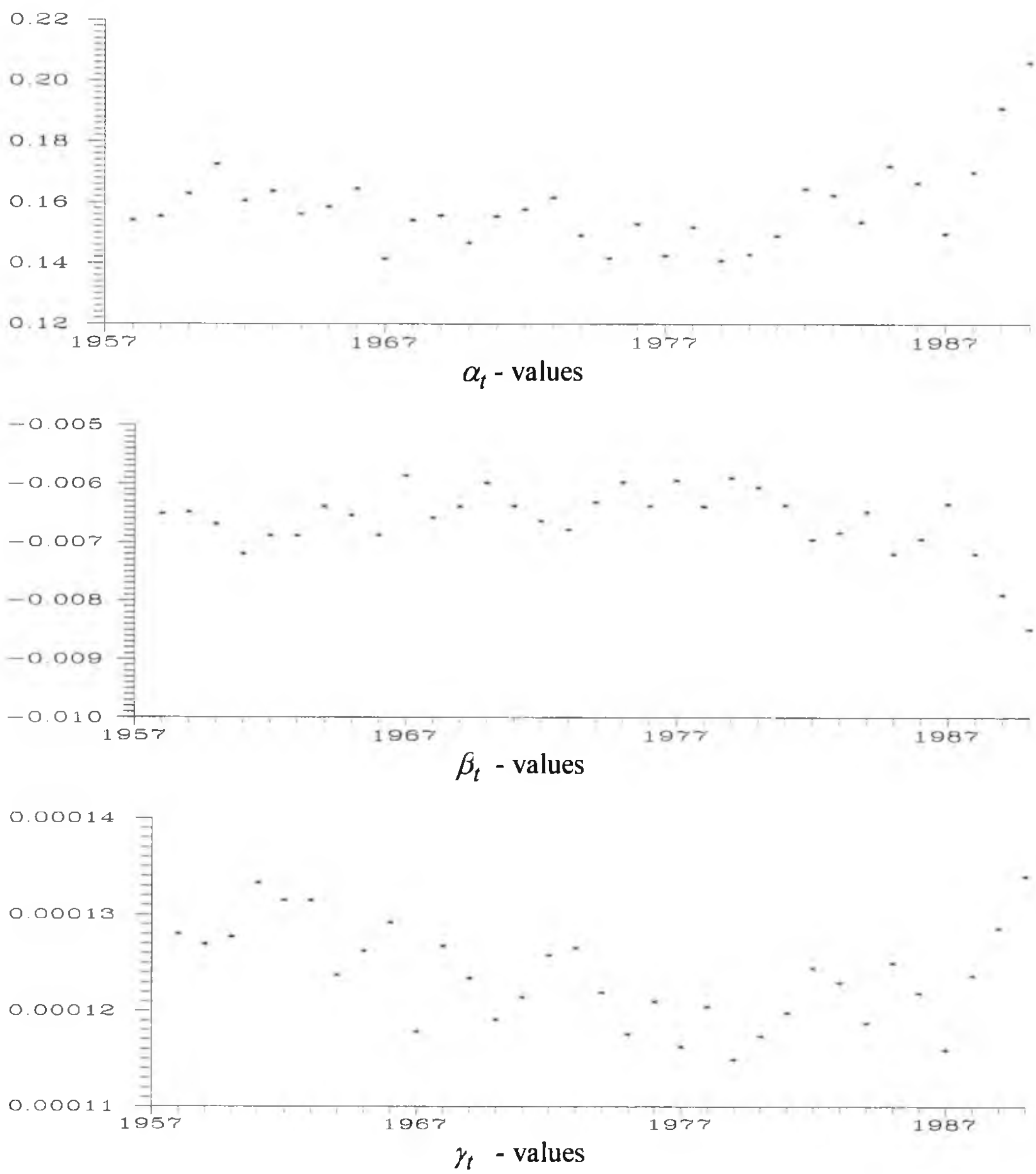


Again, the *p* - values for the statistical tests of a graduation based on each calendar year, for model *(9.2)* with p = *-0.36*, are highly supportive except for a few years where the runs test fails. This seems to be the result of the constant power parameter having somewhat less flexibility.

Now the analysis can proceed in two different ways. One possibility is to model all the parametric trends which are included in the model structure *(9.2)* through time, the other possibility is to set certain of the parameters equal to constants over time. The second approach will be discussed in Section *9.2.3*. Employing the first method we obtain the following results.

The trends in the three sets of estimated parameters for the model structure *(9.2)* with $p = -0.36$ are displayed in Figure *9.3*.

Figure 9.3 : Parameters estimates against time, model (9.2)



As a consequence, the replacement of the time dependent sets of parameters in equation *(9.2)* by quadratic polynomials in time effects was found to be very effective in reducing the excessive amount of parameterisation.

The parameter estimates, standard errors, and t - values obtained on fitting the model structure

$$\mu_{x,t}^{0.36} = (a_1 + a_2 \cdot t + a_3 \cdot t^2) + (b_1 + b_2 \cdot t + b_3 \cdot t^2) \cdot x + (c_1 + c_2 \cdot t + c_3 \cdot t^2) \cdot x^2 \quad \textbf{(9.3)}$$

where t = calendar year - *1957* (where calendar year = *1958, 1959, ...*) is given in Table *9.2*.

Table 9.2 : Parameter estimates, standard error, and t - values for model *(9.3)*

	p.e.	***s.e.***	***t - values***
a_1	0.174	0.00531	32.76
a_2	-0.003291	0.000688	-4.78
a_3	0.000154	0.00001958	7.86
b_1	-0.007233	0.000203	-35.63
b_2	0.0001322	0.00002665	4.96
b_3	-0.000004316	0.000000759	-5.68
c_1	0.0001343	0.000001907	70.42
c_2	-0.000001387	0.000000252	-5.5
c_3	0.00000003425	0.0000000072	4.75
	$\hat{\varphi} = 2.004$		

The deviance obtained is *4346* on *2169* degrees of freedom.

By way of comparison Table *9.3* contains the parameter estimates, their standard errors, and t-values, when fitting the same structure through targeting the force of mortality to death in accordance with the distribution assumptions of Section *5.2.* Again note the parameter estimates are identical under the two sets of modelling assumptions, but that the power p takes opposite signs, leading to identical graduations, see Renshaw et al (*1996b*). Note also that the corresponding standard errors differ by a factor of $\sqrt{\hat{\varphi}}$.

Table 9.3 : Parameter estimates, standard error, and t - values for model *(9.3)* based on Poisson error distribution

	p.e.	***s.e.***	***t - values***
a_1	*0.174*	*0.003801*	*45.8*
a_2	*-0.00329*	*0.0004919*	*-6.7*
a_3	*0.000154*	*0.000014*	*11.0*
b_1	*-0.007234*	*0.0001454*	*-49.8*
b_2	*0.0001321*	*0.00001905*	*6.9*
b_3	*-0.000004315*	*0.0000005426*	*-8.0*
c_1	*0.0001343*	*0.000001361*	*98.7*
c_2	*-0.000001387*	*0.0000001806*	*-7.7*
c_3	*0.00000003425*	*0.00000000515*	*6.7*

$\hat{\varphi} = 1$

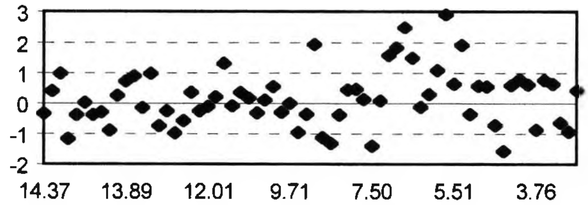
The p - values for the statistical tests of a graduation are presented next in Table *9.4,* and just some of the many standardised deviance residual plots on the constant information scale ($CIS = 2 \cdot \log(1 / \mu_{xt})$), for various calendar years, presented in Figure *9.4*.

Table 9.4 : p - values, formal graduation tests for each calendar year separately, model (9.3)

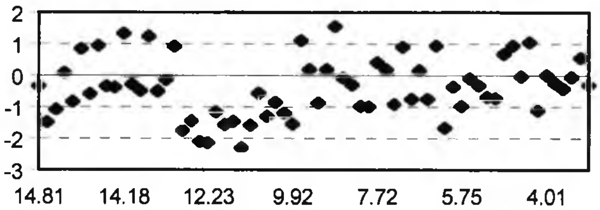
Year	ISD	Sign	Runs	Chi
1958	74	83	26	58
1959	99	40	4	50
1960	49	76	33	65
1961	97	68	31	52
1962	55	89	3	61
1963	12	1	5	31
1964	54	76	17	47
1965	3	16	5	37
1966	93	76	85	51
1967	0	99	35	65
1968	53	89	20	60
1969	3	16	35	25
1970	3	0	81	34
1971	44	89	1	56
1972	33	2	1	35
1973	77	31	1	44
1974	5	7	85	52
1975	74	83	5	64
1976	94	40	1	48
1977	32	95	65	54
1978	75	93	1	50
1979	63	93	15	56
1980	79	10	9	50
1981	89	76	11	53
1982	89	50	10	51
1983	94	23	85	43
1984	12	97	38	58
1985	60	50	30	46
1986	71	40	50	53
1987	66	77	98	69
1988	53	93	15	68
1989	46	83	12	53
1990	29	1	8	54

Although generally satisfactory it is noticeable that these graduations fail a number of the statistical tests for the period *1967 - 1973*.

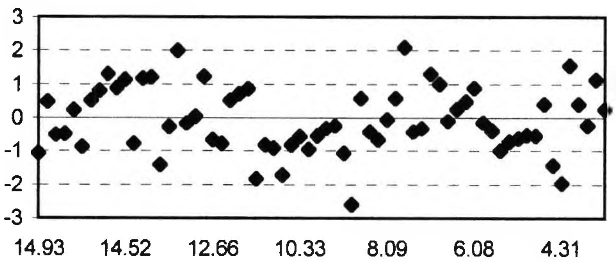
Figures 9.4 : Standardised deviance residuals vs. *CIS*, various calendar years, model *(9.3)*



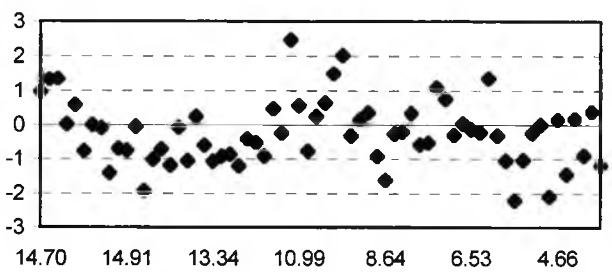
calendar year *1960*



calendar year *1970*



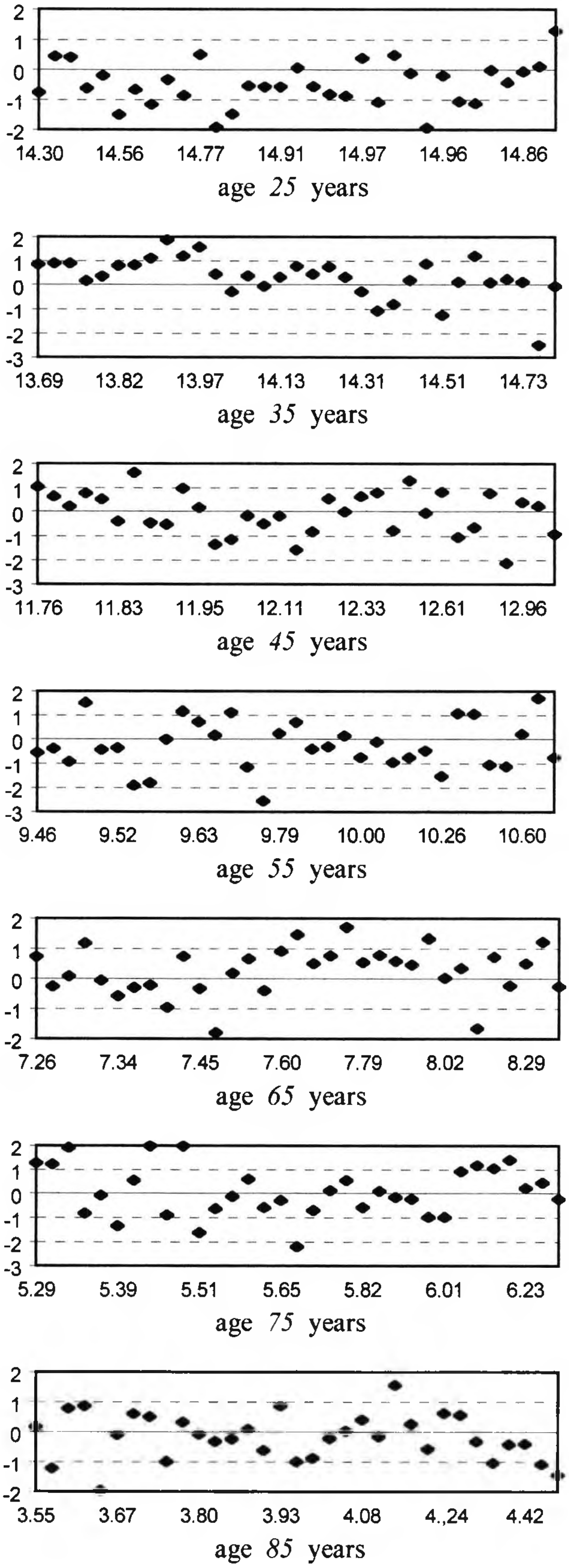
calendar year *1980*



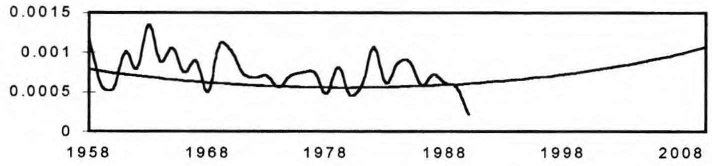
calendar year *1990*

We also reproduce the plots of the standardised deviance residuals against the constant information scale, at ten yearly age intervals, in the Figures *9.5*. The predicted force of mortality (for the time period *1958* to *2010*) is plotted against calendar year, at ten yearly age intervals, in Figure *9.6*.

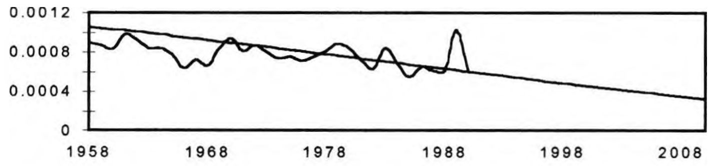
Figures 9.5 : Standardised deviance residuals vs. *CIS*, various ages, model *(9.3)*



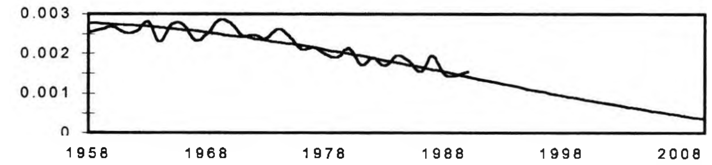
Figures 9.6 : Crude and predicted force of mortality vs. calendar year, various ages, model *(9.3)*



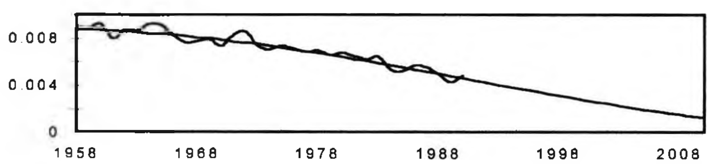
age *25* years



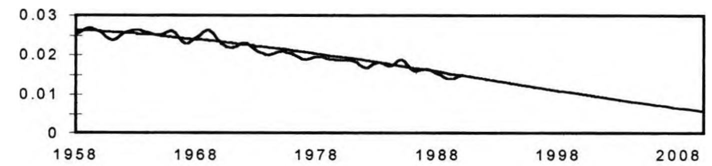
age *35* years



age *45* years

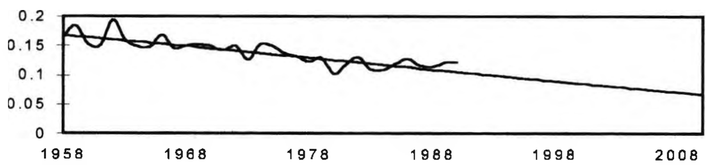


age *55* years



age *65* years

age *75* years



age *85* years

Here we have superimposed the estimated mortality curves on the corresponding crude mortality rates. This acts as a further visual check on the predictive qualities of the model.

Finally for this model, the predicted values of the force of mortality, μ_{xt}, in the age range x = *24* to *89* years, for the calendar period t = *1960* to *1990* at *10* yearly intervals, and forecast values for the years *2000* and *2010* are presented for completeness in Table *9.5*.

Table 9.5 : Predicted force of mortality, *10* yearly intervals, quinquennial ages, model *(9.3)*

	1960	1970	1980	1990	2000	2010
25	0.000743	0.000601	0.000560	0.000609	0.000763	0.001066
30	0.000785	0.000659	0.000578	0.000533	0.000521	0.000538
35	0.001036	0.000898	0.000752	0.000604	0.000461	0.000330
40	0.001613	0.001429	0.001165	0.000857	0.000549	0.000285
45	0.002759	0.002473	0.001998	0.001418	0.000837	0.000365
50	0.004905	0.004417	0.003575	0.002532	0.001486	0.000636
55	0.008744	0.007875	0.006421	0.004630	0.002814	0.001297
60	0.015324	0.013770	0.011341	0.008394	0.005377	0.002759
65	0.026155	0.023425	0.019496	0.014849	0.010065	0.005759
70	0.043331	0.038661	0.032508	0.025463	0.018220	0.011508
75	0.069656	0.061911	0.052556	0.042257	0.031764	0.021856
80	0.108792	0.096336	0.082493	0.067928	0.053351	0.039484
85	0.165403	0.145952	0.125966	0.105976	0.086517	0.068115

From the Table *9.5*, we can conclude that the force of mortality for the first ages (*25* & *30*) increases for the last years in question. Yet, comparing the $\hat{\mu}_{25,2010} = 0.001066$ value from the above Table, which is based on the model structure (*9.3*), with the $\hat{\mu}_{25,2010} = 0.00027$ value from Table *8.12*, which is based on the model structure (*8.13*), we can observe a large discrepancy between these two values. This discrepancy seems to be the result of the constant power parameter attached to the model structure (*9.3*) having somewhat less flexibility in association with the parsimonious number of parameters.

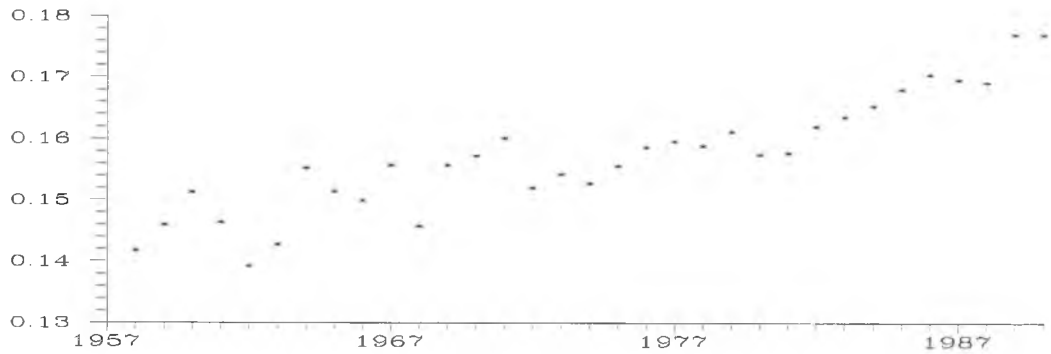
9.2.3 Modelling trends using fractional polynomial predictor structures in time effects and polynomial predictor structures in age effects

On the basis of the model structure *(9.2)* with p = *-0.36,* another effective way to reduce the excessive amount of parameterisation and produce a simple mathematical expression was found by suppressing the parameter γ_t and setting $\gamma_t = \gamma$ for all t, so that

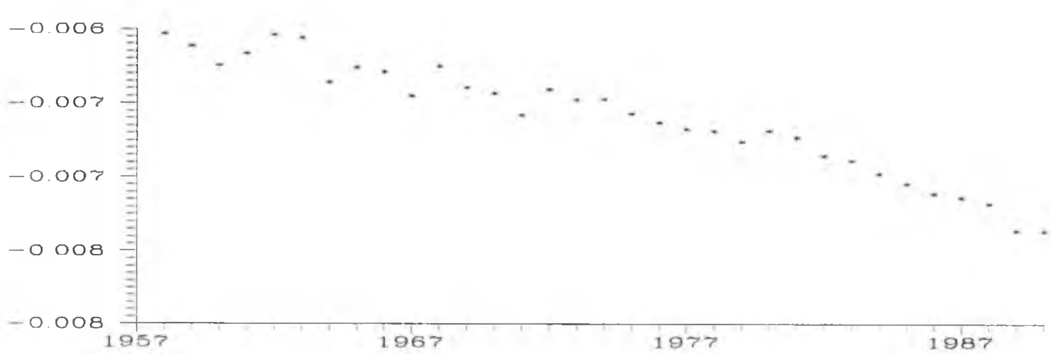
$$\mu_{x,t}^{0.36} = \alpha_t + \beta_t \cdot x + \gamma \cdot x^2 \qquad \textbf{(9.4)}$$

The trends in the other two sets of parameter estimates are displayed in Figure *9.7*.

Figure 9.7 : Parameters estimates against time, model *(9.4)*



α_t - values



β_t - values

Finally the number of parameters was further reduced by using fractional polynomials of the type

$$a + b \cdot t^k$$

to represent the variation in *both* α_t & β_t (the coefficient of correlation for the two parametric vectors has the value *-97.7%,* so that the vectors α_t & β_t are highly correlated and hence a similar formula was used to describe both parametric trends), where k is a suitable fixed index. The optimum value $k = 1.6$ was determined by looking at the appropriate deviance profile constructed by fitting the model structure for various values of k in the neighbourhood of the *1.6.*

The parameter estimates, their standard errors and t - statistics for the model structure

$$\mu_{x,t}^{0.36} = (a_1 + a_2 \cdot t^{1.6}) + (b_1 + b_2 \cdot t^{1.6}) \cdot x + c_1 \cdot x^2 \qquad \textbf{(9.5)}$$

are given in Table *9.6.*

Table 9.6 : Parameter estimates, standard error, and t - values, model *(9.5)*

	p.e.	***s.e.***	***t - values***
a_1	0.146	0.001652	88.3
a_2	0.0001028	0.000006131	16.7
b_1	-0.006113	0.00006068	-100.7
b_2	-0.000004452	0.0000001176	-37.8
c_1	0.0001235	0.00000057	216.6
	$\hat{\varphi} = 2.034$		

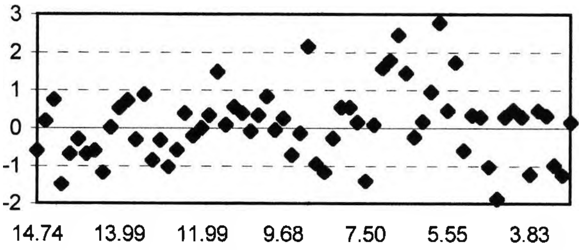
The deviance for the model structure is *4421* on *2173* degrees of freedom

The p - values for the statistical tests of a graduation are presented in Table *9.7,* and some of the many standardised deviance residual plots on the constant information scale (CIS = $2 \cdot \log(1 / \mu_{xt})$), for various calendar years, are presented in Figure *9.8.*

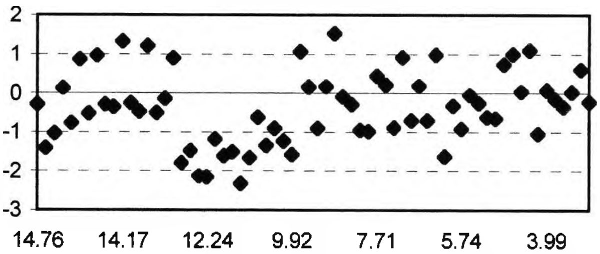
Table 9.7 : p - values, formal graduation tests for each calendar year separately, model *(9.5)*

Year	***ISD***	***Sign***	***Runs***	***Chi***
1958	*37*	*10*	***1***	*44*
1959	*65*	*7*	*15*	*41*
1960	*26*	*83*	*18*	*57*
1961	*91*	*77*	*33*	*46*
1962	*91*	*83*	*8*	*55*
1963	***3***	***0***	*35*	*28*
1964	*64*	*89*	*9*	*46*
1965	***4***	*7*	***2***	*35*
1966	*97*	*77*	*85*	*51*
1967	***0***	***100***	*33*	*65*
1968	*46*	*93*	*22*	*61*
1969	*5*	*23*	*17*	*26*
1970	*14*	***0***	*57*	*35*
1971	*43*	*89*	***1***	*58*
1972	*25*	*7*	***0***	*38*
1973	*78*	*23*	***0***	*46*
1974	*8*	*16*	*44*	*54*
1975	*79*	*83*	*5*	*64*
1976	*99*	*31*	***1***	*49*
1977	*29*	*89*	*28*	*53*
1978	*46*	***98***	***4***	*50*
1979	*23*	*95*	*7*	*56*
1980	*65*	*7*	*10*	*51*
1981	*82*	*89*	*6*	*53*
1982	*92*	*68*	*16*	*53*
1983	*86*	*10*	*95*	*44*
1984	*46*	*93*	*50*	*59*
1985	*72*	*59*	*31*	*47*
1986	*71*	*40*	*50*	*53*
1987	*66*	*77*	*99*	*69*
1988	*52*	*93*	*15*	*67*
1989	*28*	*93*	*6*	*52*
1990	*15*	***0***	*25*	*52*

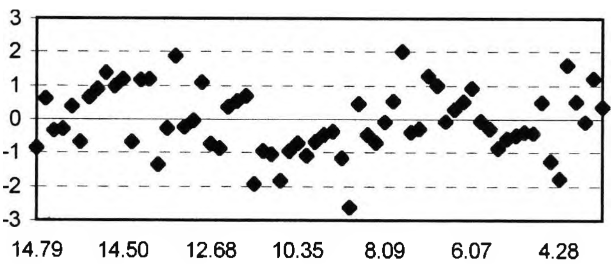
Figures 9.8 : Standardised deviance residuals vs. *CIS*, various calendar years, model *(9.5)*



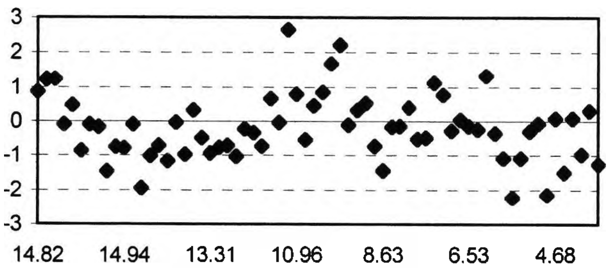
calendar year *1960*



calendar year *1970*



calendar year *1980*



calendar year *1990*

9.2.4 Analysis of age specific mortality trends

In a similar manner to Section *8.2.4*, we take a different perspective of the model by rearranging the linear predictor in expression *(9.5)* as follows

$$\mu_{x,t}^{0.36} = A(x) + B(x)\cdot t^{1.6} \qquad \textbf{(9.6)}$$

where

$$A(x) = (a_1 + b_1\cdot x + c_1\cdot x^2) \qquad \& \qquad B(x) = (a_2 + b_2\cdot x)$$

and the values of the parameters are as quoted in Table *9.5*. Thus the power of the force of mortality (index *0.36*) is represented by a fractional polynomial in time effects with age dependent coefficients.

As a further check on the model structure *(9.6)* it was decided to fit the following model structure

$$\mu_{x,t}^{0.36} = A_x + B_x\cdot t^{1.6} \qquad \textbf{(9.7)}$$

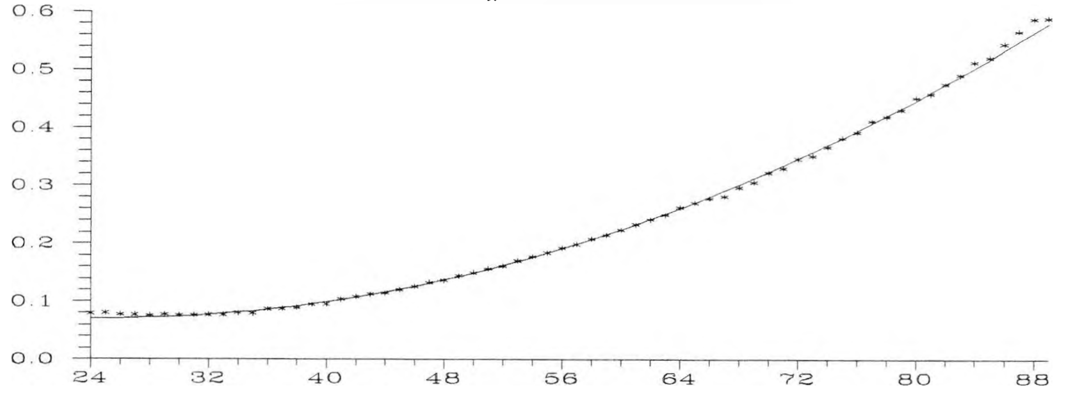
treating age as a factor. Not suprisingly this heavily parameterised structure was found to fit the data well. As evidence of this p - values for the dual statistical tests based on each age (rather than period) are presented in Table *9.8*.

