A Study of Actuarial Models for Insurance Based Applications

by

Zoltan Butt

A Dissertation Submitted to the

Faculty of Actuarial Science and Insurance
City University, London

in partial fulfillment of the
requirements for the degree of

Doctor of Philosophy
in
Actuarial Science

June 2014

CONTENTS

CONTENTS **vi**

List of Tables **ix**

List of Figures **xiii**

List of Abbreviations **xiv**

Acknowledgements **xv**

Declaration **xvi**

Abstract **xvii**

Chapter 1: Introduction **1**

1.1 Overview 1

1.1.1 Personal injury compensation in the UK and the Ogden Tables 3

1.1.2 Mortality models for heterogeneous insurance based populations 5

1.1.3 Modelling and forecasting mortality within a Lee–Carter type framework 7

1.2 Individual Contributions to Papers included in the Thesis 9

1.2.1 Contributions to the work presented in Chapters 2 and 3 (Part I) 9

1.2.2 Contributions to the work presented in Chapter 4 (Part II) 13

1.2.3 Contributions to the work presented in Chapter 5 (Part II) 16

I PERSONAL INJURY COMPENSATIONS IN THE UK AND THE OGDEN TABLES **18**

Chapter 2: Overview of Calculating Personal Injury Damages in the UK **19**

2.1 Introduction 19

2.2 Background 21

2.3 Multipliers in England and Wales: Broad Issues 23

2.4 The Ogden Tables 25

2.5 Calculating Reduction Factors for Labour Market Risks: Ogden Tables Approach based on Haberman and Bloomfield (1990) . . . 30

2.6 The Alternative Approach based on Lewis et al. (2002, 2003) . . . 34

2.7 Models of Labour Market Contingencies in the UK 39

2.7.1 Unemployment 39

2.7.2 Sickness 42

2.7.3 Industrial disputes 44

2.7.4 Early retirement 44

2.8 Other Multiple State Modelling 45

2.9 Conclusions 50

2.10 Appendix 52

Analytical comparison of traditional and alternative loss of earnings estimators (based on Bell and Taub 1998) 52

Chapter 3: Modelling Loss of Earnings Multipliers in the UK and the Ogden Tables (up to and including the 5th edition) **55**

3.1 Introduction 56

3.2 Critique of the Ogden Tables Multipliers 59

3.3 Longitudinal Labour Force Survey Data 62

3.4 Methodology 65

3.4.1 The multiple state model 67

3.4.2 Estimating the transition intensity 67
3.4.3 Estimating the transition probability 71
3.4.4 Estimating the worklife expectancy 73
3.4.5 Reduction factors 77
3.5 Additional Covariates 79
3.5.1 Economic regions 82
3.5.2 Main industry sector 83
3.5.3 Highest educational qualification 85
3.5.4 Type of disability 86
3.6 Conclusions 90
3.7 Appendix 92
Appendix A: Tables 92
Appendix B: Figures 109

II MORTALITY MODELLING AND FORECASTING OF INSURANCE BASED POPULATIONS 115

Chapter 4: Mortality Models for Heterogeneous Insurance Based Populations 116

4.1 Introduction 117
4.2 Model of Individual Differences in Frailty 118
4.3 Models and Methods of Fit 121
4.3.1 Introduction 121
4.3.2 Exploratory data analysis 122
4.3.3 Generalized linear models with parameterised link function 123
4.4 Application to Insurance Data 126
4.4.1 Immediate annuitants data set 126
4.4.2 Immediate annuitants experience 127
4.4.3 Assured lives (males) data set 136

4.4.4 Assured lives experience 136
4.5 Interpretation of Results 139
4.5.1 Identifiability 139
4.5.2 Interpretation of the Perks model parameters 140
4.5.3 Comparison of results 141
4.6 Discussion 143
4.7 Extensions and Additional Features 146
4.7.1 Comparison of two heterogeneous populations 146
4.7.2 Alternative modelling frameworks 153
4.7.3 Forecasting parameter values 156
4.8 Appendix: Frailty Distributions 159
The gamma distribution 160
The inverse Gaussian distribution 164
A generalised distributional assumption 167

Chapter 5: Modelling and Forecasting Mortality within a Lee–Carter Type Framework 169

5.1 Introduction 169
5.2 Modelling Framework 171
5.2.1 Mortality data 173
5.2.2 Basic Age–Period (AP) Lee–Carter model 173
5.2.3 Generalised family of Lee–Carter models 177
5.2.4 Stratified (or Extended) Lee–Carter model 178
5.2.5 Forecasting approach 180
5.3 Fitting Methodology 182
5.3.1 Updating cycle of the base AP (LC) fitting 184
5.3.2 Updating cycle of APC fitting 185
5.3.3 Updating cycle of SLC fitting 187
5.4 Application of the Generalised LC Models in R with ilc 190

5.4.1 Package installation 191
5.4.2 Preparing the mortality data for analysis 192
5.4.3 Fitting the mortality models and making forecasts 199
5.5 Application to Pensioners Data Set 216
5.5.1 Pensioners data set 218
5.5.2 Other mortality data sets considered 224
5.5.3 Empirical results 228
5.5.4 Conclusions 238

Chapter 6: Concluding Remarks **242**
6.1 Impact and Further Developments of the Research 243
6.1.1 Personal injury compensation in the UK and the Ogden Tables 243
6.1.2 Mortality models for heterogeneous insurance based populations 249
6.1.3 Modelling and forecasting mortality within a Lee–Carter type framework 250
6.2 Overall Summary 251

REFERENCES **253**

City, University of London
Northampton Square
London
EC1V 0HB
United Kingdom

T +44 (0)20 7040 5060

THE FOLLOWING PARTS OF THIS THESIS HAVE BEEN REDACTED FOR COPYRIGHT REASONS:

Chapter 2: Overview of Calculating Personal Injury Damages in the UK pg. 19 – 54

Chapter 3: Modelling Loss of Earnings Multipliers in the UK and the Ogden Tables (up to and including the 5th edition) pg. 55-114

Chapter 4: Mortality Models for Heterogeneous Insurance Based Populations pg.116-168

Chapter 5 (partial): Modelling and Forecasting Mortality within a Lee-Carter Type Framework pg.169-215

List of Tables

3.1 Grouping of the INECACA variable into 3 types of economic activity and the make up of the sample by gender over the period of 1993 – 2003. 64

3.2 Distribution of the respondents' current economic activity by gender across all ages (based on the definitions of the International Labour Organisation) over the period of 1993 – 2003. 92

3.3 Distribution of the respondents' main economic region by gender and economic state across all ages and the respective grouping of this into three broad regions over the period of 1993 – 2003. . . . 93

3.4 Distribution of the respondents' main industry sector by gender and economic state across all ages and the respective grouping of this into three broad occupation industries over the period of 1993 – 2003. 94

3.5 Labour Force Survey indicator of the respondents' highest qualification and the distribution of the sub-grouped variabled by gender and economic state over the period of 1998 – 2003. 95

3.6 Distribution of the respondents' current disability level by gender and economic state across all ages over the period of 1998 – 2003. 96

3.7 Cox PH model regression coefficients and standard errors for economic regions and type of transitions. 97

3.8 Cox PH model regression coefficients and standard errors for industry sectors and type of transitions. 98

3.9 Cox PH model regression coefficients and standard errors for educational qualifications and type of transitions. 99

3.10 Overall worklife expectancy to pension age 65 and 60, for males and females, respectively, over the period of 1993 – 2003. Empirical estimates. 100

3.11 Males worklife expectancy to pension age 65 conditional on the starting state (S_x) and adjusted by regions of residence over the period of 1993 – 2003. 101

3.12 Females worklife expectancy to pension age 60 conditional on the starting state (S_x) and adjusted by regions of residence over the period of 1993 – 2003. 102

3.13 Males worklife expectancy to pension age 65 conditional on the starting state (S_x) and adjusted by industry sector over the period of 1993 – 2003. 103

3.14 Females worklife expectancy to pension age 60 conditional on the starting state (S_x) and adjusted by industry sector over the period of 1993 – 2003. 104

3.15 Males worklife expectancy to pension age 65 conditional on the starting state (S_x) and adjusted by highest qualification over the period of 1998 – 2003. 105

3.16 Females worklife expectancy to pension age 60 conditional on the starting state (S_x) and adjusted by highest qualification over the period of 1998 – 2003. 106

3.17 Males worklife expectancy to pension age 65 conditional on the starting state (S_x) and adjusted by current disability level over the period of 1998 – 2003. Empirical estimates. 107

3.18 Females worklife expectancy to pension age 60 conditional on the starting state (S_x) and adjusted by current disability level over the period of 1998 – 2003. Empirical estimates. 108

4.1 Extent of the analyzed data sets (millions): 126

4.2 Examples of Gompertz–gamma model parameter settings corresponding to three low– and one high–risk populations. 151

5.1 Overall distribution of subgroups in the pooled pensioners data set and age–specific prevalence rates within subgroups (%). 219

5.2 Age– and calendar year–specific exposures recorded in the pensioners only data set (PN) over the period 1999 – 2007 (person–years). 221

5.3 Age– and calendar year–specific exposures recorded in the CMI pensioners data set over the period of 1999 – 2003 (person–years). . . 225

5.4 Fitted and projected life expectancy at age 65 by gender using the LC model with Poisson errors and ARIMA(0,1,0) time series forecasts of the κ_t parameter based on private pension scheme, CMI and E&W mortality data sets. 234

List of Figures

3.1 Diagram of data concentration for the period of Spring 1993 – Winter 2003. 63

3.2 Prevalence rates in the labour market per age across three economic states. 65

3.3 The three transient states of the labour force dynamics model. . . 68

3.4 Smoothed transition intensity between the employment (1) and unemployment (2) states for males and females over the two half periods of 1993 – 2003. 70

3.5 Worklife expectancy conditional on being in the employed state at age x for males and females over the two half periods of 1993 – 2003. 76

3.6 Worklife expectancy conditional on being in the inactive state with long–term sickness or disability at age x for males and females over the two half periods of 1993 – 2003. 77

3.7 Reduction factors (2.5 %) in the employed state conditional on the starting state (S_x) for males and females over the two half periods of 1993 – 2003. 79

3.8 Reduction factors (2.5 %) conditional on the starting state (S_x) and on the economic region for males and females over the period of 1993 – 2003. 83

3.9 Reduction factors (2.5 %) conditional on the starting state (S_x) and on the industry sector for males and females over the period of 1993 – 2003. 84

3.10 Reduction factors (2.5 %) conditional on the starting state (S_x) and on the highest qualification for males and females over the period of 1998 – 2003. 85

3.11 Reduction factors (2.5 %) conditional on the starting state (S_x) and on the type of disability for males and females over the period of 1998 – 2003. Empirical estimates. 88

3.12 Observed and smoothed transition intensity from the employed (1) state for males and females during 1993 – 2003. 109

3.13 Observed and smoothed transition intensity from the unemployed (2) state for males and females during 1993 – 2003. 110

3.14 Observed and smoothed transition intensity from the inactive (3) state for males and females during 1993 – 2003. 111

3.15 Crude and smoothed transition probability from the employed (1) state for males and females over 1993 – 2003, conditional on being alive at current age x. 112

3.16 Crude and smoothed transition probability from the unemployed (2) state for males and females over 1993 – 2003, conditional on being alive at current age x. 113

3.17 Crude and smoothed transition probability from the inactive (3) state for males and females over 1993 – 2003, conditional on being alive at current age x. 114

4.1 Deviance $D(\bar{\boldsymbol{\mu}};\, \hat{\boldsymbol{u}}_a)$ vs. fixed values of parameter a for Female Annuitants. 125

4.2 Gompertz and frailty distribution parameters for Female Annuitants Duration 1+ (age range of 40 – 100+). 129

4.3 Gompertz and frailty distribution parameters for Female Annuitants Duration 5+ (age range of 40 – 100+). 130

4.4 NLM Fitting of Perks, Modified Perks and Gompertz–inverse Gaussian models for Male Annuitants Duration 5+ 1953–56. 131

4.5 Difficulties encountered with the non–linear parameter (a) in the GLM type fitting with parameterised link function. 133

4.6 Results of GLM-P model fitted to age range 60 – 100 applied to data sets with (1) and without (2) outliers. 134

4.7 Results of GLM-P model fitted to age range 60 – 100 applied to data sets with(1) and without(2) outliers. 137

4.8 Age–specific mean frailty of populations with different degrees of gamma type heterogeneity (δ) and Gompertz individual hazard rates. 149

4.9 Age–specific relative hazard rates ($\overline{\mu}_r$ and μ_r) between a low–risk and three high–risk populations, assuming gamma distributed frailty and Gompertz individual hazard rates. 152

4.10 Probability density function of the gamma frailty distribution under different shape parameters δ ($\mu = 1$). 162

4.11 Probability density function of the inverse Gaussian frailty distribution with different shape parameters $\lambda = 2\,\psi$ ($\mu = 1$). 165

5.1 Illustration of CMI (lives) pensioners mortality experience. 194

5.2 Illustration of randomised CMI (lives) pensioners mortality experience: central exposures and log central mortality rates by additional covariate (X). 198

5.3 LC regression parameters for CMI male pensioners (lives) for age range 50 – 100 over the observation period of 1983 – 2003. 203

5.4 LC cross–classified fitted values for CMI male pensioners (lives) for age range 50 – 100 over the observation period of 1983 – 2003. 204

5.5 LC standardised deviance residuals for CMI male pensioners (lives) for age range 50 – 100 over the observation period of 1983 – 2003. 206

5.6 LC future log mortality rates values for CMI male pensioners (lives) for age range 50 – 100 over a 20–year prediction horizon. . 207

5.7 Illustration of LC forecast over a 20–year prediction horizon with 90 % CI for CMI male pensioners (lives). 208

5.8 APC regression parameters for CMI male pensioners (lives) for age range 60 – 95 over the observation period of 1983 – 2003. 211

5.9 SLC regression parameters for artificially stratified CMI male pensioners (lives) for age range 50 – 100 over the observation period of 1983 – 2003. 214

5.10 Illustration of forecast result over a 20–year period in the SLC modelling framework. 216

5.11 Pensioners only (PN) log central mortality rates for age range $50 - 110$ over the observation period of $1999 - 2007$. 222

5.12 All members (PN and beneficiaries) log central mortality rates for age range $40 - 110$ over the observation period of $1999 - 2007$. . . 223

5.13 CMI pensioners log central mortality rates for age range $50 - 110$ over the observation period of $1990 - 2003$. 225

5.14 England and Wales log central mortality rates for age range $60 - 89$ over the observation period of $1990 - 2006$. 226

5.15 LC regression parameters for private pension scheme males for age range $60 - 95$ over the observation period of $1999 - 2007$. 230

5.16 LC regression parameters for private pension scheme females for age range $60 - 95$ over the observation period of $1999 - 2007$. . . . 231

5.17 LC regression parameters for CMI male pensioners for age range $60 - 95$ over the observation period of $1990 - 2003$. 232

5.18 LC regression parameters for England and Wales for age range $60 - 89$ over the observation period of $1990 - 2006$. 233

5.19 Stratified LC regression parameters for private scheme pensioners for age range $60 - 95$ over the observation period of $1999 - 2007$. . 237

5.20 Fitted and projected life expectancy at age 65 by gender using the SLC model with Poisson errors based on private pension scheme mortality experience for age range $60 - 95$ over the observation period of $1999 - 2007$. 238

5.21 Deviance residuals of LC regression for private scheme male pensioners for age range $60 - 95$ over the observation period of $1999 - 2007$. . 239

5.22 Deviance residuals of LC regression for private scheme female pensioners for age range $60 - 95$ over the observation period of $1999 - 2007$. 240

List of Abbreviations

AP	Age–Period (Lee–Carter type) logistic model
APC	Age–Period–Cohort (Lee–Carter type) logistic model
BHPS	British Household Panel Survey
CI	Confidence Interval
CMI	Continuous Mortality Investigation
DHSS	Department of Health and Social Security
ELT	English Life Tables
ESRC	Economic & Social Research Council
EW	Worklife expectancy matrix
E&W	England and Wales
GAD	Government Actuary's Department
GLM	Generalised Linear Model
IFA	Institute and Faculty of Actuaries
ILO	International Labour Organisation
LC	Lee–Carter logistic model
LFS	Labour Force Survey
ML	Maximum Likelihood (graduation method)
PH	Proportional Hazards model
PN	Normal Pensioners
R	R programming software
RSS	Royal Statistical Society
SLC	Stratified Lee–Carter type logistic model
SVD	Singular Value Decomposition (graduation method)

Acknowledgements

This thesis is based on material that has already been published in peer reviewed journals. In particular, much of Chapters 2 and 3 were originally published in Verrall et al. 2005 and Butt et al. 2006, respectively. Similarly, a large part of Chapter 4 was published previously in Butt and Haberman (2004). The author would like to thank the anonymous referees for their insightful comments and suggestions that has led to the improvement of the published papers. The author is also grateful to the ESRC Society Today and ASTIN Bulletin for giving their permission for the papers to appear in the thesis.

The author would like to thank for the principal financial support from the ESRC (Grant RES-000-22-0883: "Quantifying involuntary non-participation in the labour market in England and Wales") and also for the additional contribution from the Institute and Faculty of Actuaries for the research work in Part I.

Further, the author wishes to thank to the Institute and Faculty of Actuaries, to the CMI Bureau and to Lucida Plc for the financial support to the research work in Part II. The author is also grateful to the CMI Bureau and Lucida Plc for providing the mortality data.

The author would like to express his appreciation to his colleague, Professor Steven Haberman, who has provided invaluable assistance and advice throughout the entire duration of this project.

The author also extends his gratitude to two of his colleagues, Professor Ben Rickayzen and Professor Celia Glass, for their continuous support, inspiration and encouragement to complete this thesis.

Last but not least, the author is ever so grateful to his family for their love and patience. I dedicate this thesis to my wonderful daughter, Chloe, who has unreservedly let me follow my path.

Declaration

I hereby grant powers of discretion to the University Librarian to allow the 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

Actuarial aspects of two important fields of insurance are considered: calculating personal injury damages for working age adults (liability insurance) and measuring the mortality in insurance based populations (e.g. life insurance and pensions). The contribution of the thesis is to demonstrate a wide array of modelling techniques and their practical implementation in these two key areas of actuarial science.

The first part considers the modelling of the labour force dynamics from the perspective of the loss of earnings multipliers in England and Wales. It reviews the estimation methods of involuntary non–participation in the labour market in relation to future loss of earnings. In response, a robust multiple state modelling methodology is developed that allows conditioning on personal characteristics of working age individuals such as disability, educational attainment and the current employment state. Applied to UK longitudinal Labour Force Survey data, it quantifies the disadvantages that plaintiffs with post–injury earnings capacity face in the labour market. This practical modelling framework leads to a set of improved loss of earnings multipliers in subsequent editions of the Ogden Tables now used in the Courts in England and Wales.