Table 9.8 : p - values, formal graduation tests for each calendar year separately, model *(9.7)*

Age	***ISD***	***Sign***	***Runs***	***Chi***
24	16	97	70	42
25	100	57	81	65
26	63	57	19	74
27	99	30	72	59
28	100	43	30	48
29	29	30	72	75
30	81	57	43	43
31	99	30	20	52
32	96	70	59	37
33	98	57	43	46
34	84	19	98	59
35	92	70	11	43
36	58	70	45	55
37	94	30	59	50
38	96	43	97	58
39	83	30	31	51
40	94	57	**3**	43
41	96	70	59	39
42	50	43	57	44
43	97	57	30	47
44	28	19	92	51
45	98	57	30	45
46	85	43	**3**	38
47	98	57	**3**	49
48	96	70	**1**	45
49	53	19	48	50
50	64	43	70	49
51	12	70	11	52
52	78	57	70	51
53	78	43	30	48
54	78	81	13	45
55	96	43	19	49
56	99	43	30	47
57	34	89	68	45
58	96	43	11	47
59	94	30	59	49
60	78	19	22	46
61	94	43	11	49
62	86	70	59	44
63	96	57	**1**	47
64	90	57	**3**	48
65	85	70	11	44
66	55	89	54	41
67	96	70	59	47
68	98	30	90	52
69	78	70	95	49
70	99	43	19	49
71	84	57	95	46
72	54	30	44	44
73	97	57	70	49
74	72	19	7	51
75	99	43	10	49
76	72	19	96	52
77	22	19	22	52
78	97	57	70	48
79	74	57	70	52
80	79	81	62	46
81	98	57	89	45
82	97	70	11	48
83	93	70	45	48
84	82	70	**1**	49
85	82	70	20	47
86	79	43	43	44
87	78	43	97	49
88	36	19	75	50
89	26	89	67	40

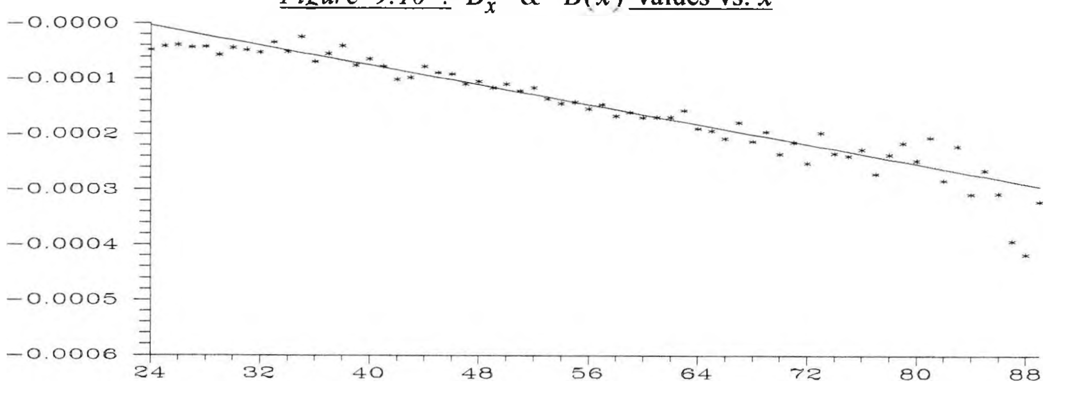
It is informative to plot the two sets of parameter estimates $\hat{A}_x$ & $\hat{B}_x$ for model *(9.7)* against the respective curves $\hat{A}(x)$ & $\hat{B}(x)$ defined above for model *(9.6)*. This is done in Figures *9.9* & *9.10* respectively. Both Figures are supportive of the choice of model *(9.6)*. Note the different scale used for Figures *9.9* & *9.10*.

Figure 9.9 : $\hat{A}_x$ & $\hat{A}(x)$ values vs. x



We also note that $\hat{A}_x$ & $\hat{A}(x)$ are similar in shape to 'crude' and 'graduated' mortality curves respectively, on the power transformation scale (index *0.36*), at time $t = 1$ (year *1958*).

Figure 9.10 : $\hat{B}_x$ & $\hat{B}(x)$ values vs. x



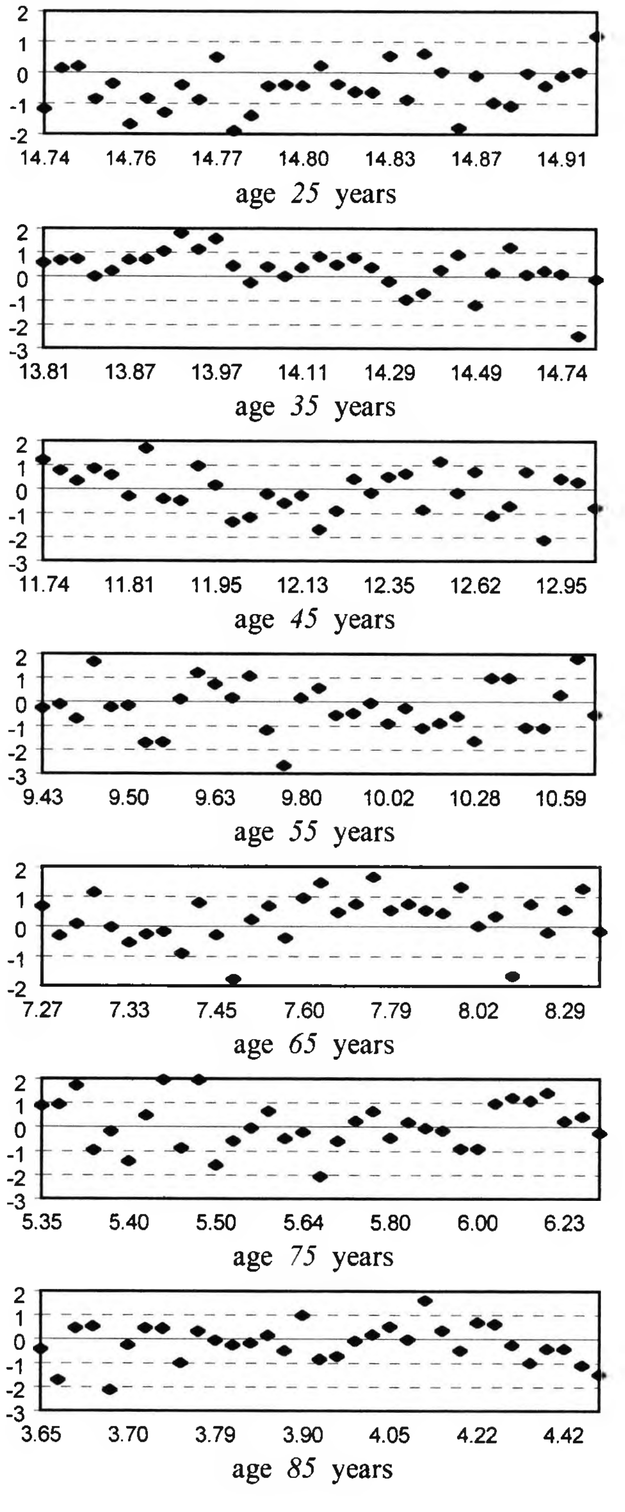
The B_x values represent the pace of mortality improvement in time, on the power (*0.36*) scale, for each age x. Lower values denote faster improvement. So, Figure *9.10* indicates that mortality improvement at the oldest ages is higher than for the youngest ages, on the power (*0.36*) scale. The different degree of closeness revealed on the two sets of graphs, is due to the different scale presented, while the p-values for the formal statistical tests reproduced in Table *9.9*, are generally supportive of the model.

Table 9.9 : p - values, formal graduation tests for each age separetely, model *(9.6)*

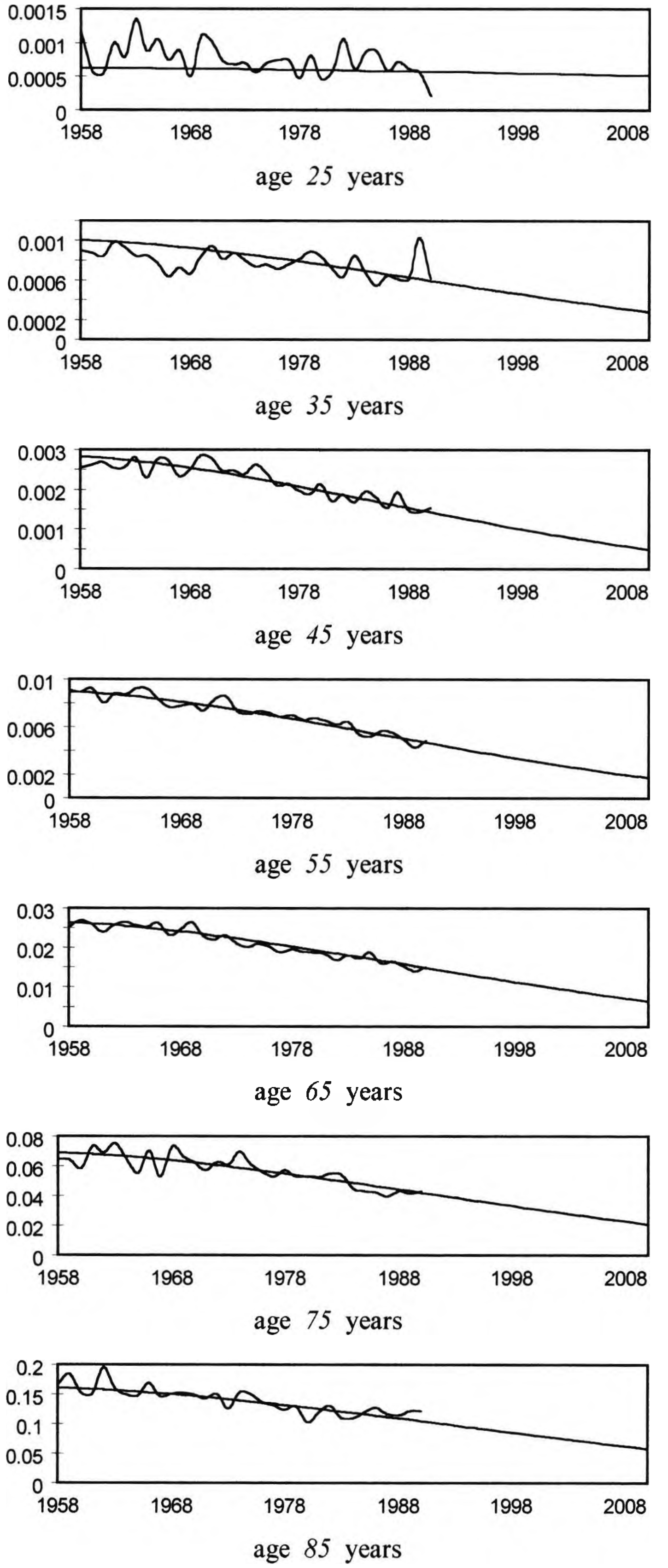
Age	***ISD***	***Sign***	***Runs***	***Chi***
24	71	19	34	18
25	**4**	**1**	17	18
26	40	3	25	38
27	61	6	31	33
28	74	19	**3**	38
29	29	6	96	72
30	99	56	19	34
31	17	30	20	58
32	57	94	5	57
33	60	88	53	58
34	17	88	53	71
35	**1**	**100**	**2**	51
36	76	88	53	63
37	60	88	67	62
38	95	43	81	54
39	60	69	**3**	64
40	**3**	**99**	29	64
41	95	43	81	36
42	10	**99**	31	52
43	49	88	38	51
44	64	43	81	55
45	93	43	57	43
46	36	6	**1**	38
47	30	19	**1**	42
48	89	69	**0**	46
49	23	6	31	43
50	21	6	45	40
51	16	6	74	45
52	89	19	34	47
53	41	5	19	43
54	46	56	30	40
55	69	6	**2**	45
56	57	19	62	43
57	86	56	57	43
58	57	80	62	50
59	66	69	31	52
60	93	69	44	48
61	64	19	12	48
62	80	19	85	43
63	93	43	10	45
64	62	19	7	46
65	62	88	38	49
66	**1**	**100**	29	56
67	**0**	**100**	80	63
68	**4**	**99**	92	64
69	**4**	**99**	92	60
70	9	88	53	58
71	18	**98**	67	58
72	**3**	**98**	20	54
73	33	94	31	59
74	58	94	45	59
75	76	43	10	51
76	88	69	59	55
77	71	19	22	51
78	95	43	70	48
79	68	69	90	50
80	6	3	25	37
81	82	11	67	39
82	87	80	12	51
83	36	19	22	37
84	50	3	25	40
85	46	11	53	40
86	20	**1**	10	33
87	95	30	72	39
88	21	11	38	35
89	67	30	82	32

Additional diagnostic evidence for model *(9.6)* is provided by the plots of standardised deviance residuals against the constant information scale $(CIS = 2 \cdot \log(1/\mu_{tx}))$, at ten yearly age intervals, in Figure *9.11*. The predicted force of mortality (for the time period *1958* to *2010*), against calendar year at ten yearly age intervals, is shown in Figure *9.12*.

<u>*Figures 9.11 :* Standardised deviance residuals vs. *CIS*, various ages, model *(9.6)*</u>



Figures 9.12 : Crude and predicted force of mortality vs. calendar year, various ages, model *(9.6)*



Here we have superimposed the estimated mortality curves on the corresponding crude mortality rates. This acts as a further visual check on the predictive qualities of the model.

Finally for this model, the predicted values of the force of mortality, μ_{xt}, in the age range $x = 24$ to 89 years, over the calendar period $t = 1960$ to 1990 at 10 yearly intervals, and the forecast values for the years 2000 and 2010 are presented for completeness in Table *9.10.*

Table 9.10 : Predicted force of mortality, *10* yearly intervals, quinquennial ages, model *(9.5)*

	1960	1970	1980	1990	2000	2010
25	0.000628	0.000616	0.000598	0.000574	0.000547	0.000516
30	0.000712	0.000668	0.000599	0.000515	0.000426	0.000336
35	0.000995	0.000902	0.000760	0.000597	0.000434	0.000285
40	0.001606	0.001426	0.001159	0.000862	0.000576	0.000332
45	0.002795	0.002464	0.001975	0.001437	0.000927	0.000505
50	0.004992	0.004403	0.003532	0.002572	0.001663	0.000909
55	0.008872	0.007861	0.006360	0.004696	0.003101	0.001755
60	0.015443	0.013772	0.011278	0.008484	0.005768	0.003420
65	0.026139	0.023476	0.019475	0.014947	0.010476	0.006517
70	0.042925	0.038822	0.032619	0.025527	0.018417	0.011977
75	0.068416	0.062284	0.052960	0.042200	0.031261	0.021146
80	0.105994	0.097081	0.083455	0.067595	0.051266	0.035893
85	0.159948	0.147308	0.127892	0.105115	0.081403	0.058721

It is of interest to investigate certain properties of model structure *(9.6)* over time. We first note that, under this model structure, the force of mortality does not increase with time, since $B(x)$ is always negative (Figure *9.10*). Further, in line with the analysis in Section *8.2.4*, it can be seen from Figure *9.12* that the predicted mortality curves change curvature during the time period involved. This feature indicates that the rate of the mortality decline through time reaches a maximum, in that time period, and afterwards diminishes. In mathematical terms, this turning point can be viewed as the time point where the second derivative, with respect to time, equals zero. That is, when

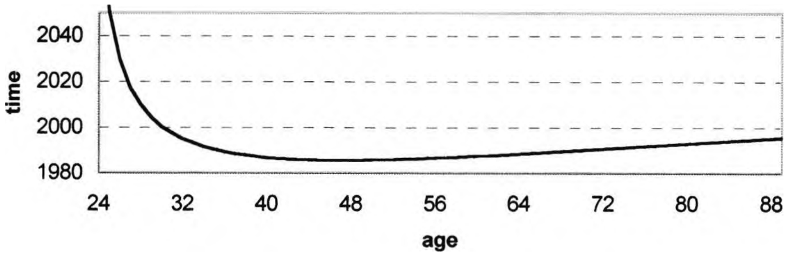
$$\frac{\partial^2}{\partial t^2}\mu_{xt} = 0$$

which leads, after some algebrical manipulations of formula *(9.6)*, to the point

$$t_x = \left\{-5.7 \cdot \frac{B(x)}{A(x)}\right\}^{-\frac{1}{1.6}}$$

The points t_x are plotted against x in Figure *9.13*.

Figures 9.13: Time-points where the second derivative equals zero, with respect to time, model *(9.6)*



It is possible to conclude from Figure *9.13* that the rate of the mortality decrease reaches its maximum during the decade of the *1980*'s for ages in the neighbourhood of *47* (more specifically for the range of ages *35* to *70*), and during the decade of *1990*'s for the ages above the age of *70*. That means that the maximum rate of improvement in mortality rates for ages above the age of *70* is expected during the *1990*'s, and for ages over *85* the maximum rate of improvement is expected towards the end of this decade.

9.3 UK male assured lives, duration 5+, period 1958 - 1990, ages 42 - 89

The power link function in association with a polynomial predictor of degree one in age effects, is used first to graduate the mortality experience for each calendar year separately. That is

$$\mu_{x,t}^{-p_t} = \alpha_t + \beta_t \cdot x \tag{9.8}$$

Further, for a fixed value of t, model *(9.8)* can be expressed as

$$\frac{d}{dx}\left(\frac{1}{\mu_x}\right) = -\left(\frac{1}{A - p \cdot x}\right) \cdot \left(\frac{1}{\mu_x}\right)$$

where

$$A = -\frac{\alpha}{\beta} \cdot p$$

Thus, when (α and β have the same sign and) p is negative, the rate at which the resistivity to death decreases with age, divided by the resistivity itself, is inversely proportional to a linear function of the age. Note, that this structure is a special case of the generalised binomial law with $A = 0$, see Gavrilov and Gavrilova (*1991*).

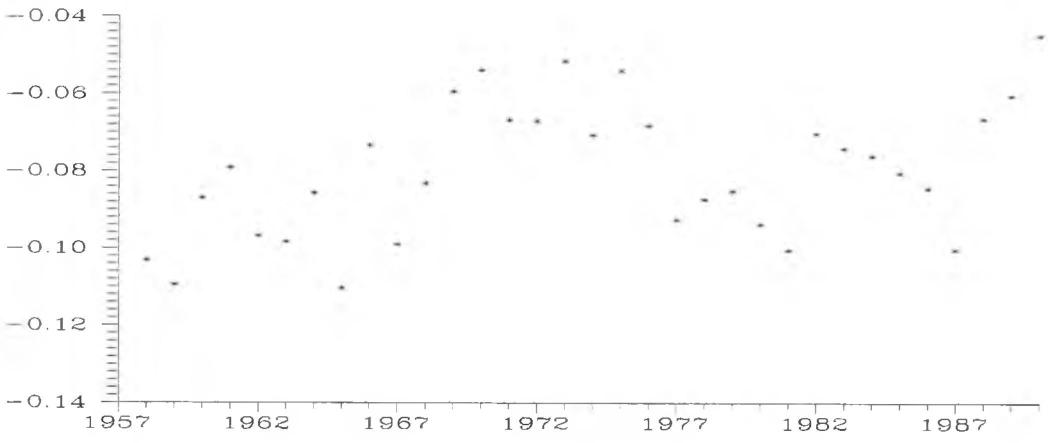
Under model structure *(9.8)*, the force of mortality automatically changes monotonically with respect to age x, for fixed t. This has been found to be consistent with the data in the restricted age range *42* to *89* years.

We again choose to target the resistivity or reciprocal of the force of mortality in accordance with the distributional assumptions of Section *5.5*.

The value of the deviance, when the model structure *(9.8)* is fitted to the data, is equal to *2754.2* on *1485* degrees of freedom. Standardised deviance residuals were then used to produce p - values for the statistical tests, which where found to justify the choice of the model structure (*9.8*).

The trend in the values of the optimum power index p_t, is illustrated in Figure *9.14*.

Figure 9.14 : Optimum p_t values against time, model *(9.8)*



The other sets of parameters $\{\alpha_t\}$ and $\{\beta_t\}$ have similar trends, confirmed by their high empirical coefficients of correlation.

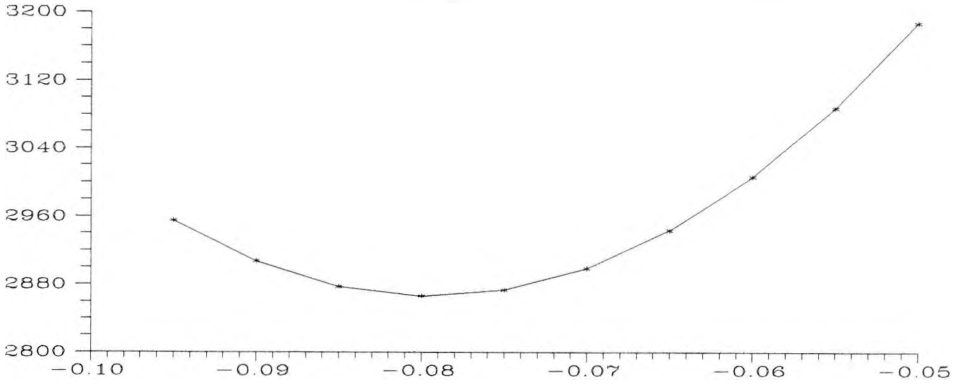
$$\hat{\rho}_{p,\alpha} = 0.9915 \qquad \hat{\rho}_{p,\beta} = -0.992 \qquad \hat{\rho}_{\beta,\alpha} = -0.9948$$

As noted previously (Section *9.2.2*) it is difficult to model p_t as a time dependent variable. Noting (Figure *9.14)* that the estimated values of p_t are banded about the value *-0.08* with no clear trend, we turn next to the model structure with constant power index

$$\mu_{x,t}^{-p} = \alpha_t + \beta_t \cdot x \qquad \textbf{(9.9)}$$

The optimum value of p is again determined by constructing the deviance profile which is displayed in Figure *9.15*. This has an optimum value at $p = -0.08$.

Figure 9.15 : Deviance profile against p, model *(9.9)*

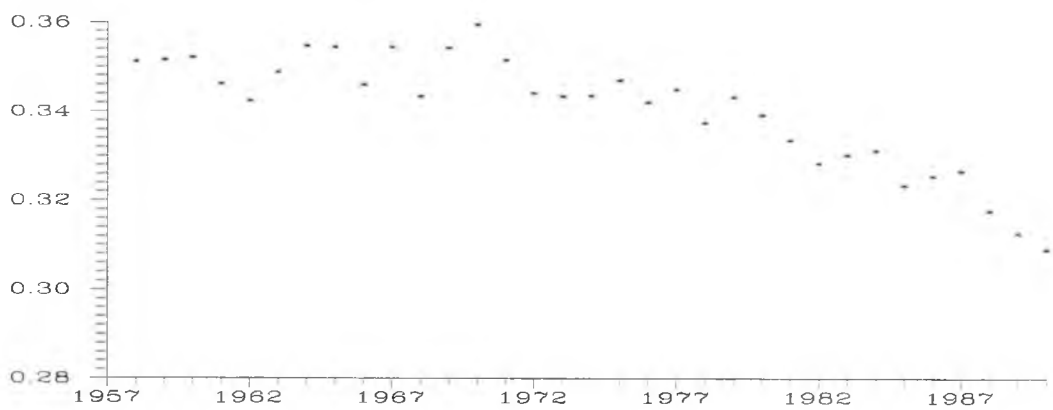


Thus the new model structure is as follows

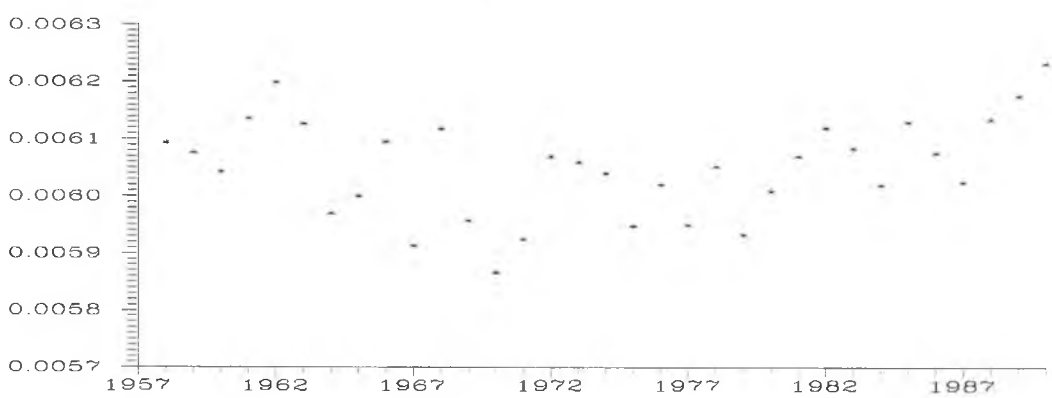
$$\mu_{x,t}^{0.08} = \alpha_t + \beta_t \cdot x \qquad \textbf{(9.10)}$$

The values of the deviance obtained was *2865.8* on *1517* degrees of freedom (an average increase in deviance of *3.5* compared with model *(9.8)*, for each calendar year). The parameter trends, for model *(9.10)*, are shown in the Figure *9.16*.

Figure 9.16 : Parameter estimates against time, model *(9.10)*



α_t - values



β_t - values

Due to the simplicity of the model structure we are able to interpret the mortality experience through the values of α_t and β_t, on the power(*0.08*) scale. Thus α_t indicates a *level* of mortality for the year t, and β_t represents the rate of increase of mortality with age, for the year t, on the power(*0.08*) scale.

Figure *9.16* shows that the parameter β_t decreased for the period *1958* to roughly *1970*, and subsequently increases. Again, as in Chapter *VIII*, we can interpret the above two figures together as implying that the mortality experience for the period *1958 - 1970* has been in favour of the oldest ages, that is there has been a faster mortality improvement for the oldest ages. But for the next period (*1970 - 1990*), the nearly quadratic decrease in the level parameter, α_t, and the nearly quadratic increase in the growth parameter, β_t, shifts the graduated curve downwards and bends the curve, at the same time, to favour the middle ages, so that there is faster mortality improvement for the middle ages.

Despite the apparently smooth progression in both sets of parameter estimates $\{\alpha_t\}$ and $\{\beta_t\}$, we attempt to simplify the model structure further by making one of the parameter sets constant over time. Thus for the model structure

$$\mu_{x,t}^{0.08} = \alpha + \beta_t \cdot x \qquad \textbf{(9.11)}$$

the deviance obtained was *3760.2* on *1549* degrees of freedom, and for the model structure

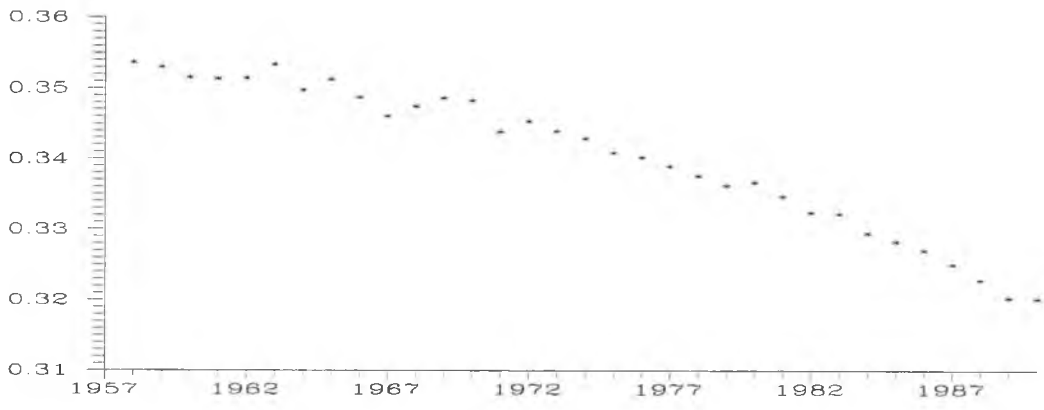
$$\mu_{x,t}^{0.08} = \alpha_t + \beta \cdot x \qquad \textbf{(9.12)}$$

the deviance obtained was *3023.5* on *1549* degrees of freedom (an average increase in deviance of *4.9* compared with model *(9.10)*, for each calendar year).

Obviously model *(9.12)* is more efficient in comparison with model *(9.11)*, due to the lower deviance. This is supported by Figure *9.16* on fitting a horizontal line to each set of plotted points.

The parametric trend of α_t for model *(9.12)* is shown in the Figure *9.17*

Figure 9.17 : α_t values against time, model *(9.12)*



This suggests that a *fractional polynomial* of the type $a + b \cdot t^k$ can be used to represent the time variation in α_t. Experiments, as before, establish an optimum k value of *1.8*. Thus the final model structure employed to analyse the mortality trends, takes the form

$$\mu_{x,t}^{0.08} = \alpha + \gamma \cdot t^{1.8} + \beta \cdot x \qquad \textbf{(9.13)}$$

The associated parameter estimates, standard errors, and t - values are as shown in Table *9.11*.

Table 9.11 : Parameters estimates, standard errors, and t - values, model *(9.13)*

	p.e.	***s.e.***	***t - values***
α	-0.3527	0.0006045	-583.4
β	-0.006052	0.00000963	-628.4
γ	0.00006094	0.000000624	97.6
	$\hat{\varphi} = 2.039$		

The deviance for the model is *3222.1* on *1580* degrees of freedom.

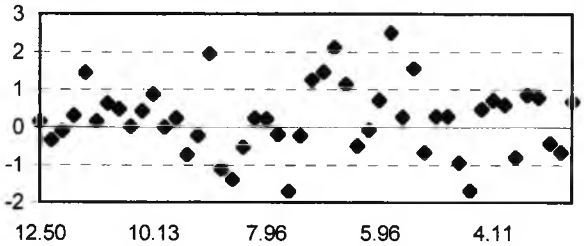
The p - values for the statistical tests of a graduation, using standardised deviance residuals, are as shown in Table *9.12*.

Table 9.12 : p - values, formal graduation tests for each calendar year separately, model *(9.13)*

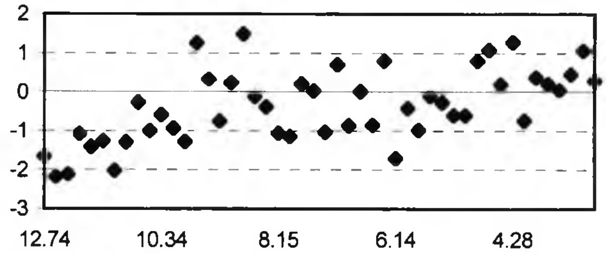
Year	***ISD***	***Sign***	***Runs***	***Chi***
1958	68	7	62	44
1959	93	28	7	48
1960	51	92	18	58
1961	94	61	72	52
1962	35	87	5	49
1963	12	**2**	50	35
1964	26	80	14	52
1965	94	61	7	44
1966	44	61	81	48
1967	**0**	**100**	92	65
1968	69	92	18	50
1969	16	38	**2**	34
1970	34	7	**1**	36
1971	**0**	95	8	62
1972	6	**0**	18	41
1973	70	12	**1**	45
1974	89	19	75	46
1975	85	80	31	54
1976	**4**	87	**3**	50
1977	46	95	67	49
1978	19	**98**	5	49
1979	61	87	5	53
1980	41	**4**	**1**	48
1981	58	71	29	51
1982	54	28	40	49
1983	28	7	96	46
1984	51	92	93	50
1985	89	80	**4**	41
1986	80	50	80	51
1987	82	50	99	53
1988	98	61	7	51
1989	15	95	**0**	53
1990	24	**2**	**3**	45

The standardised deviance residuals against the constant information scale ($CIS = 2 \cdot \log(1/\mu_{xt})$), for various calendar years, are presented in Figure *9.18*.

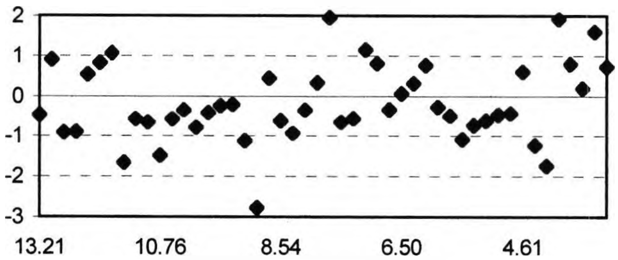
Figure 9.18 : Standardised deviance residuals vs. *CIS*, various calendar years, model *(9.13)*



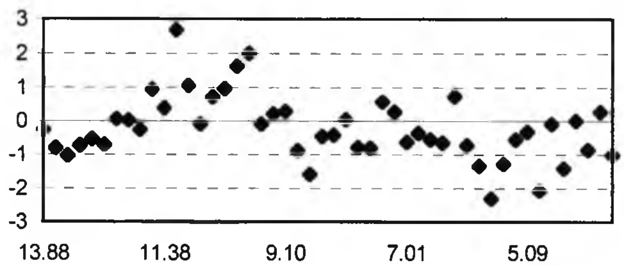
calendar year *1960*



calendar year *1970*



calendar year *1980*



calendar year *1990*

Finally for this model, the predicted values of the force of mortality, μ_{xt}, for ages $x = 45, 50, \ldots, 85$, over the calendar period $t = 1960$ to 1990 at 10 yearly intervals, and forecast values for the years 2000 and 2010 are presented for completeness in the Table *9.13*.

Table 9.13 : Predicted force of mortality, *10* yearly intervals, quinquennial ages, model *(9.13)*

	1960	1970	1980	1990	2000	2010
45	0.002785	0.002482	0.001982	0.001427	0.000926	0.000539
50	0.005031	0.004508	0.003637	0.002661	0.001764	0.001055
55	0.008850	0.007968	0.006491	0.004817	0.003256	0.001997
60	0.015190	0.013738	0.01129	0.008489	0.005840	0.003664
65	0.025496	0.023154	0.019183	0.014596	0.010205	0.006536
70	0.041921	0.038214	0.031897	0.024539	0.017412	0.011365
75	0.067628	0.061863	0.051993	0.040409	0.029065	0.019304
80	0.107195	0.098375	0.083208	0.065279	0.047547	0.032090
85	0.167150	0.153859	0.130912	0.103604	0.076344	0.052296

It is of interest to note that the estimated power link index $p = -0.08$ in these models is close to zero, which implies a log link. Comparison of Table *9.13* with Table *8.15* which is based on a log - link formula, indicates that the predicted values from the two models are comparable in size.

CHAPTER X

Additive models

10.1 Introduction

Throughout this Chapter we again target the resistivity to death using the modelling assumptions of Section *5.5*. Used in combination with the canonical reciprocal link function

$$\eta_{xt} = 1 \ / \ m_{xt}$$

it implies the fitting of mathematical formulae of the type

$$\mu_{xt} = \eta_{xt}$$

It is of interest to note that for a fixed observation period so that the suffix t is suppressed, this includes the De Moivre mortality law of *1725* as a special case (see Benjamin and Pollard, *1980*). More generally the linear predictor gives rise to additive structures in age and period effects. We view such structures as experimental.

10.2 UK male assured lives, duration 5+, period 1958 - 1990, ages 24 - 89

10.2.1 Description of the data

The methods of Section *10.1* are applied to the *UK* male assured lives data set, for duration *5+*, period *1958* to *1990* and ages *24* to *89* years, both inclusive, as described in Section *8.2.1*. The data are presented in Appendix *A*, as published by the *CMI* Bureau of the Institute and Faculty of Actuaries.

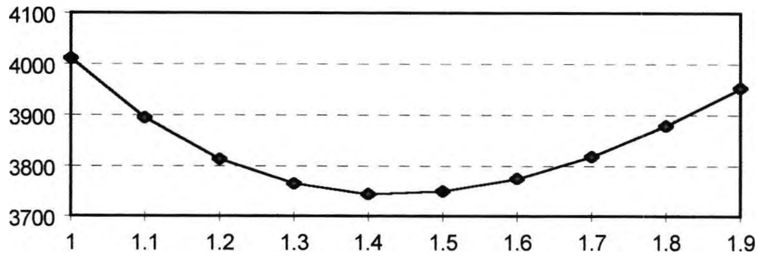
10.2.2 Analysis of age specific mortality trends

We adopt the perspective of Sections *8.2.4* and *9.2.4* at the outset and focus on fractional polynomial of the form

$$\eta_{xt} = \alpha_x + \beta_x \cdot t^k \qquad \textbf{(10.1)}$$

treating age as a factor. Fitting the above model structure, in association with the inverse link function, for different predetermined values of k (in steps of *0.1*) leads to the deviance profile displayed in Figure *10.1*.

Figure *10.1* : Deviance profile against k, model *(10.1)*



This suggests the model structure

$$\mu_{x,t} = \alpha_x + \beta_x \cdot t^{1.4} \qquad \textbf{(10.2)}$$

leading to the two sets of variable parameter estimates displayed in Figures *10.2* & *10.3*. The deviance for the model structure is *3744.4* on *2046* degrees of freedom and the estimated scale parameter $\hat{\varphi}$ = *1.83*. Again this heavily parameterised structure was found to fit the data well. As evidence of this the p - values for the dual statistical tests based on each age (rather than period) are presented in Table *10.1*.

Table 10.1 : p - values, formal graduation tests for each age separately, model *(10.2)*

Age	ISD	Sign	Runs	Chi
24	16	97	70	42
25	99	57	81	65
26	63	57	18	75
27	99	30	72	59
28	100	43	29	49
29	29	30	72	75
30	77	70	58	43
31	99	30	19	52
32	96	70	58	37
33	100	57	43	46
34	84	19	98	59
35	92	70	11	43
36	70	57	18	55
37	94	30	58	50
38	99	43	97	58
39	94	30	31	51
40	94	57	**2**	43
41	96	70	5	39
42	50	43	5	43
43	97	57	**2**	47
44	47	19	9	50
45	98	57	**2**	45
46	81	57	**1**	37
47	98	57	**2**	49
48	100	57	**0**	45
49	53	19	48	50
50	77	57	43	49
51	19	57	29	52
52	74	57	70	51
53	84	70	19	47
54	90	70	31	45
55	96	70	31	49
56	95	43	29	47
57	34	89	67	45
58	99	43	10	47
59	99	43	57	49
60	92	19	22	46
61	84	43	10	49
62	69	70	58	44
63	93	57	**1**	47
64	96	70	**0**	48
65	85	70	31	44
66	23	94	60	41
67	85	57	57	47
68	92	19	75	52
69	98	57	94	49
70	99	43	18	49
71	97	57	94	46
72	21	19	48	44
73	97	57	70	50
74	72	19	6	51
75	99	43	10	49
76	63	11	88	52
77	45	19	22	52
78	95	70	44	48
79	61	70	90	52
80	79	81	62	46
81	98	57	89	46
82	97	70	11	48
83	78	70	44	48
84	82	70	**1**	50
85	88	57	18	48
86	79	43	43	44
87	78	43	97	49
88	36	19	75	50
89	43	89	67	40

Figure 10.2 : α_x against age, model *(10.2)*

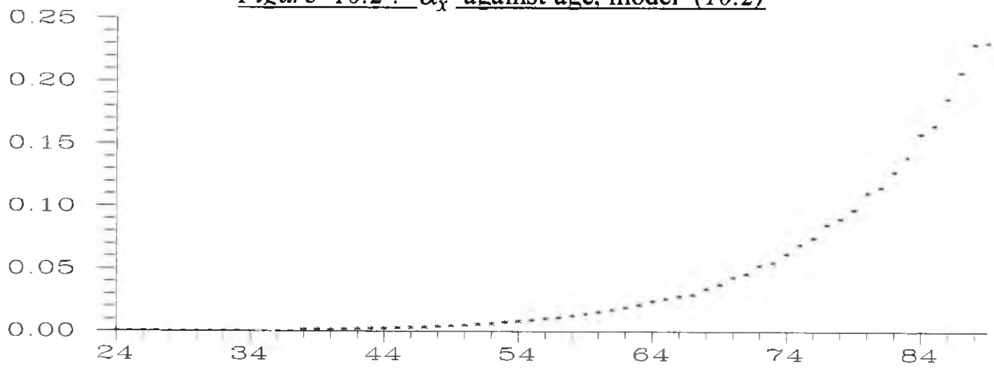
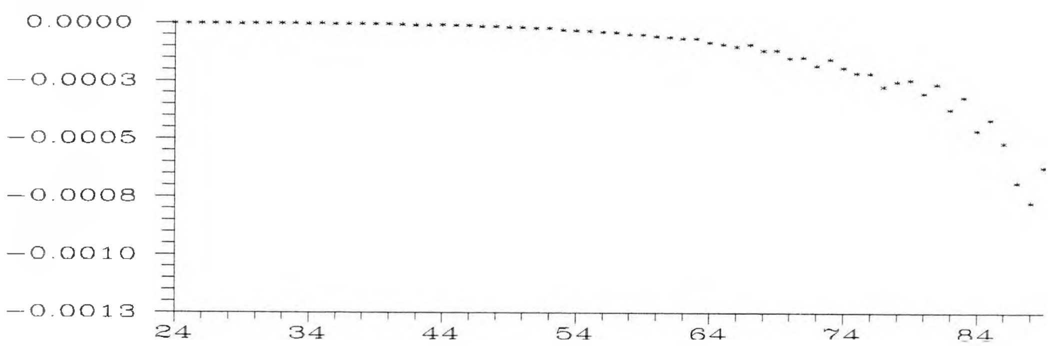


Figure 10.3 : β_x against age, model *(10.2)*



The values α_x represent the level of mortality at $t = 1$ (year *1958*), for each age $x = 24, 25, \ldots, 89$, with higher values denoting higher levels of mortality. The values β_x represent the pace at which mortality is decreasing in time, for each age x. Lower values denote a faster pace of decrease.

10.2.3 Modelling trends using fractional polynomial predictor structures in time effects and cubic spline predictor structures in age effects

Since polynomials in x were not found to give satisfactory results for modelling the parametric trends of α_x & β_x, in Figures *10.2* and *10.3* respectively, for the model structure *10.2*, spline functions were tried as an alternative, and *cubic spline functions* found to give satisfactory results. (In particular we note that mortality rates for English Life Table *No 14* were graduated using cubic spline functions - see Section *2.2*).

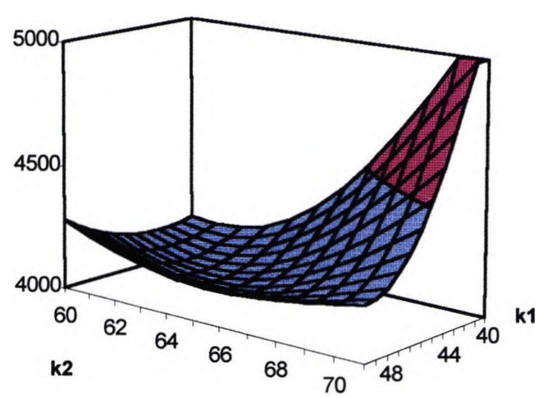
Two knots are found to be satisfactory (one knot was not sufficient to describe the patterns noted earlier) located at the ages *47* & *64*, for both cases, due to high empirical coefficients of correlation of the above parameters α_x & β_x. These knot positions were chosen by monitoring the deviance profile for different knot positions under the following model structure

$$\mu_{x,t} = a_0 + a_1 \cdot x + a_2 \cdot x^2 + a_3 \cdot x^3 + a_4 \cdot (x-k_1)_+^3 + a_5 \cdot (x-k_2)_+^3 +$$

$$\{b_0 + b_1 \cdot x + b_2 \cdot x^2 + b_3 \cdot x^3 + b_4 \cdot (x-k_1)_+^3 + b_5 \cdot (x-k_2)_+^3\} \cdot t^{1.4} \qquad \textbf{(10.3)}$$

where k_1 & k_2 denote the knot positions. The deviance profile is shown in the Figure *10.4*.

Figure 10.4 : Deviance profile against knot positions, model *(10.3)*



Thus, the final model derived is of the following form

$$\mu_{x,t} = a_0 + a_1 \cdot x + a_2 \cdot x^2 + a_3 \cdot x^3 + a_4 \cdot (x-47)_+^3 + a_5 \cdot (x-64)_+^3 +$$

$$\{b_1 \cdot x + b_2 \cdot x^2 + b_3 \cdot x^3 + b_4 \cdot (x-47)_+^3 + b_5 \cdot (x-64)_+^3\} \cdot t^{1.4} \qquad \textbf{(10.4)}$$

Note that the parameter b_0 was found to be insignificant and has been excluded from the model structure *(10.4)*.

The parameter estimates, standard errors, and t - tests for the model structure *(10.4)* are as shown in Table *10.2*.

Table 10.2 : Parameters estimates, Standard errors, and t - tests, model *(10.4)*

	p.e.	***s.e.***	***t - test***
a_0	*-0.004027*	*0.0006288*	*-6.4*
a_1	*0.0006118*	*0.00005113*	*11.9*
a_2	*-2.48E-05*	*1.37E-06*	*-18.1*
a_3	*3.24E-07*	*1.21E-08*	*26.7*
a_4	*1.03E-06*	*4.59E-08*	*22.4*
a_5	*4.85E-06*	*2.50E-07*	*19.3*
b_1	*-1.04E-06*	*1.01E-07*	*-10.3*
b_2	*6.12E-08*	*4.99E-09*	*12.2*
b_3	*-9.62E-10*	*6.03E-11*	*-15.9*
b_4	*-2.27E-09*	*4.74E-10*	*-4.7*
b_5	*-1.07E-08*	*3.10E-09*	*-3.4*
	$\hat{\varphi} = 1.877$		

The deviance for the model structure is *4067.7* on *2167* degrees of freedom.

Tables *10.3* displays the p - values for the statistical tests constructed by focusing on each age x, for the model structure *(10.4)*. The tests are very satisfactory and are supportive of the model structure.

Table 10.3 : p - values, formal graduation tests for each age separately, model *(10.4)*

Age	ISD	Sign	Runs	Chi
24	43	89	26	32
25	61	6	46	33
26	79	43	19	70
27	99	30	72	59
28	79	81	34	59
29	7	30	72	83
30	48	81	34	50
31	43	43	43	61
32	57	94	5	52
33	96	57	43	46
34	86	19	98	57
35	94	57	11	31
36	71	30	6	42
37	76	19	62	42
38	26	30	95	37
39	69	11	**1**	50
40	8	99	17	57
41	69	19	34	29
42	23	**99**	21	48
43	49	89	39	51
44	84	70	83	56
45	82	70	12	48
46	54	30	**0**	42
47	89	57	**2**	48
48	47	70	**0**	52
49	53	19	48	49
50	90	19	13	45
51	20	43	30	50
52	21	43	30	51
53	69	19	13	46
54	90	56	30	43
55	96	43	30	47
56	72	30	59	45
57	47	89	68	46
58	84	89	68	52
59	37	69	31	54
60	89	81	48	50
61	83	30	12	47
62	57	19	85	41
63	45	6	19	42
64	6	3	14	42
65	96	70	31	43
66	9	99	31	49
67	10	99	90	59
68	43	89	79	59
69	87	81	99	52
70	79	19	**1**	48
71	85	70	95	49
72	356	19	48	45
73	68	70	59	52
74	47	19	7	53
75	85	11	15	45
76	61	11	88	51
77	78	11	54	48
78	88	30	72	48
79	65	70	90	51
80	10	19	34	39
81	98	57	30	44
82	17	**99**	20	56
83	51	69	90	44
84	72	57	5	50
85	68	94	19	52
86	48	19	22	43
87	32	94	61	52
88	88	19	34	48
89	6	**99**	80	46

Additional supportive diagnostic evidence for model *(10.4)* is provided by the plots of standardised deviance residuals against the constant information scale ($CIS = 2 \cdot \log(1/\mu_{tx})$), reproduced at ten yearly age intervals in Figure *10.5*. The predicted force of mortality (for the time period *1958* to *2010*), at ten yearly age intervals is shown in Figure *10.6*.

Figure 10.5 : Standardised deviance residuals vs. *CIS*, various ages, model *(10.4)*

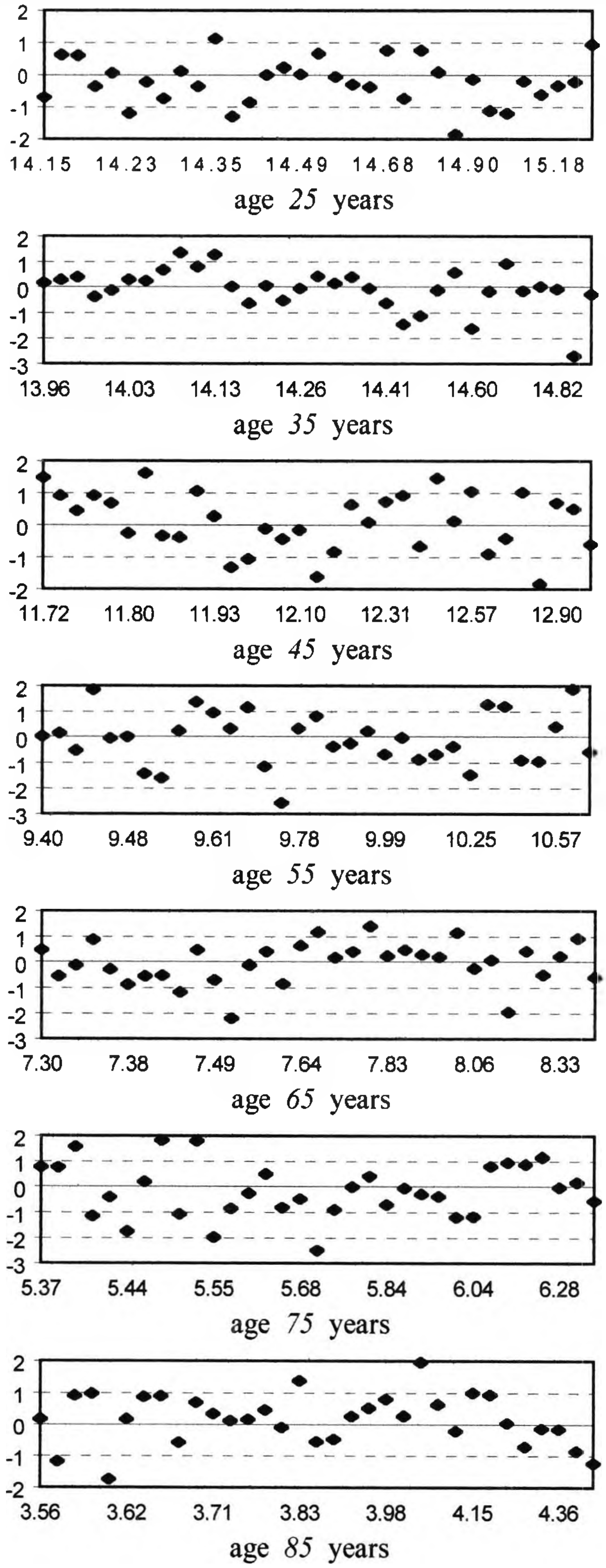
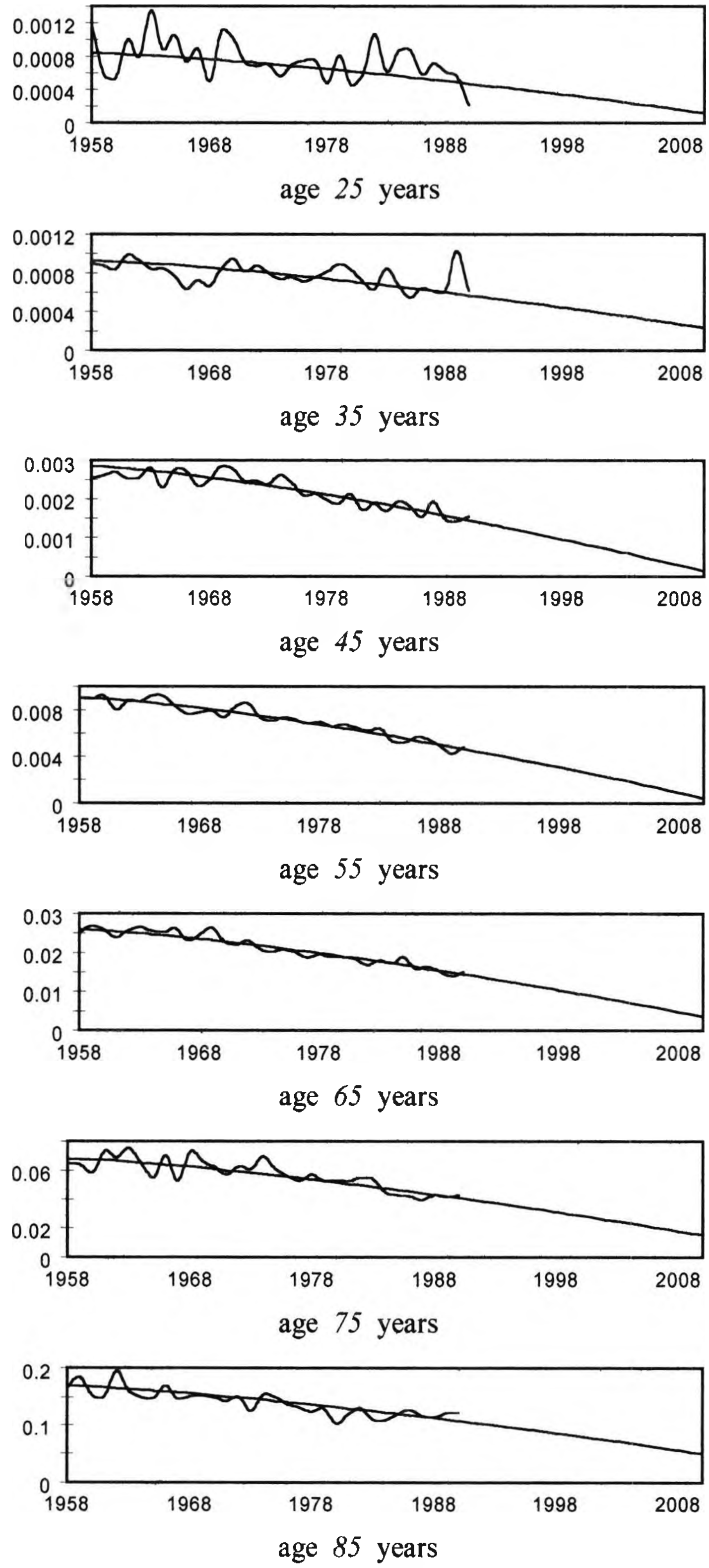


Figure 10.6 : Crude and predicted force of mortality vs. calendar year, various ages, model *(10.4)*



Here we have superimposed the estimated mortality curves on the corresponding crude mortality rates. As usual this acts as a further visual check on the predictive qualities of the model.

Finally for this model, the predicted values of the force of mortality, μ_{xt}, in the age range x = *24* to *89* years at *5* yearly ages, over the calendar period t = *1960* to *1990* at *10* yearly intervals, and the forecast values for the years *2000* and *2010* are presented for completeness in the following Table *10.4*.

Table 10.4 : Predicted force of mortality, *10* yearly intervals, quinquennial ages, model *(10.3)*

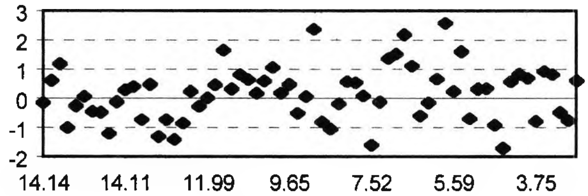
	1960	**1970**	**1980**	**1990**	**2000**	**2010**
25	0.000835	0.000746	0.000623	0.000475	0.000308	0.000124
30	0.000770	0.000704	0.000611	0.000499	0.000373	0.000235
35	0.000920	0.000835	0.000716	0.000574	0.000414	0.000237
40	0.001523	0.001357	0.001124	0.000846	0.000532	0.000186
45	0.002820	0.002487	0.002020	0.001462	0.000831	0.000138
50	0.005077	0.004467	0.003611	0.002588	0.001431	0.000160
55	0.008974	0.007923	0.006448	0.004684	0.002690	0.000500
60	0.015509	0.013776	0.011344	0.008437	0.005151	0.001542
65	0.025690	0.022958	0.019126	0.014544	0.009364	0.003675
70	0.041545	0.037349	0.031463	0.024426	0.016469	0.007731
75	0.067371	0.060937	0.051910	0.041118	0.028917	0.015518
80	0.107771	0.097994	0.084278	0.067880	0.049340	0.028979
85	0.167350	0.152795	0.132377	0.107966	0.080367	0.050057

Further supportive evidence is provided by the p - values for the statistical tests constructed by focusing on each year t, for the model structure *(10.3)*, Table *10.5*. We also display some of the many standardised deviance residual plots on the constant information scale *CIS* = $(2 \cdot \log(1/\mu_{xt}))$, for various calendar years, in Figure *10.7*.

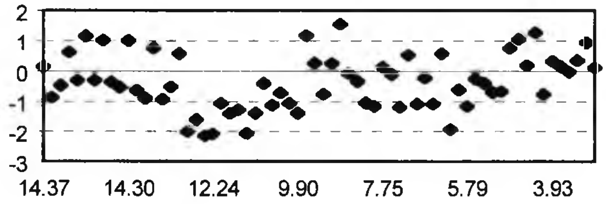
Table 10.5 : p - values, formal graduation tests for each calendar year separately, model *(10.3)*

Year	*ISD*	*Sign*	*Runs*	*Chi*
1958	94	50	59	50
1959	97	59	23	45
1960	42	89	9	59
1961	83	50	50	53
1962	93	89	**3**	59
1963	5	**1**	44	30
1964	86	59	**2**	50
1965	**5**	**1**	39	38
1966	34	31	78	52
1967	**0**	**100**	34	66
1968	28	**99**	24	59
1969	**4**	4	74	27
1970	8	**1**	29	36
1971	44	83	**3**	59
1972	64	10	5	42
1973	38	23	7	49
1974	86	40	77	51
1975	65	89	**3**	63
1976	92	50	16	50
1977	30	97	59	51
1978	7	**99**	5	50
1979	66	95	7	57
1980	55	7	**4**	52
1981	43	93	31	54
1982	95	83	64	48
1983	49	23	94	43
1984	23	95	74	56
1985	83	68	78	44
1986	56	16	35	51
1987	41	40	97	67
1988	98	68	78	62
1989	75	89	**2**	48
1990	9	**0**	14	42

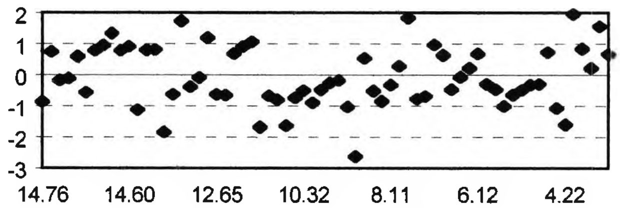
Figure 10.7 : Standardised deviance residuals vs. *CIS*, various calendar years, model *(10.4)*



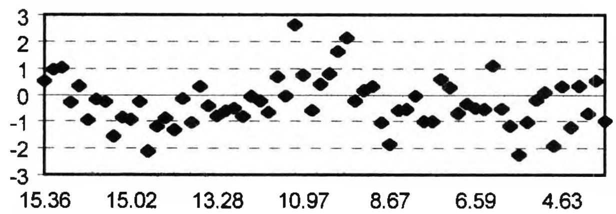
calendar year *1960*



calendar year *1970*



calendar year *1980*



calendar year *1990*

CHAPTER XI

Complementary log - log link models

11.1 Introduction

In this Chapter, we focus on the modelling assumptions of Section *6.2*, using the number of deaths P_{xt} as (over - dispersed binomial) responses, with

$$E(P_{xt}) = m_{xt} = {}^{p}R^{i}_{xt} \cdot q_{xt} \qquad \& \qquad Var(P_{xt}) = \varphi \cdot m_{xt} \cdot (1 - \frac{m_{xt}}{{}^{p}R^{i}_{xt}})$$

involving a scale parameter $\varphi > 1$, prior weights $\omega_{xt} = 1$, and variance function

$$V(m_{xt}) = m_{xt} \cdot (1 - \frac{m_{xt}}{{}^{p}R^{i}_{xt}})$$

where ${}^{p}R^{i}_{xt}$ denote the initial exposures. Both responses and exposures are based on policy counts. Used in combination with the complementary log - log link function

$$\log\{-\log(1 - q_{xt})\} = \eta_{xt}$$

we target q_{xt} rates, the probability that a life aged x dies before age $x+1$, in period t, where η_{xt} denotes the linear predictor.