The second part focuses on the modelling and estimation of mortality rates using Poisson likelihood maximisation methods. In terms of graduation, it undertakes a comprehensive assessment of the frailty models and their implications. Then it puts forward suitable parametric modelling structures in order to measure the scale of individual heterogeneity and applies generalised linear modelling graduation techniques to a large array of insurance based mortality data. In terms of forecasting, it considers the generalised Lee–Carter type modelling structures of Renshaw and Haberman (2006) and demonstrates their suitability for practical applications. Furthermore, it develops a novel stratified Lee–Carter model for the measurement of the effects of explanatory factors (other than age and time). An efficient programming package in R is provided for this class of modelling framework. Finally, a detailed analysis of the mortality trends observed in private pension scheme data serves as a case study.

Chapter 1

Introduction

1.1 Overview

The objectives of this thesis relate to the analyses and implementation of various statistical models within two important areas of insurance: liability insurance and mortality. Regarding liability insurance, the aim is to demonstrate practical ways in which the modelling and estimation of labour force dynamics can be used to investigate actuarial aspects of personal injury damages. Regarding the context of mortality, the goal is to consider the modelling and forecasting of the mortality rates of insurance based populations (e.g. annuitants, assured lives and pensioners).

Thus, the thesis falls into two Parts. Part I consists of two papers and concerns the modelling of transition intensities of multiple state models in order to calculate the loss of earnings multipliers, which in turn are used in estimating the pecuniary components of damages resulting from personal injury. Similarly, Part II is made up by two papers and involves the analysis and regression of important mortality models. The thesis has resulted in the following four research papers:

1. An investigative study on the current practice of estimating the loss of earnings in personal injury claims in England and Wales: The Ogden Tables and contingencies other than mortality
2. The impact of dynamic multi–state measurement of worklife expectancy on the loss of earnings multipliers in England and Wales
3. The application of frailty–based mortality models using generalized linear models
4. ilc: A collection of R functions for fitting a class of Lee–Carter mortality models using iterative fitting algorithms

The first two papers were both co–authored with Professor Richard Verrall and Professor Steven Haberman (my colleagues within the Faculty of Actuarial Science and Insurance) and were originally published in 2005 and 2006 as internal research papers in Cass Business School. However, they formed the nominated output reports of the ESRC funded research (Grant RES-000-22-0883) and were published again in 2006 on the ESRC "Society Today" web-site: http://www.esrcsocietytoday.ac.uk (Verrall et al. 2005, Butt et al. 2006).

The third paper was co–authored with Steven Haberman and was published in an ASTIN Bulletin in 2004 (Butt and Haberman 2004). Finally, the fourth paper was also co–authored with Steven Haberman and it was published in 2009 as an actuarial research paper in Cass Business School (Butt and Haberman 2009).

All the papers included in this thesis are self-contained publications; each has a detailed introduction and a corresponding literature review. Consequently, the reader is referred to these for a thorough background to the investigations stated therein. For the purposes of this thesis, the following sections provide a brief summary of the principal aspects of the four papers and the author's individual contributions to each.

1.1.1 Personal injury compensation in the UK and the Ogden Tables

In the papers presented in Chapters 2 and 3 (see Verrall et al. 2005 and Butt et al. 2006, respectively), we investigate the actuarial assessment of life–time labour market risks from the perspective of the loss of earnings multipliers in England and Wales. Loss of earnings multipliers are used by the legal profession to determine the financial value of future worktime, when allowing for mortality and labour market risks. In the calculation of the damages for future loss of earnings due to personal injury or wrongful death, account is taken of life time employment risks. That is, in a fair and correct compensation system, the Courts have to deduct from an individual's total future earnings an amount that is based on the length of time the claimant is likely to be out of employment, based on statistical averages observed across the working age population. The actuarial assessment of the labour market contingencies is carried out using current UK Labour Force Survey (LFS) data.

The loss of earnings multiplier constitutes for the Courts a simplified actuarial assessment of the worklife expectancy of the plaintiff discounted for the risk of mortality and early receipt of income, among other things. The worklife expectancy is the length of time a person is expected to spend economically active (i.e. employed) until the age of normal retirement (or early death). That is, it excludes from the lifespan until normal retirement all the likely non-active periods of his/her life such as unemployment, looking after a family or sickness. The loss of earnings multipliers are summarised in the so-called Ogden Tables and currently are disaggregated according to various factors including age, employment status, educational attainment and age at retirement. The historical development and the use of the Ogden Tables is discussed in detail in the above two papers included in this thesis.

In Verrall et al. (2005), we investigate the scientific literature of modelling labour market contingencies and we review the rationale and the suitability of the Ogden Tables in the light of improved methodology and data. Then in Butt et al. (2006), we construct an improved probabilistic model of estimating the worklife expectancies (i.e. loss of earnings multipliers) that capture the true dynamic nature of the labour market. The methodology takes advantage of the current LFS linked panel data sets, which allows the estimation and application of a multiple state model with three or four main economic states.

A novel aspect of the research presented in Butt et al. (2006) is that it quantifies earnings and employment risks over a life–time dimension through the application of actuarial methods in a labour economics context. The methodology usefully builds upon a previous study (Lewis et al. 2003) and addresses some potential biases in their methodology, namely the use of recall data and the use of annual (as opposed to quarterly) transitions. The methodological framework proposed by Butt et al. provides a simple and robust estimation process, which has not yet been explored in any of the labour market studies that we are aware of, and it yields results that are directly applicable to the assessment of damages in the UK Courts. There are a number of important aspects to note with respect to the methodology developed in this work:

- It maintains the practicality of the current system of multiplier–multiplicand used by the Courts in England and Wales;
- In the context of damages for personal injury it is particularly important to differentiate by disability, so our approach differentiates between the pre– and post–injury valuation of future loss of earnings;
- The empirical estimation is based on sound multiple state models methodology that captures the true dynamic nature of labour force movements in-and-out of economic states over a human lifetime;

- It allows the assessment of the effects of additional factors (such as region, industry, educational attainment and also disability) on the loss of earnings multipliers.

1.1.2 Mortality models for heterogeneous insurance based populations

The mortality modelling and analysis presented in Chapter 4 (see Butt and Haberman 2004) follows in the footsteps of Vaupel et al. (1979) and provides a detailed discussion of the issues surrounding the heterogeneity in populations and the frailty hypothesis. The topic is analysed from an actuarial standpoint, by adopting actuarial techniques and modelling frameworks to look into the extent of heterogeneity that might be present in insurance based populations. The parametric models developed and tested in this chapter hinge on the principles of frailty models and build on the basic multiplicative model introduced by Vaupel et al.. Thus, in Chapter 4 we investigate a particular family of survival models, known as random effects or dynamic models, that allow for unobserved heterogeneity. These models capture the systematic bias at individual level that results from measuring failure rates from overall population data.

Conventional mortality modelling ignores the effects of unobserved heterogeneity based on the assumption that the population is homogeneous, at least within the observable sub-groups. However, as it is demonstrated in section 4.7.1, in the case of theoretical models of failures occurring over long periods of time, there are startling consequences when this effect is overlooked. It is shown that due to the resulting population compositional dynamics (i.e. selection effects), inferences from the observable *average* population mortality rate might not reflect correctly on the individual risk of failure without some additional knowledge or assumptions about the population make-up.

The time aspect in this process is crucial, since it gives rise to a systematic downwards bias in the estimation of the individual mortality rates, due to a characteristic of the overall population rates that is often referred to as the mortality-rate plateau. Numerous recent studies in both non–human (see Carey et al. 1992, Pletcher and Curtsinger 1998, Drapeau et al. 2000) and human (see Kannisto et al. 1994, Horiuchi and Wilmoth 1998, Thatcher et al. 1998) survival have demonstrated that the rate of increase of population mortality rate slows down significantly at older ages. While alternative biological explanations do exist for the phenomenon of falling failure rates at older ages (in particular morbidity rates, whereas the organism, for instance, might *acclimatise* to the conditions causing a disease), the assumption of unobserved heterogeneity and its effects seems to be more plausible and easier to include in mathematical models. For instance, Horiuchi and Wilmoth (1998) have tested extensively the deceleration effect of the frailty hypothesis using cause-specific mortality data from Sweden and Japan and found that: "Although we cannot dismiss alternative explanations [physiological, evolutionary or reliability-theoretical], some of the findings in this study seem to support the heterogeneity hypothesis more strongly than the individual risk hypothesis."

The estimation process considered in this study is based on the ML approach applied in the usual mortality modelling methods. While, in effect, the overall (population) hazard rate is estimated in the traditional way (i.e. based on the homogeneity assumption), it is still possible to make some inferences about the mixing distribution, by taking into account the way the original parametric models of the overall hazard rate are constructed (see section 4.2).

1.1.3 Modelling and forecasting mortality within a Lee–Carter type framework

For Chapter 5, we turn our attention to graduating and projecting mortality rates in the framework of Lee–Carter (LC) type logistic models. The main goal of the fourth paper considered therein (Butt and Haberman 2009) is to make use of the good track record and popularity among researchers of this modelling approach. The graduation methodology adopted in this investigation can be classified within the same group of GLM models (with parameterised log-link and Poisson distribution of errors) that is used elsewhere in the thesis (see section 4.3.3). In terms of forecasting, however, the methodology is part of the extrapolative stochastic methods. Specifically, it makes use of ARIMA processes to project forward in time the observed past trend of the time-dependent parameters. The LC class of mortality models and forecasting approach has attracted increasing attention in the specialist literature and gained recognition for producing relatively reasonable forecasts of life expectancy in most cases (see Booth et al. 2006).

While the standard LC logistic model works quite well, it is fairly restrictive in terms of parameterisation and of choice of distribution of the residuals. To avoid these shortcomings, a GLM modelling approach is put forward by Renshaw and Haberman (2006), which allows the formulation of more elaborate parameter constructs (e.g. the addition of cohort effects) and better distributional assumptions. Nevertheless, the new logistic model is still bilinear in form, which prevents the direct maximisation of the quasi-log likelihood of the GLM model. Instead, the parameters of the model are now estimated by a Newton-Raphson iterative method that is applied to the GLM deviance function, conditional on the type of error distribution. The benefit of this approach is that we are able to generate parameter estimates based on (overdispersed) Poisson likelihood of the number of deaths (Butt and Haberman 2002).

A practical objective of the work is to implement the above modelling approach in a user friendly programming package in R software. That is, we aim to make available purpose built commands in R that could be applied to any mortality data in order to investigate the class of models described in Renshaw and Haberman (2006). The package aids the user to carry out easily a full regression analysis in three distinct stages: data preparation, computing parameter estimates and assessment of results.

The first stage involves actions such as: loading and displaying raw data sets, closing-out procedures to smooth out inconsistent data points and choosing data ranges, etc.. Then, in the main, second, stage, the package can carry out the actual graduation of six different types of log-link GLM models with Poisson or Gaussian errors, that includes, as a special case, the standard LC model too. In the final, third, stage, the user can make use of simple auxiliary methods to complete the analysis, like goodness of fit tests, compute age–specific life expectancy based on the graduated or the fitted rates and flexible plotting of results with many control parameters.

The final part of this chapter is devoted to the application of the LC graduation package to actual mortality data from a large private pension scheme. Thus, by making use of these modelling tools, we carry out a comparative analysis of the mortality trends observed in the private pension scheme data against both England and Wales national data and a pool of other pensioners data.

1.2 Individual Contributions to Papers included in the Thesis

As previously mentioned in section 1.1, this thesis is based to a varying extent on four co–authored academic papers, which have previously been published elsewhere. All of the published material included in this thesis is a result of extensive research in which I was involved from the start, making significant contributions. In the following sections, I give an in-depth account of my personal contributions to each of these studies.

1.2.1 Contributions to the work presented in Chapters 2 and 3 (Part I)

Early in 2004, Richard Verrall and Steven Haberman were contacted by the Ogden Working Party seeking collaboration with researchers in order to carry out the re-assessment of the Ogden Tables multipliers. At the time, the Working Party was concerned that the Courts may be making wrong decisions on damages because of the absence of up-to-date research into the impact of disability and unemployment on the loss of earnings multipliers and were keen to commence such work as a matter of some urgency. They highly valued the previous pivotal research of Steven Haberman in the assessment of involuntary worktime loss, which was reported in Haberman and Bloomfield (1990) and which formed the basis of the earlier editions of the Ogden Tables. They all considered that a City University team, with the full support of the actuarial profession, should make an ESRC grant application to secure funding for this research.

Richard Verrall was the leading researcher and a co–author of the working papers introduced in section 1.1 and also of the follow–up paper published in the Journal of the Royal Statistical Society (Butt et al. 2008). His main responsibility was to oversee the progress of the project and to provide theoretical insights into the applied methodology. Steven Haberman was involved in a joint supervisor capacity with expertise on both theoretical and applied aspects of the project. He was also a co–author of the above mentioned academic papers. In addition, Professor Chris Daykin was involved in an advisory capacity from the Government Actuary's Department (GAD) on the applied aspects of the project, although he did not contribute to the writing of any of the academic papers.

Being already employed at the time within City University as a Research Assistant and having good data analysis and statistical knowledge, I was approached by Richard Verrall and Steven Haberman to work on this project. My mathematical background did not include crucial actuarial concepts (e.g. life expectancy, actuarial tables, multiple state models, etc.) that were necessary for this research. Therefore, I have studied the primary reference material pointed out to me by them. Furthermore, I have carried out a broad literature review on the subject. Thus, I have consulted electronic libraries and accessed various publications in order to find relevant papers in the field of damages law, labour economics and actuarial science that provided an adequate background to this project (see Chapter 2).

Consequently, we decided that it would be a great advantage to combine their extensive experience in actuarial modelling with my practical skills of data analysis. We held regular meetings to plan future work, to discuss the key ideas (such as the estimation methods of the transition intensities of a multiple state model from longitudinal data sets) and to clarify some theoretical aspects.

My initial task was only to carry out background research into the topic and also a brief feasibility study about the available UK labour force data. Preliminary data analysis comprised of assessing what could be achieved in terms of methodology and applications. For example, I carried out a feasibility study of multiple state modelling approach by making use of the latest LFS joined panel data based on five quarterly longitudinal observations (see section 3.3).

Once I had gained sufficient knowledge of the issues involved, I became involved with drafting the initial research plan and the ESRC grant application. The application was overseen and amended by both Steven Haberman and Richard Verrall, who were jointly the team leaders of this project. In the end, the ESRC grant application was finalised in close collaboration by the three of us with some additional support of Professor Chris Daykin, who was at the time a leading member of the Ogden Working Party.

Subsequently, after the ESRC funding was obtained, I was engaged full-time throughout the remaining part of this project. I was responsible for conducting the empirical investigations and handling of the data sets. My work primarily involved the collection and preparation of the data from the UK Data Archive, carrying out statistical investigations of the regression models and writing up progress reports. In the concluding stage of the project, I participated in authoring and drafting reports and then condensing the main contributions into academic articles (e.g. end-of-award reports to ESRC). I co–authored both working papers and various other peer reviewed publications (see Butt et al. 2008, 2009, 2010).

An important aspect of my work was to carry out all the data preparations and investigations using statistical software (e.g. S-Plus and MS Excel). The analysis involved the manipulation of voluminous micro–economic data sets and the fitting of regression models. I was also responsible for the dissemination of empirical results involving the full regression analysis, tabulation and graphical presentation of the regression results, preparation of data output and summary for use in a new set of tables (for both ESRC and Ogden Tables). Most of this work was carried out using statistical software with programs developed by myself.

Data collection, tabulation and preparation for statistical analysis has involved much independent work that I had to carry out, such as:

– sorting and loading of raw data sets;

– study of the data description material (e.g. list of variables, variable definitions, etc.);

– extraction and cleaning of relevant variables (see section 3.7);

– investigation for bias and/or compatibility of different variables;

– study of the prevalence of a given set of variables in the quarterly data sets;

– analysis of statistical summaries and compilation of results, etc.

Nonetheless, other than most of the practical aspects related to data analysis, I took an active role in the theoretical aspects of the project as well as in the dissemination of the research findings. In terms of methodology, we agreed that, ideally, we should use the multiple state modelling approach based on three transient economic states (mortality could not be modelled directly from LFS). Thus, I needed to devise a way of transforming the observations on the ILO economic states in the longitudinal LFS data into transition intensities of the multiple state model dependent on additional covariates. Also, I worked out the mathematics of calculating the age–specific worklife expectancies and reduction factors from the model by using matrix algebra, which simplified and speeded up the calculations (see section 3.4.1). Then, I applied the methodology by developing suitable program codes that extracted and summarised the necessary data for the analysis. This approach permitted us to apply established regression methods (e.g. Cox

Proportional Hazards model) and I have analysed and presented the results for Steven Haberman and Richard Verrall to comment upon in our meetings.

1.2.2 Contributions to the work presented in Chapter 4 (Part II)

Originally, Steven Haberman was interested in the modelling of the time dynamics of mortality rates and to explore the effect of heterogeneity in insurance based populations. In the autumn of 1998, he proposed a practical study using existent CMI insurance data to develop the basic multiplicative modelling structure put forward by authors like Vaupel et al. (1979) and Horiuchi and Wilmoth (1998). Consequently, he has obtained necessary funding from the IFA and the CMI Bureau and he recruited me as a Research Assistant to work with him on this project part time over a term of 12 months (the other half of my time went towards building a simulation model of Income Protection insurance), which I started in early 1999.

My background in mathematical sciences helped me to face up to the theoretical challenges brought by this project. In the early stages of this project, I spent a significant part of my time studying the extensive theoretical background to mortality models, with particular focus on frailty models. Using the guidance of Steven Haberman, I was able to explore and to understand the main concepts involved. By applying adequate parameterisation of the population hazard rates (see section 4.3), I managed to fit the frailty models to the insurance data using standard non–linear regression methods. However, due to the parametric structure of the models, a direct GLM based regression methodology was not possible.

Initially, the work comprised of a laborious graduation exercise of two families of frailty models (i.e. Gompertz–gamma and Gompertz–inverse Gaussian) using a large set of insurance based mortality data (annuitants, assured lives and pensioners) provided by CMI (see section 4.4). We have found that the regression results were not entirely satisfactory and were sensitive to the choice of the fitted age range and/or to the grouping of calendar years. Further, some attempts have been made to find proof for the effect of heterogeneity in cohort based populations, but the size of the sub–grouped data seemed to be insufficient to find any conclusive evidence. Similar outcomes have been reached while trying to implement the empirical method of Horiuchi and Coale (1990), as their data was collected from national mortality investigations whereas we applied it to a smaller scale data collected by CMI from insurance companies (see section 4.3.2).