11.2 Males pensioners, period 1983 - 1990, ages 60 - 95

11.2.1 Description of the data

The data modelled in this Chapter were provided by the *CMI* Bureau, and refer to the male pensioners' experience, under *UK* life office pension schemes. They consist of initial exposures and policy totals ceasing through death, by individual calendar year, for the period *1983* to *1990* and ages *60* to *95*, both inclusive. The data are presented in Appendix *B*, as published by the *CMI* Bureau of the Institute and Faculty of Actuaries.

Since the exposures are initial and the data based on policy rather than head counts, the number of policies ceasing through death are modelled as over - dispersed binomial variables.

11.2.2 Modelling trends using polynomial predictor structures

We have investigated the complementary *log - log* link function in combination with polynomial predictor formulae of the type

$$\eta_{x,t} = \beta_0 + \sum_{j=1}^{s} \beta_j \cdot L_j(x') + \sum_{i=1}^{r} \alpha_i \cdot t'^i + \sum_{i=1}^{r} \sum_{j=1}^{s} \gamma_{ij} \cdot L_j(x') \cdot t'^i \qquad \textbf{(11.1)}$$

in which some of the parameters may be pre - set to zero. Here both the age and calendar year ranges have been mapped onto the interval [*-1, 1*] by translating the origin to the centre of the range and using the semi-range for scaling, and where $L_j(x)$ denote Legendre polynomials, as described in Section *5.7.*

An examination of the deviance profile induced by changes in the structure of the linear predictor formula *(11.1)*, coupled with copious graphical tests of the corresponding deviance residuals, leads to the adoption of the model formula

$$\log\{-\log(1-q_{x,t})\} = \beta_0 + \sum_{j=1}^{3} \beta_j \cdot L_j(x') + \sum_{i=1}^{3} \alpha_i \cdot t'^i + \gamma_{11} \cdot L_1(x') \cdot t' \qquad \textbf{(11.2)}$$

The details for the parameter estimates, standard errors, and t - values are presented in Table *11.1*. The p - values, for the corresponding statistical tests, are based on standardised deviance residuals, given by the expression *(6.4)*, and are presented in Table *11.2*. Note that the estimated scale parameter, $\hat{\varphi}_t$, is calculated for each year separately.

Table 11.1 : Parameters estimates, Standard errors, and t - tests, model *(11.2)*

	p.e.	s.e.	t - test
β_0	-2.674	0.00746	-412.5
β_1	1.637	0.01621	114.6
β_2	-1.499E-01	1.256E-02	-11.9
β_3	-3.243E-02	1.392E-02	-2.3
α_1	-4.721E-02	1.27E-02	-3.7
α_2	-3.314E-02	9.02E-03	-3.6
α_3	-3.846E-02	1.622E-02	-3.0
γ_{11}	-2.405E-02	1.394E-02	-1.7
	$\hat{\varphi} = 1.58$		

The value of the deviance is *442.28* on *280* degrees of freedom.

Table 11.2 : p - values, formal statistical tests for each calendar year separately, model *(11.2)*

Year	**ISD**	**Sign**	**Runs**	**Chi**
1983	81	75	93	53
1984	43	**2**	78	44
1985	26	95	95	59
1986	87	16	81	45
1987	66	**100**	43	48
1988	90	50	**0**	47
1989	22	9	62	53
1990	38	37	76	42

Finally, for the model *(11.2)*, the predicted values of the rate of mortality, q_{xt}, in the age range $x = 60$ to 95 years, over the calendar period $t = 1983$ to 1990 at yearly intervals, are presented for completeness in the following Table *11.3*.

Table 11.3 : Predicted q_{xt} probabilities, period *1983 - 1990*, quinquennial ages, model *(11.2)*

	1983	**1984**	**1985**	**1986**	**1987**	**1988**	**1989**	**1990**
60	0.01280	0.01245	0.01219	0.01198	0.01174	0.01141	0.01094	0.0102
65	0.02189	0.02132	0.02093	0.02061	0.02023	0.01970	0.01893	0.0178
70	0.03702	0.03613	0.03555	0.03507	0.03450	0.03367	0.03243	0.0306
75	0.06102	0.05969	0.05885	0.05817	0.05735	0.05609	0.05415	0.0513
80	0.09656	0.09468	0.09355	0.09266	0.09154	0.08974	0.08685	0.0825
85	0.14462	0.14214	0.14073	0.13969	0.13830	0.13590	0.13187	0.1257
90	0.20241	0.19940	0.19785	0.19678	0.19523	0.19230	0.18712	0.1789
95	0.26223	0.25892	0.25742	0.25651	0.25500	0.25174	0.24564	0.2357

11.2.3 Modelling trends using fractional polynomial predictor structures

We have also investigated the complementary log - log link function in combination with fractional polynomial predictor formula. Various parameterised predictor structures of the form

$$\eta_{xt}(a,b) = \alpha_t \cdot x^a + \beta_t \cdot x^b$$

were tried in combination with the complementary log - log link function in an attempt to target q_{xt} rates. The optimum values of a and b are determined by monitoring the improvement in the model deviance for different combinations of the values of a and b (in steps of *0.1*). The minimum deviance obtained is *437.2* when $a = -0.4$ and $b = 0$ or when $a = -0.3$ and $b = 0.1$. Besides, another pair of values of a and b which produces a simpler model is when $a = 1$ and $b = -1$. The value of the deviance now is *441.23* on *272* degrees of freedom (the difference occur in the deviances is statistically insignificant) and the estimated value of the scale parameter is $\hat{\varphi} = 1.622$. This combination assumes the following model structure

$$\eta_{xt} = \alpha_t \cdot x + \beta_t \cdot \frac{1}{x} \qquad \text{or} \qquad q_{x,t} = 1 - e^{-e^{\alpha_t \cdot x + \beta_t \cdot \frac{1}{x}}} \tag{11.3}$$

fitted separately for each period t, by treating t as a factor and age x as a variate, in an attempt to detect any patterns in the parameters α_t & β_t over time t.

The p - values for the statistical tests of the graduation, based on standardised deviance residuals, are applied separately for each calendar year, which are highly supportive of this heavily parameterised structure, are presented in Table *11.4*.

<u>*Table 11.4 :* p - values, formal graduation tests for each calendar year separately, model *(11.3)*</u>

Year	***ISD***	***Sign***	***Runs***	***Chi***
1983	*94*	*50*	*75*	*53*
1984	*100*	*50*	*75*	*46*
1985	*56*	*37*	*99*	*62*
1986	*96*	*37*	*64*	*49*
1987	*46*	*84*	*11*	*48*
1988	*33*	*37*	*51*	*47*
1989	*78*	*16*	*89*	*55*
1990	*43*	*16*	*11*	*44*

The trend in the parameter estimates, α_t & β_t, are displayed in Figures *11.1* & *11.2*.

Figure 11.1: α_t against t, model *(11.3)*

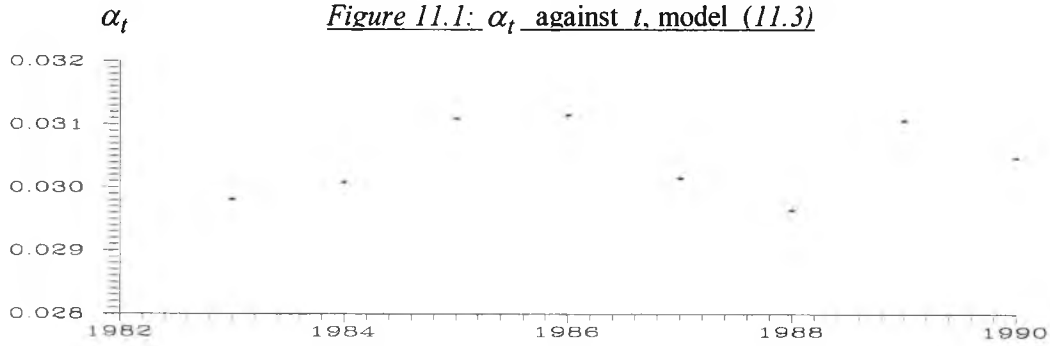


Figure 11.2 : β_t against t, model *(11.3)*

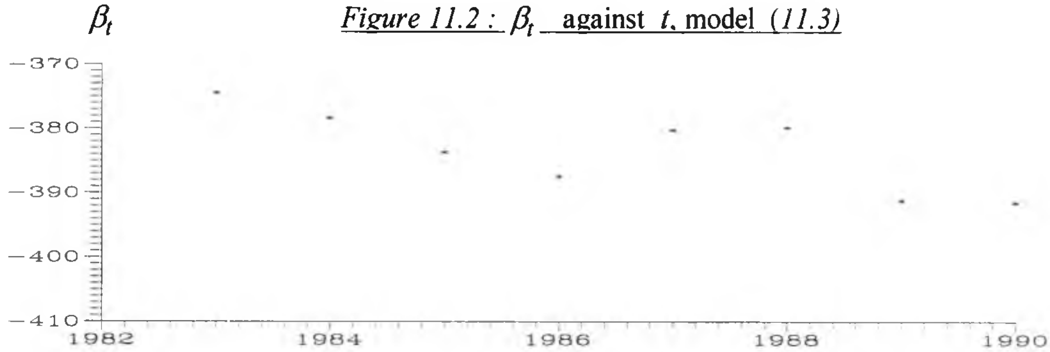


Figure *11.1* does not indicate any particular trend in the α_t values, while Figure *11.2* indicates a downward movement in the β_t. As a consequence of these graphs plus further preliminary analysis it was decided to set

$$\alpha_t = \alpha \qquad \& \qquad \beta_t = \beta + \gamma \cdot t^k$$

for different predetermined values of k. The optimum value $k = 2$ was determined by looking at the appropriate deviance profile constructed by fitting the model structure for various values of k (note that the value of $k = 2.3$ gives the minimum deviance, but for reasons of simplicity we will use the value of *2* since there is a relatively very small discrepancy in the deviance).

Thus the linear predictor finally adopted has the following mathematical expression

$$\eta_{x,t} = \alpha \cdot x + \beta \cdot \frac{1}{x} + \gamma \cdot \frac{t^2}{x}$$

The parameter estimates, standard errors, and t - values for the model structure

$$\log\{-\log(1-q_{xt})\} = \alpha \cdot x + \beta \cdot \frac{1}{x} + \gamma \cdot \frac{t^2}{x} \qquad \textbf{(11.4)}$$

where $t = 1, 2, \ldots$ (for calendar year *1983, 1984, ...*), are as shown in Table *11.5*.

Table *11.5* : Parameters estimates, standard errors, and t - tests, model *(11.4)*

	p.e.	***s.e.***	***t - test***
α	0.03042	0.000257	118.6
β	-378.6	1.542	-245.5
γ	-0.1814	0.01288	-14.1
	$\hat{\varphi} = 1.669$		

The deviance is *475.64* on *285* degrees of freedom.

The p - values for the statistical tests of a graduation, which are based on standardised deviance residuals, are presented in Table *11.6*.

Table 11.6 : p - values, formal graduation tests for each calendar year separately, model *(11.4)*

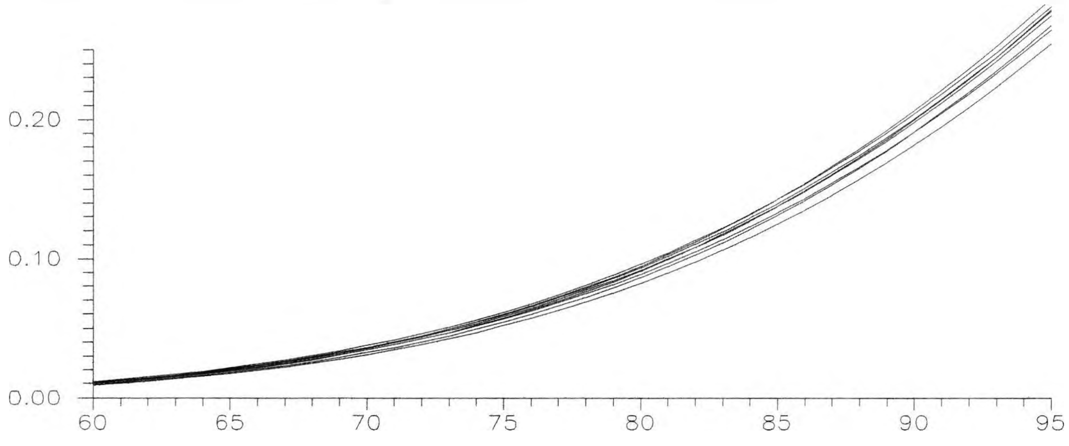
Year	*ISD*	*Sign*	*Runs*	*Chi*
1983	59	91	62	55
1984	28	**1**	61	44
1985	81	75	87	61
1986	45	**2**	87	46
1987	13	**99**	30	50
1988	49	63	9	49
1989	55	37	64	55
1990	16	5	17	42

While the model structure *(11.4)* fails some of the sign tests, it is nevertheless retained because the results for the remaining tests are satisfactory.

Further, because the data are only available for a short run of years, long term forecasting of mortality rates would be risky. However the fitted model structure *(11.4)* can be extrapolated forward a few years (up to *4* years say) to forecast mortality rates.

A display of the overall mortality trend is given in Figure *11.3,* where the curves represent the graduated mortality values for each calendar year (*1983* to *1990*) separately against age x.

Figure 11.3 : Graduated q_{xt} values against x, calendar year *1983 - 1990*, model *(11.3)*



The overall level of mortality improvement is noted by the downward movement of the graduated curve through time (starting at calendar year *1983* and ceasing at calendar year *1990*), with more improvement occurring at the oldest ages. The corresponding predicted values of q_{xt} in the age range $x = 60$ to 95 years, at 5 yearly ages, for individual calendar years $t = 1983$ to *1990*, are presented for completeness in Table *11.7*.

Table 11.7 : Predicted q_{xt} probabilities, period *1983 - 1990*, quinquennial ages, model *(11.3)*

	1983	1984	1985	1986	1987	1988	1989	1990
60	0.01118	0.01108	0.01092	0.01069	0.01040	0.01006	0.00968	0.00925
65	0.02105	0.02088	0.02059	0.02020	0.01970	0.01911	0.01844	0.01769
70	0.03686	0.03658	0.03612	0.03548	0.03468	0.03372	0.03262	0.03140
75	0.06080	0.06038	0.05967	0.05870	0.05748	0.05601	0.05432	0.05244
80	0.09529	0.09468	0.09366	0.09226	0.09048	0.08836	0.08590	0.08315
85	0.14277	0.14192	0.14053	0.13860	0.13615	0.13321	0.12981	0.12600
90	0.20526	0.20416	0.20233	0.19980	0.19659	0.19273	0.18824	0.18319
95	0.28381	0.28244	0.28017	0.27703	0.27302	0.26819	0.26257	0.25620

For this model, making use of the well - known relationship

$$q_{xt} = 1 - e^{-\mu_{xt}}$$

which is exact if μ_{xt} is assumed to be piecewise constant within each cell (x,t) it follows that the force of mortality is approximately described by the relationship

$$\log(\mu_{xt}) = \alpha \cdot x + \beta \cdot \frac{1}{x} + \gamma \cdot \frac{t^2}{x}$$

CHAPTER XII

Modelling amounts

12.1 Introduction

Following up the discussion, in Chapter *XI*, about the pensioners' experience under *UK* life office pension schemes, the graduation of the so-called 'amounts' data is addressed. These data include, in addition to the policies and exposures based on annuity counts, the total amounts associated with both the policies and exposures. Previous experience reveals that the 'amounts' based experience shows a lower mortality than the corresponding 'lives' based experiences.

The aim of this chapter is to predict the probability of death based on the 'amounts' experience, taking detailed account of the underlying structure of the data involved.

12.2 Distribution assumptions

We define, for each age x and each calendar year t:

P = the total number of policies ceasing through deaths

${}^{p}R^{i}$ = the initial exposed to risk based on policy counts

A = the total amount of pension accruing from deaths

e = the exposed to risk based on 'amounts'

$A^{(i)}$ = the amount associated with policy, i

$\overline{A}$ = the average amount accruing from deaths

q_x = the probability that a life, age x, dies before age $x+1$ based on 'lives'

q_x^{*} = the probability that a life, age x, dies before age $x+1$ based on 'amounts'

The data available for analysis comprise $(P,\ {}^{p}R^{i},\ A,\ e)$, for each member of the rectangular grid of cells $(x,\ t)$. In each cell $u = (x,\ t)$, the $A^{(i)}$ are modelled as independent, identically distributed non-negative random variables, independent of P, so that

$$A_u = \sum_{i=1}^{P_u} A_u^{(i)}$$

and hence (see equations *5.2*)

$$E(A_u) = E(A_u^{(i)}) \cdot E(P_u)$$

&

$$Var(A_u) = Var(A_u^{(i)}) \cdot E(P_u) + \{E(A_u^{(i)})\}^2 \cdot Var(P_u) \qquad \textbf{(12.1)}$$

Assumption I

Model the average amounts $\overline{A}_u$ as the independent *gamma* response variables of a *GLM* and define

$$E(\overline{A}_u) = \rho_u \qquad \& \qquad Var(\overline{A}_u) = \frac{\psi \cdot \rho_u^2}{P_u} \qquad \textit{(12.2)}$$

with scale parameter ψ, weights P_u, and variance function $Var(\rho_u) = \rho_u^2$.

Note the responses $\overline{A}_u$ are given by the ratio

$$\overline{A}_u = \frac{A_u}{P_u}$$

The above assumption pre-supposes, in part, that the individual amounts $A_u^{(i)}$ follow a gamma distribution (Renshaw and Hatzopoulos, *1996*) with

$$E(A_u^{(i)}) = \rho_u \qquad \& \qquad Var(A_u^{(i)}) = \psi \cdot \rho_u^2 \qquad \textit{(12.3)}$$

Assumption II

In keeping with Section *6.2*, the number of policies ceasing through deaths P_u, are modelled as the independent over-dispersed *binomial* response variables of a *GLM*, with

$$E(P_u) = m_u = q_u \cdot {}^{p}R_u^i \quad \& \quad Var(P_u) = \tau \cdot m_u \cdot (1 - \frac{m_u}{{}^{p}R_u^i}) \qquad \textit{(12.4)}$$

Assumption III

Then the total amounts of pension accruing from deaths A_u are modelled as the independent responses of a *GLM*, with

$$E(A_u) = q_u^* \cdot e_u \quad \& \quad Var(A_u) = (\psi + \tau) \cdot \rho_u \cdot E(A_u) - \frac{\tau}{{}^{p}R_u^i} \cdot \{E(A_u)\}^2 \qquad \textit{(12.5)}$$

Expressions *(12.5)* follow from equations *(12.1)* in combination with equations *(12.2), (12.3) & (12.4)*. Equation *(12.5)*, can be implemented in *GLIM*, by declaring $Y_u = \frac{A_u}{e_u}$ as the response variables for which

$$E(Y_u) = q_u^* \quad \& \quad Var(Y_u) = \frac{1}{\omega_u} \cdot \{E(Y_u) - \frac{\{E(Y_u)\}^2}{\kappa_u}\} \tag{12.6}$$

with weights ω_u, where

$$\omega_u = \frac{e_u}{(\psi + \tau) \cdot \rho_u} \qquad \& \qquad \kappa_u = \frac{pR_u^i}{\tau \cdot \omega_u}$$

Assumption III may be implemented in combination with any of the predictor-link relationships generally associated with binomial response *GLMs*.

12.3 Implementation

The modelling of q_u^* based on *Assumption III* requires the estimation of ψ, τ & ρ_u before equations *(12.6)* can be implemented.

Assumption I can be used to model the average amounts of pension in combination with any suitable predictor - link combination, thereby providing an estimate for the scale parameter ψ, as well as providing fitted values to estimate the ρ_u's.

For the *UK* male pensioners data set, *Assumption I* was applied using the log link in combination with a linear spline function, with seven knots positioned at ages *60, 65, 75, 76, 79, 82, 90,* for each calendar year separately. By this method the scale parameter ψ was estimated as $\hat{\psi} = 6.134$ and the predicted values ρ_{xt} are reproduced in Table *12.1* (Renshaw and Hatzopoulos, *1996*, Table *5.4*).

Table 12.1 : Predicted ρ_{xt} values (x = age, t = calendar year) with $\hat{\psi} = 6.134$

age	1983	1984	1985	1986	1987	1988	1989	1990
60	630	5293	3849	2895	956	2962	1066	4164
61	904	3868	3064	3955	1258	3188	1741	4998
62	1297	2826	2438	5403	1655	3431	2844	5999
63	1860	2065	1941	7380	2178	3693	4646	7201
64	1030	1280	1298	2578	1436	2171	2694	3232
65	570	794	869	900	947	1276	1562	1451
66	524	714	788	825	879	1151	1419	1389
67	482	642	714	755	816	1039	1289	1330
68	443	577	648	692	758	938	1171	1274
69	407	519	588	633	704	847	1065	1220
70	374	467	533	580	653	764	968	1168
71	344	420	483	531	606	690	879	1119
72	316	378	438	487	563	623	799	1071
73	290	340	398	446	523	562	726	1026
74	267	306	361	408	485	507	660	982
75	245	275	327	373	450	458	600	941
76	232	295	287	360	396	432	537	638
77	246	271	279	349	378	416	493	600
78	262	248	271	338	359	401	453	564
79	279	228	263	327	342	386	416	530
80	261	231	256	301	319	350	396	460
81	244	233	250	278	297	318	377	399
82	229	236	244	256	277	288	359	346
83	223	233	238	249	272	287	342	336
84	217	230	233	243	267	286	327	327
85	212	228	227	236	262	284	312	319
86	207	225	222	230	257	283	298	310
87	201	222	217	223	252	282	284	302
88	196	220	211	217	248	280	271	294
89	191	217	207	212	243	279	259	286
90	187	215	202	206	239	278	247	279
91	192	217	214	209	231	250	262	262
92	198	220	226	212	223	226	278	247
93	203	223	239	215	215	203	295	232
94	209	225	263	218	208	183	313	219
95	215	228	268	221	201	165	332	206

The scale parameter τ is estimated under *Assumption II* on applying this assumption to the appropriate data set based on policy counts, typically as described in Chapter *XI*. There, for the *UK* male pensioners data set, τ is estimated as $\hat{\tau} = 1.58$.

We now proceed to implement Assumption *III* using the 'own' model specification commands in *GLIM*. The results, presented in Tables *12.2* & *12.3* are based on the mathematical formula

$$\log\{-\log(1 - q^*_{x,t})\} = \beta_0 + \beta_1 \cdot x' + \beta_2 \cdot x'^2 + \alpha_1 \cdot t'$$

consisting of the complementary log - log link in combination with significant polynomial terms in x' & t', the transformed ages and periods respectively (as defined in Section *5.7*). Table *12.2* contains detail of the parameter estimates and their standard errors (Renshaw and Hatzopoulos, *1996*, Table *5.6*) and Table *12.3* lists the predicted values (Renshaw and Hatzopoulos, *1996*, Table *5.5*).

The residual plots and statistical tests for this fit, which are supportive of the model structure, are not reproduced.

Table 12.2 : Predictions based on 'amounts', parameter estimates with standard errors

	p.e.	*s.e.*	*t-test*
β_0	-2.83	0.01136	-249.1
β_1	1.839	0.01992	92.3
β_2	-0.1174	0.02918	-4.0
α_1	-0.1011	0.01144	-8.8

Table 12.3 : Predicted q^*_{xt} probabilities based on 'amounts' (x = age, t = calendar year)

Age/Year	1983	1984	1985	1986	1987	1988	1989	1990
60	0.00919	0.00893	0.00867	0.00843	0.00819	0.00796	0.00773	0.00751
65	0.01687	0.01639	0.01593	0.01548	0.01504	0.01462	0.01420	0.01380
70	0.03001	0.02917	0.02835	0.02755	0.02670	0.02603	0.02530	0.02458
75	0.05165	0.05021	0.04882	0.04746	0.04614	0.04486	0.04361	0.04240
80	0.08577	0.08343	0.08116	0.07894	0.07678	0.07468	0.07263	0.07064
85	0.13700	0.13337	0.12983	0.12638	0.12302	0.11973	0.11653	0.11341
90	0.20962	0.20430	0.19911	0.19403	0.18906	0.18421	0.17947	0.17483
95	0.30574	0.29849	0.29137	0.28439	0.27754	0.27082	0.26423	0.25778

Part 4

Comparing Mortality Experiences

CHAPTER XIII

Comparing mortality experiences and constructing mortality tables based on standard tables

13.1 Introduction

In this chapter we will investigate two types of hypotheses.

The first type considers the assumption that two mortality experiences exhibit the same underlying mortality. In other words, we examine if two mortality tables can be modelled by the same mathematical structure involving identical parametric sets. Thus, the hypothesis test takes the form :

$$H_0 : \underline{\beta}^1 = \underline{\beta}^2 \quad \text{vs} \quad H_1 : \underline{\beta}^1 \neq \underline{\beta}^2 \qquad \textit{(Hypothesis of type 1)}$$

The data sets for analysis (and comparison) consist of the male assured lives experience, for the time period *1958 - 1990* and the range of ages *23 - 62* (where there are sufficient data for analysis), for durations *0, 1, 2, 3, 4, 5+*. In particular, within these data grids, there are no cells in which zero numbers of deaths are recorded, and consequently no data cells are weighted out of the subsequent analysis in the examples presented in this Chapter.

The second type concerns the hypothesis that two mortality experiences are connected by a specific model structure of the form

$$\mu_x = f(x) \cdot \mu_x^s \qquad \textit{(Hypothesis of type 2)}$$

where $f(x)$ is a function of age x. In other words, the hypothesis that one set of mortality rates μ_x can be constructed by suitable adjustments to another set of standard mortality rates μ_x^s, see e.g. Chapter *15,* Benjamin and Pollard (*1980*). This can be a useful approach in circumstances where one of the two mortality experiences involves scanty data over part of the age range, especially at the two ends. Moreover, the above method will allow for extrapolation outside the (possible) restricted age range of one of the two data sets.

The first type of hypothesis is a special case of the second type of hypotheses when $f(x) = 1$ for all ages x.

Under the second type of hypothesis, the male assured lives mortality experience for duration 5+ is used to construct a standard table which is then used

1) to construct a life table for the pensioners lives mortality experience, for the year *1990,* in the age range *64 - 89* (where there are sufficient data), and

2) to construct a life table for the male assured lives mortality experience, at durations *0 - 1 - 2 - 3 - 4*, in the period *1958 - 1990* and the age range *23 - 62* (where there are sufficient data).

13.2 *Testing Hypotheses of the form :* $H_0 : \underline{\beta}^1 = \underline{\beta}^2$ vs $H_1 : \underline{\beta}^1 \neq \underline{\beta}^2$

13.2.1 Methodology

In this Chapter we focus on the modelling assumptions of Section *5.6* (that is, the normal approximation for the natural logarithm of the empirical resistivity to death), and which gives rise to exact statistical tests.

Focus first on a fixed period and fixed duration. Thus using $Q_x = log\ Y_x$ as (normal) responses, where $Y_x = \dfrac{{}^{p}R^{c}_{x}}{P_x}$, we have

$$E(Q_x) = m_x = log\ \frac{1}{\mu_x} \qquad \& \qquad Var(Q_x) = \rho \cdot \frac{1}{P_x}$$

with scale parameter $\varphi = \rho$, prior weights $\omega_x = P_x$, and variance function $V(m_x) = 1$.

First, we examine the particular model structure for each separate mortality experience. The class of models used is given by the following (flexible) polynomial structure in age effects

$$m_x = \sum_{i=0}^{k-1} \beta_i \cdot x^i \qquad \textbf{\textit{(13.1)}}$$

In order to determine the optimum polynomial degree, we test hypothesis of the form :

<u>Hypothesis I</u>

$$H_0\!: \ \beta_k = 0 \quad \text{vs} \quad H_1\!: \ \beta_k \neq 0$$

using the F - statistic

$$\frac{df_{k+1}}{1} \cdot \frac{Dev(k) - Dev(k+1)}{Dev(k+1)} \sim F_{1, df_{k+1}} \qquad \textbf{\textit{(13.2)}}$$

to determine the p - values, where $Dev(k)$ is the deviance and df_k the degrees of freedom for model structure (13.1) with k parameters (see also Section 3.3).

Having determined the optimum degree of the polynomial by this means, we have now to compare the mortality experiences. The following describes the procedure employed for the comparison of two mortality experiences, using a different kind of hypothesis of the form :

Hypothesis II

$$H_0: \ \beta_i^1 = \beta_i^2 \quad \text{vs} \quad H_1: \ \beta_i^1 \neq \beta_i^2 \qquad \forall \ i = 0, 1, \ldots, k\text{-}1$$

using the F - statistic

$$\frac{n-k}{k} \cdot \frac{SS_{\Theta_2} - SS_{\Theta_1}}{SS_{\Theta_1}} \sim F_{k,n-k}$$

where

$$SS_{\Theta_1} = \sum_{x=x_1}^{x_n} (Q_x - \hat{m}_x)^2 = \sum_{x=x_1}^{x_n} (Q_x - \sum_{i=0}^{k-1} \hat{\beta}_i^1 \cdot x^i)^2$$

$$SS_{\Theta_2} = \sum_{x=x_1}^{x_n} (Q_x - \sum_{i=0}^{k-1} \beta_i^2 \cdot x^i)^2$$

and Q_x are the responses for the mortality experience with fitted values $\hat{m}_x$ (Klonias, *1987*).

The statistics SS_{Θ_2} & SS_{Θ_1} are easily calculated. Thus SS_{Θ_1} is the deviance obtained when fitting the initial model, on which the inference is based. Further, SS_{Θ_2} is the 'deviance' obtained on replacing the fitted values from the initial model, in the expression for the deviance, with the fitted values under the null hypothesis H_0.

13.2.2 Grouping durations 0, 1, 2, 3 & 4 for male assured lives, period 1958 - 1990, ages 23 - 62

Quoting Puzey (*1986,* pages *126 - 127*), "Temporary initial selection is the name given to the phenomenon where mortality rates are believed to depend on the duration since passing some sort of medical process as well as on the usual age and sex. For life assurance, this medical screening takes place before the issue of a policy.

The fact that such lives have passed the medical screening means that these lives will display lighter mortality than the general population which has not undergone selection by medical screening. However the effect of having passed the medical screening changes as the duration since the medical screening increases. The effect of the medical screening is often said to *'wear off'*.

Where temporary initial selection applies, we subdivide our data according to duration since initial selection (e.g. since entry to assurance) as well as according to age and sex, to ensure that we calculate mortality rates for groups of lives who have similar characteristics with respect to mortality".

The data available for analysis, as provided by the *CMI* Bureau, have been subdivided by duration *0, 1, 2, 3, 4 and 5+,* for each calendar year *1958 - 1990* separately and for individual ages in the range *23 - 62* years.

The aim of this analysis is to investigate if it is reasonable to pool the data by duration over the whole of the observation period, in much the same spirit as the data are pooled together by duration in Section *17* of Forfar et al (*1988*) for the limited observation period *1967 - 1970.*

Following Section *13.2.1*, first we need to examine the particular model structure for all the durations and calendar years in order to determine the optimum polynomial degree. That is, applying equation *(13.1)*, for each duration ($d = 0, 1, 2, 3, 4$), and each calendar year ($t = 1958, 1959, \ldots, 1990$), we obtain the following (flexible) model structure :

$$m^{d}_{x,t} = \sum_{i=0}^{k-1} \beta^{d}_{i,t} \cdot x^{i}$$

Primary work showed that the optimum number of parameters is 5. In order to prove this, Table 13.1 gives the p - values for the F - tests (formula 13.2), for each calendar year and for each duration d separately, after comparing the model structure (13.1) with either 5 or 6 parameters, based on the hypothesis I : H_0: $\beta^d_{5,t} = 0$ vs H_1: $\beta^d_{5,t} \neq 0$

Table 13.1 : p - values for $H_0 : \beta^d_{5,t} = 0$, for each d and t

Year	**duration0**	**duration1**	**duration2**	**duration3**	**duration4**
1958	**0.040**	0.534	1.000	0.300	**0.030**
1959	**0.042**	0.672	0.469	**0.018**	0.481
1960	0.095	0.864	0.407	**0.043**	0.690
1961	**0.004**	0.356	0.871	**0.008**	**0.010**
1962	0.202	0.514	0.451	1.000	0.476
1963	0.644	0.818	**0.005**	**0.049**	**0.029**
1964	0.728	0.877	0.329	**0.020**	0.487
1965	0.069	0.081	0.070	0.838	**0.038**
1966	0.813	0.199	0.220	0.618	0.059
1967	0.544	0.100	0.118	0.846	0.075
1968	0.109	0.155	0.095	0.459	0.254
1969	0.889	0.312	0.686	0.769	0.112
1970	1.000	0.422	0.626	**0.004**	**0.003**
1971	0.835	0.207	0.679	0.636	0.309
1972	1.000	0.793	0.802	0.215	0.418
1973	0.726	0.721	0.052	0.480	**0.022**
1974	0.855	0.171	**0.033**	0.755	0.924
1975	0.422	0.630	0.774	0.145	0.763
1976	0.475	0.653	0.171	0.596	0.646
1977	0.242	0.664	0.630	0.918	0.918
1978	0.312	0.268	0.144	0.818	1.000
1979	**0.037**	0.375	1.000	0.479	**0.027**
1980	0.504	0.051	0.103	0.507	0.086
1981	0.235	0.879	0.174	0.493	0.837
1982	0.643	0.374	0.085	0.695	0.909
1983	0.211	0.641	0.874	0.605	0.070
1984	0.432	0.571	1.000	0.625	0.483
1985	0.801	0.447	**0.014**	0.199	0.780
1986	0.814	0.798	**0.017**	0.543	0.731
1987	0.167	0.460	0.659	**0.049**	0.597
1988	0.069	0.090	**0.001**	0.274	0.420
1989	1.000	0.614	0.251	0.843	0.479
1990	0.421	0.241	1.000	1.000	0.804

Significant p - values at the 5% level of significance are highlighted by bold. Table 13.1 shows an acceptable range of p - values for all the durations, on the basis of the hypothesis that the optimum number of parameters is 5.

Thus the model utilised for the further analysis takes the specific form

$$m^{d}_{x,t} = \sum_{i=0}^{4} \beta^{d}_{i,t} \cdot x^{i} \qquad \textbf{(13.3)}$$

Now having determined the optimum degree of the polynomial, for each duration and calendar year concerned, we are interested in investigating whether the parameters $\beta^{d_1}_{i,t}$ and $\beta^{d_2}_{i,t}$ are equal for different choices of d_1 and d_2, based on the hypothesis *II*. Table *13.2* gives the corresponding p - values based on the following choices.

Table 13.2 : p - values for $H_0 : \beta^1_i = \beta^2_i$

Year/d₁ - d₂	***0 - 1***	***1 - 2***	***2 - 3***	***3 - 4***	***2 - 34***
1958	*0.041*	*0.067*	*0.014*	*0.502*	*0.014*
1959	***0.000***	*0.064*	*0.068*	*0.486*	*0.068*
1960	***0.000***	*0.128*	*0.081*	***0.036***	*0.081*
1961	***0.000***	*0.508*	***0.000***	*0.225*	***0.000***
1962	***0.000***	*0.551*	*0.524*	*0.499*	*0.524*
1963	*0.142*	*0.075*	*0.331*	*0.798*	*0.331*
1964	***0.010***	***0.007***	*0.109*	***0.009***	*0.109*
1965	*0.050*	*0.052*	*0.520*	*0.688*	*0.520*
1966	***0.000***	***0.033***	***0.000***	***0.002***	***0.000***
1967	***0.000***	*0.196*	*0.301*	***0.028***	*0.301*
1968	*0.096*	*0.052*	*0.285*	*0.444*	*0.285*
1969	***0.006***	*0.407*	*0.064*	*0.222*	*0.064*
1970	***0.004***	*0.055*	***0.000***	*0.550*	***0.000***
1971	***0.003***	*0.479*	*0.179*	***0.039***	*0.179*
1972	***0.000***	*0.177*	***0.003***	***0.040***	***0.003***
1973	***0.001***	*0.187*	***0.004***	*0.191*	***0.004***
1974	***0.005***	***0.005***	*0.445*	***0.049***	*0.445*
1975	*0.067*	*0.189*	*0.161*	*0.989*	*0.161*
1976	***0.003***	***0.002***	*0.057*	*0.193*	*0.057*
1977	***0.000***	*0.875*	*0.384*	*0.064*	*0.385*
1978	***0.000***	*0.788*	***0.017***	*0.750*	***0.017***
1979	***0.025***	*0.438*	*0.061*	***0.004***	*0.061*
1980	***0.030***	***0.043***	***0.012***	***0.003***	***0.012***
1981	***0.000***	*0.139*	***0.000***	*0.555*	***0.000***
1982	***0.000***	***0.000***	*0.794*	***0.027***	*0.792*
1983	*0.931*	***0.000***	*0.097*	*0.828*	*0.097*
1984	*0.451*	*0.133*	*0.741*	*0.532*	*0.741*
1985	***0.000***	***0.014***	*0.284*	*0.527*	*0.284*
1986	***0.015***	***0.000***	*0.177*	*0.408*	*0.177*
1987	***0.045***	*0.336*	***0.024***	*0.091*	***0.024***
1988	*0.075*	*0.136*	*0.311*	***0.003***	*0.311*
1989	***0.005***	*0.083*	***0.016***	*0.061*	***0.016***
1990	***0.005***	*0.270*	***0.013***	*0.036*	***0.013***

Significant p - values at the 5% level of significance are highlighted by bold. The results reported in Table *13.2* indicate that the mortality rates at duration *0* differ significantly from duration *1*, and from the other durations, that durations *3 & 4* can be grouped together and that duration *2* seems to be closer to *3* rather than to duration *1*. Therefore, a conservative view would be to retain durations *0, 1* and *2* separately and combine *3 - 4* together.

Another acceptable grouping would be to keep the durations *0, 1* separate and combine *2 - 3 - 4* together. This is exactly the practice adopted in construction of *A 1967 / 70* with *5* years select period, and *1979 - 1982* graduations *(CMI* Report No *9, 1988*).

13.3 Testing hypotheses of the form :

$$H_0 : \mu_x = f(x) \cdot \mu_x^s \quad \text{vs} \quad H_1 : \mu_x \neq f(x) \cdot \mu_x^s \qquad \forall x$$

13.3.1 Methodology

Recall Section *13.2* involving the normal approximation for the natural logarithm of the empirical values of the resistivity to death

$$Q_x = \log Y_x = \log\left(\frac{{}^{P}R_x^c}{P_x}\right) \approx N\left(\mathrm{m}_x, \frac{\sigma^2}{w_x}\right)$$

with scale parameter σ^2 and prior weights $w_x = P_x$, where

$$E(Q_x) = m_x = \log\frac{1}{\mu_x}$$

If by μ_x^s, we denote the central rates for a standard table and by μ_x, we denote the central rates for a second mortality experience, it follows, using an obvious notation, that we can write

$$Q_x^s = \log Y_x^s \approx N\left(m_x^s, \frac{\sigma_1^2}{w_x^s}\right) \qquad \& \qquad Q_x = \log Y_x \approx N\left(m_x, \frac{\sigma_2^2}{w_x}\right)$$

Then under the assumption of independence between the two mortality experiences, it follows that

$$Q_x^s - Q_x = \log Y_x^s - \log Y_x \approx N\left(m_x^s - m_x, \frac{\sigma_1^2}{w_x^s} + \frac{\sigma_2^2}{w_x}\right)$$

or

$$\log\frac{Y_x^S}{Y_x} \approx N\left(h(x), \frac{\sigma_1^2 \cdot w_x + \sigma_2^2 \cdot w_x^s}{w_x^s \cdot w_x}\right) \qquad \textbf{(13.4)}$$

where

$$h(x) = \log\frac{\mu_x}{\mu_x^S}$$

For reasons of simplicity, we further assume the same variance σ^2 for both mortality experiences, so that

$$\log \frac{Y_x^s}{Y_x} \approx N\left(h(x)\,,\ \sigma^2 \cdot \frac{w_x + w_x^s}{w_x^s \cdot w_x}\right) \tag{13.5}$$

By analogy with Section *5.6* (approximating the logarithm for the resistivity to death as a normal distribution), we have

$$E(\log \frac{Y_x^s}{Y_x}) = h(x) = \log \frac{\mu_x}{\mu_x^s} \qquad \& \qquad Var(\log \frac{Y_x^s}{Y_x}) = \sigma^2 \cdot \frac{1}{W_x}$$

with scale parameter $\varphi = \sigma^2$, prior weights $W_x = \dfrac{w_x^s \cdot w_x}{w_x + w_x^s}$, and variance function equal to *1*.

The function $h(x)$, which is the expected value for the modelling distribution *(13.5)*, satisfies the relationship

$$m_x = \exp\{h(x)\} \cdot \mu_x^s \tag{13.6}$$

It follows that when the modelling distribution *(13.5)* is used in combination with the identity link, then $h(x)$ becomes the linear predictor.

Specifically, when

$$h(x) = \alpha \qquad \text{then } (13.6) \text{ becomes } \mu_x = A \cdot \mu_x^s$$

and when

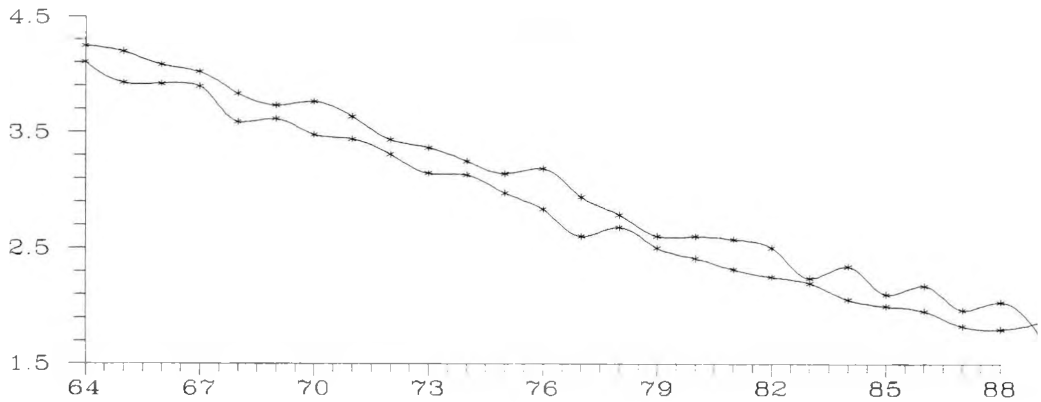
$$h(x) = \alpha + \beta x \quad \text{then } (13.6) \text{ becomes } \mu_x = A \cdot B^x \cdot \mu_x^s$$

In the remaining sections of this Chapter, we seek merely to investigate the feasibility of using these methods, without going into a detailed interpretation of any results. There are also various aspects of the method, still to be investigated, including a comparison with the approach of Currie and Waters (*1991*) for modelling the effects of select mortality.

13.3.2 Comparing male assured lives, duration 5+, and male pensioners mortality experience, year 1990, ages 64 - 89

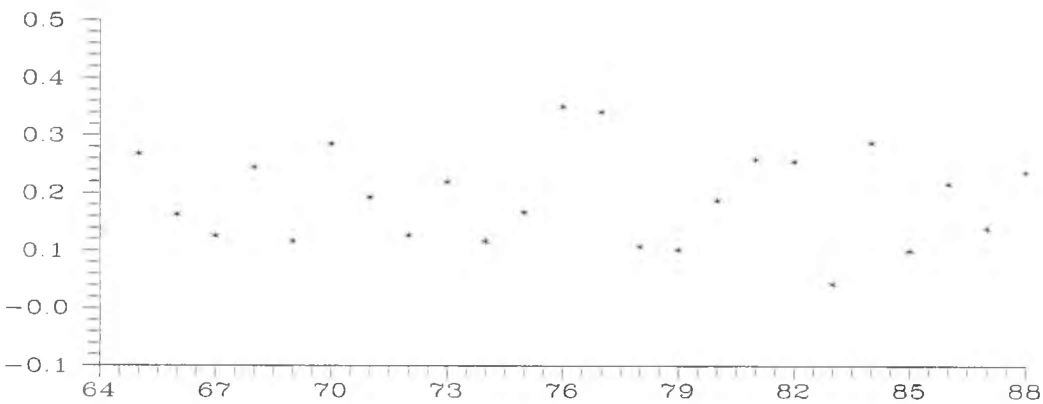
The empirical values of the resistivity to death, for both mortality experiences, are plotted against age in Figure *13.1*, with the upper curve representing the pensioners mortality experience.

Figure 13.1 : Resistivity to death vs age



The empirical responses, under the modelling assumptions *13.5*, plotted against age, are presented in Figure *13.2*.

Figure 13.2 : Model responses vs age



From Figure *13.2* it is seen that there is no particular trend in the responses, so that the function *h(x)* could potentially be modelled as a constant term, *h(x)* = *a*. This is verified by fitting the null model structure *h(x)* = *a* under the modelling assumption *(13.5)* using *GLIM*, leading to the following results

*Parameter estimate (*and *standard error)*

$a = 0.1863\ (0.01876)$

The deviance is *33.685* on *25* degrees of freedom, with scale parameter $\hat{\sigma}^2 = 1.347$.

The introduction of a second parameter, dependent on age, using the linear predictor $h(x) = a + \beta \cdot x$, proved to be insignificant.

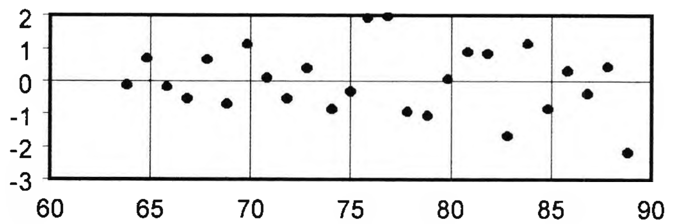
The p - values for the statistical tests based on the residuals (indicating an excellent fit) are as follows

Statistical tests : p - values

$p_{ISD} = 97 \qquad p_{sign} = 50 \qquad p_{runs} = 98 \qquad p_{chi} = 46.$

Figure *13.3*, displays the plot of deviance residuals against age, which does not show any abnormalities.

Figure 13.3 : Standardised deviance residuals against age



Thus, the male pensioners' mortality experience can be represented, in terms of the male assured lives mortality experience, by following relationship

$$\log(\mu_x) = 0.1863 + \log(\mu_x^s) \qquad \text{or} \qquad \mu_x = 1.20478 \cdot \mu_x^s$$

where μ_x *and* μ_x^s denote the force of mortality for male pensioners and for male assured lives, respectively.

The fidelity that the force of mortality for male life office pensioners is greater than that for assured lives (age for age) is as expected given the effect of selection, and has been confirmed by analyses carried out, from time to time, by the *CMI* Bureau.

13.3.3 Comparing male assured lives, grouped duration 3 - 4 with duration 5+, period 1958 - 1990, ages 23 - 62

In this Section, the construction of a model structure to represent the mortality experience for the grouped duration *3 - 4* based on the mortality experience of duration *5+* is attempted, for the period *1958 - 1990* and ages *23 - 62*.

That is, using the mathematical model structures derived in Chapters *VIII, IX, X* and employing the methodology described in Section *13.3.1*, we can construct a mathematical model structure for the grouped duration *3 - 4*, for each calendar year *1958 - 1990* and for the range of ages *23 - 62*.

The hypothesis to be tested takes the form

$$H_0 : \mu_{x,t}^{d34} = \exp\{h(x,t)\} \cdot \mu_{x,t}^{d5+} \quad vs \quad H_1 : \mu_{x,t}^{d34} \neq \exp\{h(x,t)\} \cdot \mu_{x,t}^{d5+} \qquad \forall\, x, t$$

Various linear predictor structures $h(x,t)$ have been investigated (additive in age and time effects) and the following structure is proposed following the usual exploratory analysis

$$h(x,t) = a + b \cdot x + c \cdot x^2 + d \cdot t \qquad (13.7)$$

Table *13.3* displays the estimates of the parameters, the standard errors, and the t - tests.

Table 13.3 : Parameter estimates, standard errors, & t - tests, model (13.7)

	p.e.	s.e.	t - test
a	-0.57699	0.1482	-3.89
b	0.004709	0.0008688	5.42
c	0.0265466	0.006811	3.89
d	-0.0003655	0.0000762	-4.79

Table *13.4* gives p - values based on the residuals under model *(13.7)*, which reveals an adequate fit. The residual plots are not reproduced.

Table 13.4 : p - values, model (13.7)

Year	***ISD***	***Sign***	***Runs***	***Chi***
1958	66	82	96	74
1959	12	62	38	6
1960	29	3	22	64
1961	79	89	61	43
1962	17	82	20	82
1963	87	37	51	23
1964	69	37	84	47
1965	24	3	74	68
1966	77	10	90	70
1967	80	62	17	76
1968	95	62	90	61
1969	**0**	**0**	51	**0**
1970	95	50	74	24
1971	57	37	90	65
1972	67	73	40	36
1973	62	73	11	31
1974	41	3	13	17
1975	29	10	73	42
1976	58	17	68	31
1977	11	**0**	89	25
1978	68	26	53	13
1979	**4**	**1**	57	16
1980	5	**0**	80	15
1981	95	50	17	12
1982	94	17	43	20
1983	35	10	23	52
1984	68	17	31	24
1985	12	**1**	70	33
1986	59	37	17	89
1987	89	82	56	10
1988	55	73	28	55
1989	32	73	76	65
1990	72	89	14	16

Therefore, if we choose as a standard table for duration $5+$ the mathematical expression (8.11), which involves the log link function in combination with a quadratic spline function, then the construction of mortality table(s) for duration 3 - 4 can be based on the formula

$$\mu_{x,t}^{d34} = \exp\{g(x,t)\} \cdot \mu_{x,t}^{d5+}$$

That is,

$$\mu_{x,t}^{d34} = \exp\{a + b \cdot x + c \cdot x^2 + d \cdot t\}$$
$$\cdot \exp\{\alpha + (\beta_1 + \beta_2 \cdot t^{1.8}) \cdot x + (\gamma_1 + \gamma_2 \cdot t^{1.8}) \cdot x^2 + \delta \cdot (x - 42)_+^2\}$$

where the parameter estimates are given in Table 13.3 for a, b, c & d and in Table 8.9 for α, β_1, β_2, γ_1, γ_2 & δ.

13.3.4 Comparing male assured lives, individual durations 0, 1, 2 with grouped duration 3 - 4, period 1958 - 1990, ages 23 - 62

As in the previous two sections, we limit the investigation to an examination of the feasibility of the methodology. The hypotheses to be tested this time take the form

$$H_0 : \mu^{d2}_{x,t} = \exp\{h_2(x,t)\} \cdot \mu^{d34}_{x,t} \quad \text{vs} \quad H_1 : \mu^{d2}_{x,t} \neq \exp\{h_2(x,t)\} \cdot \mu^{d34}_{x,t} \qquad \forall\, x,t \qquad \textbf{(13.8)}$$

$$H_0 : \mu^{d1}_{x,t} = \exp\{h_1(x,t)\} \cdot \mu^{d34}_{x,t} \quad \text{vs} \quad H_1 : \mu^{d1}_{x,t} \neq \exp\{h_1(x,t)\} \cdot \mu^{d34}_{x,t} \qquad \forall\, x,t \qquad \textbf{(13.9)}$$

$$H_0 : \mu^{d0}_{x,t} = \exp\{h_0(x,t)\}\ \mu^{d34}_{x,t} \quad \text{vs} \quad H_1 : \mu^{d0}_{x,t} \neq \exp\{h_0(x,t)\} \cdot \mu^{d34}_{x,t} \qquad \forall\, x,t \qquad \textbf{(13.10)}$$

where $\mu^{d0}_{x,t}$, $\mu^{d1}_{x,t}$, $\mu^{d2}_{x,t}$ & $\mu^{d34}_{x,t}$ denote the force of mortality for duration *0, 1, 2* and grouped duration *3 - 4* respectively.

Exploratory analysis using *GLIM*, based on the linear predictor $h_d(x,t)$ in association with the identity link, indicates the null structure

$$h_d(x,t) \equiv \alpha_d \qquad d = 0, 1, 2$$

for all three sets of hypotheses *(13.8)*, *(13.9)* & *(13.10)*.

The following Table *13.5* displays the parameters estimates, the standard errors, and the *t* - tests for the model structures *(13.8)*, *(13.9)* & *(13.10)*.

Table 13.5 : Parameter estimates, standard errors, & *t* - tests, models *(13.8)*, *(13.9)* & *(13.10)*

	p.e.	***s.e.***	***t - test***
α_2	-0.05131	0.01048	-4.89
α_1	-0.1098	0.01107	-9.91
α_0	-0.3102	0.01238	-25.05

This simplification means that durations *0, 1, 2* and grouped duration *3 - 4*, have similar mortality shapes. We need only to subtract a constant value on the log scale to move from one duration to another, for each calendar year *1958 - 1990*. Also the magnitudes of the estimated parameters are such that mortality increases progressively, but at a slower rate, with increasing duration, for each fixed x and t.

The t - test for duration *2* (t - value = - *4.89*) means that the parameter a_2 is significant and duration *2* can be considered different from the grouped durations *3 - 4*. Therefore, we can conclude that duration 2 can be modelled independently from the grouped durations *3 - 4*.

Tables *13.6, 13.7 & 13.8* give the p - values for the statistical tests based on residuals for models *(13.8)*, *(13.9) & (13.10)*, all of which reveal satisfactory fits.

Therefore, following up the discussion from the previous section, if we choose as a standard table the mathematical expression *(8.10)* for duration *5+*, based on the log link function in combination with a quadratic spline predictor, then the construction of mortality table(s) for duration *0, 1, 2 & 3 - 4* are based on the formula

$$\mu^{d}_{x,t} = \exp\{a_d + b \cdot x + c \cdot x^2 + d \cdot t\}$$

$$\cdot \exp\{\alpha + (\beta_1 + \beta_2 \cdot t^{1.8}) \cdot x + (\gamma_1 + \gamma_2 \cdot t^{1.8}) \cdot x^2 + \delta \cdot (x - 42)^2_+\} \qquad \textbf{(13.11)}$$

where the parameter estimates are given in Table *13.3* for a, b, c & d, in Table *13.5* for a_0, a_1 & a_2 and in Table *8.9* for α, β_1, β_2, γ_1, γ_2 & δ.

Table 13.6 : *p* - values, formal graduation tests for each calendar year separately, model *(13.8)*

Year	*ISD*	*Sign*	*Runs*	*Chi*
1958	61	26	39	39
1959	90	62	17	16
1960	91	73	28	35
1961	43	26	**3**	77
1962	98	50	50	69
1963	84	73	39	91
1964	49	37	17	54
1965	40	89	14	23
1966	57	89	47	8
1967	66	17	20	60
1968	82	50	16	96
1969	31	94	67	28
1970	73	10	14	40
1971	78	10	60	33
1972	70	10	60	47
1973	48	26	10	8
1974	86	82	42	**4**
1975	19	**1**	42	67
1976	97	26	52	38
1977	60	89	34	70
1978	53	10	89	29
1979	86	50	10	51
1980	67	94	53	23
1981	55	37	6	17
1982	36	37	74	19
1983	42	62	50	78
1984	99	37	63	58
1985	91	82	42	45
1986	13	82	55	90
1987	90	26	52	68
1988	61	82	30	84
1989	66	62	97	13
1990	86	26	85	39

Table 13.7 : *p* - values, formal graduation tests for each calendar year separately, model *(13.9)*

Year	*ISD*	*Sign*	*Runs*	*Chi*
1958	87	26	65	91
1959	**0**	62	26	7
1960	92	73	**3**	14
1961	31	17	12	12
1962	42	50	10	79
1963	88	17	30	25
1964	97	26	65	81
1965	86	37	5	79
1966	93	73	65	83
1967	73	10	82	72
1968	81	50	83	41
1969	98	50	26	25
1970	18	3	91	20
1971	80	17	12	29
1972	8	17	12	22
1973	57	5	**1**	62
1974	97	62	17	55
1975	42	3	83	57
1976	42	94	10	59
1977	86	82	78	20
1978	77	82	78	63
1979	97	50	10	64
1980	8	89	14	49
1981	95	50	62	**5**
1982	69	62	5	39
1983	40	17	20	16
1984	96	37	17	21
1985	**0**	**99**	66	**2**
1986	64	50	5	76
1987	84	50	73	64
1988	16	26	28	98
1989	**0**	82	30	68
1990	38	3	21	66

Table 13.8 : p - values, formal graduation tests for each calendar year separately, model (13.10)

Year	ISD	Sign	Runs	Chi
1958	99	50	6	55
1959	61	26	39	43
1960	76	10	14	55
1961	**4**	**1**	89	38
1962	19	62	50	34
1963	95	50	90	34
1964	93	50	50	55
1965	93	62	38	80
1966	41	10	34	57
1967	17	6	67	13
1968	32	62	10	12
1969	33	50	16	13
1970	18	**1**	**3**	49
1971	94	26	10	51
1972	**2**	**0**	35	**4**
1973	36	10	8	7
1974	13	73	39	48
1975	52	50	26	90
1976	64	50	37	86
1977	86	82	92	17
1978	84	26	76	51
1979	56	73	18	9
1980	39	82	30	22
1981	53	10	72	55
1982	97	62	38	35
1983	92	82	20	32
1984	51	94	86	71
1985	22	62	63	11
1986	**2**	82	68	93
1987	55	37	6	85
1988	98	37	10	59
1989	55	62	38	78
1990	10	89	47	90

CHAPTER XIV

Conclusions

In part I, a method for the graduation, analysis and modelling of mortality trends has been defined. This method was developed using the theory of *Generalised Linear Model's (GLMs).* In this respect *GLMs* are seen to be a beneficial tool, providing a sufficient statistical foundation for the modelling of mortality rates, a wide class of mathematical model structures, and an extensive range of diagnostic checks for confirming the plausibility of any applied model.

The statistical tests (the individual standardised deviation test, the sign test, the runs test, and the chi - square test) are based on the standardised deviance residuals, provided by *GLM's*, for any error assumption, and are complemented by analyses of the residual plots. Therefore, *GLM's* give a comprehensive framework for the statistical analysis with the potential for comparison among different model structures.

In part II, the Poisson process is confirmed to be the basis for the statistical modelling of the central mortality rates, in combination with the properties and generalisations suggested.

Emphasis is given to the gamma distribution model for the inverse of the central mortality rates (called by Gompertz the resistivity to death). Based on the gamma error for the resistivity to death we have derived the normal error distribution for the natural logarithm of the resistivity to death.

All the error distributions derived in Chapter *V*, in association with the central mortality rates, i.e. the Poisson (Section *5.2*) - gamma (Section *5.3*), compound Poisson (Section *5.4*) - gamma (Section *5.5*) and normal error structure (Section *5.6*), differ to an insignificant extent as illustrated in section *5.7*. More specifically, the estimates of the parameters differ to an insignificant extent and the deviance (residuals), in association with the log link function, are identical with the Poisson model and the compound Poisson model. The same characteristics are obtained when employing the power link function.

Initial mortality rates are associated with the (over - dispersed) binomial law distribution (Section *6.2*). Assuming a log link for the central rates, the canonical link function for central rates, the initial rates correspond to the complementary log - log model structure (Section *11.2.3*).