Therefore, it became important to explore the possibility of using non–standard GLM regression techniques with the added benefit of Poisson distributed errors and non–constant variance assumptions. Following consultation with Professor Arthur Renshaw, we decided to test two indirect GLM fitting methods: 1) using Taylor expansion of the predictor and 2) forming a parameterised link function, which have previously been applied successfully in Renshaw (1991) and Renshaw (1995), respectively. Then I managed to work out the parameterisation of our preferred population hazard rates model (Perks) using both of these approaches (see section 3.4 in Butt and Haberman 2002). Once these were verified and approved by Steven Haberman, I had to implement these in statistical software from scratch by using my own programs, as these were non–standard regression approaches.

In the latter part of this project, I also implemented, using statistical software, a series of graduation testing methods that were suggested by Steven Haberman (visual/graphical and statistical/actuarial) based on Forfar et al. (1988), which allowed us to evaluate the robustness of the choice of models and regression methods. In addition, we decided to use various sensitivity analysis techniques (e.g. choosing different age ranges, identifying outliers, grouping of calendar years, etc.). Furthermore, we also considered the modelling of future mortality trends in the presence of frailty using forecasting methods adapted to this framework (see section 4.7.3). However, this aspect could not be explored further due to time constraints on the project.

After receiving further funding from the CMI, I continued to work on the project for another year on a part–time basis. At the end of this period, I drafted the first working paper, which was verified and corrected in places by Steven Haberman and then submitted to the Faculty as an actuarial research paper in 2002 (Butt and Haberman 2002). Given the novelty of our approach and the lack of similar investigations into heterogeneity within populations of insureds, we believed that the study was relevant to the profession and it was worthy to be published in ASTIN Bulletin.

However, the Editor found the working paper to be too long and technical for the purposes of the journal and Steven Haberman has agreed to make it more concise. Thus, during 2003, he reviewed the paper and drafted a shorter version, which was commented upon and edited by myself. The revised paper was finally accepted for publication at the end of 2003 and eventually appeared in ASTIN Bulletin in 2004 (Butt and Haberman 2004). Later, I managed to further extend our investigation into the implications of the frailty models, and this work is presented in the thesis in section 4.7.1.

1.2.3 Contributions to the work presented in Chapter 5 (Part II)

In about 2008, given the ample interest that the LC type regression models described in Renshaw and Haberman (2006) had generated within the actuarial profession (e.g. particularly the age-period-cohort variant), Steven Haberman became interested in making their methodology readily available to a wider range of potential users, including students and practitioners. Since I was also keen to explore the recent advances in this field, I offered to develop a user friendly package of commands in R software, called ilc, that implemented the models and the graduation methodology applied in this paper (see section 5.4).

My initial plan was to use as templates for coding purposes some of the programs originally written by Arthur Renshaw in GLIM software, which meant that I had to learn a basic GLIM programming language. However, it quickly became apparent that it was easier and more direct to use my own programming methods and tools developed in R, due to the significant differences between R and GLIM (e.g. GLIM is a purpose built software for GLM modelling, whereas R is not).

In order for me to be able to implement the above models correctly in R, I carried out background research into the LC modelling topic and I made sure that I fully understood the Renshaw and Haberman (2006) approach. The graduation methodology is based on an iterative Newton-Raphson method applied to the deviance function of the fitted hazard model, which needed to be programmed separately for each model. Nonetheless, I wanted to find ways in which our package could be put to use most efficiently so that it integrated well with existent LC graduation tools within R. So, I made use of some of the features of an R package, called demography, which was built to graduate and forecast mortality rates within the standard LC framework (see section 5.4).

Once the ilc package was functional, we tested it against both Arthur's outputs in GLIM and the demography outputs in R. Then we made use of the package to analyse many mortality data sets (e.g. CMI pensioners in section 5.4.2). Similarly, we made use of the package to investigate a large pensioners data set provided by Lucida Plc (see section 5.5). The analysis of the pensioners' mortality was carried out jointly with Steven Haberman and Richard Verrall.

The company was also interested in us analysing the impact of additional factors (such as smoking or socio–economic indicators). However, this was not directly possible in any of the LC type models. Based on the methodology suggested in Renshaw and Haberman (2006), I have worked out the equations of adding an extra base parameter to the model and implemented this in the ilc package (see section 5.2.4). Also, I have developed tools to test this model by generating artificially randomised data. The regression tests using this approach were very encouraging (see section 5.4.3).

Originally, I drafted a short working paper, which was only intended to act as a user guide to the ilc package. However, later on, encouraged by the feedback from users, we felt that it was also helpful to provide a full description of the modelling framework alongside the package instructions. So, in 2009, under the guidance of Steven Haberman, I extended this further into an actuarial research paper that included a detailed presentation of the theoretical background and of the fitting methodology. The actuarial research paper was drafted by myself and commented upon by Steven Haberman, who was the co-author to this paper (see Butt and Haberman 2009).

Part I

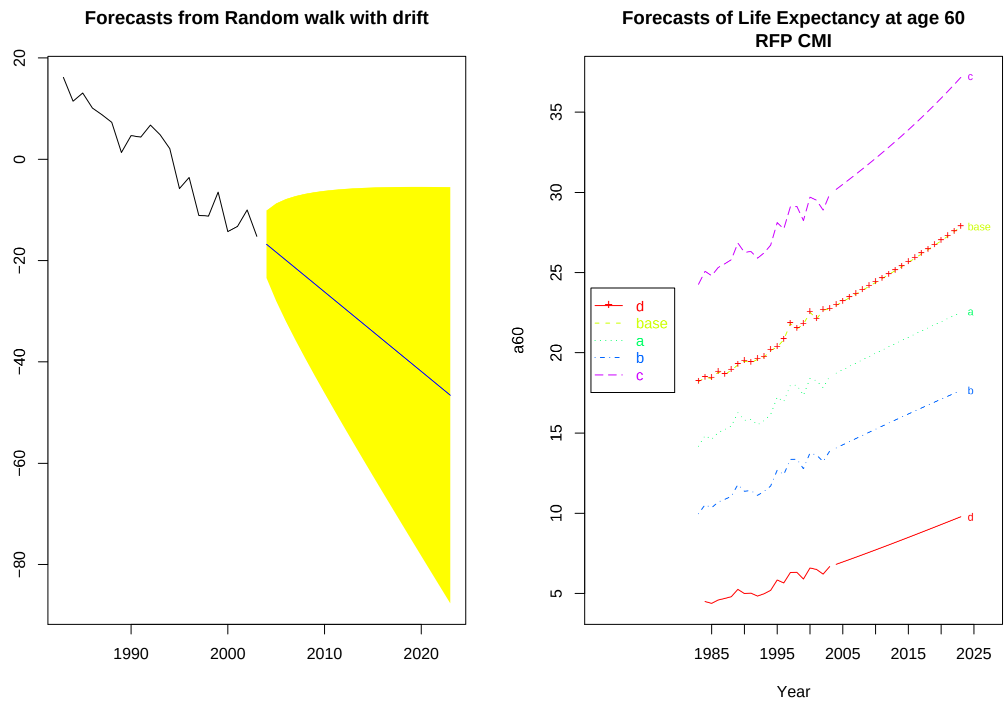
PERSONAL INJURY COMPENSATIONS IN THE UK AND THE OGDEN TABLES

The full text of this article has been removed for copyright reasons

Part II

MORTALITY MODELLING AND FORECASTING OF INSURANCE BASED POPULATIONS

The full text of this article has been removed for copyright reasons



a) Future trend parameter κ_t *b) Future life expectancy at age 60*

Figure 5.10: Illustration of forecast result over a 20–year period in the SLC modelling framework.

Thus, Figure 5.10 illustrates the resulting plots of predicted trend parameter (panel *a)*) and the future life expectancy at age 60 over a 20 year period (panel *b)*).

5.5 Application to Pensioners Data Set

The objective of the analysis presented in this section is to model the mortality improvements of pensioners belonging to a large private scheme in order to quantify the future longevity risk in terms of life–expectancy at retirement age. That is, we focus primarily on modelling the age–specific central mortality rates at retirement ages (i.e. ages above 60) when making forecasts, though, as far as the data allows it, we also consider lower (i.e. younger) age ranges to make informal comparisons. Further, the investigation makes use of the stratified LC model (see section 5.2.4) to explore the influence on longevity of additional explanatory factors other than age and period (such as region, pension amount, socio–economic group or type, etc.) available in the data set.

In the current analysis, we make use of an extensive set of demographic data that consist of over $320,000$ records of pensioners and their beneficiaries (i.e. widows, widowers and children) over the period of 1999 – 2007. In addition to the vital statistics and death records, the original data contained a large array of additional information related to region and socio–economic characteristics of the individuals in the sample.

Firstly, the raw data were checked for consistency of the individual entries, removing primarily all duplicate cases, but also any odd cases (such as those with missing dates, negative durations, child beneficiaries, etc.). Secondly, the individual experiences were transformed into age–specific counts of deaths and central exposures for each calendar year in order to calculate age–period central mortality rates for both the overall (i.e. all beneficiaries and pensioners) and the pensioners only data sets.

In the second part of the analysis, the age–specific mortality rates were examined for general features and the most suitable modelling approach was assessed based on the observed trends over calendar time. It was found that the LC type modelling framework presented the most structured and transparent approach to describe the age–period relationship of the log-mortality rates and to make stochastic forecasts of future mortality improvements. Nevertheless, following the work of Brouhns et al. (2002) and Renshaw and Haberman (2003a,b), it was considered that a theoretically more sound Poisson error assumption for the observed number of deaths provided a better alternative to the Gaussian error structure of the traditional LC framework (see section 5.2).

5.5.1 Pensioners data set

In this section, we describe the main features of the private pension scheme data set. In terms of year of pension commencement, the data span the period 1954 – 2007 and covers individual ages (including those of child beneficiaries) from as young as 2 years old up to 108 years old. The data set pools together individual vital statistics and membership dates of 321,111 pensioners and beneficiaries.[14] In addition, the data set provides detailed individual risk factors of the participants (e.g. pension amount, socio–economic group and type) and region (i.e. in the form of country of origin and post codes of main residence).

The overall sample is made up by four main subgroups with respect to the membership status of the participants at the end of the survey period (i.e. end of 2007) or at the date of death. Thus, the pooled data contain individual experiences of one of the following:

a) deceased (NL);

b) current pensioners (PN);

c) deferred pensioners (PP); and

d) widows, widowers or beneficiaries (BW).

However, given that the mortality data of the participants (i.e. the NL category)[15] were made available only with respect to the period starting from 01/01/1999 to 31/12/2007, the current mortality analysis refers only to this relatively short span of 9 calendar years (i.e. 1999 – 2007).

[14]This represent the total number of distinct cases, which excludes all the inconsistent and/or duplicate entries (between or within the subgroups). For instance, around 6.5 % of the records of the pensioners in deferment were also found in the current pensioners sample. Similarly, approximately 8 % of the (alive) beneficiary cases were also recurring among the deceased entries. However, for obvious reasons, there were negligible number of duplicate records between

Table 5.1: Overall distribution of subgroups in the pooled pensioners data set and age–specific prevalence rates within subgroups (%).

	Males (53.7)				**Females** (46.3)			
Age group (at last bd)	NL (15.9)	PN (53.0)	PP (28.3)	BW (2.7)	NL (16.8)	PN (31.0)	PP (31.9)	BW (20.2)
$\leq$ 55	6.8	4.7	85.7	30.4	3.0	4.4	88.7	9.1
56-60	4.2	15.2	14.2	5.9	1.8	6.9	11.3	4.7
61-65	6.1	23.8	0.0	6.3	2.8	17.6	0.0	7.1
66-70	9.3	17.9	0.0	7.0	5.5	15.5	0.0	9.6
71-75	14.7	14.8	0.0	10.9	10.5	16.2	0.0	13.6
76-80	20.3	11.3	0.0	13.3	20.0	15.8	0.0	18.5
81-85	19.7	7.7	0.0	13.8	25.8	13.8	0.0	19.2
86-90	13.5	3.4	0.0	9.9	20.3	7.5	0.0	13.0
91-95	4.7	1.0	0.0	2.2	8.9	1.9	0.0	4.5
$\geq$ 96	0.6	0.1	0.0	0.2	1.4	0.3	0.0	0.6

* NL – No Liability (Deceased) PN – Normal Pensioners
PP – Preserved Pensioners BW – Widow and Child Beneficiary

Table 5.1 summarises the composition of the full pensioners data set. It can be seen that the overall data are split approximately equally between male and female participants (54 % and 46 %, respectively), although the concentrations of the above mentioned subgroups are markedly different between the two sexes. That is, there are about twice as many male PNs ($0.537 \times 0.53 = 0.28$) as female PNs ($0.463 \times 0.31 = 0.14$), whereas there are about nine times as many female BWs as male BWs in the sample.

In addition, it can be observed that, in general, the female BWs tend to be distributed towards the older ages in contrast to their male counterparts (for example, 30 % of the male BWs are below age 55 compared to only 9 % in the case of female BWs). However, the prevalence of the NL and PP cases are roughly equal among the male and female participants. As a consequence, one needs to

the (alive) current/deferred pensioners and the deceased samples.

[15] Note that around 88 % and 50 % of this represents the mortality experience of male PNs and female PNs, respectively. Further, the next largest group of deceased cases is the BWs with around 7 % and 46 % prevalence of males and females, respectively.

allow for this significantly different heterogeneity between the male and female data sets when analysing the mortality characteristics within the entire sample.

The investigation has primarily focused on the PN mortality experience, although some comparative analysis was also conducted with respect to the combined (i.e. overall) pensioners data. From Table 5.1, we can see that the age distribution of the PN population seems to verify the patterns observed in other comparable data. Thus, as one might expect, there are more male early retirees than females based on an age for age comparison (i.e. in the first two age groups, below age 60), but even more so when allowing for the differential between the normal retirement ages of the two sexes. Further, there is a significantly greater concentration of female PN population at the old and very old ages than that observed in the male PN population.

The individual vital statistics were used to extract the number of deaths (y_{xt}) and the corresponding central exposures (e_{xt}) by gender and single years of age (last birthday) for ages $50 - 110$ and for $40 - 110$ based on the PN only and the combined data, respectively, for each calendar year over the survey period of $1999 - 2007$. Table 5.2 summarises the estimated central exposures based only on the PN data by gender and 5 year age groups for $1999 - 2007$.

The crude central mortality rates for the pensioners only and combined data sets between 1999 and 2007 are illustrated graphically on a logarithmic scale in Figures 5.11 and 5.12, respectively. These figures show that there is an overall linear increase in the log-mortality rates above age 60 for both males and females across all calendar years. However, there are significant variations in the log death rates from age to age, which gain in amplitude at ages below 60 and above 95, in particular in the case of females. Furthermore, it is interesting to observe that there is a distinct breaking point in the age–specific log rates at around age 60, which is most clearly observed in the case of males. This feature suggests different

Table 5.2 : Age– and calendar year–specific exposures recorded in the pensioners only data set (PN) over the period 1999 – 2007 (person–years).

Age group	1999	2000	2001	2002	2003	2004	2005	2006	2007
Males									
50-55	25659	23803	21174	17732	13648	10907	8428	6286	4494
56-60	18559	19196	19911	21267	22673	22196	20720	18740	15946
61-65	16920	17393	17618	17512	17509	17898	18460	19226	20570
66-70	14399	14597	14787	14960	15305	15786	16227	16516	16434
71-75	12787	12801	12733	12672	12603	12684	12865	13133	13380
76-80	9277	9736	9633	9779	10004	10162	10203	10262	10312
81-85	4506	4558	5113	5549	5833	6229	6582	6608	6870
86-90	1739	1919	2091	2274	2393	2358	2354	2743	3031
91-95	202	269	370	436	503	598	664	749	836
96-110	9	10	13	17	26	43	56	73	70
Females									
50-55	8028	6959	5747	4441	3292	2735	2301	1940	1589
56-60	7693	7733	7808	8023	8026	7260	6258	5158	3951
61-65	8064	7980	7819	7587	7455	7498	7508	7624	7835
66-70	8926	8667	8319	8124	7889	7713	7606	7443	7228
71-75	9128	8939	8721	8477	8299	8185	7939	7695	7539
76-80	7937	8507	8341	8197	8041	7860	7673	7544	7381
81-85	3413	3589	4383	5025	5542	6050	6466	6436	6391
86-90	1308	1549	1724	1910	2048	2115	2237	2826	3286
91-95	133	179	283	391	486	595	683	757	848
96-110	7	10	16	19	25	41	49	74	88

mortality profiles between early and normal pensioners.

Given that early retirement is usually associated with ill health, it is not unexpected to see higher mortality patterns for the group of early retirees.[16] Therefore, it is not surprising to find in the combined data (see Figure 5.12) that the normal retirees have lower mortality rates than those members (including beneficiaries) who start receiving pension at ages below 60. Although this shift in the age–specific log rates is accompanied by large variation, it nevertheless

[16]According to CMIR 21 (2004), since the mid-1980's, there are an increasing number of healthy lives taking out early retirement.

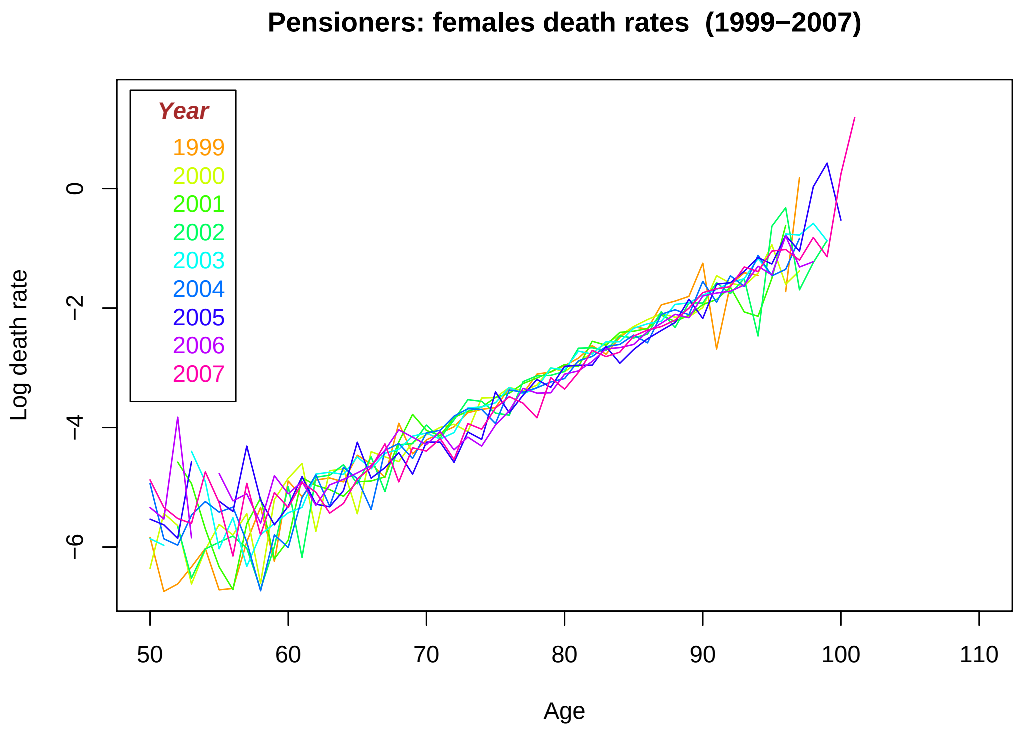


Figure 5.11 : Pensioners only (PN) log central mortality rates for age range 50 – 110 over the observation period of 1999 – 2007.