Sverdrup (*1965*) argues that there is a real loss in information by disregarding the waiting time (Section *1.2*). Therefore, it would be desirable for mortality investigations to be accomplished using the central exposed to risk and employing the associated techniques described in Chapter *I*.

In part III, a method for the construction of a mathematical models for age specific mortality trends through time is described. The method can be extended when more factors of mortality are involved (Section *7.2*).

This method gives various mathematical expressions for mortality trends when employing the multiplicative model (Chapter *VIII*), the power model (Chapter *IX*), and the additive model (Chapter *X*) for male assured lives data, or the log - log model structure for pensioners data (Chapter *XI*).

The construction of a mathematical formula with independent variables age and time can be of considerable importance to insurance companies, when taking account of the change in mortality through time (in addition to age effects). This consideration is more important for pensioners' and annuitants' portfolios, since the (expected) improvement of mortality requires an increase in the level of the premiums and consequently of the mathematical reserves.

The log link function is deemed to be the most acceptable choice for the link for the central mortality rates, justified by the smooth progression imparted to the mortality trends when the log transformation is applied. For male assured lives, duration *5+*, the log link gives the minimum deviance, in association with a polynomial predictor structure, in age and time effects, where *6* parameters are needed for each calendar year (Section *8.2.2*, model *8.4*). Further, in association with a quadratic spline predictor structure in age effects and a fractional polynomial predictor structure in time effects, a flexible model is produced with a parsimonious number of parameters. The knots are located at the age points where the mortality curve changes curvature (Section *8.2.3*, model *8.12*). In both cases, the detailed statistical results are acceptable.

The *power(2)* model structure gives the least number of parameters (equal to *5*), for male assured lives, duration *5+*, in association with the highest deviance, when employing a polynomial predictor structure in age effects and a fractional polynomial predictor structure in time effects (Section *9.2.3*, model *9.5*). Also, when employing the power model structure in association with a quadratic polynomial predictor structure, in age and time effects, we obtain a parsimonious number of parameters (*4*) for each calendar year in question (Section *9.2.2*, model *9.3*). Despite the lack of a theoretical justification for the choice of the power link function, the results produced are worthy of note.

The *additive* model produces sound results, for male assured lives, duration *5+*, when it is associated with cubic spline functions in age effects and a fractional polynomial structure in time effects. The knots are located at the age points *47* and *64* for each calendar year (Section *10.2.2*, model *10.4*).

Further, a different perspective, of the above approaches is exercised, by discussing mortality trends through time, for each age in question as regards the multiplicative model structure (Section *8.2.4*), the power model structure (Section *9.2.4*), and the additive model structure (Section *10.2.2*).

Now, focusing on the range of ages *[42, 89]* we have derived some simple mathematical expressions in association with the multiplicative and power model structures. Especially for the multiplicative model, it seems that there exists a critical point in the neighbourhood of the age of *42*, where the mortality 'development' changes curvature, according to the principle of local description in Section *2.1*. This feature is imparted to the power model structures as well. For the multiplicative model, a simple model structure is derived using a fractional polynomial structure in both age and time effects (Section *8.3*, model *8.20*). For the power model, again a simple model structure is presented, using a fractional polynomial structure in time effects and a polynomial predictor structure in age effects (Section *9.3*, model *9.13*).

In Chapter *XI*, the Complementary log - log model is applied for modelling pensioners, ages *60 - 95*, time period *1983 - 1990*, using a polynomial structure in time effects and an inverse polynomial predictor structure in age effects (Section *11.2.2*, model *11.2*).

By way of comparison we illustrate the impact of the age specific trend adjustment on male assures lives' mortality rates and we plot the fitted force of mortality for the time period *1958* to *1990* and the predicted force of mortality for the time period *1990* to *2010* against calendar year at *5* yearly age intervals in the following graphs, Figures *14.1* - *14.13*.

Figure *14.1* : Crude and predicted - forecasting force of mortality vs. calendar year, based on model structures *(8.14)* - multiplicative model, *(9.3)* - power(*1*) model, *(9.6)* - power(*2*) model & *(10.4)* - additive model, age *25* years

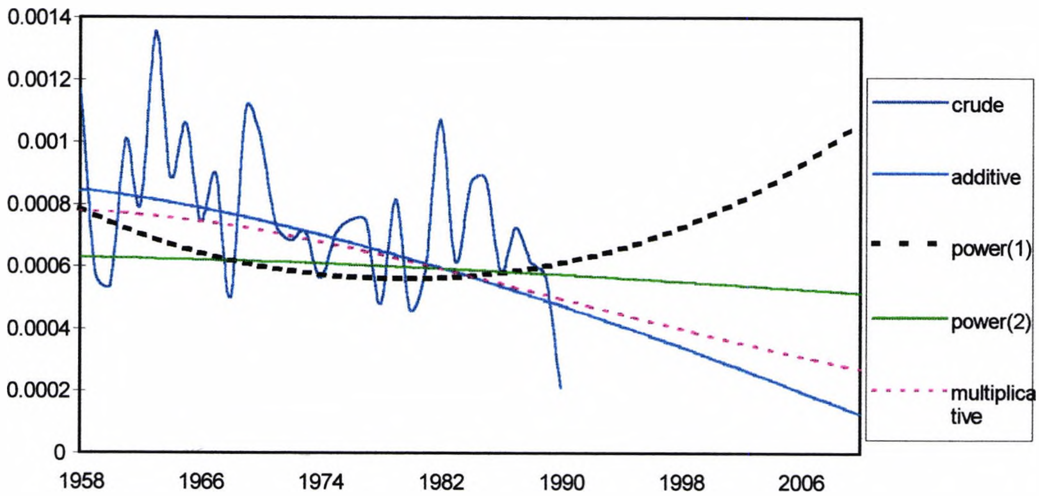


Figure *14.2* : Crude and predicted - forecasting force of mortality vs. calendar year, based on model structures *(8.14)* - multiplicative model, *(9.3)* - power(*1*) model, *(9.6)* - power(*2*) model & *(10.4)* - additive model, age *30* years

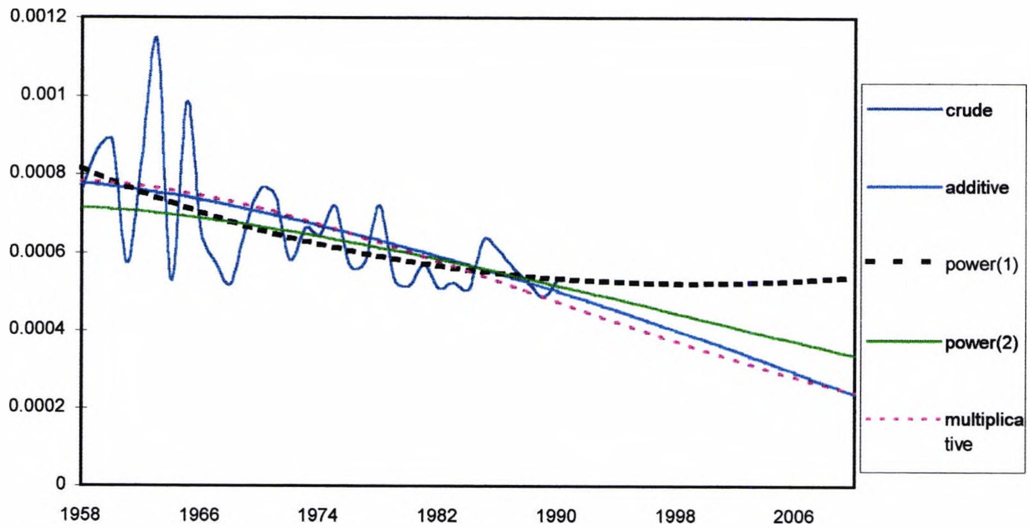


Figure *14.3* : Crude and predicted - forecasting force of mortality vs. calendar year, based on model structures *(8.14)* - multiplicative model, *(9.3)* - power(*1*) model, *(9.6)* - power(*2*) model & *(10.4)* - additive model, age *35* years

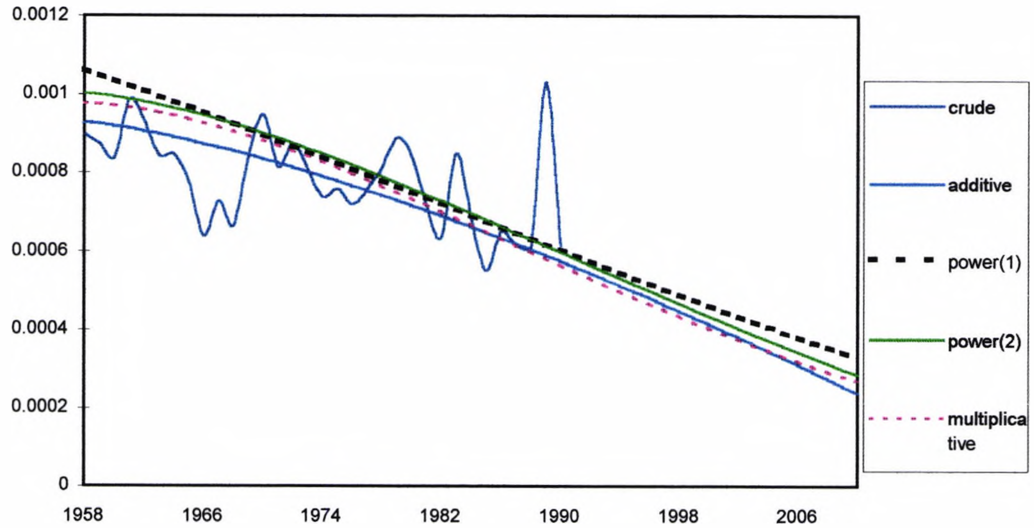


Figure *14.4* : Crude and predicted - forecasting force of mortality vs. calendar year, based on model structures *(8.14)* - multiplicative model, *(9.3)* - power(*1*) model, *(9.6)* - power(*2*) model & *(10.4)* - additive model, age *40* years

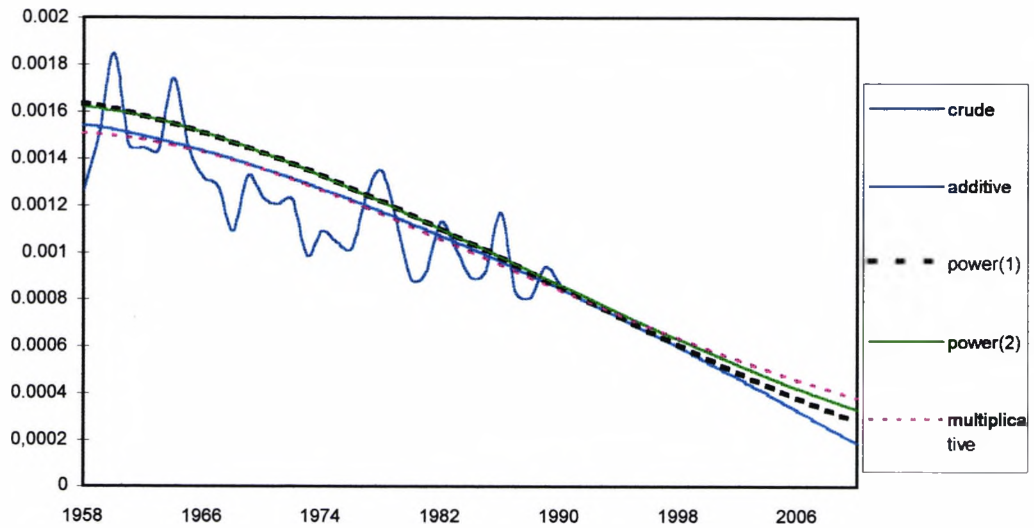


Figure *14.5* : Crude and predicted - forecasting force of mortality vs. calendar year, based on model structures *(8.14)* - multiplicative model, *(9.3)* - power(*1*) model, *(9.6)* - power(*2*) model & *(10.4)* - additive model, age *45* years

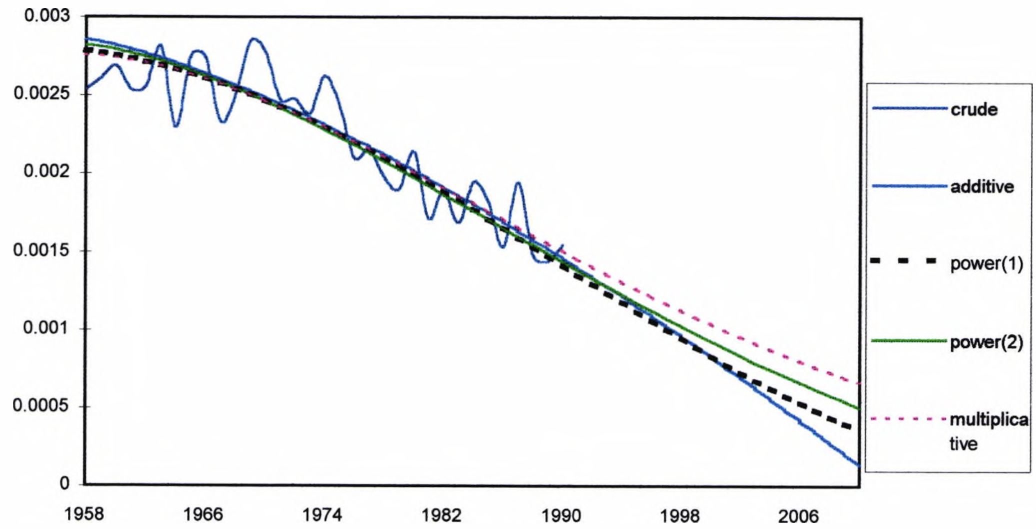


Figure *14.6* : Crude and predicted-forecasting force of mortality vs. calendar year, based on model structures *(8.14)* - multiplicative model, *(9.3)* - power(*1*) model, *(9.6)* - power(*2*) model & *(10.4)* - additive model, age *50* years

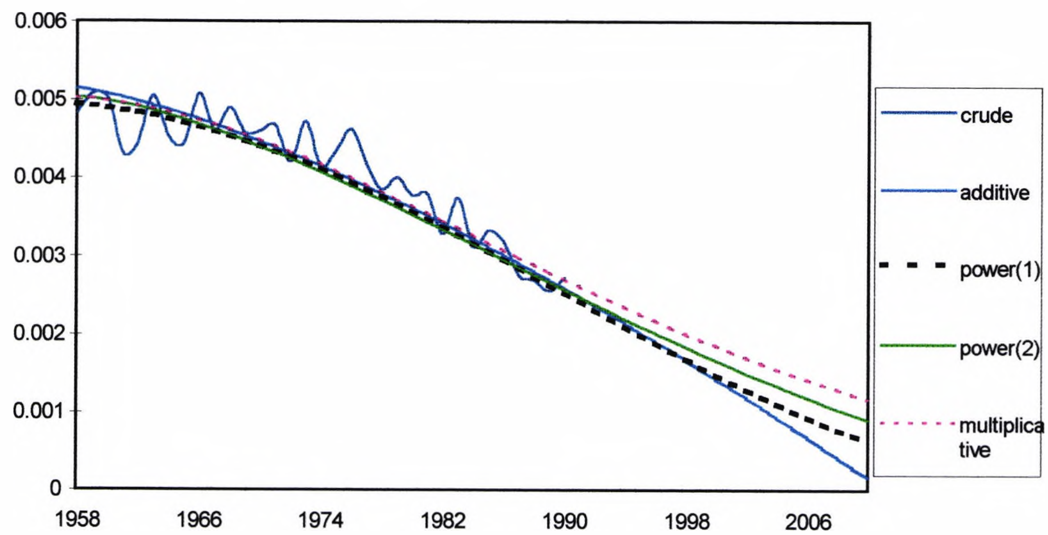


Figure *14.7* : Crude and predicted-forecasting force of mortality vs. calendar year, based on model structures *(8.14)* - multiplicative model, *(9.3)* - power(*1*) model, *(9.6)* - power(*2*) model & *(10.4)* - additive model, age *55* years

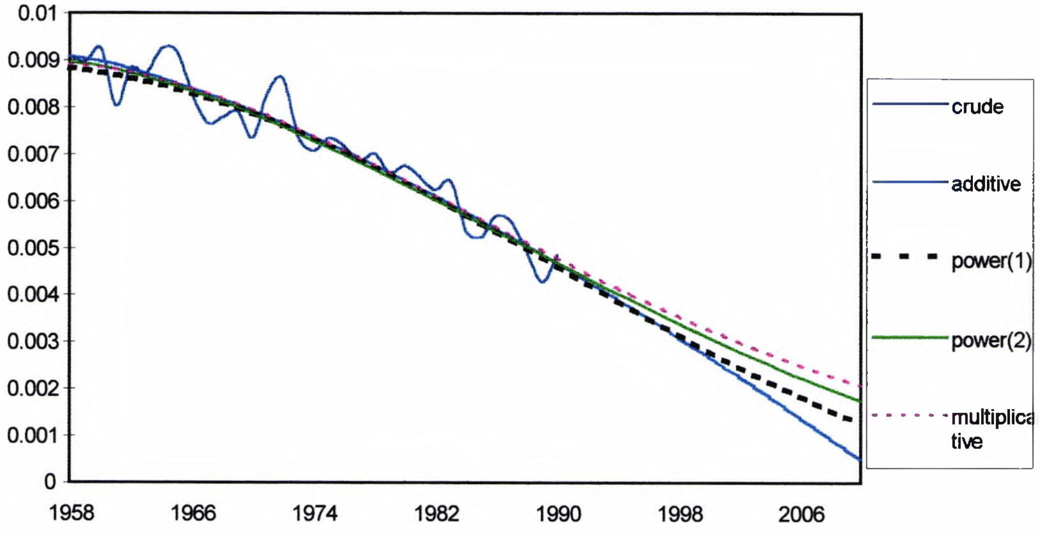


Figure *14.8* : Crude and predicted - forecasting force of mortality vs. calendar year, based on model structures *(8.14)* - multiplicative model, *(9.3)* - power(*1*) model, *(9.6)* - power(*2*) model & *(10.4)* - additive model, age *60* years

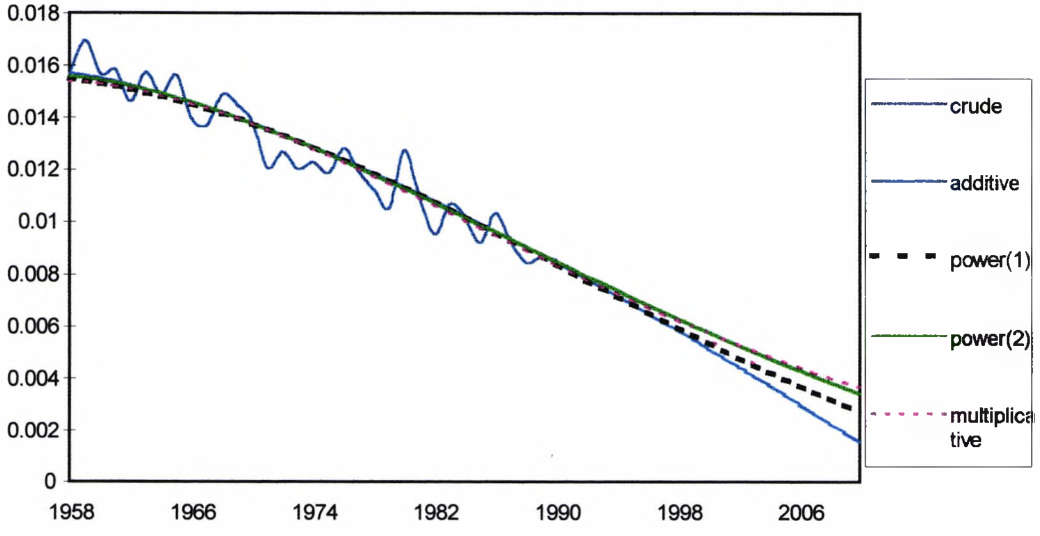


Figure *14.9 :* Crude and predicted - forecasting force of mortality vs. calendar year, based on model structures *(8.14)* - multiplicative model, *(9.3)*-power(*1*) model, *(9.6)* - power(*2*) model & *(10.4)* - additive model, age *65* years

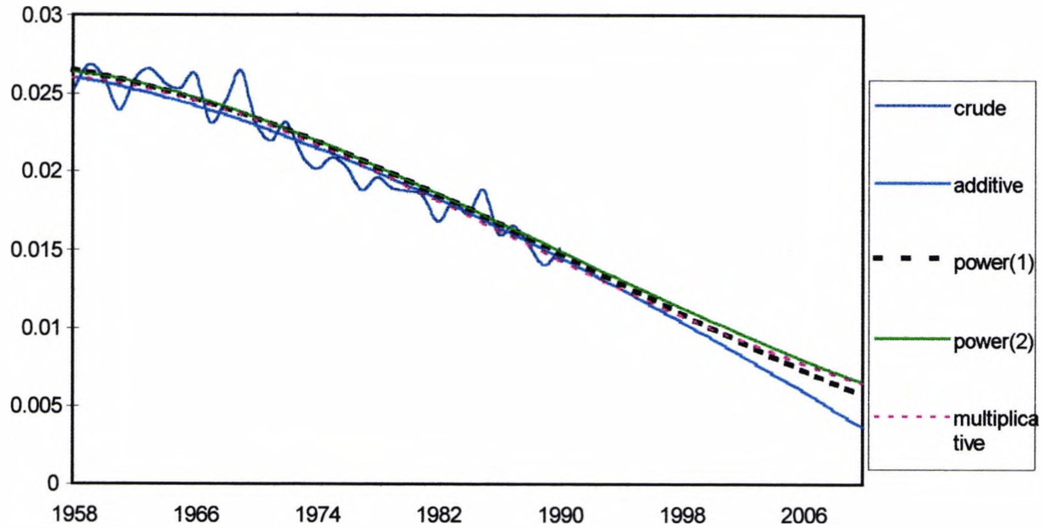


Figure *14.10 :* Crude and predicted - forecasting force of mortality vs. calendar year, ased on model structures (8.14) - multiplicative model, *(9.3)* - power(*1*) model, *(9.6)* - power(*2*) model & *(10.4)* - additive model, age *70* years

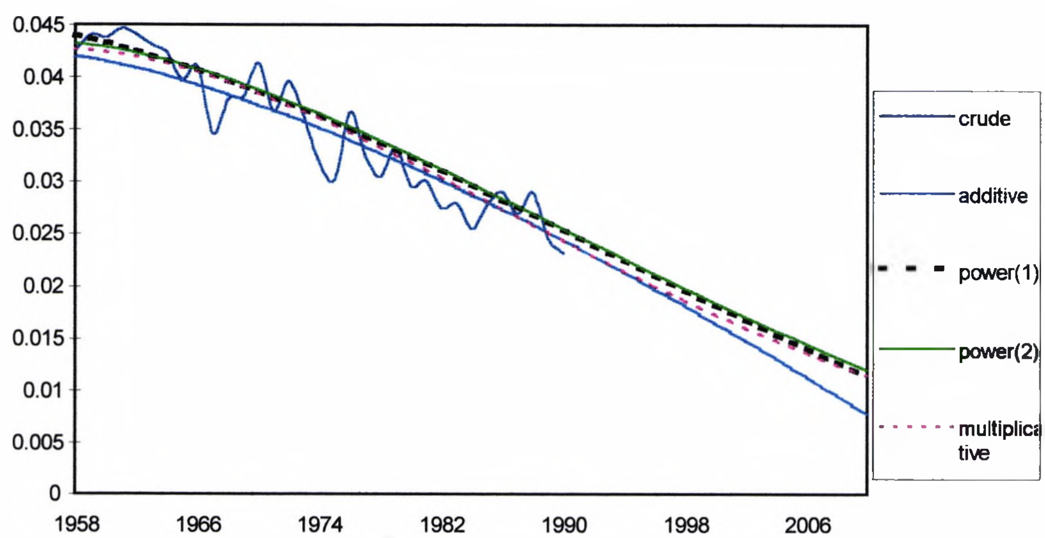


Figure *14.11* : Crude and predicted - forecasting force of mortality vs. calendar year, based on model structures (*8.14*) - multiplicative model, *(9.3)* - power(*1*) model, *(9.6)* - power(*2*) model & *(10.4)* - additive model, age *75* years

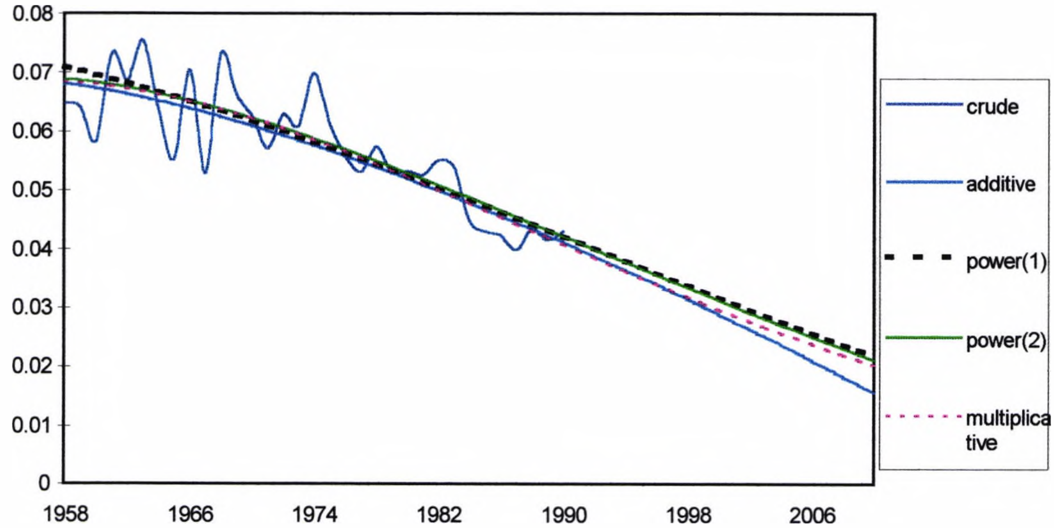


Figure *14.12* : Crude and predicted - forecasting force of mortality vs. calendar year, based on model structures (*8.14*) - multiplicative model, *(9.3)* - power(*1*) model, *(9.6)* - power(*2*) model & *(10.4)* - additive model, age *80* years

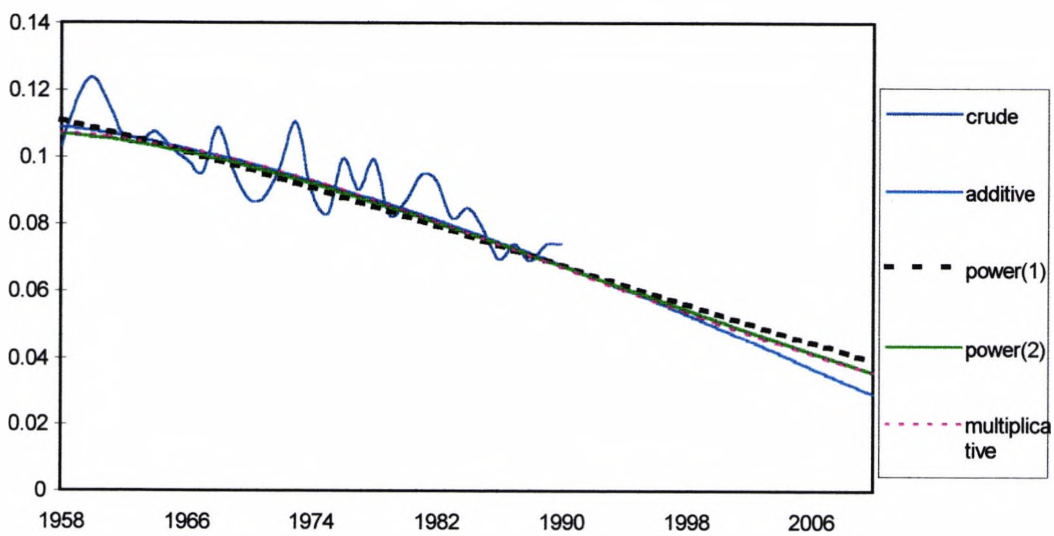
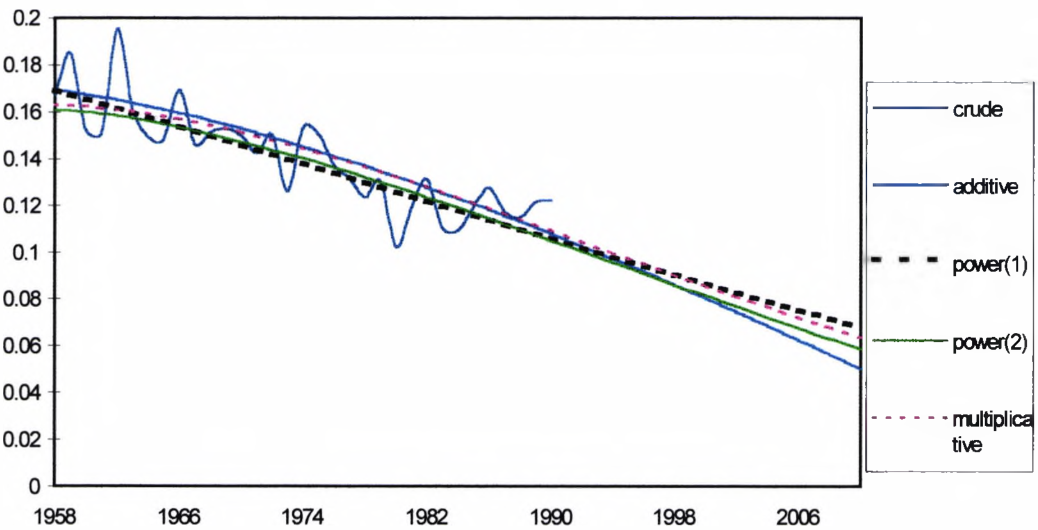


Figure *14.13* : Crude and predicted - forecasting force of mortality vs. calendar year, based on model structures *(8.14)* - multiplicative model, *(9.3)* - power(*1*) model, *(9.6)* - power(*2*) model & *(10.4)* - additive model, age *85* years



As is shown in the above graphs, in the observed time period *(1958 - 1990),* the models do not differ greatly compared with the differences that occur in the forecasting time period *(1990 - 2010).*

The poor goodness of fit, for the first ages (see for example Figure *14.1*), is granted to the high level of 'noise' in data.