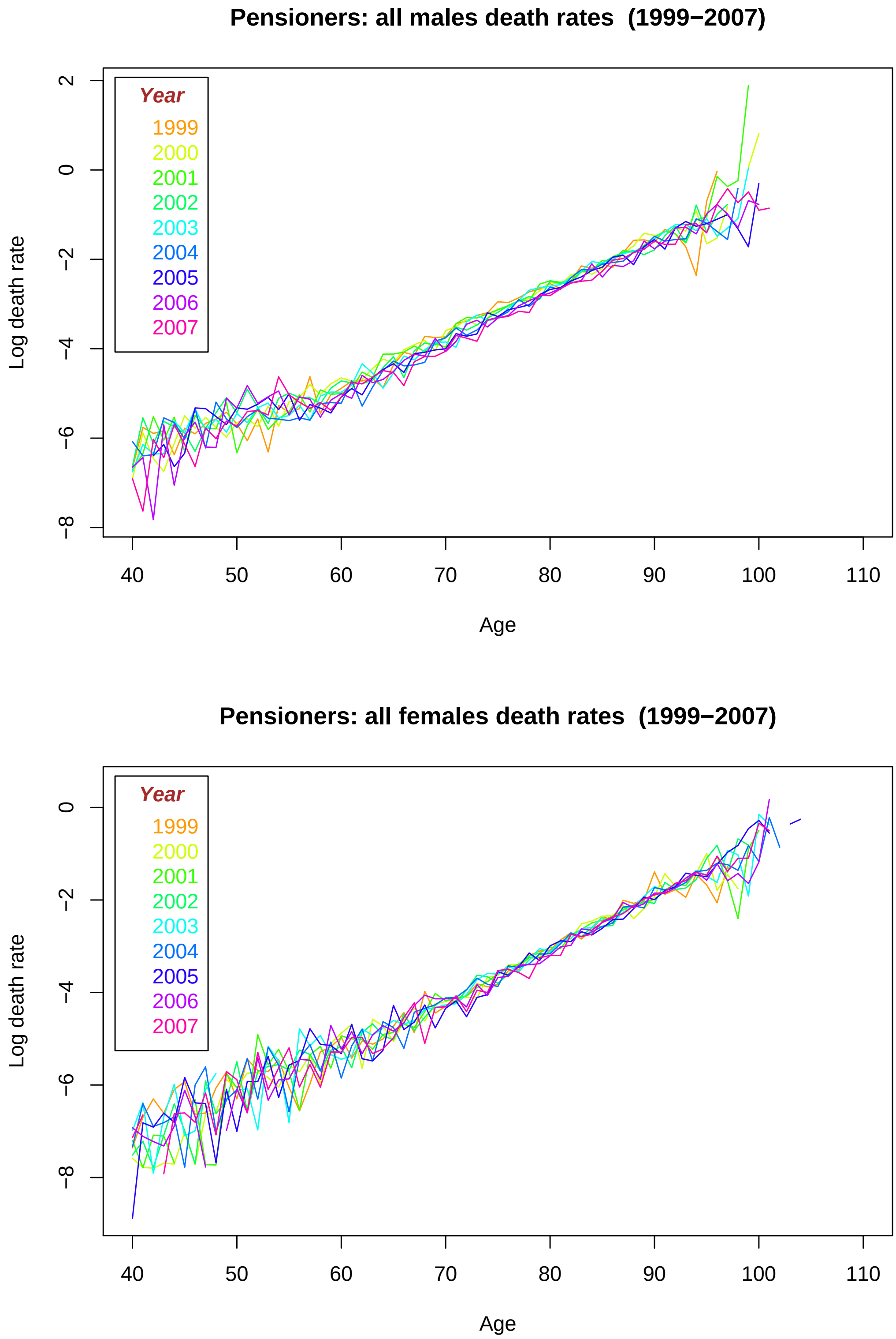


Figure 5.12 : All members (PN and beneficiaries) log central mortality rates for age range 40 – 110 over the observation period of 1999 – 2007.

suggests higher mortality rates for those retiring below age 60. We note also that the effect of early retirement is more striking in the case of males.

In order to avoid observation problems related to the inherent heterogeneity in the samples, we exclude from the analysis the age ranges where the selection effect is the strongest. Therefore, in the following parametric regression we restrict our analysis to data at ages above 60, which is found to provide the most reliable age–specific mortality features for both sexes. Similarly, in the evaluation of the erratic mortality rates at very old ages we have concluded that the most consistent results are obtained when we include the data only up to age 95.[17]

5.5.2 Other mortality data sets considered

As mentioned before, in this section we also report on the results of LC modelling to other alternative mortality data sets making use of the same age range as the one applied in the data presented in section 5.5.1. Thus, we model the CMI male pensioners mortality experience for individual ages 50 – 108 covering the period from 1983 to 2003. Similarly, we fit the LC type models to the overall population mortality experience of E&W for both genders for the retirement age range of 60 – 89 and over the period of 1990 to 2006.[18]

Both of these alternative data sets have been selected so that the age range is consistent with that used for the private scheme data in order to facilitate comparison. Given that the CMI data comprise the experience of pensioners

[17]We have attempted to cater for the data inaccuracies at the very old ages also by grouping the observations in wider age groups (e.g. 90 – 95, ≥ 95), but this approach seemed to produce less reliable results than making use of single ages between 60 – 95.

[18]We note that the available full E&W data set go back to 1980, but due to the stable nature of the data we decided to focus only on the most recent period.

Table 5.3: Age– and calendar year–specific exposures recorded in the CMI pensioners data set over the period of 1999 – 2003 (person–years).

Age group	1999	2000	2001	2002	2003
50-55	7423	7986	650	608	628
56-60	11596	12922	1370	1284	1546
61-65	35823	39292	10729	8607	9538
66-70	73389	84735	56196	40884	52951
71-75	60594	60810	52532	38984	54278
76-80	56926	57436	50180	37067	47074
81-85	38992	37230	34082	27616	35268
86-90	23500	24564	23052	18326	20976
91-95	6694	7108	7052	6286	8102
96-110	1182	1368	1420	1368	1598

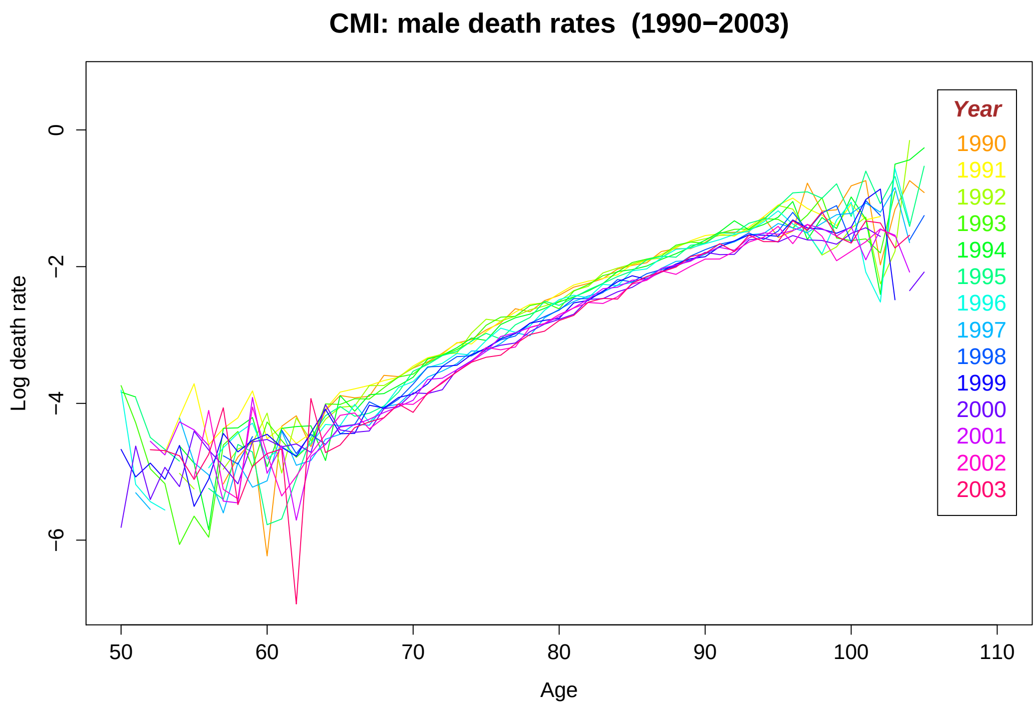


Figure 5.13: CMI pensioners log central mortality rates for age range 50 – 110 over the observation period of 1990 – 2003.

covered by policies issued by life offices in the UK, it is expected that the regression outcomes would be similar to those observed for the target data, albeit with

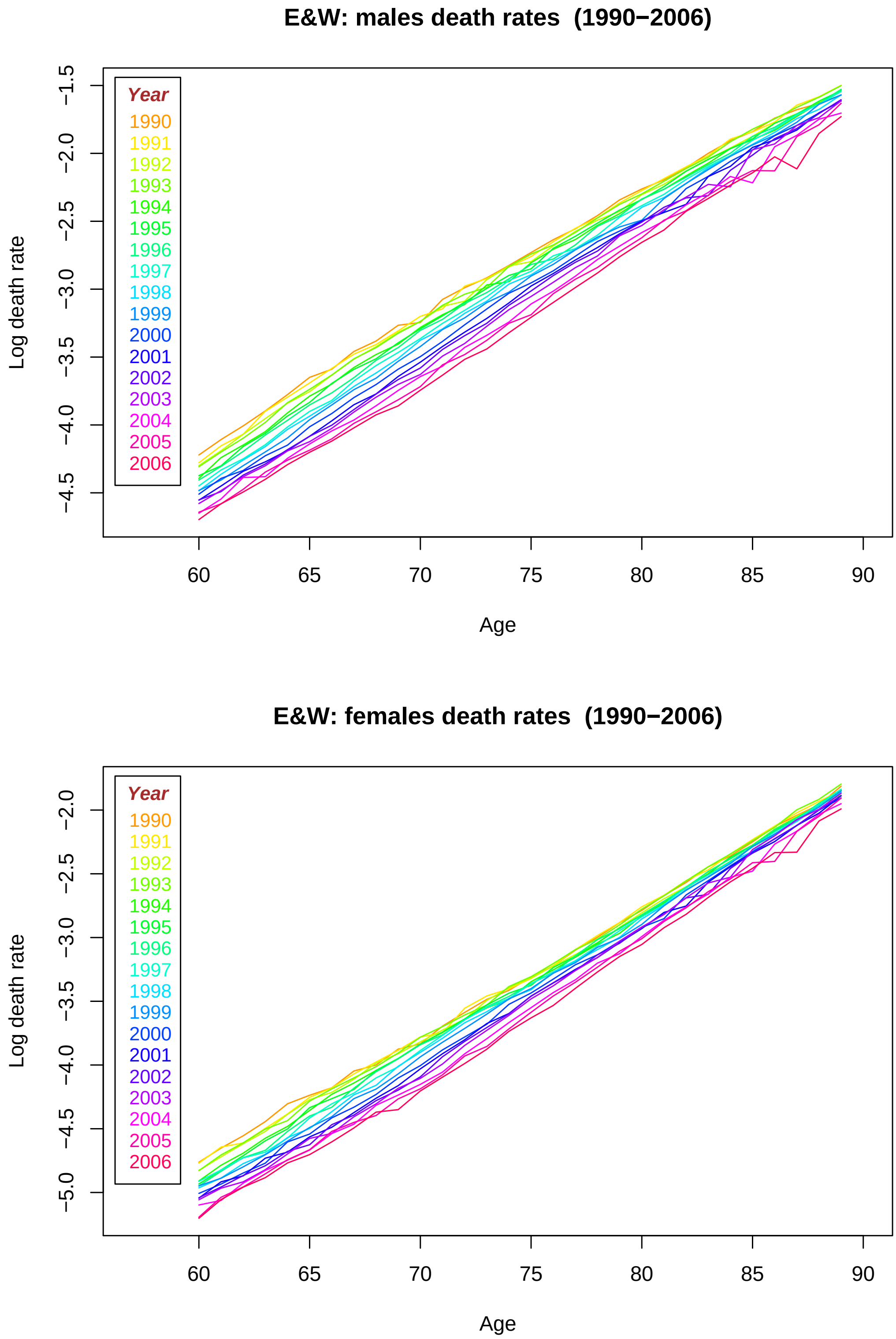


Figure 5.14 : England and Wales log central mortality rates for age range 60 – 89 over the observation period of 1990 – 2006.

smaller errors due to the larger exposures as can be seen in the data extract provided in Table 5.3. In contrast, the mortality characteristics of the national data set for E&W is likely to be markedly different from the private pension scheme experience in terms both of level and of smoothness.

Figure 5.13 presents the CMI male pensioners log-mortality curves over the period 1990 – 2003 and for all retirement ages between 50 – 110 (i.e. including both early and normal retirement life spans). The age–specific variation of the log force of mortality, in general, is less erratic and also the annual mortality improvements progress more evenly than in the case of the target data set. However, the slope and the size of the log-mortality rates are remarkably similar to those observed in the case of the corresponding male PN experience (see upper panel of Figure 5.11). Furthermore, we note that there is a slight curvature in the age–specific log-mortality rates, with a discernible *plateau* at the very high ages (not observable in the target data).

Figure 5.14 presents the E&W log-mortality rates for ages 60 – 89 over the period 1990 – 2006 differentiated by gender. The rates follow an almost linear progression by age on the logarithmic scale across the entire age range considered in this investigation. While there are some notable exceptions of uneven age–specific rates in the more recent calendar years, these occur mainly towards the older ages (i.e. above 80). Further, we can see that the mortality improvements are slightly greater over the younger ages (i.e. below 75), particularly in the case of male populations.

5.5.3 Empirical results

In this section, we consider the relative merits of the variants of the LC family of models (5.7) with respect to the private pension scheme, CMI and E&W mortality data sets. Moreover, we aim to investigate the robustness of the parameter estimates in terms of the fitted age ranges, in the light of the large variability in the mortality rates observed both at the early retirement ages (i.e. below age 60) and at the very old ages (i.e. above 90). Thus, we assess the appropriateness of the fitted parameters (for instance, one might expect a linearly increasing age effect, α_x, and a decreasing period effect, κ_t) and the forecasted life expectancies at age 65.

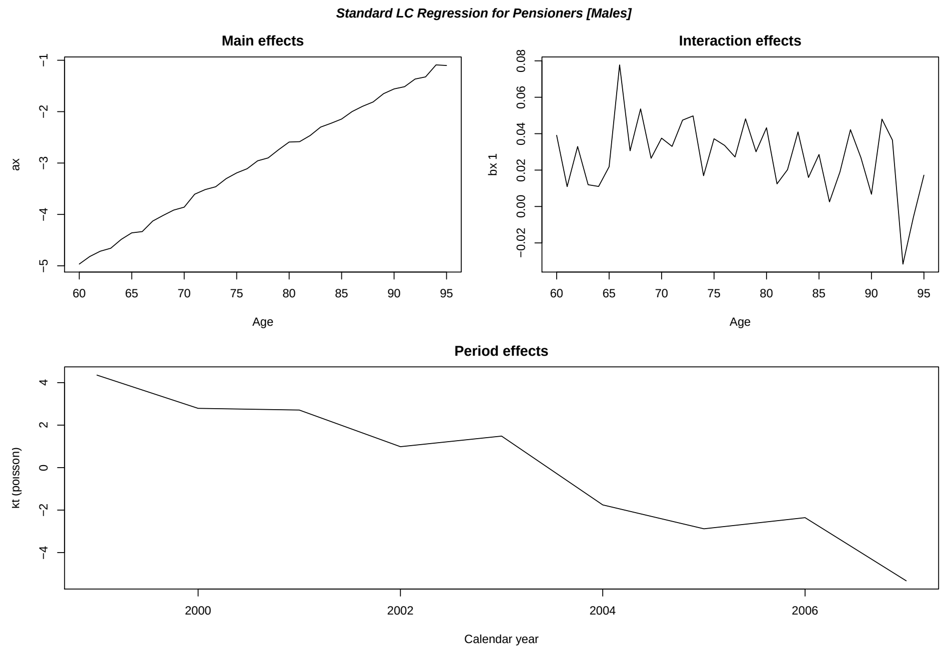
We note here that the goodness of fit of the models can reasonably be assessed by an analysis of the distribution of the standardised deviance residuals (see section 5.2.2). Thus, we plot the deviance residuals against age, period and cohort (i.e. year of birth $t-x$) in order to detect systematic patterns or a significant number of prominent outliers. The presence of any of these anomalies in the residual profiles would indicate an unsuitable model structure or parameterisation.

Judging from the estimated parameter values, it is evident that all the modelling structures containing cohort effects (i.e. $\iota_{t-x} \neq 0$) underperform the basic LC model for the given data sets (see sections 5.5.1 and 5.5.2). That is, in the case of models with cohort effects, the parameter estimates become highly erratic and show unjustifiable patterns (e.g. a sharp slump at age 65 in the age effect α_x) or take extreme values. Furthermore, the main age–period–cohort model (5.7) demonstrates very slow convergence in the deviance value and consequently the parameter estimates often depend on the chosen convergence criterion. These (undesirable) features are less significant in the case of the much larger and extensive E&W data set, which might indicate a stronger cohort effect present in the national population data than in the private scheme data (see Renshaw and Haberman 2006),[19] or reflect the relative sizes of these data sets.

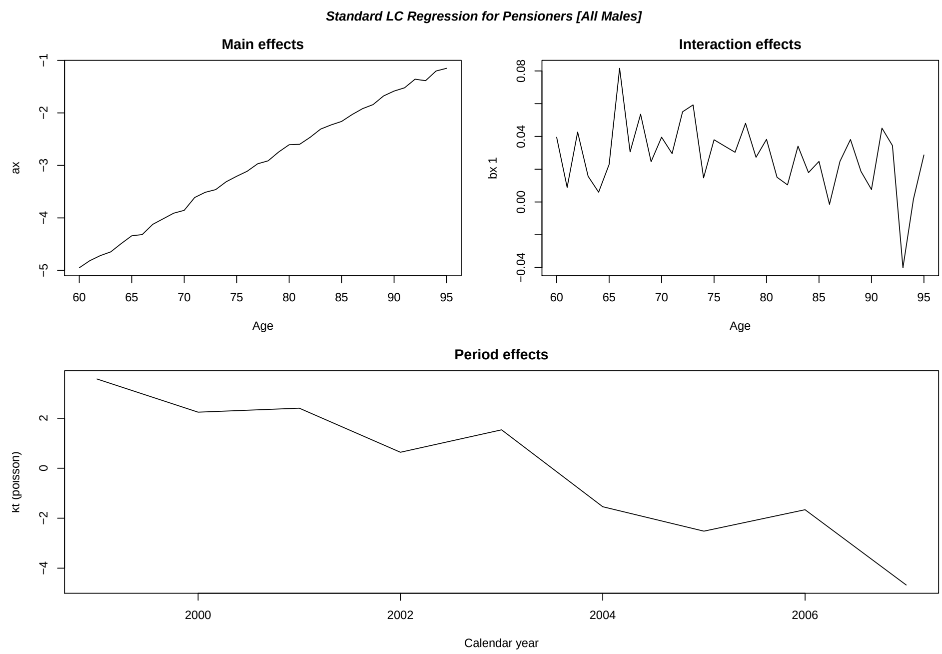
In terms of model sensitivity to the fitted age ranges, we can report that the outcomes are strongly affected by the inclusion of the data observed at ages below 60 and above 95. Thus, with the inclusion of pensioners data at younger ages, the period (κ_t) and interaction effects (β_x) become inconsistent (for instance, the regression might indicate negative improvements in mortality over calendar time). Similarly, when we make use of the data at very old ages, the interaction effects are greatly distorted, experiencing a very steep drop at ages above 95. Therefore, we choose to restrict the fitted age range to 60 – 95, which provides the most stable and reliable parameter estimates.

Thus, we illustrate the regression results for the basic LC model with a Poisson error structure for the given pensioners by gender in Figures 5.15 and 5.16 for ages between 60 and 95. Thus, looking at these fitted parameter plots, we can observe that, in the case of females, the model shows distinctively smaller improvements of the mortality rates by period compared to males. On the other hand, there is a greater variation in the age (α_x) and the interaction effects (β_x) for females than for males, which could be due to the greater heterogeneity noted in section 5.5.1 in the case of the female samples. Nevertheless, the corresponding residual plots for the pensioners by gender, shown in Figures 5.21 and 5.22, demonstrate that the overall performance of the model in terms of goodness of fit is not affected by the greater variability present in the female mortality rates. Furthermore, we can see in the residual plots that there are no distinguishable patterns by year of birth, which explains the weak performance of the models (noted earlier) when the additional cohort effect is included.

[19]While fitting the cohort effects (ι_{t-x}) after fixing the age and period effects might improve the performance of the models, we were satisfied that this would bring little benefits to the overall outcomes and have not pursued such optimisation methods further.

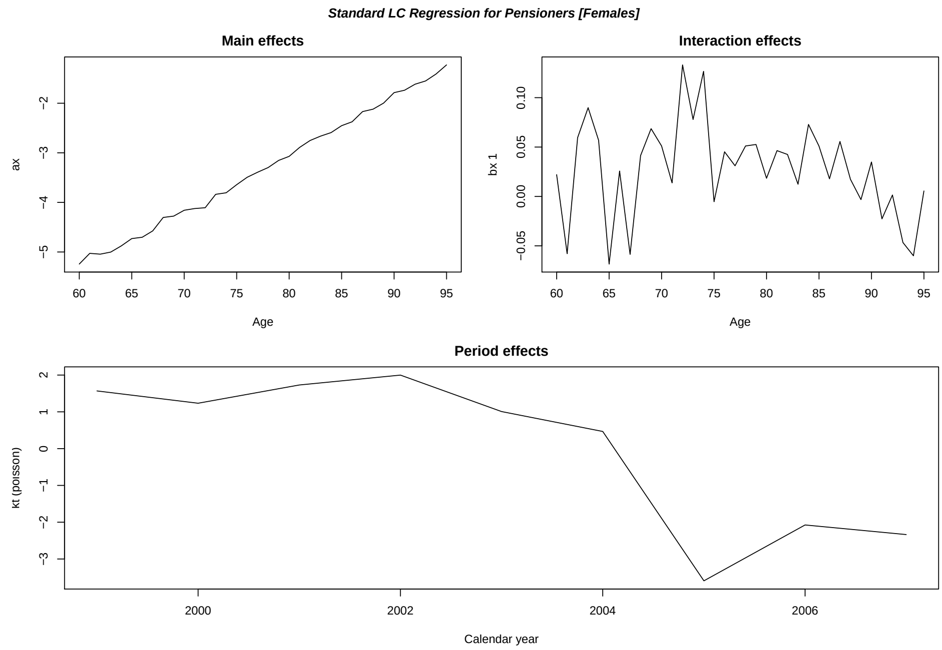


a) Private Pension Scheme: Male PN Pensioners

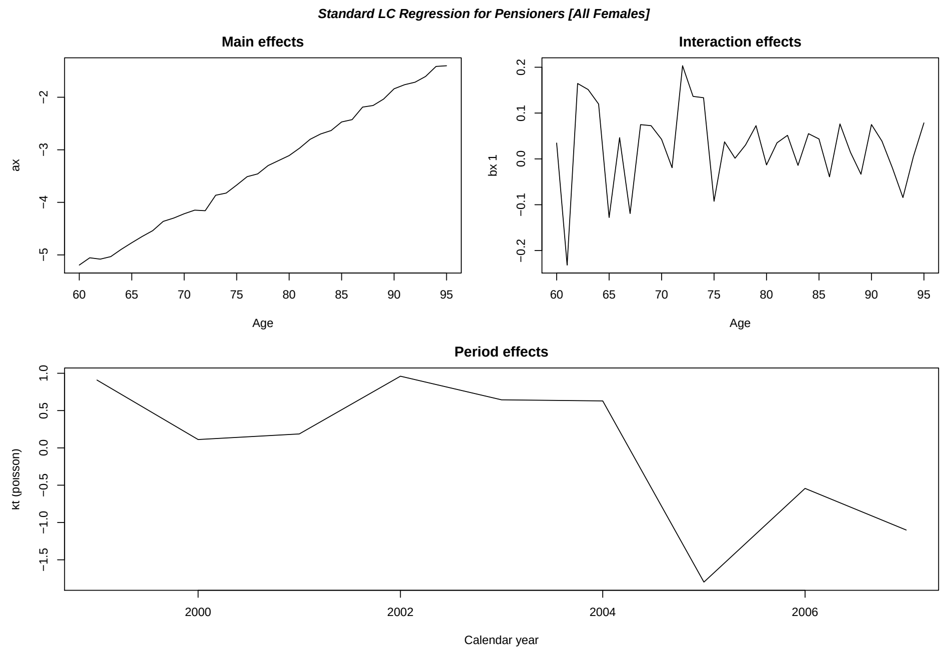


b) Private Pension Scheme: All Males

Figure 5.15 : LC regression parameters for private pension scheme males for age range 60 – 95 over the observation period of 1999 – 2007.



a) Private Pension Scheme: Female PN Pensioners



b) Private Pension Scheme: All Females

Figure 5.16 : LC regression parameters for private pension scheme females for age range 60 – 95 over the observation period of 1999 – 2007.

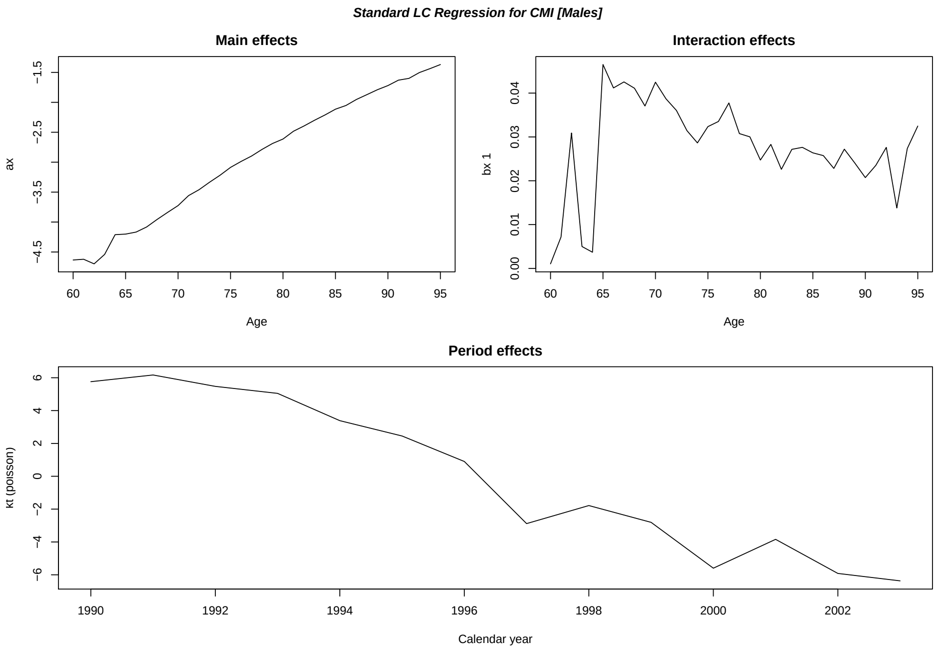
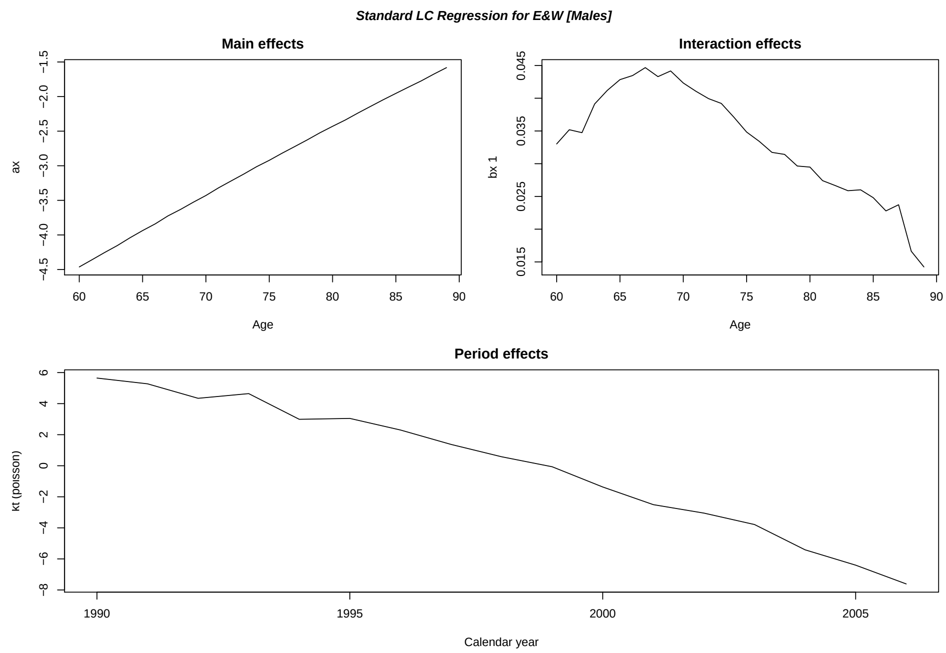
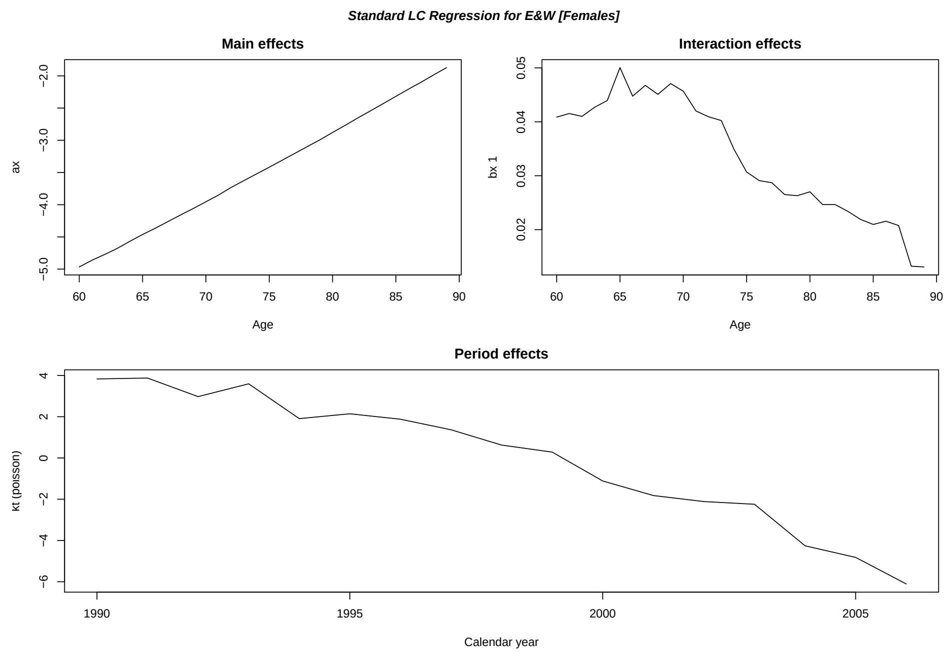


Figure 5.17: LC regression parameters for CMI male pensioners for age range 60 – 95 over the observation period of 1990 – 2003.

Additional plots of the LC parameter estimates are illustrated for CMI pensioners and E&W population in Figures 5.17 and 5.18, respectively. When comparing the regression results for the private scheme and the CMI male pensioners, we note many similarities between the corresponding parameter values. In particular, the slope of the period effects (κ_t) are approximately the same, which implies the same rate of mortality improvements over time in the two data sets. Furthermore, the interaction effects (β_x) exhibit similar patterns with respect to age (e.g. the improvements seem to be the largest for ages 65 – 75 and gradually decreasing towards the older ages). Interestingly, the same interaction pattern is also clearly visible in the results for the E&W population data, although in much smoother form. Nevertheless, we note that the slopes of the period effects in the E&W regressions tend to be smaller than those resulting from the corresponding insurance–specific data.



a) E&W Males



b) E&W Females

Figure 5.18 : LC regression parameters for England and Wales for age range 60 – 89 over the observation period of 1990 – 2006.

Table 5.4 : Fitted and projected life expectancy at age 65 by gender using the LC model with Poisson errors and ARIMA(0,1,0) time series forecasts of the κ_t parameter based on private pension scheme, CMI and E&W mortality data sets.

Year	**Pensioners**				**CMI**	**E&W**	
	All		**PN**				
	M	F	M	F	M	M	F
1990					15.7	14.2	18.4
1991					15.6	14.3	18.4
1992					15.8	14.5	18.6
1993					15.9	14.4	18.4
1994					16.3	14.9	18.9
1995					16.5	14.8	18.8
1996					16.9	15.0	18.9
1997					17.8	15.3	19.0
1998					17.6	15.5	19.2
1999	16.6	20.2	16.4	19.7	17.8	15.6	19.3
2000	16.9	20.4	16.7	19.8	18.5	16.0	19.6
2001	16.9	20.4	16.8	19.7	18.1	16.3	19.8
2002	17.3	20.2	17.2	19.6	18.6	16.4	19.9
2003	17.1	20.3	17.0	19.8	18.7	16.6	19.9
2004	17.8	20.3	17.8	20.0	*19.0*	17.0	20.4
2005	18.0	20.8	18.0	20.9	*19.2*	17.3	20.5
2006	17.8	20.5	17.9	20.6	*19.4*	17.6	20.8
2007	18.5	20.7	18.6	20.6	*19.7*	*17.8*	*21.0*
2008	*18.7*	*20.7*	*18.8*	*20.7*	*19.9*	*18.0*	*21.1*
2009	*18.9*	*20.8*	*19.1*	*20.8*	*20.2*	*18.2*	*21.3*
2010	*19.2*	*20.8*	*19.3*	*20.9*	*20.4*	*18.4*	*21.4*
2011	*19.4*	*20.8*	*19.6*	*21.0*	*20.7*	*18.6*	*21.6*
2012	*19.6*	*20.9*	*19.8*	*21.1*	*20.9*	*18.8*	*21.7*
2013	*19.8*	*20.9*	*20.1*	*21.2*	*21.2*	*19.0*	*21.9*
2014	*20.0*	*21.0*	*20.3*	*21.2*	*21.5*	*19.2*	*22.0*
2015	*20.2*	*21.0*	*20.5*	*21.3*	*21.7*	*19.4*	*22.2*
2016	*20.4*	*21.0*	*20.7*	*21.4*	*22.0*	*19.6*	*22.3*
2017	*20.6*	*21.1*	*20.9*	*21.5*	*22.3*	*19.9*	*22.5*
2018	*20.8*	*21.1*	*21.2*	*21.5*	*22.5*	*20.1*	*22.6*
2019	*20.9*	*21.1*	*21.4*	*21.6*	*22.8*	*20.3*	*22.7*
2020	*21.1*	*21.2*	*21.6*	*21.6*	*23.1*	*20.5*	*22.9*

* In the table the numbers in italics represent the forecasted life expectancies.

As shown in section 5.2.5, it is possible to make use of the LC modelling framework to forecasts future mortality by projecting the period effects (κ_t) further in time using a basic time series approach. Thus, we apply an ARIMA(0,1,0) type time series model to predict future improvements and then we make use of the projected rates to calculate future life expectancy at age 65. Table 5.4 illustrates numerically the fitted and projected life expectancy at age 65 based on the LC type regression models. Because of the short span of data available, we feel that it is not reasonable to make forecasts for a time horizon longer than a few years and the values presented are mainly for illustrative purposes up to year 2020.

Looking at the male life expectancy values contained in Table 5.4, we can see that both the current and future life expectancy of the CMI pensioners are higher than their private scheme counterparts. In turn, the male private scheme pensioners seem to have slightly greater life expectancy predictions than for the E&W population. On the other hand, judging from to the ARIMA forecast, the female private scheme pensioners are likely to experience shorter life expectancy in the future than the E&W population. Nevertheless, we should note that the forecasting of the female mortality rates is less reliable due to the unstable nature of the period effects coupled with the high level of irregularity in the interaction effects (see Figure 5.16).

In order to illustrate the impact of additional covariates on the fitted mortality rates, we make use of the stratified LC type model presented in section 5.2.4. We note that this modelling framework can be extended further to include more than one additional effects. However, from extensive empirical trials (not reported here), we have found that it would bring little improvements in terms of goodness of fit while making the model interpretation more difficult. Thus, we have fitted consecutively each of the additional factors contained in the private pension scheme data, with the exception of geographical region[20], namely:

- pension amount — factorized in 3 (or 4) levels that have been defined subjectively, using breaking points based on the distribution of the original variable;

[20]The extensive number of postcodes precluded the identification of efficient grouping factors of this variable. Therefore, further research is needed into adequate methods to deal with the subgrouping of the geographical variable, possibly based on spatial smoothing techniques.