All the graphs show a general decline in mortality rates except for the first ages *(24* to *30)* under the *power(1)* model structure where the force of mortality increases. This seems to be the result of the constant power parameter having somewhat less flexibility, in association with the parsimonious number of parameters employed. Both the *power* model structures show higher predicted mortality rates for the ranges of ages [*25, 40*] & [*65, 80*].

The *additive* model shows lower predicted mortality rates, and further it shows a faster decrease of mortality along time, for almost all the ages in question. Despite the fact that both the multiplicative and the power model structures reveal that the predicted mortality curves change their curvature during the time period involved, for all the ages in question, the additive model structure does not encompass this feature. Consequently, for the extrapolation of the mortality rates, based on this model structure, special conditions are needed, such as, for example, the presumption of the rapid decrease of the future mortality rates, in favour of the level of the premium or in favour of the level of the mathematical reserves.

The most conservative decline in predicted mortality rates seems to be for the *multiplicative* model, particularly for the range of ages [*40, 65*].

Further, it can be concluded that the rate of the mortality decrease reaches its maximum during the decade of *1980*'s for the range of ages *35* to *80*, and during the decade of *1990*'s for the ages above the age of *80*. This means that, on the basis of these models, there is expected to be a faster improvement in mortality for ages above *80* during the *1990*'s (Section *8.2.4*, Figure *8.15* & Section *9.2.4*, Figure *9.13*).

In all the model structures *(Multiplicative, Power(2), Additive)*, for the male assured lives mortality experience, the linear predictor is modelled satisfactorily by fractional polynomials in time of the form

$$\eta_{x,t} = \alpha_x + \beta_x \cdot t^k$$

for all the ages in question.

The Multiplicative model leads to the value $k = 1.8$ (Section *8.2.3*, model *8.12*), the Power(*2*) model the value $k = 1.6$, in association with the power link $p = 0.36$ (Section *9.2.3*, model *9.5*), and the Additive model the value $k = 1.4$ (Section *10.2.2*, model *10.4*). This suggest that fractional polynomials of the above form contain sufficient information needed for the mathematical modelling of the mortality trends in time effects (and this for all the available range of ages), in association with a parsimonious number of parameters. Anson (*1988*) argues that a two - dimensional mortality space is sufficient to represent the similarities and differences among human life tables, namely, the level of mortality (the rapidity with which mortality events occur, and hence in the longevity of the population), and its relative shape (the distribution of deaths at various ages). The structure of these fractional polynomials models justifies Anson's argument, since only two parameters differentiate the mortality experience among different calendar years.

In Chapter *XII*, on the modelling of amounts, the approach developed for the graduation of 'amounts' pays more attention to the intrinsic structure of the data than the approach currently advocated by the *CMI* Bureau. This approach provides some insight into the patterns of the claims amounts and of the modelling assumptions (while the *CMI* practice is simply to transform the data by dividing both the number of deaths and exposures by so - called variance ratios, before graduation proceeds). The methodology is strongly connected with the earlier work of Currie & Waters (*1991*) and of Renshaw (*1992*) on duplicate policies where the effects on the graduation approach are modelled through over - dispersion.

In part IV, Chapter *XIII*, duration is further classified in durations *0, 1, 2, 3 - 4 & 5+* for the male assured lives data. Chapter *XIII* describes the methodology for comparing mortality experiences and for constructing graduated mortality tables based on given standard tables. The analysis shows that durations *0, 1, 2 & 3 - 4* have similar mortality shapes, on the log scale, and that they are separated by the addition of a constant term, on the log scale, independent of age and time. Further, durations *0, 1, 2 & 3 - 4* can be constructed based on the mortality experience of duration *5+* by a simple mathematical formula (Section *13.3.4*).

Similar results are obtained when comparing male assured lives, duration *5+*, with pensioners mortality experience, for the calendar year *1990*. The results indicate that the mortality experiences have similar relative shapes, on the log scale, and the only difference that exists is in the terms of their levels (Section *13.3.2*).

Also, comparisons between pensioners and assured lives for durations *0, 1, 2 & 3 - 4* and assured lives for duration *5+* (taken as the given standard mortality experience), give a simple mathematical model structures for the construction of the mortality tables for pensioners and assured lives for durations *0, 1, 2 & 3 - 4* (Section *13.3.4*, model *13.11*).

The advantage of adopting this approach (rather than to model, in age and time effects, the data separately) depends on the fact that the age and (forecasting) time range for the constructed mortality tables can be extended beyond the (possible) confined ranges of age and time for the crude data alone but of course only as far as the standard mortality experiences's age and time ranges allow. As an example, this could be important for pensioners' and annuitants' mortality tables (since the mortality experience is restricted in the time period *1983 - 1990* and age range *60 - 90* years), if we advocate the above methodology and we construct mortality tables in age and time effects based on the mortality experience of the assured lives at durations *5+*.

CHAPTER XV

Appendix A

In the following Tables (*15.1 - 15.7*) the data (central exposed to risk based on policies - policy totals ceasing through death) for male assured lives, duration *5+*, ages *24 - 89*, for each calendar year (*1958 - 1990*) separately, are presented, as published by the *CMI* Bureau of the Institute and Faculty of Actuaries.

Table 15.1 : Central exposed to risk, for male assured lives, duration 5+, based on policies - policy totals ceasing through death, ages *24 - 89*, calendar year *1958 - 1962*

	1958		1959		1960		1961		1962	
24	5637.0	4	5983.5	10	6305.0	6	6772.0	3	7789.5	12
25	6857.5	8	6985.0	4	7400.5	4	7947.5	8	8806.0	7
26	10749.0	7	10445.0	7	10880.5	4	11602.5	11	12335.0	5
27	15506.5	13	15151.5	10	15060.5	18	15898.0	19	17233.0	12
28	19644.0	22	19701.0	13	19572.0	17	19849.0	15	21301.5	10
29	24778.5	21	24601.0	14	25075.0	19	25439.5	14	26145.5	23
30	30104.0	23	30018.0	26	30471.5	27	31446.5	18	32142.5	27
31	35299.0	35	35266.5	27	35825.5	32	36828.5	31	38217.0	26
32	40176.5	22	40380.5	38	40904.5	44	42024.5	47	43412.5	25
33	44063.0	39	45243.0	50	45993.5	39	46842.5	33	48336.5	37
34	47489.0	49	49026.5	50	50808.0	41	51816.5	47	52946.5	53
35	51095.0	46	52625.5	46	54754.0	46	56699.5	56	57843.0	54
36	55451.0	68	56021.5	73	58080.5	67	60497.0	59	62631.5	38
37	61963.0	67	60080.5	76	61110.5	61	63346.0	79	65836.5	68
38	67680.0	76	66628.0	73	65043.0	98	66253.0	98	68485.0	77
39	62296.5	71	72305.0	109	71679.0	109	70043.5	106	71145.5	96
40	55168.5	70	65551.0	99	76727.5	142	76020.0	111	73956.0	107
41	61547.5	103	58250.5	125	69704.0	135	81482.0	116	80486.0	152
42	70796.0	119	64570.5	114	61659.5	116	73715.5	160	86116.5	149
43	77043.0	170	73829.0	163	67837.5	155	64889.0	157	77336.5	164
44	82491.5	204	80129.5	186	77366.0	193	71064.0	179	67748.0	119
45	86635.0	220	84886.5	222	83169.5	224	80301.0	204	73433.5	188
46	88146.5	252	89106.0	251	88115.5	238	86228.5	257	82935.0	272
47	89647.5	323	90532.0	293	92205.0	322	91143.0	311	88757.0	329
48	92859.0	353	91714.5	316	93288.5	352	94843.5	328	93408.0	361
49	94786.0	447	94309.0	434	93795.0	404	95362.5	473	96741.0	376
50	93833.5	453	94277.0	479	94751.5	476	94289.5	407	95509.5	424
51	91337.5	549	93775.0	542	94824.0	519	95199.0	452	94513.0	507
52	87577.5	522	91337.0	611	94546.5	563	95547.0	564	95565.5	547
53	84911.5	621	87277.0	605	91628.0	649	94740.0	651	95545.0	612
54	81714.0	697	84025.5	714	87200.0	680	91473.5	680	94116.0	721
55	76256.5	693	78460.5	704	81482.0	754	84545.5	678	88218.5	780
56	71171.5	766	73447.0	712	76226.0	763	79204.0	811	81911.0	826
57	66284.0	763	69198.5	764	72025.0	704	74721.5	808	77442.0	800
58	61078.5	784	64174.5	844	67496.5	882	70394.5	910	72880.5	837
59	53745.5	793	57968.5	839	61398.5	906	64709.0	840	67355.5	924
60	39848.0	628	42770.0	726	46643.0	732	49601.5	785	52092.5	761
61	33166.0	587	33866.5	624	36723.0	610	40134.0	682	42602.5	756
62	29804.5	616	30603.5	618	31557.0	582	34332.5	620	37476.5	733
63	26708.5	561	27520.0	591	28528.5	600	29572.5	632	32046.5	726
64	23902.5	572	24033.0	575	25030.5	646	26044.5	650	26916.5	673
65	16220.5	410	15831.0	425	16016.5	416	16684.5	399	16995.0	439
66	11848.0	349	11560.0	326	11331.5	283	11427.5	316	11751.5	397
67	10130.5	292	9978.5	275	9899.0	267	9749.0	281	9739.5	278
68	9044.0	317	8876.5	330	8906.5	247	8900.5	308	8660.5	289
69	8305.5	302	8070.5	322	8030.5	273	8067.0	304	8012.0	294
70	7592.0	323	7351.0	324	7249.0	318	7252.0	324	7202.5	318
71	7009.5	341	6845.5	282	6688.5	311	6624.5	291	6588.5	282
72	6565.5	363	6379.5	337	6284.0	299	6166.0	329	6063.5	289
73	6115.5	320	5947.5	314	5853.0	255	5799.0	304	5665.5	298
74	5699.0	354	5549.5	375	5431.5	326	5386.5	306	5331.5	297
75	5294.5	343	5119.5	328	5003.5	292	4905.0	360	4832.0	332
76	4954.5	383	4757.0	307	4652.0	366	4531.5	315	4427.0	325
77	4697.5	422	4452.5	352	4283.0	340	4191.0	359	4073.0	363
78	4397.0	453	4213.5	335	4017.5	350	3853.0	326	3766.5	320
79	3926.5	392	3852.5	372	3757.5	398	3585.0	313	3431.5	379
80	3382.5	349	3331.5	391	3266.0	404	3190.0	369	3053.0	322
81	2928.0	302	2948.0	322	2892.0	325	2814.5	332	2719.5	307
82	2543.5	320	2550.0	318	2565.5	308	2531.5	361	2456.0	300
83	2227.0	354	2183.0	301	2197.0	291	2168.0	335	2157.0	296
84	1908.0	287	1861.5	292	1844.0	305	1851.5	298	1823.0	315
85	1624.0	271	1571.0	291	1534.0	234	1539.0	233	1520.5	297
86	1297.0	228	1330.5	253	1308.5	219	1290.0	242	1261.0	231
87	988.5	225	1041.0	192	1090.0	227	1064.0	191	1046.5	235
88	756.5	187	774.5	171	831.5	195	882.5	204	834.0	189
89	577.5	126	579.0	170	594.0	127	630.5	154	685.0	135

Table 15.2 : Central exposed to risk, for male assured lives, duration *5+*, based on policies - policy totals ceasing through death, ages *24 - 89*, calendar year *1962 - 1967*

	1963		1964		1965		1966		1967	
24	9051.0	7	9888.0	6	11111.0	10	13445.0	6	15770.0	12
25	10343.5	14	12392.0	11	14170.5	15	16083.0	12	18988.0	17
26	13391.5	10	14817.0	10	17588.0	17	20077.0	16	22104.0	20
27	18542.5	16	20108.0	12	22384.0	22	25479.0	13	27712.0	15
28	23203.0	18	25070.5	18	27751.5	27	30614.0	30	33461.0	29
29	28184.5	23	30721.0	29	33689.0	32	36778.0	15	39122.0	32
30	33230.5	38	35896.0	19	39546.5	39	42842.0	28	45545.0	26
31	39059.0	35	40522.5	36	44280.5	43	48170.0	39	50798.5	29
32	44960.0	43	46111.5	42	48586.0	39	52369.0	47	55213.0	35
33	50039.0	32	51825.5	49	53871.0	45	56129.0	45	58694.5	45
34	54645.0	44	56573.0	57	59489.5	54	61409.0	55	62123.0	47
35	59224.0	50	61333.5	52	64532.0	50	67236.0	43	67405.5	49
36	63711.5	87	65286.0	64	68675.5	61	71815.0	101	72970.0	74
37	67987.5	75	69359.0	91	72183.0	104	75529.0	83	77123.5	81
38	70860.5	69	73376.5	90	76095.5	87	78820.0	81	80487.0	100
39	73259.0	97	75961.0	105	79818.0	104	82469.0	116	83379.0	114
40	74952.0	108	77431.5	135	81815.0	119	85405.0	113	85683.0	110
41	78039.5	138	79307.0	148	83335.5	141	87588.0	149	89035.5	127
42	85022.0	160	82587.0	175	85189.5	166	88886.0	223	91030.5	148
43	90316.5	192	89454.5	200	88356.5	163	90687.0	193	92091.0	204
44	80519.0	187	94399.0	229	95404.0	238	93798.0	188	93376.5	184
45	69973.5	197	83502.5	192	99816.0	275	100272.0	275	95478.5	223
46	75825.0	211	72450.0	278	88222.0	266	104505.0	294	102143.5	272
47	86646.0	323	79583.5	288	75975.0	279	91539.0	261	105706.0	378
48	90913.0	377	87427.0	353	81676.0	335	78931.0	288	92400.0	312
49	95032.0	370	92741.0	356	90933.0	345	84321.0	378	79057.5	329
50	96893.0	490	95568.5	433	95210.5	423	92645.0	471	83264.5	384
51	95696.0	556	97461.5	511	98117.0	548	96990.0	429	91549.0	468
52	94852.5	588	96291.0	556	100005.0	593	99949.0	556	95833.0	530
53	95505.0	639	95045.0	580	98546.5	669	101475.0	732	98379.0	663
54	94755.5	743	95102.5	720	96840.0	786	99643.0	688	99302.0	717
55	90619.0	790	91496.5	847	94024.5	865	95217.0	790	95044.0	729
56	85319.5	868	88006.0	800	90897.0	873	92744.0	835	91133.0	874
57	79925.0	1006	83479.0	913	88066.5	903	90277.0	904	89207.0	901
58	75270.0	906	77867.5	945	83145.5	978	87096.0	1003	86669.5	991
59	69737.5	996	72191.0	915	76203.0	1064	80818.0	951	82182.0	1003
60	54101.5	853	56619.5	844	60269.0	942	63237.0	883	64802.0	889
61	44815.5	831	46757.5	764	50095.5	904	53417.0	870	54702.0	827
62	39783.5	800	42100.0	745	45165.5	937	47944.0	880	49262.5	857
63	34884.0	787	37332.5	712	40412.5	878	42933.0	809	44281.0	783
64	29273.0	696	32057.5	764	34934.0	868	37498.0	907	38671.0	837
65	17229.5	458	18671.0	480	20528.0	520	21560.0	566	21983.0	509
66	11950.5	351	12152.5	326	13155.0	341	14064.0	380	14156.0	347
67	10018.0	311	10229.0	281	10460.5	334	11154.0	310	11557.5	279
68	8599.5	307	8897.0	274	9115.0	303	9182.0	314	9601.5	304
69	7793.0	287	7762.0	302	8119.5	285	8209.0	297	8013.0	296
70	7019.5	303	6794.0	288	6842.0	272	7069.0	290	6984.5	242
71	6504.5	270	6339.5	290	6216.0	274	6216.0	250	6230.0	264
72	6024.0	302	5950.0	298	5838.5	304	5723.0	281	5527.0	270
73	5532.0	328	5495.0	312	5510.0	271	5392.0	281	5139.5	259
74	5162.5	309	5029.0	280	5072.0	305	5058.0	303	4846.0	279
75	4726.5	357	4617.0	296	4561.5	252	4498.0	317	4409.0	233
76	4347.5	315	4248.0	309	4212.0	270	4123.0	294	3981.0	229
77	3945.5	353	3884.0	303	3832.0	300	3794.0	280	3674.0	252
78	3634.5	330	3527.0	298	3508.5	294	3461.0	295	3356.0	287
79	3356.0	304	3250.5	304	3173.0	296	3154.0	311	3069.5	228
80	2945.0	308	2871.5	309	2772.0	285	2718.0	270	2652.5	253
81	2582.5	296	2585.5	254	2577.5	260	2446.0	301	2318.5	275
82	2365.5	291	2275.5	252	2253.5	286	2214.0	274	2093.5	212
83	2134.0	303	2071.5	249	1995.0	258	1918.0	273	1869.5	233
84	1810.5	317	1807.0	226	1777.5	263	1705.0	232	1602.0	228
85	1469.5	237	1502.5	225	1530.0	227	1476.0	250	1399.0	205
86	1228.0	271	1215.0	224	1258.5	212	1279.0	209	1228.5	182
87	1013.5	195	981.0	170	973.5	213	1004.0	180	1036.0	160
88	804.5	209	811.0	157	789.5	175	785.0	167	808.5	151
89	661.0	150	626.0	165	637.0	141	629.0	138	625.5	121

Table 15.3 : Central exposed to risk, for male assured lives, duration *5+*, based on policies - policy totals ceasing through death, ages *24 - 89*, calendar year *1968 - 1972*

	1968		1969		1970		1971		1972	
24	16888.0	11	17329.0	13	19772.0	14	21307.0	19	19635.0	21
25	21999.0	11	23459.5	26	23717.0	24	27146.0	20	29223.0	20
26	25743.0	18	29708.5	23	31554.0	27	31931.0	32	36567.0	32
27	30208.0	25	35519.5	19	40585.0	24	42668.0	38	43101.0	32
28	35448.0	25	38509.5	21	44992.0	35	50797.0	36	53013.0	32
29	41825.0	33	43759.5	47	47086.0	23	54410.0	42	61211.0	29
30	48117.0	25	51003.0	34	52265.0	40	55257.0	41	63691.0	37
31	53405.0	44	56731.5	50	59593.0	32	60050.0	37	63389.0	42
32	57699.0	41	61135.0	36	64845.0	51	67378.0	41	67342.0	45
33	61559.0	38	64960.0	46	68762.0	58	72372.0	44	74740.0	52
34	64523.0	34	68475.5	59	72468.0	45	76074.0	70	79635.0	55
35	67656.0	45	71218.0	60	75940.0	72	79718.0	65	83444.0	73
36	72674.0	82	73816.0	68	77973.0	83	82439.0	66	86252.0	74
37	77602.0	83	78390.0	63	79920.0	68	83676.0	77	88735.0	78
38	81406.0	83	83121.0	105	84550.0	107	85531.0	85	89479.0	102
39	84276.0	93	75677.5	115	77980.0	103	89996.0	124	91229.0	120
40	85951.0	94	77542.5	103	80727.0	100	93683.0	113	95140.0	117
41	88451.0	133	80372.5	180	84006.0	167	96822.0	153	98950.0	144
42	91659.0	151	82678.0	183	85278.0	180	98417.0	145	101753.0	176
43	93050.0	162	85174.5	156	87276.0	218	99521.0	209	103062.0	194
44	93613.0	195	86514.0	208	89515.0	247	100659.0	187	103521.0	191
45	93748.0	235	85936.0	245	90225.0	251	102560.0	254	103995.0	258
46	96213.0	275	86336.5	303	89658.0	289	103232.0	293	105782.0	297
47	102337.0	298	88319.0	322	89609.0	319	102165.0	348	105850.0	334
48	105829.0	362	94317.0	383	91368.0	387	101982.0	360	104727.0	366
49	91847.0	381	107268.0	485	107020.0	474	103231.0	436	103813.0	419
50	77241.0	379	91934.0	421	108217.0	498	107272.0	502	103896.0	437
51	81508.0	484	77069.0	377	93807.0	511	109843.0	581	107951.0	568
52	89510.0	506	80878.0	474	77555.0	461	93655.0	472	109857.0	650
53	93383.0	628	88642.5	548	81070.0	551	77209.0	443	94093.0	593
54	95181.0	671	92015.0	663	88248.0	682	80230.0	532	76873.0	545
55	93834.0	731	91545.0	725	89910.0	662	85818.0	717	78020.0	671
56	90063.0	817	90516.0	848	89869.0	785	87682.0	778	83896.0	704
57	86794.0	841	87448.0	860	89184.0	923	87970.0	830	85772.0	841
58	84670.0	958	83788.0	985	85736.0	938	86847.0	948	85492.0	904
59	81141.0	996	80657.0	1034	80718.0	920	81968.0	862	83112.0	958
60	65316.0	973	65533.0	950	66101.0	915	66454.0	803	67844.0	861
61	55833.0	818	57162.5	926	58125.0	907	58145.0	863	58318.0	810
62	50186.0	855	51821.5	872	53646.0	964	54502.0	875	54388.0	865
63	45327.0	870	46755.0	903	48754.0	976	50323.0	911	51040.0	922
64	39662.0	858	40972.0	875	42465.0	877	43962.0	898	45344.0	934
65	22105.0	544	22550.0	597	23950.0	554	24827.0	546	25489.0	591
66	14166.0	344	14199.5	370	14354.0	401	14547.0	346	14817.0	353
67	11584.0	353	11658.5	333	11636.0	309	11667.0	274	11882.0	313
68	9911.0	302	9980.5	317	10123.0	343	10051.0	264	10105.0	323
69	8357.0	251	8687.0	294	8784.0	304	8899.0	247	8819.0	312
70	6822.0	259	7196.5	275	7514.0	311	7595.0	280	7723.0	306
71	6147.0	274	6062.5	242	6440.0	250	6699.0	321	6735.0	290
72	5487.0	263	5524.5	235	5507.0	302	5810.0	272	5992.0	261
73	4951.0	254	5000.0	259	5017.0	268	4997.0	213	5303.0	313
74	4587.0	267	4459.0	258	4565.0	283	4594.0	255	4555.0	259
75	4229.0	310	4033.5	268	3978.0	249	4079.0	233	4103.0	258
76	3878.0	262	3720.5	281	3599.0	252	3564.0	224	3641.0	255
77	3542.0	293	3466.5	275	3364.0	267	3235.0	233	3185.0	221
78	3237.0	264	3159.0	257	3080.0	268	2977.0	255	2894.0	231
79	2941.0	294	2835.0	249	2812.0	232	2745.0	187	2642.0	263
80	2588.0	282	2523.0	242	2441.0	213	2431.0	215	2366.0	231
81	2225.0	256	2214.5	238	2211.0	233	2137.0	234	2111.0	213
82	1979.0	233	1929.0	239	1934.0	197	1918.0	219	1854.0	225
83	1800.0	199	1710.0	213	1668.0	233	1662.0	189	1655.0	208
84	1559.0	220	1529.5	215	1446.0	195	1413.0	181	1429.0	190
85	1309.0	198	1276.0	195	1273.0	191	1198.0	171	1164.0	175
86	1156.0	210	1087.5	185	1049.0	175	1025.0	158	960.0	163
87	1015.0	223	939.5	205	870.0	150	845.0	168	839.0	150
88	822.0	173	783.5	164	723.0	124	684.0	137	667.0	144
89	645.0	124	649.5	126	623.0	130	584.0	103	544.0	116

Table 15.4 : Central exposed to risk, for male assured lives, duration *5+*, based on policies - policy totals ceasing through death, ages *24 - 89*, calendar year *1973 - 1977*

	1973		1974		1975		1976		1977	
24	18204.0	9	17460.0	11	17233.3	12	17512.8	13	17972.9	11
25	26780.0	19	24889.0	14	24113.0	17	24066.8	18	24060.9	18
26	39492.0	32	36167.0	25	33787.0	16	32941.2	17	32232.0	22
27	49405.0	42	53227.0	38	48592.7	33	45428.4	20	43393.0	23
28	53979.0	26	62055.0	34	67113.5	41	62048.1	41	57404.8	33
29	64357.0	52	65278.0	44	75294.0	39	82059.2	51	75325.4	49
30	72353.0	48	75669.0	49	76489.5	55	89814.5	51	97427.1	56
31	73390.0	41	82759.0	38	86443.1	40	88361.1	65	103128.4	58
32	71420.0	53	82335.0	44	92711.3	54	97551.2	55	99023.1	65
33	74913.0	47	79175.0	59	90831.5	58	103328.0	68	108234.4	71
34	82806.0	68	82288.0	65	86258.0	58	100287.2	75	113199.3	79
35	87832.0	71	90736.0	67	89445.5	68	94609.3	68	109273.4	83
36	91051.0	90	95416.0	87	97960.5	62	97241.2	91	101785.6	99
37	93369.0	85	98007.0	80	102383.0	104	105872.1	120	104089.1	67
38	95470.0	95	99915.0	120	104321.6	85	109789.6	118	112308.5	124
39	95915.0	101	101805.0	107	106032.7	94	111424.2	143	115823.1	146
40	97134.0	96	101472.0	111	107324.1	112	112697.4	115	116874.9	146
41	100936.0	142	102610.0	132	106814.9	146	113720.4	172	117912.9	152
42	104717.0	173	106452.0	144	107540.0	166	112604.2	178	118549.5	170
43	106970.0	215	109697.0	221	111115.1	210	113012.3	193	116890.9	200
44	107900.0	226	111561.0	224	114198.7	241	116159.9	214	116666.6	223
45	107585.0	257	111584.0	293	114979.9	280	118474.0	249	119146.1	257
46	107810.0	288	111292.0	286	115130.5	286	118948.7	306	121018.2	269
47	108936.0	323	110678.0	326	114069.3	354	118552.1	362	121085.3	335
48	108749.0	384	111613.0	351	113237.7	341	117067.7	353	119915.3	351
49	106939.0	437	110736.0	384	113616.2	414	115647.1	412	118025.6	421
50	104770.0	495	107882.0	448	111784.9	487	114754.8	531	115126.7	481
51	104950.0	473	105878.0	509	109003.8	543	112696.7	495	113976.4	486
52	108733.0	561	105684.0	569	106914.7	590	110272.2	491	112562.5	559
53	110843.0	755	109282.0	678	106427.5	614	107587.6	638	109450.0	595
54	94230.0	582	110917.0	829	109498.7	754	106611.4	719	106305.5	757
55	75014.0	553	92272.0	653	109142.6	804	107581.9	771	103436.8	709
56	76688.0	679	73337.0	616	91013.1	719	107507.0	916	104297.2	790
57	81830.0	832	74774.0	707	72197.5	567	89630.0	781	104588.0	1018
58	83734.0	870	80186.0	833	73626.9	706	70876.1	635	86933.3	804
59	82145.0	955	80616.0	936	77647.6	832	70974.1	720	67663.0	758
60	69340.0	835	68733.0	845	68043.0	809	65646.3	842	59427.0	710
61	59998.0	773	61540.0	911	61920.8	886	61125.7	882	58247.4	846
62	54793.0	906	56607.0	841	58683.2	923	58740.3	897	57125.8	869
63	51119.0	918	51610.0	914	53840.6	854	55545.5	913	54808.6	890
64	46399.0	934	46550.0	959	47349.9	872	49129.1	886	50039.5	879
65	26386.0	557	26491.0	534	26533.1	555	26599.4	539	27029.0	509
66	15179.0	342	15311.0	341	15604.0	353	15162.4	333	14491.1	305
67	12171.0	324	12304.0	310	12541.5	328	12589.9	313	12141.9	285
68	10375.0	303	10557.0	276	10819.5	267	10938.5	296	10821.5	280
69	8903.0	283	9147.0	321	9486.5	288	9660.8	275	9647.6	250
70	7708.0	279	7792.0	248	8190.2	248	8463.3	311	8512.5	278
71	6864.0	261	6800.0	294	7127.2	251	7494.9	293	7567.2	269
72	6112.0	246	6296.0	275	6434.9	288	6623.4	288	6836.8	281
73	5494.0	286	5639.0	285	5967.4	273	5935.8	246	6061.1	288
74	4805.0	262	4977.0	250	5262.1	287	5494.4	239	5457.1	260
75	4069.0	248	4258.0	298	4588.1	281	4745.3	265	4933.5	262
76	3679.0	264	3640.0	237	3939.1	247	4146.4	285	4262.0	284
77	3265.0	278	3297.0	197	3437.4	248	3578.0	250	3687.3	253
78	2826.0	241	2916.0	226	3132.1	218	3141.6	245	3213.1	224
79	2585.0	223	2530.0	211	2688.0	208	2795.7	253	2814.1	212
80	2237.0	247	2169.0	195	2266.2	188	2344.8	233	2378.5	215
81	2055.0	212	1971.0	198	2045.4	235	1976.6	185	2012.0	205
82	1844.0	226	1794.0	212	1805.0	186	1775.6	185	1726.2	156
83	1612.0	207	1585.0	205	1646.7	187	1573.7	163	1537.7	168
84	1415.0	208	1358.0	186	1420.2	189	1396.9	193	1316.7	151
85	1165.0	147	1153.0	177	1209.0	182	1150.6	158	1124.5	148
86	958.0	155	963.0	160	1025.5	165	975.2	150	921.2	143
87	784.0	135	787.0	134	885.4	149	849.4	130	801.5	131
88	676.0	136	627.0	128	679.5	111	733.1	152	693.6	101
89	521.0	121	538.0	105	569.9	109	552.8	106	570.4	102