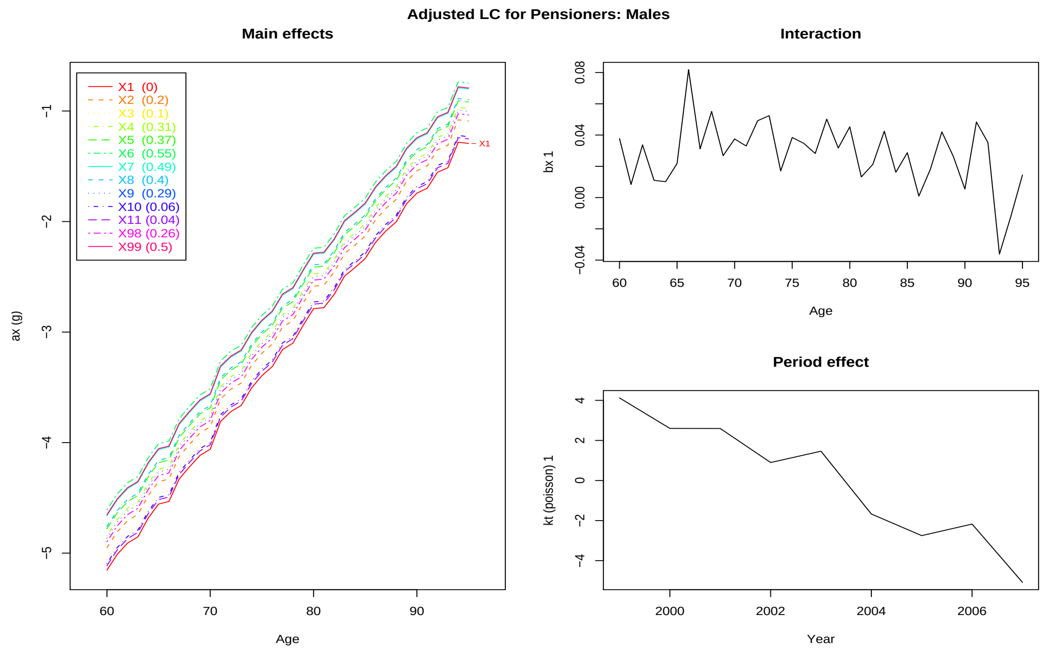
- socio–economic group — an indicator of 13 distinct levels;
- socio–economic type — an indicator of 63 distinct levels.

The regression results suggest that (individually) both the pension amount and the socio–economic group variables have a significant impact on the overall mortality of the observed pensioners. In contrast, the effects of the socio–economic type variable are not so distinguishable and also present practical problems in terms of over–segmentation of the mortality experience, giving rise to a large number of empty data cells. As a consequence, the latter might need to be further subgrouped in order to be useful for modelling purposes.[21] However, given the subjective nature of the subgrouping of the pension amount variable and the fact that the socio–economic group factor already contains some information related to personal wealth, we opted to focus on the latter.

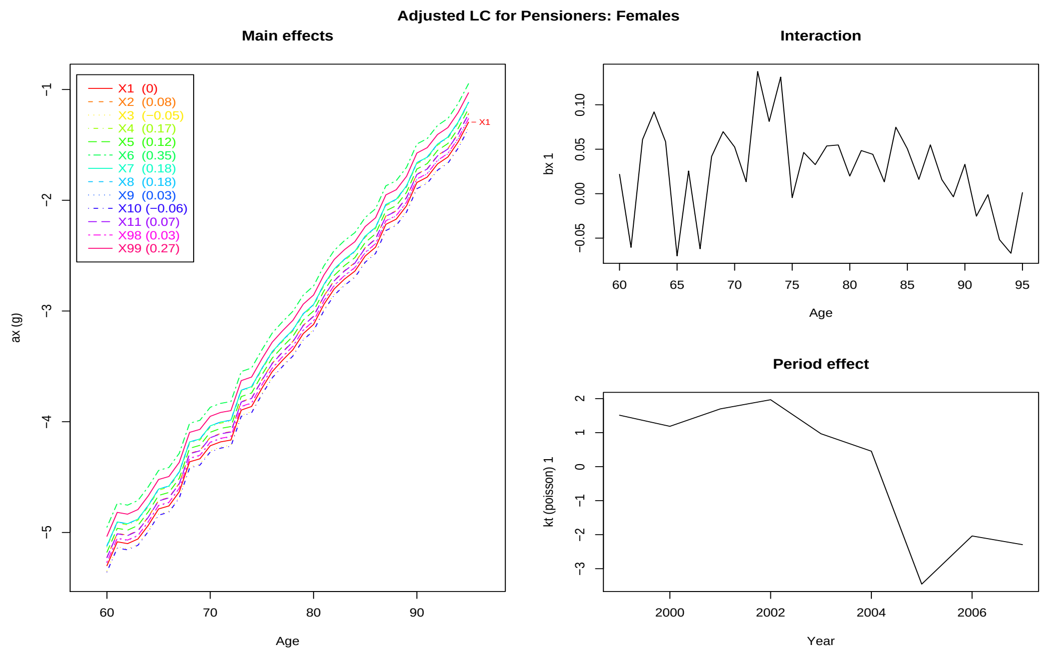
Figure 5.19 illustrates the fitted parameter values by gender of the stratified LC modelling in the presence of the socio–economic group effect. When compared to the basic LC model outcomes (see Figures 5.15 and 5.16), we can see that the period and interaction effects are almost identical in the two models, whereas the main age effect is now stratified in order to represent the relative differences in mortality between the subgroups. We note that the fitted additional effects are rescaled such that the first level is always taken as the base value, effectively corresponding to the original age effect α_x in the basic LC model (see section 5.3.3). Thus, the additional effects represent the overall (absolute) deviations from the base mortality level on the log scale.

Assuming that the additional effects remain constant in time (similarly to the age and the interaction effects), forecasting of future longevity can proceed based on the same time series methods used in the basic LC approach (see section 5.2.5).

[21] Nevertheless, this is not necessary for the current data given the relatively good performance of the socio–economic group variable.



a) Private Scheme Pensioners: Males PN



b) Private Scheme Pensioners: Females PN

Figure 5.19 : Stratified LC regression parameters for private scheme pensioners for age range 60 – 95 over the observation period of 1999 – 2007.

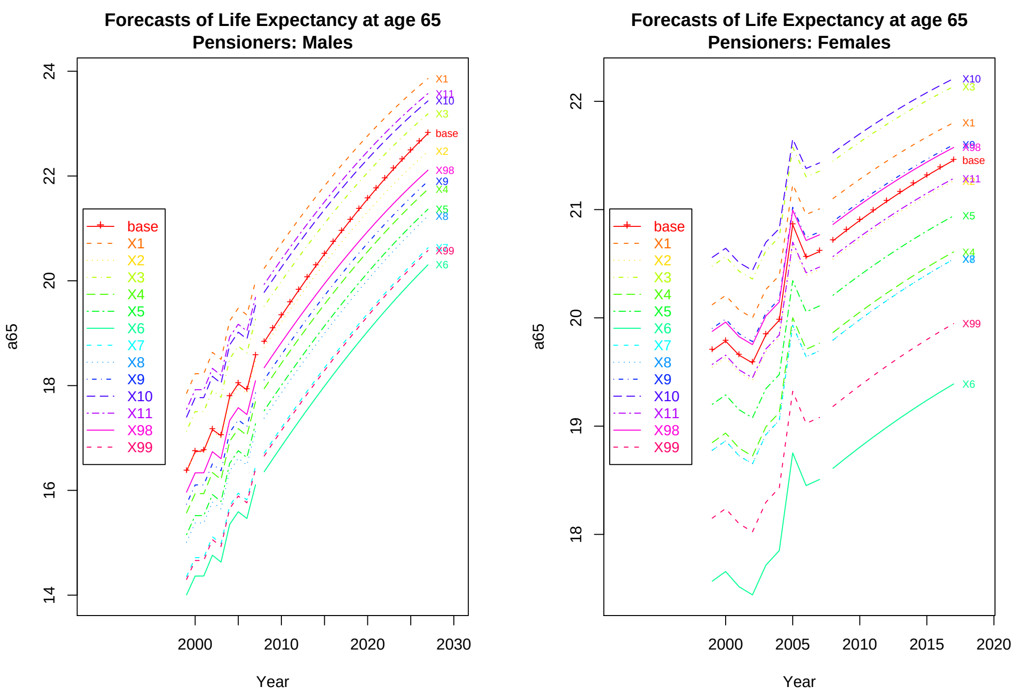


Figure 5.20 : Fitted and projected life expectancy at age 65 by gender using the SLC model with Poisson errors based on private pension scheme mortality experience for age range 60 – 95 over the observation period of 1999 – 2007.

Thus, forecasted $\dot{\kappa}_{t_n+s}$ mortality improvements are applied together with the additional effects in order to calculate the corresponding life expectancies of the subgroups at age 65 in year $t_n + s$ $(s > 0)$. The results of this procedure are illustrated in Figure 5.20.

5.5.4 Conclusions

The modelling of the age– and period–specific mortality observed in the private pension scheme data was carried out within a LC type family of functions with Poisson error structures. Thus, in the current analysis, we applied a unified modelling framework and iterative fitting methodology proposed by Renshaw and Haberman (2006). This is a GLM setting that makes use of theoretically sound

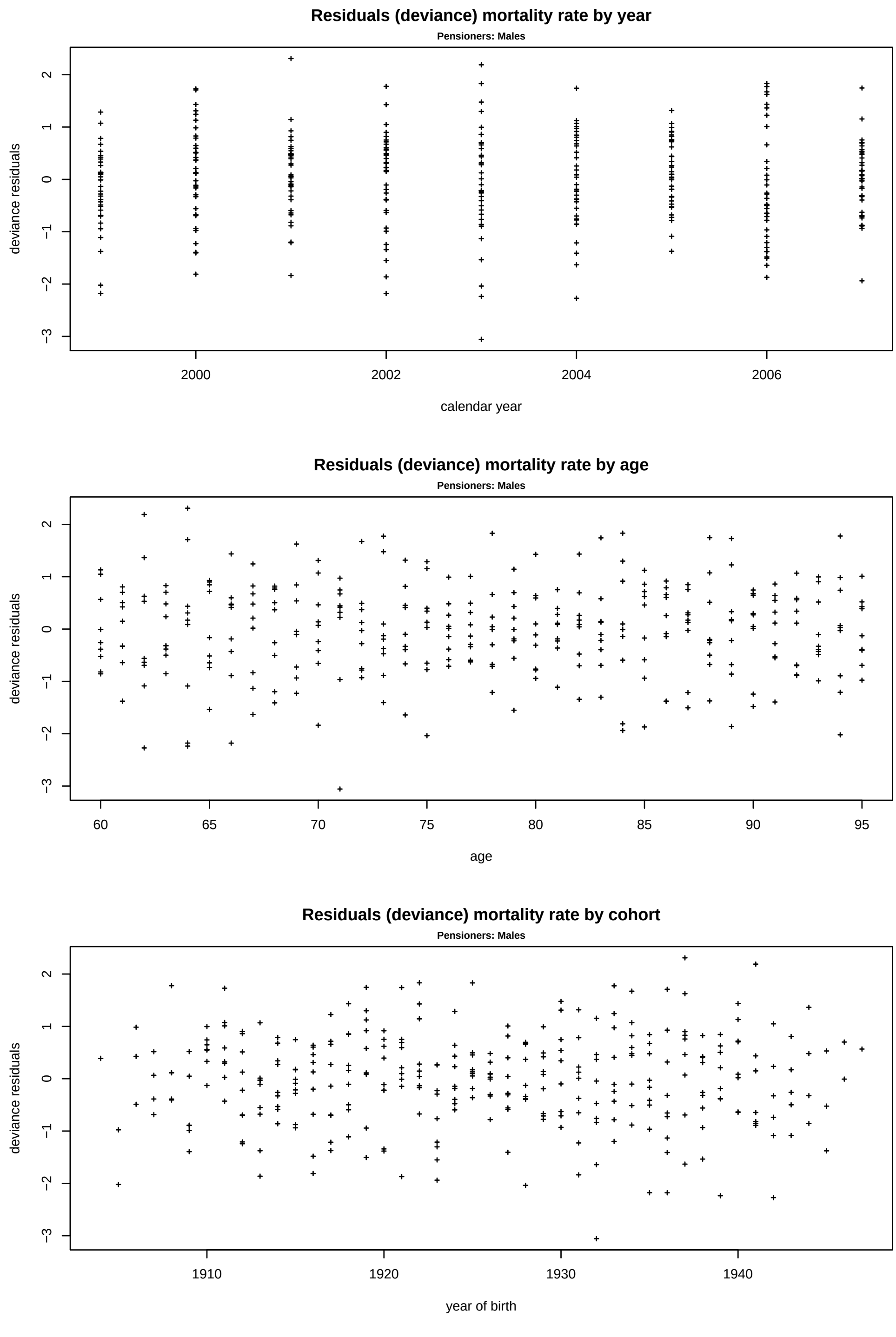


Figure 5.21: Deviance residuals of LC regression for private scheme male pensioners for age range 60 – 95 over the observation period of 1999 – 2007.

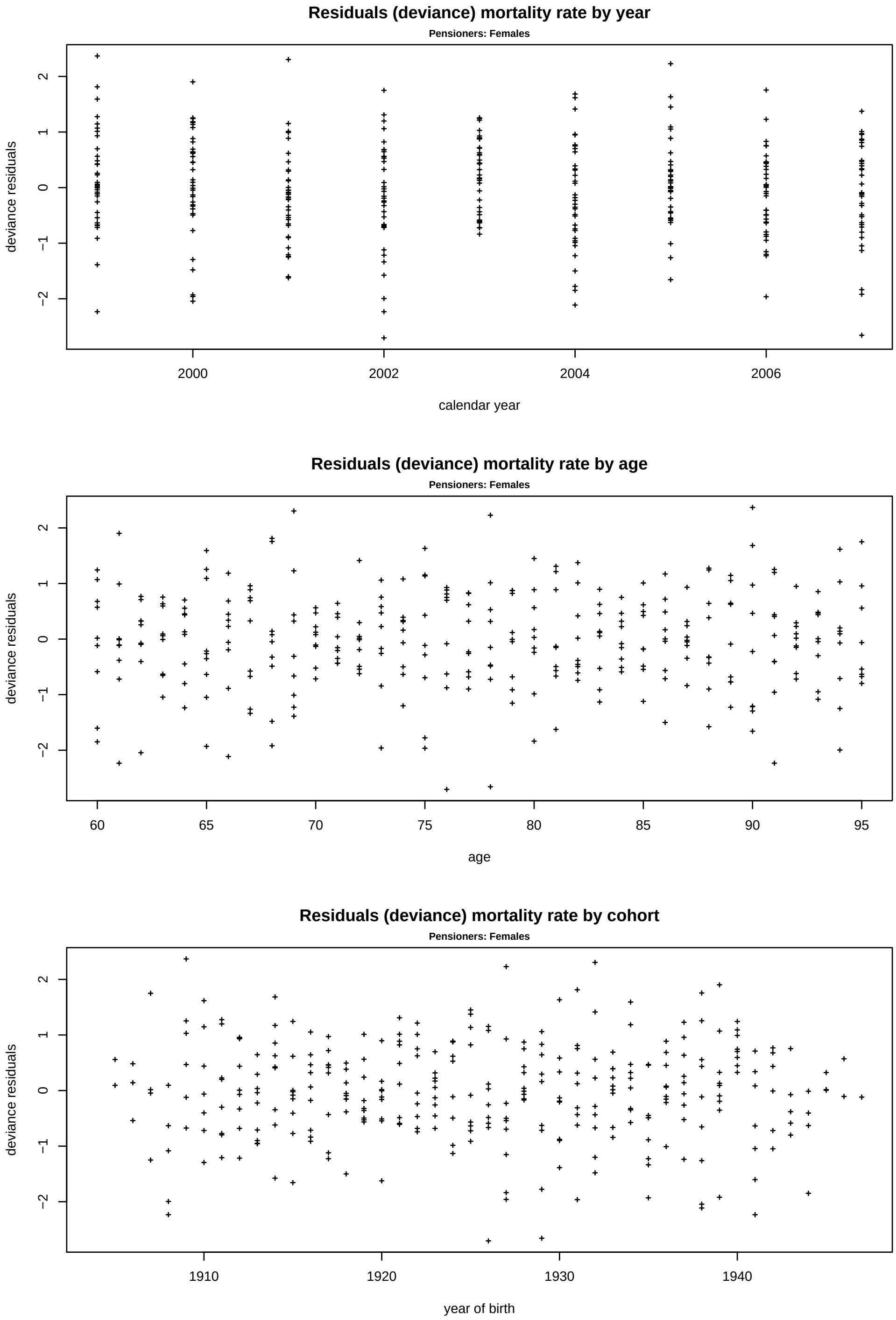


Figure 5.22: Deviance residuals of LC regression for private scheme female pensioners for age range $60 - 95$ over the observation period of $1999 - 2007$.

ML estimation methods to fit the model parameters. The overall assessment of the model is that it yields a reasonable fit, although the robustness of the parameters was a key issue. That is, the limited span of the observation period and the high variability of the mortality rates at lower ages resulted unstable parameter estimates. Nevertheless, we have found that, in general, the parameter estimates exhibit similar features between the mortality data sets investigated. On the other hand, the evaluation of the model variants showed that the cohort effects are consistently of minor significance within the insurance–specific data sets.

In terms of forecasting, we have made use of standard univariate time series methods to extrapolate the period parameter of the LC family and to predict future mortality. Life expectancy at age 65 based on the predicted rates provides a basis of comparison between the future mortality of the private scheme pensioners and the other data sets. Thus, the observed male pensioners seem to have somewhat higher life expectancy than the general E&W male population, but smaller than those based on the CMI pensioners data. However, given that the analysis is limited to the available short term data, it is clear that longer term predictions are not appropriate. This makes it difficult to make conclusive predictions of future improvements, although it seems quite reasonable to assume that the observed trends will continue at least in the nearest future.

An advantage of the ML estimation methodology is that it can be readily adapted to allow for additional effects within the predictor structure. This feature was exploited to estimate the impact of the additional explanatory factors contained in the private pension scheme data set. We have found that, among the potential additional effects shaping the pensioners mortality experience, the socio–economic group variable (a proxy for personal wealth) provides the most satisfactory adjustment of the base LC model.

Chapter 6

Concluding Remarks

In this thesis, we have demonstrated various modelling techniques and their practical implementation in two key areas of actuarial science. In particular, we have provided examples related to the estimation of future loss of earnings in the context of liability insurance and to the modelling and forecasting of mortality rates of populations of life insurance and pensions. All the empirical analyses presented here were carried out using real UK based data.

Thus, Part I is concerned with the modelling of labour force dynamics from the perspective of the loss of earnings multipliers in England and Wales, whereas Part II is related to the theoretical and practical aspects of modelling and forecasting of mortality rates. Since each of the chapters that make up the thesis is a self–contained paper, providing its own set of conclusions and discussions, we have avoided revisiting those in this chapter. Instead, we will direct attention to the impact of the work on both academic and other audiences. We also describe the ways in which the work has facilitated further research carried out by the author, and by others.

6.1 Impact and Further Developments of the Research

6.1.1 Personal injury compensation in the UK and the Ogden Tables

Following the publication of the papers presented in Chapters 2 and 3, there has been a great deal of interest both in the methodology and in the outcome of the research. The research has lead to improvements in the calculation method of damages due to personal injury that were incorporated in the 6$^{\text{th}}$ edition Ogden Tables (Actuarial Tables 2007). Thus, the impact of this research was very significant and consequent developments were generally more substantial than that of the research presented in Chapters 4 and 5. Therefore, in recent years the author of this thesis has tended to direct more of his attention to the research contained in Part I, and in the following we will examine the main aspects of this work.

As mentioned in section 1.1, the papers presented in the first part of this thesis were the nominated output reports of an ESRC grant and were published on the ESRC Society Today website. The execution of the research and the reports have received from the independent peer reviewers an overall classification of outstanding. The work has generated a considerable interest from legal practitioners and statisticians alike, and has given rise to the following co-authored publications:

1. Butt, Haberman, Verrall and Wass (2008)
2. de Wilde R., Wass, Verrall, Haberman and Butt (2008)
3. Butt, Haberman, Verrall and Wass (2009)
4. Butt, Haberman, Verrall and Wass (2010)

Butt et al. (2008) provided an inter–disciplinary approach between Actuarial Science and Economics in order to estimate age–specific worklife expectancies (and corresponding reduction factors) for the calculation of future loss of earnings. In this paper, the three–state Markov chain model used in Chapter 3 was replaced by a two–state alternative, whereas the 'unemployed' and 'out of labour force' states were merged into a single 'non–employed' state. This simplification allowed the results to become more transparent and better suited for practical use in courts.

Further, in order to validate the results, the above actuarial methodology was compared against an empirical econometric modelling approach using cross–sectional LFS data. While the results were broadly similar for the two models, there were significant differences due to measurement bias and reduced precision in the case of the econometric approach. Therefore, the actuarial method was adopted as the more reliable alternative and the final results were presented corresponding to the multiple state modelling approach. In addition, the paper has also reviewed briefly the US forensic economics literature, which was found to be more prolific in terms of dynamic modelling and measurement of future loss of earnings than the UK literature.

Similarly as in the original analysis, baseline WLEs were estimated as a function of age and sex and the analysis was extended to allow for stratification on a number of additional variables that are the most relevant in the context of personal injury compensations. Thus, the final results were presented in the form of a set of age–specific WLEs (and corresponding RFs) disaggregated by sex, initial employment status, disability and level of education. The calculated values were also used in a case study illustration of evaluating the personal injury compensation for future loss of earnings for a plaintiff with post–injury earnings capacity. This practical illustration was presented in comparison to the outcomes resulting when using the 5$^{\text{th}}$ edition Ogden Tables multipliers and the previous methodology. Strikingly, it was found that the damages award could be almost 43 % higher when using the revised method of calculation. The main reason for this discrepancy was due to the fact that the previous methodology did not provide adequate compensations for the disadvantages faced in the labour market by disabled workers with residual earnings potential.

The paper was presented in an *Ordinary Meeting* of the *Royal Statistical Society* (RSS) and appeared in the Journal of RSS Series A together with the discussions and the authors' replies. All the commentators found the methodology and the results described in the paper very significant and potentially important in the context of damages for personal injury. The paper has clearly demonstrated that educational attainment and disability are the most significant explanatory variables for valuing future loss of earnings and the revised calculation approach has lead to a fairer system of personal injury compensations. Consequently, the WLE estimates published in this paper were used in a simplified and restructured format by GAD in the followup editions of the Ogden Tables (i.e. 6$^{\text{th}}$ and 7$^{\text{th}}$: Actuarial Tables 2007, 2011).

In de Wilde R. et al. (2008), the merits of the approach in Butt et al. (2008) and the resulting 6$^{\text{th}}$ edition Ogden Tables were clarified in response to a critical paper that appeared in the *Association of Personal Injury Lawyers Personal Injury Focus* (APIL PI Focus) after the publication of the above JRSS paper (see Barling 2008). The author of that critique has raised a number of issues about the use of the latest set of RFs and the updated methodology, in particular related to the retirement age, the use of the highest educational attainment for younger age groups and the classification of disability that resulted from the LFS data. The first two of these issues were mainly technical and had reduced implications as they could easily be taken into account in practice by the Courts. While the issues regarding the disability variable had some valid foundations, they could be resolved by a more precise description of what constituted as disability in the new set of tables. Furthermore, de Wilde R. et al. provided supporting evidence that the use of the LFS disability variable was the best available measure and matched closely the results given by a benchmark disability study (see Burchardt 2000). The authors have argued that while the new methodology and data have some unavoidable deficiencies, they still represent a significant improvement in estimating life–time labour market risks compared to the previous approach.

The aim of Butt et al. (2009) and Butt et al. (2010) was to disseminate the results to a wider specialist audience from both the US and the UK. Both of these papers have been written by invitation from the respective publishers. In particular, the US forensic economics audience appeared to show interest in the discounted WLE type calculations for the purposes of estimating future loss of earnings. Typically, the US–type approach would not express the life–time loss of employment as a discounted value and could potentially overestimate the future loss of earnings. Thus, in the US framework the age–specific WLE is calculated using a probabilistic Markov chain model, based on Alter and Becker (1985), that focuses on estimating the age–specific transition probabilities between the employment states. The above papers presented a simplified version of the UK methodology based on transition intensities and emphasised the importance of using additional explanatory factors in the estimation of future loss of earnings multipliers.

Therefore, it can be seen that this series of research has generated a considerable interest in the specialist literature, within both the UK and the US. The methodological framework and the results provided a simple and robust compensation system that had a wide ranging implications in many aspects of the UK tort system and beyond. These include:

- The 6^{th} edition of the Ogden Tables were constructed based on our suggestions of estimating future loss of earnings multipliers (Actuarial Tables 2007). The new set of tables and the improved methodology are in use since 2008 in most court cases in E&W involving personal injury litigation.
- The use of the tables by the legal profession demonstrate that the new methodology and data have succeeded in providing greater accuracy in the estimation of pre– and post–injury employment risks.
- The public has benefited from an improved assessment of court awards by a more reliable and fairer system of capturing the true dynamic nature of the labour market (based on pre– and post–injury earnings potential).
- The liability insurance industry has benefited by a fairer and more accurate set of Ogden Tables multipliers.
- Researcher communities have found new grounds to initiate further studies to assess the bias in labour market data and to evaluate better measurement systems.
- The research has demonstrated suitable econometric methodology for the measurement of life–time employment risk factors (other than age and starting employment state).
- While some authors are reporting that the gap between the life expectancy of general population and those seriously disabled (by injuries) is reducing, this research is providing tangible evidence that the length of time spent in employment is often not the same. This poses potential implications for policymakers in terms of making adequate social and health provisions in order to improve employment prospects of the disabled populations with earnings potential.
- Therefore, the research has made an important contribution in three main

areas: the calculation of future losses in personal injury litigation, actuarial science and labour economics.

In addition to the above mentioned impacts, the study provides a scope to answer some interesting questions that emerge and to carry out further research in many related areas. These might include:

- Evaluate the impact of potential bias in the measurement of disability in the current UK LFS data based on existing studies.
- An interdisciplinary review of other factors that could affect current employment risks predictions (like medical improvements, standard of care or changes in economic conditions) could be carried out. Similarly, with some development of the results in relation to the disability variable, sociologists would be interested in the impact of disability on life–time employment.
- Duration dependence of the transition intensities and their relation to other covariates (like prospective wage, sickness/unemployment benefit, marital status, number of dependent children, micro–economic indicators, etc.) should be considered in order to smooth out crude estimates. The estimated transition probabilities could be compared to the results from studies of unemployment spells in Britain (e.g. Arulampalam and Stewart 1995 or Narendranathan and Stewart 1993*b*,*a*).
- Since the non–employment categories are currently defined in an unconventional way, it is difficult to carry out direct comparative work with that of labour economists. Therefore, it would be informative to apply the model to provide estimates based upon conventional ILO definitions of employment, unemployment and inactivity.
- Develop the analysis in relation to the disability covariate to include within the multiple state model the risk of becoming disabled. This would provide

superior estimates of pre–injury employment risks as it would also include the risk of becoming disabled and the associated increased risk to future employment.

- There is considerable interest in estimating the distribution of the working age population across areas of economic activities both from a social and from an economic standpoint. For example, evaluating the consequences of the gender and race differentials in labour market participation are key issues for sociologists. Whereas forecasting labour market dynamics is the primary tool in economic planning and policy making.

- In general, the insurance industry have a vested interest in examining the relationship between the impairment and the employment risks, and consequently in the development of corresponding life–time risk measures.

6.1.2 Mortality models for heterogeneous insurance based populations

A detailed account of the work presented in Chapter 4 has also appeared as an actuarial research paper within City University London in 2002, which was followed by a shorter version published in an ASTIN Bulletin in 2004 (see Butt and Haberman 2002 and Butt and Haberman 2004). The work has attracted some attention in the specialist literature and it was referred to in papers appearing in various journals worldwide (see Debon et al. 2005, Delwarde et al. 2006, Olivieri 2006, Hosseini 2008, Ramsay and Arcila 2013 and Willemse and Kaas 2007).

Since the formulation of the multiplicative frailty model, there has been a proliferation of applications in various fields other than mortality. This concept can be readily extended to areas like contracting diseases or component deterioration of complex systems, but also to topics where *frail* is applied in a much wider context, like fertility, migration or leaving unemployment (see for example Vaupel and Yashin 1985*b*). In addition, it is worth noting that frailty models have been applied extensively in actuarial studies related to mortality and life insurance, yet these were somehow avoided in general insurance mathematics and insurance economics.

6.1.3 Modelling and forecasting mortality within a Lee–Carter type framework

The work contained within Chapter 5 has also been published as an actuarial research paper within City University London in 2009 (see Butt and Haberman 2009). The R programming package ilc that was developed during this work has been used in many mortality studies carried out by the author and also by others. Among other investigations, the package was applied by the author in the mortality analysis of England and Wales and also of various groups of pensioners, which is presented partly in Chapter 5. In addition, the author of this thesis has also supervised many final year projects of undergraduate actuarial students, whom were interested to learn more about the Lee–Carter type modelling. Thus, the students made use of the functionalities of this package in order to carry out more easily the graduations of the Lee–Carter type models. In this way, they were able to focus more on the modelling structure and to achieve a better understanding of the results. It appears that the package has generated some interest among some external academics and actuarial practitioners too, but due to time constraints it was not possible to fully follow up this information.

Appropriate mortality modelling and forecasting is an all important tool for governments and for insurance companies alike in order to administer and to plan the financial development of pension schemes, and also to make adequate health and social policy provisions. Recently, it has been demonstrated that the Lee–Carter type stochastic forecasting method has successfully aided the forecasting of mortality rates of national and insurance populations all over the world (for instance, see Renshaw and Haberman 2003*a*,*b*, Lundström and Qvist 2004, Haberman and Renshaw 2007, Denuit et al. 2011).

6.2 Overall Summary

This thesis covers topics concerning two important actuarial areas that are organised in two respective parts. In the first part of the thesis, we illustrate the estimation and forecasting of the transition intensities of multiple state models in the context of calculating personal injury damages for working age adults (liability insurance). Then in the second part of the thesis, we consider the modelling and measuring the mortality rates of insurance based populations (annuitants, life insurance and pensions).

In the case of liability insurance, the main contribution of the thesis is to provide a systematic empirical analysis of the factors that affect future labour market risks by making use of longitudinal UK LFS data. The benefit of the demonstrated dynamic modelling method is twofold. On one hand, it builds on a sound scientific approach with consideration given to particular circumstances (e.g. starting economic state, disability status, socio-economic characteristics, etc.), rather than to average population values. On the other hand, it allows the investigation of the true labour force dynamics by making use of longitudinal observation data, which potentially increases the accuracy of the estimation. Also, it provides a single measure of life–time labour market risks of non–participation in the labour market that is well suited for the use of calculating damages for future loss of earnings.

In the case of mortality analysis, this dissertation can be viewed as an exploration of both theoretical and practical aspects of mortality modelling. On one hand, it attempts to answer theoretical questions related to the appropriateness of commonly applied homogeneity assumption when modelling the mortality rates of insurance based populations. In this respect, the results indicate that, subject to issues of identifiability, there is evidence of low heterogeneity in the insurance based populations, suggesting that a pre-selection effect might play some role. On the other hand, it demonstrates practical ways in which generalised linear modelling techniques can be used to graduate various types of mortality models based on (overdispersed) Poisson likelihood assumption. In addition, we have developed the necessary programming tools for the analysis of logistic mortality models that can factor in cohort and also non–age related effects. The methodologies developed allow both the graduation and forecasting of mortality rates.

REFERENCES

Aalen, O. O. (1988), "Heterogeneity in survival analysis", *Statistics in Medicine* **7**, 1121–1137.

Aalen, O. O. (1992), "Modelling heterogeneity in survival analysis by the Compound Poisson distribution", *The Annals of the Applied Probability* **2**(4), 951–972.

Actuarial Tables (2004), With explanatory notes for use in Personal Injury and Fatal Accident Cases, Technical Report Fifth Edition, Government Actuary's Department, London, UK. TSO.

Actuarial Tables (2007), With explanatory notes for use in Personal Injury and Fatal Accident Cases, Technical Report Sixth Edition, Government Actuary's Department, London, UK. TSO.

Actuarial Tables (2011), With explanatory notes for use in Personal Injury and Fatal Accident Cases, Technical Report Seventh Edition, Government Actuary's Department, London, UK. TSO. (Updated in 2013).

Ainslie, R. (2000), "Annuity and insurance products for impaired lives". Presented to the *Staple Inn Actuarial Society, Institute of Actuaries*, London.

Alter, G. and Becker, W. (1985), "Estimating lost future earnings using the new worklife tables", *Monthly Labour Review* **108**(2), 39–42.

Andersen, P. K., Klein, J. P. and Rosthøj, S. (2003), "Generalised linear models for correlated pseudo-observations, with applications to multi-state models", *Biometrika* **90**(1), 15–27.

Artola, C. and Bell, U.-L. (2001), Identifying labour market dynamics using Labour Force Survey data, Discussion Paper No. 01–44, Centre for European Economic Research, Mannheim.

Arulampalam, W., Booth, A. and Taylor, M. (2000), "Unemployment persistence", *Oxford Economic Papers* **52**, 24–50.

Arulampalam, W. and Stewart, M. B. (1995), "The determinants of individual unemployment durations in an era of high unemployment", *The Economic Journal* **105**(March), 321–332.

Barling, L. (2008), "Applying the sixth edition of the Ogden Tables A–D: a view from the sharp end", *Association of Personal Injury Lawyers PI Focus* **18**(1), 23–27.

Barmby, T. A., Ercolani, M. G. and Treble, J. G. (1999), "Using the Labour Force Surveys to construct an absence rate series for Britain, 1984-1997 (revised)", Institute for Labour Research. University of Essex.

Bell, E. B. and Taub, A. J. (1998), "Expected working life, transition probabilities and the size of an award", *Journal of Forensic Economics* **11**(2), 91–102.

Benjamin, B. and Pollard, J. H. (1980), *The analysis of mortality and other actuarial statistics*, Second Edition, Heinemann, London.

Benjamin, B. and Pollard, J. H. (1993), *The analysis of mortality and other actuarial statistics*, The Institute of Actuaries and the Faculty of Actuaries, London.

Blumen, I., Kogan, M. and McCarthy, P. J. (1955), The industrial mobility of labour as a probability process, *in* "Cornell Studies of Industrial and Labour Relations", Vol. vi, Cornell University Press, Ithaca, New York.

Blundell, R. and Johnson, P. (1997), Pensions and retirement in the UK, Working Paper No. 6154, National Bureau of Economic Research, Cambridge MA.

Blundell, R. and Johnson, P. (1998), "Pensions and labor-market participation in the United Kingdom", *The American Economic Review* **88**(2), 168–172. Papers and Proceedings of Hundred and Tenth Annual Meeting of the American Economic Association.

Böheim, R. and Taylor, M. P. (2000), Unemployment duration and exit states in Britain. Institute for Social and Economic Research. University of Essex.

Booth, H. (2006), "Demographic forecasting: 1980 to 2005 in review", *International Journal of Forecasting* **22**(3), 547–581.

Booth, H., Hyndman, R. J., Tickle, L. and de Jong, P. (2006), "Lee–Carter mortality forecasting: A multi–country comparison of variants and extensions", *Demographic Research* **15**(Article 9), 289–310.

Bradley, S., Crouchley, R. and Oskrochi, R. (2001), Social exclusion and labour market transitions: A multi-state multi-spell analysis using the BHPS, Working Paper No. 2001/03, Department of Economics, Lancaster University.

Bradley, S., Crouchley, R. and Oskrochi, R. (2003), "Social exclusion and labour market transitions: A multi-state multi-spell analysis using the BHPS", *Labour Economics* **10**(6), 659–679. Elsevier Science.

Brouhns, N., Denuit, M. and Vermunt, J. (2002), "A Poisson log–bilinear regression approach to the construction of projected life tables", *Insurance: Mathematics and Economics* **31**, 373–393.

Burchardt, T. (2000), "Enduring economic exclusion: Disabled people, income and work", Joseph Rowntree Foundation. York, UK.

Butt, Z. and Haberman, S. (2002), Application of frailty-based mortality models to insurance data, Actuarial Research Paper No. 142, Department of Actuarial Science and Statistics, City University London.

Butt, Z. and Haberman, S. (2004), "Application of frailty–based mortality models using generalized linear models", *ASTIN Bulletin* **34**(1), 175–197.

Butt, Z. and Haberman, S. (2009), ilc: A collection of R functions for fitting a class of Lee–Carter mortality models using iterative fitting algorithms, Actuarial Research Paper No. 190, Faculty of Actuarial Science and Insurance, City University London.

Butt, Z., Haberman, S. and Verrall, R. (2006), The impact of dynamic multi–state measurement of worklife expectancy on the loss of earnings multipliers in England and Wales, Working Paper No. 2, Cass Business School, City University London. Published on ESRC Society Today website.