Table 15.5 : Central exposed to risk, for male assured lives, duration *5+*, based on policies - policy totals ceasing through death, ages *24 - 89*, calendar year *1978 - 1982*

	1978		1979		1980		1981		1982	
24	18673.3	11	19187.4	15	17233.3	12	18898.3	10	17089.8	10
25	25033.5	12	25788.9	21	24113.0	17	25447.0	15	23324.5	25
26	32468.0	16	33138.5	16	33787.0	16	32649.5	26	29946.5	17
27	42406.8	25	41830.5	27	48592.7	33	41462.8	21	37802.8	20
28	55001.5	32	52591.3	27	67113.5	41	51216.3	25	46948.3	20
29	70722.8	43	66387.3	39	75294.0	39	61403.8	33	56913.3	37
30	90123.0	65	82899.3	44	76489.5	55	73408.3	42	66963.0	34
31	113237.0	76	102976.8	58	86443.1	40	87952.3	42	78883.3	38
32	117360.8	60	126274.9	59	92711.3	54	104647.5	57	92983.5	75
33	111354.3	57	128857.8	96	90831.5	58	124723.8	84	109355.8	79
34	120231.3	86	121252.8	78	86258.0	58	149156.5	116	129431.3	67
35	125212.8	102	130162.5	116	89445.5	68	149511.3	107	154166.0	98
36	119531.8	119	134398.4	97	97960.5	62	138930.5	95	153700.8	111
37	110773.0	109	127313.9	116	102383.0	104	146951.8	127	141387.0	97
38	112358.3	94	117268.0	85	104321.6	85	149666.8	126	149525.8	168
39	120549.8	133	118593.0	113	106032.7	94	140519.5	120	151491.8	112
40	123568.8	167	126209.8	143	107324.1	112	128133.0	118	141513.3	160
41	124447.3	146	129372.4	170	106814.9	146	128269.0	146	128585.5	154
42	125118.5	181	129717.8	185	107540.0	166	135762.5	174	128459.0	168
43	125124.5	200	129817.4	195	111115.1	210	138184.3	206	135847.3	179
44	122935.3	204	129345.8	224	114198.7	241	137460.3	222	137628.3	237
45	121676.3	241	126068.8	240	114979.9	280	136109.3	233	135705.0	257
46	123863.3	277	124599.8	270	115130.5	286	134224.5	305	134062.5	292
47	125329.5	325	126316.1	315	114069.3	354	130095.0	282	131680.8	289
48	124518.0	334	127421.5	378	113237.7	341	127735.8	297	127435.5	310
49	122927.8	359	126014.3	475	113616.2	414	128381.5	400	124456.8	356
50	119629.8	461	123159.4	494	111784.9	487	127595.5	485	123940.8	408
51	116722.8	542	119819.9	506	109003.8	543	125012.8	552	123187.5	502
52	115654.3	598	116988.4	563	106914.7	590	121944.3	535	120570.0	536
53	113638.8	716	115516.5	685	106427.5	614	118421.5	606	117457.3	588
54	109899.5	629	112949.8	721	109498.7	754	114745.5	651	113770.3	601
55	104859.0	737	107393.1	711	109142.6	804	110778.3	723	108065.3	675
56	102241.8	823	102758.0	767	91013.1	719	106696.8	814	104632.0	738
57	103148.8	872	100546.8	802	72197.5	567	102034.8	876	101153.0	769
58	103448.3	1036	101367.5	933	73626.9	706	97998.0	900	96515.3	819
59	84941.3	820	100322.3	1076	77647.6	832	94520.3	1014	91519.5	901
60	58063.8	654	73185.3	771	68043.0	809	83599.5	914	78662.0	752
61	53886.5	675	52390.3	614	61920.8	886	76923.8	1009	73163.0	895
62	55504.0	770	51243.1	778	58683.2	923	61575.0	805	70715.8	1006
63	54434.3	921	52612.5	875	53840.6	854	46377.0	767	56201.5	825
64	50383.0	907	49650.1	791	47349.9	872	43477.8	693	40697.3	685
65	27923.3	549	27473.3	521	26533.1	555	25362.0	468	22014.0	370
66	14942.3	324	15089.0	293	15604.0	353	14318.0	260	13199.3	288
67	12018.3	249	12144.4	263	12541.5	328	11903.0	222	11231.3	217
68	10719.8	286	10417.8	255	10819.5	267	10555.3	264	10108.5	224
69	9799.8	322	9549.3	264	9486.5	288	9318.5	261	9213.0	243
70	8653.3	264	8653.3	289	8190.2	248	8266.3	249	8138.8	224
71	7795.8	257	7821.8	277	7127.2	251	7702.3	275	7370.0	236
72	7083.0	276	7203.4	258	6434.9	288	7221.0	269	6915.3	239
73	6395.8	281	6573.9	290	5967.4	273	6695.0	310	6533.5	242
74	5616.5	266	5922.4	280	5262.1	287	6178.8	293	6017.3	275
75	4971.5	285	5066.0	269	4588.1	281	5550.3	292	5500.3	303
76	4506.8	258	4517.4	268	3939.1	247	4852.0	262	4897.5	291
77	3883.8	261	4093.0	280	3437.4	248	4224.5	279	4314.0	251
78	3391.0	284	3499.0	219	3132.1	218	3708.5	213	3760.3	250
79	2934.0	232	3053.0	236	2688.0	208	3360.8	243	3332.8	287
80	2467.0	245	2559.3	212	2266.2	188	2784.8	264	2762.5	257
81	2083.3	194	2145.4	204	2045.4	235	2371.8	248	2307.3	206
82	1779.0	171	1842.8	188	1805.0	186	2004.5	194	1987.0	201
83	1534.5	168	1579.3	157	1646.7	187	1675.8	182	1698.5	208
84	1295.5	150	1314.0	142	1420.2	189	1376.8	175	1348.5	133
85	1107.3	137	1180.0	154	1209.0	182	1092.5	131	1065.8	140
86	940.5	157	982.8	125	1025.5	165	927.5	140	854.3	116
87	755.3	133	807.6	101	885.4	149	774.0	107	737.8	118
88	652.0	103	659.1	106	679.5	111	639.5	93	603.0	102
89	577.8	96	580.5	120	569.9	109	532.0	89	500.8	86

Table 15.6 : Central exposed to risk, for male assured lives, duration *5+*, based on policies - policy totals ceasing through death, ages *24 - 89*, calendar year *1983 - 1987*

	1983		1984		1985		1986		1987	
24	15503.0	8	13898.3	4	13063.8	6	11643.5	9	9400.3	5
25	21098.3	13	18364.8	16	16995.5	15	15312.3	9	12435.5	9
26	27169.3	13	23600.3	17	21498.8	16	19147.8	14	15615.3	6
27	34092.3	30	29458.8	16	27057.8	7	23672.8	15	19167.3	17
28	42159.5	24	36301.3	29	33340.0	26	29459.8	16	23197.8	6
29	51322.3	33	44202.5	27	40489.0	16	35755.8	22	28450.0	7
30	61257.0	32	53217.8	27	48835.0	31	42797.5	26	33790.3	19
31	71159.5	42	62809.8	35	58330.0	39	51190.5	22	40176.5	23
32	82972.0	51	72361.3	35	67554.0	36	60001.8	29	47227.5	25
33	97241.5	60	83996.3	39	77571.8	64	69372.8	39	54786.8	34
34	113668.5	74	97958.5	55	89235.0	47	78854.0	54	62905.5	39
35	134331.0	114	114753.0	78	103681.3	57	90521.8	59	71434.5	44
36	159572.0	102	135133.3	81	120870.8	102	105022.3	74	81360.3	47
37	158037.8	134	160225.0	118	141784.8	109	121915.8	98	93797.5	68
38	145712.5	123	158627.0	156	167277.3	160	142669.8	147	108726.5	104
39	153503.3	133	145546.5	115	163539.8	138	166632.5	141	126725.8	99
40	155320.8	158	152933.0	137	150652.8	139	163142.8	191	147468.0	123
41	144959.0	181	154727.3	156	158193.0	221	150365.5	162	146111.5	153
42	131932.0	163	144183.0	154	159559.5	174	157620.0	188	133678.5	110
43	131599.8	191	130790.0	172	148057.3	186	158330.0	163	139933.0	194
44	139099.0	229	130484.3	235	134425.5	219	146777.0	252	140964.8	190
45	140379.8	237	137137.5	268	133718.8	243	132885.5	204	130230.8	254
46	138693.3	298	138471.5	290	140443.3	281	131926.3	269	117778.5	219
47	136754.0	357	136332.5	292	141531.0	290	138323.3	283	116122.8	232
48	134219.5	340	134407.0	349	139262.5	315	139347.5	350	121834.5	285
49	129335.5	403	131579.0	405	137017.0	357	136546.8	325	122200.5	293
50	125563.3	471	125681.8	392	133229.8	443	133561.5	427	119085.8	329
51	125332.8	461	122472.3	430	127383.3	442	129855.0	433	116519.5	364
52	124710.5	452	122599.5	503	124617.0	526	124776.0	509	113640.5	461
53	121889.8	525	121932.5	557	124658.8	557	121987.0	499	109423.8	429
54	118369.0	658	118974.5	624	123999.5	582	121780.0	578	106391.8	458
55	112648.3	726	113591.0	610	118704.0	622	119114.0	678	104574.0	581
56	107410.3	775	108392.3	711	113837.0	686	114582.5	678	102815.8	587
57	104566.8	802	103937.5	752	109152.5	776	110284.8	729	99458.8	735
58	100872.0	898	101175.8	795	104973.0	832	105980.0	839	95796.5	694
59	95061.5	876	96386.3	828	100454.8	868	99926.0	894	90630.0	775
60	80208.3	857	80880.0	823	85803.3	791	86334.5	893	76679.5	707
61	72180.8	835	71421.3	814	74862.3	874	76349.8	856	68610.0	745
62	70192.5	926	67606.3	845	69459.0	865	70030.3	853	63705.0	781
63	67666.3	1117	65458.3	1022	65707.0	969	64563.8	916	58093.3	787
64	52280.8	806	60960.3	1027	59814.8	947	56455.5	893	50515.0	755
65	22210.5	403	27822.0	481	33556.5	634	32070.3	514	27921.5	462
66	12455.3	229	11933.0	202	15426.8	284	18144.0	289	15932.8	301
67	11067.8	247	10055.5	185	10327.8	247	12940.8	250	13522.8	266
68	10097.5	195	9714.8	242	9400.8	209	9184.3	194	10338.3	203
69	9371.5	290	9165.5	228	9329.3	238	8563.8	193	7465.3	175
70	8533.5	239	8542.3	218	8965.8	249	8718.3	254	7188.0	194
71	7712.0	266	7923.8	259	8484.8	238	8460.5	209	7282.0	222
72	7046.5	232	7245.3	227	7930.5	276	8114.3	253	7171.0	220
73	6721.5	286	6688.5	282	7293.5	262	7537.3	294	6751.8	265
74	6263.0	242	6268.0	225	6675.5	271	6874.0	299	6297.0	271
75	5689.3	307	5786.0	258	6172.8	266	6161.8	261	5629.3	224
76	5115.5	257	5216.8	295	5756.0	286	5784.5	314	5037.3	257
77	4601.8	268	4698.0	249	5188.8	316	5404.3	323	4725.8	205
78	4089.5	282	4205.5	242	4720.8	300	4823.8	288	4444.8	269
79	3515.8	257	3685.0	251	4228.5	289	4413.3	313	3954.3	231
80	2947.3	241	3066.3	260	3514.3	275	3695.3	257	3485.3	259
81	2529.0	241	2604.0	209	2961.5	304	3090.3	243	2946.3	217
82	2100.8	212	2216.5	193	2527.3	220	2629.3	217	2516.3	205
83	1828.5	223	1837.0	211	2143.5	206	2269.8	260	2116.5	202
84	1484.8	168	1519.8	173	1732.5	204	1851.5	211	1720.0	146
85	1156.5	128	1185.3	130	1375.3	165	1418.3	181	1361.3	160
86	924.8	122	969.5	137	1112.5	158	1153.3	130	1077.3	131
87	741.5	116	763.8	91	918.3	126	936.5	124	895.3	107
88	621.3	90	624.3	116	732.5	98	798.3	123	735.0	113
89	522.0	116	493.0	72	564.5	93	600.5	79	595.0	86

Table 15.7 : Central exposed to risk, for male assured lives, duration *5+*, based on policies - policy totals ceasing through death, ages *24 - 89*, calendar year *1988 - 1990*

	1988		1989		1990	
24	*8158.0*	*11*	*7305.5*	*3*	*6939.0*	*2*
25	*11393.5*	*7*	*10603.5*	*6*	*9523.3*	*2*
26	*14545.8*	*8*	*13999.8*	*11*	*12892.3*	*3*
27	*17753.0*	*6*	*17234.5*	*11*	*16097.5*	*9*
28	*21218.8*	*18*	*20509.5*	*13*	*19124.8*	*8*
29	*25463.0*	*5*	*24225.5*	*6*	*22491.8*	*16*
30	*30553.5*	*16*	*28861.3*	*14*	*26169.8*	*14*
31	*35914.5*	*24*	*34225.3*	*23*	*30659.8*	*17*
32	*41841.8*	*19*	*39797.5*	*17*	*36162.3*	*29*
33	*48290.3*	*25*	*45912.0*	*28*	*41508.8*	*28*
34	*55446.8*	*33*	*52523.0*	*26*	*47326.0*	*33*
35	*63456.5*	*39*	*60101.0*	*62*	*53768.3*	*33*
36	*71528.3*	*30*	*68307.5*	*39*	*61017.8*	*57*
37	*81055.8*	*50*	*76757.8*	*63*	*69190.3*	*57*
38	*92998.5*	*54*	*86632.0*	*83*	*77299.8*	*64*
39	*107286.8*	*91*	*99022.8*	*83*	*87220.3*	*83*
40	*124850.5*	*101*	*113532.0*	*107*	*99177.3*	*86*
41	*146965.0*	*142*	*131712.8*	*144*	*113412.8*	*122*
42	*144688.8*	*138*	*154268.8*	*151*	*131393.3*	*132*
43	*131854.3*	*159*	*151069.8*	*177*	*153582.5*	*187*
44	*138010.8*	*212*	*137413.8*	*153*	*150053.8*	*212*
45	*138685.5*	*205*	*142753.5*	*206*	*136043.5*	*211*
46	*128142.8*	*233*	*143112.0*	*248*	*141513.8*	*243*
47	*115223.0*	*226*	*131776.0*	*267*	*141558.8*	*279*
48	*113625.5*	*270*	*118358.5*	*278*	*130448.5*	*270*
49	*118546.0*	*333*	*116241.3*	*237*	*116943.3*	*275*
50	*118266.8*	*319*	*119946.8*	*307*	*114326.5*	*312*
51	*115769.3*	*366*	*119750.3*	*370*	*118530.3*	*327*
52	*113615.0*	*346*	*117428.3*	*402*	*118867.3*	*390*
53	*111111.0*	*465*	*115084.3*	*430*	*116743.3*	*355*
54	*106608.3*	*428*	*111965.3*	*476*	*114211.3*	*452*
55	*101932.8*	*503*	*105759.0*	*453*	*109591.3*	*532*
56	*101483.3*	*560*	*102368.8*	*505*	*104260.0*	*537*
57	*100721.0*	*593*	*102757.0*	*634*	*101979.8*	*579*
58	*97412.3*	*640*	*101580.5*	*636*	*102100.8*	*624*
59	*92678.3*	*711*	*96863.5*	*753*	*99634.3*	*668*
60	*78664.5*	*664*	*82588.3*	*716*	*85365.3*	*729*
61	*69407.8*	*743*	*72553.3*	*667*	*75709.3*	*709*
62	*65136.0*	*799*	*67042.0*	*791*	*69852.8*	*726*
63	*60429.5*	*793*	*62662.3*	*772*	*64520.0*	*801*
64	*53037.8*	*789*	*55791.0*	*727*	*58014.5*	*834*
65	*28719.3*	*439*	*29878.3*	*420*	*32234.3*	*487*
66	*16049.3*	*262*	*16591.0*	*256*	*18006.5*	*304*
67	*13498.5*	*248*	*13637.5*	*233*	*14743.8*	*265*
68	*12326.0*	*261*	*12355.0*	*219*	*12940.5*	*280*
69	*9607.5*	*220*	*11672.3*	*261*	*12102.0*	*290*
70	*7054.5*	*205*	*9115.8*	*224*	*11457.8*	*266*
71	*6814.0*	*176*	*6769.0*	*221*	*9037.3*	*239*
72	*7065.5*	*241*	*6632.0*	*182*	*6762.8*	*218*
73	*6872.5*	*236*	*6844.8*	*217*	*6658.8*	*230*
74	*6434.5*	*224*	*6627.0*	*276*	*6757.5*	*262*
75	*5888.3*	*256*	*6083.5*	*253*	*6378.3*	*275*
76	*5264.5*	*251*	*5506.8*	*257*	*5876.5*	*243*
77	*4780.8*	*260*	*4970.0*	*233*	*5330.0*	*281*
78	*4512.8*	*267*	*4522.3*	*278*	*4810.5*	*296*
79	*4206.8*	*294*	*4249.5*	*289*	*4338.5*	*321*
80	*3600.8*	*249*	*3756.0*	*278*	*3815.3*	*283*
81	*3194.5*	*254*	*3297.8*	*289*	*3393.3*	*258*
82	*2707.5*	*202*	*2906.0*	*247*	*3003.5*	*245*
83	*2296.3*	*229*	*2469.5*	*223*	*2698.3*	*287*
84	*1861.5*	*215*	*2002.8*	*188*	*2223.3*	*213*
85	*1451.3*	*167*	*1535.5*	*187*	*1709.5*	*209*
86	*1129.0*	*132*	*1242.0*	*171*	*1304.3*	*148*
87	*906.8*	*107*	*966.8*	*98*	*1066.5*	*149*
88	*760.8*	*108*	*799.8*	*104*	*860.5*	*112*
89	*600.8*	*102*	*640.8*	*95*	*704.5*	*122*

CHAPTER XVI

Appendix B

In the following Tables (*16.1 - 16.2*) the data (initial exposed to risk based on policies - policy totals ceasing through death) for male pensioners' experience, ages *60 - 95*, for each calendar year (*1983 - 1990*) separately, are presented, as published by the *CMI* Bureau of the Institute and Faculty of Actuaries.

Table 16.1 : Central exposed to risk, for male pensioners, based on policies - policy totals ceasing through death, ages *60 - 95*, calendar year *1983 - 1986*

	1983		1984		1985		1986	
60	384	5	408.5	6	461.6	12	490	10
61	730	7	745.7	5	829.4	13	906.7	9
62	896	11	888.2	17	928.3	17	1002.7	16
63	1091.5	21	1089.5	15	1107.8	15	1146.3	17
64	1034.5	35	1257.3	26	1294.9	42	1292.7	25
65	7235	155	7724.9	161	8831.6	211	8420.3	178
66	16944.5	409	14686.1	357	16948.8	361	18430	396
67	21111	637	17089.1	436	14767.1	365	16513	387
68	24543	796	20762.5	594	16676.3	507	13969.8	408
69	26192.5	848	23961.9	780	20010.5	643	15617.5	431
70	26896	983	25543.5	930	23040.2	788	18799.4	669
71	25840	1122	26113.6	1037	24565.6	946	21618.8	842
72	23677	1016	24754.3	1076	24865.7	1145	22688.4	975
73	22746.5	1160	22590.9	1077	23325.5	1131	22741.4	1030
74	21219	1176	21550.2	1217	21156.5	1090	21267.1	1102
75	19368	1180	20055.1	1197	20147.6	1209	19204	1008
76	17781	1111	18172.3	1192	18627.6	1226	18214.3	1190
77	16351.5	1161	16507.6	1113	16742.1	1286	16610.7	1162
78	14945.5	1212	15148	1193	15166.7	1171	14840.2	1097
79	13165.5	1196	13638	1146	13759.5	1211	13528.3	1126
80	11323	1108	11877.2	1117	12284.1	1124	12261	1152
81	9815	1011	10155.5	1122	10568.3	1117	10776.2	1034
82	8080.5	933	8681.6	1019	8891.5	977	9261.8	1059
83	6333	755	7090.6	901	7557.8	952	7663.1	925
84	4813	665	5486.5	678	6106.1	804	6451.7	821
85	3565	490	4143	579	4757.2	699	5146.3	692
86	2795	462	2994.2	432	3498.2	549	3945.4	580
87	2146	369	2307.6	357	2502.4	402	2867.8	503
88	1562	276	1774.8	326	1913.4	337	2018.6	342
89	1141	226	1249.8	215	1439.8	295	1525.1	283
90	857.5	163	904.2	173	1013.5	193	1110.2	237
91	671	144	680.3	121	723.8	167	783.5	159
92	460.5	90	525.1	106	548.2	123	534.9	101
93	327.5	84	363.8	100	398.4	90	411.4	88
94	215	72	232.1	48	275.6	81	298.1	68
95	122.5	35	145.1	48	171.7	42	192.9	52

Table 16.2 : Central exposed to risk, for male pensioners, based on policies - policy totals ceasing through death, ages *60 - 95*, calendar year *1986 - 1990*

	1987		1988		1989		1990	
60	566.4	10	621.9	4	554	6	477.5	1
61	1004.7	15	1082.5	18	1003.5	8	847.5	12
62	1176	18	1235.7	24	1124.1	17	962.8	12
63	1310	27	1431.3	26	1300.4	20	1035.9	11
64	1413.3	35	1516.5	31	1472.3	42	1100.2	18
65	8087.9	183	7563.8	146	6969.7	131	5164.9	101
66	18166.4	417	16668.6	365	15103.3	294	11035.1	217
67	19894.6	508	18349.2	437	16088.5	344	11397.2	230
68	17307.2	433	19555.7	520	17324.7	440	11639.1	317
69	14482.4	443	16865.8	444	18308.1	530	12342.8	328
70	16175.1	567	14183.9	500	15731.2	468	12926.2	393
71	19220.8	760	15710.2	636	13288.5	516	10867.2	343
72	21763.4	959	18345.1	813	14479.9	558	8654.5	311
73	22717	1141	20628.3	993	16848.6	735	9138.7	385
74	22606.5	1191	21384.4	1127	18918.2	991	10415.1	444
75	20986.1	1210	21219.8	1195	19551.7	1047	11518.8	572
76	18877.3	1169	19630.9	1225	19438.1	1210	11717.5	668
77	17597.2	1246	17529.9	1174	17761.7	1211	11363.5	812
78	15953	1228	16259.2	1269	15700.2	1084	10326.3	684
79	14192	1138	14682.6	1172	14510.3	1140	9281.4	730
80	12693.3	1188	12893.1	1125	13058.2	1107	8400.3	719
81	11230.5	1175	11337	1076	11352	1091	7531.4	706
82	9813.5	1118	9947.3	1046	9919.5	1011	6646.4	664
83	8253.3	949	8605.2	978	8626.8	944	5717.2	601
84	6825.6	877	7194.4	861	7365.9	916	4709	565
85	5676.6	818	5873.4	793	6124	800	3772.6	477
86	4483	618	4814.5	683	4917.3	703	3005.9	395
87	3401.3	597	3795	633	3973.4	583	2380.4	353
88	2379.4	362	2797.1	454	3084.3	553	1875.1	285
89	1694.4	341	1970.4	330	2267.6	376	1424.1	205
90	1220.3	233	1349.8	242	1593.5	298	1016	184
91	876.5	154	981	210	1082	205	698.5	136
92	628	129	698.2	152	757	178	459.9	89
93	432.8	104	481.9	113	518.2	116	297.8	57
94	335.6	89	330.6	80	360.6	101	207.2	48
95	226.7	71	241.6	62	235.6	57	147.6	29

CHAPTER XVII

References

- Anson J. (*1988*), The parameters of death : A consideration of the quantity of information in a Life Table using a Polynomial representation of the Survivorship Curve, Statistics in Medicine, Vol *7*, *895 - 912.*

- Beard R. E. (*1951*), Some notes on graduation, *J.I.A., 77, 382.*

- Benjamin B. and Pollard J. (*1980*), The Analysis of mortality and other Actuarial Statistics, Heinmann, London.

- Chadburn R. G. (*1991*), Notes on Mortality, The City University, London, Volume 2, Graduation and Miscellaneous Topics.

- *CMI* Committee (*1976*), The Graduation of Pensioners' and of Annuitants Mortality Experience *1967 - 70, CMIR, 2, 57.*

- *CMI* Report No *9* (*1988*), The graduation of the *1979 - 1982* Mortality Experiences, Institute and Faculty of Actuaries.

- Coale A. J. and Demeny P. (*1966*), Regional Model Life Tables and Stable Populations, University Press, Princeton, *NJ.*

- Copas J. B. and Haberman S., (*1983*). Non - parametric graduation using Kernel methods, *J.I.A., 110, 135 - 156.*

- Currie I. D. and Waters H. R. (*1991*), On Modelling Select Mortality, *J.I.A., 118, 453 - 488.*

- Forfar D., McCutcheon, J. and Wilkie A. (*1988*), On graduation by mathematical formula, *J.I.A., 115, 1 - 149.*

- Francis B., Green M. and Payne C. (*1993*), Eds, The *GLIM* system. Release *4* Manual, Clarendon Press, Oxford.
- Gavin J. B., Haberman S. and Verrall R. J. (*1993*), Moving weighted average graduation using Kernel estimation. Insurance : Mathematics and Economics, *12, 113-126.*
- Gavin J. B., Haberman S. and Verrall R. J. (*1994*), On the Choice of Bandwidth for Kernel graduation. *J.I.A., 121, 119 - 134.*
- Gavrilov L. and Gavrilova N. S. (*1991*), The Biology of life span: a quantitative approach, Academician V.P. Skulachev.
- Gerber H.U. (*1990*), Life Insurance Mathematics, Swiss Association of Actuaries Zurich.
- Gompertz, B (*1825*), On the nature of the function of the law of human mortality and on a new model of determining the value of life contingencies. Phil. Trans. of Royal Society, *115, 513 - 85.*
- Haberman S. and Sibbett T. (*1995*), The History of Actuarial Science, Pickering and Chatto.
- Hartmann M. (*1980*), Laws of Mortality on the Logit scale, *2*nd Edition, Wiley.
- Hatzopoulos P. (*1992*), Graduation through Generalised Linear Models with the statistical package *GLIM*, A dissertation submitted for the award of the degree of Master of Science in Actuarial Science, The City University, London.
- Heligman L. and Pollard J. H. (*1980*), The Age Pattern of Mortality, *J.I.A., 107, 49 - 80.*
- Kakulos N. (*1990*), Stochastic Processes, *3*rd Edition, Athens.
- Keyfitz N. (*1982*), Choice of Functions for Mortality Analysis : Effective Forecasting Depends on a Minimum Parameter Representation. Theoretical Population Biology, *21, 329 - 352.*
- Klonias V. (*1987*), Applied Statistics, Notes at Crete University.
- London D. (*1985*), Graduation : The Revision of Estimates. Actex Publications, Winsted and Abington, Conn.

- Makeham W. M. (*1860*), On the law of mortality, *J.I.A., 13, 325 - 358.*

- McCullagh P. and Nelder J.A. (*1983* & *1989*), Generalised Linear Models, *1*st & *2*nd Edition, Chapman and Hall.

- McCutcheon J. J. (*1984*), Spline Graduation with Variable Knots, Proceedings of *22*nd *I.C.A.*, *4, 47.*

- *OPCS* (*1987*), Life Tables *1980 - 1982.* Series *DS* No. *7. HMSO.*

- Perks W. F. (*1932*), On some experiments in the graduation of mortality statistics, *J.I.A. 63, 12 - 40.*

- Puzey A. S. (*1986*), Exposed to Risk, Special Note, Institute of Actuaries and Faculty of Actuaries - Education Service.

- Redington F. M. (*1969*), An Explanation into Patterns of Mortality, *J.I.A., 95, 243 - 298.*

- Renshaw A. E. (*1991a*), Actuarial Graduation Practice and Generalized Linear & Non-Linear Models, *J.I.A., 118, 295.*

- Renshaw A. E. (*1991b*), Graduation by Generalised Linear Modelling Techniques, Actuarial Research Paper No *26*, The City University, London.

- Renshaw A. E. (*1992*), Joint Modelling for Actuarial Graduation and Duplicate Policies, *J.I.A., 119, 69.*

- Renshaw A. E., Haberman S. and Hatzopoulos P. (*1996a*), Recent Mortality Trends in Male Assured Lives, *B.A.J., 2, 449 - 477*

- Renshaw A. E., Haberman S. and Hatzopoulos P. (*1996b*), On the Duality of Assumptions Underpinning the Construction of Life Tables, *ASTIN* Bulletin, to appear.

- Renshaw A. E. and Hatzopoulos P. (*1996*), On the Graduation of 'Amounts', *B.A.J., 2, 185 - 205.*

- Royston P. and Altman D. G. (*1994*), Regression using Fractional Polynomials of Continuous Covariates : Parsimonious Parametric Modelling, Applied Statistics, *43*, *429 - 468*.

- Steffensen J. F. (*1934*), Forsikrings matematik. Gads, Copenhagen.

- Sverdrup E. (*1965*), Estimates and Test - Procedures in Connection with Stochastic Models for Deaths, Recoveries and Transfers between Different States of Health. Skandinavisk Aktuarielidskrift, *48, 184.*

- Thatcher A. R., Some results on the Gompertz and Heligman and Pollard Laws of Mortality, *J.I.A., 117, 135 - 151.*

- Thiele P. (*1872*), On a mathematical formula to express the rate of mortality throughout the whole of life, *J.I.A., 16, 313.*

- Verrall R. J. (*1992*), A State Space Formulation of Whittaker - Henderson Graduation with Extensions, Actuarial Research Paper No *33*, The City University, London.