Butt, Z., Haberman, S., Verrall, R. and Wass, V. (2008), "Calculating compensation for loss of future earnings: Estimating and using work life expectancy", *Journal of the Royal Statistical Society, Series A* **171**(4), 763–805.

Butt, Z., Haberman, S., Verrall, R. and Wass, V. (2009), *Personal Injury and Wrongful Death Damages Calculations: A Trans–Atlantic Dialogue*, Contemporary Studies in Economic and Financial Analysis, Emerald Press, USA, chapter 8 (*Estimating and Using Work Life Expectancy in the UK*), pp. 103–134.

Butt, Z., Haberman, S., Verrall, R. and Wass, V. (2010), "Worklife expectancy: calculating compensation for loss of future earnings", *Measurement and Control (by invitation)* **43**(5), 146–151.

Cappellari, L. and Jenkins, S. P. (2003), "Transitions between unemployment and low pay", Institute for Social and Economic Research. University of Essex.

Carey, J. R., Liedo, P., Orozco, D. and Vaupel, J. W. (1992), "Slowing of mortality rates at older ages in large medfly cohorts", *Science* **258**, 457–463.

Chib, S. and Greenberg, E. (1998), "Analysis of multivariate probit models", *Biometrika* **85**, 347–361.

Ciecka, J. E., Donley, T. and Goldman, J. (1995), "A Markov process model of worklife expectancies based on Labor Market Activity in 1992–93", *Journal of Legal Economics* **5**(3), 17–41.

Ciecka, J. E., Donley, T. and Goldman, J. (1997), "Regarding median years to retirement and worklife expectancy", *Journal of Forensic Economics* **10**(3), 297–310.

Ciecka, J. E., Donley, T. and Goldman, J. (2000), "A Markov process model of work-life expectancies for ages 66–75 based on Labor Market Activity in 1997–98", *Journal of Legal Economics* **10**(2), 27–36.

CMIR 21 (2004), Continuous mortality investigation reports, Technical Report 21, The Faculty of Actuaries and Institute of Actuaries, UK.

Congdon, P. (1994), "Analyzing mortality in London: Life Tables with frailty", *Journal of the Royal Statistical Society, Series D* **43**(2), 277–308.

Consultation Paper (2000), "Damages: The discount rate and alternatives to lump sum payments", Department for Constitutional Affairs. London, UK.

Consultation Paper (2002), "Damages for future loss: Giving the courts the power to order periodical payments for future loss and care costs in personal injury cases", Department for Constitutional Affairs. London, UK.

Cramer, H. and Wold, H. (1935), "Mortality variations in Sweden: A study in graduation and forecasting", *Skandinavisk Aktuarieidskrift* **18**, 161–241.

Crouchley, R. and Oskrochi, R. (2000), Testing for sample selection bias in the labour market behaviour of respondents from the British Household Panel Survey, Working Paper No. 2000/02, Centre for Applied Statistics. Lancaster University, Lancaster, UK.

de Wilde R., Wass, V., Verrall, R., Haberman, S. and Butt, Z. (2008), "Applying the sixth edition of the Ogden Tables: A response from the Ogden Working Party and the tables' authors", *Association of Personal Injury Lawyers PI Focus* **18**(3), 14–17.

Debon, A., Montes, F. and Sala, R. (2005), "A comparison of parametric models for mortality graduation. application to mortality data for the valencia region (spain)", *Statistics and Operations Research Transactions* **29**(2), 269–288.

Delwarde, A., Denuit, M., Guillén, M. and Vidiella-i-Anguera, A. (2006), "Application of the Poisson log-bilinear projection model to the G5 mortality experience", *Belgian Actuarial Bulletin* **6**, 54–68.

Denuit, M., Haberman, S. and Renshaw, A. (2011), "Longevity–indexed annuities", *North American Actuarial Journal* **15**(1), 97–111.

Drapeau, M. D., Gass, E. K., Simison, M. D., Mueller, L. D. and Rose, M. R. (2000), "Testing the heterogeneity theory of late–life mortality plateaus by using cohorts of drosophila melanogaster", *Experimental Gerontology* **35**, 71–84.

Ercolani, M. G. (2000), "A simple empirical model of sickness absence", Institute for Social and Economic Research. University of Essex.

Forfar, D. O., McCutcheon, J. J. and Wilkie, A. D. (1988), "On graduation by mathematical formula", *Journal of the Institute of Actuaries* **115**, 1–149.

Forfar, D. O. and Smith, D. M. (1988), "The changing shape of English Life Tables", *Transactions of the Faculty of Actuaries* **40**, 98–134.

Gini, C. (1924), Premières recherches sur la fécondabilité de la femme, *in* "Proceedings of the International Mathematical Congress", Vol. 2, pp. 889–892.

Grime QC, S. (2003), Multiplier issues briefing note, Technical report, Deans Court Chambers.

Gutterman, S. and Vanderhoof, I. (1998), "Forecasting changes in mortality", *North American Actuarial Journal* **2**(4), 135–138.

Haberman, S. and Bloomfield, D. (1990), "Work time lost to sickness, unemployment and stoppages: Measurement and application", *Journal of the Institute of Actuaries* **117**, 533–595.

Haberman, S. and Pitacco, E. (1999), *Actuarial Models for Disability Insurance*, Chapman and Hall/CRC Press, Boca Raton.

Haberman, S. and Renshaw, A. E. (1996), "Generalised linear models and actuarial science", *Journal of the Royal Statistical Society, Series D* **45**, 407–436.

Haberman, S. and Renshaw, A. E. (2007), "Discussion of *'Pension plan valuation and mortality projection: A case study with mortality data'* ", *North American Actuarial Journal* **11**(4), 148–150.

Haberman, S. and Renshaw, A. E. (2009), "On age–period–cohort parametric mortality rate projections", *Insurance: Mathematics and Economics* **45**(2), 255–270.

Harvey, A. C. and Durbin, J. (1986), "The effects of seat belt legislation on British road casualties: A case study in structural time series modelling", *Journal of the Royal Statistical Society, Series A* **149**, 187–227.

Hoem, J. M. (1977), "A Markov chain model of working life tables", *Scandinavian Actuarial Journal* pp. 1–20.

Hoem, J. M. (1990), "Identifiability in hazard model with unobserved heterogeneity: The compatibility of two apparently contradicting results", *Theoretical Population Biology* **37**, 124–128.

Horiuchi, S. and Coale, A. J. (1990), "Age patterns of mortality for older women: An analysis using the age–specific rate of mortality change with age", *Mathematical Population Studies* **2**(4), 245–267.

Horiuchi, S. and Wilmoth, J. R. (1998), "Deceleration in the age pattern of mortality at older ages", *Demography* **35**(4), 391–412.

Hosseini, R. (2008), Adverse selection in the annuity market and the role for Social Security, 2008 Meeting Papers 264, Society for Economic Dynamics. http://EconPapers.repec.org/RePEc:red:sed008:264.

Hougaard, P. (1984), "Life Table methods for heterogeneous populations: Distributions describing heterogeneity", *Biometrika* **71**, 75–83.

Hougaard, P. (1986), "Survival models for heterogeneous populations derived from stable distributions", *Biometrika* **73**, 387–396.

Jackson, C., Sharples, L., Thompson, S., Duffy, S. and Couto, E. (2003), "Multistate Markov models for disease progression with classification error", *Journal of the Royal Statistical Society, Series D* **52**(2), 193–209.

Jones, B. L. (1993), Modelling multi–state processes using a Markov assumption, *in* ARCH, ed., "Proceedings of the 27$^{\text{th}}$ Annual Research Conference, Iowa City, 1992", pp. 239–248.

Jones, B. L. (1998), "A model for analysing the impact of selective lapsation on mortality", *North American Actuarial Journal* **2**(1), 79–86.

Judicial Studies Board (2002), "Guidelines for the assessment of general damages in personal injury cases", Oxford University Press. 6$^{\text{th}}$ Edition.

Judicial Studies Board (2004*a*), "Civil Bench Book: Section 17. Damages", Judicial Studies Board, London, UK.

Judicial Studies Board (2004*b*), "Guidelines for the assessment of general damages in personal injury cases", Oxford University Press. 7th Edition.

Kannisto, V., Lauritsen, J., Thatcher, A. and Vaupel, J. (1994), "Reductions in mortality at advanced ages: Several decades of evidence from 27 countries", *Population and Development Review* **20**(4), 793–810.

Keyfitz, N. (1982), "Choice of functions for mortality analysis: Effective forecasting depends on a minimum paramater representation", *Theoretical Population Biology* **21**, 329–352.

Keyfitz, N. (1985), *Applied mathematical demography*, Second Edition, Springer, New York.

Lancaster, T. (1990), *The econometric analysis of transition data*, Cambridge University Press, Cambridge.

Le Bras, H. (1976), "Lois de mortalité et age limité", *Population* **31**, 655–692.

Lee, L. (2000), "The Lee–Carter method for forecasting mortality, with various extensions and applications", *North American Actuarial Journal* **4**, 80–93.

Lee, R. and Carter, L. (1992), "Modelling and forecasting U.S. mortality", *Journal of the American Statistical Association* **87**, 659–671.

Lee, R. and Nault, F. (1993), *Modeling and forecasting provincial mortality in Canada.* Paper presented at the World Congress of International Union for the Scientific Study of Population, Montreal. Unpublished.

Lewis, R., McNabb, R., Robinson, H. and Wass, V. (2002), "Court awards of damages for loss of future earnings: An empirical study and an alternative method of calculation", *Journal of Law and Society* **29**(3), 406–435.

Lewis, R., McNabb, R., Robinson, H. and Wass, V. (2003), "Loss of earnings following personal injury: Do the courts adequately compensate injured parties?", *The Economic Journal* **113**(November), 568–584.

Li, N. and Lee, R. D. (2005), "Coherent mortality forecasts for a group of populations: an extension of the Lee–Carter method", *Demography* **42**(3), 575–594.

Lundström, H. and Qvist, J. (2004), "Mortality forecasting and trend shifts: an application of the Lee–Carter model to Swedish mortality data", *International Statistical Review* **72**(1), 37–50.

Manton, K. G. and Stallard, E. (1984), "Recent trends in mortality analysis", *Academic press, Inc.* pp. 236–299.

Manton, K. G., Stallard, E. and Vaupel, J. W. (1981), "Methods for comparing the mortality experience of heterogeneous populations", *Demography* **18**, 389–410.

Manton, K. G., Stallard, E. and Vaupel, J. W. (1986), "Alternative models for the heterogeneity of mortality risks among the aged", *Journal of the American Statistical Association* **81**, 635–644.

McCullagh, P. and Nelder, J. (1989), *Generalized linear models*, Second Edition, Chapman and Hall, London.

McNown, R. and Rogers, A. (1989), Time–series analysis forecasts of a parameterised mortality schedule, *in* P. Congdon and P. Batey, eds, "Advances in regional demography: Information, forcasts, models", Bellhaven Press, London, pp. 107–123.

Millimet, D. L., Nieswiadomy, M., Ryu, H. and Slottje, D. (2003), "Estimating worklife expectancy: An econometric approach", *Journal of Econometrics* **113**, 83–113.

Narendranathan, W. and Stewart, M. B. (1993*a*), "How does the benefit effect vary as unemployment spells lengthen?", *Journal of Applied Econometrics* **8**, 361–381.

Narendranathan, W. and Stewart, M. B. (1993*b*), "Modelling the probability of leaving unemployment: Competing risks models with flexible baseline hazards", *Journal of the Royal Statistical Society, Series C* **42**(1), 63–83. Applied Statistics.

Olivieri, A. (2006), "Heterogeneity in survival models. applications to pensions and life annuities", *Belgian Actuarial Bulletin* **6**, 23–39.

Paull, G. (2002), Biases in the reporting of labour market dynamics, Working Paper No. 02/10, The Institute for Fiscal Studies, London, UK.

Pletcher, S. D. and Curtsinger, J. W. (1998), "Mortality plateaus and the evolution of senescence: Why are old-age mortality rates so low?", *Evolution* **52**(2), 454–464.

Pollard, J. H. (1970), "Random mortalilty fluctuations and the binomial hypothesis", *Journal of the Institute of Actuaries* **96**, 251–264.

Ramsay, C. M. and Arcila, L. D. (2013), "Pricing funeral (burial) insurance in a microinsurance world with emphasis on Africa", *North American Actuarial Journal* **17**(1), 63–81.

Renshaw, A. E. (1991), "Actuarial graduation practice and generalised linear and non-linear models", *Journal of the Institute of Actuaries* **118**, 295–312.

Renshaw, A. E. (1995), Graduation and generalized linear models: An overview, Actuarial Research Paper No. 73, Department of Actuarial Science and Statistics, City University London.

Renshaw, A. E. and Haberman, S. (2003*a*), "Lee–Carter mortality forecasting: A parallel generalized linear modelling approach for England and Wales mortality projections", *Journal of the Royal Statistical Society, Series C* **52**(1), 119–137.

Renshaw, A. E. and Haberman, S. (2003*b*), "Lee–Carter mortality forecasting with age specific enhancement", *Insurance: Mathematics and Economics* **33**, 255–272.

Renshaw, A. E. and Haberman, S. (2006), "A cohort-based extension to the Lee–Carter model for mortality reduction factors", *Insurance: Mathematics and Economics* **38**, 556–570.

Renshaw, A. E. and Haberman, S. (2008), "On simulation–based approaches to risk measurement in mortality with specific reference to Poisson Lee–Carter modelling", *Insurance: Mathematics and Economics* **42**(2), 797–816.

Richards, H. (2000), "Worklife expectancies: Increment-decrement less accurate than conventional", *Journal of Forensic Economics* **13**(2), 271–289.

Sithole, Z. T., Haberman, S. and Verrall, R. J. (2000), "An investigation into parametric models for mortality projections, with applications to immediate annuitants' and life office pensioners' data", *Insurance: Mathematics and Economics* **27**(3), 285–312.

Skoog, G. and Ciecka, J. (2001*a*), "The Markov (increment-decrement) model of labor force activity: Extended tables of central tendency, variation, and probability intervals", *Journal of Legal Economics* **11**(1), 23–87.

Skoog, G. and Ciecka, J. (2001*b*), "The Markov (increment-decrement) model of labor force activity: New results beyond work-life expectancies", *Journal of Legal Economics* **11**(1), 1–21.

Skoog, G. and Ciecka, J. (2002), "Probability mass functions for additional years of labor market activity induced by the Markov (increment-decrement) model", *Economics Letters* **77**, 425–431.

Smith, S. J. (1982), "New worklife estimates reflect changing profile of labor force", *Monthly Labor Review* **105**(3), 15–20.

Smith, S. J. (1986), "Worklife estimates: effects of race and education", *Bureau of Labor Statistics Bulletin* **2254**. US Department of Labor.

Stewart, M. and Swaffield, J. (1999), "Low pay dynamics and transition probabilities", *Economica* **66**, 23–42.

Tabeau, E. (2001), A review of demographic forecasting models for mortality, *in* E. Tabeau, A. Van Den Berg Jeths and C. Heathcote, eds, "Forecasting mortality in developed countries: insights from a statistical, demographic and epidemiological perspective", Dordrecht: Kluwer Academic Publishers, pp. 1–32.

Thatcher, A. R. (1999), "The long–term pattern of adult mortality and the highest attained age", *Journal of the Royal Statistical Society, Series A* **162**, 5–30.

Thatcher, A. R., Kannisto, V. and Vaupel, J. W. (1998), The force of mortality at ages 80 to 120, *in* "Odense Monographs on Population Aging", Vol. 5, Odense University Press, Centre for Health and Social Policy.

Tobin, J. (1958), "Estimation of relationships for limited dependent variables", *Econometrica* **26**, 24–36.

Vaupel, J. W. (1988), “Inherited frailty and longevity”, *Demography* **25**, 277–287.

Vaupel, J. W. (1999), “Contribution to the discussion of *‘The long–term pattern of adult mortality and highest attained age’*, by A. R. Thatcher”, *Journal of the Royal Statistical Society, Series A* **162**, 31–32.

Vaupel, J. W., Manton, K. G. and Stallard, E. (1979), “The impact of heterogeneity in individual frailty on the dynamics of mortality”, *Demography* **16**(3), 439–454.

Vaupel, J. W. and Yashin, A. I. (1985*a*), “The deviant dynamics of death in heterogeneous populations”, *Sociological Methodology* pp. 179–211. Edited by N B Tuma, Jossey–Bas, San Francisco.

Vaupel, J. W. and Yashin, A. I. (1985*b*), “Heterogeneity’s ruses: Some surprising effects of selection on population dynamics”, *The American Statistician* **39**(3), 176–185.

Vaupel, J. W., Yashin, A. I. and Manton, K. G. (1988), “Debilitation’s aftermath: Stochastic process models of mortality”, *Mathematical Population Studies* **1**, 21–48.

Venables, W. N., Smith, D. M. and the R Development Core Team (2005), “An introduction to R. Notes on R: A programming environment for data analysis and graphics”. Version 2.1.1 (2005-06-20). ISBN 3–900051–12–7.

Verrall, R., Haberman, S. and Butt, Z. (2005), An investigative study on current practice of estimating the loss of earnings in personal injury claims in England and Wales: The Ogden Tables and contingencies other than mortality, Working Paper No. 1, Cass Business School, City University London. Published on ESRC Society Today website.

Wang, S. S. and Brown, R. L. (1998), “A frailty model for projection of human mortality improvement”, *Journal of Actuarial Practice* **6**, 221–241.

Willemse, W. J. and Kaas, R. (2007), “Rational reconstruction of frailty-based mortality models by a generalisation of Gompertz law of mortality”, *Insurance: Mathematics and Economics* **40**(3), 468–484.

Wilmoth, J. (1993), Computational methods for fitting and extrapolating the Lee–Carter model of mortality change, Technical report, Berkeley: Department of Demography, University of California.

Woodbury, M. A. and Manton, K. G. (1977), "A random–walk model of human mortality and ageing", *Theoretical Population Biology* **11**, 37–48.

Yashin, A. I. and Iachine, I. A. (1997), "How frailty models can be used for evaluating longevity limits: Taking advantage of an interdisciplinary approach", *Demography* **34**, 31–48.

Yashin, A. I., Manton, K. G. and Vaupel, J. W. (1985), "Mortality and ageing in a heterogeneous population: A stochastic process model with observed and unobserved variables", *Theoretical Population Biology* **27**, 154–175.

Yashin, A. I., Vaupel, J. W. and Iachine, I. A. (1994), "A duality in ageing: The equivalance of mortality models based on radically different concept", *Mechanisms of Ageing and Development* **74**, 1–14.

Yashin, A. I., Vaupel, J. W. and Iachine, I. A. (1995), "Correlated individual frailty: An advantageous approach to survival analysis of bivariate data", *Mathematical Population Studies* **5**, 145–159.